Electrodynamics

Calculate fields from sources and boundaries, then use Maxwell's equations to explain radiation, field forces and the propagation of pulses in a guide or cable.

Read the subject in depth

Start with electrostatics if this is your first field theory course. Each volume develops its arguments in sequence, with worked problems and links to experiments.

- ### Electrostatic field theory and measurement

  Learn to infer a field from charges, potentials or boundary measurements. Build the mathematics of flux, uniqueness and Green functions as each problem requires it.

- ### Magnetism, induction and field dynamics

  Begin with a steady current and follow what changes when a source, a conductor or a field moves. Track energy and momentum with explicit system boundaries.

- ### Waves, radiation and relativity

  Follow a disturbance from its source to a measurement. Work through polarisation, dispersion, scattering and changes of observer using measurable quantities.

- ### Fields and the response of matter

  Connect electric and magnetic measurements to relaxation times, carrier motion, thermal fluctuations and collective order. Use frequency, temperature and geometry to distinguish competing explanations.

- ### Electromechanics, optics and measurement

  Follow the field into an instrument. Predict which mechanical displacement, temperature change or optical signal can be observed and determine what can be learned from it.

- ### Numerical field laboratory

  Turn a prediction into a computational experiment. Derive a reference, solve the field problem, challenge the model with a diagnostic and interpret the residual.

How the course fits together

1. **Chapters 1–4**

   Charges and currents specify sources. Use symmetry and boundary conditions to turn their integral laws into field predictions, then ask how changing flux produces a voltage around a loop.

2. **Chapters 5–8**

   Combine the electric and magnetic energy balances to follow energy through space. A time-dependent source can radiate. Far-field expansions and interface conditions determine which parts of a source remain observable at a distance.

3. **Chapters 9–12**

   Conductivity, walls and material interfaces change how a field propagates. Follow a signal through a guide and along a transmission line, then integrate the field stress to find the force on matter.

Study electrodynamics by question

Each route connects an introductory calculation, a field method and an application. Follow one question across the book or use the numbered contents to work through the full course.

1. What can a field measurement reveal about its source?

Begin with the flux through a sphere. Next replace a symmetry argument by a boundary-value problem. Finish by asking how much detail about its source survives at a distance.

1. Charge, potential and field energy → 39. A ring, a disk and an electric-field maximum in empty space → 13. Local field laws at points, sheets and holes → 38. Recovering a vector field from divergence, curl and boundary data → 2. Conductors and boundary-value problems → 14. Building a potential from point responses → 40. The same metal sphere connected to ground or electrically isolated → 15. How far does an electrode pattern reach? → 41. How a circular wall selects Bessel modes → 16. Spherical harmonics as measurements of boundary voltage → 7. Multipole fields → 34. Capacitance as a matrix of electrostatic energies

Textbook comparison · Griffiths, ch. 1–3 · Fließbach, ch. 2–3, 7–12

Read the university derivation · David Tong · electrostatic fields (PDF) · Markus Zahn · rectangular and spherical boundary problems (MIT OpenCourseWare)

2. How do microscopic sources produce fields in matter?

Infer bound charge from polarisation and bound current from magnetisation. Check the interface condition before using a bulk constitutive constant in a calculation.

8. Polarization of a dielectric sphere → 18. How microscopic dipoles produce macroscopic charge → 36. Static screening and collective charge oscillations → 3. Magnetic flux and material response → 17. From a measured current to its magnetic field → 19. How a magnet produces an exterior field → 42. Forces and torques from a matrix of flux linkages

Textbook comparison · Griffiths, ch. 4–6 · Fließbach, ch. 13–15, 27–32

Read the university derivation · David Tong · electromagnetism in matter (PDF) · Markus Zahn · dielectric response and interfaces (MIT OpenCourseWare)

3. Where do energy and momentum travel?

Keep track of work, field storage, stress and transport in the same control volume. Then find the scale at which a circuit can represent those balances with a few variables.

4. Induction in moving circuits → 43. Where current and energy go in a charging capacitor → 5. Electromagnetic waves and energy flux → 12. Force carried by a field → 20. Where momentum waits between source and receiver → 29. When a circuit element must be treated as a field

Textbook comparison · Griffiths, ch. 7–8 · Fließbach, ch. 17–19, 25–26

Read the university derivation · David Tong · time-dependent fields (PDF) · Markus Zahn · induction and stored field energy (MIT OpenCourseWare)

4. ## How does a material or a wall select a wave?

   Start at one interface. Introduce material dispersion separately from waveguide dispersion. Use a cavity, a ray and an aperture to test the consequences of increasingly different length scales.

   9. Charge relaxation and magnetic diffusion → 35. How a conductor admits a field step and a sinusoid → 44. A lossy half-space seen through its surface impedance → 21. What an interface does to phase, field and power → 37. Reconstructing a partially polarised electromagnetic wave → 45. Two polarisations moving through a uniaxial crystal → 22. Why dispersion and absorption belong together → 10. Waves in a hollow guide → 11. Pulses on a transmission line → 27. Cavity eigenmodes, measured linewidth and ring-down → 30. From wavefronts to rays, focusing and caustics → 32. Diffraction from an aperture of finite width

   Textbook comparison · Griffiths, ch. 9–10 · Fließbach, ch. 28, 31, 33–38

   Read the university derivation · David Tong · propagation and material response (PDF) · Markus Zahn · waves confined to lines and guides (MIT OpenCourseWare)

5. ## When does an accelerating charge radiate?

   Distinguish an observation time from a source time, integrate energy across a distant sphere and compare free and bound charges. Close the energy balance at the emitting charge.

   6. Radiation from a small electric dipole → 23. Potentials, gauge freedom and the arrival of a signal → 24. A moving charge, its retarded position and its radiation → 28. Why a small transparent sphere scatters light → 46. Radiation when the electric dipole cancels → 31. Radiation reaction and the energy stored near a charge

   Textbook comparison · Griffiths, ch. 10–11 · Fließbach, ch. 21–24

   Read the university derivation · David Tong · radiation from moving charges (PDF) · Markus Zahn · fields of time-dependent sources (MIT OpenCourseWare)

6. ## What do two observers agree on?

   Transform charges, fields and forces between frames. Use invariants to check the signs, then calculate an observable orbit using relativistic momentum.

   25. Changing observers and classifying electromagnetic fields → 26. From local symmetry to Maxwell's equations → 33. Energy, momentum and a charged particle's worldline

   Textbook comparison · Griffiths, ch. 12 · Fließbach, ch. 4, 17–19

   Read the university derivation · David Tong · covariance and particle motion (PDF)

Contents

1. Chapter 1 · Charge, potential and field energy
2. Chapter 2 · Conductors and boundary-value problems
3. Problem sheet 1 · Fields from charge and metal boundaries
4. Chapter 3 · Magnetic flux and material response
5. Chapter 4 · Induction in moving circuits
6. Problem sheet 2 · Flux linkage and mechanical work
7. Chapter 5 · Electromagnetic waves and energy flux

8. Chapter 6 · Radiation from a small electric dipole
9. Problem sheet 3 · Energy flux and radiation
10. Chapter 7 · Multipole fields
11. Chapter 8 · Polarization of a dielectric sphere
12. Problem sheet 4 · Far fields and induced moments
13. Chapter 9 · Charge relaxation and magnetic diffusion
14. Chapter 10 · Waves in a hollow guide
15. Problem sheet 5 · Metal penetration and guide cutoff
16. Chapter 11 · Pulses on a transmission line
17. Chapter 12 · Force carried by a field
18. Problem sheet 6 · Reflections, force and stored energy
19. Chapter 13 · Local field laws at points, sheets and holes
20. Chapter 14 · Building a potential from point responses
21. Chapter 15 · How far does an electrode pattern reach?
22. Chapter 16 · Spherical harmonics as measurements of boundary voltage
23. Chapter 17 · From a measured current to its magnetic field
24. Chapter 18 · How microscopic dipoles produce macroscopic charge
25. Chapter 19 · How a magnet produces an exterior field
26. Chapter 20 · Where momentum waits between source and receiver
27. Chapter 21 · What an interface does to phase, field and power
28. Chapter 22 · Why dispersion and absorption belong together
29. Chapter 23 · Potentials, gauge freedom and the arrival of a signal
30. Chapter 24 · A moving charge, its retarded position and its radiation
31. Chapter 25 · Changing observers and classifying electromagnetic fields
32. Chapter 26 · From local symmetry to Maxwell’s equations
33. Chapter 27 · Cavity eigenmodes, measured linewidth and ring-down
34. Chapter 28 · Why a small transparent sphere scatters light
35. Chapter 29 · When a circuit element must be treated as a field
36. Chapter 30 · From wavefronts to rays, focusing and caustics
37. Chapter 31 · Radiation reaction and the energy stored near a charge
38. Chapter 32 · Diffraction from an aperture of finite width
39. Chapter 33 · Energy, momentum and a charged particle’s worldline
40. Chapter 34 · Capacitance as a matrix of electrostatic energies
41. Chapter 35 · How a conductor admits a field step and a sinusoid
42. Chapter 36 · Static screening and collective charge oscillations
43. Chapter 37 · Reconstructing a partially polarised electromagnetic wave
44. Chapter 38 · Recovering a vector field from divergence, curl and boundary data
45. Chapter 39 · A ring, a disk and an electric-field maximum in empty space
46. Chapter 40 · The same metal sphere connected to ground or electrically isolated
47. Chapter 41 · How a circular wall selects Bessel modes
48. Chapter 42 · Forces and torques from a matrix of flux linkages
49. Chapter 43 · Where current and energy go in a charging capacitor
50. Chapter 44 · A lossy half-space seen through its surface impedance
51. Chapter 45 · Two polarisations moving through a uniaxial crystal
52. Chapter 46 · Radiation when the electric dipole cancels

53. Chapter 47 · Four contacts on a conductor of finite thickness
54. Problem sheet 7 · Project · Infer a charge distribution from a grounded tube
55. Problem sheet 8 · Project · Trace energy from a wave to the material supporting it
56. Problem sheet 9 · Project · Measure a cavity and identify two different loss mechanisms
57. Electrostatic field theory and measurement
58. Magnetism, induction and field dynamics
59. Waves, radiation and relativity
60. Fields and the response of matter
61. Electromechanics, optics and measurement
62. Numerical field laboratory
63. Applied study
64. Mathematical field guide to electrodynamics
65. Notation and assumptions
66. Further reading

Chapter 1 · Charge, potential and field energy

Reconstruct both the local electric field and the total energy of a charged body from the same continuous charge distribution.

FROM FIRST PRINCIPLES

Count the charge within a spherical surface

A potential probe gives a number at a point. To predict that reading from charge density, begin with two physical inputs in vacuum. Gauss's law relates the electric flux through a closed surface to its enclosed charge. In a static field the work per unit charge between two points is independent of path. For a spherically symmetric density $\rho(r)$, contributions in tangential directions cancel. The radial field has the same magnitude everywhere on a sphere of radius r. A shell of radius s and thickness ds contains charge $4\pi s^2 \rho(s)\, ds$. Hence

$$Q(r) = 4\pi \int_0^r \rho(s) s^2 \, ds,$$

$$4\pi r^2 E_r(r) = Q(r)/\epsilon_0.$$

For a sphere of radius R carrying uniform density, the enclosed charge grows as r^3 inside and stops growing outside. The surface through which the flux passes continues to grow as r^2. Dividing gives two different field laws:

$$E_r(r) = \frac{\rho r}{3\epsilon_0} \ (r \le R),$$

$$E_r(r) = \frac{\rho R^3}{3\epsilon_0 r^2} \ (r \ge R).$$

Choose zero potential at infinity and integrate the work inward. For an interior probe split the path at the surface so that the appropriate field is used on each segment:

$$V(r) = \int_R^\infty E_{\text{out}}(s)\, ds + \int_r^R E_{\text{in}}(s)\, ds = \frac{\rho}{6\epsilon_0}(3R^2 - r^2).$$

The centre has a finite voltage even though its field vanishes. The probe measures the accumulated work along a path, while the field is the local slope of that voltage.

Background · Potential and electric field

$$\mathbf{E} = -\nabla V$$

A volt per metre is a newton per coulomb. This relation describes an electrostatic field. Time-dependent induction introduces a vector-potential term.

DERIVATION AND APPLICATIONS

How an interior voltage reveals a radial charge gradient

An outside probe cannot distinguish spherically symmetric charge distributions with the same total charge. To see what an inside probe adds, suppose an insulating sphere of radius R holds an immobile charge density $\rho(r) = \rho_0(1 - r/R)$ in vacuum. Ignore polarisation of the material and set potential zero at infinity. A spherical shell of thickness dr contributes its density times $4\pi r^2 dr$. The resulting enclosed charge is

$$q(r) = 4\pi\rho_0 \left(\frac{r^3}{3} - \frac{r^4}{4R} \right),$$

$$Q = q(R) = \frac{\pi \rho_0 R^3}{3}.$$

Rotational symmetry fixes the direction of the field and makes its magnitude constant over the Gaussian sphere. Substituting the enclosed charge, then eliminating the density amplitude using the measured total charge, gives

$$E_r(r) = \frac{q(r)}{4\pi\epsilon_0 r^2} = \frac{kQ}{R^2}(4x - 3x^2),$$

$$x = \frac{r}{R},$$

$$k = \frac{1}{4\pi\epsilon_0}.$$

At the surface this expression joins the exterior Coulomb field. Positive charge concentrated towards the centre makes the field peak before reaching the surface. Potential instead accumulates the contributions of the field along the whole radial path. Starting from the known surface potential fixes the integration constant even when the voltmeter only accesses the interior.

$$V(r) = \frac{kQ}{R} + \int_r^R E_r(s)\, ds = \frac{kQ}{R}(2 - 2x^2 + x^3).$$

WORKED EXAMPLE

Identify an unknown charge profile

Change the inputs and investigate this derivation.

A specimen of radius 60.0 mm has a surface voltage of 30.0 V relative to a distant reference. The inferred charge is $Q = VR/k = 0.200$ nC. The gradient model predicts a centre voltage of $2(30.0) = 60.0$ V, so a differential probe between centre and surface should read 30.0 V. A uniformly charged sphere with that same outside voltage gives 45.0 V at the centre. A sufficiently small probe could distinguish these source profiles with a single additional reading.

SECOND APPLICATION

Find where a field probe reads its maximum

Change the inputs and investigate this derivation.

For another sphere, $R = 90.0$ mm and $Q = 0.900$ nC. Differentiate $4x - 3x^2$ with respect to x. Its derivative vanishes at $x = 2/3$, with negative second derivative. The peak is therefore 60.0 mm from the centre and has magnitude $(4/3)kQ/R^2 = 1.33$ kV/m. The surface value is only 0.999 kV/m. The enclosed charge still grows beyond the peak, but increasingly dilute outer shells provide too little additional flux to keep pace with the increasing area.

Try it · Graded charge

Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Sphere radius (mm) **60**

Total charge (nC) **0.2**

Linearly falling volume-charge density

Field maximum at r=2R/3 = 40 mm. Surface/centre potentials 29.959 / 59.917 V. Density is proportional to 1−r/R inside.

Start from this chapter. For concentric conductors, use a cylindrical Gaussian surface and integrate the radial field to get a voltage. Compare the capacitance per length with the numerical charge on the centre conductor.

OneSimulation laboratory · Recover cable impedance from a cross-section

Can an exterior field tell you where the charge sits?

A spherical cloud and a thin shell each carry charge Q and have the same outer radius R. Before moving either slider, sketch the two fields along a radius. Which part of the sketch could distinguish the sources in a measurement?

A Gaussian sphere of radius r encloses $Q(r/R)^3$ in the solid ball for $r < R$. Dividing by its area gives

$$E = Qr/(4\pi\epsilon_0 R^3).$$

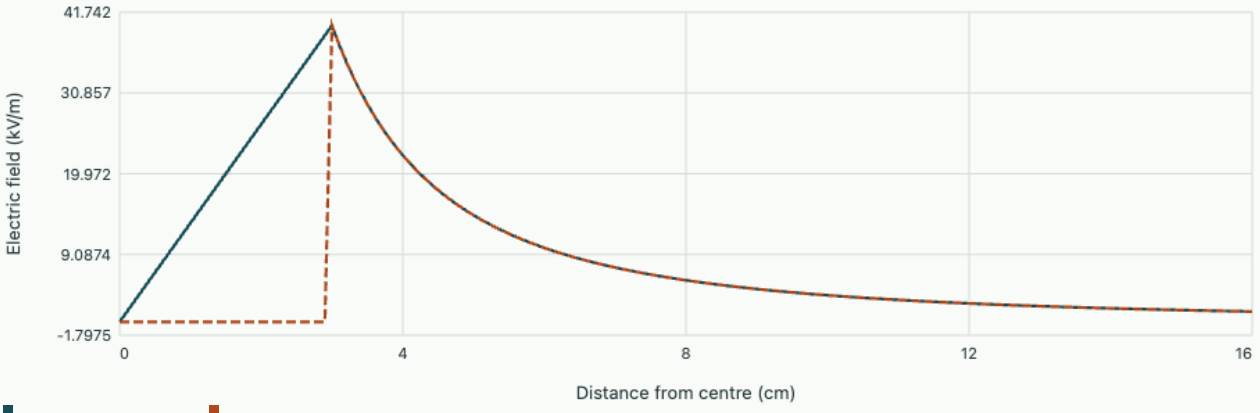
The shell encloses zero until its surface. Outside, both enclose the same Q. The shell curve shows the exterior value at the discontinuity.

Open this experiment online

Try the calculation

Charge Q (nC) 4

Cloud radius R (cm) 3



Uniform ball | Thin spherical shell

At the surface the exterior limit is 39.945 kV/m. Halfway out, the solid ball gives 19.972 kV/m.

Keep Q fixed and double R. The maximum field falls by a factor of four, while the interior slope falls by a factor of eight. Explain why those are distinct measurements.

Why a charge in an ionic solution has a finite electrostatic range

Put a weakly charged small bead in a large bath of two mobile ion species carrying charges $+e$ and $-e$. Each species has equilibrium number density n_0 far away. The solvent is a continuum with uniform permittivity ϵ, and a heat bath fixes temperature T. In a vacuum the measured potential of a spherical charge falls as $1/r$, even very far from it. In this liquid the electric field redistributes material, and the

material changes the field. We can determine the decay law and the cloud that produces it using the same Gauss law. Fix the potential reference at infinity, where the liquid is electrically neutral.

Let thermal motion compete with electrical energy

An ion's local electrostatic energy is $+eV$ or $-eV$. In dilute thermal equilibrium its local density acquires a Boltzmann factor, $n_\pm = n_0 \exp[\mp eV/(k_B T)]$, because the far-away solution acts as a reservoir of particles. A positive bead repels positive ions and attracts negative ions. The mobile charge density is therefore

$$\rho_m = e(n_+ - n_-) = -2en_0 \sinh[eV/(k_B T)].$$

This is a source in Poisson's equation, not a new constitutive law for the solvent. We have separately included the solvent's polarisation in ϵ.

If $|eV| \ll k_B T$, expand to first order. The symmetric mixture cancels even powers. This gives $\rho_m = -(2e^2 n_0/k_B T)V$, and an externally prescribed bead charge density ρ_b now satisfies

$$\nabla^2 V - \kappa^2 V = -\rho_b/\epsilon,$$

$$\kappa^2 = \frac{2e^2 n_0}{\epsilon k_B T}.$$

The inverse length κ came from comparing the field-curvature coefficient with the charge response. Increasing either mobile species' density shortens the distance over which a disturbance survives. Increasing temperature weakens the concentration change caused by a given voltage. Within these assumptions that length grows as $\sqrt{T}$ at fixed permittivity and density.

Solve the exterior and test its near-source assumption

In the liquid outside a spherical bead of radius a, put $v = rV$. The radial Laplacian becomes v''/r, so

$$v'' - \kappa^2 v = 0.$$

Discard the exponentially increasing solution at infinity. For a bead with fixed free charge q and no mobile ions inside, take a spherical Gaussian surface just outside the surface. Spherical symmetry and the boundary condition give

$$-4\pi a^2 \epsilon V'(a^+) = q.$$

Solving for the remaining coefficient yields

$$V(r) = \frac{q}{4\pi\epsilon(1+\kappa a)} \frac{e^{-\kappa(r-a)}}{r},$$

$$E_r(r) = \frac{q}{4\pi\epsilon(1+\kappa a)} e^{-\kappa(r-a)} \left(\frac{1}{r^2} + \frac{\kappa}{r}\right), \quad r \geq a.$$

The maximum potential magnitude in the liquid occurs at the bead surface. Its value $q/[4\pi\epsilon a(1+\kappa a)]$ therefore tests the linearisation everywhere outside. Shrinking a bead at fixed q ultimately violates this local condition even if the distant tail remains small. An ideal point-source expression can guide a far-field fit without faithfully describing the molecular neighbourhood of the source.

Account for the apparent disappearance of charge

A collector of radius r measures free enclosed charge $4\pi\epsilon r^2 E_r$, counting the bead and the mobile ions. It therefore encloses the fraction

$$\frac{Q_{\mathrm{enc}}(r)}{q} = \frac{1+\kappa r}{1+\kappa a} e^{-\kappa(r-a)}.$$

At the bead it reads one. At infinity it tends to zero. The missing charge is precisely the countercharge in the surrounding liquid. It can be verified independently without differentiating the field: integrate $4\pi r^2 \rho_m = -4\pi\epsilon\kappa^2 r^2 V$ from a to infinity. The radial integral of $re^{-\kappa(r-a)}$ is $a/\kappa + 1/\kappa^2$, cancelling the factor $1+\kappa a$. The resulting total mobile charge is $-q$. A screening length describes a decay scale; charge is not abruptly cancelled on one particular shell.

For $\kappa a = 1$, at $r = a + \kappa^{-1} = 2a$ the voltage has fallen to $e^{-1}/2 = 18.4\%$ of its surface value. Yet the collector still encloses $3/(2e) = 55.2\%$ of the bead charge. Voltage and enclosed charge are different operations on the solution. A simulation or probe returning the voltage ratio cannot be read directly as the fraction of ions that have neutralised the bead.

Predict what changes when the ion valence changes

For several ionic species with charge $z_i e$, the Boltzmann expansion gives

$$\rho_m = \sum_i z_i e n_{i0} - (e^2/k_B T)\sum_i z_i^2 n_{i0} V + \cdots.$$

Bulk electroneutrality cancels the first sum. The screening coefficient instead contains squared valences,

$$\kappa^2 = e^2 \sum_i z_i^2 n_{i0}/(\epsilon k_B T).$$

The sign of a species determines which way it moves in a given field, while its contribution to linear screening is positive for either sign. A divalent ion both feels double the electrical energy change and carries double the charge when it rearranges.

Compare two salts at the same formula-unit number density n. A fully dissociated 1:1 salt contributes $2n$ to the squared-valence sum. A 2:1 salt with one divalent cation and two singly charged anions contributes $4n + 2n = 6n$. The predicted length is smaller by $1/\sqrt{3}$, even though both baths are electrically neutral. The comparison assumes dilute ideal ions with the same solvent response. At higher concentrations, ion-size and correlation effects can remove this simple dependence.

Recover a screening length from a pair of readings

A bead has radius 8.0 nm. Relative to infinity two probes read 3.0 mV at 12 nm and 0.90 mV at 20 nm. Infer the screening length and surface voltage within the model. How could a third probe falsify the interpretation?

▼ **Derive the result and check it independently**

The ratio of rV, not V alone, isolates the exponential. Here $r_2V_2/(r_1V_1) = 18/36 = 1/2$, so

$$\kappa^{-1} = 8.0/\ln 2 = 11.54\,\mathrm{nm}.$$

Extrapolating to a gives

$$V(a) = V(12)(12/8)e^{4/11.54} = 6.36\,\mathrm{mV}.$$

At room temperature k_BT/e is approximately 25.7 mV. The largest ratio is 0.247 and the cubic correction in $\sinh x/x$ there is about 1%, making linearisation a plausible first approximation.

Plot $\ln[rV(r)]$ against r for a third and further radii. A single straight line with slope $-\kappa$ must describe all of them, without any external calibration of q. Curvature or dependence on the chosen pair of radii would challenge the isolated spherical linear model. Before interpreting a fitted slope as an ionic strength, independently measure or specify ϵ, T, species charges and background electrostatic conditions.

Cambridge Electromagnetism, chapter 6 · screening and the Debye–Hückel model

Enclose the source before integrating the field

Imagine an insulating sphere charged uniformly throughout its volume. A probe outside it measures the field of a point charge, yet that representation fails at its centre. Our first objective is to see how a local charge distribution determines which measurements can distinguish the two sources. Use electrostatics in vacuum, with a sphere of radius R and total free charge Q. A conductor in electrostatic equilibrium would distribute its excess charge differently.

Gauss's law states that the electric flux through any closed surface is the enclosed charge divided by ϵ_0. Flux is the integral of the component normal to the surface. Pulling the field magnitude outside that integral requires symmetry in addition to the law itself. For this sphere rotational symmetry forces the field to be radial and of equal magnitude everywhere on a concentric spherical surface. The test surface can pass through the charged material, since it is a geometrical construction and does not insert a conductor.

Background · Can a zero-flux surface contain a strong field?

Imagine a point charge outside a closed spherical Gaussian surface. Its field is nonzero on every point of the sphere. Near the charge the field enters the sphere and elsewhere it leaves. A small cone of field lines cuts the surface twice: at each crossing its flux magnitude is fixed by the same solid angle at the charge. The two oriented contributions cancel.

$$\oint_{\partial V} \mathbf{E} \cdot d\mathbf{A} = 0, \qquad \oint_{\partial V} |\mathbf{E} \cdot d\mathbf{A}| > 0.$$

Place a charge at the centre instead and there is only an outward crossing. Spherical symmetry then permits taking the field magnitude outside the integral. Any external charge destroys that simplification for the total field even though the enclosed charge and the signed integral stay exactly the same.

Match the interior and exterior solutions

Uniform volume density gives $\rho = 3Q/(4\pi R^3)$. A Gaussian sphere inside the charged body encloses the same fraction of its charge as of its volume. Outside, it encloses the entire charge. Substituting the enclosed charge into Gauss's law gives

$$4\pi r^2 E_r(r) = \frac{Q_{\mathrm{encl}}(r)}{\epsilon_0},$$

$$E_r(r) = \begin{cases} \dfrac{Qr}{4\pi\epsilon_0 R^3}, & 0 \le r \le R, \\ \dfrac{Q}{4\pi\epsilon_0 r^2}, & r > R. \end{cases}$$

Both branches have units of volts per metre and give the same surface value. The field vanishes at the centre. A finite change in volume density changes the slope at the surface, while a sheet of surface charge could produce a jump in the normal field itself.

Since a static field has zero curl, a scalar potential can be chosen such that $\mathbf{E} = -\nabla V$. Take $V(\infty) = 0$ and integrate the radial field inward. The exterior potential equals $Q/(4\pi\epsilon_0 r)$. To reach an interior point, continue the integral from the surface with the interior field. This gives

$$V(r) = \frac{Q}{8\pi\epsilon_0 R}\left(3 - \frac{r^2}{R^2}\right), \qquad 0 \le r \le R.$$

Matching at the surface supplies the additive constant which an interior integration alone cannot fix. A zero electric field at the centre means that the slope of the potential vanishes there. Its value need not vanish.

Assemble the sphere and account for the work

Build the charge by adding successive spherical shells while holding the final density fixed. When the assembly reaches radius r, its charge is $q(r) = Q(r/R)^3$. Bringing a small additional shell $dq = 3Qr^2dr/R^3$ from infinity costs $dW = q(r)dq/(4\pi\epsilon_0 r)$. Integration gives

$$W = 3Q^2/(20\pi\epsilon_0 R).$$

This positive work is stored in the electrostatic configuration. The shell’s interaction with itself is higher order in the infinitesimal added charge.

A complementary description uses the energy density $\epsilon_0 E^2/2$. Integration over the exterior alone gives $Q^2/(8\pi\epsilon_0 R)$, which is five sixths of the total. The remainder resides inside the charge distribution. A thin spherical conducting shell of equal radius and charge produces the same outside field, yet it has no interior field in its empty cavity. Its total energy is therefore lower.

Interpret three probe readings

For $R = 0.10$ m and $Q = 2.0$ nC, the surface field is 1.80×10^3 V/m. At half the radius the field halves to about 899 V/m. At twice the radius it is roughly 449 V/m. The surface potential is 180 V, while the centre is near 270 V. Their difference predicts the work per unit charge required to carry a small positive test charge from surface to centre. A sufficiently large test charge would disturb the source distribution and violate the assumed measurement model.

The assembly energy is only 2.16×10^{-7} J. Doubling Q multiplies the field and potential by two but multiplies the stored energy by four. A comparison of fields outside two spherical samples probes their total charge. A measurement of assembly work at fixed radius can reveal how that charge was distributed inside them.

If the charge is reduced gradually while its spatial shape is preserved, the surface voltage is proportional to total charge. The incremental work needed to assemble the next element therefore grows with charge already present. Integrating that growth explains the familiar factor of one half in electrostatic self-energy expressions.

Insert a dielectric into a spherical capacitor

An enclosed-charge construction also works when different radial regions store different electric energies. Put a conducting sphere of radius a inside a concentric thin conducting shell of radius b. Give the inner sphere free charge $+Q$ and the outer shell $-Q$. Fill $a < r < c$ with a linear dielectric of permittivity ϵ_1 and $c < r < b$ with ϵ_2. There is no free charge at the interface. Assume the shells and both materials share a common centre so that a spherical Gaussian surface has uniform radial displacement.

Gauss's law for free charge gives $D_r = Q/(4\pi r^2)$ throughout the gap. The electric field in material i is D_r/ϵ_i. Thus the normal field can jump at c even though normal displacement is continuous. To predict the measured voltage, integrate the radial field outward from a to b; the positive quantity obtained is the inner voltage minus the outer voltage. This yields

$$\frac{1}{C} = \frac{V_a - V_b}{Q} = \frac{1}{4\pi}\left[\frac{1/a - 1/c}{\epsilon_1} + \frac{1/c - 1/b}{\epsilon_2}\right].$$

The sum is naturally a sum of inverse capacitances since the same flux passes through both layers and their voltage drops add. Replacing the two dielectric constants by a simple average would weight every radial thickness equally, despite the field being stronger near the small inner sphere. Setting the two permittivities equal cancels the arbitrary interface location, a useful check on the integration.

Decide where a small amount of dielectric matters most

Choose $a = 10.0$ mm, $c = 20.0$ mm, $b = 40.0$ mm,

$$\epsilon_1 = 4\epsilon_0,$$
$$\epsilon_2 = \epsilon_0.$$

The radial contributions after factoring out $1/(4\pi\epsilon_0)$ are $(100 - 50)/4 = 12.5$ m⁻¹ and $50 - 25 = 25.0$ m⁻¹. Hence

$$C = 4\pi\epsilon_0/(37.5\,\mathrm{m}^{-1}) = 2.97\,\mathrm{pF}.$$

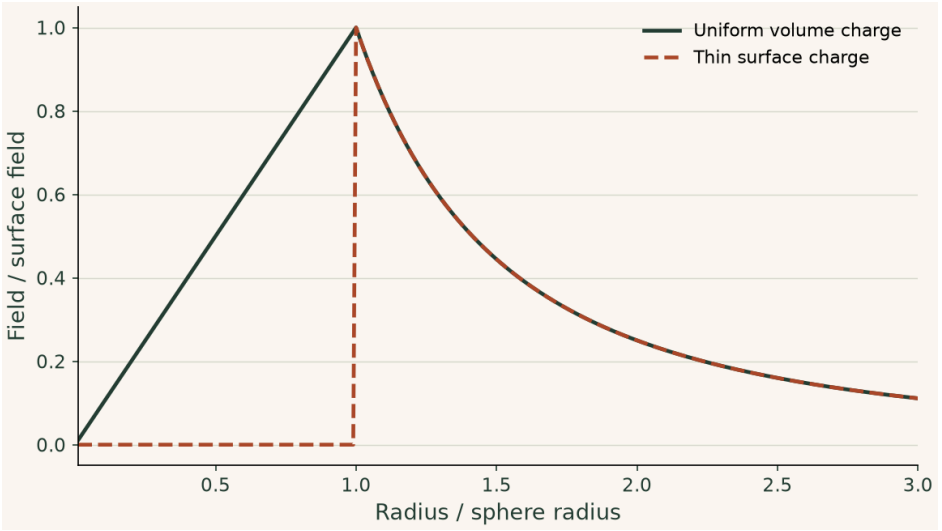
A charge of 0.300 nC produces about 101 V, one third across the inner material and two thirds across the outer one. An all-vacuum device of the same radii has $C = 1.48$ pF.

For a further check integrate local energy density $u = \mathbf{E}\cdot\mathbf{D}/2$ over each layer. The integration element $4\pi r^2 dr$ reduces the energy integral to $Q^2(1/r_{\mathrm{in}} - 1/r_{\mathrm{out}})/(8\pi\epsilon_i)$. The two contributions are again in the ratio one to two, summing to $Q^2/(2C) = 15.2$ nJ. Moving dielectric towards the inner conductor usually has a larger effect per unit volume because the undisturbed energy density is higher there. Near metal contact surfaces roughness or air pockets can make precisely this region differ from its intended permittivity.

At the interface voltage is continuous but radial field just inside the outer material is four times its value just inside the inner material for these permittivities. A finite jump in voltage would imply an infinitely strong field over zero thickness, absent any imposed dipole sheet. Inserting these interfacial fields into the constitutive law gives identical normal displacement from either side. This is a more local test than comparing the terminal capacitance: a calculation can accidentally reproduce a similar integrated voltage while placing the field discontinuity on the wrong side of the boundary.

Infer a charge distribution from a radial field scan

Two spheres of equal radius have the same positive total charge. One has uniform volume density, the other stores charge in a thin spherical shell. Apply Gauss’s law to spheres of varying radius and scale the result by $Q/(4\pi\epsilon_0 R^2)$.



At half radius the filled sphere has half the surface field and the shell has zero. All exterior readings coincide. Even a perfect exterior scan cannot distinguish these two spherically symmetric charge distributions.

▶ **Read values from the plot**

Develop this calculation further

Chapter 13 · Local field laws at points, sheets and holes · Chapter 18 · How microscopic dipoles produce macroscopic charge

Study the full derivation

From a force reading to Poisson’s equation · Electrostatic field theory and measurement

Imagine trying to infer an invisible charge distribution with a force probe and a voltmeter. Coulomb’s law, Gauss’s law and the electrostatic potential describe the same experiment in three different ways. Deriving the connections also exposes the measurements a single probe cannot make.

Chapter 2 · Conductors and boundary-value problems

Use the uniqueness of the potential to replace a grounded conductor by an image charge and then recover its induced surface charge.

FROM FIRST PRINCIPLES

Why the conductor voltage determines the exterior field

In electrostatic equilibrium a mobile charge cannot have a tangential electric force along a conductor. The surface is therefore at one potential. A grounded wire fixes that potential at zero. In the surrounding vacuum $\mathbf{E} = -\nabla V$. Apply Gauss's law to a small volume and use the divergence theorem:

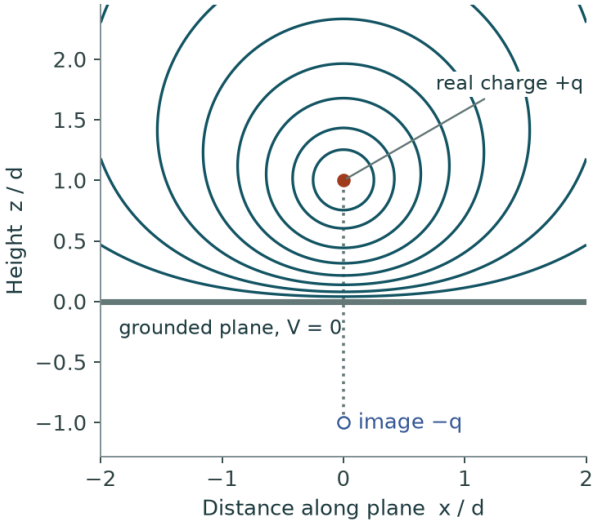
$$\oint_{\partial\Omega} \mathbf{E}\cdot d\mathbf{a} = \int_{\Omega} \rho/\epsilon_0 \, dV,$$

$$\nabla^2 V = -\rho/\epsilon_0.$$

Why is it enough to find one potential with the correct sources and boundary values? Let $w = V_1 - V_2$ for any two candidates with the same charges. The source cancels, leaving $\nabla^2 w = 0$. Their common prescribed voltage gives $w = 0$ on the boundary. Integrate the product rule for $\nabla\cdot(w\nabla w)$:

$$\int_{\Omega} |\nabla w|^2 \, dV = \oint_{\partial\Omega} w\,\partial_n w \, dA - \int_{\Omega} w\nabla^2 w \, dV = 0.$$

The integrand is nonnegative, so the two gradients coincide throughout the domain. For an exterior problem we also prescribe decay at infinity. An image charge placed outside the physical solution domain can therefore generate the exact field inside it, provided the potential satisfies all of these conditions. The induced surface charge then follows from the outward normal field just outside the metal.



Why the image charge satisfies an entire boundary condition

Every point on the plane has the same distance from the real and reflected charges. Their potentials cancel there. Above the plane the image contributes a harmonic potential, so the real charge supplies the only singularity in the physical domain. Uniqueness selects this solution when the potential also decays at infinity.

Try it · Plane image

Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Charge height (cm) 2

Point charge (nC) 1

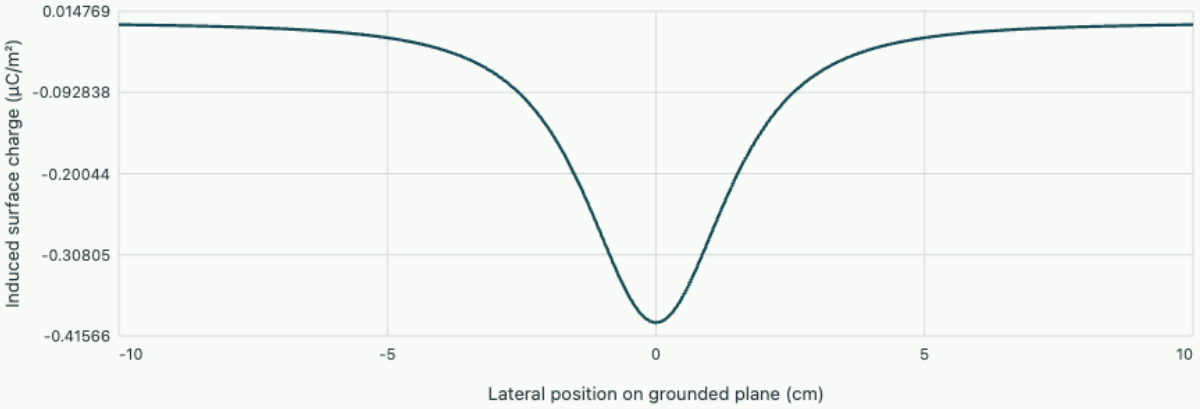


Image-charge surface density

The profile is a radial cut through the two-dimensional surface density. Integrating over the entire plane gives –q; integrating this line alone does not.

DERIVATION AND APPLICATIONS

Find the image charge of a grounded sphere

Curvature changes both the location and the size of an electrostatic image. Place a point charge q a distance a from the centre of a grounded conducting sphere of radius $R < a$. Take vacuum outside, zero potential at infinity and no other sources. Axial symmetry suggests placing a trial image q' on the same axis at distance $b < R$. At a surface point making angle θ with the source axis, let d and d' denote the distances to charge and image respectively.

$$d^2 = R^2 + a^2 - 2Ra\cos\theta,$$
$$d'^2 = R^2 + b^2 - 2Rb\cos\theta.$$

To cancel the potentials at every angle, the two distances must be proportional by one constant. Choose $b = R^2/a$. Both the constant and cosine terms in d'^2 then equal those in d^2 times R^2/a^2. This fixes the

required image strength through the grounded boundary condition.

$$\frac{d'}{d} = \frac{R}{a},$$
$$\frac{q}{d} + \frac{q'}{d'} = 0,$$
$$q' = -\frac{R}{a}q.$$

This construction has the physical charge as its only exterior singularity. It decays at infinity and satisfies the prescribed surface voltage, so uniqueness fixes the exterior potential. A Gaussian surface enclosing the metal but excluding the source encloses exactly the image in the mathematical field. Its flux therefore identifies the net induced charge with q'. The force on the source points towards the centre. Using $k = 1/(4\pi\epsilon_0)$ gives its radial component

$$F_a = \frac{kqq'}{(a-b)^2} = -\frac{kq^2Ra}{(a^2-R^2)^2}.$$

WORKED EXAMPLE

Predict the balance reading near a metal sphere

Change the inputs and investigate this derivation.

Let $R = 50.0$ mm, $a = 100$ mm and $q = 2.00$ nC. The image is -1.00 nC at 25.0 mm from the centre. The source and image are consequently 75.0 mm apart. The attractive force is $8.99 \times 10^{9}(2.00 \times 10^{-18})/(0.0750)^2 = 3.20$ µN. A position clamp must supply the same magnitude outwards to hold the charge stationary.

SECOND APPLICATION

Recover a source distance from the grounding lead

Change the inputs and investigate this derivation.

A sphere of radius 30.0 mm acquires -0.900 nC when a positive 3.00 nC source is introduced from infinity. Inverting the induced-charge ratio gives $a = R(3.00/0.900) = 100$ mm, or 70.0 mm clear of the surface. The limit $a \gg R$ explains the small charge transfer from ground at long range. An isolated, initially neutral sphere would instead require an additional central image to keep its net induced charge zero.

Try it · Sphere force

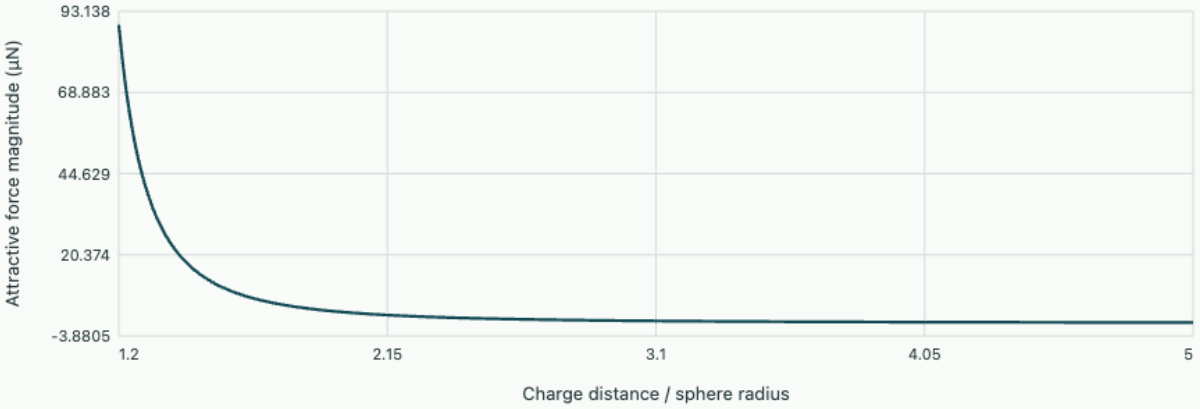
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Sphere radius (mm) **50**

External charge (nC) **2**



Grounded conducting sphere

Image charge q′=−q/(a/R); force acts towards the sphere. At a=2R the image lies at R/2 and carries −q/2.

Start from this chapter. Move the centre conductor off axis while holding both conductor potentials fixed. The radial logarithmic solution loses its symmetry. Which integral of the numerical field still defines capacitance?

OneSimulation laboratory · Recover cable impedance from a cross-section

Measure the voltage that a grounded enclosure can hold inside

Imagine a long grounded square tube, one metre on each side. Uniformity along the tube reduces the electrostatic problem to its square cross section. A prescribed volume charge $\rho_0 \sin(\pi x/L)\sin(\pi y/L)$, with $L = 1\ \mathrm{m}$, vanishes at all four walls. Ground fixes the boundary voltage. Does that also fix the interior voltage to zero?

Gauss's law with $\mathbf{E} = -\nabla V$ gives

$$-\nabla^2 V = \rho/\epsilon_0.$$

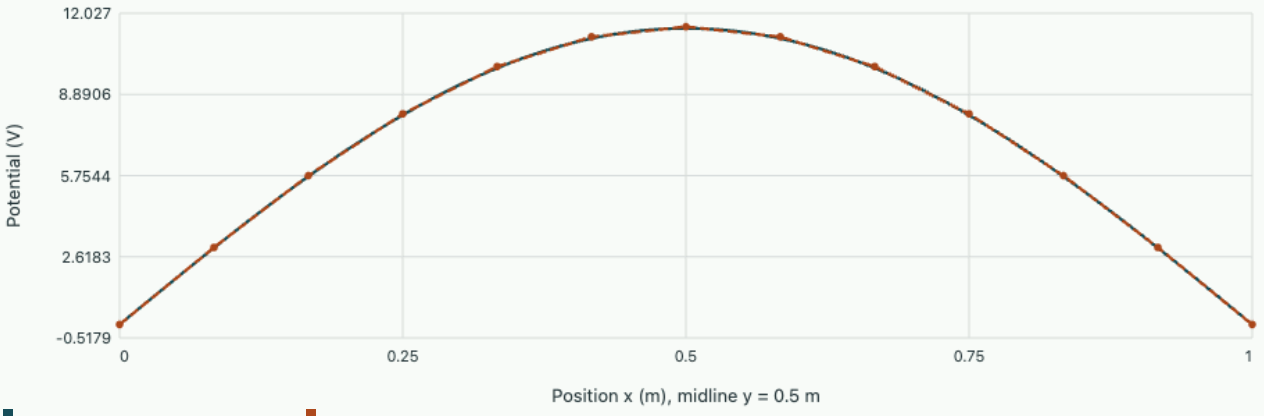
Each second derivative of the sine product contributes $-(\pi/L)^2$ times that product. Thus $V_{\text{centre}} = \rho_0 L^2/(2\pi^2\epsilon_0)$. The numerical experiment replaces $-\nabla^2 V$ at each interior node by $(4V_C - V_N - V_S - V_E - V_W)/h^2$. The charge is a single eigenmode of that stencil, so one conjugate-gradient step from zero solves the discrete system exactly in exact arithmetic.

Open this experiment online

Try the calculation

Cells per side **12**

Charge amplitude ρ_0 (nC/m³) **2**



Centre potential 11.50889 V on the grid, 11.44331 V in the continuum. Nodal error 0.06558 V.

Download the OneSimulation model

In your OneSimulation environment, run the model with these inputs.

```
python grounded-square.py --cells 12 --charge 2
```

At 2 nC/m³, 12 cells per side give 11.50889 V at the centre, compared with 11.44331 V in the continuum. Double the grid resolution. Explain why the error falls close to fourfold even though both discrete solves have converged. Then double the charge and predict which of absolute and relative error changes.

The CSV contains every midline grid value. Compare its centre row with the plot after choosing the same charge and grid. The browser applies the five-point operator directly. The Python version assembles and solves the coefficient-form PDE with OneSimulation.

How a right-angle conducting corner shares an induced charge

Two grounded conducting planes meet at a right angle, occupying the faces $x = 0,\ y > 0$ and $y = 0,\ x > 0$. A long uniformly charged thin filament of line charge $\lambda > 0$ runs parallel to their intersection through position (a, b), with $a, b > 0$. The vacuum region is the first quadrant. Quantities below are per unit length along the filament; its radius is negligible compared with both distances to the metal. Find the force on the filament and how the grounded faces divide the induced charge. Knowing that the two metal faces have identical voltage does not tell us how much charge each supplies.

Construct a potential that satisfies both metal faces

For an isolated line, Gauss's law gives

$$E_R = \lambda/(2\pi\epsilon_0 R);$$

integration gives

$$V = -[\lambda/(2\pi\epsilon_0)]\ln(R/R_0).$$

A single line's potential has no zero reference at infinity. Our image system will have zero total line charge, so all arbitrary R_0 factors cancel. Reflect the real source in the vertical face and reverse its sign. Reflect it independently in the horizontal face and reverse its sign again. Neither added source can be left without its own partner at the other face. The fourth line at $(-a, -b)$ must be positive. Denote distances to the real and image lines by $R_{++}, R_{-+}, R_{+-}, R_{--}$ following their coordinates. Then

$$V = \frac{\lambda}{2\pi\epsilon_0}\ln\frac{R_{-+}R_{+-}}{R_{++}R_{--}}.$$

At $x = 0$ distances cancel pairwise in the ratio, and the same happens at $y = 0$. The solution tends to zero at large distance and its only singularity in the physical domain is at the real filament. The grounded Dirichlet boundary and appropriate behaviour at the corner and infinity fix the field uniquely. If the faces were instead free to hold different voltages, this particular superposition alone would not solve their boundary conditions.

Read the signed force from three image fields

At the source position remove its own singular field. The vertical reflection attracts horizontally with field $-\lambda/(4\pi\epsilon_0 a)$; the horizontal reflection attracts vertically by the corresponding expression with b. The doubly reflected positive image repels along (a, b), partly compensating both attractions. Multiplying by λ gives

$$\frac{\mathbf{F}}{L} = \frac{\lambda^2}{4\pi\epsilon_0}\left[\left(-\frac{1}{a} + \frac{a}{a^2+b^2}\right)\hat{\mathbf{x}} + \left(-\frac{1}{b} + \frac{b}{a^2+b^2}\right)\hat{\mathbf{y}}\right].$$

In the symmetric case $a = b$ the two components are equal. Each is exactly half as large as the attraction to its corresponding isolated face. The second grounded conductor changes the charge distribution of the first; simply adding two single-plane forces would overestimate the force. If $a \gg b$, the horizontal force

towards the remote vertical wall decays as b^2/a^3, while the downward force approaches its single-plane value. The charge and its nearest opposite image look more like a dipole to the distant face.

Find the charge split by two independent routes

On the bottom face the surface normal from metal into vacuum is $+\hat{\mathbf{y}}$. Computing $-\epsilon_0 \partial_y V(x, 0^+)$ gives

$$\sigma_{\text{bottom}}(x) = -\frac{\lambda b}{\pi}\left[\frac{1}{(x-a)^2+b^2} - \frac{1}{(x+a)^2+b^2}\right], \quad x > 0.$$

The second term reduces the charge near the corner and forces the density to vanish there. Integration from zero to infinity uses the antiderivative $b^{-1}\arctan[(x \mp a)/b]$. The two terms give $-2\lambda \arctan(a/b)/\pi$. Exchanging a, b gives the vertical face charge. Their sum is $-\lambda$ because the two acute arctangents add to $\pi/2$. For $a = 20$ mm and $b = 10$ mm the bottom and vertical faces receive respectively 70.5% and 29.5% of the induced negative charge. The closer face collects the greater magnitude, although both are perfectly grounded.

A reciprocity argument provides a check with no surface differentiation. Remove the filament and solve an auxiliary problem with the vertical face at 1 V and the bottom at 0 V. In polar coordinates the bounded angular solution is $w = 2\theta/\pi$ after division by 1 V, because its second angular derivative vanishes and its face values agree. The singular junction between ideal differently biased electrodes belongs only to this auxiliary problem. The induced charge per length on the real vertical face is $-\lambda w(a, b) = -2\lambda \arctan(b/a)/\pi$. This reproduces the integral result and interprets the measured charge split as an angular coordinate.

Anticipate the field near a corner before refining a mesh

Empty space immediately around the wedge vertex contains no real source if $a, b > 0$. In planar polar coordinates a separated harmonic function vanishing on the two faces has form $r^\nu \sin(\nu\theta)$, with $\nu = m\pi/\alpha$ for opening angle α and positive integer m. For the first-quadrant region $\alpha = \pi/2$, the leading generic term is $r^2 \sin(2\theta)$. Its field tends to zero linearly with distance from the vertex. Expanding the explicit bottom-face density above around $x = 0$ indeed gives a term proportional to x, providing a local check on the image construction.

It is easy to transfer an intuition about sharp metal tips to the wrong side of the metal. A conducting projection surrounded by a vacuum angle greater than π has a leading exponent below unity. Differentiating its potential then gives a field that can diverge as the ideal corner is approached. The inside corner of the two grounded planes in this problem has the opposite local behaviour. Mesh refinement near either corner should be informed by which opening angle belongs to the field domain.

Infer what two electrode meters can measure

The induced vertical charge accounts for exactly one quarter of the total magnitude. A second experiment scales both source coordinates by a common factor of two. Determine the locus of the original source, the new charge split and the force ratio. Which observation could distinguish the two radii?

▼ **Derive the result and check it independently**

The auxiliary potential gives

$$\theta = \pi/8 = 22.5^\circ.$$

Hence $b/a = \tan(\pi/8) = \sqrt{2} - 1$ and any distance along that ray gives the same ideal charge split. Doubling a, b preserves that ratio. In the force formula every term scales inversely with length, so both force components halve at unchanged λ. A force measurement or a resolved surface charge profile can distinguish the distances. Integrating each face first discarded that spatial information. In a numerical solution, compare both electrode integrals and the local force; matching a global charge split alone leaves the distance scale untested.

Griffiths, 5th ed., §§3.1–3.2 · uniqueness and image methods

Trade unknown surface charge for known potential

A point charge held above a large grounded metal plane attracts charge towards the region beneath it. The induced surface density is nonuniform and initially unknown. Direct integration of Coulomb's law would therefore introduce the very quantity one wants to determine. In a boundary-value formulation the potential of the metal is known instead. Assume a perfectly conducting plane at $z = 0$, vacuum above it and a positive point charge q at $z = a > 0$. Grounding fixes the plane at zero potential and allows charge to enter or leave it.

Away from the point source the potential obeys Laplace's equation, and at the charge its singular behaviour is the Coulomb potential. These properties with the boundary values determine a unique solution. To see why, subtract two candidate solutions with the same charges and Dirichlet boundary conditions. Their difference w has zero Laplacian and vanishes on the boundary. Integrating $\nabla \cdot (w \nabla w) = |\nabla w|^2 + w \nabla^2 w$ over the domain gives zero net surface term, including the limit at infinity for appropriately decaying fields. Hence the integral of $|\nabla w|^2$ vanishes and $w = 0$ throughout.

Construct the image

Put a fictitious charge $-q$ at $z = -a$. At any point on the plane its distance equals the distance to the actual charge, so the two potentials cancel pointwise. In the upper half-space the fictitious source lies outside the solution domain. The combined potential has the correct single singularity, is zero on the metal and decays at infinity. By uniqueness it is the upper-half-space solution. With s the horizontal distance from the projection of the charge,

$$V(s, z) = \frac{q}{4\pi\epsilon_0} \left[\frac{1}{\sqrt{s^2 + (z - a)^2}} - \frac{1}{\sqrt{s^2 + (z + a)^2}} \right].$$

It is important to keep the domain in view. The expression continued below $z = 0$ describes an auxiliary mathematical problem. Inside the real conductor electrostatic equilibrium gives a constant potential and zero field. The fictitious charge is a way to reproduce the valid exterior data, rather than a literal localisation of the induced charge.

Recover the induced charge and force

The field just outside a conducting surface has no tangential component. Its outward normal component is σ_s/ϵ_0 because the field inside is zero. Differentiate the potential with respect to height and set $z = 0^+$ to obtain

$$\sigma_s(s) = -qa/[2\pi(s^2 + a^2)^{3/2}].$$

Integrating over rings $2\pi s\,ds$ from zero to infinity gives total induced charge $-q$. A nearby grounded conductor connected to a reservoir can supply that total charge even though its potential remains fixed.

At the position of the real charge its own singular field is excluded when calculating force. The induced field equals that of the image at distance $2a$, so the attraction has magnitude $q^2/(16\pi\epsilon_0 a^2)$. Integrating the force from infinity gives interaction energy $-q^2/(16\pi\epsilon_0 a)$. Simply multiplying the image potential at the charge by q would double the magnitude because the induced distribution adjusts as the charge moves. The work integral automatically includes the changing response.

Quantify the localisation

Take $q = 1.0$ nC and $a = 20$ mm. The force magnitude is 5.62 µN. The surface density directly beneath the charge is $-q/(2\pi a^2) = -0.398$ µC/m². Within a disk of radius S the integral gives $-q[1 - a/\sqrt{S^2 + a^2}]$. Setting the fraction in brackets equal to one half yields

$$S = \sqrt{3}\,a = 34.6\,\mathrm{mm}.$$

The density peaks directly below the charge, but a disk wider than its height is needed to collect half the total.

Bring the charge twice as close while keeping q fixed. Both the peak density and the attraction quadruple, while the radius containing half the induced charge halves. This spatial prediction is sensitive to finite plate size. A real plate with a radius comparable to a cannot reproduce the infinite-plane surface distribution far from the centre, even when the potential at its surface is tightly controlled.

Only the half-space accessible to the physical charge enters the field-energy comparison. Extending the mathematical solution through the grounded boundary introduces a second singular source whose mechanical motion has no counterpart in the laboratory. This is especially important when computing energy by moving the real charge toward the plate.

How a patterned electrode is smoothed by distance

A grounded plane is one boundary condition. Another laboratory problem specifies a different potential at each position along a plane, using many narrow electrodes separated by small insulating gaps. Model the ideal limit by a prescribed function $V(x, 0)$ in vacuum occupying $z > 0$, uniform along y. Away from charges electrostatic potential obeys $\partial_x^2 V + \partial_z^2 V = 0$. First impose $V(x, 0) = V_k \cos(kx)$ with $k > 0$ and require decay at large z. Separability suggests $V(x, z) = V_k \cos(kx) Z(z)$. Then

$$Z'' - k^2 Z = 0.$$

Discard the growing exponential and set $Z(0) = 1$.

$$V(x, z) = V_k e^{-kz} \cos(kx),$$

$$E_z = k V_k e^{-kz} \cos(kx).$$

Each sinusoidal boundary component is transmitted with a factor e^{-kz}. By linearity a general periodic pattern decomposes into such harmonics, each decaying according to its own wavenumber. The constant component fixes the voltage reference or background field separately. Smoothing with height follows directly from Laplace's equation in a source-free half-space. It places a physical limit on how well a distant probe can resolve narrow electrode features even before instrument noise is considered.

Predict a non-sinusoidal surface measurement

Let a fabricated pattern have period $d = 4.00$ mm with a fundamental voltage amplitude of 2.00 V and a third harmonic of 0.600 V, both cosine waves with maxima at $x = 0$. At height $z = 0.500$ mm the attenuation exponents are

$$(2\pi/d)z = 0.785,$$

$$3(2\pi/d)z = 2.356.$$

Hence the voltage amplitudes there are

$$2.00 e^{-0.785} = 0.912\,\mathrm{V},$$

$$0.600 e^{-2.356} = 0.0569\,\mathrm{V}.$$

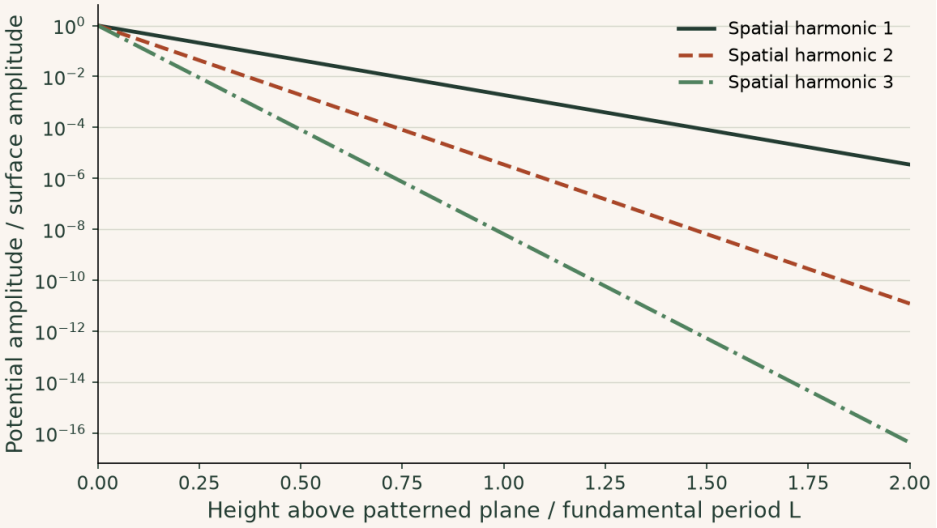
The third-to-first amplitude ratio has fallen from 0.300 at the metal to 0.0624.

A probe sensitive instead to the normal field differentiates in z and multiplies each voltage component by its wavenumber. At the same height the fundamental contributes 1.43×10^3 V/m and the third harmonic contributes 268 V/m at the shared maximum. Reporting a field ratio as though it were a voltage ratio would therefore exaggerate the surviving high spatial frequencies by a factor of three. Inverting the voltage law to recover the fabrication pattern multiplies measurement noise by e^{kz}. Repeating a scan at two known heights allows a direct test: the ratio for each Fourier component should follow its own exponential independently of the unknown electrode voltage. A ground plane placed nearby on the other side would change the far boundary condition and require a different transfer factor.

For the same apparatus, suppose detector noise permits relative voltage harmonics down to one percent of their surface amplitudes. A harmonic remains individually detectable only if $e^{-kz} > 0.01$. At a height of 0.500 mm this requires spatial wavelength greater than $2\pi z / \ln 100 = 0.682$ mm at that noise level. The limit is a measurement criterion, not a strict cutoff in Laplace's equation: every finite wavelength has a nonzero transfer at finite height. An improved signal-to-noise ratio extends the accessible range only logarithmically. Reducing the probe height can therefore improve recovery more effectively than merely accumulating many additional scans.

Which electrode details reach a remote probe?

Set the potential on an infinite plane to a spatial sinusoid of wavelength L/n, with no free charge above it. Separation in Laplace's equation gives an exponentially decaying height envelope $e^{-2\pi n y/L}$ for the bounded solution.



At a height L/2 the first harmonic retains 4.32% of its surface amplitude, whereas the third retains about 0.0081%. Trying to infer fine surface detail from a distant measurement therefore amplifies even a small voltage error.

Try it · Boundary harmonics

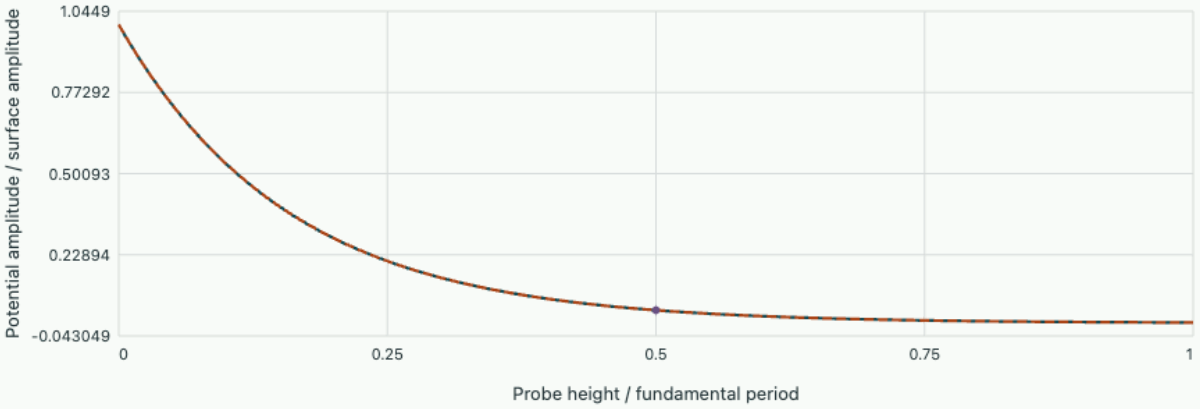
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Spatial harmonic number **1**

Selected probe height / fundamental period **0.5**



Amplitude at the probe 0.043214. Recovering the boundary coefficient multiplies amplitude errors by 23.141.

▶ **Read values from the plot**

Develop this calculation further

Chapter 14 · Building a potential from point responses · Chapter 15 · How far does an electrode pattern reach? · Chapter 16 · Spherical harmonics as measurements of boundary voltage

Study the full derivation

When do boundary measurements determine a unique electrostatic field? · Electrostatic field theory and measurement

A boundary-value solver may converge while solving the wrong electrical experiment. Establishing which potentials, charges and fluxes are fixed gives a uniqueness proof and a way to turn an approximate field into an upper bound.

Problem sheet 1 · Fields from charge and metal boundaries

Test the symmetry and boundary-value arguments of Lectures 1 and 2 by predicting fields, energies and induced charges.

Hollow out a charged sphere

An insulating sphere of radius $2a$ contains uniform charge density ρ except in a concentric empty cavity of radius a. Derive the radial field in all three regions. Match it at the two boundaries and calculate the potential difference between the cavity and the outer surface for $\rho = 1.0$ µC/m³ and $a = 0.10$ m.

▼ **Enclose the remaining volume charge**

In the cavity the concentric Gaussian sphere encloses zero charge and spherical symmetry gives zero field. In the charged material the enclosed charge is $4\pi\rho(r^3 - a^3)/3$, so $E_r = \rho(r - a^3/r^2)/(3\epsilon_0)$. Outside it becomes $7\rho a^3/(3\epsilon_0 r^2)$. The two expressions agree at $2a$ and the material field tends to zero at a. Integrating outward gives $V(a) - V(2a) = \rho a^2/(3\epsilon_0)$, about 376 V. An off-centre cavity would break the symmetry used in this proof.

Place a grounded plane below a charge

A positive 4.0 nC charge sits 50 mm above a large grounded plate. Find the force magnitude and the radius collecting 80% of the negative induced charge. How do both change if the charge doubles at fixed height?

▼ **Use the image and integrate the surface charge**

The point charge interacts with an image of magnitude 4.0 nC at separation 0.100 m. Its attractive force is $kq^2/(2a)^2 = 14.4$ µN. The fraction within a disk is $1 - a/\sqrt{S^2 + a^2}$. Requiring 0.8 gives

$$S = \sqrt{24}\, a = 0.245\,\mathrm{m}.$$

Doubling the charge leaves the fractional distribution and this disk radius unchanged, doubles the density at every point and quadruples the force. The comparatively large disk helps explain why finite plates can distort measurements of total induced charge.

Recover a potential in a charge-free gap

Two coaxial conducting cylinders of radii 5 and 20 mm are held at 100 and 0 V in vacuum. Neglect end effects. Solve Laplace’s equation in the gap, predict the reading at 10 mm, and find how far from the inner wall a probe must move to see half the inner-wall field.

▼ **Integrate the radial Laplacian**

With cylindrical symmetry Laplace's equation is $(rV')' = 0$. One integration gives $rV' = \text{constant}$ and the second gives a logarithm. Matching the wall potentials yields $V = 100\ln(20\,\text{mm}/r)/\ln 4$ V and $E_r = 100/(r\ln 4)$ V/m with r in metres. At 10 mm the potential is 50 V. Since the field is proportional to $1/r$, moving from 5 to 10 mm, a distance of 5 mm, halves the field. The halfway-potential position is the geometric mean of the two radii.

Explore these relations

- For the grounded plane, vary charge height and inspect the image-charge field and force. Open the live calculation.
- For the charge-free gap, combine harmonic boundary data and inspect their decay into the interior. Open the live calculation.

Chapter 3 · Magnetic flux and material response

Predict how a narrow air gap changes a toroidal coil's flux, stored energy and inductance from Ampère's law and the magnetic boundary conditions.

FROM FIRST PRINCIPLES

How turns, permeability and a gap set the flux

Wind N turns around a closed magnetic core and drive a steady current I. Use a core of length ℓ_c and constant cross-section A, interrupted by an air gap of length g. Assume a linear core of permeability μ_c, a nearly uniform field through each cross-section and negligible leakage and fringing. Ampère's law counts the linked current along a loop. Flux continuity through a thin pillbox at the interface gives a common normal B:

$$H_c\ell_c + H_g g = NI,$$
$$B = \mu_c H_c = \mu_0 H_g.$$

The material relation $B = \mu H$ is experimental input. Insert each material's field into the loop balance, then multiply by area to obtain the flux $\Phi = BA$:

$$\Phi\left(\frac{\ell_c}{\mu_c A} + \frac{g}{\mu_0 A}\right) = NI.$$

The bracket is the reluctance. Each turn links the same flux, so the flux linkage is $N\Phi$. For a linear core the coil inductance follows from flux linkage per current:

$$L = \frac{N\Phi}{I} = \frac{N^2 A}{\ell_c/\mu_c + g/\mu_0}.$$

A gap can control L even when it occupies little of the magnetic path. Compare g/μ_0 with ℓ_c/μ_c before neglecting it. If the core saturates, the field-to-current ratio changes with the operating current.

DERIVATION AND APPLICATIONS

Derive the inductance measured by a small alternating signal

A winding on a biased core can give different answers to two seemingly ordinary inductance measurements. Consider a narrow toroid with N turns, uniform cross-section A, core path length ℓ and one short air gap g. Assume negligible leakage, fringing and hysteresis. Allow the core law $H_c(B)$ to be nonlinear but single-valued. Flux conservation makes the induction B common to gap and core, while Ampère's law adds their distinct field integrals.

$$NI = \ell H_c(B) + \frac{g}{\mu_0} B,$$

$$\Lambda = NAB.$$

Here Λ is flux linkage. A slowly varying current superposed on a constant bias probes only a small interval of this curve. Differentiate the circulation equation about the operating induction B_0 and define the differential permeability $\mu_d = (dH_c/dB)^{-1}$ evaluated there.

$$N\,dI = \left(\frac{\ell}{\mu_d} + \frac{g}{\mu_0}\right) dB,$$

$$L_d = \frac{d\Lambda}{dI} = \frac{N^2 A}{\ell/\mu_d + g/\mu_0}.$$

The induced terminal voltage after subtracting winding resistance is $d\Lambda/dt$. Thus the locally measured inductance is the slope. A DC current and a flux measurement instead give the chord joining the bias point to the origin. Writing $\mu_s = B_0/H_c(B_0)$ makes that expression

$$L_s = \frac{\Lambda_0}{I_0} = \frac{N^2 A}{\ell/\mu_s + g/\mu_0}.$$

These definitions coincide for a straight constitutive curve through the origin. A large air gap also makes the two estimates converge because the same linear air term then dominates each denominator.

WORKED EXAMPLE

Predict the response to a sinusoidal probe

Change the inputs and investigate this derivation.

Take $N = 100$, $A = 100$ mm², $\ell = 100$ mm and $g = 0.500$ mm. At the selected bias suppose $\mu_d = 200\mu_0$ and $\mu_s = 1000\mu_0$. The incremental effective gap is $0.500 + 100/200 = 1.00$ mm, giving $L_d = 1.26$ mH. The chord yields 2.09 mH. A 1.00 kHz current of 10.0 mA peak therefore produces a reactive voltage of $2\pi f L_d I = 79.0$ mV peak, provided the excursion stays within the nearly constant local slope.

SECOND APPLICATION

Infer the material slope from a voltage ramp

Change the inputs and investigate this derivation.

A second winding has $N = 80$, $A = 150$ mm², $\ell = 120$ mm and $g = 0.200$ mm. An inductive voltage of 96.0 mV during a 40.0 A/s ramp gives $L_d = 2.40$ mH. Reversing the slope formula yields $\ell/(\mu_d/\mu_0) = \mu_0 N^2 A/L_d - g = 0.303$ mm, so $\mu_d/\mu_0 \simeq 396$. The positive residual is a check on subtracting the gap contribution before attributing the measured slope to the core.

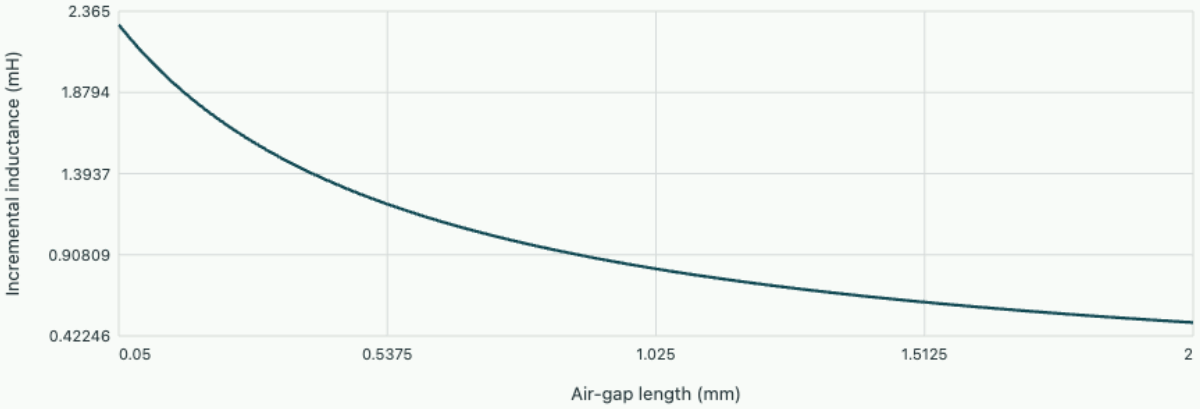
Try it · Differential inductance

Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Differential relative permeability	200
Core length (mm)	100



Incremental magnetic circuit

N=100 turns and A=100 mm². Differential permeability is the local slope dB/dH at the bias; substituting a secant B/H gives a different inductance.

When does improving an iron core stop helping?

A narrow magnetic circuit has 20 cm of linear core material and a single adjustable air gap. The two regions have equal area. Neglect leakage and fringing. Predict whether buying material with twice the permeability can double the useful gap field.

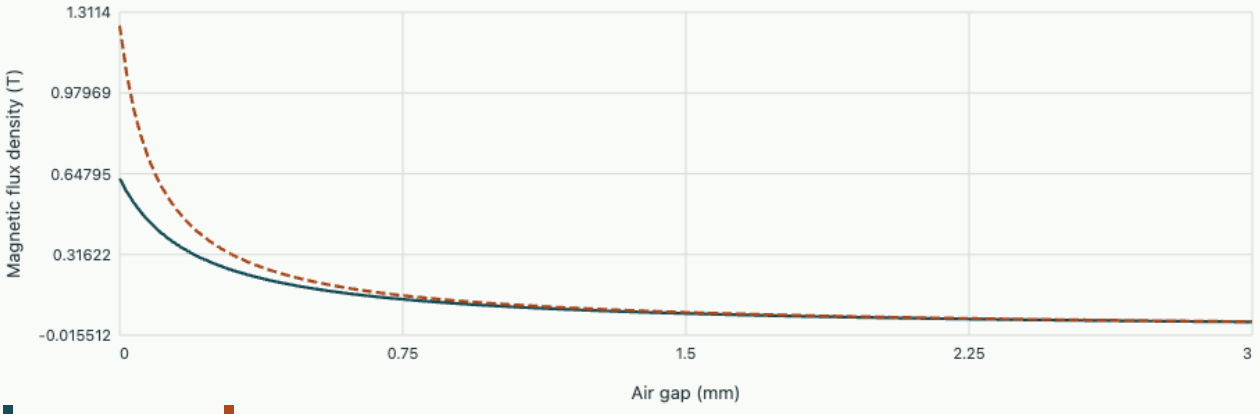
Ampère's circuital law gives $NI = B\ell_c/(\mu_0\mu_r) + Bg/\mu_0$, because the normal flux density is continuous and the area is constant. At zero gap the first term dominates. At $g = \ell_c/\mu_r$ the two drops are equal. Notice that a millimetre of air can cost more magnetomotive force than the entire core.

Open this experiment online

Try the calculation

Relative permeability **1000**

Magnetomotive force NI (A turns) **100**



At 1 mm the field is 0.1047 T. Equal core and gap reluctances occur at 0.200 mm.

Read the two curves at 1 mm for $\mu_r = 1000$. Explain why a change in gap width can matter more than a large change in core material. The ideal zero-gap end may exceed the range of a physically linear core.

A thick toroidal core carries different flux densities at its inner and outer edges

A closed toroidal core occupies inner radius a, outer radius b and axial height h, with rectangular cross section. An ideal closely distributed N-turn winding carries steady current I. Assume negligible leakage outside the core, an axisymmetric toroidal field and a linear isotropic reversible constitutive law at first. A technician estimates the inductance by putting the full cross-sectional area on a circle halfway between the two radii. The estimate is convenient. The underlying field problem tells us where it is accurate, how early the inner material starts to saturate, and which region contributes most strongly to a measured flux linkage.

Use the correct circulation at each radius

For every circular path entirely within the core, Ampère's law for $\mathbf{H}$ encloses the same free current NI, independent of its radius. Cylindrical symmetry makes the tangential field constant on that particular circle. Hence $H_\phi(r) = NI/(2\pi r)$. The inner material experiences the larger magnetising field. It is inappropriate to impose a uniform H across the full radial thickness simply because a single value of current feeds the winding.

Each radial strip of width dr presents area $h\,dr$ normal to the flux. For uniform permeability μ, adding these strips gives the flux of one turn and linkage of the whole winding:

$$\Phi = \frac{\mu N I h}{2\pi}\ln(b/a),$$

$$L = \frac{N\Phi}{I} = \frac{\mu N^2 h}{2\pi}\ln(b/a).$$

Contrast this with the arithmetic mean radius $r_m = (a+b)/2$, which predicts $L_m = \mu N^2 h(b-a)/(2\pi r_m)$. Their ratio is $\ln(b/a)/[(b-a)/r_m]$. For a narrow core write $a = r_m(1-\delta), b = r_m(1+\delta)$. Expanding the logarithm gives

$$L/L_m = 1 + \delta^2/3 + \delta^4/5 + \cdots.$$

The leading error is quadratic in fractional half-thickness. The symmetry of the cross section cancels the first-order error around its midpoint while leaving the inner edge particularly important for local field limits.

Test inductance using the energy instead of flux linkage

The magnetic energy density in a reversible linear material is $BH/2$. The volume of one radial shell is $2\pi r h\,dr$. Insert the local H, integrate and obtain

$$W_m = \mu N^2 I^2 h \ln(b/a)/(4\pi) = LI^2/2.$$

The radial integral includes $1/r^2$ from squared field and r from the volume. Missing either yields the wrong radial weighting. A magnetostatic finite-element result can therefore be checked by both an integral of flux over a radial cross section and an integral of stored energy over the full material.

For $a = 10$ mm, $b = 30$ mm, $h = 8.0$ mm, $N = 80$ and $\mu_r = 1000$, the ideal small-signal inductance is 11.25 mH. The mean-radius estimate is 10.24 mH, about 9% below the exact result. At $I = 50$ mA the inner field is $B(a) = 0.0800$ T while the outer field is 0.0267 T. A fluxmeter reports the cross-sectional average, here 0.0439 T. It cannot on its own certify that the inner material stays below a specified knee on its magnetisation curve.

Put different materials at different radii

Suppose the radial region $a < r < s$ has permeability μ_1 and the outer region $s < r < b$ has μ_2, with neither material carrying a free interface current. The azimuthal $\mathbf{H}$ is tangential to the material interface and remains continuous there as required. The tangential B may jump. Each material carries a parallel contribution to total flux:

$$L = \frac{N^2 h}{2\pi}\left[\mu_1 \ln(s/a) + \mu_2 \ln(b/s)\right].$$

The layers sit side by side across the cross section. Adding their reluctances in series would describe a different geometry, in which the same flux crosses successive materials along its path. For a uniform material the radius that divides the flux equally solves $\ln(s/a) = \ln(b/s)$, so $s = \sqrt{ab}$. In the 10–30 mm example this is 17.3 mm; half the flux occupies only 7.3 mm of the total radial thickness. Swapping a better and a poorer material between two equal-width radial bands therefore changes the measured inductance.

How a radial average can reveal a local material curve

Introduce $X = NI/(2\pi)$. For one homogeneous reversible material the exact flux per height is

$$\Phi/h = \int_a^b B(X/r)\,dr = X\int_{X/b}^{X/a} B(H)H^{-2}\,dH.$$

A terminal flux measurement samples a whole interval of magnetising field. The endpoints both move as current changes. Differentiating the last expression yields

$$X^2\,d[\Phi/(hX)]/dX = aB(X/a) - bB(X/b).$$

Thus a family of flux readings relates the material values on the inner and outer edges. If the lower-field portion of the material curve has been calibrated independently, this relation can help infer higher-field response from the toroidal measurement.

This inverse calculation differentiates measured data, amplifying flux noise and any systematic drift. A preferable check on a proposed $B(H)$ curve is often to integrate it forward across the annulus and compare directly with the flux measurements over multiple currents. Agreement at just one bias does not constrain the full material slope. A second toroid machined from the same material with a different ratio of outer to inner radius changes the radial weighting and provides an additional test.

Identify the first loss of linearity

A measured reversible $B(H)$ curve follows the stated linear permeability until $B = 0.20$ T, after which its differential slope decreases. At what current does the first radial shell reach the knee? Derive the differential inductance above that current without assigning one uniform permeability to the whole core.

▼ **Derive the result and check it independently**

The first shell is at a. Solve $\mu NI/(2\pi a) = 0.20$ T, giving $I = 0.125$ A. Farther out the local field is still lower. In a nonlinear reversible material the exact flux linkage is $Nh\int_a^b B[NI/(2\pi r)]\,dr$. Differentiate with respect to I at fixed geometry and obtain $L_d = N^2h(2\pi)^{-1}\int_a^b [dB/dH]_{H(r)}dr/r$. The derivative falls progressively as a larger radial region samples the reduced material slope. A sharp knee in the local curve can therefore become a gradual bend in a terminal inductance measurement.

Integrate digitised $B(H)$ over successively finer radial rings and compare the local slope result with a symmetric finite difference of the total flux linkage. Disagreement can indicate inadequate radial resolution right where the inner shell first enters the knee. If the material has hysteresis, specify the branch and excitation protocol before treating a measured small-signal slope as a reversible derivative.

Griffiths, 5th ed., §§6.3–6.4 and 7.2.3 · magnetic response and inductance

Specify the path taken by the field

A toroidal winding wraps a ring of magnetic material. Large permeability can guide flux around the ring. Cutting a small gap into it can greatly reduce inductance. To understand this change we need to distinguish the field produced by free current from the magnetic response of the material. Write $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$, where $\mathbf{M}$ is magnetic moment per unit volume. For a linear isotropic material,

$$\mathbf{B} = \mu_0 \mu_r \mathbf{H}.$$

The approximation of a single constant μ_r requires operation below strong saturation and negligible hysteresis over the field excursion.

Treat the cross-section as narrow compared with the mean radius and assume tightly distributed turns. Let the material path length be ℓ_c, gap length g and common area A. Neglect stray flux and fringing as a first approximation when the gap is short compared with the transverse dimensions. These choices reduce a spatially varying field to one representative flux through each cross-section. They do not imply equal values of H in the two materials.

Calculate a coil field without an assumed magnetic path

Ampère's circulation law simplifies when symmetry makes the field uniform along a useful closed path. For a circular wire loop in free space, symmetry does not supply such a path through its aperture. In vacuum a steady thin-wire current instead contributes according to Biot–Savart,

$$d\mathbf{B} = (\mu_0 I/4\pi) d\boldsymbol{\ell} \times \mathbf{R}/R^3,$$

where $\mathbf{R}$ points from the wire element to the observation point. The formula assumes a complete steady circuit. For a laboratory coil the distant return leads should contribute little to the region being observed.

Take a circular loop of radius a and observe on its axis at coordinate z. All wire elements share distance $\sqrt{a^2 + z^2}$. Their transverse contributions cancel in opposite pairs. The axial component of $d\boldsymbol{\ell} \times \mathbf{R}$ is $a\, d\ell$, and $\oint d\ell = 2\pi a$. For N tightly spaced turns the sum yields

$$B_z(z) = \frac{\mu_0 N I a^2}{2(a^2 + z^2)^{3/2}}.$$

A 50-turn coil of radius 40.0 mm carrying 0.800 A produces 0.628 mT at the centre. One radius away the ratio is $2^{-3/2} = 0.354$, giving 0.222 mT. Far along the axis the field falls as the inverse cube and is proportional to the magnetic moment $NI\pi a^2$. Such a falloff is absent from a lumped magnetic circuit with a single prescribed uniform field. Which model is appropriate depends on whether a core guides most of the flux or the winding produces a field through open space.

Practise on sheet 2 · Choose the separation of a uniform-field coil pair. Further reading · University of Cambridge · David Tong, Electromagnetism · §3.2.3 Biot-Savart Law.

Apply the two circulation and flux laws

Magnetic flux continuity follows from $\nabla \cdot \mathbf{B} = 0$. With negligible leakage, the same $\Phi = BA$ crosses the core and the gap. Ampère's law around the ring gives $H_c \ell_c + H_g g = NI$ for N turns carrying current I. Substitute $H_c = B/(\mu_0 \mu_r)$ and $H_g = B/\mu_0$ to find

$$B = \frac{\mu_0 N I}{g + \ell_c/\mu_r},$$

$$\Phi = \frac{NI}{\ell_c/(\mu_0 \mu_r A) + g/(\mu_0 A)}.$$

Each denominator contribution in the flux expression is called a reluctance. The sum follows from the distinct magnetic potential drops along one flux path. A gap as short as ℓ_c/μ_r accounts for half the total circulation even though it occupies only a small fraction of the length. For a ring of appreciable radial width, the ungapped toroidal field varies as $1/r$ and flux should be integrated over the cross-section.

Link field energy to an electrical measurement

A coil's flux linkage is $N\Phi$ when each turn links the same flux. For a fixed, linear geometry $L = N\Phi/I$. The voltage needed to raise the current is $v = L\, dI/dt$ in the passive sign convention, omitting winding resistance. The power delivered to the magnetic field is vi. Integrating as the current grows gives stored energy $W = LI^2/2$. Equivalently, integrating $B^2/(2\mu)$ over each region apportions this energy between the gap and the core.

Since the same B passes through both regions but their permeabilities differ, a unit volume of gap stores μ_r times the energy of the same volume of core in this model. The ratio of total gap energy to total core energy is $g\mu_r/\ell_c$. If permeability depends on field, a small-signal inductance measured around a bias is the derivative $d(N\Phi)/dI$. It can differ from the ratio $N\Phi/I$ inferred from a single point.

Estimate a coil with a machined gap

Choose $N = 100$, $I = 0.40$ A, $\ell_c = 0.20$ m, $\mu_r = 2000$, $A = 1.0$ cm² and $g = 0.50$ mm. The effective length in the field denominator is

$$0.50 + 0.10 = 0.60\,\text{mm}.$$

Thus $B = 0.0838$ T, $\Phi = 8.38$ µWb and $L = 2.09$ mH. The gap contributes five sixths of the circulation and stores five sixths of the 0.168 mJ total energy.

Closing the gap at the same current would predict 0.503 T from the linear law. The calculation itself supplies a reason to consult the real material curve before accepting that number. A measured inductance that departs from the prediction may reflect core saturation, imperfect contact at the joint or fringing that changes the effective gap area. Measuring a small current sweep at several gap widths helps distinguish material response from the geometry of a single fixed gap.

The current-to-flux ratio and the incremental response need not agree near magnetic saturation. If λ is the linked flux, a slow voltage measurement detects $d\lambda/dt = (d\lambda/dI)\dot{I}$, so small AC perturbations about a DC operating point report the local slope. In contrast, taking a single ratio λ/I effectively joins that point to the origin. The two are equal in the linear magnetic-circuit model derived above. Deliberately sweeping the bias provides an operational way to find where that model stops describing the core.

Two windings sharing a flux path

A second coil wound on the same magnetic core allows the flux model to predict mutual induction. Choose both positive winding directions so currents create flux around the core in the same sense. In a linear core of reluctance $\mathcal{R}$ with negligible leakage, the magnetic flux is

$$\Phi = (N_1 I_1 + N_2 I_2)/\mathcal{R}.$$

Each turn links that same flux. Multiplying by the number of turns in winding 1 or 2 gives flux linkages λ_1, λ_2, hence

$$\lambda_1 = L_1 I_1 + M I_2,$$
$$\lambda_2 = M I_1 + L_2 I_2,$$
$$M = \frac{N_1 N_2}{\mathcal{R}},$$
$$L_i = \frac{N_i^2}{\mathcal{R}}.$$

Equality of the two mutual coefficients expresses reciprocity in this linear, stationary arrangement. With imperfect flux linkage one often writes $M = k_c\sqrt{L_1 L_2}$. The stored energy $W = (L_1 I_1^2 + 2M I_1 I_2 + L_2 I_2^2)/2$ must be nonnegative for arbitrary currents, requiring $|M| \le \sqrt{L_1 L_2}$. In the perfectly coupled limit opposing ampere-turns can cancel the common core flux. Leakage and winding resistance then control the small residual voltages seen in real instruments.

Open the secondary so that $I_2 = 0$. Faraday's law in the passive terminal-voltage convention gives

$$v_i = d\lambda_i/dt = N_i\, d\Phi/dt.$$

Their voltage ratio is the turns ratio, regardless of sinusoidal amplitude, provided winding drops are negligible. The input current still has to build the necessary flux. Inferring it requires the primary magnetizing inductance, which the turns ratio alone cannot supply.

Set a sinusoidal operating point

Use $N_1 = 400$, $N_2 = 100$, core area $A = 4.00$ cm² and linear reluctance $\mathcal{R} = 1.00 \times 10^6$ A/Wb. Then $L_1 = 0.160$ H, $L_2 = 0.0100$ H and $M = 0.0400$ H. Drive the primary with 8.00 V RMS at 100 Hz with secondary open. Integrating the sinusoidal voltage gives peak flux $\Phi_{\mathrm{pk}} = \sqrt{2}V_{\mathrm{rms}}/(2\pi f N_1) = 45.0$

μWb. The predicted peak flux density is 0.113 T, secondary voltage 2.00 V RMS and magnetizing current $V_{\mathrm{rms}}/(2\pi f L_1) = 79.6$ mA RMS.

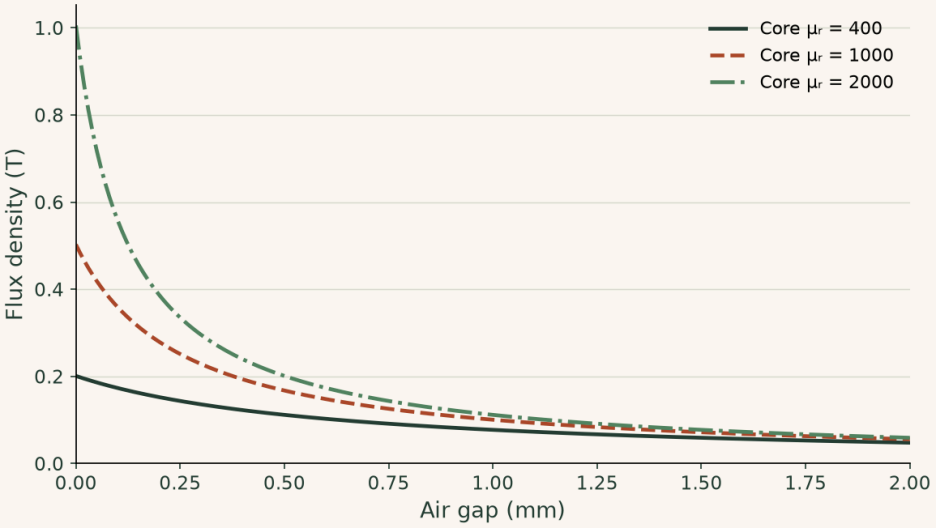
At unchanged voltage but half the frequency, voltage ratio is unchanged while peak flux density and current double in this model. A measured current waveform that develops sharp peaks at low frequency can thus signal core saturation before any dramatic change in the nominal turns ratio. A fixed DC terminal voltage on an ideal resistance-free winding would instead keep changing the flux in time, outside the assumed sinusoidal operating condition.

A short electrical pulse can determine the sense of the windings before a sinusoidal test. Mark one terminal on each coil and increase primary current through the first marked terminal. The sign of the measured open-secondary pulse with respect to its marked terminal identifies whether the two positive winding conventions reinforce or oppose one another. Reversing one winding changes the sign of the mutual coefficient and of that pulse but leaves both self-inductances and the magnitude of the voltage ratio unchanged. A sign convention overlooked here becomes important when combining two coils in series, because their mutual induced voltages can either add or subtract.

Winding resistance also introduces voltage even at steady current, so a four-wire resistance measurement helps separate ohmic voltage from linkage changes during a slow pulse.

When does an air gap dominate a magnetic circuit?

A 0.25 m core with uniform cross-section carries 100 ampere-turns. In the linear, negligible-leakage model, the core and air gap share the same B. Their magnetomotive drops add, giving $B = \mu_0 N I/(\ell/\mu_r + g)$.



At $\mu_r = 1000$ the equivalent core length in air is 0.25 mm. Adding a 0.25 mm gap halves the zero-gap flux density. Beyond that scale, improving core permeability brings diminishing returns. Fringing and saturation require separate checks.

▶ **Read values from the plot**

Develop this calculation further

Chapter 17 · From a measured current to its magnetic field · Chapter 19 · How a magnet produces an exterior field

Study the full derivation

Interfaces, demagnetising fields and the work of magnetisation · Magnetism, induction and field dynamics

Two samples made of the same material can respond very differently to an externally applied field if one is a long rod and the other is a thin disc. The induced moment depends on the internal field, which the magnetised sample helps create. The same distinction reappears when one tries to infer a force from stored magnetic energy.

Chapter 4 · Induction in moving circuits

Separate transformer and motional electromotive forces, then use the mechanical and electrical power balances to check a moving conductor.

FROM FIRST PRINCIPLES

Follow a charge around a moving loop

A straight conducting bar of length ℓ slides right at speed v on two fixed rails. Close the circuit at the left and apply a uniform magnetic field B out of the page. Orient the loop counterclockwise as seen by the reader. In a moving conductor the Lorentz force per unit charge contains both an electric field and a motional term. Its circulation defines the signed electromotive force:

$$\mathcal{E} = \oint_{C(t)} (\mathbf{E} + \mathbf{v}_{\text{wire}} \times \mathbf{B}) \cdot d\boldsymbol{\ell}.$$

In a static applied field the electric-field circulation vanishes. Only the right bar moves, and its motional force points down although positive circulation on the right edge points up. Its contribution is $-B\ell v$. Independently, a bar at distance x encloses flux $B\ell x$, so

$$\mathcal{E} = -B\ell v = -\frac{d}{dt}(B\ell x).$$

If instead the field varies while the bar moves, the fixed-path Faraday term adds $-\ell x \dot{B}$. The two mechanisms combine by the product rule:

$$\mathcal{E} = -\ell x \dot{B} - B\ell \dot{x} = -\frac{d\Phi_B}{dt}.$$

For a resistive circuit with negligible self-inductance, $I = \mathcal{E}/R$. The negative sign predicts clockwise current during expansion and an opposing magnetic force on the bar. The mechanical work required to keep it moving supplies the resistive heating.

DERIVATION AND APPLICATIONS

Close the mechanical power balance of a rotating coil

A rotating loop drives current through an inductive circuit even while the magnetic flux through the loop is reversing. Consider N tightly wound turns of area A rotating at imposed angular speed ω in a uniform static field B. Take the angle between the loop normal and the field to be $\theta = \omega t$. Use a constant total circuit resistance R and self-inductance L. Neglect brushes, friction and radiation. The external flux linkage and its time derivative give the motional source voltage

$$\Lambda_{\text{ext}} = NBA\cos\theta,$$

$$e = -\frac{d\Lambda_{\text{ext}}}{dt} = NBA\omega\sin\theta.$$

Self-induction supplies a second voltage that depends on the rate of change of current. After the initial transient dies away, the loop equation has a sinusoidal solution with peak current I_0 and lag ϕ:

$$L\dot{i} + Ri = e,$$

$$I_0 = \frac{NBA\omega}{\sqrt{R^2 + (\omega L)^2}},$$

$$\tan\phi = \frac{\omega L}{R}.$$

The field torque along the rotation coordinate is $i\,\partial\Lambda_{\text{ext}}/\partial\theta$. An external motor balances that torque, so its instantaneous power equals ei. Multiplying the loop equation by current now separates irreversible heating from reversible storage.

$$P_{\text{motor}} = ei = Ri^2 + \frac{d}{dt}\left(\frac{1}{2}Li^2\right),$$

$$\langle P_{\text{motor}}\rangle = \frac{1}{2}RI_0^2.$$

The storage term integrates to zero over a complete cycle, though for part of the cycle a falling current can return energy. The mean opposing torque magnitude follows by dividing mean mechanical power by the imposed angular speed.

WORKED EXAMPLE

Measure the mean shaft torque

Change the inputs and investigate this derivation.

A teaching generator uses $N = 100$, $A = 20.0$ cm², $B = 0.200$ T and $\omega = 100$ rad/s. It produces 4.00 V peak. For $R = 3.00$ Ω and $L = 40.0$ mH, the reactance is 4.00 Ω and the impedance magnitude is 5.00 Ω. The 0.800 A peak current dissipates 0.960 W on average. The mean shaft torque required by the electrical load is therefore $0.960/100 = 9.60$ mN m.

SECOND APPLICATION

Read an inductance from the timing of a current maximum

Change the inputs and investigate this derivation.

In a separate apparatus the current maximum follows the voltage maximum by 2.50 ms at 50.0 Hz. The lag is $2\pi(50.0)(0.00250) = \pi/4$, so $\omega L = R$. With measured total resistance 12.0 Ω this gives

$$L = 12.0/(2\pi\,50.0) = 38.2\,\text{mH}.$$

A 6.00 V rms source would then deliver $6.00^2/(2 \times 12.0) = 1.50$ W of mean heating. Using only the resistance to infer current would double this power estimate.

Try it · Generator

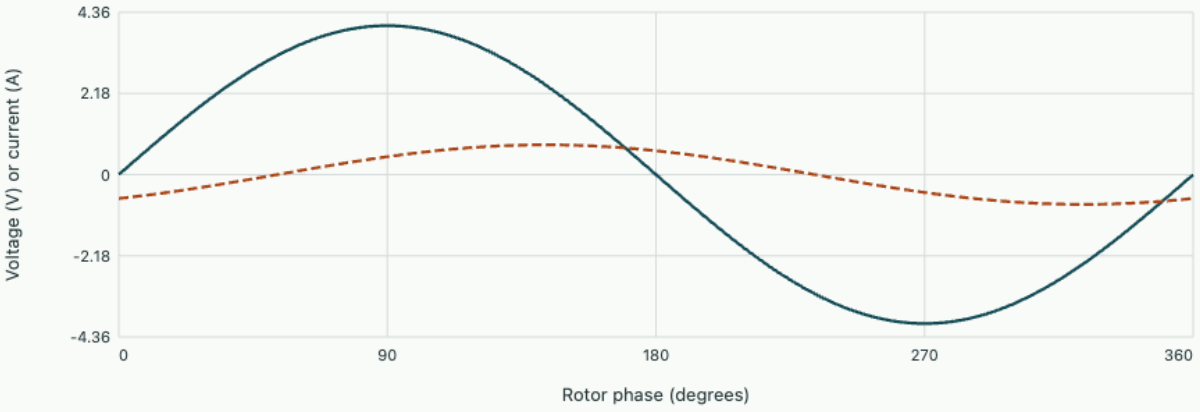
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Angular speed (rad/s) **100**

Coil inductance (mH) **40**



NAB=0.04 Wb, R=3 Ω. Phase lag 53.13°; mean shaft power 0.96 W and mean torque 0.0096 N m. Voltage and current curves carry separately named units.

Where does a moving bar's energy go?

A 40 g freely sliding conductor of active length 20 cm starts at 2 m/s. A uniform transverse field and a closed circuit turn its kinetic energy into heat. Ignore friction and circuit inductance. Predict whether doubling the magnetic field doubles the initial braking force.

The motional emf $B\ell v$ drives $I = B\ell v/R$. Lenz's law fixes the force opposite the velocity, so $m\dot{v} = -(B\ell)^2 v/R$. Separating $dv/v = -dt/\tau$ gives

$$\tau = mR/(B\ell)^2.$$

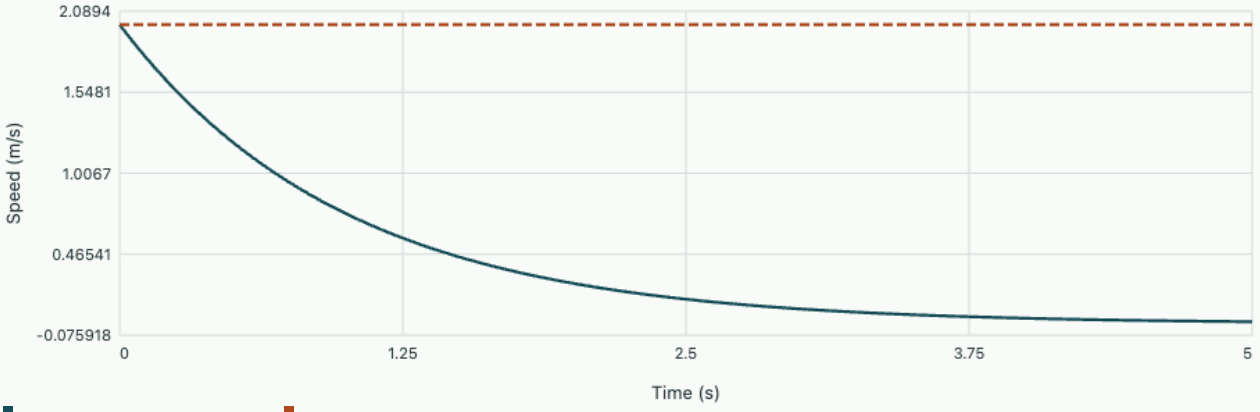
Multiplying the motion equation by v gives the independent power check $-d(mv^2/2)/dt = I^2 R$.

Open this experiment online

Try the calculation

Magnetic field (T) **0.5**

Circuit resistance (Ω) **0.25**



Decay time 1.000 s. At one decay time the speed is 0.736 m/s and 86.5% of the original kinetic energy has gone into Joule heat.

Read the time to fall from 2 m/s to 1 m/s and compare it with $\tau \ln 2$. Increase the resistance. Explain why a more resistive circuit brakes the bar less strongly.

How a spinning disk produces a voltage in an unchanging magnetic field

A conducting disk of thickness t rotates at angular velocity $\omega\hat{\mathbf{z}}$ in a uniform stationary axial field $B\hat{\mathbf{z}}$. A stationary brush contacts a circular hub at radius a, another contacts the rim at $b > a$, and a stationary lead connects them through a resistor. Model the contacts as full concentric rings so the disk current is radial and axisymmetric. The metal moves but the external circuit and the magnetic field pattern do not. A contour drawn on a sheet of paper encloses the same magnetic flux before and after one revolution. Nevertheless a voltmeter across the brushes reads a steady voltage. Which physical contour belongs in the emf calculation?

Separate the charge velocity from motion of the conductor

The Lorentz force on a charge in moving metal involves the full laboratory velocity. For motional emf around a conducting path, integrate the force per unit charge on charges carried with each local piece,

$$\mathcal{E} = \oint (\mathbf{E} + \mathbf{v}_{\text{material}} \times \mathbf{B}) \cdot d\boldsymbol{\ell}.$$

In steady operation the electrostatic contribution has zero circulation. On the disk $\mathbf{v}_{\text{material}} = \omega r\hat{\boldsymbol{\phi}}$, giving $\mathbf{v}_{\text{material}} \times \mathbf{B} = \omega r B\hat{\mathbf{r}}$; the stationary return lead contributes zero motional term. Following the path outwards through the disk and back through the load gives

$$\mathcal{E} = \int_a^b \omega B r \, dr = \tfrac{1}{2}\omega B(b^2 - a^2).$$

A material loop with every point tracked consistently obeys the moving-loop flux rule. The brush circuit continuously exchanges material at its sliding contacts; a line frozen into the disk does not stay joined to the same stationary wire. Using the derivative of flux through a fixed drawing of the apparatus omits that relative motion. This explains why the measured voltage remains steady even though the rotating charge carriers individually move in circles.

Derive the internal resistance directly from current continuity

Let disk conductivity be σ and positive current I flow radially outward inside it. Radial symmetry and charge conservation require

$$J_r = I/(2\pi r t).$$

The laboratory form of Ohm's law at nonrelativistic speed is $J_r = \sigma(E_r + \omega r B)$ for this component. Therefore $E_r = I/(2\pi\sigma t r) - \omega r B$. Integrating $-E_r$ between hub and rim gives the actual terminal voltage:

$$V_{\text{rim}} - V_{\text{hub}} = \mathcal{E} - I R_d,$$

$$R_d = \frac{\ln(b/a)}{2\pi\sigma t}.$$

At open circuit charges accumulate until $E_r = -\omega r B$, leaving zero radial conduction current and a positive rim voltage. With a load R_L, $I = \mathcal{E}/(R_L + R_d)$, neglecting contact and return-lead resistance. For a real narrow brush the spreading resistance is different; the full-ring symmetry was necessary for this logarithmic one-dimensional expression. Letting a shrink to exactly zero at finite current also makes this ideal radial resistance diverge. A real contact must have finite area.

Calculate the braking torque and test the energy balance

Current-driven charges experiencing $\mathbf{J} \times \mathbf{B}$ transmit mechanical force to the metal. Here $\hat{\mathbf{r}} \times \hat{\mathbf{z}} = -\hat{\boldsymbol{\phi}}$, so the magnetic force opposes the rotation. A shell of volume $2\pi r t \, dr$ carries torque $dT_z = -r\, J_r B (2\pi r t)\, dr = -IBr\, dr$. Integrating gives

$$T_z = -IB(b^2 - a^2)/2.$$

The positive shaft power needed to maintain constant speed is $-T_z\omega = I\mathcal{E}$, precisely the sum $I^2R_L + I^2R_d$. Independently integrating Joule heating J_r^2/σ over the disk gives $I^2 \ln(b/a)/(2\pi\sigma t)$. This independent volume calculation is sensitive to the same radial concentration of current that matters at the hub.

For $a = 5.0$ mm, $b = 50.0$ mm, $B = 0.20$ T and $\omega = 100$ rad/s, the open voltage is 24.75 mV. Let the entire internal path, including contacts, measure 20.0 mΩ and choose $R_L = 80.0$ mΩ. The current becomes 0.2475 A, the load receives 4.90 mW, and internal loss is 1.23 mW. The braking torque magnitude is 61.3 µN m, requiring 6.13 mW shaft power. Under the same geometry an open voltmeter can read the same source emf while producing essentially no magnetic braking through this external circuit.

Predict the electrical signature of a freely slowing rotor

Disconnect the motor while retaining the resistive load. Neglect bearing friction, mechanical windage and circuit inductance on the mechanical time scale. Let $g = B(b^2 - a^2)/2$, so $\mathcal{E} = g\omega$ and $T_z = -gI$. With total series resistance R the rotor's moment of inertia J obeys $J\dot{\omega} = -(g^2/R)\omega$. Its speed and open-emf scale therefore decay exponentially with time constant JR/g^2. Its kinetic energy decays twice as fast in the exponential because it is proportional to ω^2.

For the preceding disk $g = 2.475 \times 10^{-4}$ V s. If $J = 2.00 \times 10^{-6}$ kg m² and $R = 0.100\ \Omega$, the predicted speed e-folding time is 3.26 s. The initial kinetic energy at 100 rad/s is 10.0 mJ. With the 80 mΩ load present, 80% of that energy ultimately arrives at the load under the stated assumptions. A measured shorter decay even with the load disconnected points to mechanical or other electromagnetic loss. Comparing loaded and unloaded decay rates tests the braking law independently of an absolute shaft-torque measurement.

Use reversal and loading to distinguish a motional emf

Two runs at opposite angular velocity give open-circuit voltages $+27.0$ mV and -23.0 mV, with equal speed magnitudes and field. Under an 80.0 mΩ load, the speed-odd terminal voltage is 20.0 mV. Infer the motional emf, a speed-independent offset, and the effective internal resistance. Could one load reading locate where heat is generated?

▼ Derive the result and check it independently

A speed-independent contact or thermoelectric offset survives averaging, giving $(27.0 - 23.0)/2 = 2.0$ mV. The half-difference gives a motional emf of 25.0 mV. Under load $I = 20.0\ \mathrm{mV}/80.0\ \mathrm{m\Omega} = 0.250$ A. The speed-odd loss is 5.0 mV, so the inferred internal resistance is 20.0 mΩ. The measured terminal characteristic only identifies the total series loss. Separate voltage taps, a contact-resistance measurement, or a spatial heating measurement is needed to divide that loss among the disk, contacts and leads.

Reverse just B as a further test: the motional signal reverses again. If both speed and field reverse, its sign should remain unchanged. These paired observations test the bilinear ωB dependence without using an assumed calibration of the brush offset.

Griffiths, 5th ed., §§7.1.2–7.1.3 and 8.3 · emf and mechanical work

Follow a charge around a circuit

A conducting rod slides on two rails in a static magnetic field. Although a flux meter sees the enclosed flux changing, a laboratory observer finds no time-varying applied magnetic field. The measured voltage comes from the force on charges carried with the moving rod. In a stationary loop beside an AC coil the situation is different: a changing magnetic field produces an electric field with nonzero circulation. A consistent treatment includes both mechanisms.

Choose a positive direction around a thin material circuit. The associated surface normal follows the right-hand rule. If a segment moves at velocity $\mathbf{v}_{\mathrm{wire}}$, the electromotive force is the integral of $(\mathbf{E} + \mathbf{v}_{\mathrm{wire}} \times \mathbf{B}) \cdot d\boldsymbol{\ell}$ around it. In the quasistatic circuit description, the drift motion along a thin wire adds no tangential magnetic force. The integral has units of work per charge. Its sign fixes the direction of the current for a passive resistance.

Background · What does changing the route of a voltmeter lead change?

In a stationary electrostatic experiment a scalar voltage difference fixes the reading across two terminals. A time-dependent magnetic flux can make the lead routing part of the measurement. For two fixed paths from A to B define the path reading as $V_\gamma = -\int_A^B \mathbf{E} \cdot d\boldsymbol{\ell}$. Orient the comparison loop by going out on path 1 and returning on path 2. Faraday's law gives

$$V_1 - V_2 = -\oint \mathbf{E} \cdot d\boldsymbol{\ell} = \frac{d\Phi_B}{dt}.$$

A uniform normal field changing at 0.20 T/s and an enclosed area of 10 cm² give a difference of 0.20 mV. Reversing the loop orientation reverses the sign. Pairing the measurement leads closely reduces the enclosed area and therefore the inductive pickup. This estimate assumes stationary conductors and a high-impedance instrument.

Recover the same fields from different potentials

Zero magnetic divergence permits $\mathbf{B} = \nabla \times \mathbf{A}$ in a contractible region, such as the interior of a ball. Combining this with Faraday's law yields $\nabla \times (\mathbf{E} + \partial_t \mathbf{A}) = 0$ and hence $\mathbf{E} = -\nabla\phi - \partial_t \mathbf{A}$. The scalar potential alone cannot represent a nonzero electric circulation around a stationary loop. Still, individual values of the potentials are not unique. For any smooth single-valued χ, take $\mathbf{A}' = \mathbf{A} + \nabla\chi$ and $\phi' = \phi - \partial_t \chi$. The curl of a gradient vanishes, and the two new derivative terms in $\mathbf{E}'$ cancel, so both fields and the Lorentz force stay fixed.

Inside a long ideal solenoid assume $\mathbf{B} = B(t)\mathbf{e}_z$ and axial symmetry. Choose $\phi = 0$ and $\mathbf{A} = B(-y, x, 0)/2$. The resulting azimuthal electric field has magnitude $r|\dot{B}|/2$. Another allowed choice uses $\chi = Bxy/2$, which gives

$$\mathbf{A}' = (0, Bx, 0),$$

$$\phi' = -\dot{B}xy/2.$$

Computing only $-\partial_t \mathbf{A}'$ would misleadingly give a field pointing along the y axis everywhere. Including the scalar gradient restores $(\dot{B}y, -\dot{B}x, 0)/2$.

At $x = 0.0400$ m, $y = 0.0300$ m and $\dot{B} = 0.500$ T/s both calculations give $(0.00750, -0.0100, 0)$ V/m. A circle of radius 50.0 mm encloses a flux growing at 3.93 mWb/s and has circulation -3.93 mV in the positive sense around z. The measured loop emf is gauge independent. A voltage measured along a specified open path in a changing magnetic field also depends on the routing of its leads.

Practise on sheet 2 · Test a proposed change of potentials. Further reading · University of Cambridge · David Tong, Electromagnetism · §3.2.2 Gauge Transformations.

Differentiate the linked flux

Let the loop boundary move during a short interval. Its magnetic flux changes both because $\partial_t \mathbf{B}$ changes across the original surface and because the perimeter sweeps a thin ribbon. Faraday's field equation relates the first contribution to $-\oint \mathbf{E} \cdot d\boldsymbol{\ell}$. The oriented ribbon gives the second contribution $-\oint (\mathbf{v}_{\text{wire}} \times \mathbf{B}) \cdot d\boldsymbol{\ell}$. Adding them gives

$$\mathcal{E} = -d\Phi/dt.$$

The flux rule holds here because the moving contour is the circuit being followed.

Take a rectangular loop in the horizontal plane with uniform field B upwards. Its right side is a rod of length ℓ, moving outwards at speed v so the upward flux increases. Positive circulation seen from above is counterclockwise. The induced electromotive force is therefore negative and the current is clockwise. Equivalently, on the moving rod $\mathbf{v} \times \mathbf{B}$ points down the page. This local force gives the same sign as the global flux law.

Close the energy balance

For negligible self-inductance and total resistance R, the magnitude is $I = B\ell v/R$. A current-carrying rod experiences force $\mathbf{F} = I\boldsymbol{\ell} \times \mathbf{B}$, opposite to its velocity. Constant-speed operation needs a pulling force $B\ell I$. Multiplying by speed gives

$$P_{\text{mech}} = B\ell I v = I^2 R.$$

The magnetic interaction redirects the transfer. The external pull supplies the energy dissipated as heat.

If the wire has appreciable self-inductance L, its own flux changes when current changes. With constant applied speed the magnitude equation becomes $L\,dI/dt + RI = B\ell v$ after the initial step in velocity. The current builds with time constant L/R. At each instant the mechanical power balances Joule heating plus $d(LI^2/2)/dt$. Predicting an immediate full drag force at the moment of starting would neglect the energy required to establish current.

Compare two measurements

Let $B = 0.40$ T, $\ell = 0.25$ m, $v = 3.0$ m/s and $R = 0.60$ Ω. The open-circuit emf magnitude is 0.300 V. Closing the circuit gives steady current 0.500 A and drag force 0.0500 N. Mechanical power and electrical heating both equal 0.150 W. A voltmeter with large internal resistance measures nearly the open-circuit value but draws little current, so it produces little braking. The induced voltage alone does not establish how much energy is transferred.

With $L = 0.12$ H the time constant is 0.20 s. At that time after closing, $I = 0.500(1 - e^{-1}) = 0.316$ A. The mechanical power is 0.0948 W, of which 0.0599 W heats the resistor. The remaining 0.0349 W increases magnetic field energy. If the pulling force is removed entirely, speed can no longer be prescribed; the mechanical inertia must enter an additional equation. Under those conditions a single exponential computed at constant speed describes only part of the response.

A simple reversal checks all sign conventions in the experiment. Reverse the external magnetic field while keeping the prescribed velocity fixed. The induced voltage changes sign, and so does the steady current in a passive resistor. Their product in the electrical power is unchanged. The mechanical force depends on the product of current and magnetic field and therefore still opposes motion. The same circuit can operate as a motor if an independent supply forces current against its generator direction. The supply then contributes to the energy balance as a separate external port.

A changing flux can slow a moving magnet

Consider a small conducting loop of resistance R and inductance L fixed beside the path of an external magnet. Let x be the magnet's displacement and $\Phi(x)$ its flux through the loop in a specified positive direction. Its motion supplies an electromotive force $-\Phi'(x)\dot{x}$. Over a small region approximate the gradient by the constant $K = \Phi'(x)$ and take the loop's self-inductance to be independent of position. The current equation is

$$L\dot{I} + RI = -K\dot{x}.$$

It follows that a static magnet produces no sustained current in a resistive loop, even if its absolute magnetic flux is large.

Multiply the equation by I. Inductive energy changes at rate $d(LI^2/2)/dt$, and heat appears at rate RI^2. Together these equal $-KI\dot{x}$. Mechanical power supplied by the loop to the magnet is therefore $KI\dot{x}$, giving electromagnetic force $F = KI$. At sufficiently low frequency the inductive term is negligible and $F \simeq -(K^2/R)\dot{x}$. A magnetic interaction has become an effective viscous damper through the combined action of induction and resistance.

For harmonic velocity with complex amplitude u, the force amplitude is $-K^2 u/(R + i\omega L)$. Its component opposing velocity has effective damping $c_{\mathrm{em}} = K^2 R/(R^2 + \omega^2 L^2)$. The imaginary component returns stored energy during part of each cycle. Equating cycle-mean mechanical loss to mean resistor heating gives another way to verify which component really dissipates.

Compare motion below and above the electrical corner

Near a chosen working point let $K = 0.200$ Wb/m, $R = 2.00\ \Omega$, $L = 20.0$ mH and peak speed 0.100 m/s. The corner frequency in angular units is $R/L = 100$ rad/s. At $\omega = 10.0$ rad/s the induced emf amplitude is 0.0200 V and current amplitude is

$$0.0200/\sqrt{2.00^2 + 0.200^2} = 9.95\,\mathrm{mA}.$$

The average heat is $RI_{\mathrm{pk}}^2/2 = 99.0$ µW. The low-frequency approximation gives $K^2/R = 0.0200$ N s/m and predicts 100 µW, sufficiently close here.

At 200 rad/s the same velocity amplitude gives current 4.47 mA and heat 20.0 µW. The smaller loss despite unchanged peak speed comes from a greater fraction of the emf supporting inductive voltage. At fixed nonzero frequency the limit $R \to 0$ eliminates dissipation while retaining induced current. The apparent divergence of the low-frequency approximation as resistance tends to zero is resolved because its assumption $\omega L \ll R$ then fails. Measuring both force phase and heating provides independent tests of this simple localized flux model.

The approximation of constant flux gradient can itself be checked experimentally. Translate the magnet slowly through several known positions while measuring open-circuit emf and speed. Dividing the signed emf by the signed speed estimates the local gradient, without needing an absolute calibration of the magnet’s total flux. If the gradient changes sign across the motion, the induced current can have a distorted waveform even under perfectly sinusoidal displacement. The small-amplitude force response derived here should then be measured about a fixed mean position, or the full position-dependent gradient should be retained in the coupled electrical and mechanical equations.

Develop this calculation further

Chapter 20 · Where momentum waits between source and receiver · Chapter 29 · When a circuit element must be treated as a field

Study the full derivation

Motional emf and transformer emf in moving circuits · Magnetism, induction and field dynamics

Two laboratory demonstrations produce the same voltmeter trace. In one, a stationary loop sees a changing magnetic field. In the other, a conductor moves through a stationary field. The charges experience different local forces, yet under identifiable conditions both readings follow from the rate of change of magnetic flux.

Problem sheet 2 · Flux linkage and mechanical work

Combine the magnetic-circuit model of Lecture 3 with Faraday's law from Lecture 4.

Determine a gap from an inductance measurement

A 200-turn toroidal coil has cross-section 80 mm², core length 0.16 m and relative permeability 1600. A low-current test gives $L = 8.0$ mH. Infer the gap and determine the fraction of magnetic energy held there. State one reason this inference can misestimate the physical gap.

▼ **Convert the inductance to an effective length**

For uniform area and negligible leakage,

$$L = \mu_0 N^2 A/(g + \ell_c/\mu_r).$$

The inferred effective length is 0.503 mm, of which 0.100 mm is contributed by the core. The gap is about 0.403 mm and the energy fraction is

$$0.403/0.503 = 0.801.$$

Fringing gives the gap an effective area larger than the core area. Variations in actual incremental permeability would likewise bias a gap determined solely from an inductance reading.

Drop a sliding rod under gravity

Two vertical rails and their conducting top link form a closed circuit through a sliding rod of mass 0.020 kg and length 0.25 m. A uniform 0.50 T field is normal to the loop. The total resistance is 0.10 Ω. Derive the speed from rest, find the terminal speed and compare electrical power with gravitational power there. Neglect inductance and friction.

▼ **Solve the coupled force and current laws**

The current magnitude is $B\ell v/R$, giving upward drag $B^2\ell^2 v/R$. Thus $m\dot{v} = mg - B^2\ell^2 v/R$, with solution $v = v_t(1 - e^{-t/\tau})$, $v_t = mgR/(B^2\ell^2) = 1.26$ m/s and $\tau = mR/(B^2\ell^2) = 0.128$ s. At terminal speed the current is 1.57 A and heating is 0.246 W, matching mgv_t. Changing resistance changes the terminal speed; the mechanically available power per second therefore also changes.

Drive a stationary loop with a changing field

A one-turn circular loop of area 0.012 m² and resistance 3.0 Ω lies normal to $B(t) = 0.080\cos(2\pi 50 t)$ T. Find the peak emf, mean resistive power and total charge passing a reference point during a half-cycle from maximum positive to maximum negative flux. What changes if a second identical turn is added in series?

▼ **Distinguish derivative and integrated flux**

The peak emf is $AB_0\omega = 0.302$ V. With negligible inductance, mean power is

$$\mathcal{E}_0^2/(2R) = 15.2\,\mathrm{mW}.$$

Integrating $I = -\dot{\Phi}/R$ gives signed transported charge $2AB_0/R = 0.640$ mC in the selected positive direction. Two identical series turns double both the emf and the resistance. Current and half-cycle charge remain the same, while total mean power doubles. This assumes the same field threads each turn.

Choose the separation of a uniform-field coil pair

Place two identical coaxial coils at $z = \pm d/2$ with the same current orientation. Use the single-coil axial field to choose d so that both the first and second derivatives vanish at the midpoint. Find its field for 100 turns per coil, radius 0.100 m and current 0.500 A.

▼ **Differentiate the axial field twice**

Reflection symmetry makes the first derivative zero for every spacing. A single factor $(a^2 + z^2)^{-3/2}$ has second derivative $3(4z^2 - a^2)/(a^2 + z^2)^{7/2}$. Both coils contribute zero curvature at the midpoint if $d/2 = a/2$, hence $d = a$. Adding their fields gives

$$B(0) = (4/5)^{3/2}\mu_0 NI/a = 0.450\,\mathrm{mT}.$$

The first surviving midpoint variation is then fourth order in displacement.

Revisit chapter 3 · Calculate a coil field without an assumed magnetic path.

Test a proposed change of potentials

In a region of zero physical fields take $\chi = sxt$ with $s = 0.300$ V/m. Transform from zero potentials, recover the fields, and determine the false acceleration inferred for a charge-to-mass ratio of 2.00 C/kg if the transformed scalar potential is forgotten.

▼ **Differentiate both potentials**

The transformation gives

$$A'_x = st,$$

$$\phi' = -sx.$$

The spatially uniform vector potential has zero curl. Its time derivative contributes $-s\mathbf{e}_x$ to the electric field, and the scalar gradient contributes $+s\mathbf{e}_x$, giving zero acceleration. Keeping only the vector term would give -0.600 m/s². Changing representation cannot cause a particle to start accelerating.

Revisit chapter 4 · Recover the same fields from different potentials.

Explore these relations

- For the inductance inference, vary the magnetic gap and compare its reluctance with the core contribution. Open the live calculation.
- For the falling rod, vary electrical loading and compare magnetic braking with the terminal motion. Open the live calculation.

Chapter 5 · Electromagnetic waves and energy flux

Derive the electromagnetic energy balance and use it to infer field amplitudes and transmitted energy from a measured wave intensity.

FROM FIRST PRINCIPLES

Derive the local account of field energy

A material charge moving with velocity $\mathbf{v}$ gains power $q\mathbf{v}\cdot\mathbf{E}$. Summing gives the rate of work per volume $\mathbf{J}\cdot\mathbf{E}$. To find the energy reservoir supplying this work, take the scalar product of the Ampère–Maxwell law with $\mathbf{E}$ and of Faraday's law with $\mathbf{H}$. Assume linear, time-independent, nondispersive material with scalar ϵ, μ.

$$\mathbf{E}\cdot(\nabla\times\mathbf{H}) = \mathbf{E}\cdot\mathbf{J} + \epsilon\mathbf{E}\cdot\partial_t\mathbf{E},$$
$$\mathbf{H}\cdot(\nabla\times\mathbf{E}) = -\mu\mathbf{H}\cdot\partial_t\mathbf{H}.$$

Subtract the first identity from the second. The left side is the divergence of the cross product $\mathbf{E}\times\mathbf{H}$. On the right each field times its own derivative is half the derivative of its square:

$$\frac{\partial}{\partial t}\left(\frac{\epsilon E^2 + \mu H^2}{2}\right) + \nabla\cdot(\mathbf{E}\times\mathbf{H}) = -\mathbf{J}\cdot\mathbf{E}.$$

Integrate over a volume. Positive outward flux and positive work on matter both decrease the field energy remaining inside. The cross product has units W/m² and gives the direction of energy travel. This provides a way to calculate power crossing the insulation of a cable even when no conduction current crosses that surface.

Background · Poynting flux

$$\mathbf{S} = \mathbf{E}\times\mathbf{H}$$

Integrating its outward normal component gives field power leaving a closed surface. Work on matter and changing stored energy complete the balance.

DERIVATION AND APPLICATIONS

Follow the energy into a charging capacitor

A capacitor can store energy in an insulating gap while its supply leads carry conduction current. Consider circular vacuum plates of radius a separated by $d \ll a$. Use cylindrical coordinates, with a positive, nearly uniform electric field E along $+z$. Charge slowly enough to neglect radiation and propagation delays. The edge of the uniform-field model represents a surface just inside the rim. During charging, the increasing electric displacement contributes to Ampère's circulation law around any concentric circle of radius $r \le a$.

$$2\pi r H_\phi = \frac{d}{dt}\int \epsilon_0 E\, dA = \epsilon_0 \pi r^2 \dot{E},$$

$$H_\phi = \frac{\epsilon_0 r}{2}\dot{E}.$$

The magnetic field grows linearly with radius because the enclosed displacement current scales with area while the circumference scales with radius. When E and $\dot{E}$ are positive, the cross product of axial electric field and azimuthal magnetic field points radially inward. Integrating over the side of the chosen cylinder gives the incoming electromagnetic power.

$$P_{\text{in}} = -2\pi r d\, S_r = 2\pi r d\, E H_\phi = \epsilon_0 \pi r^2 d\, E\dot{E}.$$

Express the electric energy in this same cylindrical part of the gap through its volume. Its time derivative agrees with the surface result at every radius. At the rim, the voltage and total charging current give a third way of measuring this rate.

$$\frac{d}{dt}\left(\frac{\epsilon_0 E^2}{2}\pi r^2 d\right) = P_{\text{in}},$$

$$P_{\text{in}}(a) = VI,$$

$$V = Ed,$$

$$I = \epsilon_0 \pi a^2 \dot{E}.$$

WORKED EXAMPLE

Trace a voltage ramp

Change the inputs and investigate this derivation.

For $a = 40.0$ mm and $d = 1.00$ mm, the capacitance is

$$\epsilon_0 \pi a^2/d = 44.5\,\text{pF}.$$

A voltage ramp of 2.00 V/µs draws 89.0 µA. At 10.0 V, electrical input is 0.890 mW. The inward flux density at the rim is then

$$0.000890/(2\pi\, 0.0400\, 0.00100) = 3.54\,\text{W/m}^2.$$

A cylinder of half the radius receives one quarter of the power because it encloses one quarter of the changing field energy.

SECOND APPLICATION

Check a discharge with a local magnetic measurement

Change the inputs and investigate this derivation.

For a separate capacitor, $d = 2.00$ mm and $E = 3.00$ kV/m. A probe at $r = 20.0$ mm reads $H_\phi = -2.00$ mA/m. The negative sign places the Poynting vector outwards. Its magnitude is $3000(0.00200) = 6.00$ W/m², so the energy leaves that cylinder at $2\pi(0.0200)(0.00200)(6.00) = 1.51$ mW. The axial voltage is still positive. The outward flux therefore identifies a decreasing field magnitude rather than a reversal of plate polarity.

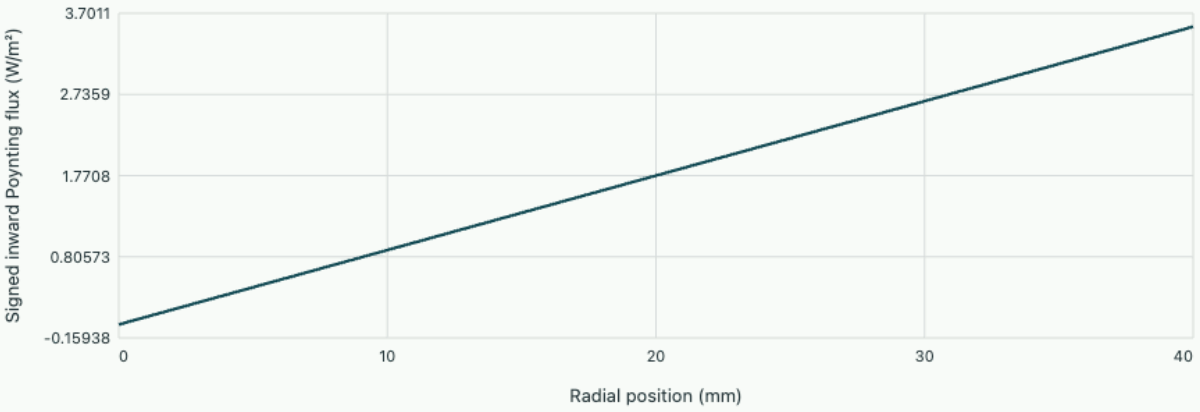
Try it · Capacitor flux

Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Voltage (V)	10
Voltage ramp (V/µs)	2
Plate gap (mm)	1



Inward energy flux

40 mm plate radius. Electrical input VI = 0.00089012 W. A negative ramp reverses energy flow at fixed positive voltage.

Start from this chapter. At fixed applied voltage, doubling the stored electric energy doubles capacitance. Use the energy and charge estimates from the same cable cross-section as two

> independent checks.
>
> OneSimulation laboratory · Recover cable impedance from a cross-section

Can a standing field contain energy without transporting it?

Let a normally incident plane wave meet a reflector. The electric-amplitude reflection coefficient r is real and may be negative. At perfect reflection, first predict the electric energy and the net power at the boundary. Vary the observation phase through a cycle.

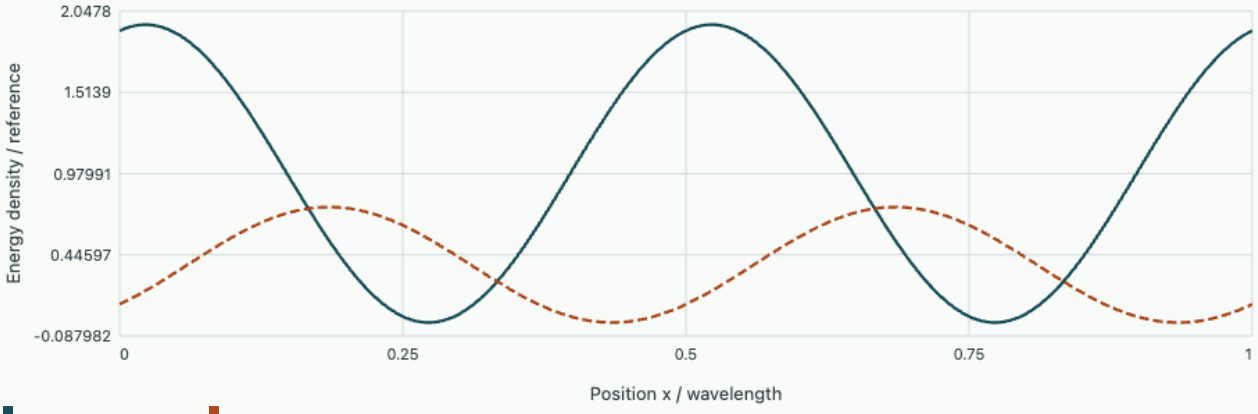
A right-going wave obeys $ZH_i = E_i$. A left-going wave obeys $ZH_r = -E_r$. In units of $\epsilon E_0^2/2$, the electric and magnetic energy densities are therefore the squares of the sum and difference of the two scalar waveforms. In their product the cross terms cancel, giving cycle-averaged flux proportional to $1 - r^2$.

Open this experiment online

Try the calculation

Reflected electric amplitude / incident **0.6**

Phase of incident wave (degrees) **30**



The cycle-averaged net power is 64.0% of the incident power. Reversing the sign of reflection swaps the pattern at the boundary.

Try $r = -1$ and zero phase. Locate a magnetic maximum where the electric field is zero. Advance by a quarter cycle and trace where the field energy has gone.

Follow energy through the bright and dark fringes of two crossing waves

Two coherent monochromatic plane waves of equal frequency and equal electric amplitude E_0 travel in vacuum. Their wavevectors make angles $+\theta$, $-\theta$ with the z axis in the xz plane. Both electric fields point

along y, and the waves are in phase at the origin. An intensity camera on the plane normal to z sees alternating bright and dark bands. Is a dark band devoid of electromagnetic energy? Inspecting the electric field alone would suggest yes. The magnetic field, Poynting flux and the observer's integration area give three complementary tests of that answer.

Construct electric and magnetic fields from the same phases

Write $K = k\sin\theta$, $k_z = k\cos\theta$, $k = \omega/c$, and use complex amplitudes multiplying $e^{-i\omega t}$. Then the two phases are $e^{ik_z z}e^{\pm iKx}$. Their electric fields add to $E_y = 2E_0 e^{ik_z z}\cos Kx$. For each wave separately $c\mathbf{B} = \hat{\mathbf{k}} \times \mathbf{E}$. Its x components add with the same sign, while the z components have opposite signs. The total fields are therefore

$$cB_x = -2E_0\cos\theta\, e^{ik_z z}\cos Kx,$$

$$cB_z = 2iE_0\sin\theta\, e^{ik_z z}\sin Kx.$$

At an electric node $\cos Kx = 0$ the magnetic field still oscillates in the longitudinal direction. Near the nodes it cannot be reconstructed by replacing the whole superposition with one plane wave propagating in $+z$. Even though both constituents have $E = cB$, that equality of pointwise magnitudes is not inherited by their sum. The phase i in the longitudinal magnetic term matters when calculating time-averaged transport.

Compute what the camera and a local energy measurement see

For complex peak-amplitude phasors, time-averaged flux is $\langle \mathbf{S} \rangle = \mathrm{Re}(\mathbf{E} \times \mathbf{B}^*)/(2\mu_0)$, and mean energy density is

$$\langle u \rangle = \epsilon_0(|E|^2 + c^2|B|^2)/4.$$

Substitute the fields. The x flux contains a purely imaginary product between E_y and B_z^*, so its cycle average vanishes. The remaining observables are

$$\langle S_z \rangle = 2\epsilon_0 c E_0^2\cos\theta\,\cos^2 Kx,$$

$$\langle u \rangle = \epsilon_0 E_0^2[(1+\cos^2\theta)\cos^2 Kx + \sin^2\theta\sin^2 Kx].$$

At a dark band $\langle S_z \rangle = 0$ but $\langle u \rangle = \epsilon_0 E_0^2\sin^2\theta$. The oscillatory magnetic field there stores positive energy. At a bright band the ratio of mean flux to mean energy density reaches $2c\cos\theta/(1+\cos^2\theta) \le c$. Between them the ratio varies with x; neither its bright-band value nor its spatial average is the velocity of an individual light ray in free space. Local instantaneous transverse energy exchange can occur even though the mean x-flux is zero.

Integrate over a complete fringe and compare with incoming power

The bright-to-bright period along the camera is $\pi/K = \lambda/(2\sin\theta)$, obtained from the squared electric field. Over a whole period both squared trigonometric terms have mean $1/2$. The area-averaged axial flux and energy density become $\epsilon_0 c E_0^2\cos\theta$ and $\epsilon_0 E_0^2$. Independently, each incident wave carries mean intensity $I_0 = \epsilon_0 c E_0^2/2$ along its own direction, whose projection normal to the camera is $I_0\cos\theta$. Their

sum exactly reproduces the integrated fringe measurement. Interference redistributes energy crossing the screen without changing this integrated power.

Take $\theta = 30^\circ$, $\lambda = 600$ nm and incident $E_0 = 100$ V/m peak per wave. The fringe spacing is 600 nm. A bright band receives mean normal power density 46.0 W/m², while integrating over one complete period gives 23.0 W/m². A dark band contains mean energy density 2.21×10^{-8} J/m³. A sufficiently small electric-dipole detector can be blind at the node even though a magnetic probe responds there. An instrument that averages over many fringe periods cannot reconstruct this local distinction from the total collected power.

Distinguish an imperfect dark band from lost energy

Change one incident electric amplitude to E_1 and the other to E_2, leaving directions and polarisation as before. Each forward field still gives a magnetic x-component proportional to its electric field. Consequently $\langle S_z \rangle = (\epsilon_0 c \cos\theta/2)[E_1^2 + E_2^2 + 2E_1E_2\cos(2Kx)]$. The intensity extrema now scale as $(E_1 + E_2)^2$ and $(E_1 - E_2)^2$. Even a perfectly resolving coherent detector will report shallow minima if the beam powers are unbalanced.

If one electric amplitude is twice the other, the peak-to-trough power ratio is nine, and the visibility is

$$2E_1E_2/(E_1^2 + E_2^2) = 0.80.$$

Independently block each beam in turn and measure its mean power. Their ratio is four, not two, because the sensor measures power. From those two blocked-beam readings predict the coherent fringe visibility before attributing any discrepancy to an unstable relative phase. Unequal amplitudes also give a nonzero area-averaged lateral Poynting flux proportional to $(E_1^2 - E_2^2)\sin\theta$, consistent with the unequal transverse components of incident power.

Distinguish visibility from coherence and from spatial resolution

Replace the ideal point detector with a rectangular pixel of width w, centred at x_c, uniformly sensitive to normally incident energy flux. Derive the contrast reduction, and find the largest pixel width giving at least 90% of ideal visibility in the first lobe.

▼ **Derive the result and check it independently**

Write $\cos^2 Kx = [1 + \cos(2Kx)]/2$. The pixel average of the oscillating term is $\cos(2Kx_c)\sin(Kw)/(Kw)$. Thus measured visibility $(S_{\max} - S_{\min})/(S_{\max} + S_{\min}) = |\sin(Kw)/(Kw)|$. In the first lobe the equation $\sin y/y = 0.90$ gives $y \simeq 0.787$, so $w \leq 0.250$ fringe periods, approximately 150 nm for these waves. A pixel covering exactly one fringe period has zero measured visibility, despite perfect mutual coherence.

As an independent test, move a narrow aperture across the pixel or change the crossing angle. The predicted visibility depends on $w\sin\theta/\lambda$. Poor coherence and unequal beam intensities can also suppress visibility, but those effects need not follow this spatial averaging law. Conservation of power and the fringe-contrast measurement test different properties of the field.

Griffiths, 5th ed., §§8.1.2 and 9.2.3 · Poynting flux and wave energy

Charge conservation requires a changing field term

Ampère's magnetostatic law relates magnetic circulation to current. A surface passing through a wire detects a conduction current, whereas a second spanning surface between the plates of a charging capacitor may intercept none. The evolving electric field supplies the missing current. In vacuum,

$$\nabla \times \mathbf{B}/\mu_0 = \mathbf{J} + \epsilon_0 \partial_t \mathbf{E}.$$

Taking the divergence of both sides and using Gauss's electric law gives $\nabla \cdot \mathbf{J} + \partial_t \rho = 0$, the continuity equation for charge.

Together with Faraday's law and the two divergence laws, this term permits waves even in a region containing no material charges. In source-free vacuum take the curl of $\nabla \times \mathbf{E} = -\partial_t \mathbf{B}$. Use $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ and substitute the Maxwell–Ampère law. The result is

$$\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 \partial_{tt} \mathbf{E}.$$

The coefficient sets $c = 1/\sqrt{\mu_0 \epsilon_0}$.

Read a propagating field

Take a sinusoidal electric field $\mathbf{E} = E_0 \cos(kz - \omega t)\,\hat{\mathbf{x}}$. The wave equation gives $\omega = ck$, and Faraday's law fixes $\mathbf{B} = (E_0/c)\cos(kz - \omega t)\,\hat{\mathbf{y}}$. The divergence equations require these plane-wave fields to be transverse. Their relative orientation selects the positive direction of travel. Flipping the sign of the magnetic field alone with the same electric waveform would reverse the local energy flow but would violate Faraday's law for this chosen propagation phase.

Wave speed and field amplitude answer different questions. A stronger sinusoidal source increases the energy in a given volume without changing its speed through linear vacuum. Changing to a lossless homogeneous nondispersive material changes the propagation speed to $1/\sqrt{\mu\epsilon}$ and the impedance relating electric and magnetic fields to $\sqrt{\mu/\epsilon}$. Near a source, the simple wave relation between amplitudes need not hold.

Identify work and transport locally

Dot the Maxwell–Ampère equation with $\mathbf{E}$ and Faraday's law with $\mathbf{B}/\mu_0$. Use the identity $\nabla \cdot (\mathbf{E} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{B})$ to combine them. Rearranging gives Poynting's theorem

$$\partial_t u + \nabla \cdot \mathbf{S} = -\mathbf{J} \cdot \mathbf{E},$$

$$u = \frac{\epsilon_0 E^2}{2} + \frac{B^2}{2\mu_0},$$

$$\mathbf{S} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0}.$$

Integrated over a volume, the equation says that net incoming electromagnetic power increases stored field energy or performs work on matter. In a vacuum plane wave the electric and magnetic energy densities are equal and $\mathbf{S} = c u \, \hat{\mathbf{z}}$. In a standing wave counterpropagating energy flows can cancel on average while local fields and stored energy remain large. A detector of heating and a point probe of electric amplitude need not then respond in the same way.

Convert intensity into field and pulse energy

A broad, normally incident vacuum wave has mean intensity 2.0 W/m². For a sinusoid the cycle mean of $\cos^2$ is one half. Therefore $\langle S \rangle = E_0^2 / (2 Z_{\mathrm{vac}})$ with $Z_{\mathrm{vac}} = \mu_0 c \simeq 377\ \Omega$. The electric peak is 38.8 V/m and the magnetic peak is 1.30×10^{-7} T. Using the RMS electric field instead gives

$$\langle S \rangle = E_{\mathrm{rms}}^2 / Z_{\mathrm{vac}}.$$

Mixing a peak-valued field with that RMS formula would overestimate the power by a factor of two.

An absorbing disk of area 0.010 m² intercepts a 3.0 ms pulse with this mean intensity throughout its flat top. Neglect edge effects and switching intervals. The energy received is $2.0 (0.010) (0.0030) = 60$ µJ. In the lossless TEM line, the same flux principle becomes the circuit power VI. The coaxial-line experiment reports $Z_0 = 47.45\ \Omega$ for the baseline cross-section. A 1 V RMS forward sinusoid there carries about 21.1 mW. A 1 V peak sinusoid carries half that mean power, which gives a direct check on the convention used by a voltage probe.

For a travelling plane-wave pulse, integrate the instantaneous Poynting flux over a plane normal to propagation and over the time during which the pulse crosses. The result is the total transported energy. Equivalently integrate field energy density throughout the pulse volume at one instant. The length of the pulse is its speed multiplied by its duration, so the two constructions agree. This provides a check when a detector reports time-averaged irradiance but an independent instrument reports energy per pulse; the averaging interval and illuminated area enter explicitly.

Which observer can see a purely magnetic field?

A charged particle responds to $q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ in one inertial frame. A frame moving alongside the particle initially sees no magnetic part of the force on it, so the assignment of a field as electric or magnetic cannot be frame independent. For a frame moving at velocity $\mathbf{u}$ relative to the laboratory, let parallel and perpendicular denote components relative to $\mathbf{u}$. The parallel fields remain unchanged. Transforming the electromagnetic field tensor gives the perpendicular components in SI units,

$$\mathbf{E}'_\perp = \gamma(\mathbf{E}_\perp + \mathbf{u}\times\mathbf{B}),$$
$$\mathbf{B}'_\perp = \gamma(\mathbf{B}_\perp - \mathbf{u}\times\mathbf{E}/c^2),$$
$$\gamma = (1-u^2/c^2)^{-1/2}.$$

For crossed fields $\mathbf{E}\cdot\mathbf{B}=0$, choose $\mathbf{u}=\mathbf{E}\times\mathbf{B}/B^2$. The vector triple product then gives

$$\mathbf{u}\times\mathbf{B} = -\mathbf{E}.$$

If $E<cB$, this is an admissible subluminal boost and the moving observer sees $\mathbf{E}'=0$. The remaining magnetic field has magnitude $\sqrt{B^2-E^2/c^2}$. A charge initially moving at $\mathbf{u}$ in the laboratory has zero Lorentz force, in agreement with a stationary charge in the new frame.

Let $\mathbf{E}=6.00$ MV/m along x and $\mathbf{B}=0.0500$ T along z. The required velocity is 1.20×10^{8} m/s along negative y, about $0.400c$. The moving frame measures $B'=0.0458$ T. At $E=cB$ the required speed would be c and no inertial rest frame eliminates either field. This is the relation for a vacuum plane wave. Its electromagnetic field cannot be removed by trying to travel beside the wave.

Practise on sheet 3 · Choose a frame for electric-dominated crossed fields. Further reading · University of Cambridge · David Tong, Electromagnetism · §5.3.2 The Electromagnetic Tensor.

Locate the energy entering a resistor

A steady resistor has constant fields and still receives a continuous supply of energy. Model a long straight cylindrical wire of radius a, conductivity σ and resistive length ℓ. A DC current I flows in the positive axial direction. Assume a uniform current density, negligible end effects and a coaxial return sufficiently far away that the magnetic field immediately outside the wire has cylindrical symmetry. The internal axial electric field is

$$E_z = I/(\sigma\pi a^2).$$

Continuity of tangential electric field gives the same axial value immediately outside its lateral surface.

Ampère's law fixes the azimuthal magnetic field there at $H_\phi = I/(2\pi a)$. Since $\mathbf{e}_z\times\mathbf{e}_\phi = -\mathbf{e}_r$, the Poynting flux points radially inward. Multiplying its magnitude by the lateral area gives

$$P_{\text{in}} = E_z H_\phi (2\pi a\ell) = \frac{I^2\ell}{\sigma\pi a^2} = I^2R.$$

The result is the familiar circuit power but obtained from fields on the surrounding surface. It accords with the local energy theorem: field energy within the chosen volume is steady while the inward flux supplies the material work $\int \mathbf{J}\cdot\mathbf{E}\,dV$. A radial electrostatic field can also be present because of surface charge establishing a voltage relative to the return. It produces longitudinal Poynting transport in the gap. Its direction and magnitude depend on the wider circuit, whereas the inward component above accounts for the resistive loss in the selected segment.

Audit the flux at different radii

Use $\ell = 1.00$ m, $a = 0.500$ mm, $\sigma = 1.00 \times 10^{6}$ S/m and $I = 1.00$ A. The cross-sectional area is 7.85×10^{-7} m² and the resistance is $1.273\ \Omega$. The axial electric field is 1.273 V/m and the surface magnetic field is 318 A/m. Their product gives an inward energy flux density of 405 W/m². Integrating over $2\pi a\ell = 0.00314$ m² yields 1.27 W, matching the measured electrical input.

Now evaluate the field on an imaginary cylinder of radius $r = a/2$ inside the wire. Uniform current density encloses only $I/4$. Its magnetic field is half the surface value, and its area is also halved. The integrated inward power is therefore one quarter, 0.318 W. It supplies precisely the heat produced in one quarter of the wire volume. As r tends to zero the local inward flux vanishes with r while volumetric heating density remains finite. Taking the divergence in cylindrical coordinates reconciles these limits. A skin-effect current profile at high frequency would invalidate the uniform-field calculation and redistribute both heating and energy input toward the surface.

Measured current and terminal voltage determine the inward energy rate without measuring either field directly, but they cannot locate the path of the incoming energy elsewhere in the circuit. To calculate that path one can choose a coaxial geometry with known conductor and return potentials. The radial field in the dielectric then combines with the azimuthal magnetic field to give longitudinal flux toward the load. Approaching the resistive segment, part of this flux turns inward as demonstrated above. Choosing a different return path may change the exterior pattern substantially while preserving the same resistance and the same total power budget.

Carry energy from source to load through a coaxial gap

The same energy theorem closes the circuit picture if the wire is the inner conductor of a coaxial delivery line. In the uniform, low-loss part take inner radius a, return-conductor inner radius b, potential difference V and equal opposing currents of magnitude I. A cylindrical Gaussian surface gives radial field $E_r = V/[r\ln(b/a)]$ in the dielectric. Ampère's law gives

$$H_\phi = I/(2\pi r).$$

The axial Poynting flux density is therefore proportional to $1/r^2$. Integrate over the dielectric annulus,

$$P_z = \int_a^b E_r H_\phi \, 2\pi r \, dr = \frac{VI}{\ln(b/a)} \int_a^b \frac{dr}{r} = VI.$$

The familiar terminal formula has emerged by integrating power moving through the dielectric. It remains compatible with power crossing the lateral surface into a resistor at the line's end. The outer return conductor also matters: it sets the radial electric field geometry and provides the return current defining where the external magnetic field cancels in the coaxial model.

Take $V = 10.0$ V, $I = 0.200$ A, $a = 1.00$ mm and $b = 4.00$ mm. Delivered power is 2.00 W. The fraction flowing between a and any intermediate radius r is $\ln(r/a)/\ln(b/a)$. At $r = 2.00$ mm this is one half. The outer part of the gap has greater cross-sectional area yet carries the other half because flux density falls outward. If the inner conductor develops finite resistance per length R' and leakage is negligible, a small

segment loses $I^2 R'\,dz$ of longitudinal power. The longitudinal difference equals its radial inward flux, providing a local field interpretation of the ordinary circuit voltage drop.

Develop this calculation further

Chapter 20 · Where momentum waits between source and receiver · Chapter 21 · What an interface does to phase, field and power · Chapter 22 · Why dispersion and absorption belong together

Study the full derivation

Energy flow and mechanical force from vacuum fields · Magnetism, induction and field dynamics

An external support holds a charged capacitor plate or a current-carrying coil in place. One can infer the support force from fields just outside the body, without summing forces on every microscopic charge. A second surface integral tracks the energy feeding the same apparatus. Both methods start with Maxwell's equations, but they transport different conserved quantities.

Chapter 6 · Radiation from a small electric dipole

Relate a time-dependent electric dipole moment to its radiation pattern, total power and the conditions under which a far-field detector measures them.

FROM FIRST PRINCIPLES

From a time-dependent current to dipole radiation

Take a localized, slowly moving source much smaller than the wavelength, in vacuum. At an observation distance r large compared with both source size and wavelength, the part of the retarded vector potential that survives at order $1/r$ samples the whole small source at approximately one retarded time $t_r = t - r/c$. Define its dipole moment $\mathbf{p} = \int \mathbf{r}'\rho\, dV'$. Charge conservation with vanishing current at the outer boundary gives, by integration by parts,

$$\dot{\mathbf{p}} = -\int \mathbf{r}'\nabla'\cdot\mathbf{J}\, dV' = \int \mathbf{J}\, dV',$$

$$\mathbf{A} \simeq \frac{\mu_0}{4\pi r}\dot{\mathbf{p}}(t_r).$$

When taking a curl far from the source, differentiating the retarded time supplies a factor $1/c$ and retains $1/r$. Differentiating the geometrical prefactor gives a smaller $1/r^2$ term. With $\hat{\mathbf{n}}$ pointing to the detector, the leading fields are

$$\mathbf{B}_{\mathrm{rad}} = \frac{\mu_0}{4\pi c r}\,\ddot{\mathbf{p}}(t_r)\times\hat{\mathbf{n}},$$

$$\mathbf{E}_{\mathrm{rad}} = c\,\mathbf{B}_{\mathrm{rad}}\times\hat{\mathbf{n}}.$$

The first cross product removes the longitudinal dipole acceleration. Thus a sinusoidal dipole has a dark radiation axis and maximum intensity broadside. The flux decreases as $1/r^2$, just cancelling the area growth of a distant sphere. A constant dipole or one increasing at a constant rate has no such dipole radiation term.

DERIVATION AND APPLICATIONS

Calculate how much dipole radiation a collector intercepts

An angular scan gives intensity along one line of sight. Turning it into collected power requires integrating the radiation pattern over the angles accepted by the optics. Assume a pointlike linearly oscillating electric dipole in vacuum, an observer many wavelengths away and a lossless collector. Its axisymmetric far-field power per unit solid angle is $K\sin^2\theta$, where K sets the strength and θ is measured from the oscillation axis. The solid angle of a narrow polar ring includes a factor of $\sin\theta$ arising from its circumference.

$$dP = 2\pi K\sin^3\theta\, d\theta,$$

$$P_{\mathrm{tot}} = 2\pi K\int_0^{\pi}\sin^3\theta\, d\theta = \frac{8\pi K}{3}.$$

Consider a collector accepting the full azimuth from the positive axis out to half-angle α. Set $u = \cos\theta$ in the integral. The term $1 - u^2$ is the angular weight and the limits in u run from $\cos\alpha$ to one.

$$f_{\mathrm{cap}} = \frac{P(0 \le \theta \le \alpha)}{P_{\mathrm{tot}}} = \frac{3}{4}\int_{\cos\alpha}^{1} (1 - u^2)\, du = \frac{2 - 3\cos\alpha + \cos^3\alpha}{4}.$$

For a collector surrounding the equator, take the band between $\pi/2 - \delta$ and $\pi/2 + \delta$. Its omitted power resides in two identical polar caps. Subtracting their contributions gives

$$f_{\mathrm{band}} = \frac{3\sin\delta - \sin^3\delta}{2}.$$

A hemisphere admits exactly one half of the power by symmetry. Near the axis the small-cap result begins as $3\alpha^4/16$ for α in radians, because both intensity and available solid angle shrink there.

WORKED EXAMPLE

Collect fluorescence through a polar cone

Change the inputs and investigate this derivation.

A fixed dipole emits 8.00 nW and an objective in vacuum accepts a polar cone with $\alpha = 60.0^\circ$. Substituting $\cos\alpha = 0.500$ gives

$$f_{\mathrm{cap}} = (2 - 1.50 + 0.125)/4 = 0.15625.$$

The incident optical power is 1.25 nW. If a downstream filter transmits 80.0 percent throughout this cone, the detector receives 1.00 nW. Randomly rotating emitters would require averaging the orientation before applying a collection factor.

SECOND APPLICATION

Infer total power from an equatorial band

Change the inputs and investigate this derivation.

An annular monitor accepts all azimuths from 60.0° to 120.0°. Here $\delta = 30.0^\circ$ and $f_{\mathrm{band}} = (1.50 - 0.125)/2 = 0.6875$. A calibrated reading of 5.50 µW implies total emitted power $5.50/0.6875 = 8.00$ µW. The band covers only half of the sphere's solid angle, so treating this dipole as isotropic would overestimate the total by 3.00 µW.

How much radiation fits inside a cone?

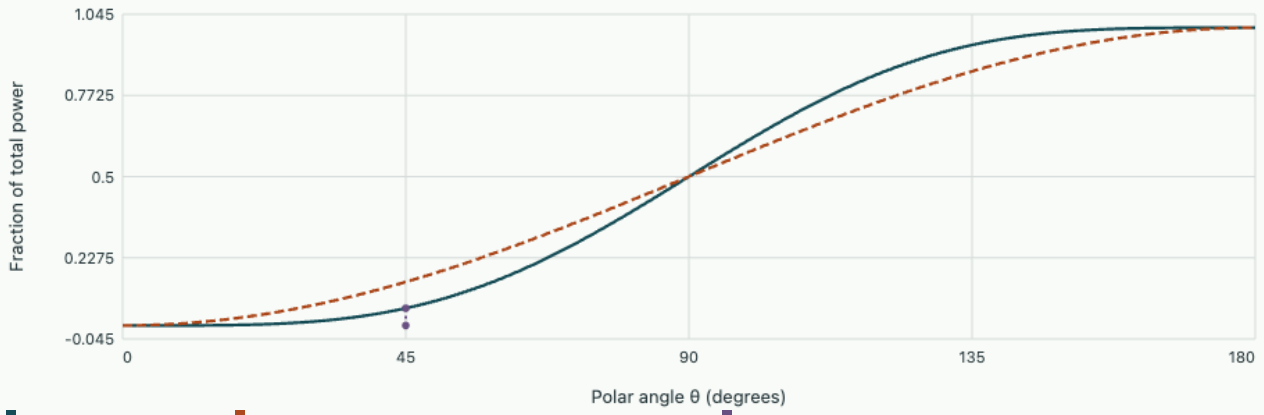
An electrically short dipole radiates a far-zone intensity proportional to $\sin^2\theta$. A detector subtends a circular cap of half-angle α about the dipole axis. Guess whether a 45-degree cap intercepts one quarter of the radiated power.

A ring on the viewing sphere has area $2\pi r^2 \sin\theta\, d\theta$. The extra sine comes from area, even for an isotropic source. Integrating $\sin^3\theta$ from 0 to α and dividing by its integral from 0 to π gives $(2 - 3\cos\alpha + \cos^3\alpha)/4$.

Open this experiment online

Try the calculation

Polar-cap half angle (degrees) 45



One polar cap | Isotropic source, same total power | Selected cone

One polar cap collects 5.806% of all dipole power. Both caps together collect 11.612%.

Move the cap to 90 degrees and explain the agreement of the two curves at this point. At a small angle the dipole fraction starts at order α^4, while the isotropic fraction starts at order α^2. Recover both leading powers using the ring argument.

A dipole pulse has different clocks for displacement, radiation and colour

A small neutral source changes its electric dipole moment along a fixed axis. Unlike a harmonic antenna operating indefinitely, this source begins and ends with negligible dipole moment, current and acceleration. Let its history in the source frame be $p(t) = p_0 \exp[-t^2/(2T^2)]$ along z. Assume that the source size is much smaller than cT, that the detector is far outside the source and its near zone over the main spectral band, and that the environment is vacuum. The aim is to infer both the time profile and total energy of the radiated pulse from this smooth change in charge distribution. At no point do we create or destroy net charge.

Translate the mechanical history into a detector trace

For a localized conserved source at leading electric-dipole order, the far electric field at fixed polar angle has magnitude proportional to $\ddot{p}(t - r/c)\sin\theta/r$. The sign of a polar component depends on basis convention, but relative lobes in a single fixed detector keep their signs. Introduce reduced retarded time $u = (t - r/c)/T$. Two derivatives give

$$\dot{p} = -\frac{p_0}{T} u e^{-u^2/2},$$

$$\ddot{p} = \frac{p_0}{T^2}(u^2 - 1)e^{-u^2/2}.$$

The source moment is largest at $u = 0$, its first derivative vanishes there, and its second derivative is negative with maximal magnitude. A detector sensitive to signed electric field observes a central lobe bracketed by oppositely signed weaker lobes. Its zero crossings are $u = \pm 1$; the side extrema are at $u = \pm\sqrt{3}$, where their magnitude relative to the central peak is

$$2e^{-3/2} = 0.446.$$

A power detector squares the waveform and loses its signs. Inferring the original $p(t)$ from a power-only trace therefore needs additional phase or source information.

Integrate the radiated energy in the time domain

The instantaneous angular power per emission time in the nonrelativistic electric-dipole approximation is

$$dP/d\Omega = \ddot{p}^{\,2} \sin^2\theta/(16\pi^2\epsilon_0 c^3).$$

Integrating over directions gives

$$P = \ddot{p}^{\,2}/(6\pi\epsilon_0 c^3).$$

Use $dt = T\,du$, and expand $(u^2-1)^2 = u^4 - 2u^2 + 1$. Gaussian moments are $\int e^{-u^2}du = \sqrt{\pi}$, $\int u^2 e^{-u^2}du = \sqrt{\pi}/2$, and $\int u^4 e^{-u^2}du = 3\sqrt{\pi}/4$ over the whole line. The time integral becomes

$$W = \frac{p_0^2}{6\pi\epsilon_0 c^3 T^3}\,\frac{3\sqrt{\pi}}{4}.$$

Stretch the event to twice its duration at fixed maximum moment. The radiative field peak decreases by four, its peak power decreases by sixteen, but the pulse lasts twice as long, giving one eighth the total radiated energy. This strong time-scale dependence is a useful diagnostic: a trace with unchanged integrated energy following such a stretch would challenge the assumed fixed dipole amplitude or the dipole-only description.

Check the total with a frequency-domain measurement

Define $\widetilde{p}(\omega) = \int_{-\infty}^{\infty} p(t) e^{i\omega t} dt = p_0 T\sqrt{2\pi}\exp[-(\omega T)^2/2]$. The transform of $\ddot{p}$ is $-\omega^2\widetilde{p}$, with no endpoint terms. Parseval's identity with this convention gives a one-sided energy spectrum

$$\frac{dW}{d\omega} = \frac{\omega^4 |\widetilde{p}(\omega)|^2}{6\pi^2\epsilon_0 c^3}, \qquad \omega > 0.$$

The factor of two for negative frequencies has already been included. Although $|\widetilde{p}|$ peaks at zero frequency, radiation has no zero-frequency component for this pulse. Maximising $\omega^4 e^{-(\omega T)^2}$ gives

$$\omega_{\max} = \sqrt{2}/T.$$

Integrating the spectrum over $\omega > 0$ reproduces the time-domain result because $\int_0^\infty v^4 e^{-v^2} dv = 3\sqrt{\pi}/8$. This is an independent check of both the dipole differentiation and the Fourier convention. A spectrum plotted per unit wavelength would need the Jacobian $|d\omega/d\lambda|$ and generally has a different peak.

For $T = 0.50$ ns, the two signed field zero crossings are 1.00 ns apart while maximum spectral energy per angular-frequency interval occurs near 450 MHz when converted to ordinary frequency. If the detector is placed 0.60 m farther along the same far-field ray, every feature arrives 2.00 ns later. After multiplying the signed fields by their respective radii and compensating this flight time, the traces should match. Failure of the scaling can reveal a near-field contribution or a frequency-dependent transfer function in the receiving system.

Two integral tests for the sign of a measured waveform

For a single far-field detector at fixed angle outside a radiation node the signed voltage, after calibrating its response, is proportional to $\ddot{p}(t - r/c)$. The area of that trace is zero when $\dot{p}$ has the same value before and after the event. Integrate the first time moment by parts:

$$\int t\ddot{p}\, dt = [t\dot{p}]_{-\infty}^{\infty} - [p]_{-\infty}^{\infty}.$$

It too vanishes for the specified Gaussian pulse. These identities constrain both the relative signs and the wing amplitudes. A clipped recording can violate them even when the main central peak is accurately sampled.

The second time moment survives. Two integrations by parts give $\int t^2 \ddot{p}\, dt = 2\int p\, dt$, equal to $2\sqrt{2\pi} p_0 T$ for the model. A common delay shifts t^2 into a sum of quadratic, linear and constant terms. The lower two moments vanish, so the second moment is unchanged by that delay. This supplies an independent way to compare two radius-corrected detector traces without first finding the centre of a noisy lobe. Its usefulness depends on retaining the wings long enough to control the quadratic time weighting.

Recover a pulse timescale without identifying a carrier

An equatorial detector records a signed pulse whose central lobe crosses zero at 6.2 and 7.8 ns. A spectrometer, calibrated per unit ordinary frequency, finds the strongest signal near 280 MHz. Decide whether the observations fit the Gaussian dipole model. What should a detector on the dipole axis measure?

▼ **Derive the result and check it independently**

The two time zeros bracket $2T$, giving $T = 0.80$ ns and the source-centred retarded event at detector time 7.0 ns. The predicted spectral peak is

$$f_{\max} = \sqrt{2}/(2\pi T) = 281\,\mathrm{MHz}.$$

Converting from angular to ordinary frequency changes the spectral density by a constant 2π, so the peak is at the frequency just calculated. Both measurements agree within the stated precision. They still do not fix the absolute pulse energy without amplitude calibration and distance.

Along the dipole axis the transverse radiation field vanishes in this approximation. A finite nearby axial detector can respond to electrostatic or induction fields and a finite optical aperture can capture off-axis radiation. Changing distance and angular aperture separately can distinguish these effects. A signed time integral of the far electric trace should also vanish because it is proportional to the difference between final and initial $\dot{p}$.

Cambridge Electromagnetism, chapter 5 · retarded fields and electric dipole radiation

From potentials to a distant detector

An alternating current in a short conductor moves positive and negative charge relative to one another. A current probe at the source and a field probe well outside it measure different aspects of the same charge motion. Consider a localised neutral source of size d, oscillating at angular frequency ω in vacuum. Assume $d \ll \lambda = 2\pi c/\omega$ so that phase changes across the source are small. The leading electric multipole is its dipole moment $\mathbf{p} = \int \mathbf{r}'\rho(\mathbf{r}', t)d^3r'$.

A signal emitted at source time $t_r = t - r/c$ reaches an observation point a distance r away at time t. Keeping the leading distant term of the retarded vector potential gives

$$\mathbf{A} \simeq \mu_0 \int \mathbf{J}(\mathbf{r}', t_r)d^3r'/(4\pi r).$$

By charge continuity the current integral equals $\dot{\mathbf{p}}$. One sees this by differentiating the integral defining a component of $\mathbf{p}$, substituting $-\nabla' \cdot \mathbf{J}$ for $\partial_t \rho$ and integrating by parts. No surface term remains for a localised source whose current vanishes at the enclosing boundary.

Extract the transverse radiation field

Curling the potential differentiates both $1/r$ and the retarded time. The former produces a term falling as $1/r^2$. Differentiating t_r produces a term proportional to $\ddot{\mathbf{p}}/(cr)$, which dominates sufficiently far away. The resulting magnetic field is transverse to the direction of observation. The radiative electric field can be written as the transverse vector $\hat{\mathbf{r}} \times (\hat{\mathbf{r}} \times \ddot{\mathbf{p}}(t_r))/(4\pi\epsilon_0 c^2 r)$. The magnetic field is $\hat{\mathbf{r}} \times \mathbf{E}/c$.

Take $\mathbf{p} = p_0 \cos(\omega t)\hat{\mathbf{z}}$. The electric amplitude is proportional to $\omega^2 p_0 \sin\theta/r$, where θ is the angle to the dipole axis. Along the axis no transverse component survives. In the equatorial plane the radiation is strongest. These conclusions concern the energy-carrying far field. The electrostatic dipole field and the induction field decay with higher inverse powers of radius and may dominate a probe whose distance is small on the wavelength scale.

Integrate over a sphere

Use the plane-wave expression for the cycle-averaged Poynting flux locally in the far zone. Multiplying by r^2 gives the power per unit solid angle. Integration uses $d\Omega = \sin\theta\, d\theta\, d\varphi$ and $\int \sin^2\theta\, d\Omega = 8\pi/3$. For a peak dipole moment the result is

$$\frac{d\langle P\rangle}{d\Omega} = \frac{\omega^4 p_0^2}{32\pi^2\epsilon_0 c^3}\sin^2\theta,$$

$$\langle P\rangle = \frac{\omega^4 p_0^2}{12\pi\epsilon_0 c^3}.$$

The sphere has more area at larger radius while the intensity falls as r^{-2}, so total power is independent of the chosen far-field radius. At fixed dipole moment, doubling frequency increases the radiated power sixteenfold. If the experiment instead fixes the amplitude of current through a short element, the resulting dipole moment falls with increasing frequency because current is the time derivative of charge. The change in what is controlled matters.

Check an angular scan

A calibrated small dipole radiates 3.0 mW in total. At $r = 10$ m in its far zone, the equatorial mean intensity is $3\langle P\rangle/(8\pi r^2) = 3.58$ µW/m². At 30° from the axis it is one quarter of that value, 0.895 µW/m². Treating the source as isotropic would conceal both the axial null and the enhanced equatorial intensity. A dipole probe receiving the field must also be correctly polarised to measure the predicted amplitude.

Move a correctly oriented equatorial probe from 10 to 20 m. Its measured electric amplitude should halve and its measured power per area should fall by four. Failure of this scaling could indicate reflections, inadequate distance from the source or a directional response in the probe. Combining this distance test with an angular scan distinguishes a radiation measurement from a mapping of the near electrostatic field.

An oscillating electric dipole with fixed orientation emits linearly polarized radiation at a generic far-field point. The electric field points along the projection of the dipole axis onto the plane transverse to the line of sight. Turning a linear receiving antenna through a right angle can therefore remove its signal even at a location of maximum emitted intensity. To distinguish antenna misalignment from a source-pattern null, repeat the scan with two orthogonal receiving orientations and add the received powers.

Radiation from a driven bound charge

A dipole need not be fed by a wire. An incident optical field can drive a bound charge, which reradiates part of the incident power. Model a charge of magnitude q and effective mass m attached to a harmonic restoring force $-m\omega_0^2 x$. Expose it to a spatially uniform sinusoidal electric field of peak amplitude E_0 at angular frequency ω. Uniformity here requires the oscillator size to be small compared with the wavelength. Include linear damping rate γ through the term $m\gamma\dot{x}$. Steady harmonic response gives a dipole amplitude

$$|p| = \frac{q^2 E_0/m}{\sqrt{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}}.$$

The radiated power of a sinusoidal electric dipole with peak moment p is $\omega^4 |p|^2/(12\pi\epsilon_0 c^3)$. Divide by the incident plane-wave intensity $\epsilon_0 c E_0^2/2$ to obtain a scattering cross section. Well below resonance, where the restoring term dominates, define the static polarizability $\alpha_0 = q^2/(m\omega_0^2)$ in SI units. Then

$$\sigma_{\mathrm{sc}} \simeq \frac{k^4}{6\pi}\left|\frac{\alpha_0}{\epsilon_0}\right|^2,$$

$$k = \omega/c.$$

The strong fourth-power dependence originates in the dipole's radiating acceleration. The amplitude of the induced moment varies little in the low-frequency regime, but twice differentiating it in time produces a field amplitude proportional to ω^2. Squaring that field gives the fourth power. Near an internal resonance the polarizability itself varies rapidly and this simple colour dependence no longer applies.

Predict the ratio and magnitude of scattered light

Suppose a small isotropic object has $\alpha_0/(4\pi\epsilon_0) = 1.00 \times 10^{-30}$ m³, and both 450 nm and 650 nm light are far from its absorption bands. At 450 nm, $k = 1.40 \times 10^{7}$ m⁻¹. Substitution gives $\sigma_{\mathrm{sc}} = (8\pi/3)k^4(10^{-30}\,\mathrm{m}^3)^2 \simeq 3.18 \times 10^{-31}$ m². At 650 nm the cross section is smaller by $(650/450)^4 = 4.35$, about 7.32×10^{-32} m². Equal incident irradiances therefore give that same ratio of total scattered powers.

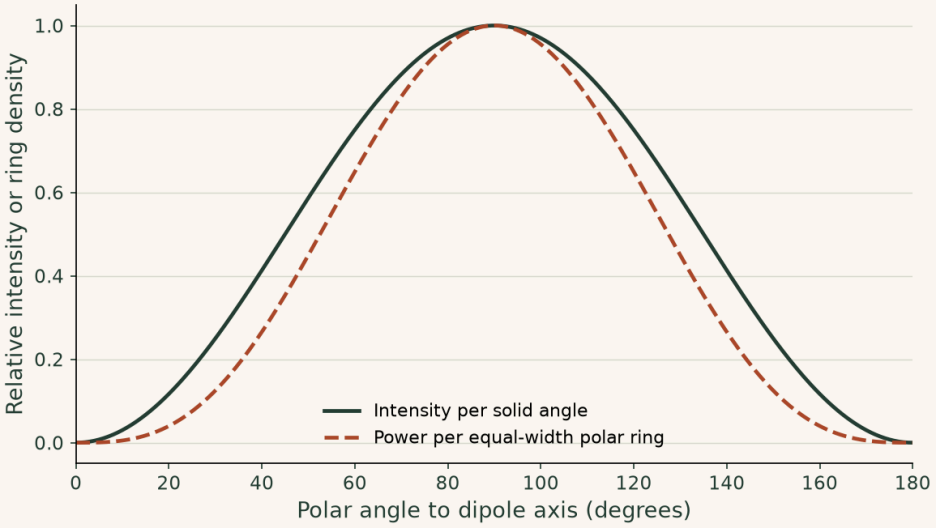
The angular pattern remains dipolar, so a small detector also needs a known collection angle and polarization geometry. At a direction parallel to the incident electric polarization the ideal driven dipole has a null, irrespective of wavelength. A distribution of particle sizes with some dimensions comparable to wavelength can fill such a null and alter the colour ratio. Comparing angular scans and wavelength scans separates these two experimental predictions rather than attributing every departure to a changed number of scatterers.

An ideal absorption-free dielectric particle still removes some power from the original forward beam by redirecting radiation. For a dilute ensemble one might infer scattering cross section from the weak reduction in collimated transmitted power, provided light scattered into the detector aperture is excluded. If the material also absorbs, the reduction measures extinction and is larger than scattering alone. Measuring total scattered power and forward-beam attenuation separately then distinguishes reradiated energy from internal dissipation. Comparing measurements at several concentrations checks whether neighbouring objects remain independent scatterers or instead interfere appreciably across the ensemble.

The polarizability here is in SI units. Tabulated polarizability volumes must first be multiplied by $4\pi\epsilon_0$ before substitution in the SI cross-section expression.

Translate a dipole pattern into collected power

An electrically short dipole radiates far-field intensity proportional to $\sin^2\theta$. A detector spanning a narrow band of polar angles and the full azimuth subtends $2\pi \sin\theta\, d\theta$. Compare the two peak-normalised angular profiles.



A polar ring at 30° collects only 1/8 of the power of an equally wide equatorial ring to first order in its width. Its intensity per unit solid angle is 1/4 of the equatorial value. The extra sine comes from the area of the ring.

▶ **Read values from the plot**

Develop this calculation further

Chapter 23 · Potentials, gauge freedom and the arrival of a signal · Chapter 24 · A moving charge, its retarded position and its radiation · Chapter 28 · Why a small transparent sphere scatters light

Study the full derivation

Radiating multipoles and what a detector actually integrates · Waves, radiation and relativity

A small oscillating source looks point-like from a distant detector, yet its angular pattern reveals its structure. Derive electric and magnetic dipole radiation from a conserved current, identify what survives when electric dipoles cancel, and turn a detector aperture into a calibrated estimate of total power.

Problem sheet 3 · Energy flux and radiation

Use the energy balance of Lecture 5 and the angular pattern of Lecture 6 to interpret remotely measured field data.

Reconcile two field instruments

A normally incident plane wave in vacuum gives a measured RMS magnetic field of 12 nT. Find the RMS electric field and mean intensity. An ideal absorber covers 25 cm²; determine its mean heating rate and explain how a standing wave would complicate the inference from a single magnetic probe.

▼ **Reconstruct the Poynting flux**

For a forward plane wave $E_{\mathrm{rms}} = cB_{\mathrm{rms}} = 3.60$ V/m and $\langle S \rangle = E_{\mathrm{rms}}^2 / Z_{\mathrm{vac}} = 0.0344$ W/m². The area is 0.0025 m², so mean power is 86 µW. A standing wave can have a magnetic antinode with substantial stored energy even though oppositely directed powers cancel. Then a measurement of the magnetic field alone fails to determine the net power delivered.

Infer total dipole radiation

A far-zone probe oriented correctly at $r = 8.0$ m and $\theta = 45^\circ$ from an electric dipole axis reads 1.20 µW/m². Infer the equatorial intensity at the same distance, total radiated power and the reading at 16 m and 90°. State a test for an unwanted reflected wave.

▼ **Integrate the angular pattern**

At 45° the factor $\sin^2\theta$ is one half. The equatorial intensity is therefore 2.40 µW/m². Total power follows from $P = (8\pi/3) r^2 S_{\mathrm{eq}}$, giving 1.29 mW. Doubling radius at the equator gives 0.600 µW/m². A radial scan at fixed angle should make $r^2 S$ constant. Reproducible oscillations over fractions of a wavelength can reveal interference with a reflection.

Which quantity is fixed as frequency changes?

A very short current element of effective length ℓ has $\dot{p} = I\ell$, with sinusoidal current of peak amplitude I_0. Express dipole radiation power in terms of current amplitude. Double the frequency at fixed current and then at fixed dipole moment. Compare and interpret the two power ratios.

▼ **Trace the frequency dependence**

For a sinusoid $p_0 = I_0 \ell / \omega$. Inserting this into $P = \omega^4 p_0^2 / (12\pi\epsilon_0 c^3)$ gives

$$P = \omega^2 I_0^2 \ell^2 / (12\pi\epsilon_0 c^3).$$

At fixed current it grows fourfold, while at fixed dipole moment it grows sixteenfold. If the original power were 0.50 mW, the outcomes would be 2.0 and 8.0 mW. At the higher frequency the same current displaces less charge per half-cycle. This explains why the two experiments impose different source conditions.

Choose a frame for electric-dominated crossed fields

In a laboratory $E_x = 12.0$ MV/m and $B_z = 0.0200$ T. Decide which field can vanish in an inertial frame. Determine the velocity of that frame and the magnitude of the remaining field. Use $c = 3.00 \times 10^8$ m/s.

▼ **Transform the magnetic component first**

Here $cB = 6.00$ MV/m, less than E. Set $\mathbf{u} = c^2\mathbf{E} \times \mathbf{B}/E^2$ so

$$\mathbf{u} \times \mathbf{E}/c^2 = \mathbf{B}.$$

The velocity is 1.50×10^8 m/s along negative y. The remaining electric field is $\sqrt{E^2 - c^2B^2} = 10.4$ MV/m, consistent with invariance of $E^2 - c^2B^2$. Eliminating the electric field instead would require

$$E/B = 6.00 \times 10^8\,\mathrm{m/s}.$$

Revisit chapter 5 · Which observer can see a purely magnetic field?.

Explore these relations

- For the two field instruments, compare travelling and standing waves before inferring flux from one local field. Open the live calculation.
- For dipole radiation, vary the source and inspect the angular power distribution and integrated power. Open the live calculation.

Chapter 7 · Multipole fields

Predict the leading far field of a localized charge distribution and use symmetry to decide how rapidly it decays.

FROM FIRST PRINCIPLES

Why far-field measurements resolve moments in order

Place a static charge distribution inside a ball of radius a. A distant detector at $\mathbf{r} = r\hat{\mathbf{n}}$, $r \gg a$, sums contributions with very similar travel distances. Begin with Coulomb's potential, and factor the common r out of the denominator:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0 r}\int \rho(\mathbf{r}')\left(1 - 2\frac{\hat{\mathbf{n}}\cdot\mathbf{r}'}{r} + \frac{|\mathbf{r}'|^2}{r^2}\right)^{-1/2} dV'.$$

For a small dimensionless z,

$$(1+z)^{-1/2} = 1 - z/2 + 3z^2/8 + \cdots.$$

Keeping terms through first order in a/r takes the detector direction outside the second integral:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\left(\frac{Q}{r} + \frac{\hat{\mathbf{n}}\cdot\mathbf{p}}{r^2} + \cdots\right),$$

$$Q = \int \rho\, dV',$$

$$\mathbf{p} = \int \mathbf{r}'\rho\, dV'.$$

In a neutral sample the leading coefficient vanishes exactly. The dipole term distinguishes viewing directions and gains one extra inverse power of distance. The electric field is the negative spatial gradient, so its monopole and dipole pieces fall as r^{-2} and r^{-3}. An inversion-symmetric neutral distribution also has zero dipole and requires the next, quadrupole term. The omission of higher terms is controlled by source size relative to observation distance.

DERIVATION AND APPLICATIONS

Translate the origin of a multipole expansion

A tabulated dipole moment can change when the coordinate origin moves. To track what a distant measurement actually fixes, consider a bounded static distribution of point charges q_j in vacuum. Let $\mathbf{r}_j$ denote their positions relative to origin O, and observe at $\mathbf{r} = r\mathbf{n}$ with r much larger than every source distance. Expand the reciprocal distance in the Coulomb potential to first order in each source coordinate.

$$\frac{1}{|\mathbf{r}-\mathbf{r}_j|}=\frac{1}{r}+\frac{\mathbf{n}\cdot\mathbf{r}_j}{r^2}+O\!\left(\frac{|\mathbf{r}_j|^2}{r^3}\right).$$

The coefficients of the first two terms are total charge Q and dipole moment $\mathbf{p}$. Move the origin by vector $\mathbf{a}$ without moving the apparatus. Subtract $\mathbf{a}$ from every source coordinate and sum the moments again.

$$Q=\sum_j q_j,$$

$$\mathbf{p}'=\sum_j q_j(\mathbf{r}_j-\mathbf{a})=\mathbf{p}-Q\mathbf{a}.$$

A neutral system consequently has an origin-independent dipole. For a charged system the observation vector also changes. To the same order, $r'=|\mathbf{r}-\mathbf{a}|$ and the monopole at the new origin contributes precisely the apparent dipole that was removed from the source moment.

$$\frac{Q}{r'}+\frac{\mathbf{n}'\cdot\mathbf{p}'}{r'^2}=\frac{Q}{r}+\frac{Q\mathbf{n}\cdot\mathbf{a}+\mathbf{n}\cdot(\mathbf{p}-Q\mathbf{a})}{r^2}+O(r^{-3}).$$

Multiplying by $k=1/(4\pi\epsilon_0)$ restores voltage units. The cancellation holds for a distant observer relative to both origins. A finite-order numerical truncation can still differ at the first omitted order.

WORKED EXAMPLE

Choose the centre of charge for two positive beads

Change the inputs and investigate this derivation.

Charges 2.00 nC at $x=0$ and 1.00 nC at $x=30.0$ mm have $Q=3.00$ nC and $p_x=30.0$ pC m. The shift $a=p_x/Q=10.0$ mm sets the new dipole to zero. At the physical position $x=1.00$ m the shifted monopole predicts

$$kQ/0.990=27.24\,\mathrm{V}.$$

The exact two-charge expression $k(2.00+1.00/0.970)10^{-9}$ gives 27.24 V to that precision. Higher moments explain the small residual.

SECOND APPLICATION

Locate a neutral dipole from an axial voltage

Change the inputs and investigate this derivation.

Two opposite charges of magnitude 1.50 nC are separated by 8.00 mm. Their dipole moment is 12.0 pC m, pointing towards the positive charge. At a distance of 0.400 m from the midpoint on the positive axis, the leading voltage is

$$kp/r^2 = 0.674\,\mathrm{V}.$$

Moving the coordinate origin 20 mm towards the positive charge leaves the moment at 12.0 pC m because $Q = 0$. Higher terms about the shifted origin account for the change in the numerical observer distance.

Try it · Multipole origin

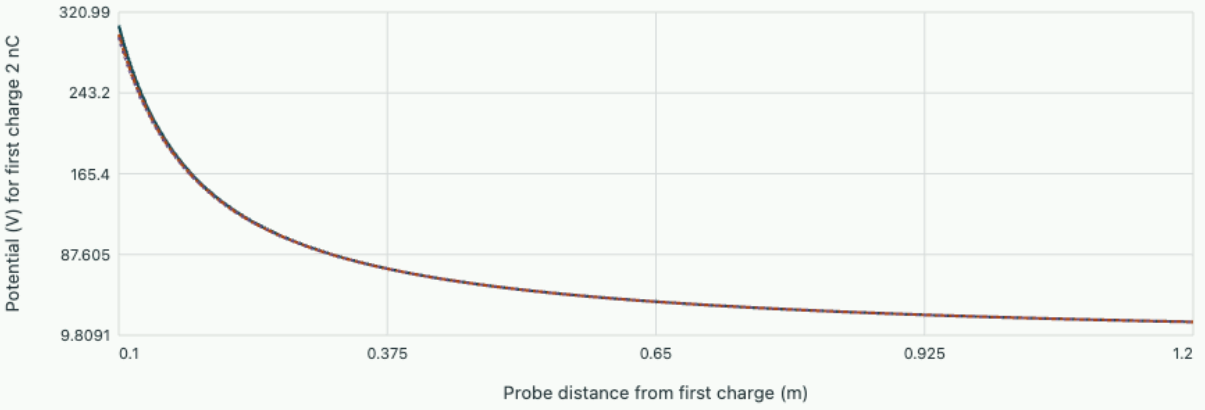
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Second / first charge ratio **0.5**

Charge separation (mm) **30**



Two exact point charges · Monopole at centre of charge · Monopole + dipole at old origin

Centre of charge 10 mm. Shifting the expansion origin removes the dipole for nonzero total charge; the exact field stays unchanged.

How far away is far enough to treat two charges as a dipole?

Take a positive and a negative charge, each of magnitude 1 nC, a distance $2a$ apart. Observe on the positive axis at $z > a$. Choose a tolerance of one per cent before looking at the graph and predict the required distance.

Superposition gives

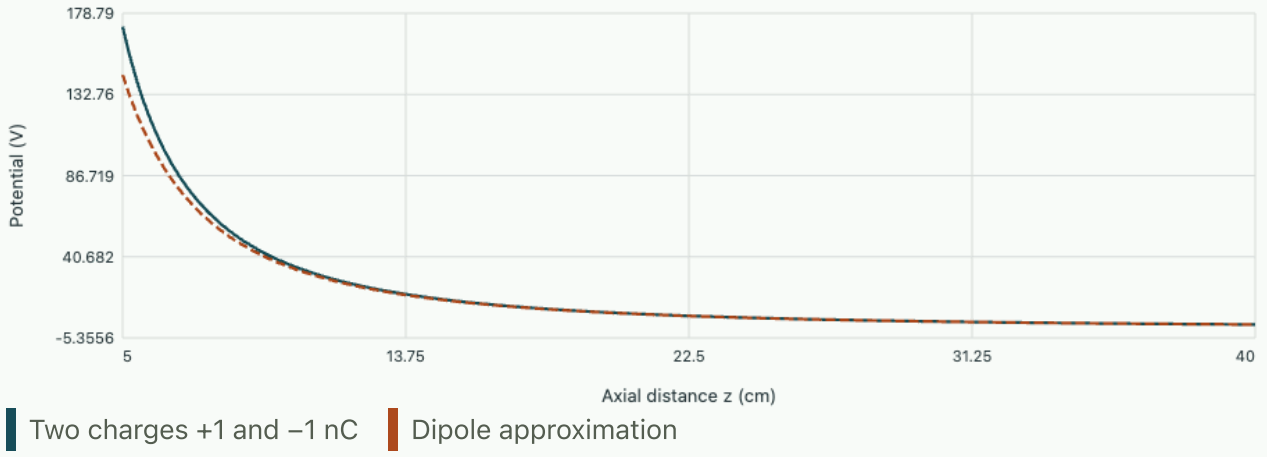
$$V = kq[(z-a)^{-1} - (z+a)^{-1}] = 2kqa/(z^2 - a^2).$$

Replacing the denominator by z^2 gives the dipole term. Dividing its difference from the exact expression by the exact expression leaves exactly $(a/z)^2$ on this axis.

Open this experiment online

Try the calculation

Half separation a (cm) 2



The dipole underestimates the exact axial potential by 4.00% at 10 cm. Doubling that distance divides this relative error by four.

The one-per-cent distance is $10a$. Use two separations to check it on the curves. If both charges are doubled, explain separately what happens to the absolute error and to the required distance for a relative tolerance.

Reconstruct the orientation of a quadrupole from six voltage probes

Suppose a small charge distribution is confined near a known origin and has zero total charge and zero electric dipole moment. A device surrounds it with probes at a common radius R much larger than the source diameter. Each probe reports the electrostatic potential relative to infinity. Axial measurements are easy to interpret when the source is known to share the instrument's axes. Here the internal orientation is unknown. The task is to infer a tensor from a set of readings, predict another reading, and identify which unseen source features the data leave undetermined.

Extract the angular form from the Coulomb integral

For a source coordinate $\mathbf{s}$ and observation direction $\mathbf{n}$, expansion of the Coulomb kernel at $\mathbf{r} = R\mathbf{n}$ gives $R^{-1} + (\mathbf{n}\cdot\mathbf{s})/R^2 + [3(\mathbf{n}\cdot\mathbf{s})^2 - s^2]/(2R^3) + \cdots$. The first two terms integrate to zero by our source assumptions. Define

$$Q_{ij} = \int \rho(\mathbf{s})(3 s_i s_j - s^2 \delta_{ij})\, d^3 s,$$

$$V(R\mathbf{n}) = \frac{1}{8\pi\epsilon_0 R^3}\mathbf{n}^T Q \mathbf{n} + O(R^{-4}).$$

This convention absorbs the factor of three into Q; a tensor defined using the raw second moment would need a trace removal and a different numerical factor. The stated tensor is symmetric. Its trace vanishes identically because $3s^2 - s^2\delta_{ii} = 0$ in three spatial dimensions. Thus the six independent entries of a general symmetric matrix reduce to five. The scalar $\mathbf{n}^T Q \mathbf{n}$ is the quadrupole measured along the chosen direction, in C m². It has the same value for opposite probe directions.

Construct an invertible set of probe directions

Let each calibrated reading be $w(\mathbf{n}) = 8\pi\epsilon_0 R^3 V(R\mathbf{n})$. Along the three Cartesian axes, $w_x = Q_{xx}$, $w_y = Q_{yy}$, $w_z = Q_{zz}$. Their sum should vanish for a pure quadrupole. To obtain an off-diagonal component, point a probe at $\mathbf{n}_{xy} = (\hat{\mathbf{x}} + \hat{\mathbf{y}})/\sqrt{2}$. Expanding the quadratic form gives

$$w_{xy} = (Q_{xx} + Q_{yy})/2 + Q_{xy}.$$

Two analogous diagonal-of-a-plane directions give Q_{xz}, Q_{yz}. Six measurements determine the five independent values with one trace check. More directions provide independent residuals for noise or the influence of unmodelled moments.

For a concrete observation, report readings in units of $Q_* = 1.00 \times 10^{-14}$ C m². Suppose the axial readings are $0.50, 0.50, -1.00$, while the positive bisectors of xy, xz, yz give $2.00, -0.25, -0.25$. The trace check passes. The reconstructed off-diagonal entries are $Q_{xy} = 1.50 Q_*$ and $Q_{xz} = Q_{yz} = 0$. The xy block has eigenvectors $(1,1,0)/\sqrt{2}$ and $(1,-1,0)/\sqrt{2}$ with eigenvalues $2Q_*, -Q_*$. The z axis also has eigenvalue $-Q_*$. The source's principal symmetry axis lies on the positive xy bisector in this quadrupolar description. Rotating the other two axes within their degenerate plane yields the same tensor.

Predict a reading, then test the radial dependence

For $\mathbf{n} = (1,1,1)/\sqrt{3}$, the diagonal contribution is the trace divided by three and vanishes. The off-diagonal contribution is $2Q_{xy}/3 = Q_*$. At $R = 0.10$ m the corresponding predicted potential is

$$Q_*/(8\pi\epsilon_0 R^3) = 44.9\,\mathrm{mV}.$$

Put a second independently calibrated probe on the same ray at 0.20 m. Its quadrupole potential should be one eighth of that value, 5.62 mV. This is a new prediction using both an angle and a radius absent in the reconstruction.

Measure also at the opposite direction. Their difference cancels the quadrupole and all even angular orders while retaining odd moments, which can expose a residual dipole or octupole. Their average cancels the odd contributions but can still contain the monopole and higher even orders. Multiplying by R^3 and seeking a common plateau over several radii is stronger than a single pairwise symmetry test.

Because voltage signal weakens as R^{-3}, a distant probe can reduce truncation bias yet amplify instrumental error when reconstructing Q.

Estimate which principal axis is actually determined by the measurements

For a real symmetric xy block, an eigenaxis angle φ obeys $\tan(2\varphi) = 2Q_{xy}/(Q_{xx} - Q_{yy})$. Evaluate it using the signed numerator and denominator to select the appropriate quadrant. Perturbing our measured difference $Q_{xx} - Q_{yy}$ from zero to $+0.10Q_*$, while leaving $Q_{xy} = 1.50Q_*$, shifts the inferred principal axis from 45° to about 44.0°. The large separation between the exceptional eigenvalue and the two smaller ones stabilises this axis.

Within the plane belonging to the two equal negative eigenvalues, even a very small measurement perturbation can pick out almost any pair of apparent principal axes. That numerical rotation does not establish two new physical directions in the source. Fit the unique axis and the nearly degenerate subspace separately, and compare changes in predicted probe voltages under plausible uncertainty. A common 1% radius calibration error also changes the inferred magnitude by approximately 3% because of the R^3 factor, but leaves the inferred angular orientation unchanged if it affects every probe equally.

Identify what the reconstruction cannot see

All six readings receive the same unknown voltage offset. Can the trace check remove it? Would that procedure distinguish a meter offset from an actual nonzero monopole? What changes if the absolute reference is reliable but the source may also have a dipole?

▼ **Derive the result and check it independently**

At a common radius a uniform voltage offset adds the same constant to each rescaled w. A genuine quadrupole has zero axial sum, so under the stated zero-charge assumption the mean of the three axial readings estimates that common offset. Subtract it from every reading before the angular inversion. An actual monopole contributes the same angular offset on a single sphere and cannot be distinguished from an instrumental zero by this operation. Measurements at multiple radii separate a constant voltage offset, a $1/R$ monopole and a $1/R^3$ quadrupole, provided the potential reference stays stable when moving the probes.

A dipole is odd under reversal of the observation direction. Pairing each measurement with its antipodal partner and averaging first removes its contribution exactly in the multipole expansion. After that, fit the symmetric traceless tensor and predict more withheld directions. Distinct microscopic charge arrangements can share all five adjusted moments; the measured tensor describes the distant field to this order and generally does not identify a unique arrangement of internal charges.

Griffiths, 5th ed., §3.4 · multipole expansions and source moments

Reading a charge distribution from far away

The radiation of an oscillating dipole was described through its time-dependent moment. A static distribution also has a hierarchy of moments. Consider fixed charges q_a inside a sphere of radius a and an observation point $\mathbf{r} = r\mathbf{n}$ in vacuum. Measuring the complete charge map from far away would require extraordinary precision. Two distributions sharing their first few moments produce nearly the same voltage there. We can quantify which differences survive.

Superposition gives the exact potential as a sum of $q_a/(4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_a|)$. Factor out r and expand $(1+x)^{-1/2}$ with $x = -2\mathbf{n}\cdot\mathbf{r}_a/r + r_a^2/r^2$. Keep both the first-order contribution from r_a^2 and the squared contribution from $\mathbf{n}\cdot\mathbf{r}_a$. They have the same order in source size.

$$\frac{1}{|\mathbf{r}-\mathbf{r}_a|} = \frac{1}{r} + \frac{\mathbf{n}\cdot\mathbf{r}_a}{r^2} + \frac{3(\mathbf{n}\cdot\mathbf{r}_a)^2 - r_a^2}{2r^3} + O(a^3/r^4).$$

The total charge multiplies the first term. The second depends on the dipole moment $\mathbf{p} = \sum q_a \mathbf{r}_a$. Translation of the coordinate origin by $\mathbf{d}$ changes the dipole to $\mathbf{p} - Q\mathbf{d}$. A neutral system therefore has an origin-independent dipole. When charge and dipole vanish, the leading quadrupole is likewise independent of the chosen origin.

From potential to a vector measurement

For an axial dipole the leading potential is $p\cos\theta/(4\pi\epsilon_0 r^2)$. A probe measures a force per unit charge, requiring both radial and angular derivatives. The angular gradient contains $1/r$ because the physical arc length is $r\,d\theta$. This produces

$$E_r = \frac{2p\cos\theta}{4\pi\epsilon_0 r^3},$$

$$E_\theta = \frac{p\sin\theta}{4\pi\epsilon_0 r^3}.$$

A voltage reading of zero on the equatorial plane thus coexists with a nonzero transverse field. The vector result also supplies a laboratory diagnostic. Doubling distance reduces an ideal dipole field by eight, while a residual net charge produces a fourfold reduction. A small net charge can dominate sufficiently far from a nominally neutral device.

A quadrupole on the bench

Put $+1.0$ nC at $z = \pm 1.0$ cm and -2.0 nC at the origin. Charge, dipole and all odd moments vanish. On the positive axis the exact result simplifies before expansion,

$$V(z) = kq\left(\frac{1}{z-a} + \frac{1}{z+a} - \frac{2}{z}\right) = \frac{2kqa^2}{z(z^2-a^2)}, \qquad z > a.$$

At $z = 0.100$ m the leading term $2kqa^2/z^3$ gives 1.798 V. The exact value is 1.816 V, so the relative error referred to the exact value is

$$a^2/z^2 = 1.0\%.$$

In the equatorial plane the same arrangement gives $2kq[1/\sqrt{r^2 + a^2} - 1/r]$, with a negative leading voltage of half the axial magnitude. Along directions where the angular factor of a leading moment vanishes, a relative-error estimate based on that term alone loses meaning. Compare absolute errors there or retain the next surviving moment.

A simple way to detect an experimental origin error is to compare voltages on opposite sides at equal nominal distances. For a pure axial dipole these voltages have equal magnitude and opposite signs. Translating either probe along the axis changes the magnitude because of the radial gradient. An apparent even contribution can then be caused by mismatched distances as easily as by a physical quadrupole. Reverse the source orientation while leaving the two probes fixed to separate a geometric offset from a source component that retains its sign.

A field gradient exerts a net force on a neutral dipole

Far-field multipole moments also determine how a localized source responds to an imposed field that varies little across it. Consider two charges $\pm q$ separated by a small vector $\mathbf{d}$ pointing from negative to positive charge. Their dipole moment is

$$\mathbf{p} = q\mathbf{d}.$$

Expand the external potential around their midpoint $\mathbf{R}$. The constant contributions cancel, leaving interaction energy $q[V(\mathbf{R} + \mathbf{d}/2) - V(\mathbf{R} - \mathbf{d}/2)] \simeq -\mathbf{p} \cdot \mathbf{E}(\mathbf{R})$. The next terms are negligible when the external field varies on a scale much longer than the separation.

For a fixed dipole orientation the force follows by taking the gradient with respect to the midpoint. In components $F_i = p_j \partial_i E_j$. The electrostatic field is curl free, so this equals $p_j \partial_j E_i$, the difference of the forces on the two ends obtained by direct expansion. A uniform field consequently exerts torque $\mathbf{p} \times \mathbf{E}$ but no net force on a fixed neutral dipole.

An induced dipole changes its magnitude as the field is raised. For a linear polarizable object with $\mathbf{p} = \alpha \mathbf{E}$, the work integrated over field buildup gives energy $U = -\alpha E^2/2$. The one-half also follows by adding the internal energy cost $p^2/(2\alpha)$ to the coupling $-\mathbf{p} \cdot \mathbf{E}$ and minimizing over $\mathbf{p}$. The force is then $\mathbf{F} = (\alpha/2)\nabla E^2$ for real constant α in a static field. Simply substituting αE into the permanent-dipole energy before differentiating would double the answer.

Predict a force on a dielectric bead

A neutral bead in vacuum has static polarizability $\alpha = 4.00 \times 10^{-21}$ C m²/V. Locally on a symmetry axis the field is $E_z = 2.00 \times 10^4$ V/m and its axial gradient is 3.00×10^6 V/m². Assume the bead is much smaller than the variation length $E/|\partial_z E| = 6.67$ mm and other field components vanish on that axis. The

induced moment is 8.00×10^{-17} C m. Its axial force is $\alpha E_z \partial_z E_z = 2.40 \times 10^{-10}$ N, directed toward increasing squared field.

If the field amplitude of the electrodes is doubled without changing their geometry, both field and gradient double, so the force rises fourfold. A fixed, mechanically aligned permanent dipole would show only a twofold force increase. This scaling and a reversal of electrode polarity offer practical discrimination between the two models. Near a field zero an induced dipole and its leading force both vanish; finite-size and higher-moment effects must be considered if a measurable force persists there.

The validity condition on field gradients can be phrased in terms of an observable length scale for a real bead. Moving it by a fraction of a millimetre along the axis and measuring its induced moment probes the local derivative of field. A moment tracking field linearly while force follows its squared-field gradient supports the same polarizability in both tests. If the bead size becomes comparable to the variation length, different parts of it experience different applied fields. Higher induced moments and local field distortion then enter the force balance, so one scalar dipole polarizability is no longer sufficient.

Develop this calculation further

Chapter 16 · Spherical harmonics as measurements of boundary voltage · Chapter 28 · Why a small transparent sphere scatters light

Study the full derivation

Normalised spherical harmonics and the field of an arbitrary three-dimensional source · Electrostatic field theory and measurement

An axial voltage scan can determine Legendre coefficients only if the source is rotationally symmetric. A general spherical electrode or molecule needs angular modes with two indices. Their normalisation fixes the conversion from a measured multipole coefficient to a physical voltage.

Chapter 8 · Polarization of a dielectric sphere

Match electrostatic fields at a curved interface and infer a material's induced dipole from its exterior potential.

FROM FIRST PRINCIPLES

How a dielectric surface determines its internal field

Put a uniform, isotropic dielectric sphere of radius a and permittivity ϵ in a uniform static field $E_0\hat{\mathbf{z}}$ in vacuum. No free charge is supplied to the interface. Gauss's law for the electric displacement therefore requires continuous normal D, while electrostatic circulation requires continuous tangential E. Write the potentials using the axial angular dependence $\cos\theta$:

$$V_{\text{in}} = -Ar\cos\theta,$$
$$V_{\text{out}} = -E_0 r\cos\theta + B\cos\theta/r^2.$$

These are harmonic functions. The interior expression is regular at the centre and the exterior correction decays at infinity. Continuity of potential at $r = a$ gives

$$A = E_0 - B/a^3.$$

Obtain the radial fields by differentiating both potentials with respect to radius and reversing their signs. Normal displacement continuity then gives a second independent equation:

$$\epsilon A = \epsilon_0(E_0 + 2B/a^3), \qquad (\epsilon + 2\epsilon_0)A = 3\epsilon_0 E_0.$$

The interior field is uniform and has magnitude A. The exterior correction is that of a dipole of moment $4\pi\epsilon_0 B$ along the applied field. In a material indistinguishable from vacuum, $A = E_0$ and the dipole vanishes. This limit checks both the radial derivative's factor of two and the sign of the induced dipole.

DERIVATION AND APPLICATIONS

Find the field inside a spherical cavity in a dielectric

A hollow pore samples the field inside a bulk insulator. A sphere makes the difference between the pore field and the surrounding macroscopic field analytically accessible. Let the host have static permittivity ϵ_h, the pore filling have permittivity ϵ_c, and their spherical interface have radius a. Both media are linear, isotropic and homogeneous. There is no free interfacial charge. Far inside the host an applied uniform field E_∞ points along z. Choose potentials with the same angular dependence as this forcing.

$$V_c = -Cr\cos\theta,$$
$$V_h = -E_\infty r\cos\theta + \frac{D}{r^2}\cos\theta.$$

The regular interior solution gives a uniform pore field of magnitude C. The exterior term proportional to D is the decaying dipolar disturbance. Matching potential at $r = a$ ensures the tangential field is continuous. Matching normal displacement uses $E_r = -\partial_r V$ and introduces a factor of two in the exterior term.

$$C = E_\infty - \frac{D}{a^3},$$

$$\epsilon_c C = \epsilon_h \left(E_\infty + \frac{2D}{a^3} \right).$$

Eliminate the exterior coefficient with the first equation. The result depends on both permittivities even when the pore itself is empty. The radius controls the dipole term through its volume, while the uniform cavity field depends only on the contrast.

$$C = \frac{3\epsilon_h}{\epsilon_c + 2\epsilon_h} E_\infty,$$

$$D = a^3 \frac{\epsilon_c - \epsilon_h}{\epsilon_c + 2\epsilon_h} E_\infty.$$

Equal permittivities recover a perfectly uniform field everywhere. A vacuum cavity in a host of very large positive permittivity has C/E_∞ approaching $3/2$, with the far field inside the host held fixed.

WORKED EXAMPLE

Read the voltage across an air pore

Change the inputs and investigate this derivation.

A transparent host has $\epsilon_h = 4.00\epsilon_0$ and a spherical air pore of diameter 0.600 mm. Under $E_\infty = 1.50$ kV/m, the field ratio is $12/9 = 4/3$. The pore field is 2.00 kV/m and the magnitude of the potential difference between opposite axial poles is $2000(0.000600) = 1.20$ V. A diameter perpendicular to the applied field has both endpoints at the same potential.

SECOND APPLICATION

Infer a pore filling from the measured field

Change the inputs and investigate this derivation.

In a separate host of relative permittivity 3.00, a calibrated probe finds a uniform cavity field equal to 0.750 of the host's distant field. Inverting the ratio gives

$$\epsilon_c/\epsilon_0 = 3(3.00)/0.750 - 2(3.00) = 6.00.$$

The filling is more polarisable than the host. As a check, the exterior coefficient is

$$D/(a^3 E_\infty) = (6 - 3)/(6 + 6) = 0.250,$$

which yields the observed 0.750 field ratio by the potential condition alone.

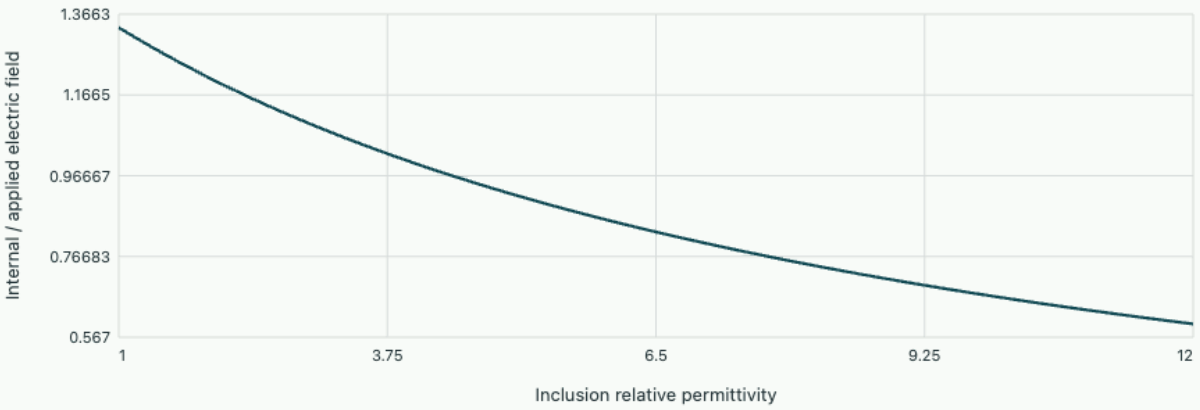
Try it · Dielectric pore

Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Host relative permittivity 4



Spherical inclusion in dielectric host

An air pore has field ratio 1.3333. The formula depends on both host and inclusion, with continuous potential and normal displacement.

Why a leaky sphere can reverse its drift in an alternating field

A neutral polarised particle can move toward a stronger electric field. If charge can also migrate through the particle and its surroundings, changing the frequency alone can reverse the leading gradient force. Calculate when the reversal exists.

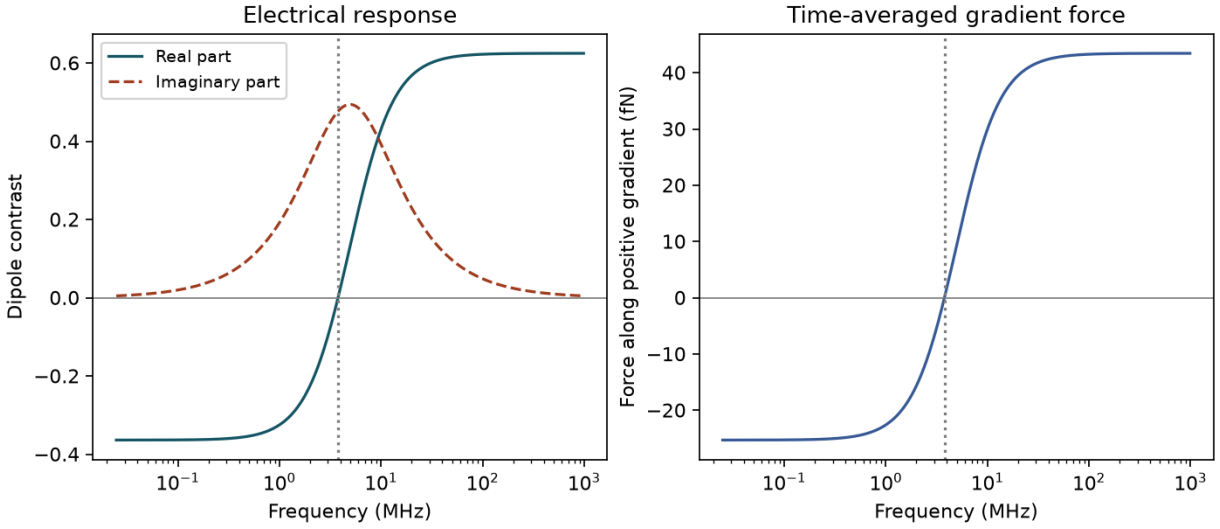
Follow the boundary current through the interface

Take a sphere of radius a in a homogeneous medium. At frequencies for which its diameter is tiny compared with the wavelength and external field-variation length, the field near it can be treated quasistatically. Adopt time dependence $e^{i\omega t}$. Continuity of total normal current involves $(\sigma + i\omega\epsilon)E_n = i\omega\epsilon^* E_n$ with $\epsilon^* = \epsilon - i\sigma/\omega$. Matching the scalar potential and this boundary flux is algebraically the same sphere problem as in electrostatics, with the appropriate complex material ratios.

The dipole contrast $K = (\epsilon_p^* - \epsilon_m^*)/(\epsilon_p^* + 2\epsilon_m^*)$ has low-frequency limit $K_0 = (\sigma_p - \sigma_m)/(\sigma_p + 2\sigma_m)$ and high-frequency limit $K_\infty = (\epsilon_p - \epsilon_m)/(\epsilon_p + 2\epsilon_m)$. Defining $\tau = (\epsilon_p + 2\epsilon_m)/(\sigma_p + 2\sigma_m)$, multiplying numerator and denominator by the complex conjugate of the latter gives

$$\operatorname{Re} K = \frac{K_0 + K_\infty (\omega\tau)^2}{1 + (\omega\tau)^2}.$$

The time-averaged gradient force in the ideal standing or locally in-phase field is $2\pi\epsilon_m a^3 \operatorname{Re} K \, \nabla E_{\mathrm{rms}}^2$, with ϵ_m the real permittivity of the host. This prefactor uses rms electric field. A particle with positive real contrast is drawn towards increasing rms intensity. A crossover occurs only when the two limiting contrasts have opposite signs, in which case $\omega_c = \sqrt{-K_0/K_\infty}/\tau$. If both contrasts vanish the particle is electrically matched in this model at all frequencies and no unique crossover can be inferred.



The vertical line marks the zero of the in-phase dipole. Loss and out-of-phase polarisation can remain finite there.

Try it · Dep contrast

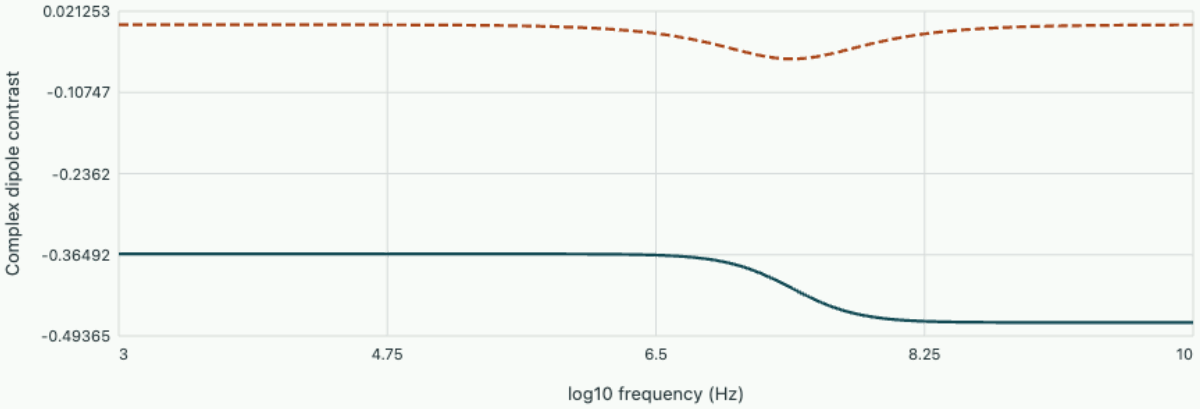
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Particle conductivity (S/m) **0.02**

Medium conductivity (S/m) **0.1**



Particle/medium relative permittivities are 3/80. The sign of the real contrast controls the time-averaged force in a field-intensity gradient; a zero crossing need not remove dielectric loss.

Download the corresponding editable notebook. The browser recalculates the reduced equations directly; the notebook exposes the full calculation and its checks.

Choose relative permittivities 60 and 10 and conductivities 0.002 and 0.010 S/m for particle and medium, respectively. At low frequency conduction makes the particle less responsive than its environment, $K_0 = -0.364$. At high frequency its larger dielectric constant gives $K_\infty = 0.625$. The predicted sign change is near 3.77 MHz. For a 5 micrometre radius sphere in a 10^{12} V²/m³ gradient, the limiting forces have magnitudes of order tens of femtonewtons. They are signed predictions along the field gradient, not measurements of a terminal velocity; inferring velocity also requires the drag law and the fluid motion.

Change only the particle conductivity so that it exceeds the medium value. The root disappears while the dielectric relaxation remains visible. The spherical single-interface model ignores ionic double layers, field-dependent material response, electrothermal flow and interactions with nearby walls. A frequency-dependent measured displacement alone can therefore have more than one cause, which makes the predicted zero and radius scaling useful independent checks.

Two spheres with the same material properties have radii 3 and 6 micrometres. What should doubling the radius change at fixed frequency and field gradient, and which parameter should remain the same?

▼ **Work through the solution**

The leading induced dipole and gradient force scale with a^3, so the larger sphere has eight times the predicted force. Both limiting contrasts and the relaxation time depend on material parameters, giving the same crossover frequency. A systematic radius dependence of that frequency suggests another length-dependent contribution or invalidity of the small-sphere approximation.

Open editable notebook Download .ipynb

Which field jumps at the pole of a dielectric sphere?

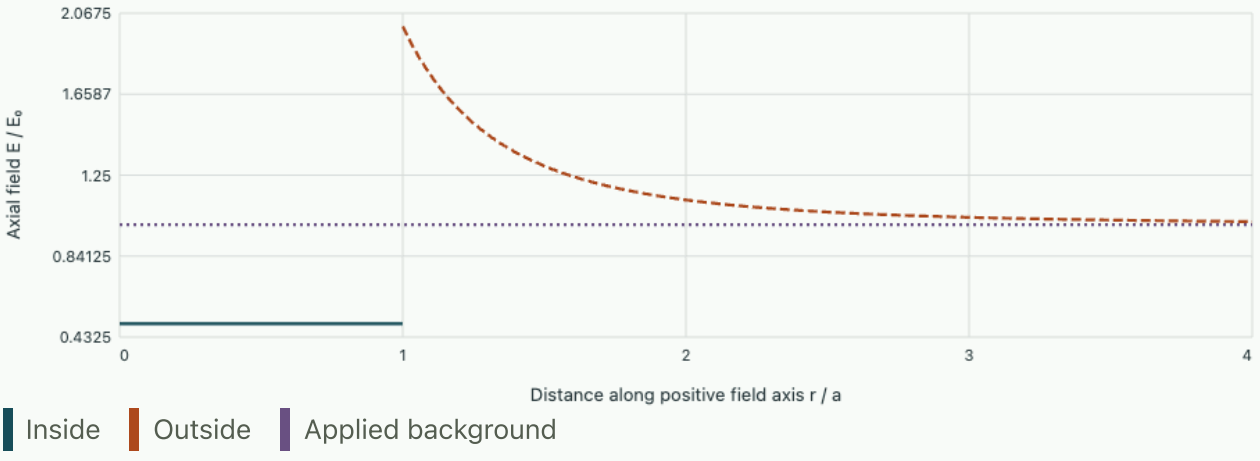
A linear dielectric sphere in vacuum sits in an otherwise uniform electric field. Follow the positive polar axis. Use the normal-displacement condition to predict the ratio of the fields before solving for either of them.

The regular interior potential is $-Ar\cos\theta$, while the exterior adds a dipole to the background potential. Continuity of potential and of ϵE_n gives $A/E_0 = 3/(\epsilon_r + 2)$ and exterior dipole coefficient $(\epsilon_r - 1)/(\epsilon_r + 2)$ in units of $E_0 a^3$. The two curves show both one-sided limits at $r = a$.

Open this experiment online

Try the calculation

Relative permittivity ε_r 4



Just inside, $E/E_0 = 0.5000$. Just outside at the pole, $E/E_0 = 2.0000$. Their ratio is 4.00.

Set the permittivity ratio to one. Then increase it until the interior field is below one fifth of the applied field. Does the exterior field at the pole grow without bound? Take the limit of the coefficient to explain

your result.

A coated inclusion can have zero external dipole and a finite internal field

A spherical particle of radius a and relative permittivity ϵ_c is surrounded by a concentric shell extending to b, with relative permittivity ϵ_s. The host has relative permittivity ϵ_h and supports an imposed uniform electrostatic field $E_0\hat{\mathbf{z}}$ far from the particle. All three materials are homogeneous, linear and isotropic, with positive permittivities and no free interfacial charge. Can one choose the amount of coating so that a distant probe sees exactly the undisturbed field, even though the core still changes the internal field? The answer requires matching fields at two interfaces under the same applied excitation.

Choose radial functions before assigning interface conditions

With axial symmetry and a uniform applied field, only the $\ell = 1$ sector is driven. Spherical boundaries preserve this angular dependence. The two radial solutions of Laplace's equation are r and r^{-2}. Reject the divergent term in the solid core, keep both in the shell, and allow a decaying dipole in the host. Write

$$V_c = -Ar\cos\theta,$$

$$V_s = (-Br + C/r^2)\cos\theta,$$

$$V_h = (-E_0 r + D/r^2)\cos\theta.$$

Here D is a potential coefficient, distinct from electric displacement $\mathbf{D}$. The latter enters through continuity of its normal component ϵE_r, while continuity of potential fixes the common tangential electric field at each radius. At $r = a$ these two equations reduce to

$$A = B - C/a^3,$$

$$\epsilon_c A = \epsilon_s(B + 2C/a^3).$$

Define $m = (\epsilon_c - \epsilon_s)/(\epsilon_c + 2\epsilon_s)$. Eliminating A gives

$$C = a^3 m B,$$

$$A = (1 - m)B.$$

This first interface problem determines how the core changes the mixture of regular and singular solutions in the shell, even before we choose a host.

Condense the inner sphere and shell into a surface response

Let $f = (a/b)^3$ be the core volume fraction. At the outer surface the shell potential divided by $-b\cos\theta$ is $B(1 - fm)$, and its normal displacement divided by $\epsilon_0\cos\theta$ is $\epsilon_s B(1 + 2fm)$. Their ratio defines the relative permittivity of a homogeneous sphere having the same boundary response to this excitation:

$$\epsilon_{\mathrm{eq}} = \epsilon_s \frac{1 + 2fm}{1 - fm},$$

$$\frac{D}{b^3 E_0} = \frac{\epsilon_{\mathrm{eq}} - \epsilon_h}{\epsilon_{\mathrm{eq}} + 2\epsilon_h}.$$

The second relation follows by matching host potential and normal displacement at b just as for a single homogeneous sphere. It also checks limiting cases. At $f = 0$ the equivalent object has shell permittivity; at $f = 1$ substitution reduces it to ϵ_c. If the core and shell are identical, $m = 0$ and changing their nominal boundary does nothing.

To eliminate the exterior dipole choose $\epsilon_{\mathrm{eq}} = \epsilon_h$, giving $f = (\epsilon_h - \epsilon_s)/[m(\epsilon_h + 2\epsilon_s)]$. A physical value requires

$$0 \le f \le 1;$$

for positive distinct permittivities the host must lie between the shell and core responses. An arbitrary mismatched shell cannot always cancel the polarisation of its core. The condition concerns a uniform applied field in electrostatics. A nearby point source drives multiple angular orders with different radial responses.

Read all three fields in a neutral inclusion

Choose $\epsilon_c = 8, \epsilon_s = 2, \epsilon_h = 3$. Then $m = 1/2$, $f = 2/7$, and $a/b = 0.6586$. Under the matched condition $D = 0$, so continuity at b demands $B(1 - fm) = E_0$, giving $B = 7E_0/6$ and $A = 7E_0/12$. The core still sees a uniform field only 58.3% as large as the host's undisturbed field. At b the shell-side radial electric field at the north pole is

$$B(1 + 2fm) = 1.50 E_0.$$

Multiplying by shell permittivity 2 recovers $3E_0$, exactly the required host displacement. At the equator the tangential field agrees directly. These surface tests can reveal a wrong sign in C even if an exterior dipole fit accidentally vanishes.

For $b = 30$ µm in $E_0 = 3000$ V/m, $a = 19.8$ µm and the voltage difference across the core diameter has magnitude $2aA = 69.2$ mV. The host would give $2aE_0 = 119$ mV over the same distance. A far-field perturbation measurement alone cannot establish a particle's internal electric exposure. A small capacitive probe inserted into the core would perturb the solution, so a practical comparison must estimate that additional boundary disturbance.

Check the coating with a different spatial pattern

To test how specific the cancellation is, impose an axisymmetric $\ell = 2$ potential on a distant enclosing electrode. More generally the radial functions are r^ℓ and $r^{-\ell-1}$ for each angular order. Repeating the inner matching gives $m_\ell = \ell(\epsilon_c - \epsilon_s)/[\ell\epsilon_c + (\ell + 1)\epsilon_s]$ in place of m. The outer surface now weighs the decaying term by $f_\ell = (a/b)^{2\ell+1}$. The equivalent boundary permittivity for this order is $\epsilon_s[1 + (\ell + 1)f_\ell m_\ell/\ell]/[1 - f_\ell m_\ell]$. Setting $\ell = 1$ reproduces the expression derived directly above.

For the neutral-dipole geometry at $\ell = 2$, $m_2 = 12/22$ and $f_2 = (2/7)^{5/3}$. These give an equivalent response near 2.36, clearly distinct from the host value of 3. A field-gradient experiment can thus reveal the inclusion even when the uniform-field exterior measurement gives an exact cancellation. This is a direct test of the assumption that the imposed field varies slowly on the scale of the coated sphere.

Quantify tolerance and distinguish neutrality from zero polarisation

Keep the outer radius fixed but let the manufactured core radius be 1% larger than the neutral value. Predict the sign and approximate size of the exterior dipole coefficient $D/(b^3 E_0)$. Why does summing all bound surface charges give a weaker check?

▼ **Derive the result and check it independently**

A small fractional radius error gives $\delta f \simeq 3f(0.01) = 0.00857$. Differentiate the equivalent permittivity:

$$d\epsilon_{\mathrm{eq}}/df = 3\epsilon_s m/(1 - fm)^2,$$

which here is

$$49/12 = 4.083.$$

Thus $\delta\epsilon_{\mathrm{eq}} \simeq 0.0350$. Near the matched point the dipole coefficient varies as $\delta\epsilon_{\mathrm{eq}}/(3\epsilon_h)$, giving $+0.00389$. The slightly increased high-permittivity core gives a weak dipole along the applied field. Inserting the actual perturbed radius in the full rational formula checks the linearisation.

Bound charge integrated over each complete spherical interface vanishes because its angular pattern is proportional to $\cos\theta$. This is true for both matched and mismatched coatings. It guarantees absence of a monopole but says nothing about the dipole. A numerical verification should check continuity of radial displacement, tangential field and the exterior dipole coefficient separately; one integrated neutral charge reading cannot replace those tests.

Griffiths, 5th ed., §4.4.2 · boundary problems with linear dielectrics

An inclusion in a uniform field

Place a homogeneous dielectric sphere of radius a and permittivity ϵ in vacuum. Distant electrodes produce a static uniform field $E_0\mathbf{e}_z$ before insertion. Take the dielectric to be linear, isotropic and free of mobile charge. Bound charge can develop at its surface even though the object is electrically neutral. The monopole measurement from the previous lecture consequently misses the material response. The dipole moment provides a measurable alternative.

In each constant-permittivity region Gauss's law for $\mathbf{D}$ and $\mathbf{E} = -\nabla V$ give Laplace's equation. The background potential is $-E_0 r\cos\theta$, an angular pattern of degree one. A spherical interface does not couple different angular degrees. The allowed radial functions for this degree are r and r^{-2}. Finiteness at the centre prohibits a singular interior term. Matching the imposed field far away fixes the growing exterior term.

$$V_{\text{in}} = -Ar\cos\theta,$$
$$V_{\text{out}} = -E_0 r\cos\theta + B\cos\theta/r^2.$$

Solve at the surface

Continuity of voltage gives

$$A = E_0 - B/a^3.$$

It also guarantees continuity of tangential field, as a tangential displacement compares points along the same surface potential. Normal displacement is continuous in the absence of free surface charge. Differentiating radially changes the sign and supplies a factor of two for the decaying term, so $\epsilon A = \epsilon_0(E_0 + 2B/a^3)$. Elimination gives

$$A = \frac{3\epsilon_0}{\epsilon + 2\epsilon_0}E_0,$$
$$B = a^3\frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0}E_0.$$

The uniform internal field polarizes the sphere along the field direction. Its bound surface charge is $\sigma_b = (\epsilon - \epsilon_0)A\cos\theta$. Opposite hemispheres cancel in total charge. Their separation gives an induced dipole $p = 4\pi\epsilon_0 B$. Equivalently integrate the uniform polarization over the volume. Both constructions give the same dipole, one from exterior voltage and one from microscopic charge displacement. When $\epsilon = \epsilon_0$ the exterior perturbation vanishes. A very large positive permittivity suppresses the internal field and approaches the conductor result.

Predict a voltage difference

Take $a = 2.00$ mm, $\epsilon = 4\epsilon_0$ and $E_0 = 30.0$ kV/m. The internal field is 15.0 kV/m, and $B = 1.20 \times 10^{-4}$ V m². Multiplication by $4\pi\epsilon_0$ gives $p = 1.34 \times 10^{-14}$ C m. At an axial probe 10.0 mm from the centre the sphere increases voltage relative to the uniform-field baseline by $B/r^2 = 1.20$ V. The opposite axial probe sees a shift of -1.20 V. A differential measurement doubles the signal and removes a common voltage offset.

For the assumed uniform applied field the exterior dipole solution is exact at every point outside the sphere. A nearby field-producing electrode can spoil that simplification. A field that varies across the particle contains higher angular orders in its local expansion and their induced responses survive at the surface. In an experiment, repeating the measurement with a smaller sphere tests the predicted a^3 law while reducing sensitivity to field curvature.

The assumption of real, constant permittivity is experimentally testable. A slowly varying applied field should give a dipole that follows the field with negligible lag. If the measured signal shifts in phase while frequency is raised, the response must be described by a frequency-dependent complex polarization. A static voltage estimate alone cannot supply that extra timescale.

Build an effective permittivity from many inclusions

A single dielectric sphere in a smooth applied field develops a dipole. A dilute suspension of such spheres can therefore alter the bulk capacitance of a sample. Suppose identical spheres of radius a and permittivity ϵ_i occupy a small fraction f of a matrix with permittivity ϵ_m. The matrix is homogeneous, both materials are isotropic, and no conducting path or interfacial free charge is present. Each sphere sees nearly the matrix field when it is far from all neighbours. The single-sphere boundary calculation gives an excess dipole $p = 4\pi\epsilon_m a^3 b E_m$, with $b = (\epsilon_i - \epsilon_m)/(\epsilon_i + 2\epsilon_m)$.

The number of inclusions per unit sample volume is $n = f/(4\pi a^3/3)$. At first order the excess polarization density is

$$np = 3f\epsilon_m b E_m.$$

The difference between E_m and the sample-average field is itself first order in f, so replacing one by the other in this correction changes only higher-order terms. Comparing average displacement with $\epsilon_{\mathrm{eff}} E$ gives

$$\epsilon_{\mathrm{eff}} = \epsilon_m[1 + 3fb + O(f^2)].$$

This is a dilute effective-medium prediction. The absolute radius cancels at fixed volume fraction because each dipole scales as a^3 while number density scales as a^{-3}. The cancellation requires particles small compared with the scale of the imposed field. Aggregation, closely spaced pairs and a large occupied fraction can change local fields and produce a nonlinear dependence on concentration.

Measure a composite using two capacitor readings

A parallel-plate measurement cell has electrode area 25.0 cm² and gap 1.00 mm. Filling it with matrix alone gives $\epsilon_m = 2.00\epsilon_0$, hence $C_m = 44.3$ pF after subtraction of stray capacitance. Disperse spheres with $\epsilon_i = 8.00\epsilon_0$ at $f = 0.0200$. The contrast factor is

$$(8 - 2)/(8 + 4) = 0.500.$$

Consequently

$$\epsilon_{\mathrm{eff}}/\epsilon_m \simeq 1 + 3(0.0200)(0.500) = 1.030.$$

The expected reading is 45.6 pF, an increase of approximately 1.33 pF.

A second sample with the same volume fraction but twice the sphere radius should give the same leading result. A systematic change might indicate settling, aggregation, finite-gap effects or frequency-dependent interface polarization. In the equal-permittivity limit the contrast factor vanishes and no capacitive signal should remain. In the high-permittivity limit b approaches one, so the fractional increase saturates at approximately $3f$ to first order. This bounded prediction gives a useful check when a simple arithmetic mixing rule appears to predict an arbitrarily large capacitance increase from an arbitrarily dilute concentration.

The hypothesis that the suspension is dilute can be tested without a microscopic field probe. Make a concentration series by adding material with known volume at fixed particle size. The predicted initial slope of relative capacitance versus volume fraction for the numerical pair is $3b = 1.50$. A linear fit restricted to the lowest concentrations estimates the contrast parameter and hence the inclusion permittivity if the matrix is known. Strong curvature at higher concentration then measures the breakdown of the independent-inclusion approximation. Extrapolating that curvature back as a new single-particle dielectric constant would blend collective and material effects.

Develop this calculation further

Chapter 18 · How microscopic dipoles produce macroscopic charge · Chapter 21 · What an interface does to phase, field and power

Study the full derivation

Anisotropic and nonlinear dielectrics from local measurements to global fields · Electrostatic field theory and measurement

A single dielectric constant cannot describe a crystal whose displacement tilts away from its electric field, or a reversible material whose incremental response changes with bias. Starting with the constitutive measurement clarifies which parts of ordinary electrostatics survive.

Problem sheet 4 · Far fields and induced moments

Infer moments and dielectric properties from potential measurements, including a measurement taken where a leading potential vanishes.

The axis and the midplane

Put charges $+2.00$ nC and -2.00 nC at $z = +3.0$ mm and -3.0 mm. Obtain the leading voltage at $z = 0.120$ m and the electric field at the same radial distance in the midplane. Explain which observable remains on the midplane.

▼ **Show the dipole calculation**

The separation is 6.0 mm, giving $p = 1.20 \times 10^{-11}$ C m along positive z. The axial potential is

$$kp/r^2 = 7.49\,\mathrm{V}.$$

The midplane potential vanishes exactly by symmetry. There the leading field is $-kp\,\mathbf{e}_z/r^3$, or 62.4 V/m towards negative z. Although tangent derivatives in the midplane vanish, the voltage changes in the perpendicular direction. The small source ratio $0.003/0.12 = 0.025$ makes the neglected relative finite-size corrections of order 0.025^2.

Measure the dielectric constant

A sphere of radius 3.0 mm in a field of 20.0 kV/m creates an additional axial voltage of 1.20 V at 15.0 mm. Determine its relative permittivity and predict the interior field. State one geometry change that tests this interpretation.

▼ **Show the inversion**

The dipole coefficient is $B = \Delta V r^2 = 2.70 \times 10^{-4}$ V m². Thus

$$s = B/(a^3 E_0) = 0.500.$$

From $s = (\epsilon_r - 1)/(\epsilon_r + 2)$ obtain

$$\epsilon_r = (1 + 2s)/(1 - s) = 4.00.$$

The uniform interior field is $3E_0/(\epsilon_r + 2) = 10.0$ kV/m. At twice the axial distance the extra voltage should fall to 0.300 V. Measuring the extra voltage requires subtracting the no-sphere reading, which increases in magnitude away from the origin.

A nearly neutral inclusion

A polarized sphere has induced moment 3.0×10^{-13} C m and a residual free charge of $+1.0$ pC. Find the axial distance where charge and dipole produce equal potential magnitudes. For two probes placed at opposite axial positions, derive combinations that isolate the two contributions.

▼ **Show the sum and difference**

Equating kQ/r and kp/r^2 gives $r = p/Q = 0.30$ m, assuming the probes are outside the sphere. After subtracting imposed background voltage, the two readings are $V_{\pm} = kQ/r \pm kp/r^2$. Their average isolates the monopole while their half difference isolates the dipole. At the crossover each isolated term has magnitude 0.0300 V. An even quadrupole would also enter the average, so repeat at a second radius to distinguish its faster r^{-3} potential decay.

Explore these relations

- For the axis and midplane, compare the exact source with its multipole approximation away from the charges. Open the live calculation.
- For the dielectric inference, vary relative permittivity and compare the sphere’s internal and external response. Open the live calculation.

Notebook practice · Why a leaky sphere can reverse its drift in an alternating field

Chapter 9 · Charge relaxation and magnetic diffusion

Distinguish two conductor timescales and estimate how deeply an alternating magnetic field penetrates a solid.

FROM FIRST PRINCIPLES

Trace two relaxation processes back to charge and current

A homogeneous ohmic conductor obeys the measured material law $\mathbf{J} = \sigma\mathbf{E}$. Local conservation says that charge decreases where current flows out. Take the divergence of the material law, then use Gauss's law for constant permittivity:

$$\partial_t \rho = -\nabla \cdot \mathbf{J} = -\sigma \nabla \cdot \mathbf{E} = -\frac{\sigma}{\epsilon}\rho,$$

$$\rho(t) = \rho(0) e^{-t\sigma/\epsilon}.$$

This equation describes local free-volume charge relaxation within the constant-conductivity model. The spatial redistribution of a magnetic field in the same body is a different measurement. At frequencies where displacement current is small compared with conduction current, write $\nabla \times \mathbf{B} = \mu\sigma\mathbf{E}$ and take a curl. Use Faraday's law and the absence of magnetic monopoles to cancel the gradient of $\nabla \cdot \mathbf{B}$:

$$-\nabla^2 \mathbf{B} = -\mu\sigma\, \partial_t \mathbf{B},$$

$$\partial_t \mathbf{B} = \frac{1}{\mu\sigma}\nabla^2 \mathbf{B}.$$

A magnetic disturbance varying on the length L has a diffusion time of order $\mu\sigma L^2$, which grows as the square of the size. A charge-relaxation time ϵ/σ contains no length. These estimates rely on the separate constitutive and frequency assumptions above. Extrapolating a static conductivity through arbitrarily high frequencies exceeds those assumptions.

DERIVATION AND APPLICATIONS

Follow a magnetic field step into a metal

A switched field probes a conductor over a growing depth. Take a conducting half-space $x > 0$ with constant permeability μ and conductivity σ. Initially its magnetic field is zero. At $t = 0$ impose a uniform tangential field B_s at its planar surface and maintain it thereafter. Work at times longer than charge relaxation and switching transients, neglect displacement current, and assume no variation parallel to the surface. Ampère's and Faraday's laws, together with Ohm's law, reduce to a diffusion equation for the tangential induction.

$$\partial_t B = D\partial_x^2 B,$$

$$D = \frac{1}{\mu\sigma}.$$

Since D has dimensions of area per time and the half-space supplies no length scale, try $B = B_s f(\eta)$ with $\eta = x/(2\sqrt{Dt})$. Differentiating at fixed x gives $\partial_t \eta = -\eta/(2t)$ and differentiating at fixed t gives

$$\partial_x \eta = 1/(2\sqrt{Dt}).$$

Common factors cancel, leaving

$$f'' + 2\eta f' = 0,$$
$$f' = Ce^{-\eta^2}.$$

The surface imposes $f(0) = 1$ and the unaffected interior gives $f(\infty) = 0$. Integrate the Gaussian slope from η to infinity and normalise its integral at the surface. The result fixes both the shape and the rate at which any chosen field contour advances.

$$\frac{B(x,t)}{B_s} = \operatorname{erfc}\left(\frac{x}{2\sqrt{Dt}}\right),$$
$$\operatorname{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^\infty e^{-s^2} ds.$$

This contour speed falls with time. A sinusoidal skin depth describes a steady, single-frequency measurement, while the step response has a penetration length that grows with the square root of elapsed time.

WORKED EXAMPLE

Time the half-field crossing in aluminium

Change the inputs and investigate this derivation.

Take $\sigma = 3.50 \times 10^7$ S/m and $\mu = \mu_0$. The diffusivity is 0.0227 m²/s. Half the applied field corresponds to

$$\operatorname{erfc}^{-1}(0.5) = 0.47694.$$

At $x = 2.00$ mm its arrival time is

$$x^2/[4D(0.47694)^2] = 0.193\,\text{ms}.$$

A sensor at 4.00 mm would reach the same fraction at 0.773 ms as long as the far surface remains sufficiently distant.

SECOND APPLICATION

Estimate conductivity from a single contour

Change the inputs and investigate this derivation.

In a thick nonmagnetic sample the 15.73 percent contour reaches 3.00 mm after 0.180 ms. Since $\operatorname{erfc}(1) = 0.1573$, this contour has $x = 2\sqrt{Dt}$. Its measured depth gives $D = (0.00300)^2/(4 \times 0.000180) = 0.0125$ m²/s and hence $\sigma = 1/(\mu_0 D) = 6.37 \times 10^7$ S/m. A conductivity inferred this way depends on knowing the permeability independently, since the field trace determines their product.

Try it · Magnetic step

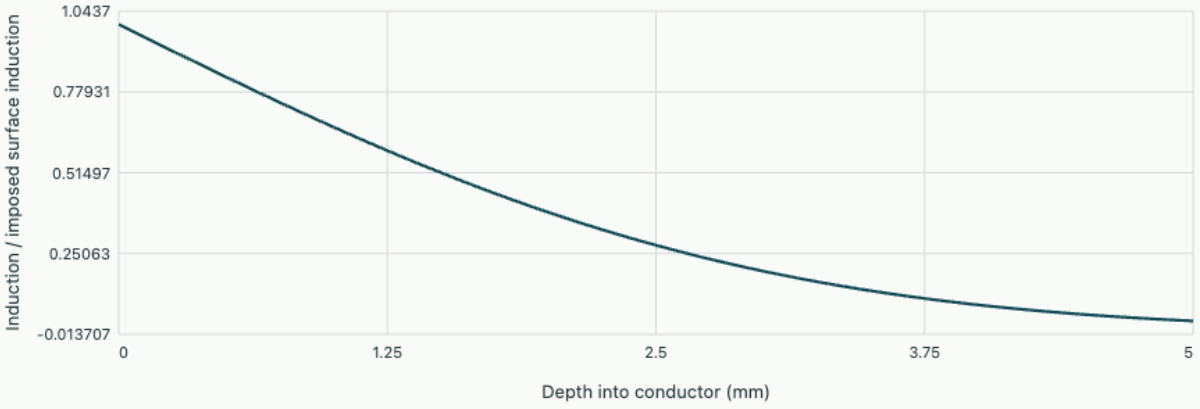
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Time after switching (ms) **0.2**

Conductivity (MS/m) **60**



Diffusion of a step field

D = 0.013263 m²/s; half-amplitude depth 1.5535 mm. Nonmagnetic uniform half-space with a suddenly imposed surface field.

How deep does a time-dependent field penetrate?

A plane harmonic field enters a homogeneous good conductor. Assume $\mu = \mu_0$ and conduction current much greater than displacement current. Mark the depth where amplitude is reduced to $1/e$ before comparing the magnitude with the in-phase component.

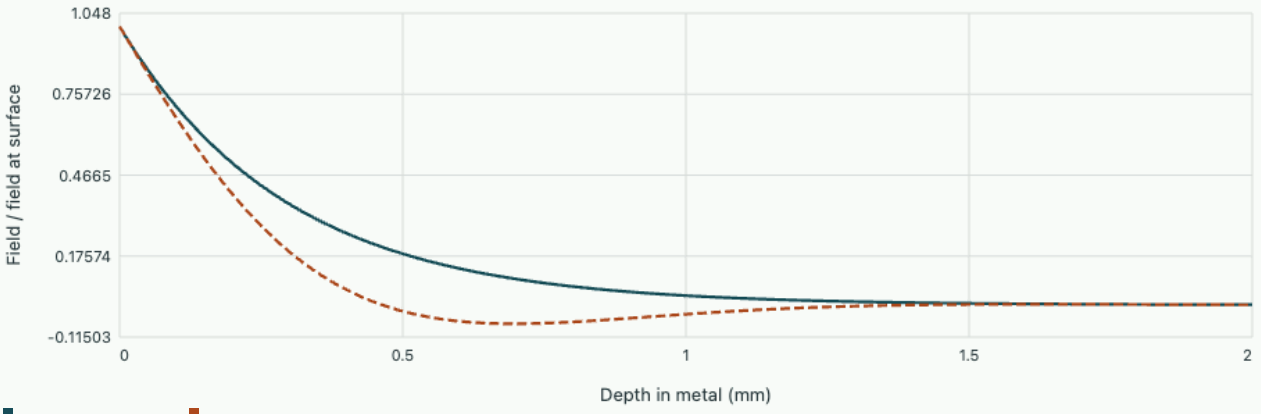
Combining Faraday's law with $\nabla \times H = \sigma E$ gives the diffusion equation $\partial_t E = (\mu\sigma)^{-1}\partial_z^2 E$. With time factor $e^{i\omega t}$ the decaying spatial root is $(1+i)/\delta$ where $\delta = \sqrt{2/(\omega\mu\sigma)}$. Only its real part controls attenuation. Its imaginary part shifts the phase.

Open this experiment online

Try the calculation

Frequency (kHz) 50

Conductivity (MS/m) 58



Skin depth 0.2955 mm. At this depth the magnitude has fallen to 1/e and the phase lags the surface by one radian.

Raise frequency fourfold and predict the shift of the first zero of the in-phase field. Use the graph to distinguish this zero from disappearance of the full harmonic amplitude.

Why charge accumulates at the interface between two weak conductors

A pair of plane electrodes sandwiches two slightly conducting dielectric layers in series. A voltage step is applied and held for a long time. At the first instant the voltage divides as in a capacitor. Much later it divides as in two resistors. The terminal voltage is unchanged throughout, so what physical quantity moved between these limits? A probe located at the internal interface reveals the charge, its timescale and the field that can persist even after the two outer electrodes are shorted. This is a simple example of interfacial, or Maxwell–Wagner, relaxation.

Match total current rather than conduction current alone

Let layer i have thickness d_i, absolute permittivity ϵ_i and conductivity $\sigma_i > 0$. Set positive field from electrode 1 through layer 1 and layer 2 toward electrode 2. Large area and negligible edge effects make the fields $E_1(t), E_2(t)$ uniform within each layer. Both media are linear, isotropic and nondispersive over the measurement band apart from their specified Ohmic currents. Assume no initial space charge, no interfacial reactions and a voltage rise slow compared with field transit time but fast compared with interfacial relaxation. When the applied voltage has become the constant V,

$$d_1E_1 + d_2E_2 = V,$$

$$\sigma_1E_1 + \epsilon_1\dot{E}_1 = \sigma_2E_2 + \epsilon_2\dot{E}_2.$$

The second equation is continuity of the sum of conduction and displacement current densities. The conduction currents alone can differ while free charge accumulates on the interface. Define $\sigma_f = \epsilon_2E_2 - \epsilon_1E_1$. Differentiating it and using the second equation yields $\dot{\sigma}_f = \sigma_1E_1 - \sigma_2E_2$, an independent local balance that fixes the sign of any predicted sheet charge.

Derive the response without guessing a circuit

Eliminate $E_2 = (V - d_1E_1)/d_2$. Since V is constant after the rise,

$$\dot{E}_2 = -d_1\dot{E}_1/d_2.$$

Collecting the terms gives

$$(\epsilon_1d_2 + \epsilon_2d_1)\dot{E}_1 + (\sigma_1d_2 + \sigma_2d_1)E_1 = \sigma_2V.$$

Before conduction has had time to move appreciable charge, $\sigma_f(0^+) = 0$, so

$$E_1(0^+) = \epsilon_2V/(\epsilon_1d_2 + \epsilon_2d_1).$$

In steady state the displacement derivatives vanish and $E_1(\infty) = \sigma_2V/(\sigma_1d_2 + \sigma_2d_1)$. The first-order equation therefore gives

$$E_1(t) = E_1(\infty) + [E_1(0^+) - E_1(\infty)]e^{-t/\tau},$$

$$\tau = \frac{\epsilon_1d_2 + \epsilon_2d_1}{\sigma_1d_2 + \sigma_2d_1}.$$

Substitution into the first voltage equation gives $E_2(t)$, and subtraction of the two normal displacements gives

$$\sigma_f(t) = \sigma_f(\infty)(1 - e^{-t/\tau}).$$

Its final value can have either sign even when both applied fields always point in the same direction. If $\sigma_1/\epsilon_1 = \sigma_2/\epsilon_2$, the initial and final voltage divisions coincide. Both materials conduct and dissipate power, but this particular interfacial charge does not build up.

Reconstruct an interface charge from a voltage trace

Choose $d_1 = d_2 = 1.00$ mm, $\epsilon_1 = 2\epsilon_0, \epsilon_2 = 6\epsilon_0$, $\sigma_1 = 1.00 \times 10^{-9}$ S/m and $\sigma_2 = 3.00 \times 10^{-10}$ S/m. With $V = 10.0$ V, the first layer initially drops 7.50 V but finally drops only 2.31 V. The time constant is 54.5 ms. After one time constant the first-layer drop is 4.22 V. The final sheet is positive, $\epsilon_0[6(7692) - 2(2308)] \simeq 3.68 \times 10^{-7}$ C/m². Equal steady conduction currents $2.31\ \mu\mathrm{A/m^2}$ through both layers independently confirm the final voltage division.

An alternative calculation uses admittances per area in the Laplace domain:

$$Y_i/A = (\sigma_i + s\epsilon_i)/d_i.$$

The fraction of total voltage across layer 1 is $Y_2/(Y_1 + Y_2)$. Its pole is at $s = -(\sigma_1 d_2 + \sigma_2 d_1)/(\epsilon_1 d_2 + \epsilon_2 d_1)$, precisely the time constant found without using a circuit. Its high- and low-frequency limits give the two electrode readings above. The familiar parallel resistor–capacitor representation of each layer now has an identifiable field origin and a way to test its validity.

Estimate conductivities from a step without measuring a tiny dc current

If the thicknesses and permittivities are known, the late-time voltage fraction and the measured relaxation time constrain two different combinations of the conductivities. In the equal-thickness example write $f = V_1(\infty)/V = \sigma_2/(\sigma_1 + \sigma_2)$. The transient gives the sum $\sigma_1 + \sigma_2 = (\epsilon_1 + \epsilon_2)/\tau$. Using $f = 0.231$ and $\tau = 54.5$ ms yields a sum of 1.30×10^{-9} S/m, with contributions $\sigma_2 \simeq 3.0 \times 10^{-10}$ S/m and $\sigma_1 \simeq 1.0 \times 10^{-9}$ S/m. The initial capacitive voltage division remains an independent check on the assumed permittivities and geometry.

This inversion becomes ill-conditioned if the initial and final voltage fractions nearly coincide. Although the formal time constant is still finite, the measured exponential then has very small amplitude and its decay time is hard to estimate against offsets or drift. A static leakage-current measurement can provide a more reliable conductivity constraint in that case. Fitting an arbitrary time constant to a trace that is experimentally flat would confuse a model parameter with a measured observable.

Short the electrodes after the charge has accumulated

Hold the previous setup until steady, then reduce the voltage quickly to zero and leave the outer plates shorted. Determine the two signed internal fields immediately after shorting, and decide what a late-time interface probe will record.

▼ **Derive the result and check it independently**

The fast change cannot instantaneously remove free charge trapped on the internal interface, so the displacement jump stays at its pre-existing value. By linearity superpose a fast -10 V step whose capacitive division changes E_1 by -7500 V/m and E_2 by -2500 V/m. Starting from the preceding steady fields leaves $E_1 = -5192$ V/m and $E_2 = +5192$ V/m. Equal thicknesses make their line integral zero, as required by the terminal short.

With $V = 0$ the same differential equation now has zero steady fields. Both internal fields decay to zero with 54.5 ms time constant, fed by the discharging interfacial sheet. Hence zero terminal voltage coexists temporarily with nonzero local electric fields, stored electrical energy and Ohmic heating. A numerical calculation enforcing continuity of normal displacement at all times would erase precisely the free interface charge that controls this memory. Check each field at the voltage step, charge conservation at the interface, and the eventual return to zero.

Fließbach, chapter 30 · applications of macroscopic field equations

What relaxes when a field is applied?

A metal plate inserted in a static electric field rapidly rearranges its free surface charge. A metal plate subjected to a changing magnetic field can carry eddy currents for an appreciable time. The two observations concern different field equations. Model a uniform material using $\mathbf{J} = \sigma\mathbf{E}$, $\mathbf{D} = \epsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$, treating the coefficients as constants over the frequencies used.

Free charge conservation gives

$$\partial_t \rho + \nabla \cdot \mathbf{J} = 0.$$

Within the uniform bulk, $\nabla \cdot \mathbf{E} = \rho/\epsilon$ gives

$$\partial_t \rho = -(\sigma/\epsilon)\rho.$$

A prepared bulk charge perturbation therefore decays exponentially with characteristic time $\tau_q = \epsilon/\sigma$. The derivation says nothing about the disappearance of total charge from an isolated body. Charge can move to its boundary, where the bulk equation with constant coefficients across the interface no longer applies. Extremely short formal times in good metals also probe frequencies beyond the constant-conductivity model.

Curl fields obey a different evolution

Use Ampère's law without displacement current when $\omega\epsilon \ll \sigma$ and the frequency lies within the material model. Combining $\nabla \times \mathbf{B} = \mu\sigma\mathbf{E}$ with Faraday's law gives

$$\nabla \times (\nabla \times \mathbf{B}) = -\mu\sigma\partial_t \mathbf{B}.$$

Because magnetic flux has zero divergence, the left-hand side is $-\nabla^2\mathbf{B}$. Thus

$$\partial_t \mathbf{B} = D_m \nabla^2 \mathbf{B},$$

$$D_m = \frac{1}{\mu\sigma}.$$

A spatial Fourier component with wavenumber k decays as $\exp(-D_m k^2 t)$. Long spatial scales persist longer because the electric field required to change a large magnetic flux drives stronger circulating currents. Dimensional analysis gives a diffusion time of order $\mu\sigma L^2$, with a geometric factor determined by the actual boundaries.

OneSimulation · Does a thin sheet have a field-free interior?

Place a 2 mm copper sheet between two coils that impose the same harmonic tangential magnetic field on its two faces. Neglect displacement current and edge fringing. For a phasor proportional to $e^{i\omega t}$, diffusion across the thickness becomes a complex boundary-value problem. Reflection symmetry fixes the slope at the centre.

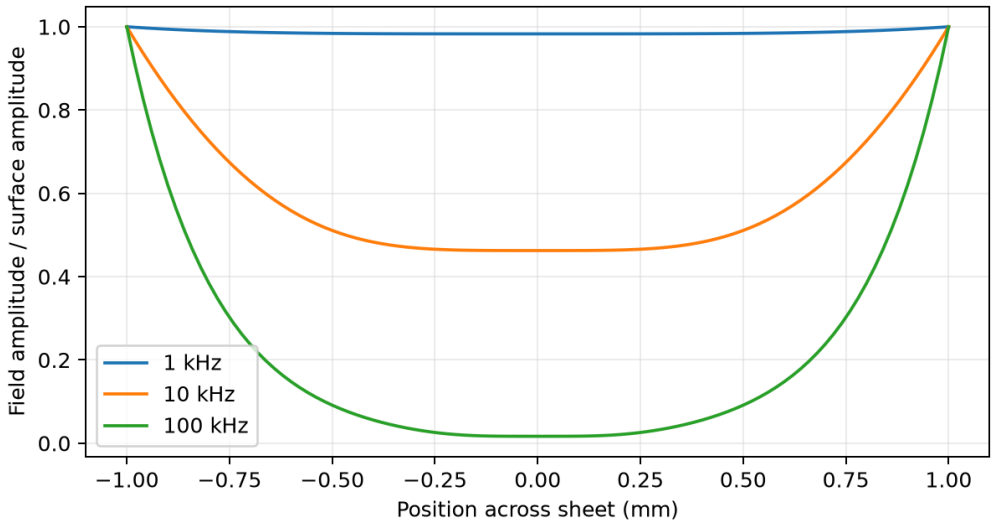
$$H'' = i\omega\mu\sigma H,$$

$$\frac{H(x)}{H_s} = \frac{\cosh[(1+i)x/\delta]}{\cosh[(1+i)a/\delta]},$$

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}}.$$

Here a is the half-thickness. The half-space law would keep a single decaying exponential. In this experiment the second face drives a second field with a definite phase. OneSimulation solves for both quadratures on a thickness grid. Compare the calculated centre amplitude with the two-face solution before using a measured attenuation to infer conductivity.

```
"""Harmonic magnetic diffusion in a copper sheet with two driven faces."""
import matplotlib.pyplot as plt
import numpy as np
from onesim.numerics.linear.matrices.dense_linear_algebra import solve_dense

mu, sigma = 4e-7 * np.pi, 5.8e7  # H/m, S/m
half_width = 1e-3  # m
x = np.linspace(-half_width, half_width, 101)
h = x[1] - x[0]
n = len(x) - 2
identity = np.eye(n)
minus_laplacian = 2 * identity - np.eye(n, k=1) - np.eye(n, k=-1)
boundary = np.zeros(n)
boundary[[0, -1]] = 1.0  # unit field at each face
fig, ax = plt.subplots(figsize=(6.6, 3.5), layout="constrained")

for frequency in (1e3, 1e4, 1e5):
    omega = 2 * np.pi * frequency
    alpha = omega * mu * sigma * h**2
    # Real and imaginary parts of the same complex diffusion equation.
    matrix = np.block([[minus_laplacian, -alpha * identity],
                       [alpha * identity, minus_laplacian]])
    u, v = np.split(solve_dense(matrix, np.r_[boundary, np.zeros(n)]), 2)
    field = np.r_[1, u + 1j * v, 1]
    delta = np.sqrt(2 / (omega * mu * sigma))
    exact = np.cosh((1+1j) * x / delta) / np.cosh((1+1j) * half_width / delta)
    assert np.max(np.abs(field - exact)) < 8e-4
    print(f"{frequency:g} Hz: centre amplitude {abs(field[n//2+1]):.4f}")
    ax.plot(1000 * x, abs(field), label=f"{frequency/1000:g} kHz")
ax.set(xlabel="Position across sheet (mm)", ylabel="Field amplitude / surface amplitude")
ax.legend()
ax.grid(alpha=0.2)
fig.savefig("electrodynamics.png", dpi=180)
```



Both surfaces have unit amplitude. At 10 kHz the interior still reaches 46% of that field. At 100 kHz the centre reaches less than 2%.

Try it · Magnetic sheet

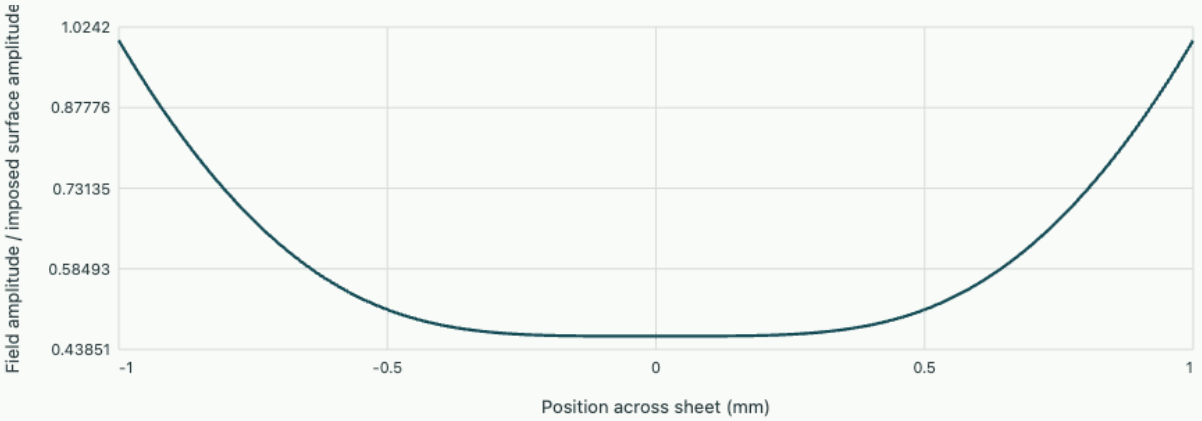
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Frequency (kHz) **10**

Sheet half-thickness (mm) **1**



Two faces driven in phase

Copper conductivity 5.8×10⁷ S/m; skin depth 0.66085 mm. Centre/surface amplitude 0.46269. The notebook solves the complex diffusion boundary-value problem independently.

Download the Python example

An alternating field outside a copper block

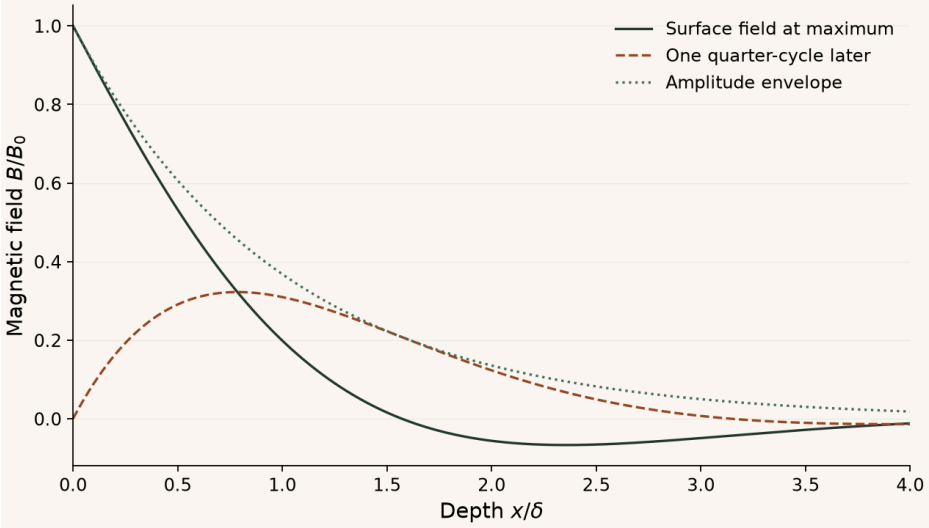
Take a tangential field at the plane $x = 0$ bounding a conductor occupying $x > 0$. Substitute $B(x,t) = \mathrm{Re}\{B_0 e^{i\omega t - \kappa x}\}$ into diffusion. The admissible square root of $\kappa^2 = i\omega\mu\sigma$ has positive real part,

$$\kappa = (1+i)/\delta,$$

$$\delta = \sqrt{2/(\omega\mu\sigma)}.$$

At depth δ the amplitude has dropped to $1/e$ and the phase lags by one radian. These are independent observables for testing the semi-infinite, constant-property solution. The surface excitation sets the boundary amplitude. It need not equal the field that an external coil would generate with the metal absent.

For copper take $\sigma = 5.8 \times 10^7$ S/m, $\mu = \mu_0$ and $f = 1.00$ kHz. Then $\delta = 2.09$ mm. At $x = 4.00$ mm the field amplitude is 0.148 of its surface value and the phase lag is 1.91 rad. Raising frequency by a factor of four halves the penetration depth. If a plate thickness is comparable to δ, fields from the opposite face also affect the solution and a second boundary condition must replace the one-sided decay condition. This geometric check matters as much as the conductivity value when using an amplitude ratio to infer a material property.



A conducting half-space in a sinusoidal field. Both the attenuation and the phase shift are set by the depth measured in skin depths.

The oscillating solution also predicts current concentrated near the surface, since Ampère's law makes the current density proportional to the local spatial derivative of magnetic field. As frequency rises, less of the metal cross-section transports the alternating current at a given field amplitude. This mechanism helps explain why a conductor's measured AC resistance can exceed its DC value even when the local conductivity is unchanged.

Relate surface fields to the power dissipated in a metal

The skin depth gives the distance over which current flows, yet a circuit measurement usually returns an impedance or a heating rate. Use the conducting half-space $x > 0$ and time convention $e^{i\omega t}$. Choose magnetic field along z, with phasor $H_z = H_0 e^{-(1+i)x/\delta}$. From Ampère's law without displacement current, the only current component is

$$J_y = -\partial_x H_z = (1+i) H_z/\delta.$$

Ohm's law therefore fixes the tangential electric field immediately inside the conductor.

$$Z_s = \frac{E_y(0)}{H_z(0)} = \frac{1+i}{\sigma\delta},$$

$$R_s = \operatorname{Re} Z_s = \sqrt{\frac{\pi f \mu}{\sigma}}.$$

The surface resistance R_s is expressed in ohms even though neither field in this ratio is a terminal voltage or current. Current integrated over depth is a sheet current with units A/m equal to the surface magnetic amplitude. The real part of the surface impedance relates mean heating per unit area to this current, while the imaginary part describes reversible magnetic energy storage near the surface.

For peak phasors, mean inward Poynting flux is

$$\operatorname{Re}[E_y H_z^*]/2 = R_s |H_0|^2/2.$$

The same result must follow from local heating. Insert $|J_y|^2 = 2|H_0|^2 e^{-2x/\delta}/\delta^2$ into $\int_0^\infty |J_y|^2/(2\sigma)\,dx$. The integral supplies $\delta/2$, reproducing the surface formula. This equality is a sensitive check on the distinction between amplitude decay over δ and squared-amplitude decay over $\delta/2$.

Estimate loss in a microwave wall

Take copper with constant $\sigma = 5.8 \times 10^7$ S/m, $\mu = \mu_0$ at $f = 10.0$ MHz. The skin depth is $\delta = 20.9$ µm and $R_s = 1/(\sigma\delta) = 0.825$ mΩ. A nearly uniform tangential peak magnetic field of 80.0 A/m illuminates a wall patch of area 0.0200 m². Surface heating density is $0.000825(80.0)^2/2 = 2.64$ W/m², corresponding to 52.8 mW for the patch. If the supplied field were an RMS amplitude the division by two would be inappropriate.

Increasing frequency fourfold at constant peak magnetic field doubles the loss and halves the skin depth. A coating 10 µm thick at the original frequency is less than one skin depth, so a calculation assuming all current decays within that coating is unreliable; substrate conductivity and the second interface matter. A very thick low-loss conductor can instead be represented by the surface condition without resolving the field throughout its interior. A numerical mode solver using that substitution should recover both the inferred loss and its square-root frequency dependence.

The complex sign in the impedance also makes a phase prediction. For the stated time convention the surface electric field leads the tangential magnetic field by 45 degrees. Squaring the electric amplitude

alone and dividing by the bulk resistivity would give a volumetric rather than an areal heating rate. To obtain an areal quantity one must integrate over its penetration profile. A measured phase substantially different from 45 degrees can indicate a magnetic permeability with appreciable losses, a finite film thickness, or a displacement current no longer negligible compared with conduction.

A calibrated surface-field phase measurement tests this condition independently of absolute field amplitude.

Develop this calculation further

Chapter 22 · Why dispersion and absorption belong together · Chapter 29 · When a circuit element must be treated as a field

Study the full derivation

How a magnetic field relaxes into a finite conductor · Magnetism, induction and field dynamics

A conducting sheet initially shields its interior from an applied magnetic step. Given time, a normal resistive metal admits a static tangential magnetic field. The same equations give a finite slab a discrete relaxation spectrum and a periodic drive a complex penetration depth. These are different initial and boundary conditions for one diffusion operator.

Chapter 10 · Waves in a hollow guide

Derive a rectangular guide's cutoff and use its dispersion relation to distinguish phase delay from signal transit time.

FROM FIRST PRINCIPLES

Why a conducting guide cuts off long waves

In vacuum, take a curl of Faraday's law and substitute the source-free Ampère–Maxwell law. Use $\nabla \cdot \mathbf{E} = 0$. Each Cartesian electric-field component then obeys

$$\nabla^2 E_y = c^{-2} \partial_t^2 E_y.$$

Consider a rectangular guide with perfectly conducting walls at $x = 0, a$. Take a TE mode with E_y independent of y and a travelling factor $e^{i(k_z z - \omega t)}$. The tangential electric field must vanish on the two side walls, so the lowest nonzero profile is $\sin(\pi x / a)$. Differentiating that sine twice contributes $-(\pi/a)^2$; differentiating the travelling factors gives

$$k_z^2 = \omega^2 / c^2 - (\pi/a)^2,$$
$$f_c = \frac{c}{2a}.$$

The boundary condition consumes part of the squared wave number in the transverse direction. Below f_c, k_z is imaginary and the solution bounded far down the guide decays. Above cutoff it is real. Differentiating the dispersion relation gives the speed of a narrowband envelope,

$$d\omega / dk_z = c^2 k_z / \omega.$$

Near cutoff it can be small even though the field between the metal walls still obeys the vacuum wave equation.

Background · Wave impedance

$$Z = E_t / H_t$$

The ratio has units of ohms. The field direction and mode determine the relevant tangential components. In a guide it can differ from the intrinsic impedance of the filling material.

DERIVATION AND APPLICATIONS

Measure penetration below waveguide cutoff

A guided TE mode can have nonzero field past a junction even below its propagation threshold. Consider the fundamental TE mode of an air-filled rectangular guide with broad dimension a and perfectly conducting walls. Let the section beyond the junction be uniform and extend far enough that there is no appreciable returning field. Its standing transverse profile fixes the transverse wavenumber at π/a. Inserting this profile into the homogeneous wave equation leaves an ordinary equation for longitudinal amplitude $F(z)$.

$$\frac{d^2F}{dz^2} + \left[\left(\frac{2\pi f}{c}\right)^2 - \left(\frac{\pi}{a}\right)^2\right]F = 0,$$

$$f_c = \frac{c}{2a}.$$

Below f_c the bracket is negative. Define its negative square root in magnitude as the positive decay coefficient α. Of the two exponential solutions only the one decaying towards the remote end satisfies the assumed boundary condition.

$$\alpha = \frac{\pi}{a}\sqrt{1 - (f/f_c)^2},$$

$$\frac{F(z)}{F(0)} = e^{-\alpha z}.$$

Suppose two identical weak probes respond to local squared electric-field amplitude at identical transverse coordinates. Forming their ratio cancels the unknown excitation and the transverse mode profile. Their readings in decibels then supply a slope from which cutoff can be recovered.

$$\frac{S(z_2)}{S(z_1)} = e^{-2\alpha(z_2 - z_1)},$$

$$10\log_{10}\frac{S(z_2)}{S(z_1)} = -\frac{20\alpha}{\ln 10}(z_2 - z_1).$$

A single decaying mode in a lossless semi-infinite guide stores reactive energy and has zero mean axial power. These probe readings track a local field. In a short section bounded by a second discontinuity, both exponentials can be required and axial power can pass through the combination.

WORKED EXAMPLE

Predict the depth dependence of a probe signal

Change the inputs and investigate this derivation.

For $a = 30.0$ mm and $c = 3.00 \times 10^8$ m/s, $f_c = 5.00$ GHz. At 4.00 GHz the decay coefficient is $(\pi/0.0300)(0.600) = 62.8$ m⁻¹. Over 20.0 mm the field amplitude falls to $e^{-1.257} = 0.285$ of its initial

value. A square-law probe reads 0.0810 of its initial signal, corresponding to a drop of 10.9 dB.

SECOND APPLICATION

Recover the guide width from two readings

Change the inputs and investigate this derivation.

At 6.00 GHz two identical probes 10.0 mm apart have signal ratio 0.250. Their field ratio is 0.500 and $\alpha = \ln 2/0.0100 = 69.3$ m⁻¹. The vacuum wavenumber is $2\pi f/c = 125.7$ m⁻¹, so $\pi/a = \sqrt{125.7^2 + 69.3^2} = 143.5$ m⁻¹. This gives $a = 21.9$ mm and cutoff 6.85 GHz, above the measuring frequency as required.

Try it · Below cutoff

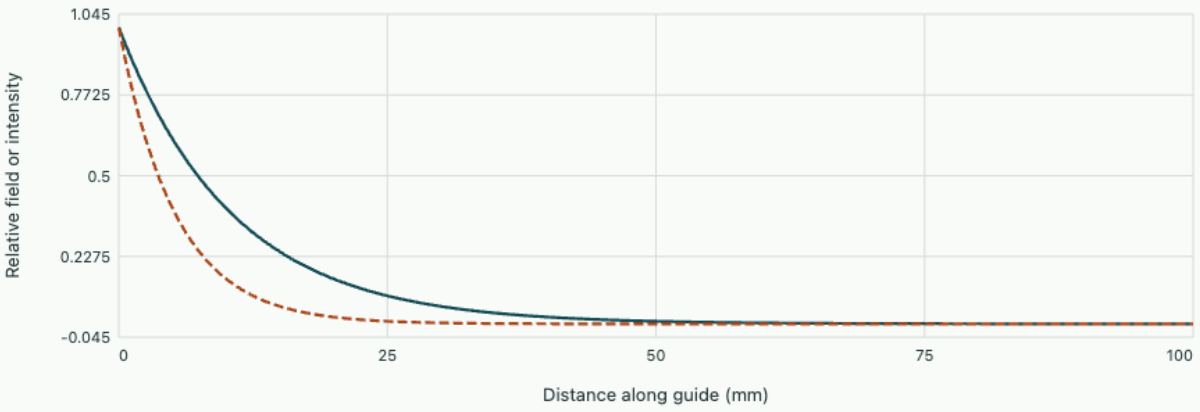
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Frequency / cutoff **0.8**

Guide width (mm) **20**



Field amplitude Squared-amplitude detector response

Decay constant 94.248 m⁻¹. An isolated evanescent mode carries no propagating axial power; the squared amplitude describes a local probe response.

What changes at the cutoff of a rectangular guide?

For the lowest transverse electric mode of an empty perfectly conducting rectangular guide, the wider side sets $k_c = \pi/a$. Move the width while holding the source frequency fixed. Predict where axial propagation becomes possible.

Separation of the transverse and longitudinal coordinates gives

$$k_z^2 = (2\pi f/c)^2 - k_c^2.$$

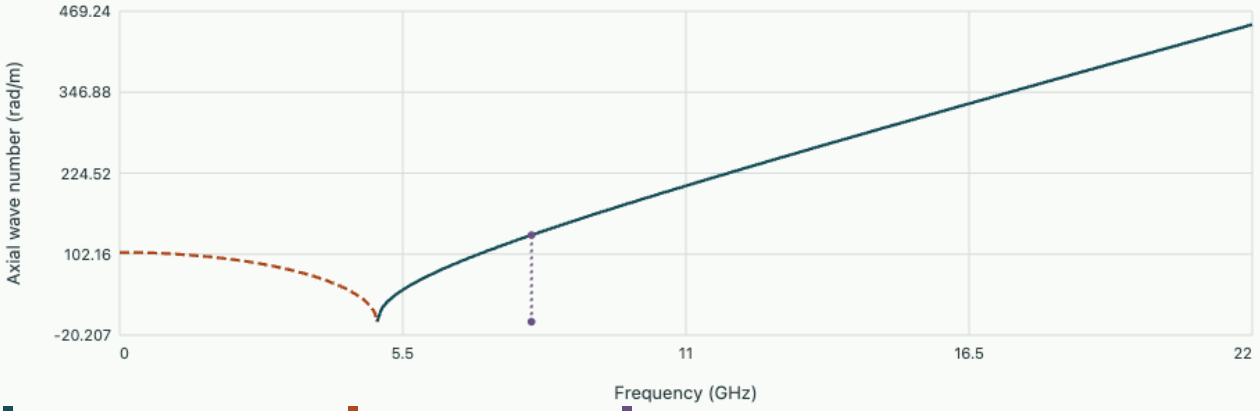
Above $f_c = c/(2a)$ the positive square root is the phase constant. Below it the magnitude of the imaginary root is the spatial decay rate. Both have units of inverse length and refer to different solutions.

Open this experiment online

Try the calculation

Guide width a (cm) 3

Operating frequency (GHz) 8



Propagation constant β Decay constant α Selected operating point

Cutoff 4.997 GHz. At 8.0 GHz the mode propagates with β = 130.94 m⁻¹.

Use 3 cm and 4 GHz. Read the attenuation constant and infer the distance for a factor e drop in amplitude. Widen the guide until 4 GHz propagates. Locate the boundary with the algebra and then with the operating-point marker.

How a narrow waveguide lengthens a pulse even with perfectly conducting walls

Launch a short microwave pulse into an empty, straight rectangular guide of width a, with only its lowest transverse electric mode excited. The walls are ideal conductors and the receiver is matched, so there is neither Ohmic loss nor a reflected wave to explain any change in its shape. Yet a received pulse can be longer and lower in peak amplitude than the transmitted one. Group delay explains the arrival of the

pulse's centre. To explain the altered width we must also find how that delay varies within its frequency band.

Follow the phase of each frequency separately

For the dominant mode the longitudinal propagation constant above cutoff is $k(\omega) = c^{-1}\sqrt{\omega^2 - \omega_c^2}$, with $\omega_c = \pi c/a$. The transverse mode shape is set by the walls; the frequency dependence below comes from its longitudinal wavenumber. With convention $e^{ikz - i\omega t}$, each spectral component accumulates phase $k(\omega)L$ in a length L. Let $\Omega = \omega - \omega_0$ be detuning around the positive carrier. For a spectrum concentrated well above cutoff, expand

$$k(\omega_0 + \Omega) = k_0 + k_0'\Omega + \tfrac{1}{2}k_0''\Omega^2 + \cdots,$$

$$k_0' = \frac{\omega_0}{c\sqrt{\omega_0^2 - \omega_c^2}},$$

$$k_0'' = -\frac{\omega_c^2}{c(\omega_0^2 - \omega_c^2)^{3/2}}.$$

The term $k_0' L$ shifts every frequency by the same first-order group delay. The negative k_0'' means higher frequencies have a smaller delay than lower frequencies. Thus the fast and slow parts of the spectrum separate in time even if the magnitude of each transmitted Fourier coefficient is unchanged. The expansion also predicts its own failure near cutoff, where the higher derivatives increase and some input spectral components may fall below cutoff altogether.

Carry a Gaussian envelope through the Fourier integral

Choose analytic input $E^+(0,t) = E_p e^{-t^2/(2T_0^2)} e^{-i\omega_0 t}$, with the physical field given by its real part. The spectral envelope is proportional to $e^{-T_0^2\Omega^2/2}$. Multiply by the propagation phase and set $s = t - k_0' L$. Completing the square in the inverse transform requires integrating $\exp[-(T_0^2 - ik_0''L)\Omega^2/2 - i\Omega s]$. The result, apart from the carrier, is

$$A(L,t) = \frac{E_p T_0}{\sqrt{T_0^2 - ik_0''L}} \exp\left[-\frac{(t - k_0'L)^2}{2(T_0^2 - ik_0''L)}\right].$$

Taking the magnitude gives $T_L = T_0\sqrt{1 + (k_0''L/T_0^2)^2}$, using the same width convention as the input. The magnitude peak ratio is $\sqrt{T_0/T_L}$. The factor compensates the broadening so that the time integral of $|A|^2$ stays fixed. The complex exponent also contains a quadratic phase in s; thus the instantaneous frequency varies across the outgoing envelope. A power meter sees the broader envelope, while a phase-sensitive receiver can test the chirp. One should distinguish an amplitude width T in $e^{-t^2/(2T^2)}$ from an intensity full width at half maximum $2\sqrt{\ln 2}\,T$.

Predict a flight time and time spread independently

Take $a = 20.0$ mm, $L = 4.00$ m and $f_0 = 9.00$ GHz. Then $f_c = 7.495$ GHz, $v_g = 1/k_0' = 0.554c$ and mean arrival delay is 24.1 ns. The dispersion coefficient is $k_0'' = -0.241$ ns²/m. For $T_0 = 1.00$ ns the width grows to 1.39 ns and the magnitude peak falls to about 85% of its input value. The guided carrier advances in phase at speed greater than c, but the group arrival and the causal front cannot be inferred from that phase speed.

Check the frequency sorting by evaluating L/v_g independently at two nearby frequencies. For weak bandwidth, their delay separation per angular-frequency separation must approach $k_0'' L$. A direct numerical inverse transform provides a further test: sample the input spectrum densely enough to resolve the rapidly varying phase, multiply by $e^{ik(\omega)L}$, and evaluate the time signal. Refine both bandwidth and frequency spacing. If the exact transform remains asymmetric after convergence, the quadratic expansion has reached its limit even when the solver itself has converged.

Estimate the frequency band in which a single dispersion coefficient suffices

The neglected cubic term in accumulated phase is $L k_0''' \Omega^3/6$. Differentiating the expression for k'' yields

$$k_0''' = 3\omega_0 \omega_c^2 / [c(\omega_0^2 - \omega_c^2)^{5/2}].$$

Relative to the quadratic term its magnitude is $\omega_0 |\Omega| / (\omega_0^2 - \omega_c^2)$. At a typical spectral detuning of $1/T_0$ this ratio is about 0.058 for the 9 GHz example. The ratio doubles at twice that detuning, although the spectral amplitude there is already smaller. Testing the accumulated cubic phase itself is also necessary if the guide is sufficiently long.

Thus a nominally narrow fractional bandwidth $\Delta f / f_0$ can still give a poor quadratic approximation near cutoff, where the smaller scale is $f_0 - f_c$. The analytic Gaussian also has arbitrarily early and late tails. Its fitted peak delay describes a prepared modulation and cannot time the first arrival of new information. A sharply switched source contains frequencies far outside the fitted Gaussian band, including components whose propagation approaches the vacuum light speed.

Test how close a measurement can approach cutoff

A second run uses $f_0 = 8.00$ GHz, $T_0 = 2.00$ ns and the same guide length. Estimate delay and width. How could a measured lower peak be distinguished from attenuation, filtering at cutoff or a reflection?

▼ **Derive the result and check it independently**

At 8.00 GHz one finds $v_g = 0.350c$, $k_0'' = -1.36$ ns²/m, delay 38.2 ns and $T_L = 2\sqrt{1 + (5.45/4)^2} = 3.38$ ns. Measure the magnitude of the frequency-domain transmission after correcting for port response. Pure dispersion changes phase while preserving each above-cutoff spectral magnitude in this model. A systematic magnitude reduction over the band indicates losses or coupling, a sharp suppression near the cutoff indicates spectral filtering, and a periodic ripple or time-separated echo suggests a reflection.

A second guide length supplies another check: delay varies linearly with L and $T_L^2 - T_0^2$ varies quadratically at fixed band and input width. If only a short time window is recorded, a broader pulse can be partly clipped and look as if transmission lost energy. Extend the integration window before evaluating that claim.

Griffiths, 5th ed., §9.5.2 · transverse electric modes in a rectangular guide

Fields constrained by metal walls

Can a hollow metal tube carry an arbitrarily slowly varying electromagnetic wave? Take perfectly conducting walls, a uniform vacuum interior, width a along x, height b along y and translation symmetry along z. Tangential electric field must vanish on each wall. In free space a plane wave can have uniform transverse fields. Inside a hollow simply connected guide a nonzero transverse electrostatic potential difference between separate conductors cannot be maintained. The lowest guided pattern consequently has transverse variation.

For the simplest transverse-electric mode set $E_y = E_m \sin(\pi x/a)\exp[i(\beta z - \omega t)]$, taking real parts for physical fields. At the side walls E_y is tangential and the sine vanishes. At the top and bottom the electric field is normal, as allowed by surface charge. Its zero divergence and accompanying magnetic field follow from Maxwell’s equations. In particular the curl includes an x derivative, giving a longitudinal magnetic component and the conventional label TE_{10}.

Separate transverse and longitudinal wavenumbers

Substitute the field into the vacuum wave equation. The two spatial derivatives contribute $-(\pi/a)^2$ and $-\beta^2$, while the time derivative contributes $-\omega^2/c^2$. Hence

$$\beta^2 = \frac{\omega^2}{c^2} - \left(\frac{\pi}{a}\right)^2,$$

$$f_c = \frac{c}{2a}.$$

Below cutoff the longitudinal dependence is exponential. In a long uniform section the bounded solution decays along the guide. A finite closed-off section can also have a growing exponential determined by reflections at its far end. Above cutoff the phase velocity is ω/β, while differentiating the dispersion relation gives the narrowband envelope velocity $d\omega/d\beta = c^2\beta/\omega$. Their product is c^2. Phase fronts move faster than c because the fixed transverse field variation contributes to total wavenumber. A modulation envelope in this lossless guide propagates more slowly than c.

Choose a usable frequency band

For a guide with $a = 22.86$ mm and $b = 10.16$ mm, the first cutoff is 6.56 GHz. Higher-order patterns have transverse wavenumber squared $(m\pi/a)^2 + (n\pi/b)^2$. The next candidate cutoffs are 13.11 GHz for

TE_{20} and 14.75 GHz for TE_{01}. Consequently a 10.0 GHz drive supports one propagating mode in this ideal guide.

At 10.0 GHz the group velocity is

$$c\sqrt{1-(6.56/10)^2} = 0.755c.$$

A 2.00 m section delays a narrow envelope by 8.84 ns. Its guided wavelength is $c/[f\sqrt{1-(f_c/f)^2}] = 39.7$ mm, exceeding the 30.0 mm free-space wavelength. Near cutoff the group velocity varies rapidly across a finite bandwidth, so quoting one envelope delay can conceal pulse distortion. Separately, real metal walls introduce attenuation and bends can excite evanescent modes even in the quoted single-mode band.

A practical way to measure the guided wavelength is to create a partial reflection at the termination. The forward and backward fields interfere and their stationary intensity maxima are separated by half the guided wavelength. At the example frequency the spacing should be about 19.9 mm. That measurement tests the propagation constant at one frequency. To test the group velocity, launch a short modulated burst or measure how transmitted phase changes between two nearby frequencies.

Close a guide to make a cavity

Placing conducting end plates at $z = 0$ and $z = L$ turns a propagating guide mode into a standing field with discrete frequencies. For the TE pattern of the previous derivation, the electric field has only E_y and is tangential to both end plates. The conductor therefore requires it to vanish there. A consistent mode has $E_y \propto \sin(\pi x/a)\sin(p\pi z/L)$ with positive integer p. The transverse electric field still points normal to the walls at fixed y, satisfying the original conditions. Combining longitudinal and transverse wavenumbers in the wave equation gives

$$f_{10p} = \frac{c}{2}\sqrt{\frac{1}{a^2} + \frac{p^2}{L^2}}.$$

The lowest such cavity resonance lies above the guide's cutoff even though no net power travels through the end plates. In a standing mode electric and magnetic energies exchange, and cycle-average net axial transport vanishes. Increasing cavity length at fixed p brings the frequency progressively nearer the guide cutoff. Conversely, if several longitudinal nodes fit into a short cavity, the longitudinal contribution can dominate the resonance frequency.

Weak finite conductivity or an output coupler causes stored energy to leave the mode. Define $Q = \omega U/P_{\mathrm{loss}}$ with U the cycle-averaged stored energy and P_{loss} the average dissipated or extracted power. When Q is approximately constant after switching off the drive, $dU/dt = -\omega U/Q$ gives an energy decay time Q/ω. A field-amplitude measurement follows the square root of energy and consequently has twice that time constant.

Identify a resonance and estimate a loaded quality factor

Choose $a = 30.0$ mm, $L = 40.0$ mm and cavity height 12.0 mm. The lowest member of this family has $f_{101} = (c/2)\sqrt{(1/0.0300)^2 + (1/0.0400)^2} = 6.25$ GHz using $c = 3.00 \times 10^8$ m/s. The parent guide's cutoff is 5.00 GHz. Raising longitudinal index to two gives $f_{102} = 9.01$ GHz. Measured field nodes along the cavity distinguish these patterns from other possible modes near the same frequency.

After terminating the drive at the first resonance, suppose a weak probe's voltage amplitude falls by a factor e in 2.00 µs. Then

$$Q = \omega \tau_{\text{amplitude}}/2 = \pi f \tau_{\text{amplitude}} = 3.93 \times 10^4.$$

The corresponding full resonance width is approximately $f/Q = 159$ kHz for a single weakly damped mode. A fit to detected power would return a 1.00 µs time constant instead. Increasing the probe coupling shortens the measured lifetime, so this ring-down gives the loaded quality factor. Measurements over several coupling strengths help separate internal wall losses from energy deliberately extracted by the probe.

The energy decay also predicts how much source power is needed to sustain a measured field. For example, if the resonant cavity in this measurement stores 2.00 nJ, total time-averaged loss is

$$U/\tau_U = 2.00\,\text{mW}.$$

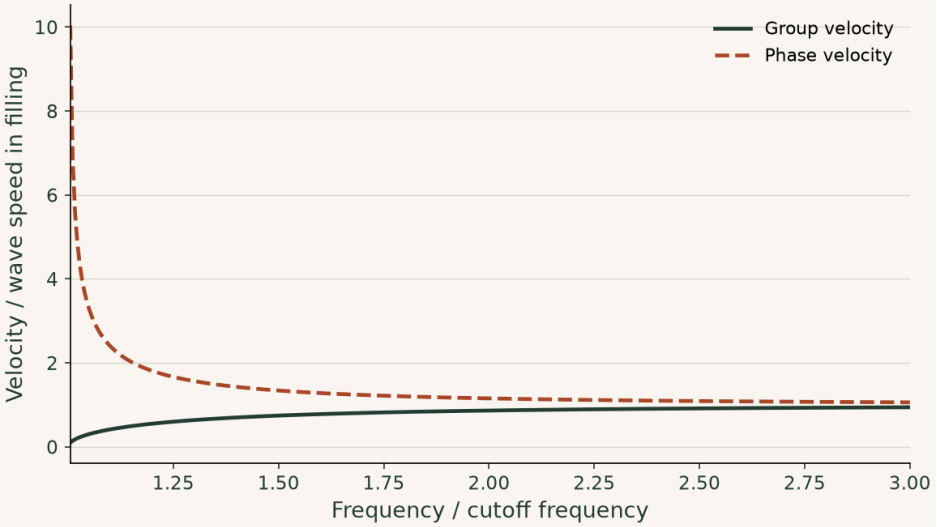
Source power actually coupled into the mode must replenish that loss in steady state. The source's forward power can be larger because some incident energy reflects at the coupling port. An independently measured coupling efficiency is therefore needed before converting a power-meter reading at the source into a stored field or a wall heating rate.

Read a waveguide delay near cutoff

In a lossless uniform guide with one propagating mode,

$$\beta^2 = (\omega^2 - \omega_c^2)/c_m^2.$$

Compute $v_p = \omega/\beta$ and $v_g = d\omega/d\beta$. The bandwidth of a pulse must remain inside this propagating band for its group-delay estimate to be useful.



At twice cutoff the group velocity is 0.866 times the medium wave speed and the phase velocity is 1.155 times it. A pulse delay follows the first of these. Both curves approach unity far above cutoff within the single-mode model.

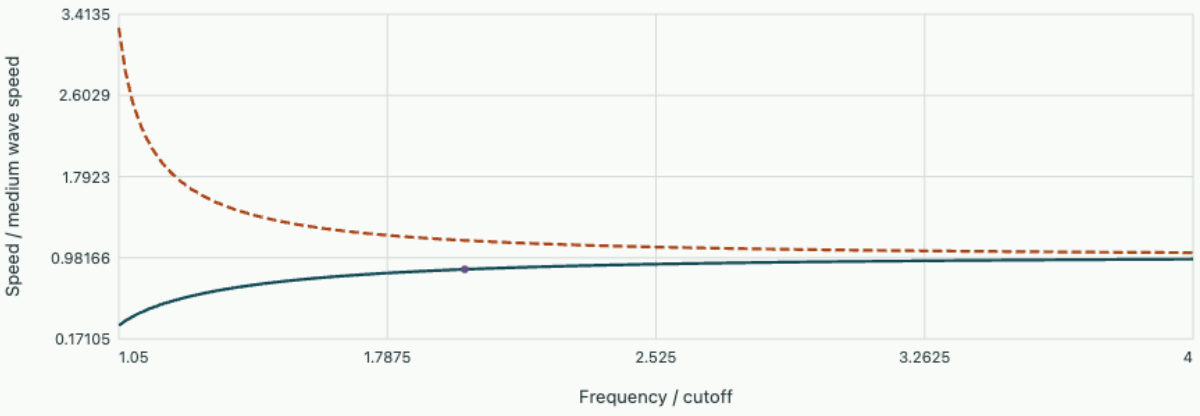
Try it · Waveguide speeds

Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Selected frequency / cutoff 2



Group velocity Phase velocity Selected group speed

Group × phase speed = medium wave speed squared. At the selected frequency the group ratio is 0.86603. Phase velocity exceeding c carries no superluminal information.

▶ **Read values from the plot**

Develop this calculation further

Chapter 27 · Cavity eigenmodes, measured linewidth and ring-down · Chapter 30 · From wavefronts to rays, focusing and caustics

Study the full derivation

A guided mode, a cavity line and the energy they carry · Waves, radiation and relativity

A microwave pulse sent through a rectangular guide arrives late even though the guide is evacuated. Terminating the guide converts the travelling pattern into a family of resonances. Use the same transverse eigenproblem to predict its cutoff, pulse delay, cavity frequency and measured linewidth.

Problem sheet 5 · Metal penetration and guide cutoff

Distinguish magnetic diffusion from waveguide cutoff by calculating their frequency and length dependence.

Invert a penetration measurement

At 400 Hz the tangential magnetic amplitude falls to $1/e$ over 4.00 mm in a homogeneous nonmagnetic half-space. Infer conductivity, predict the phase lag at 6.00 mm, and find the frequency that halves the penetration depth.

▼ **Show the diffusion calculation**

The measured length is δ. Inverting $\delta^2 = 1/(\pi f \mu_0 \sigma)$ gives $\sigma = 3.96 \times 10^7$ S/m. Depth divided by δ is the phase lag in radians, giving 1.50 rad with amplitude ratio $e^{-1.5} = 0.223$. Since δ scales with $f^{-1/2}$, a 1.60 kHz drive gives 2.00 mm. A second nearby boundary or frequency-dependent permeability would spoil this inference.

A microwave pulse between walls

An empty rectangular guide has width 30.0 mm and height 12.0 mm. Find the first two cutoffs. At 8.00 GHz determine the group velocity and envelope transit time over 3.00 m. Use $c = 3.00 \times 10^8$ m/s.

▼ **Show the allowed band**

The first cutoff $c/(2a)$ is 5.00 GHz. Among the next candidates $c/a = 10.0$ GHz and $c/(2b) = 12.5$ GHz, the former is lower. Thus only TE_{10} propagates at 8 GHz. Its group velocity is $c\sqrt{1-(5/8)^2} = 2.34 \times 10^8$ m/s and transit takes 12.8 ns. The free-space value would be 10.0 ns. This applies to a narrowband excitation with negligible end-reflection delays.

Two decaying fields

Use the same guide below cutoff at 4.00 GHz. Derive its amplitude decay length, compute the amplitude ratio after 30.0 mm and explain how increasing frequency distinguishes this measured decay from conductor skin effect.

▼ **Show the field comparison**

Subcutoff $\beta = i\alpha$ with $\alpha = (2\pi/c)\sqrt{f_c^2 - f^2} = 62.8$ m^{-1}. Its amplitude decay length is 15.9 mm, giving $e^{-62.8(0.030)} = 0.152$ after 30 mm in a long uniform guide. Raising frequency towards cutoff increases this length, whereas a good conductor's skin depth decreases. An ideal evanescent mode in a semi-infinite empty guide does not dissipate power. Metal penetration with finite conductivity does dissipate power through $\mathbf{J} \cdot \mathbf{E}$.

Explore these relations

- For the penetration measurement, vary frequency and conductivity and compare attenuation with phase lag. Open the live calculation.
- For the microwave pulse, compare phase and group speeds above cutoff. Open the live calculation.

Chapter 11 · Pulses on a transmission line

Use charge and flux conservation to predict the first voltages seen when a pulse meets a mismatched load.

FROM FIRST PRINCIPLES

From a short cable segment to two travelling waves

Consider voltage $V(z,t)$ from inner conductor to return, and current $I(z,t)$ towards increasing z. A short, uniform, lossless line segment of length Δz stores charge $C'\Delta z\,V$ and links magnetic flux $L'\Delta z\,I$. No current leaks through the dielectric. The current difference must charge the segment, and the voltage drop must drive its changing flux:

$$I(z) - I(z+\Delta z) = C'\Delta z\,\partial_t V,$$
$$V(z) - V(z+\Delta z) = L'\Delta z\,\partial_t I.$$

Divide by Δz and take its limit. Differentiate the voltage equation once more in space and eliminate the current derivative using the charge balance:

$$\partial_z V = -L'\partial_t I,$$
$$\partial_z I = -C'\partial_t V,$$
$$\partial_z^2 V = L'C'\partial_t^2 V.$$

For a forward wave $V = F(t - z/v)$, comparing second derivatives gives

$$v = 1/\sqrt{L'C'}.$$

Substituting in the first-order relation shows the forward current is

$$I = V/Z_0,$$

with $Z_0 = vL' = \sqrt{L'/C'}$ for the travelling part. A backward wave reverses that current sign. The line can therefore have a finite characteristic impedance even when its conductors dissipate no heat.

DERIVATION AND APPLICATIONS

Sum repeated returns on a pulse transmission line

A step applied to a finite line can produce a staircase at the load. Model a uniform lossless transmission line with real characteristic impedance Z_0, one-way travel time τ, resistive termination R_L and a source step V_s in series with resistance R_s. Initially the line is unenergised. Until a reflection comes back, the source sees the characteristic impedance, giving the launched voltage V_+. On reaching either termination an incident voltage wave produces a reflected fraction set by that termination.

$$V_+ = V_s \frac{Z_0}{R_s + Z_0},$$
$$\Gamma_L = \frac{R_L - Z_0}{R_L + Z_0},$$
$$\Gamma_s = \frac{R_s - Z_0}{R_s + Z_0}.$$

For the source reflection, keep its step height fixed. The incremental wave therefore encounters just R_s. The first arriving step at the load changes its voltage by incident plus reflected amplitudes. Each further round trip multiplies the corresponding incident step by $\Gamma_L\Gamma_s$. Index the first arrival by $n = 0$.

$$\Delta V_{L,n} = (1 + \Gamma_L)V_+(\Gamma_L\Gamma_s)^n,$$
$$t_n = (2n + 1)\tau.$$

The voltage after the nth arrival is the partial sum of these steps. Provided the magnitude of the round-trip multiplier is below unity, the series converges. Substitution of the two voltage reflection coefficients recovers the expected DC divider exactly.

$$V_{L,\infty} = \frac{(1 + \Gamma_L)V_+}{1 - \Gamma_L\Gamma_s} = \frac{V_s R_L}{R_s + R_L}.$$

Successive steps alternate when the reflection coefficients have opposite signs. The spacing of changes measures travel time even when delays in the measuring electronics obscure the absolute launch time.

WORKED EXAMPLE

Predict overshoot at a resistive receiver

Change the inputs and investigate this derivation.

Use $V_s = 3.00$ V, $R_s = 25.0\ \Omega$, $Z_0 = 50.0\ \Omega$, $R_L = 100\ \Omega$ and $\tau = 20.0$ ns. The launched step is 2.00 V. The load and source coefficients are $+1/3$ and $-1/3$. The receiver rises to 2.667 V at 20.0 ns, then falls by $2.667/9 = 0.296$ V at 60.0 ns. Its eventual 2.400 V follows from the divider. Measuring only the first plateau would overestimate the steady voltage.

SECOND APPLICATION

Read the source resistance from staircase increments

Change the inputs and investigate this derivation.

A separate 75.0 Ω line has a 225 Ω termination. Two consecutive load increments are $+0.800$ V and $+0.0800$ V, 50.0 ns apart. Their ratio is $0.100 = \Gamma_L\Gamma_s$ and $\Gamma_L = 0.500$, yielding $\Gamma_s = 0.200$. Inverting gives

$$R_s = 75.0(1.20/0.800) = 112.5\,\Omega.$$

The one-way time is 25.0 ns, corresponding to 5.00 m at a measured propagation speed of 2.00×10^8 m/s.

Try it · Line steps

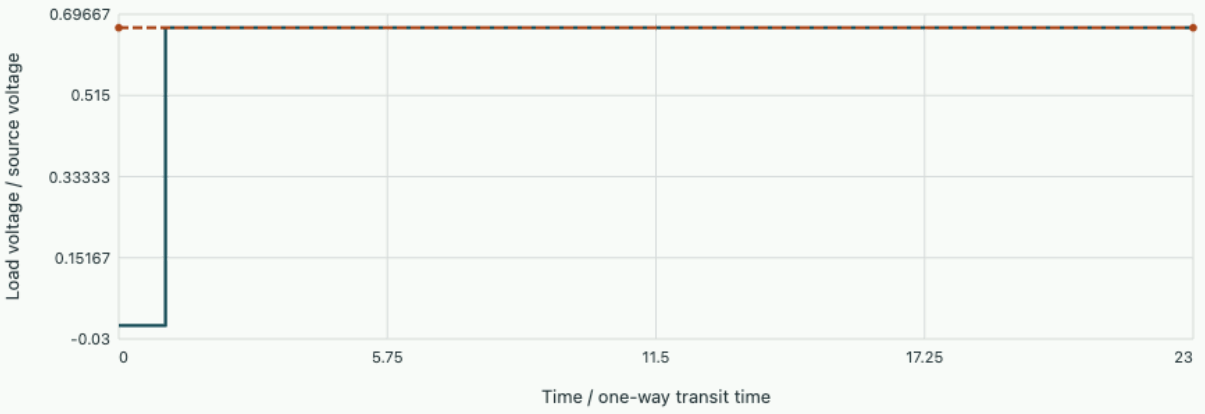
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Source / line resistance 1

Load / line resistance 2



ΓL=0.33333, Γs=0. Each returning increment is multiplied by ΓLΓs; the final load voltage is RL/(Rs+RL).

Start from this chapter. Predict how changing only the permittivity alters both wave speed and characteristic impedance. The OneSimulation solve supplies transverse line constants for the pulse calculation.

OneSimulation laboratory · Recover cable impedance from a cross-section

Watch a pulse reflect from a load

A lossless metre of transmission line has speed 2.0×10^8 m/s and characteristic impedance 50 Ω. Launch a Gaussian voltage pulse from a matched source. It peaks at the source at 3 ns. Predict when it peaks at the far end and when the return reaches the source.

At the load, voltage adds and currents subtract. Substituting $V_L = Z_L I_L$ gives

$$\Gamma = (Z_L - Z_0)/(Z_L + Z_0).$$

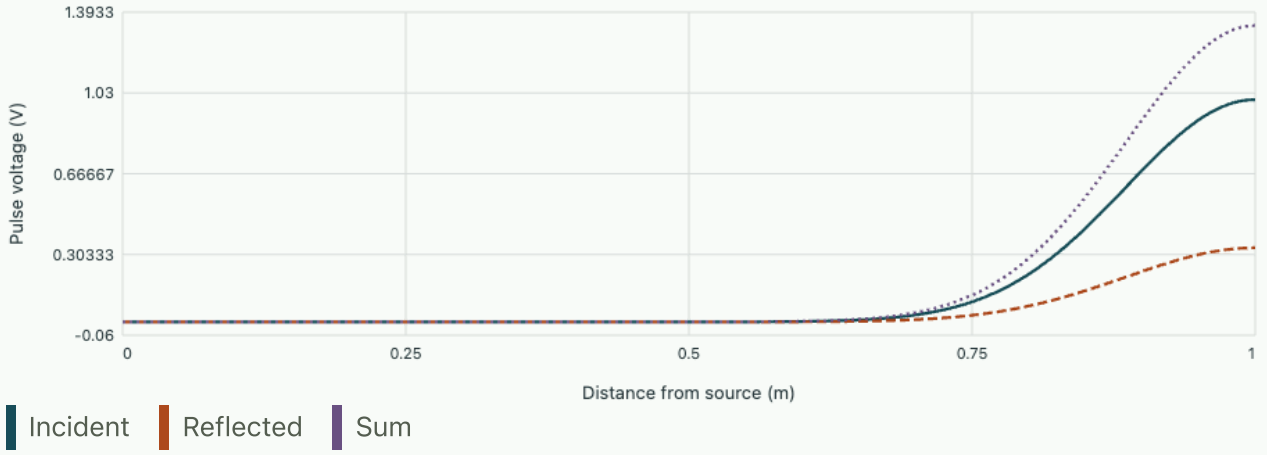
An observer at position x sees the incident shape delayed by x/v, and the reflected shape by $(2L - x)/v$. This supplies an independent travel-time check for the moving curves.

Open this experiment online

Try the calculation

Load (Ω) 100

Time after launch (ns) 8



Load reflection coefficient 0.333. At the load the voltage is 1.333 V. The matched source absorbs the return.

Sweep time past 8 ns and 13 ns. At a short, explain why zero load voltage coexists with a finite reflected pulse. At a matched termination, trace the energy carried by the incident pulse.

A cable can attenuate a pulse without distorting its shape

A laboratory sends two pulses down a long uniform cable. A shorter pulse has more high-frequency content, so it is often expected to distort more. That expectation implicitly assumes a particular frequency dependence of the propagation constant. We will derive a case in which an entire waveform is simply delayed and reduced in amplitude, even though the line converts energy to heat. The calculation also identifies when adding a second loss mechanism changes the phase response and how to distinguish pure attenuation from imperfect matching at the ends.

Give both energy-storage mechanisms their own losses

Let L', C' be the series inductance and shunt capacitance per length. Include a series resistance R' per length and a shunt conductance G' per length. Assume that these four coefficients are positive constants

over the pulse band, the line is uniform, and one transverse electromagnetic mode suffices. Charge and voltage balances on a short segment give

$$-\partial_z V = R' I + L' \partial_t I,$$
$$-\partial_z I = G' V + C' \partial_t V.$$

With time convention $e^{-i\omega t}$ and a forward spatial dependence $e^{-\Gamma z}$, their product and ratio yield

$$\Gamma^2 = (R' - i\omega L')(G' - i\omega C'),$$
$$Z_{\mathrm{line}}^2 = (R' - i\omega L')/(G' - i\omega C').$$

Choose the forward branch $\Gamma = \alpha - i\beta$, $\alpha \geq 0, \beta \geq 0$, with the impedance branch satisfying the original first-order equations for a forward wave. In general both amplitude and phase change differently for different frequencies. A matched receiver must account for the actual line impedance rather than just a nominal lossless value.

Find the exact condition for undistorted propagation

Suppose $R'/L' = G'/C' = \nu$, with units of inverse seconds. Both complex factors then share $\nu - i\omega$, so $\Gamma = \sqrt{L'C'}(\nu - i\omega)$ and $Z_{\mathrm{line}} = \sqrt{L'/C'}$ is real and frequency independent. Define $u = 1/\sqrt{L'C'}$ and $\alpha = \nu/u$. The forward transfer function over distance z multiplies each voltage component by $e^{-\alpha z} e^{+i\omega z/u}$. Fourier inversion thus gives

$$V(z,t) = e^{-\alpha z} V(0, t - z/u),$$
$$I(z,t) = V(z,t)/Z_{\mathrm{line}}.$$

The delay sign can be checked by inserting the result in the time-domain equations. In fact V satisfies

$$\partial_z^2 V = L'C'(\partial_t + \nu)^2 V.$$

One-way solutions satisfy the first-order propagation operator $\partial_z + u^{-1}(\partial_t + \nu)$. Direct substitution gives zero for any differentiable input shape. Piecewise smooth pulses follow by continuity or by their Fourier representations. This is the Heaviside distortionless condition. It permits attenuation; it is a relationship between two loss-to-storage ratios.

Verify that the missing power becomes heat

Multiply the two original equations by I and V respectively, then add. The product rule yields

$$-\partial_z (VI) = R' I^2 + G' V^2 + \partial_t \left(\frac{1}{2} L' I^2 + \frac{1}{2} C' V^2 \right).$$

The first two terms are positive local heating rates per length. For a forward pulse $I = V/Z_{\mathrm{line}}$. At the matched loss ratios,

$$R'/Z_{\mathrm{line}}^2 = G'.$$

Conductor heating and dielectric leakage heating are therefore equal at fixed local voltage, although the resistances of the two physical paths play different roles. Integrating over all time removes the storage derivative if the line begins and ends quiet. The pulse energy crossing each section decays as $e^{-2\alpha z}$, the square of its voltage attenuation. This provides a local and an integrated measurement of the same physical loss.

Take $L' = 0.25$ µH/m and $C' = 100$ pF/m. Then $u = 2.00 \times 10^{8}$ m/s and $Z_{\mathrm{line}} = 50.0\ \Omega$. For $R' = 0.050\ \Omega/\mathrm{m}$, choose $G' = 20.0$ µS/m, making both ratios 2.00×10^{5} s^{-1}. A 100 m line delays every part of a matched pulse by 0.500 µs and multiplies its voltage by $e^{-0.10} = 0.905$. Its transmitted energy is 81.9% of the launched energy, leaving 18.1% as loss along the line. A pulse measured to be ten per cent lower in voltage has therefore lost more than ten per cent of its transported energy.

Show why a constant attenuation and a linear phase are restrictive

The sufficient condition can also be obtained as a necessity if exact distortionless propagation is required over a continuous frequency range in this constant-coefficient model. Suppose $\Gamma = \alpha - ib\omega$ with constant attenuation and delay per length b. Squaring and matching constant, linear and quadratic terms against $(R' - i\omega L')(G' - i\omega C')$ yields $\alpha^2 = R'G'$, $b^2 = L'C'$ and $2\alpha b = R'C' + L'G'$. Eliminating α, b leaves $(R'C' - L'G')^2 = 0$. A line with only series resistance and zero shunt loss cannot satisfy that exact identity unless it is lossless altogether.

This test uses the functional dependence on frequency, rather than one measured attenuation value. In a real cable the skin effect makes R' increase with frequency, and a dielectric may have frequency-dependent polarisation loss. The two traces from two cable lengths therefore test a specified band of operation. An approximately straight phase line over a limited band can be entirely consistent with the measurements while extrapolating that fit to a much shorter pulse gives appreciable waveform error.

Estimate the effect of loss mismatch

A second cable has the same R', L', C' but negligible G'. Does weak loss recover the same exact propagation law? Derive its leading high-frequency attenuation and give a procedure for detecting the missing exact cancellation.

▼ Derive the result and check it independently

When $\omega L' \gg R'$ and $\omega C' \gg G'$, expand the square root around its lossless form. To first order the attenuation is $\alpha \simeq [R'/Z_{\mathrm{lossless}} + G' Z_{\mathrm{lossless}}]/2$, with $Z_{\mathrm{lossless}} = \sqrt{L'/C'}$; the phase is approximately ω/u. With negligible G' the asymptotic attenuation is 0.00050 m^{-1}. This is an approximation over a band. At sufficiently low frequency R' competes strongly with $\omega L'$; then Γ behaves like the square root of frequency and the waveform distorts.

Send the same calibrated pulse through two lengths, compare complex spectra and form their ratio to suppress the source response. Under distortionless propagation the logarithm of its magnitude is constant across frequency and its unwrapped phase is linear. Periodic ripples can indicate connector reflections. Time-gating a clean first arrival, or matching the relevant band and checking a second cable length, helps

distinguish distributed distortion from an echo. A single rise-time reading cannot by itself identify the underlying loss mechanism.

Griffiths, 5th ed., §9.5.3 · fields of a coaxial transmission line

A line resolves the travel time

A cable can store energy all along its length. When a change of source voltage travels an appreciable fraction of the cable during its rise time, a single lumped capacitor cannot encode both its arrival and its reflection. Model a uniform two-conductor line using inductance L' and capacitance C' per unit length. Omit conductor resistance and dielectric leakage initially. Let $V(z,t)$ be voltage between conductors and $I(z,t)$ current toward increasing z.

A short section dz experiences an inductive voltage difference $L'dz\,\partial_t I$. Its net incoming current charges $C'dz$ at the rate $\partial_t V$. Accounting for the positive current direction gives the pair

$$\partial_z V = -L'\partial_t I,$$

$$\partial_z I = -C'\partial_t V.$$

Differentiating once more gives wave equations with $v = 1/\sqrt{L'C'}$. Substitution of a travelling waveform shows that a forward wave has $I_+ = V_+/Z_0$, while a backward wave has $I_- = -V_-/Z_0$, where $Z_0 = \sqrt{L'/C'}$. The characteristic impedance thus describes the voltage-to-current ratio of one travelling component. A resistance meter at the end of a finite DC cable ordinarily measures something different.

How the load launches a return wave

At a resistive termination R_L the total voltage equals R_L times total current. Write $V_- = \Gamma V_+$ and use the opposite current signs above. Rearrangement gives

$$\Gamma = \frac{R_L - Z_0}{R_L + Z_0},$$

$$V_L = (1+\Gamma)V_+.$$

An open end doubles voltage and cancels current. A shorted end cancels voltage and doubles current. With positive real Z_0 the reflected fraction of incoming wave power is Γ^2. A voltage reflection of either sign carries energy back toward the source. A matched resistor absorbs all the incident energy, a condition also useful for interpreting the first step in a time-domain voltage trace.

Locate and characterize a mismatch

A line with $L' = 250$ nH/m and $C' = 100$ pF/m has $v = 2.00 \times 10^8$ m/s and $Z_0 = 50.0$ Ω. A 2.00 V source step with series resistance 50.0 Ω launches 1.00 V into a 30.0 m cable. Initially the source sees the characteristic impedance because no reflection has had time to return. At 150 ns the leading edge reaches a 150 Ω load, where $\Gamma = 0.500$. The load rises to 1.50 V and draws 10.0 mA.

The return step of 0.50 V reaches the source at 300 ns. The source resistance is matched, so it absorbs the returning wave with no further reflection and the source-end voltage becomes 1.50 V. The result agrees with the final DC divider. A measured return time locates the discontinuity through $d = v t_{\mathrm{return}}/2$, while its signed amplitude identifies an impedance increase or decrease. Distributed losses must be estimated before using this amplitude ratio on a long real cable.

Changing the source resistance leaves the first arrival time unchanged while altering the incident amplitude. It also changes what happens when the first reflection returns. At a source whose internal resistance differs from the line impedance, part of that returning pulse is launched outward again. Repeated round trips then generate a staircase of voltages separated by twice the propagation time.

Match unequal impedances with a finite line segment

A resistive load can be matched at a selected frequency even when its resistance differs from the feeding line’s characteristic impedance. Insert a uniform lossless section of impedance Z_t and length ℓ between the original line Z_0 and the load R_L. Inside the added section decompose voltage into forward and backward waves. With the load plane at $z = 0$, a convenient phasor convention gives

$$V(z) = V_+ e^{-i\beta z} + V_- e^{i\beta z},$$

$$I(z) = (V_+ e^{-i\beta z} - V_- e^{i\beta z})/Z_t.$$

At $z = 0$ use $V(0)/I(0) = R_L$ to remove the unknown reflection coefficient.

Evaluate the ratio at the input plane $z = -\ell$. Collecting the exponentials into cosines and sines gives

$$Z_{\mathrm{in}} = Z_t \frac{R_L \cos(\beta\ell) + i Z_t \sin(\beta\ell)}{Z_t \cos(\beta\ell) + i R_L \sin(\beta\ell)}.$$

For an electrically short segment the input simply approaches the load. At a quarter wavelength $\beta\ell = \pi/2$ and $Z_{\mathrm{in}} = Z_t^2/R_L$. Setting this equal to Z_0 gives

$$Z_t = \sqrt{Z_0 R_L}.$$

Reflections still occur inside the added section, but their summed contribution at its input cancels at the design frequency. This is a phase-dependent match, so its behaviour under a suddenly applied step differs from a resistor matched at every frequency.

Design a section and test its bandwidth

A 50.0 Ω source line drives a 100 Ω resistor at 100 MHz. Available cable has phase speed 2.00×10^8 m/s. The required matching impedance is 70.7 Ω and the length is

$$v/(4f) = 0.500\,\mathrm{m}.$$

At design frequency the input impedance equals 50.0 Ω, whereas directly attaching the resistor would produce reflection amplitude $(100 - 50)/(100 + 50) = 1/3$ and reflect 11.1 percent of the incident power.

To test the frequency limit without dividing by a nearly infinite tangent, evaluate the sine–cosine form at 50.0 MHz. The electrical length is $\pi/4$, so the trigonometric factors cancel. Substituting the three real impedances gives $Z_{\text{in}} = 66.7 - 23.6i$ Ω. The feeding-line reflection magnitude $|(Z_{\text{in}} - 50)/(Z_{\text{in}} + 50)|$ is about 0.243, reflecting 5.9 percent of incident power. At 200 MHz the section is half a wavelength and the input again equals the bare load. A frequency sweep should therefore show a matched centre frequency and increasing reflected power away from it. Loss or an appreciably reactive load changes both the depth and position of that minimum.

A frequency-domain match can also be understood through a short series of time-domain reflections. The impedance first rises from 50.0 to 70.7 ohms and then rises again from 70.7 to 100 ohms. The two interface reflections have the same sign when considered locally. At a quarter wavelength, propagation through the matching section and back adds a half-cycle phase, making the returning contributions cancel after all internal reflections are summed. An abrupt step arrives at those interfaces at different times, so the earliest observed reflected step need not vanish. A narrowband sinusoid and a broadband step perform different tests of the same lossless device.

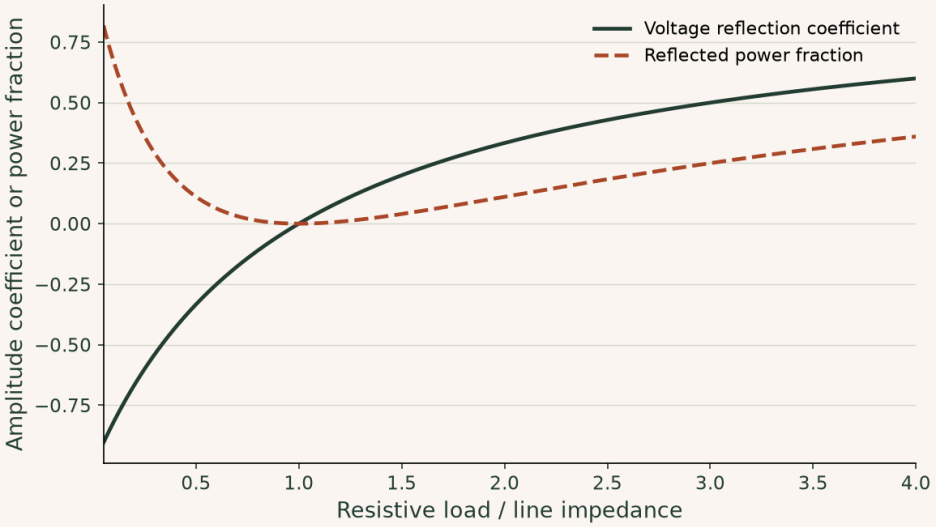
An actual connector of finite electrical length introduces another discontinuity that may be visible in either test.

Recover the load from a reflected pulse

Send a voltage step down a lossless line into a real load. Continuity gives

$$\Gamma = (Z_L - Z_0)/(Z_L + Z_0).$$

Use the original wave impedance for both incident and backward power, which makes reflected power scale as Γ^2.



Loads one half and twice the line impedance both reflect 1/9 of the incident power. Their reflected voltages have opposite signs. Measuring polarity as well as magnitude separates those two possibilities.

Try it · Reflection

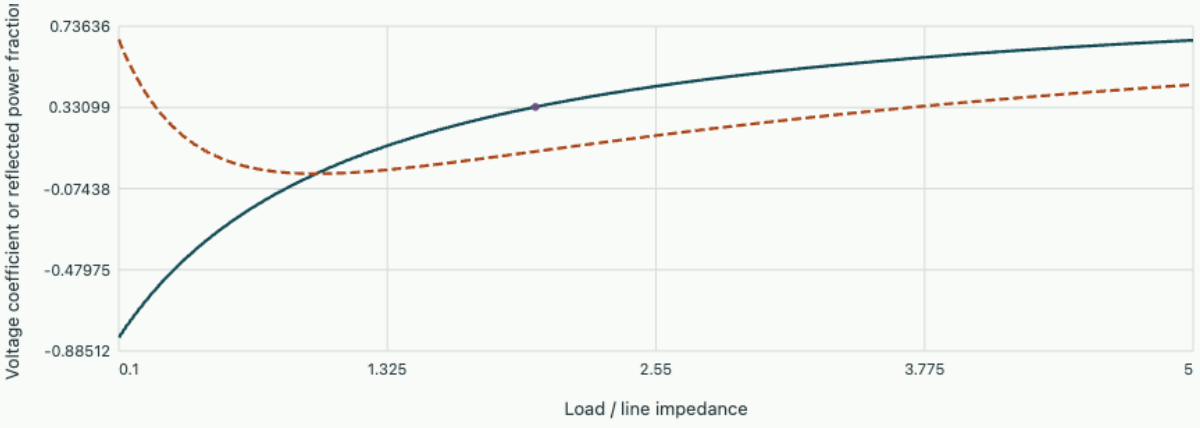
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Load / line impedance

2



Voltage reflection Γ Reflected power Γ² Selected voltage reflection

Voltage reflection 0.33333; reflected power fraction 0.11111. Loads r and 1/r give opposite voltage signs and equal reflected power.

▶ **Read values from the plot**

Develop this calculation further

Chapter 21 · What an interface does to phase, field and power · Chapter 29 · When a circuit element must be treated as a field

Study the full derivation

Oblique interfaces and films whose reflections cancel · Waves, radiation and relativity

A film only one quarter of a wavelength thick can suppress reflection from a transparent surface. Follow the boundary conditions all the way from an oblique incident wave to the input admittance of a stack, then calculate how much reflection returns when the wavelength, angle or layer order changes.

Chapter 12 · Force carried by a field

Compute electrical pressure from a local stress balance and reconcile two energy calculations for a moving capacitor plate.

FROM FIRST PRINCIPLES

Recover electrostatic pressure from virtual work

Place two large parallel, perfectly conducting plates of area A a distance d apart in vacuum. Ignore edge fields and disconnect the voltage source so that charge Q stays fixed. A pillbox at the metal surface gives

$$E = Q/(\epsilon_0 A).$$

To assemble the charge slowly, bring each increment dq across the voltage $V(q) = qd/(\epsilon_0 A)$ already established:

$$U(Q,d) = \int_0^Q V(q)\,dq = \frac{Q^2 d}{2\epsilon_0 A}.$$

Take a small displacement increasing d. With the leads disconnected there is no electrical work input. The work done by the field is minus the change in stored energy. Thus the signed force tending to increase the gap and the magnitude of its pressure are

$$F_d = -\left(\frac{\partial U}{\partial d}\right)_Q = -\frac{Q^2}{2\epsilon_0 A},$$

$$p = \frac{|F_d|}{A} = \frac{\epsilon_0 E^2}{2}.$$

The minus sign gives attraction. The factor one half also has a direct force interpretation. At one plate, only the field of the other sheet acts on its net charge. That sheet produces half the total gap field. If a voltage source remains connected, charge moves during the displacement and the source's work must enter the same energy account.

DERIVATION AND APPLICATIONS

Find the force drawing a dielectric into a capacitor

A partially inserted dielectric slab feels a lateral force even though the strong field between broad plates points across the gap. Use a plate width w, overlap length ℓ, gap d and a slab filling the gap for an inserted distance x. The slab has constant relative permittivity κ. Assume quasistatic motion, uniform plate separation and d much smaller than both occupied lengths and the width. The two regions share a voltage and their charges add, so the leading capacitance and its position derivative are

$$C(x) = \frac{\epsilon_0 w}{d}[\ell + (\kappa - 1)x],$$

$$C' = \frac{\epsilon_0 w}{d}(\kappa - 1).$$

First disconnect the supply, leaving total free charge Q fixed. A small inward displacement lowers the field energy. Energy conservation assigns that decrease to mechanical work, giving the inward electric force

$$F_Q = -\frac{d}{dx}\frac{Q^2}{2C(x)}\bigg|_Q = \frac{Q^2}{2C^2}C'.$$

With a battery attached instead, moving the slab draws additional charge $dQ = V\,dC$. The battery does work $V dQ$, while the field energy increases by half that amount. The difference supplies the mechanical work. This accounts for the same instantaneous force in either experiment when the initial voltage is the same.

$$F_V\,dx = V\,dQ - d\left(\frac{1}{2}CV^2\right) = \frac{1}{2}V^2\,dC,$$

$$F_V = \frac{1}{2}V^2 C'.$$

The microscopic lateral force arises where the field bends near the dielectric boundary. The capacitance method recovers the integrated force without needing that local field profile. It loses accuracy near the positions where the slab first enters or completely fills the capacitor.

WORKED EXAMPLE

Predict the pull at constant voltage

Change the inputs and investigate this derivation.

Let $w = 50.0$ mm, $d = 0.500$ mm, $\kappa = 3.00$ and $V = 100$ V. Then $C' = 1.77$ nF/m and the force is 8.85 µN. During 10.0 mm of insertion, capacitance increases by 17.7 pF. The battery supplies $V^2\Delta C = 177$ nJ, of which 88.5 nJ becomes mechanical work and the remainder increases the stored electric energy.

SECOND APPLICATION

Follow the force after disconnecting the supply

Change the inputs and investigate this derivation.

A different isolated apparatus starts at 80.0 pF and 60.0 V. Inserting the slab by 20.0 mm raises its capacitance to 100 pF, giving $C' = 1.00$ nF/m. The conserved charge is 4.80 nC and final voltage is 48.0 V. Initial and final forces are 1.80 and 1.15 µN. The work delivered is the drop in field energy, $(Q^2/2)(1/C_i - 1/C_f) = 28.8$ nJ. Its average force of 1.44 µN lies between the endpoint values.

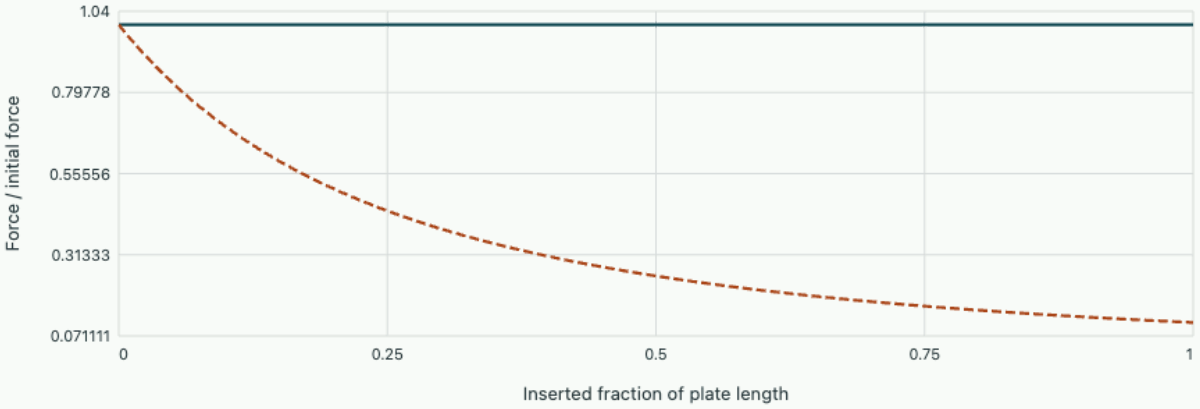
Try it · Dielectric force

Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Relative dielectric permittivity	3
Initial capacitor voltage (V)	500



For 20 mm plate width and 1 mm gap the initial force is 44.271 µN. Battery work must be included in the fixed-voltage balance.

Does separating capacitor plates weaken the force?

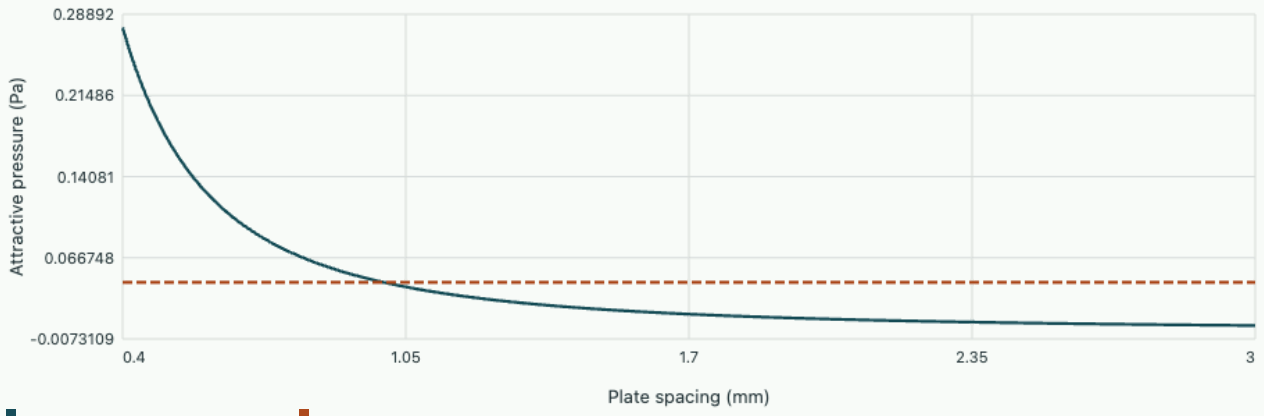
Set an air-filled parallel-plate capacitor to the reference voltage at a 1 mm gap. In one run leave the source attached and in the other disconnect it. Neglect edge fields. Both start with the same charge and pressure.

Write $E = V/d$ with a voltage supply and $E = \sigma/\epsilon_0$ for fixed free charge per unit area. The electrical pressure $p = \epsilon_0 E^2/2$ then gives different distance dependences. On the fixed-voltage path the source exchanges energy with the field as the capacitance changes.

Open this experiment online

Try the calculation

Reference voltage at 1 mm (V) **100**



Voltage maintained Charge trapped

At 2 mm the voltage supply gives 0.01107 Pa. With the original charge trapped, the pressure stays at 0.04427 Pa.

Find the distance at which the fixed-voltage pressure falls to one quarter of its initial value. Double the reference voltage and explain why the crossing distance of the two curves stays fixed.

Why the two halves of a charged sphere pull apart while its net force is zero

A thin spherical conducting shell of radius a carries total free charge Q in vacuum. Its two hemispheres are joined by a mechanically thin seam near the equator. Everything is initially at rest; any slit in the surface is negligibly narrow compared with a. The whole isolated shell has no preferred direction and experiences zero net electromagnetic force. A force gauge in the seam can nevertheless register a finite load. We will calculate the force on the upper hemisphere by two different choices of surface and then determine what changes if the charge fills the volume instead of sitting on the metal skin.

Start with the force on a charged surface element

Spherical symmetry gives $\sigma = Q/(4\pi a^2)$, $\mathbf{E}_{\text{out}} = \sigma\hat{\mathbf{r}}/\epsilon_0$ immediately outside and $\mathbf{E}_{\text{in}} = 0$. Multiplying a sheet's charge directly by its outside field would include the force the tiny element's own field would exert on itself. Near a smooth element approximate its local patch as planar. Its own fields on its two sides are opposite, $\pm\sigma\hat{\mathbf{r}}/(2\epsilon_0)$; the field of the rest of the assembly is continuous over that infinitesimal separation. Averaging fields on the two sides therefore extracts the field acting on the element. The force density per area is $+\sigma^2\hat{\mathbf{r}}/(2\epsilon_0)$, an outward pressure p.

Each patch acts along its own radial normal. Integrating only the vertical component $p\cos\theta$ over $0 \le \theta \le \pi/2$ gives

$$2\pi a^2 p \int_0^{\pi/2} \sin\theta\cos\theta\, d\theta = \pi a^2 p.$$

The result is pressure times projected circular area, not pressure times the whole curved hemisphere area. Thus

$$F_{\text{upper}} = \frac{Q^2}{32\pi\epsilon_0 a^2},$$

$$F_{\text{lower}} = -F_{\text{upper}}$$

where the signs refer to the z axis. Reversing the charge signs reverses the electric fields, but the mechanical load stays outward. The whole object's zero net force states only that these internal electromagnetic forces have an equal and opposite partner.

Move the integration surface away from the charges

Instead surround the entire upper half-space by an infinite hemisphere closed by the equatorial plane. In electrostatics the total force on enclosed matter is the outward flux of the Maxwell stress tensor. On the closure at infinity the stress falls as $1/r^4$, so area growth cannot leave a finite contribution. On the equatorial plane the outward surface normal for the upper region is $-\hat{\mathbf{z}}$. The electric field lies parallel to this plane outside the metal and vanishes inside. Thus $T_{zz} = -\epsilon_0 E^2/2$ for radial distance $s > a$, yielding upward traction on the upper half-space. Its integral is

$$F_z = \int_a^\infty \frac{\epsilon_0}{2}\left(\frac{Q}{4\pi\epsilon_0 s^2}\right)^2 2\pi s\, ds = \frac{Q^2}{32\pi\epsilon_0 a^2}.$$

This surface passes through an empty region with no electric charge and gives the same finite force. It resolves an apparent puzzle: the equatorial disk inside the conductor has zero field, but the entire cut plane includes an exterior annulus where the field transmits stress. A numerical integration on a small flat disk confined to the hollow interior would report zero because it omitted nearly all of the relevant control surface.

Estimate the load and the error of a finite stress integral

At $a = 30.0$ mm and $Q = 5.0$ nC, the pressure is 0.0110 Pa and the upward force is 31.2 µN. Consider checking this value in a field calculation restricted to equatorial radii $a < s < R$. Since the stress integrand is proportional to $s^{-3}ds$, that annulus captures fraction $1 - a^2/R^2$ of the analytic plane integral. Extending to $5a$ captures only 96%. A finite closed control surface would supply the missing contribution on its outer cap. An open truncated plane can instead be extrapolated using the measured $1/R^2$ tail. Both checks are independent of the surface-pressure argument.

A third consistency check uses uniform expansion of the whole conducting shell with fixed charge, for which $U = Q^2/(8\pi\epsilon_0 a)$. Then $-dU/da = Q^2/(8\pi\epsilon_0 a^2) = 4\pi a^2 p$, the outward pressure integrated over all area during a common radial displacement. It is larger than the hemispherical seam load by a factor of

four because these are different virtual displacements. Differentiating a known energy requires matching the change of geometry to the actual force measurement.

Translate the electromagnetic force into a seam measurement

For a narrow continuous join on the equator, let the tensile load transmitted across it be expressed as force per unit seam length. Rotational symmetry makes it uniform, so $t_{\text{seam}} = F_z/(2\pi a) = Q^2/(64\pi^2\epsilon_0 a^3)$. In the numerical example this is 1.66×10^{-4} N/m. Four equally loaded supports in an ideal symmetric mount would each carry about 7.8 µN if they carry the entire vertical separating load. This mechanical conversion uses the projection integral for the actual displacement; the radial pressure at a single point is not itself the measured force.

At fixed Q doubling a cuts the total seam force by four. Connect the shell instead to a voltage supply holding $V = Q/(4\pi\epsilon_0 a)$ fixed as a changes and the charges supplied by the source change proportionally with a. Substitution in the same static force formula now gives $F_z = \pi\epsilon_0 V^2/2$, independent of radius. The differing trends arise from comparing different prepared charges. A local field-stress calculation remains valid for either experimental protocol at its specified instantaneous state.

Repeat the force balance for a solid insulating sphere

Fill a sphere of the same radius with uniform fixed volume charge $\rho = 3Q/(4\pi a^3)$. Determine the force on all charge in the upper half and explain how the equatorial stress surface supplies the difference from the hollow metal case.

▼ **Derive the result and check it independently**

Inside the uniformly charged sphere $\mathbf{E} = \rho\mathbf{r}/(3\epsilon_0)$. Its vertical force density is $\rho^2 z/(3\epsilon_0)$. The geometric moment of the upper half is

$$\int z\,dV = \int_0^a z\pi(a^2 - z^2)\,dz = \pi a^4/4.$$

Integration gives $F_z = \pi\rho^2 a^4/(12\epsilon_0) = 3Q^2/(64\pi\epsilon_0 a^2)$, one and a half times the shell result. Self-interaction of the charge in the upper half contributes zero net force by pair symmetry, allowing the volume integral over the total smooth field.

The exterior annulus still yields exactly the original shell force, since the exterior field depends only on total charge. Now the equatorial disk inside the insulating sphere also contains a radial field $\rho s/(3\epsilon_0)$. Its stress adds $\int_0^a [\rho^2 s^2/(18\epsilon_0)]2\pi s\,ds = \pi\rho^2 a^4/(36\epsilon_0)$. Combining disk and exterior reproduces the direct volume integral. An exterior voltage scan alone cannot distinguish these two charge distributions, but a force measurement on the separating halves can.

Griffiths, 5th ed., §§2.5.3 and 8.2.2 · electrostatic pressure and Maxwell stress

Momentum at a boundary

A capacitor can pull its plates together across vacuum. The ordinary mechanical contact stress of a solid is absent in the gap, yet a surrounding surface still records a force. In vacuum, the electric field contributes the electrostatic stress tensor $T_{ij} = \epsilon_0(E_iE_j - \delta_{ij}E^2/2)$. Integrating the traction $\mathbf{T}\cdot\mathbf{n}$ around a body gives the electric force on enclosed material in a static configuration. An enclosing surface wholly in vacuum avoids having to specify internal material stresses.

To see the origin of this expression, use $\nabla\cdot\mathbf{E} = \rho/\epsilon_0$ and $\nabla\times\mathbf{E} = 0$. The derivative of E_iE_j contains a field-gradient term $E_j\partial_jE_i$. Electrostatic symmetry of derivatives makes this $\partial_i(E^2/2)$, which cancels the isotropic term in the divergence. The remaining term is ρE_i, the local electric force density. The divergence theorem converts the volume force integral into the stated surface integral.

Pressure on a conducting plate

For a planar metal surface the field immediately outside is normal. The magnitude of traction is then

$$p = \epsilon_0E^2/2.$$

Another derivation splits the surface-sheet field symmetrically between its two sides. The charge feels the average of the limiting fields, excluding the sheet’s own force. With zero interior field the result is $\sigma E/2$, identical since $\sigma = \epsilon_0E$. Using the full outside field as the field acting on every surface charge would double the force.

For plate area A and separation d, neglect fringing and put $C = \epsilon_0A/d$. If the plates are isolated with charge magnitude Q, stored energy $U = Q^2/(2C)$ increases linearly with separation. The signed force in the direction of increasing d is $-\partial U/\partial d = -Q^2/(2\epsilon_0A)$. The force tends to close the gap, consistent with the traction calculation.

Include the supply in a fixed-voltage experiment

When a voltage source remains connected, Q varies as d changes. The source supplies work $V\,dQ$. The relevant mechanical energy at fixed V is $U - VQ = -CV^2/2$. Differentiating this expression gives

$$F_d = (V^2/2)dC/dd = -\epsilon_0AV^2/(2d^2).$$

The electrical stress formula gives the same instantaneous force for either boundary condition. What changes during motion is the relationship between field and separation.

Take $A = 0.0100$ m², $d = 1.00$ mm and $V = 500$ V. The field magnitude is 5.00×10^5 V/m, the pressure is 1.11 Pa and attraction is 0.0111 N. Doubling the gap at fixed voltage quarters the force. Disconnecting first holds Q fixed and leaves force unchanged in the ideal wide-plate limit. A direct classroom test is to measure reaction force while controlling either voltage or charge. Departures as separation approaches plate width reveal fringing fields, which redistribute both stress and capacitance.

The planar calculation can be checked by dimensions before using any geometry. Electric-field energy density has units of joules per cubic metre, identical to newtons per square metre. It therefore has the units of pressure. This equality of dimensions alone cannot fix the direction of a force in a nonuniform field. Around the edge of a plate, for example, field lines bend and the tensor traction must be projected onto the direction being measured before integration.

Pressure from an incident and a reflected wave

For time-dependent fields the force on matter combines a flux of electromagnetic momentum through its boundary with the change of field momentum inside. A useful steady special case is a normally incident plane wave on a macroscopic stationary surface in vacuum. Its cycle-averaged energy flux is I, and transport of energy dU carries momentum dU/c in the propagation direction. Let the surface reflect fraction $\mathcal{R}$, transmit fraction $\mathcal{T}$ into vacuum on the far side, and absorb fraction $\mathcal{A} = 1 - \mathcal{R} - \mathcal{T}$. Assume all beams are collinear and neglect thermal reradiation in the mechanical measurement.

For an area A over time dt, the incoming axial momentum is $I A\, dt/c$. The reflected beam leaves with negative momentum and the transmitted beam with positive momentum. The difference delivered to the surface gives

$$p_{\mathrm{rad}} = \frac{I}{c}(1 + \mathcal{R} - \mathcal{T}) = \frac{I}{c}(\mathcal{A} + 2\mathcal{R}).$$

A fully absorbing sheet receives I/c even though there is no exiting optical wave. A perfect mirror receives $2I/c$ because it reverses photon momentum. A perfectly transmitting transparent sheet with no reflection takes no average axial momentum in this stationary, vacuum-to-vacuum accounting. Interfaces bounding different media require a joint treatment of material and electromagnetic momentum; the simple outgoing-vacuum relation has intentionally fixed that complication.

The same mirror result follows from the stress at a conducting boundary. At normal incidence the tangential electric field cancels while magnetic field amplitudes add. Using the cycle-averaged magnetic pressure at the boundary then gives $2I/c$. This provides an independent check on the momentum-flux interpretation without assigning a mechanical collision trajectory to a wave.

Predict a torsion-balance signal

Direct a continuous 2.00 W beam entirely onto a small mirror at normal incidence. Measured fractions are $\mathcal{R} = 0.960$, $\mathcal{T} = 0.0200$, leaving $\mathcal{A} = 0.0200$. Integrated axial force is

$$(2.00/c)(0.0200 + 1.920) = 1.29 \times 10^{-8}\,\mathrm{N}.$$

Mount the mirror so its force acts 40.0 mm from a torsion fibre axis with a perpendicular lever arm. Torque is 5.18×10^{-10} N m. A torsion constant 2.00×10^{-7} N m/rad gives a static twist of 0.00259 rad or 0.148°.

The absorbed power is 0.0400 W. Heating can drive air currents or slow distortions that also deflect the balance, particularly outside vacuum. Reversing the mechanical lever arm while leaving absorbed power

constant reverses the optical torque. Modulating the beam and measuring both delay and amplitude provides an additional way to distinguish the rapid radiation force from a slower thermal response.

For a plane mirror at oblique incidence the reflected momentum reverses only its component normal to the surface. A beam of known total power fully intercepted by the mirror imparts normal force $2P\cos\alpha/c$ when it is perfectly reflected at angle α to the normal. If the controlled quantity is irradiance per area perpendicular to the beam and the physical mirror area is fixed, intercepted power adds another projection factor. The pressure per unit mirror area is then $2I\cos^2\alpha/c$. Specifying whether power or surface footprint is held fixed is essential before comparing an angular torque scan with either formula.

Develop this calculation further

Chapter 20 · Where momentum waits between source and receiver · Chapter 25 · Changing observers and classifying electromagnetic fields · Chapter 26 · From local symmetry to Maxwell’s equations

Study the full derivation

Capacitance matrices, floating nodes and controlled approximation bounds · Electrostatic field theory and measurement

A multi-electrode device generally has no single capacitance until the other terminals have been connected or isolated. An energy identity determines the matrix that relates voltages to charges, and a complementary flux identity can certify estimates even when the exact field is unknown.

Problem sheet 6 · Reflections, force and stored energy

Determine a cable fault and use energy accounting to diagnose electrical forces with different experimental controls.

A reflected step

A line of characteristic impedance 75 Ω and propagation speed 2.00×10^8 m/s carries a 1.20 V step. A 0.360 V downward return step appears 480 ns after launch. Find the distance, reflection coefficient, load resistance and total first voltage at the load.

▼ **Show the fault reconstruction**

The wave made a round trip, so the mismatch is $vt/2 = 48.0$ m away. The signed ratio is

$$\Gamma = -0.360/1.20 = -0.300.$$

Invert the reflection law to get $R_L = 75(1+\Gamma)/(1-\Gamma) = 40.4\ \Omega$. Voltage at the load rises to $1.20(1-0.300) = 0.840$ V on first arrival. Its total current, $(1.20+0.360)/75 = 20.8$ mA, gives the same resistance.

Weigh the electric pull

Horizontal vacuum plates of area 40.0 cm² are separated by 0.500 mm. Find the voltage producing 2.00 mN of attraction. If disconnected and separated quasistatically to 1.00 mm, determine mechanical work supplied and final voltage.

▼ **Show the two controlled stages**

Pressure is $F/A = 0.500$ Pa and $E = \sqrt{2p/\epsilon_0} = 3.36 \times 10^5$ V/m. The supply must provide $V = Ed = 168$ V. After disconnection the field stays fixed for broad plates because charge and area are fixed. The opposing force remains 2.00 mN, so the applied mechanical work is $F\Delta d = 1.00$ µJ. The final voltage is 336 V. The increased electrical field volume accounts for the stored energy.

Check a field-energy inventory

A uniform 50 Ω lossless cable has propagation speed 2.00×10^8 m/s. A rectangular 2.0 V forward pulse is 100 ns long. Find its length, the total travelling energy and its separate electric and magnetic parts. Where does the energy go at an ideal open termination?

▼ **Show the travelling inventory**

The pulse occupies $v\Delta t = 20.0$ m. Its power is $V^2/Z_0 = 0.080$ W and energy is 8.0 nJ. The distributed parameters follow from Z_0 and v, giving $C' = 1/(vZ_0) = 100$ pF/m and $L' = Z_0/v = 250$ nH/m. Electric

energy $\frac{1}{2}C'\ell V^2$ is 4.0 nJ and magnetic energy is equal since $I = V/Z_0$. At an open circuit the voltage reflection coefficient is $+1$ and current cancels at the end. The reflected pulse ultimately carries back all 8.0 nJ; instantaneous electric and magnetic shares can change during overlap.

Account for the energy of a charging capacitor

Two circular plates of radius 5.00 cm, separated by 1.00 mm of vacuum, are charged by a constant current of 1.00 mA. At the instant their voltage is 10.0 V, find the displacement current, magnetic field at the plate rim, increase in electric energy per second and direction of electromagnetic energy flow. Neglect fringe fields in the gap.

▼ **Worked solution**

The field between the plates is V/d = 10^4 V/m. The capacitance is $\varepsilon_0\pi R^2/d$ = 69.5 pF. Since dV/dt = I/C, the electric flux changes at precisely the rate needed to give ε_0 dΦ/dt = I through the gap. Thus a circular Ampèrian path returns the same result whichever spanning surface one chooses.

Inside the disk only the fraction r^2/R^2 of the changing flux is enclosed. At the rim B = $\mu_0 I/(2\pi R)$ = 4.00 nT. The magnetic field is azimuthal. With the displacement current along the electric field, E × B points radially inward. The Poynting flux at the rim has magnitude EB/μ_0 = 31.8 W/m².

The side area is $2\pi Rd = 3.14 \times 10^{-4}$ m², giving 0.0100 W inward. The derivative of ½CV² is CV dV/dt = VI = 0.0100 W. Both calculations close the same energy balance. The uniform-field approximation describes the interior and integrated flux well for d much less than R. It cannot resolve the local fringe-field pattern right at the metal edges.

Explore these relations

- For the reflected step, vary the load ratio and compare voltage reflection with transmitted power. Open the live calculation.
- For the electric pull, vary the field and inspect the Maxwell pressure. Open the live calculation.

Can impedance identify a dielectric by itself?

A coaxial line with fixed radii and relative permittivity 2.25 measures 47.4537 Ω. A second sample measures 44.0381 Ω. Assuming concentric geometry and the same permeability, infer its relative permittivity. Use the laboratory’s offset-conductor run to construct a second possible explanation. Which additional measurement distinguishes these two explanations?

▼ **Work through the solution**

For a homogeneous concentric TEM line $Z_0 = \sqrt{L'/C'}$ and C' scales with permittivity while L' is fixed. Thus

$$\epsilon_{r,2} = 2.25(47.4537/44.0381)^2 = 2.613.$$

An unchanged dielectric with a 0.40 mm offset gives the same impedance in the numerical study. Wave speed distinguishes them. Geometry cancels in $L'C' = \mu\epsilon$ for a homogeneous TEM line. A dielectric change would reduce the speed by a factor $\sqrt{2.25/2.613} = 0.928$, while the offset alone leaves it unchanged.

Use the OneSimulation laboratory to check your answer

Compare the original concentric cable with a 0.40 mm offset at unchanged permittivity. Read both C and L to distinguish a dielectric change from a geometry change. Open the corresponding live laboratory reference.

Chapter 13 · Local field laws at points, sheets and holes

Derive divergence and curl from what a probe encloses. Track a point charge through a regularisation and work out when a potential can be defined over an entire region.

What would a very small instrument measure?

Imagine six square flux meters on the faces of an empty cube of side h. Each one measures the field component along its outward normal. A uniform field produces equal and opposite readings on opposite faces. A field whose x-component grows towards the right gives a positive difference between those two readings. To leading order their difference is $h^2[E_x(x+h/2)-E_x(x-h/2)]$. Adding the three pairs and dividing by volume gives

$$\lim_{h\to 0}\frac{1}{h^3}\oint_{\partial V_h}\mathbf{E}\cdot\mathbf{n}\,dS=\partial_x E_x+\partial_y E_y+\partial_z E_z.$$

The limit exists for a smooth field. Tiling a region with cubes then cancels all the interior-face readings. Only its outer surface remains. This explains why a local divergence controls a global flux even for an irregular boundary. In spherical symmetry a sphere of radius r makes the analogous measurement $4\pi r^2 E_r$. Taking the difference between two nearby concentric spheres and dividing by the shell volume gives

$$\nabla\cdot\mathbf{E}=r^{-2}d(r^2E_r)/dr.$$

The factor r^2 records the changing area of the meter.

A circulation meter is a small oriented loop. For a rectangle in the xy-plane, positive traversal about $+\hat{\mathbf{z}}$ compares E_y on its two vertical edges and E_x on its horizontal edges. Division by area gives

$$(\nabla\times\mathbf{E})_z=\partial_x E_y-\partial_y E_x.$$

Neighbouring loops cancel shared edges, giving the surface-to-boundary rule known as Stokes' theorem. Reversing the loop normal reverses the circulation. Curl measures local circulation density, even at a point where the vector itself happens to vanish.

Recover the charge lost by a pointwise derivative

The field of a point charge is $q\hat{\mathbf{r}}/(4\pi\epsilon_0 r^2)$. Its radial divergence formula gives zero at every $r>0$, whereas a sphere of any positive radius has flux q/ϵ_0. To make sense of both observations, replace the point by a smooth model potential of width $a>0$,

$$V_a(r)=\frac{q}{4\pi\epsilon_0\sqrt{r^2+a^2}},$$

$$E_a(r)=\frac{qr}{4\pi\epsilon_0(r^2+a^2)^{3/2}}.$$

Differentiate r^2E_a before dividing by r^2. The two derivative terms combine to leave a sharply concentrated charge density:

$$\rho_a(r) = \frac{3qa^2}{4\pi(r^2+a^2)^{5/2}},$$

$$Q_a(R) = q\frac{R^3}{(R^2+a^2)^{3/2}}.$$

For each fixed $R>0$, $Q_a(R)\to q$ as $a\to 0$. For each fixed $r>0$, $\rho_a(r)\to 0$. A probe with a smooth spatial response f measures $\int f\rho_a\, d^3r$, which tends to $qf(0)$. The object defined by this last operation is the three-dimensional delta distribution. Poisson's equation for Coulomb's potential therefore reads $\nabla^2(1/r) = -4\pi\delta^3(\mathbf{r})$. The singular part supplies exactly the missing flux.

Worked measurement with two collecting spheres

For $a = 2.0$ mm, a spherical meter at $R=a$ encloses $q/2^{3/2} = 0.354q$. At $R=5a$ it encloses $125q/26^{3/2} = 0.943q$. A mesh representing the smooth source needs cells considerably narrower than a to reproduce its peak. An integrated flux can already be close to the point-charge answer well outside that core. Reporting a peak and reporting an enclosed charge are distinct resolution tests.

Treat an interface as a source with a measured sign

An infinite uniform charge sheet at $z=0$ generates $E_z = \sigma\ \mathrm{sgn}(z)/(2\epsilon_0)$ when reflection symmetry makes both sides equivalent. Integrating its derivative over a thin interval gives the jump σ/ϵ_0 irrespective of the interval's thickness. Thus $\partial_z E_z = (\sigma/\epsilon_0)\delta(z)$. The surface density has units C m^{-2} and the one-dimensional delta contributes m^{-1}, as a volume density requires.

A finite-volume calculation can use that integral directly. For a cell of area A straddling the sheet the face-flux difference is $A(E_z^+ - E_z^-) = A\sigma/\epsilon_0$. As the cell becomes thinner its density estimate grows, while its charge stays fixed. If fields of $+600$ and -200 V/m are recorded just above and below a sheet, its density is

$$\epsilon_0(800\ \mathrm{V/m}) = 7.08\ \mathrm{nC/m^2}.$$

Their mean, 200 V/m, is an independently applied background compatible with the same jump.

Ask what lies inside the spanning surface

Consider $\mathbf{F} = C\hat{\boldsymbol{\phi}}/s$ outside a cylindrical excluded region, with s the distance from its axis. Its local curl vanishes. A circular traversal still has circulation $2\pi C$. A disk spanning that circle cuts the excluded region, where the assumed smooth field is unavailable. Every sufficiently small contractible loop has zero circulation. A single-valued global scalar potential on the entire punctured region would require zero circulation on the large loop as well. This is why the topology of the domain is part of a potential formulation.

Measure the flux that a finite sphere captures

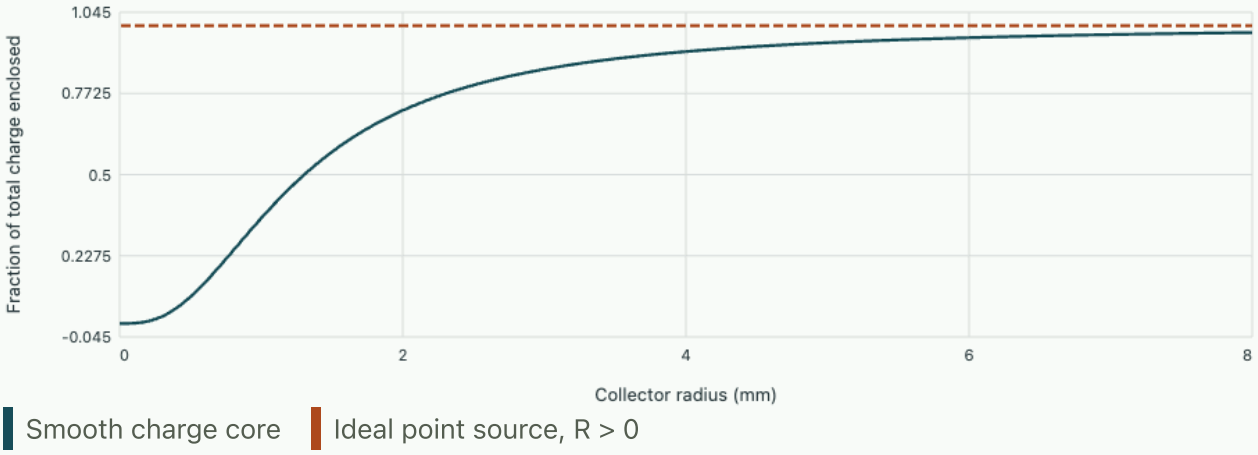
A point source would put all its charge inside every positive collector radius. For a 1 mm smooth core, a 2 mm collector encloses 71.6%. Increase the radius and infer which assumption an experimental flux measurement can resolve.

Open this experiment online

Try the calculation

Core radius a (mm) 1

Collector radius R (mm) 2



At 2 mm the collector encloses 71.6% of the charge. Doubling its radius would enclose 91.3%.

Test the instrument before using its reading

Recover a core width from one flux fraction

A calibrated spherical probe of radius R encloses $q/2$ for the smooth source above. Determine a/R and predict the fraction enclosed at radius $2R$.

▼ **Work through the solution**

Raise the measured fraction to the power $2/3$. It becomes $R^2/(R^2+a^2) = 2^{-2/3}$, so

$$a/R = \sqrt{2^{2/3} - 1} = 0.766.$$

The second meter collects $[1 + (0.766/2)^2]^{-3/2} = 0.814$ of the total, to three significant figures. Disagreement with both readings suggests an incorrect source shape or an off-centre probe.

Diagnose a circulation measurement

A loop encloses a very long changing flux tube. Outside it the magnetic field is negligible. Can the induced electric field have zero local curl at every point of the loop and a nonzero integral around it?

▼ **Work through the solution**

Yes. Outside the tube $-\partial_t \mathbf{B} = 0$, and a local curl probe records zero. A spanning disk reaches inside the tube. Its integrated flux derivative gives

$$\oint \mathbf{E} \cdot d\boldsymbol{\ell} = -d\Phi_B/dt.$$

Doubling the loop radius leaves this integral fixed for the same enclosed flux; cylindrical symmetry then halves the azimuthal field on the larger loop.

Further reading

- Griffiths, fifth edition, §§1.2–1.6 and 2.2 (publisher contents)
- Fließbach, seventh edition, Distributionen

Study the full derivation

Vector calculus through flux, circulation and singular sources · Electrostatic field theory and measurement

A closed-surface flux and a closed-path circulation answer different experimental questions. Starting from small physical collecting surfaces shows why the associated differential operators look different in Cartesian, cylindrical and spherical coordinates.

Chapter 14 · Building a potential from point responses

Solve one source problem exactly, reuse it for a distributed charge, and find what an electrode measures without solving for its surface charge point by point.

Start with the response to one thin sheet

Two large grounded conducting plates occupy $x = 0$ and $x = L$. A charge sheet of surface density σ sits at $x = s$, far from plate edges. Translational symmetry along each plate makes the potential a function of x alone. Integrating Poisson's equation across the charge shows that V is continuous and $V'(s^+) - V'(s^-) = -\sigma/\epsilon_0$. The potential is linear in each empty gap. It is helpful to remove the scale σ/ϵ_0 and call the remaining function $G(x, s)$.

On the left write $G = Ax$ to satisfy the first plate. On the right write $G = B(L - x)$ to satisfy the second. Continuity imposes $As = B(L - s)$. The derivative jump gives $-B - A = -1$, so $A = (L - s)/L$ and $B = s/L$. A single expression covers both segments:

$$G(x, s) = \frac{\min(x, s)[L - \max(x, s)]}{L},$$

$$-\partial_x^2 G(x, s) = \delta(x - s).$$

This Green function has units of length in this one-dimensional reduction. A source near the left plate produces a steep left slope and a shallow right slope. The left conductor acquires the greater magnitude of induced charge. Exchanging detector and source coordinates leaves G unchanged. This symmetry will turn out to be a physical consequence of the linear field equation.

> **Background · Reciprocity**
>
> $$\int \rho_1 V_2 \, dV = \int \rho_2 V_1 \, dV$$
>
> The displayed form assumes compatible homogeneous boundary conditions and a reciprocal linear medium. Boundary work terms matter for separately driven electrodes.

Assemble many sheets and retain the boundary voltages

A slice of thickness ds has sheet charge $\rho(s)\, ds$. Linearity allows its contribution $\rho(s)G(x, s)\, ds/\epsilon_0$ to be added to every other slice. If the plate voltages are V_0, V_L, a straight line accounts for the source-free solution satisfying those voltages:

$$V(x) = V_0(1 - x/L) + V_L x/L + \frac{1}{\epsilon_0}\int_0^L G(x, s)\rho(s)\, ds.$$

For uniform positive ρ, split the integral at $s = x$. The left piece is $(L - x)x^2/(2L)$, and the right piece is $x(L - x)^2/(2L)$, before multiplication by ρ/ϵ_0. Together they yield

$$V_{\text{charge}} = \rho x(L - x)/(2\epsilon_0).$$

Twice differentiating recovers the volume charge everywhere, and at either plate the charge contribution is zero.

Worked field measurement between unequal plates

Let $L = 10$ mm, $V_0 = 0$, $V_L = 2.0$ V, and $\rho = 1.00\,\mu\text{C/m}^3$. The straight-line background contributes 1.00 V at the midpoint. The uniform space charge contributes $\rho L^2/(8\epsilon_0) = 1.41$ V. The maximum potential can exceed both conductor voltages in the presence of a positive source. At the midpoint the charge contribution to the field vanishes by symmetry, so the field there still has its background value -200 V/m. This gives two independent observables for the same solution.

Prove that source and detector can exchange roles

For two electrostatic states occupying the same region with a linear, symmetric permittivity tensor, define potentials V_1, V_2 and charges ρ_1, ρ_2. Multiply the first Poisson equation by V_2, the second by V_1, then subtract and integrate. The two terms containing products of gradients cancel after integration by parts because of tensor symmetry. With outward normal from the dielectric region one obtains

$$\int_\Omega (V_2\rho_1 - V_1\rho_2)\, d^3x = \oint_{\partial\Omega} \left[V_1 \mathbf{n}\cdot\epsilon\nabla V_2 - V_2 \mathbf{n}\cdot\epsilon\nabla V_1\right] dA.$$

With both outer boundary potentials zero, the surface term vanishes. Choosing a point source for each state shows that the potential per unit charge at the other point is the same in either order. Fixed geometry, material response and boundary type matter. Replacing a grounded electrode by an isolated electrode between the two measurements changes the problem.

Worked electrode reading from a weighting potential

Take state 1 to be the real charge q inside a cavity whose electrodes are all grounded. For state 2 remove the charge, set the right electrode to 1 V and the others to zero, and divide this potential by 1 V to define w. The surface normal in the identity points into the metal, so $\mathbf{n}\cdot\epsilon\nabla V_1$ is the free charge density on the metal face. The identity now gives

$$Q_R = -q\, w(\mathbf{x}_q).$$

For parallel plates $w = x/L$. Moving 12 pC from $0.20L$ to $0.65L$ changes the right-electrode charge by -5.4 pC. Its change on the other plate is $+5.4$ pC, as an independent sum check. Device designers can reuse one computed w for many charge positions.

Move a charge sheet between grounded electrodes

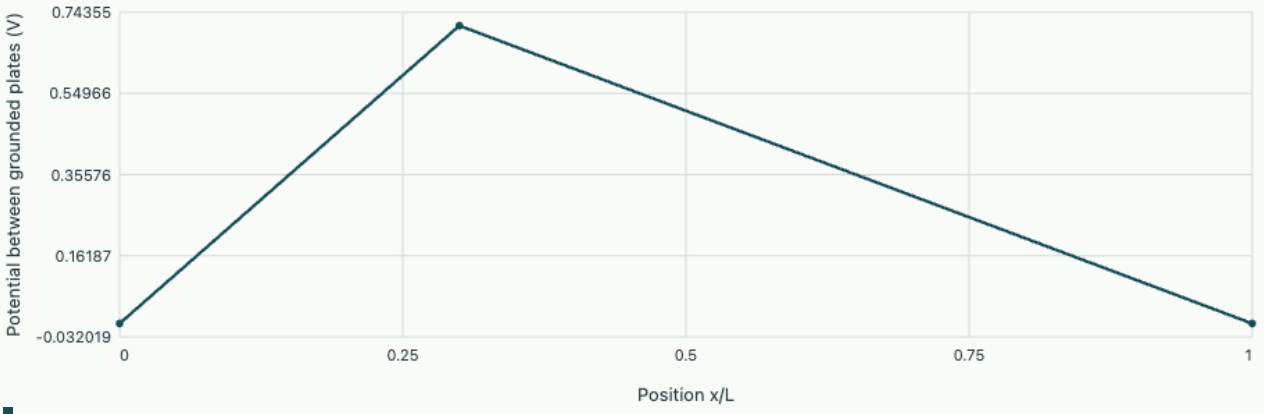
Keep the plate separation at 10 mm. At 30% of the gap, a positive sheet of 3 nC/m² induces −2.1 and −0.9 nC/m² on the left and right plates. Move the sheet through the midpoint, then reverse its charge. Predict the sign of the slope jump each time.

Open this experiment online

Try the calculation

Sheet position s/L **0.3**

Sheet charge density (nC/m²) **3**



Grounded 10 mm gap

Induced densities are -2.1 nC/m² on the left and -0.9 nC/m² on the right. Potential at the sheet is 0.712 V.

Practise splitting the integral and reading an induced signal

Locate a hidden sheet using both electrodes

A grounded pair of plates has induced charges of -3.0 pC on the left and -7.0 pC on the right from a single positive sheet of finite illuminated area in the large-plate approximation. Infer its charge and dimensionless position. Does the pair determine the width if several sheets may be present?

▼ Work through the solution

All induced charge sums to $-q$, so $q = 10.0$ pC. The right fraction 0.700 places the sheet at $s = 0.700L$. For a cloud, linearity gives

$$Q_R = -\int (x/L)\, dq.$$

The two readings yield total charge and first moment only. Two equal sheets at $0.60L$ and $0.80L$ give identical readings.

Bound a numerical influence coefficient

A discretisation gives

$$G(0.2L, 0.7L) = 0.062L,$$
$$G(0.7L, 0.2L) = 0.058L.$$

Which exact value should both approach? Does averaging establish convergence?

▼ **Work through the solution**

Their target is

$$0.2(1 - 0.7)L = 0.060L.$$

The discrepancy may come from depositing source charge and sampling voltage with different weights. The two errors happen to cancel in the mean here. Refine the grid and perform each measurement separately. Reciprocity alone also cannot reveal a common bias in both entries.

Further reading

- Fließbach, Randwertprobleme, pp. 59–67
- Griffiths, §§3.1–3.2, boundary conditions, uniqueness and images

Study the full derivation

Green functions, the Poisson kernel and what a small electrode can influence · Electrostatic field theory and measurement

A Green function is the response of an entire boundary-value problem to a unit source. Once that response is known, a continuous charge distribution and a patterned boundary voltage can both be treated by integration. The two integrations require different kernels.

Chapter 15 · How far does an electrode pattern reach?

Solve Laplace's equation in a strip, read decay lengths straight from boundary data and decide when a truncated series can predict an interior measurement.

Specify all the boundaries before searching for a formula

A long rectangular cross-section occupies $0 < x < a,\ y > 0$. The side walls are grounded, the bottom edge carries a prescribed voltage $f(x)$, and the potential decays as y grows. Physically the bottom edge may be an array of narrow driven electrodes separated by insulating gaps. The idealisation prescribes voltage in the limit of small gaps. There is no free charge in the interior and the dielectric is homogeneous. The problem is thus

$$V_{xx} + V_{yy} = 0.$$

A single field line or Gaussian surface does not tell us how a complicated f propagates upwards.

Try one product $X(x)Y(y)$. Dividing the equation by the product gives

$$X''/X = -Y''/Y.$$

Each side depends on a different variable, so both must be constant. Write the first constant as $-k^2$. Testing zero and positive values for X''/X yields only the trivial solution satisfying both grounded sides. The negative choice yields

$$X = A\sin kx + B\cos kx.$$

The left wall sets $B = 0$; the right wall requires

$$k_n = n\pi/a,$$
$$n = 1, 2, \ldots.$$

For a fixed n, Y is a sum of growing and decaying exponentials. The far boundary eliminates the growing branch.

Determine the coefficients by asking one question per mode

No individual mode can fit general data at the base. Add the admissible products and enforce the bottom voltage. Multiply by $\sin(m\pi x/a)$ and integrate over the width. Products of distinct sine modes average to zero, while the squared sine integrates to $a/2$. These measurements recover independent coefficients:

$$b_n = \frac{2}{a}\int_0^a f(x)\sin(n\pi x/a)\,dx,$$

$$V(x,y) = \sum_{n=1}^{\infty} b_n \sin(n\pi x/a)e^{-n\pi y/a}.$$

Physically every extra oscillation across the base shortens the vertical penetration length. For $a = 40$ mm the first mode decays by a factor e every 12.7 mm; the third takes only 4.24 mm. Complicated boundary structure can leave a very simple interior field.

Worked cancellation that moves with height

Drive the base with $f = 3\sin(\pi x/a) + 3\sin(3\pi x/a)$ volts. At the centre the two boundary voltages cancel since $\sin(3\pi/2) = -1$. Above it the voltage becomes $3(e^{-t} - e^{-3t})$ with $t = \pi y/a$. Differentiate to locate its maximum: $-e^{-t} + 3e^{-3t} = 0$ gives

$$t = \frac{1}{2}\ln 3.$$

Thus $y = 0.175a$ and the reading reaches $2/\sqrt{3} = 1.155$ V, even above a point set to zero volts. A full solution at the centre depends on the surrounding boundary data.

Predict the influence of a finite electrode

Place a 5 V electrode on the central half of the base and ground the remainder. Direct integration yields

$$b_n = \frac{10\ \mathrm{V}}{n\pi}\left[\cos(n\pi/4) - \cos(3n\pi/4)\right].$$

Even modes vanish by reflection symmetry. The first two nonzero coefficients are $b_1 = 4.50$ V and $b_3 = -1.50$ V. At $x = a/2, y = a/4$, their combined central reading is

$$4.50e^{-\pi/4} + 1.50e^{-3\pi/4} = 2.19\,\mathrm{V}.$$

A conservative bound on the discarded tail uses $|b_n| \le 20/(n\pi)$ V. Summing a geometric bound from $n = 5$ gives a maximum error of $20e^{-5\pi/4}/[5\pi(1 - e^{-\pi/4})] \approx 0.046$ V. Away from the base this is an inexpensive error certificate.

At the electrode edge itself the series approaches the midpoint of the voltage jump and overshoots nearby when truncated. A real electrode has a gap and finite edge radius. Estimating the maximum surface field from this discontinuous idealisation makes the result depend on the chosen resolution. An interior voltage at a specified positive height is much better conditioned.

Change a physical boundary and rebuild the mode

An enclosure of finite height H with a grounded lid needs a mode which vanishes at H and equals one at the base. Solving the same vertical ODE now gives

$$Y_n = \sinh[k_n(H - y)]/\sinh(k_n H).$$

The transverse eigenvalues stay fixed because the side walls have not moved. For $H = a/2$ and $y = H/2$, the fundamental is transmitted with factor $\sinh(\pi/4)/\sinh(\pi/2) = 0.377$ instead of the open-strip

value $e^{-\pi/4} = 0.456$. This difference is a way to diagnose an unmodelled nearby conductor.

Move a measurement plane away from patterned electrodes

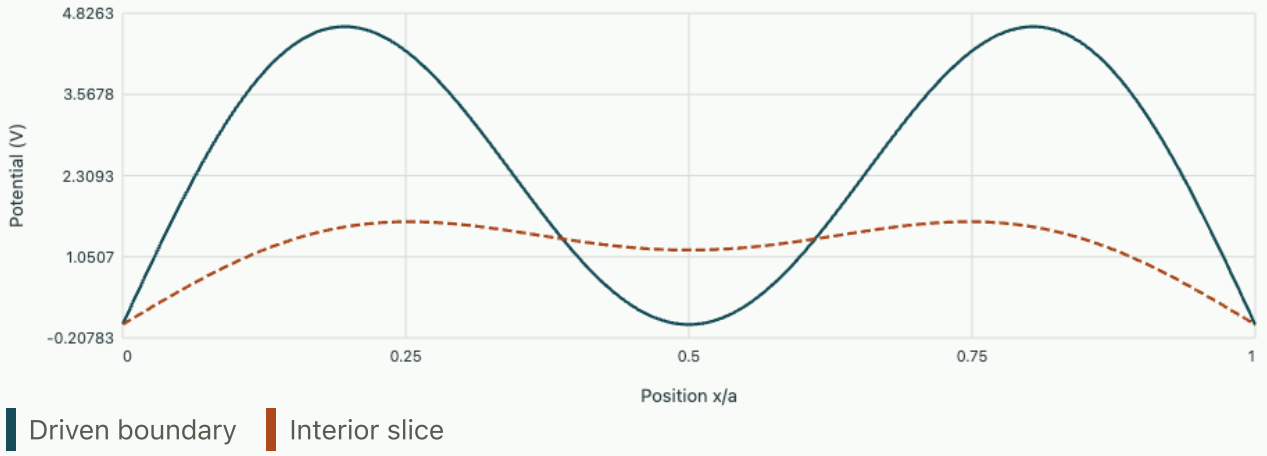
The first boundary harmonic is fixed at 3 V. Set the third to 3 V as well. At the midpoint the boundary values cancel; 0.18 strip widths above it the two terms leave about 1.15 V. Follow the two curves to see which boundary details survive propagation into the gap.

Open this experiment online

Try the calculation

Height y/a **0.18**

Third harmonic at the boundary (V) **3**



At the middle of the strip the potential is 1.15 V. The third harmonic retains 18.3% of its boundary amplitude.

Use a measurement to judge which boundary details survive

How accurately could the third harmonic be recovered?

At $y = a/2$ an instrument can resolve a 10 mV Fourier amplitude. How large must the third harmonic at the base be? Compare the fundamental.

▼ **Work through the solution**

Third-mode attenuation is $e^{-3\pi/2} = 0.00898$, so about 1.11 V is needed at the boundary. The fundamental needs only $0.010 e^{\pi/2} = 0.0481$ V. Back-propagating a noisy field magnifies its finest scales first. Improving the number of model harmonics cannot compensate for a missing signal.

Check the energy of one harmonic

The base drives $b\sin(kx)$ for one admissible k. Integrate the electric energy per unit length perpendicular to the strip and check its dimensions.

▼ **Work through the solution**

The two field components are $-kb\cos(kx)e^{-ky}$ and $kb\sin(kx)e^{-ky}$, so their squares sum to $k^2b^2e^{-2ky}$. Integrating $\epsilon E^2/2$ across a and upwards to infinity gives $U' = \epsilon akb^2/4$, in J/m. At fixed boundary amplitude a shorter wavelength stores more energy near the base despite penetrating less far.

Further reading

- Griffiths, §3.3.1, Cartesian separation of variables
- Fließbach, Randwertprobleme

Study the full derivation

Rectangular geometries, orthogonality and sources that drive individual modes · Electrostatic field theory and measurement

Fourier modes are solutions selected by their electrode walls. On a rectangle they turn a boundary-voltage problem into independent vertical equations. When a charge fills the rectangle the same eigenfunctions give a different denominator and a different physical response.

Chapter 16 · Spherical harmonics as measurements of boundary voltage

Reconstruct interior and exterior potentials from an angular voltage scan. Use orthogonality to isolate a dipole or quadrupole and quantify how rapidly fine angular structure disappears.

Why a sphere selects angular patterns

Consider a spherical surface of radius a covered by closely spaced insulated electrodes. They can prescribe a voltage varying with position, whereas a single connected conductor in electrostatic equilibrium can have only one voltage. Suppose there are no other charges, the surrounding permittivity is ϵ_0, and the reference is zero at infinity. At first let the applied voltage be independent of azimuth. In spherical coordinates the charge-free equation becomes

$$\frac{1}{r^2}\partial_r(r^2\partial_r V)+\frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\,\partial_\theta V)=0.$$

Writing $V=R(r)P(u)$ with $u=\cos\theta$ and multiplying by $r^2/(RP)$ separates the radial and angular parts. The angular operator becomes $d[(1-u^2)dP/du]/du$; regularity at both poles selects eigenvalues $\ell(\ell+1)$ for nonnegative integer ℓ. The regular eigenfunctions are the Legendre polynomials. The first three are $1, u, (3u^2-1)/2$. Substitution of any of these into the angular equation verifies its eigenvalue.

For each angular order the radial equation reads $(r^2R')'=\ell(\ell+1)R$. Insert $R=r^p$ to obtain

$$p(p+1)=\ell(\ell+1).$$

The two choices are $p=\ell$ and $p=-\ell-1$. The first stays regular at the centre; the second decays at infinity. These choices describe different physical domains for the same measured surface voltage.

Infer coefficients and the surface charge they require

On the spherical surface write $f(u)=\sum_\ell b_\ell P_\ell(u)$. Legendre polynomials are orthogonal under the integral du from -1 to 1; for equal indices the integral is $2/(2\ell+1)$. The weight has a geometric origin: a narrow latitude strip subtends solid angle $2\pi\sin\theta\,d\theta=-2\pi\,du$. Uniform steps in polar angle must be weighted by $\sin\theta$ before estimating the angular coefficients.

$$b_\ell=\frac{2\ell+1}{2}\int_{-1}^{1}f(u)P_\ell(u)\,du,$$

$$V_{\text{in}}=\sum_\ell b_\ell (r/a)^\ell P_\ell(u),$$

$$V_{\text{out}}=\sum_\ell b_\ell (a/r)^{\ell+1} P_\ell(u).$$

At the boundary both sides have the prescribed potential. Their normal derivatives generally differ. For each order the outward fields at a are $+(\ell+1)b_\ell P_\ell/a$ outside and $-\ell b_\ell P_\ell/a$ inside. Gauss's pillbox therefore assigns the thin electrode array a net sheet charge density

$$\sigma(u) = \frac{\epsilon_0}{a} \sum_{\ell} (2\ell + 1) b_\ell P_\ell(u).$$

Worked inverse problem with three voltmeter readings

Suppose independent measurements establish that orders above two are negligible. At the north pole, equator and south pole a scan gives 4, 1 and 2 V. The first and last readings differ only in the odd term, hence

$$b_1 = (4 - 2)/2 = 1\,\mathrm{V}.$$

Their average gives $b_0 + b_2 = 3$ V and the equator gives

$$b_0 - b_2/2 = 1\,\mathrm{V}.$$

Solving yields $b_0 = 5/3$ V and $b_2 = 4/3$ V. On the north axis at $r = a/2$ the predicted voltage is

$$5/3 + 1/2 + 1/3 = 2.50\,\mathrm{V}.$$

The three readings alone would not exclude a higher-order angular pattern. Additional sample positions are needed to test that assumption.

Use radial decay to design a measurement

The exterior radial factor contains one extra power of a/r compared with the relative angular order. At $r = 2a$ a quadrupole potential has fallen to one eighth of its surface value. An order-six pattern has fallen to $1/128$. A 20 mV order-six component at a distant detector would require 2.56 V at the surface. A 20 mV instrumental uncertainty has exactly the same amplification if one attempts a surface reconstruction from that single radius.

Worked cancellation and a charge measurement

For a pure quadrupole set $a = 50$ mm and $b_2 = 3.0$ V. The north-pole surface density is $5\epsilon_0 b_2/a = 2.66$ nC/m²; at the equator it is half that magnitude with opposite sign. The net charge vanishes because $\int_{-1}^{1} P_2\, du = 0$. A charge meter that captures the entire array would read zero while a local probe measures finite voltage and field. Net charge measures precisely the $\ell = 0$ coefficient.

Recognise when a latitude scan is insufficient

A painted voltage pattern can vary with longitude as well. A periodic azimuthal factor has form $e^{im\phi}$ with integer m. The extra angular term in Laplace’s equation introduces $-m^2/(1 - u^2)$ into the Legendre equation. Regular solutions are associated Legendre functions with $|m| \le \ell$. Combined with the azimuthal factor and normalisation they form $Y_{\ell m}$. For an orthonormal convention the coefficient is the surface

integral of $Y_{\ell m}^{*} f$ over solid angle. The radial powers remain unchanged. In particular, $f \propto \sin\theta\cos\phi$ is a dipolar pattern oriented along the x direction. Averaging every latitude ring would completely erase it.

Continue three boundary harmonics through a sphere

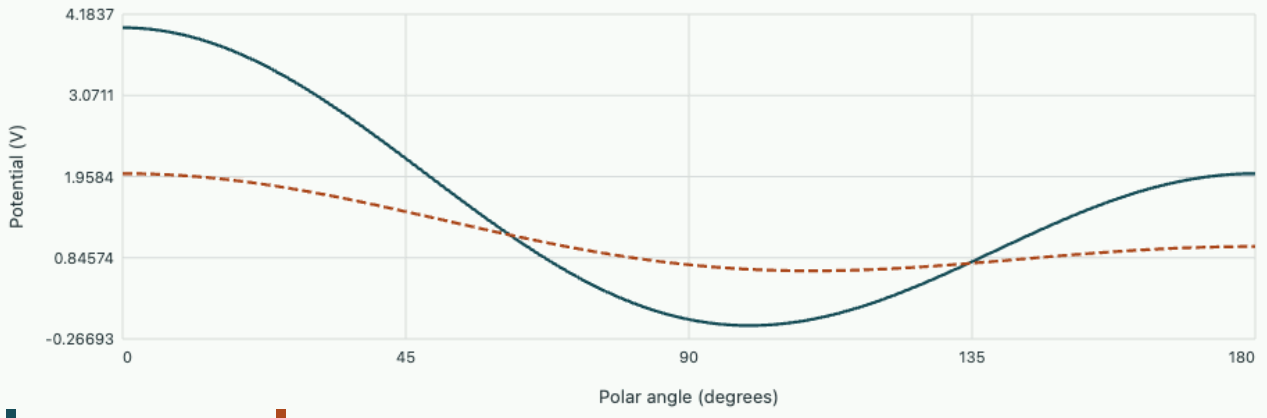
The boundary has a constant 1 V term, a 1 V dipole coefficient and an adjustable quadrupole. At half the sphere radius with a 2 V quadrupole, the pole reads 2.00 V and the equator 0.75 V. Compare inside and outside measurements with the same angular resolution.

Open this experiment online

Try the calculation

Observation radius r/a **0.5**

Quadrupole on the sphere (V) **2**



The north pole and equator read 2 V and 0.75 V at this radius. Both branches coincide on the boundary.

Distinguish angular resolution from total charge

Locate a field null inside a sphere

Take $b_1 = 1.2$ V, $b_2 = -1.0$ V, $a = 60$ mm, with all other coefficients zero. Find a point on the positive axis at which the full electric field vanishes.

▼ **Work through the solution**

On the axis, symmetry sets transverse field to zero. The potential is $1.2 z/a - (z/a)^2$ volts. Its derivative vanishes at $z/a = 0.6$, or 36 mm. The local axial extremum is compatible with Laplace's equation because transverse curvature has the opposite sign. The point is a saddle of the three-dimensional potential.

Is an interior maximum consistent with the measured boundary?

A code sampling a smooth, charge-free sphere reports 6.0 V at its centre; the largest voltage applied anywhere on the boundary is 4.0 V. Explain the discrepancy using the harmonics.

▼ **Work through the solution**

At the centre all terms with $\ell > 0$ vanish. The remaining b_0 equals the angular average of the boundary voltage and cannot exceed 4 V. The check isolates a cause to investigate: interior sources, incorrect boundary data or a defective discretisation. It does not require knowing the other coefficients.

Further reading

- Griffiths, §§3.3.2 and 3.4, spherical separation and multipoles
- Fließbach, Legendrepolynome, pp. 80–88
- Fließbach, Kugelfunktionen, pp. 99–107

Study the full derivation

Normalised spherical harmonics and the field of an arbitrary three-dimensional source · Electrostatic field theory and measurement

An axial voltage scan can determine Legendre coefficients only if the source is rotationally symmetric. A general spherical electrode or molecule needs angular modes with two indices. Their normalisation fixes the conversion from a measured multipole coefficient to a physical voltage.

Chapter 17 · From a measured current to its magnetic field

Construct the magnetic field of a steady source, calculate a coil pair's uniform region and extract the first magnetic moment that survives at a large distance.

Identify a physically possible source

A wire carries a measured steady current. Its magnetic effect depends on the complete current path, including the return conductor. Charge conservation requires $\nabla \cdot \mathbf{J} = 0$ for a stationary distribution. A finite wire segment with current suddenly ending at an insulated tip would accumulate charge. The missing return path can be neglected in an approximation only after comparing its distance and field with those of the measured segment.

The magnetostatic field equations in vacuum are

$$\nabla \cdot \mathbf{B} = 0,$$
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}.$$

The first suggests writing $\mathbf{B} = \nabla \times \mathbf{A}$. Adding any gradient to $\mathbf{A}$ leaves $\mathbf{B}$ unchanged. In unbounded space we may choose $\nabla \cdot \mathbf{A} = 0$ together with decay for a localised source. Applying curl to the potential gives

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A};$$

the chosen divergence removes the first term.

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{J},$$
$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d^3 r'.$$

Each Cartesian component solves a scalar Poisson problem. Taking the observation-coordinate curl under the integral differentiates the distance kernel. Since $\nabla(1/R) = -\mathbf{R}/R^3$, exchanging the order of the cross product yields the Biot–Savart integrand $\mathbf{J}(\mathbf{r}') \times \mathbf{R}/R^3$. Ampère's circulation law remains true for all steady currents. It gives a field value immediately only when symmetry constrains the field along a chosen path.

Turn the vector integral into an axial measurement

For a thin circular loop of radius a in the xy-plane, $\mathbf{J}\, d^3 r'$ becomes $I\, d\boldsymbol{\ell}'$. At a point z on the axis every element has the same distance $\sqrt{a^2 + z^2}$. Opposite elements produce cancelling transverse fields. The surviving axial numerator integrates to $2\pi I a^2$ because each tangent crosses the inward radial component of $\mathbf{R}$ with the same orientation. Thus

$$B_z(z) = \frac{\mu_0 I a^2}{2(a^2 + z^2)^{3/2}}.$$

For a 40 mm loop carrying 0.80 A, the central field is $12.57\,\mu\mathrm{T}$. At $z = a$ it is smaller by $2^{3/2}$, or $4.44\,\mu\mathrm{T}$. Reversing current reverses both readings. No field singularity exists on the axis; evaluating the ideal filament field directly on the wire instead requires a model of its finite cross-section.

Design a locally uniform field by cancelling curvature

Place two identical coaxial loops at $z = \pm d/2$ carrying the same current. Their combined field is even in z, so its first derivative at the midpoint is automatically zero. The second derivative sets the leading departure from uniformity. For a loop profile $b(z) = C(a^2 + z^2)^{-3/2}$, differentiation gives

$$b''(z) = 3C\frac{4z^2 - a^2}{(a^2 + z^2)^{7/2}},$$

$$B''_{\mathrm{pair}}(0) = 2b''(d/2).$$

Choosing $d = a$ cancels the curvature, leaving a fourth-order axial variation near the centre. The midpoint field then equals $\mu_0 I (4/5)^{3/2}/a$ for one turn on each loop. For $a = 60.0$ mm and $I = 0.50$ A this gives $7.49\,\mu\mathrm{T}$. Identical compact N-turn windings scale that reading by N while their finite winding dimensions eventually change the profile. A ratio of two nearby field readings is often a better test of spacing than an absolute reading affected by current calibration.

Find which source information a distant detector retains

For $r \gg a$, the exact axial loop field approaches $\mu_0 I a^2/(2r^3)$. Define the dipole moment $\mathbf{m} = I\pi a^2 \hat{\mathbf{n}}$. The axial far field is then $\mu_0 m/(2\pi r^3)$. Expanding the vector-potential kernel for a general localised steady source explains the same dependence: the term proportional to $\int \mathbf{J}\,dV$ vanishes for a closed current. Indeed, applying the divergence theorem to $x_i \mathbf{J}$ gives $\int J_i\,dV = 0$ when currents vanish at the outer boundary. The antisymmetric part of the next moment produces

$$\mathbf{m} = \tfrac{1}{2}\int \mathbf{r}' \times \mathbf{J}\, d^3 r',$$

$$\mathbf{A}_{\mathrm{dip}} = \frac{\mu_0}{4\pi}\frac{\mathbf{m} \times \hat{\mathbf{r}}}{r^2}.$$

For a planar loop the line-integral moment reduces to current times signed area. A detector far away therefore has difficulty distinguishing small loops with the same moment but different shapes. Closer readings, which retain the higher moments, are needed.

Find a pair of coils with a flat central field

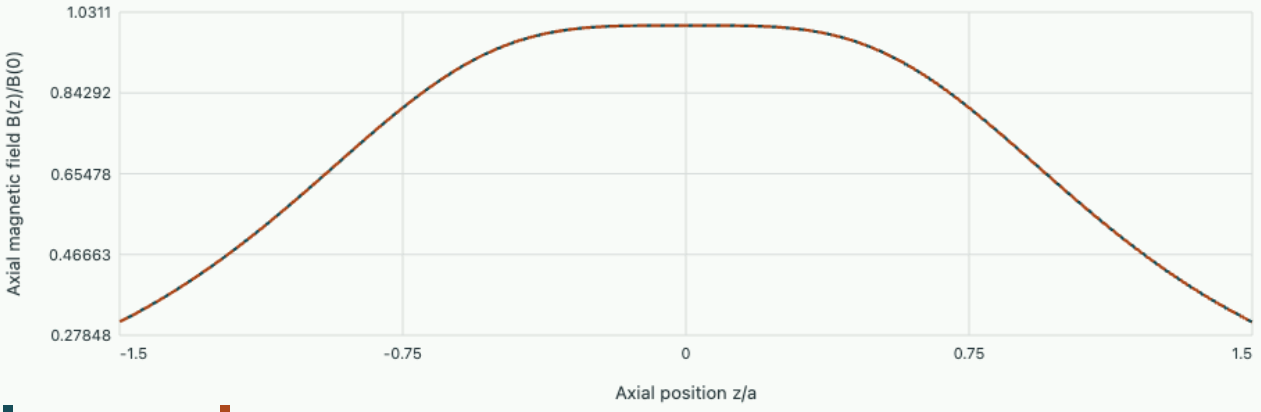
Change the separation of two identical current loops. Each curve is divided by its own central field so that the graph tests uniformity. At Helmholtz spacing, a probe one quarter-radius off centre reads about 0.42% below the centre.

Open this experiment online

Try the calculation

Coil separation / radius **1**

Axial test position / radius **0.25**



Selected pair | Helmholtz spacing d/a = 1

At z/a = 0.25 the selected field differs from the central field by -0.42%. The sign distinguishes a central peak from a central dip close to the origin.

Check which coil measurement identifies the source

Test the dipole approximation at a specified distance

At $z = 3a$, how large is the error if the exact loop field is replaced by its axial dipole formula? Express it relative to the exact result.

▼ Work through the solution

The approximation-to-exact ratio is

$$[1 + (a/z)^2]^{3/2} = (10/9)^{3/2} = 1.171.$$

The distant-source expression overpredicts by about 17.1%. At $z = 10a$ the excess falls to 1.50%. A distance of a few radii may be insufficient for precision work.

Choose a coil polarity for a gradient probe

The second coil of a symmetric pair is reversed. Predict the midpoint field and identify which derivative can now be nonzero there.

▼ Work through the solution

The two equal axial fields cancel at the midpoint. With opposite currents the sum is an odd function, allowing a linear gradient. For a pair $b(z - d/2) - b(z + d/2)$, its central derivative is $-2b'(d/2)$. With

$d > 0$, $b'(d/2) < 0$, so this wiring gives a positive axial gradient. Checking polarity distinguishes this arrangement from the uniform-field pair.

Further reading

- Griffiths, §§5.2–5.4, current sources and vector potential
- Fließbach, Magnetischer Dipol

Study the full derivation

The vector potential as an instrument for computing flux · Magnetism, induction and field dynamics

An experimenter can measure a flux change with a coil without measuring the magnetic field over every square millimetre of the coil area. The vector potential gives a corresponding computational tool. Its line integral determines flux, while the freedom to choose its gauge can turn a difficult boundary problem into a manageable one.

Chapter 18 · How microscopic dipoles produce macroscopic charge

Derive the volume and surface charges of a polarised body, determine the fields in layers and calculate which electrical work a capacitance measurement includes.

Replace a dense collection of dipoles by a smooth field

A small neutral molecule can have positive and negative charge centres separated by a displacement. Its dipole moment $\mathbf{p}$ points from negative to positive. A volume containing many molecules but small compared with the macroscopic variation defines a dipole moment per volume $\mathbf{P}$, in C/m². For observation points outside the body the sum of dipole potentials becomes

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int_\Omega \mathbf{P}(\mathbf{r}')\cdot\nabla'\frac{1}{|\mathbf{r}-\mathbf{r}'|}\,d^3r'.$$

The prime matters: differentiation with respect to source position has the opposite sign from differentiation with respect to observation position. Apply the product rule to $\nabla'\cdot(\mathbf{P}/R)$ and integrate the divergence over the body. The result has exactly the Coulomb form for a volume distribution $\rho_b = -\nabla\cdot\mathbf{P}$ and a surface distribution $\sigma_b = \mathbf{P}\cdot\mathbf{n}$. Together the two carry zero total charge for a bounded polarised neutral object. The volume and surface terms can each be nonzero.

Background · Electric displacement

$$\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$$

Its flux measures free charge. Polarisation is dipole moment per volume. A discontinuity of normal displacement depends on free surface charge.

Worked slab with a polarisation gradient

In $0 < z < d$, prescribe $P_z = P_0 z/d$ over a broad slab. The bound volume charge is uniform at $-P_0/d$. The lower face carries zero and the upper face $+P_0$. Per unit area the negative volume contribution is $-P_0$, cancelling the positive face. In an isolated large-plane approximation with no applied background, $D_z = 0$ outside and inside. Thus $E_z = -P_0 z/(\epsilon_0 d)$ in the slab. Voltage rises from bottom to top by $P_0 d/(2\epsilon_0)$. With $P_0 = 10$ nC/m² and $d = 2.0$ mm the rise is 1.13 V. Uniformly polarised material is the special case in which only the two faces carry charge.

Separate what an external circuit supplies

Gauss's law says $\epsilon_0\nabla\cdot\mathbf{E} = \rho_{\text{free}} + \rho_b$. Substituting the polarisation divergence and collecting terms defines $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$ with $\nabla\cdot\mathbf{D} = \rho_{\text{free}}$. A small pillbox then gives the jump in the normal component of $\mathbf{D}$ as the free surface density. A narrow electrostatic circulation loop gives continuity of tangential $\mathbf{E}$. These two statements specify different components. A spatially varying material can make $\nabla\times\mathbf{D} \neq 0$ even when $\nabla\times\mathbf{E} = 0$.

For a reversible linear isotropic dielectric, measurements can supply $\mathbf{P} = \epsilon_0 \chi_e \mathbf{E}$. Then $\epsilon = \epsilon_0(1 + \chi_e)$ converts between $\mathbf{D}$ and $\mathbf{E}$. Permanent polarisation and history-dependent dielectrics require additional constitutive information. The prescribed-gradient example deliberately specified $\mathbf{P}$ directly.

Turn boundary conditions into a measurable capacitance

Place two clean dielectric layers of thicknesses d_1, d_2 between broad electrodes. There is no free sheet at their interface. Both layers therefore have the same normal $D = Q/A$ away from edges. Each voltage drop is $D d_i/\epsilon_i$. Adding the drops gives

$$\frac{C}{A} = \frac{1}{d_1/\epsilon_1 + d_2/\epsilon_2},$$

$$\frac{V_1}{V_2} = \frac{d_1/\epsilon_1}{d_2/\epsilon_2}.$$

Take one millimetre of relative permittivity 2 and one millimetre of relative permittivity 8. With 10 V applied, the first layer takes 8 V and the second 2 V. Their fields are 8 and 2 kV/m. The free plate density is 1.42×10^{-7} C/m². A simple arithmetic mean of the permittivities would give a wrong voltage split. The layers share normal displacement in this orientation. For strips lying side by side across the electrode face, the voltage is common and their capacitances add in parallel.

Track the source that holds the voltage

At fixed geometry, adding a small plate charge dQ requires reversible work $V\,dQ$. With constant linear capacitance, integrating from zero to Q gives

$$U = Q^2/(2C).$$

In the layered capacitor this equals $(A/2)\sum_i d_i E_i D_i$. Changing a layer while Q stays fixed changes the mechanical work by the corresponding change in U. Holding voltage fixed connects a source that supplies or removes charge. If insertion at fixed V increases capacitance by dC, source work is $V\,dQ = V^2 dC$; the field stores only $V^2 dC/2$. The other half can appear as mechanical work by the electric force. Specifying the electrical control variable resolves an apparent sign ambiguity in capacitor force calculations.

Move a dielectric interface between voltage-controlled plates

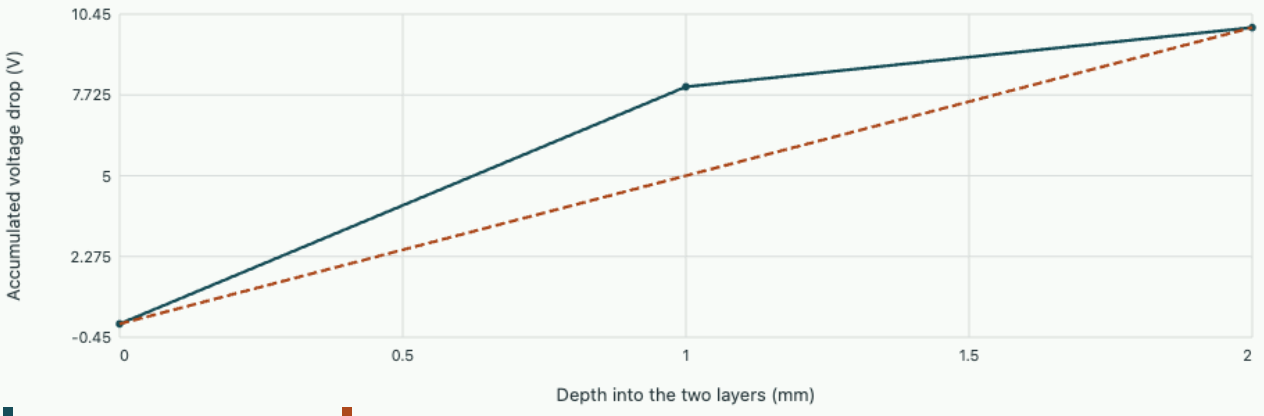
The pair has a total thickness of 2 mm and an applied voltage of 10 V. The first dielectric has relative permittivity 2. Equal layer thicknesses and a second permittivity of 8 give slopes of 8 and 2 V/mm. Which physical charge would be missed by forcing those slopes to agree?

Open this experiment online

Try the calculation

Second relative permittivity 8

First layer thickness (mm) 1



Layer 1 drops 8 V. The joint bound interface density is -53.1 nC/m²; the free interface density is zero.

Infer a hidden material property

Recover a permittivity from a voltage tap

Two layers of thickness 0.80 and 1.20 mm share 9.0 V. A tap shows a 3.0 V drop across the first layer of relative permittivity 4. Find the second relative permittivity.

▼ **Work through the solution**

Equal normal displacement gives

$$\epsilon_1 V_1/d_1 = \epsilon_2 V_2/d_2.$$

The second drop is 6.0 V, hence

$$\epsilon_{r2} = 4(3.0/0.80)(1.20/6.0) = 3.0.$$

A free interfacial sheet would supply a jump in D and invalidate this inversion.

Predict which bound charge a probe detects at the internal interface

In the equal-thickness, 10 V example choose the field direction positive from layer 1 to layer 2. Find the combined bound interface density and check against the field jump.

▼ **Work through the solution**

With $E_1 = 8000$, $E_2 = 2000$ V/m, $P_1 = \epsilon_0 E_1$ and $P_2 = 7\epsilon_0 E_2$. The top face of layer 1 and bottom face of layer 2 together carry $P_1 - P_2 = -6000\epsilon_0 = -53.1$ nC/m². This agrees with $\epsilon_0(E_2 - E_1)$ from Gauss's law for total charge.

Practicum · locate free and bound charge from a buried voltage tap

Two broad metal electrodes at $z = 0$ and $z = d_1 + d_2$ are held at 10.0 V and 0 V. Between them lie linear isotropic dielectrics, first $d_1 = 1.00$ mm with $\epsilon_{r1} = 2.00$, then $d_2 = 2.00$ mm with $\epsilon_{r2} = 6.00$. A very thin immobile sheet of free charge may be trapped at their interface. An ideal nonperturbing probe there reads 4.00 V. Neglect fringing. On paper, sketch a candidate potential and label the electric field, displacement, and source of each surface charge. Can all of your labels fit the three voltage readings simultaneously?

Reconstruct the source from field jumps

Determine both field vectors, both displacement vectors, the free charge sheet and the bound interface charge. State the direction associated with a positive sheet density.

▼ **Follow the calculation and check the result**

Set the interface normal from layer 1 to layer 2 along $+\hat{\mathbf{z}}$. The potentials in the two charge-free interiors must be affine because $\partial_z(\epsilon\partial_z V) = 0$ there. The slopes are -6.00 V/mm and -2.00 V/mm. Therefore $E_1 = 6000$ V/m, $E_2 = 2000$ V/m and both point along the normal. Their displacements agree: $D_1 = 2\epsilon_0(6000) = D_2 = 6\epsilon_0(2000) = 106.25$ nC/m². Gauss's law for free charge gives

$$\sigma_f = D_2 - D_1 = 0.$$

Compute the polarisation separately. $P_1 = \epsilon_0(6000)$ and $P_2 = 5\epsilon_0(2000)$. Facing surfaces contribute $+P_1$ and $-P_2$, so $\sigma_b = P_1 - P_2 = -35.42$ nC/m². A vacuum pillbox that counts all charge gives $\epsilon_0(E_2 - E_1) = -35.42$ nC/m², the same result by a different route. The signed density indicates charge per area, while the oriented normal fixes the sign in each jump equation.

Change one reading without changing either electrode

After a treatment, the tap reads 5.00 V with both plate voltages unchanged. A colleague assumes that equal voltage drops imply equal charge on facing electrodes. Test the inference, find the trapped charge, and close a global charge budget.

▼ **Follow the calculation and check the result**

Now $E_1 = 5000$ V/m and $E_2 = 2500$ V/m because the equal voltage drops occur across different thicknesses. So $D_1 = 88.54$ nC/m², $D_2 = 132.81$ nC/m² and the trapped free sheet is $+44.27$ nC/m². The conductor face charges are $+88.54$ nC/m² on the low-z plate and -132.81 nC/m² on the high-z plate.

Added to the sheet they sum to zero, consistent with zero exterior displacement for the ideal screened setup.

At the internal interface the bound charge is $P_1 - P_2 = \epsilon_0(5000 - 12500) = -66.41$ nC/m². Total charge there is -22.14 nC/m² even though the trapped free sheet is positive. Check it using $\epsilon_0(E_2 - E_1) = -22.14$ nC/m². An instrument sensitive only to the electric-field jump would give the wrong sign if its reading were interpreted as free charge.

Test what the measurements can actually resolve

Suppose the second tap reading has uncertainty ±0.05 V and the other inputs are known exactly. Propagate the uncertainty to the inferred free surface charge, then propose a null experiment.

▼ **Follow the calculation and check the result**

With V_m the tap reading,

$$\sigma_f = \epsilon_0[\epsilon_{r2}V_m/d_2 - \epsilon_{r1}(10 - V_m)/d_1].$$

Its slope is $\epsilon_0(3000 + 2000)$ per volt in SI units, namely 44.27 nC/m² per volt. The inferred sheet is 44.27 ± 2.21 nC/m². The uncertainty in the same tap is correlated between both drops, so treating the two inferred fields as independent would give a different uncertainty.

Remove the trapped charge and repeat with both electrodes at 0 V. The unique solution has zero field, and a residual tap voltage would indicate an offset or a violation of the material assumptions. A stronger differential test keeps the sheet and changes the applied voltage: the inferred jump in displacement should remain fixed while the two plate charges adjust. In a numerical boundary-value calculation, compare the reconstructed fluxes on either side of the interface and the integrated free charge, using the physical face permittivity.

Further reading

- Griffiths, §§4.2–4.4, bound charge, displacement and energy
- Tong, Electromagnetism in Matter, polarisation and displacement

Study the full derivation

Anisotropic and nonlinear dielectrics from local measurements to global fields · Electrostatic field theory and measurement

A single dielectric constant cannot describe a crystal whose displacement tilts away from its electric field, or a reversible material whose incremental response changes with bias. Starting with the constitutive measurement clarifies which parts of ordinary electrostatics survive.

Chapter 19 · How a magnet produces an exterior field

Translate microscopic loop moments into volume and surface currents. Derive a shape-dependent internal field and learn which magnetic response a coil measurement can recover.

Add the fields of tiny current loops

A small current loop produces a dipole vector potential $\mu_0 \mathbf{m} \times \hat{\mathbf{R}}/(4\pi R^2)$ far from that loop. Coarse-grain a material into a dipole moment per volume $\mathbf{M}$, in A/m. With $\nabla'(1/R) = \hat{\mathbf{R}}/R^2$, the potential becomes an integral of $\mathbf{M} \times \nabla'(1/R)$. Apply the curl product rule to $\nabla' \times (\mathbf{M}/R)$, then integrate the curl over the body. For outward normal $\mathbf{n}$ this gives

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi}\left[\int_\Omega \frac{\nabla' \times \mathbf{M}}{R}\, d^3r' + \oint_{\partial\Omega} \frac{\mathbf{M} \times \mathbf{n}}{R}\, dS'\right].$$

Comparison with the vector potential of conduction currents identifies $\mathbf{J}_b = \nabla \times \mathbf{M}$ and $\mathbf{K}_b = \mathbf{M} \times \mathbf{n}$. These currents are a representation of the material dipoles. Neighbouring equal loop currents cancel where they share a boundary. At the outer surface a current remains unpaired. The curl measures how cancellation fails within a nonuniform specimen.

Worked magnetised cylinder

Let a very long cylinder of radius a have uniform axial $M_0\hat{\mathbf{z}}$. Its bound volume current vanishes. Its side carries $K_\phi = M_0$, and its flat end faces carry none in this model. The remaining current acts like a long solenoid. The internal field far from the ends is $B_z = \mu_0 M_0$, while the exterior field tends to zero in the infinite-length limit. With $M_0 = 4.0 \times 10^4$ A/m this gives 50.3 mT. A finite specimen has a return field near and beyond its ends; the far field of a finite magnet must eventually be dipolar.

What a loop linking a winding actually counts

Subtracting $\mathbf{J}_b$ from magnetostatic Ampère's law gives $\nabla \times \mathbf{H} = \mathbf{J}_{\text{free}}$ with $\mathbf{H} = \mathbf{B}/\mu_0 - \mathbf{M}$. A rectangular loop across an interface yields $\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{K}_{\text{free}}$, while $\nabla \cdot \mathbf{B} = 0$ guarantees continuous normal B. In a region free of conduction currents and with suitable topology one can introduce a magnetic scalar potential $\mathbf{H} = -\nabla\phi_m$. Its divergence follows from $\nabla \cdot \mathbf{H} = -\nabla \cdot \mathbf{M}$. This alternative calculation uses the same magnetisation as the current description. Both predict identical $\mathbf{B}$.

Derive the demagnetising field of a sphere

Put a sphere of radius a with fixed uniform $\mathbf{M} = M\hat{\mathbf{z}}$ in vacuum and assume no applied field. Inside and outside, $\nabla \cdot \mathbf{M} = 0$. The angular structure at the surface is $M\cos\theta$, so use $\phi_i = Cr\cos\theta$ and $\phi_o = Ca^3\cos\theta/r^2$. They are harmonic, regular at the origin or at infinity, and continuous at a; continuity ensures equal tangential H. At the surface their outward fields are $H_{i,r} = -C\cos\theta$ and $H_{o,r} = 2C\cos\theta$. With normal B continuous, $2C = -C + M$, hence

$$\mathbf{H}_i = -\tfrac{1}{3}\mathbf{M},$$

$$\mathbf{B}_i = \tfrac{2}{3}\mu_0\mathbf{M}.$$

The factor $1/3$ encodes specimen shape. A very long axial cylinder with the same magnetisation has almost zero internal H far from its ends. The sphere's exterior scalar potential corresponds to a dipole moment equal to magnetisation times volume.

Worked susceptibility seen from outside

For a reversible linear sphere with $M = \chi H_i$ in a uniform applied field H_0, superposition adds the applied field and the demagnetising contribution: $H_i = H_0 - M/3$. Solving yields

$$M/H_0 = \chi/(1 + \chi/3).$$

If a moment measurement gives $M/H_0 = 1.50$, the intrinsic susceptibility is 3.0. As χ grows, M/H_0 approaches 3, and the inversion becomes very sensitive to moment uncertainty. Using $B = \mu H$ without specifying where the field is measured can conceal this shape dependence.

Separate reversible energy from cycle loss

When ferromagnetic domains rearrange, B can depend on the prior excursion of H. A single constant susceptibility then cannot describe an entire cycle. In a slow, approximately uniform closed magnetic circuit the source supplies incremental energy density $H\,dB$. Returning to the initial magnetic state leaves the signed cycle integral $\oint H\,dB$; its dissipated energy per volume is the positive area for the driven loop orientation. A measured loop area of 180 J/m³ at 50 cycles/s would dissipate 9.0 kW/m³, or 0.090 W in 10 cm³. Eddy-current loss can also enlarge a measured dynamic loop, so a rate sweep helps interpret that result.

Infer material response from an external field

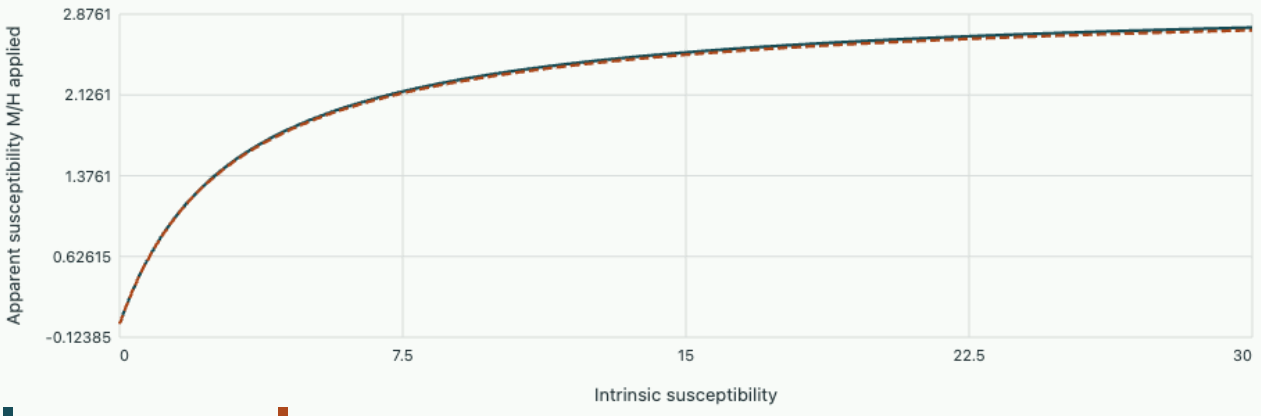
The sphere reference curve gives an apparent susceptibility of 2.25 at intrinsic susceptibility 9. The adjustable ellipsoid starts at N = 0.33, giving 2.27. Try smaller demagnetising factors for elongated ellipsoids magnetised along their long axes. Determine when an apparent response would severely underestimate the intrinsic one.

Open this experiment online

Try the calculation

Demagnetising factor N **0.33**

Intrinsic susceptibility **9**



Selected ellipsoid Sphere N = 1/3

At the selected susceptibility the apparent value is 2.27; the internal H field is 25.2% of the applied H field.

Test a surface-current picture

What current remains if magnetisation falls towards the rim?

Take a long cylinder with $M_z = M_0(1 - s^2/a^2)$ inside. Find the azimuthal bound currents and the field on its axis.

▼ **Work through the solution**

Cylindrical curl gives

$$J_{b,\phi} = -dM_z/ds = 2M_0 s/a^2.$$

At $s = a$ the surface current is zero. Regard a shell ds as a small solenoidal sheet $K = J_{b,\phi} ds$. At the axis all shells contribute, giving $B_z(0) = \mu_0 \int_0^a J_{b,\phi}\, ds = \mu_0 M_0$. Far from the ends the radial profile inside is $B_z(s) = \mu_0 M_z(s)$.

Invert a nearly saturated shape response

If $M/H_0 = 2.70$ for a linear sphere, what is χ? Estimate how an absolute uncertainty of 0.03 in the measured ratio propagates.

▼ **Work through the solution**

Writing $r = M/H_0$, one finds

$$\chi = r/(1 - r/3) = 27.$$

Its derivative is $(1 - r/3)^{-2} = 100$, so the local uncertainty is about 3. Distinguishing two high susceptibilities requires precise specimen geometry and much better relative accuracy than the measured moment alone suggests.

Further reading

- Griffiths, §§6.1–6.4, magnetisation, bound currents and H
- Tong, Electromagnetism in Matter, magnetic fields in matter

Study the full derivation

From microscopic moments to measured magnetisation · Magnetism, induction and field dynamics

A magnetised sample can make a compass turn without a power supply attached. The macroscopic theory assigns it bound currents in its bulk and on its surface. These are calculated sources for the magnetic field. We will reconstruct them from magnetic moments, predict signals from finite samples, and ask what an exterior field measurement can identify.

Chapter 20 · Where momentum waits between source and receiver

Derive the electromagnetic momentum balance, trace light pressure through a finite pulse, and account for the angular impulse when a magnetic flux decays.

Give a region a mechanical and a field account

Charges respond to Lorentz force density $\mathbf{f} = \rho\mathbf{E} + \mathbf{J} \times \mathbf{B}$. To predict the total force on enclosed matter, one can measure the fields on a surrounding vacuum surface. First eliminate the sources. In vacuum Gauss's law gives $\rho = \epsilon_0 \nabla \cdot \mathbf{E}$, and the Ampère–Maxwell law gives

$$\mathbf{J} = (\nabla \times \mathbf{B})/\mu_0 - \epsilon_0 \partial_t \mathbf{E}.$$

Introduce $\mathbf{g} = \epsilon_0 \mathbf{E} \times \mathbf{B} = \mathbf{S}/c^2$. Its time derivative supplies both $\epsilon_0 \partial_t \mathbf{E} \times \mathbf{B}$ and $\epsilon_0 \mathbf{E} \times \partial_t \mathbf{B}$. The first cancels the displacement-current term in the force; Faraday's law rewrites the second.

For example, the electric terms now read $\epsilon_0[\mathbf{E}(\nabla \cdot \mathbf{E}) - \mathbf{E} \times (\nabla \times \mathbf{E})]$. Component i is $\partial_j[\epsilon_0(E_i E_j - \delta_{ij} E^2/2)]$, with repeated Cartesian indices summed. The magnetic terms have the same form because $\nabla \cdot \mathbf{B} = 0$. Integration gives

$$T_{ij} = \epsilon_0(E_i E_j - \tfrac{1}{2} E^2 \delta_{ij}) + \frac{1}{\mu_0}(B_i B_j - \tfrac{1}{2} B^2 \delta_{ij}),$$

$$\mathbf{F}_{\text{matter}} = \oint T\mathbf{n}\, dA - \frac{d}{dt}\int \mathbf{g}\, dV.$$

The normal points out of the chosen volume. Positive T_{xx} denotes tension along x in this convention. For a wave travelling in $+x$ with transverse fields, $T_{xx} = -u$, where u is energy density. On an inlet whose outward normal is $-\hat{\mathbf{x}}$, the normal traction therefore points in $+x$.

Measure the time before light reaches a detector

A finite plane pulse in vacuum has $|\mathbf{S}| = cu$. Its total momentum is U/c along the propagation direction. Surround a freely absorbing target and a length of initially empty space upstream with a vacuum control surface. While a pulse enters but has yet to reach the target, the inlet stress integral is nonzero. The electromagnetic momentum inside grows at exactly the same rate. The minus time derivative in the total force cancels the inlet contribution and the target has zero force during this interval.

Worked optical impulse

A 0.60 mJ pulse is fully absorbed after crossing the inlet. Initially and finally the region contains no pulse. Integrating the balance over that full time removes the stored-field term. The target gains $0.60 \times 10^{-3}/c = 2.00 \times 10^{-12}$ N s. For ideal normal reflection the exiting pulse carries the opposite momentum, giving twice that impulse. For a transparent window with parallel faces at normal incidence the outgoing pulse may eventually have the same vacuum momentum as the incoming pulse, while the two faces still experience transient forces.

Convert distributed field momentum into a turning impulse

Now take a long coaxial arrangement: an inner narrow cylindrical shell of radius a carries free charge $+\lambda$ per length; an outer shell of radius b carries $-\lambda$. There is a uniform axial magnetic field B_0 over the annular gap, with end effects negligible. For a definite preparation take an ideal long solenoid producing uniform field throughout $r < b$, and confine the initial electric field to the gap. Between shells $E_r = \lambda/(2\pi\epsilon_0 r)$. The field momentum density points towards negative azimuth because $\hat{\mathbf{r}} \times \hat{\mathbf{z}} = -\hat{\boldsymbol{\phi}}$.

$$g_\phi = -\lambda B_0/(2\pi r),$$

$$L'_{\mathrm{field},z} = \int_a^b r g_\phi \, 2\pi r \, dr = -\tfrac{1}{2}\lambda B_0 (b^2 - a^2).$$

The prime denotes angular momentum per length. While the coil current decays slowly, Faraday’s law gives $2\pi r E_\phi = -\pi r^2 \dot{B}$ for radii up to b. The torque on each shell per length is $r\lambda_{\mathrm{shell}} E_\phi(r)$. Integrating from B_0 to zero yields $+\lambda B_0 a^2/2$ on the inner shell and $-\lambda B_0 b^2/2$ on the outer. Their sum reproduces the initial field angular momentum. Clamps would receive those impulses if the shells were held fixed.

Worked sign and magnitude

With $\lambda = 20$ nC/m, $B_0 = 0.25$ T, $a = 10$ mm and $b = 30$ mm, the initial $L'_z = -2.0 \times 10^{-12}$ kg m/s. That is angular momentum per metre of apparatus. The outer shell receives nine times the inner impulse magnitude and rotates in the opposite sense. Torque from the induced azimuthal electric field supplies this transfer even though both shells started at rest.

Transfer angular momentum from a field into two shells

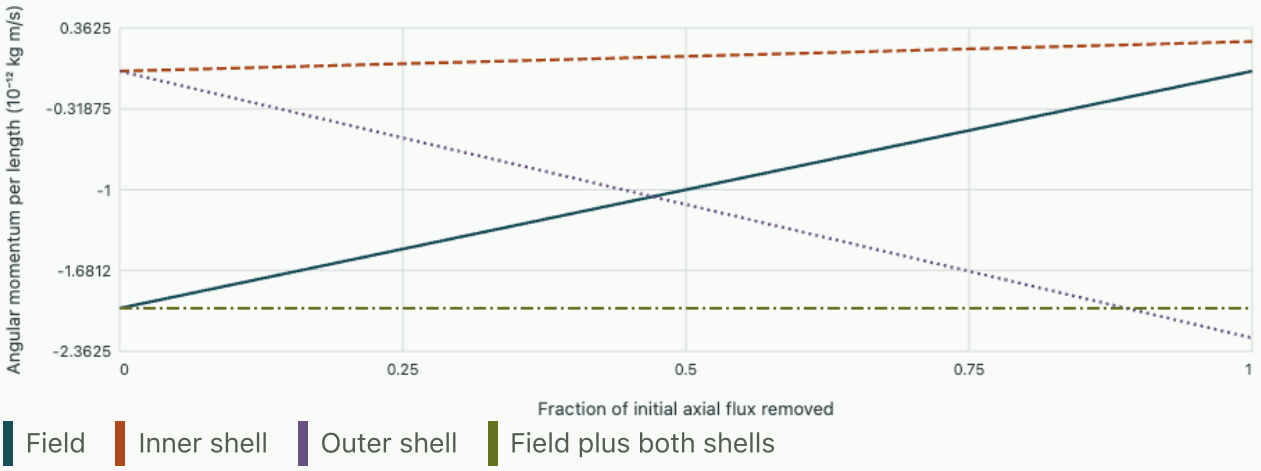
The inner radius is 10 mm and its line charge is +20 nC/m. For a 30 mm outer radius and an initial field of 0.25 T, the final shell impulses are +0.25 and −2.25 in units of 10⁻¹² kg m/s. Trace the four accounts during switch-off.

Open this experiment online

Try the calculation

Outer radius b (mm) **30**

Initial field B (T) **0.25**



After switch-off the inner shell carries 0.25 and the outer shell -2.25 in plotted units. Their sum equals the initial field value -2.

Choose a control surface for momentum

Find a pressure when transmission and reflection coexist

A normal beam of intensity 300 W/m² reaches a free-standing film in vacuum. Reflected power fraction is 0.20 and transmitted fraction is 0.50. Find the time-averaged pressure and identify the momentum carried away.

▼ **Work through the solution**

With incident direction positive, outgoing momentum per area and time is $(T - R)I/c$. The deposited amount is

$$(1 - T + R)I/c = 0.70(300)/c = 0.700\,\mu\text{Pa}.$$

The absorbed part contributes $0.30I/c$ and reflection adds $0.40I/c$. The result assumes forward and backward beams remain normal to the film.

Reverse just the charges

In the coaxial example reverse both shell charges while keeping the axial field and current decay unchanged. Which quantities reverse?

▼ **Work through the solution**

The radial field, stored azimuthal momentum and angular momentum reverse. Faraday's induced E_ϕ follows the same history because the flux decay is unchanged. Each electric force and angular impulse reverses since the charge signs reverse. This independent sign argument checks the cross product in the original momentum density.

Practicum · audit a pulse bouncing inside a partly absorbing cavity

A shutter has just prepared a short $U_0 = 1.20$ mJ pulse travelling to the right in an otherwise empty one-dimensional cavity. Start the accounting after preparation. The left mirror reflects perfectly. At every encounter with the right mirror 60% of incident pulse energy is reflected, 30% passes into vacuum on the right, and 10% is absorbed. Pulse duration is short compared with the flight time across the cavity. Mirrors are mounted on massive force probes; neglect their displacement and delay any re-emission of heat. The challenge is to reconcile two large, oppositely signed readings with a small net impulse.

Audit one encounter and decide which surface terms matter

For energy U just before an encounter with the right mirror, calculate its mechanical impulse. Then find the impulse when the reflected packet hits the left mirror. Which momentum is still present between those events?

▼ Follow the calculation and check the result

The incoming momentum is $+U/c$. The two outgoing packets carry $+0.30U/c$ and $-0.60U/c$. Momentum balance therefore assigns the right mirror $+(1 - 0.30 + 0.60)U/c = +1.30U/c$. This includes reversing the reflected packet and stopping the absorbed packet. The left mirror receives $-2(0.60U)/c = -1.20U/c$ at the next collision. Between those events the cavity still contains $-0.60U/c$ of field momentum. The complete system also contains the $+0.30U/c$ transmitted packet outside the cavity.

Stop the movie at a specified time

Freeze the experiment just after the third right-mirror encounter and before the next left-mirror encounter. Find the remaining energy, radiated output, deposited heat, both probe impulses, and total field momentum.

▼ Follow the calculation and check the result

Three incident energies in units of U_0 form the sequence $1, 0.6, 0.36$. Their sum is 1.96. The output energy is $0.30(1.96)U_0 = 0.588U_0$, heat is $0.196U_0$, and $0.216U_0$ remains. The three fractions sum to unity. Right-probe impulse is

$$+1.30(1.96)U_0/c = +2.548U_0/c.$$

Only two left reflections have occurred, so the left probe reads $-2(0.60+0.36)U_0/c=-1.920U_0/c$. Their sum is $+0.628U_0/c$.

The escaped packets carry $+0.588U_0/c$, but the packet still inside carries $-0.216U_0/c$. Including these two contributions gives measured plus field momentum $(0.628+0.588-0.216)U_0/c=U_0/c$, the initial value. Counting a third left-hand reflection at this time would violate the preselected event boundary. Since $U_0/c=4.00\times10^{-12}$ N s, the probe readings are approximately $+10.2$ and -7.68 pN s.

Sum the full history in two independent ways

Continue until no macroscopic pulse remains. Find the impulse on each mirror, then derive their combined reading directly from exterior measurements without examining any reflection.

▼ **Follow the calculation and check the result**

The geometric series of right hits sums to $U_0/(1-0.6)=2.5U_0$. The right impulse is $3.25U_0/c$. Summing every left hit gives

$$-2(0.60)U_0/[(1-0.60)c]=-3.00U_0/c.$$

Thus the pair receives $+0.25U_0/c$. All transmitted packets travel right and together carry $0.30U_0/(1-0.60)=0.75U_0$. An external vacuum control surface therefore gives $(1-0.75)U_0/c$ immediately. The absorbed $0.25U_0$ accounts for the energy stored as heat.

A detector without enough time resolution can recover the integrated exterior momentum even though it cannot resolve the alternating internal forces. Conversely, two large internal probe signals can coexist with a modest net exterior impulse. In a numerical time march record separate signed transfers and the energy still in flight after each boundary update, then vary the stopping time. A budget that closes only after the very last encounter cannot diagnose an intermediate sign mistake.

Further reading

- Griffiths, §§8.2.2–8.2.4, stress, momentum and angular momentum
- Tong, Electromagnetism and Relativity, energy-momentum and angular momentum

Study the full derivation

Energy flow and mechanical force from vacuum fields · Magnetism, induction and field dynamics

An external support holds a charged capacitor plate or a current-carrying coil in place. One can infer the support force from fields just outside the body, without summing forces on every microscopic charge. A second surface integral tracks the energy feeding the same apparatus. Both methods start with Maxwell's equations, but they transport different conserved quantities.

Chapter 21 · What an interface does to phase, field and power

Match fields at an optical surface, derive reflection for each polarisation and follow energy when the refracted wave becomes evanescent.

Read refraction from the clock shared by two media

A plane wave strikes a flat interface $z = 0$ between transparent isotropic dielectrics. Their permeabilities are μ_0; the refractive indices n_1, n_2 are real at the chosen frequency. Write every field as the real part of a complex amplitude times $e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$. Tangential electric and magnetic fields must match at every point x and every time t on the surface. This is impossible for independent exponential factors except where their temporal frequencies and tangential wave numbers agree. Therefore $\omega_i = \omega_r = \omega_t$, $\theta_r = \theta_i$, and

$$k_{\parallel} = \frac{\omega}{c} n_1 \sin\theta_i = \frac{\omega}{c} n_2 \sin\theta_t.$$

The interface fixes phase matching before it fixes amplitudes. Which amplitudes match depends on polarisation. Choose the xz-plane of incidence so that s polarisation has electric field along y and p polarisation has magnetic field along y.

Use one amplitude convention that works for both polarisations

Use the tangential electric component for all amplitudes. Let Y be the ratio of the tangential magnetic magnitude to that electric component for a forward wave, with the corresponding sign reversing for a backward wave. From $\mathbf{H} = (n/Z_0)\hat{\mathbf{k}} \times \mathbf{E}$, $Z_0 = \sqrt{\mu_0/\epsilon_0}$, the forward admittances are $Y_s = n\cos\theta/Z_0$ and $Y_p = n/(Z_0\cos\theta)$. The factors differ because an oblique p field has only a fraction $\cos\theta$ of its magnitude tangential to the surface.

$$E_i + E_r = E_t,$$

$$Y_1(E_i - E_r) = Y_2 E_t,$$

$$r = \frac{Y_1 - Y_2}{Y_1 + Y_2},$$

$$t = \frac{2Y_1}{Y_1 + Y_2}.$$

For propagating waves the normal component of mean energy flux is $Y|E_{\parallel}|^2/2$. Hence $R = |r|^2$ and $T = (Y_2/Y_1)|t|^2$. In a lossless interface $R + T = 1$. Some conventions reverse the reflected p-polarisation unit vector and therefore reverse the sign called r_p. Comparing tangential components explicitly removes that ambiguity.

Worked oblique reading

For air to glass with $n_1 = 1.0$, $n_2 = 1.5$ at 45°, Snell's law gives

$$\theta_t = 28.13^\circ.$$

Dropping the common factor $1/Z_0$, the s admittances are 0.707 and 1.323. Their reflection amplitude is -0.303, giving 9.20% reflected power. The p admittances are 1.414 and 1.701. Their tangential reflection amplitude is -0.0920, giving 0.847%. Giving only one reflectance for an unpolarised beam would hide this nearly elevenfold difference. At normal incidence the two admittances coincide for each medium, providing a sign and magnitude check.

Find the angle that cancels one reflected component

An interface is reflection-free for p polarisation when $n_1/\cos\theta_i = n_2/\cos\theta_t$. Combine this condition with Snell's law for distinct indices to get

$$\theta_i + \theta_t = \pi/2,$$

$$\tan\theta_B = n_2/n_1.$$

At the air–glass interface this angle is 56.31°. A small index change shifts that angular minimum, making a polarised measurement a way to infer index. Absorption introduces a complex admittance and generally turns the exact zero into a finite minimum.

Keep phase matching beyond the critical angle

For incidence from large to small index, $k_\parallel$ can exceed the magnitude $n_2\omega/c$ allowed for medium 2. Solve the dispersion relation there for the normal wave number:

$$k_{z,2}^2 = (n_2\omega/c)^2 - k_\parallel^2,$$

$$k_{z,2} = i\kappa,$$

$$\kappa = \frac{2\pi}{\lambda_0}\sqrt{n_1^2\sin^2\theta_i - n_2^2}.$$

The positive imaginary branch gives decay into $z > 0$. For a single such interface Y_2 is imaginary, so $|r| = 1$; a reactive magnetic-to-electric ratio means zero average normal flux there. A field still extends into the second medium and can carry energy parallel to the surface. Another material brought within a decay length changes the boundary problem and can enable transmission.

Worked length-scale distinction

For $n_1 = 1.5, n_2 = 1.0, \theta_i = 50^\circ, \lambda_0 = 600$ nm, the critical angle is 41.81°. The amplitude length $1/\kappa$ is 169 nm. Intensity follows squared field and therefore falls by e over 84.3 nm. A detector using intensity must state which length it reports. Close to the critical angle the penetration length becomes large; finite beams then sample a range of angles around the ideal plane-wave value.

Find a polarisation zero and the onset of total reflection

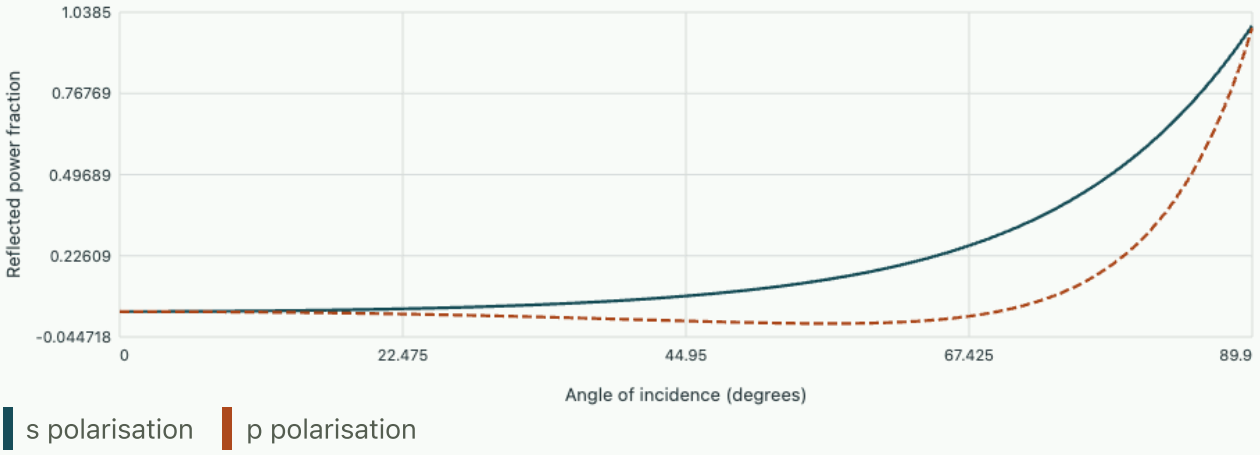
Air into glass of index 1.5 gives a p-polarised reflection zero at 56.3°. Reverse the indices, predict the critical angle, then inspect what happens to each reflected power fraction as incidence becomes more oblique.

Open this experiment online

Try the calculation

Incident refractive index **1**

Transmitted refractive index **1.5**



The Brewster minimum occurs at 56.3°. At 45°, reflected power is 9.2% (s) and 0.847% (p).

Read an amplitude measurement correctly

How can a transmitted field exceed the incident amplitude?

At normal incidence from index 1.8 to 1.2, calculate the reflected and transmitted tangential electric amplitudes and powers.

▼ **Work through the solution**

Here $r = (1.8 - 1.2)/(1.8 + 1.2) = 0.20$, $t = 1.20$. Reflected power fraction is 0.04. The transmitted power fraction is

$$(1.2/1.8)(1.20)^2 = 0.96.$$

The reduced wave admittance converts the larger electric amplitude to a smaller power than the incident beam.

Find the angular acceptance of a total-reflection measurement

A liquid has unknown index n. Light is incident from known glass of index 1.60, and total internal reflection first appears at $57.0^\circ \pm 0.2^\circ$. Infer n and a first-order uncertainty.

▼ **Work through the solution**

The grazing transmitted wave obeys $n = 1.60 \sin 57.0^\circ = 1.342$. Differentiation gives

$$dn = 1.60 \cos 57.0^\circ d\theta,$$

with $d\theta$ in radians. The quoted angular interval corresponds to about 0.0030 in index. Index uncertainty of the glass would add another contribution.

Further reading

- Griffiths, §9.3, plane waves crossing dielectric interfaces
- Tong, Electromagnetism in Matter, reflection and refraction

Study the full derivation

Oblique interfaces and films whose reflections cancel · Waves, radiation and relativity

A film only one quarter of a wavelength thick can suppress reflection from a transparent surface. Follow the boundary conditions all the way from an oblique incident wave to the input admittance of a stack, then calculate how much reflection returns when the wavelength, angle or layer order changes.

Chapter 22 · Why dispersion and absorption belong together

Obtain complex permittivity from a damped charge, relate its imaginary part to heating and distinguish wave crests, envelopes and a causal change in a signal.

Replace an instantaneous spring constant by a delayed response

A static dielectric constant is measured after the material has relaxed. At optical frequencies the bound charges have inertia and damping. Model a density N of independent bound charges of mass m, charge q, natural angular frequency ω_0 and damping rate $\gamma > 0$. A weak field gives a displacement satisfying $m\ddot{x} + m\gamma\dot{x} + m\omega_0^2 x = qE(t)$. The resulting polarisation is $P = Nqx$. This model assumes small displacements and an isotropic response. A crystal with several resonances needs additional modes and oscillator strengths.

For a sinusoidal field with the convention $e^{-i\omega t}$, differentiation multiplies a phasor by $-i\omega$. Dividing the equation by its dynamic stiffness gives

$$\chi(\omega) = \frac{P}{\epsilon_0 E} = \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega},$$

$$\omega_p^2 = \frac{Nq^2}{m\epsilon_0}.$$

Rationalise the denominator using $A = \omega_0^2 - \omega^2,\ B = \gamma\omega$. Then $\chi' = \omega_p^2 A/(A^2 + B^2)$ and $\chi'' = \omega_p^2 B/(A^2 + B^2)$. The real part changes sign across the resonance. The imaginary part stays positive at positive frequencies with our convention. In the low-frequency limit the static spring law reappears,

$$\chi(0) = \omega_p^2/\omega_0^2.$$

Decide which part of the response dissipates work

The field delivers power density $E\dot{P}$ to the bound charges. The polarisation current phasor is $-i\omega\epsilon_0\chi E$. For peak-amplitude phasors the cycle average of two real oscillating quantities is half the real part of one phasor times the complex conjugate of the other. Thus

$$\langle E\dot{P}\rangle = \tfrac{1}{2}\omega\epsilon_0\chi''|E|^2.$$

This result also follows by multiplying the oscillator equation by $N\dot{x}$. The kinetic and spring energy changes integrate to zero over one steady period, leaving $Nm\gamma\langle\dot{x}^2\rangle$ as heat. Positive loss therefore has a mechanical interpretation independent of any phasor notation. Changing to an $e^{+i\omega t}$ convention reverses the sign attached to the imaginary part, while the heat remains positive.

Worked response at resonance

Take $\omega_p^2 = 0.20\,\omega_0^2$ and $\gamma = 0.10\,\omega_0$. At $\omega = \omega_0$, $A = 0$ and $\chi = 2i$. If $\omega_0 = 2.0 \times 10^{12}$ rad/s and $E = 100$ V/m peak, the absorbed power density is 1.77×10^5 W/m³. This is local power where the field

has that amplitude. Applying it throughout a thick absorbing slab would neglect attenuation of the field with depth.

Recover the actual lengths and speeds in a material

For relative permeability one and a plane wave $e^{ikz-i\omega t}$,

$$k = (\omega/c)\sqrt{1+\chi}.$$

Choose $n = n' + i\kappa$ with $\kappa \geq 0$ for a passive medium. The field amplitude decays as $e^{-\omega\kappa z/c}$ and intensity as $e^{-2\omega\kappa z/c}$. At the resonance of the previous example, $1+\chi = 1+2i$. Equating real and imaginary parts of the squared index gives

$$n'^2 - \kappa^2 = 1,$$

$$2n'\kappa = 2,$$

$$n' = \sqrt{\tfrac{1}{2}(\sqrt{5}+1)},$$

$$\kappa = \sqrt{\tfrac{1}{2}(\sqrt{5}-1)}.$$

Here $n' = 1.272,\ \kappa = 0.786$. At the stated angular frequency the intensity attenuation length $c/(2\omega\kappa) = 95.3\,\mu\mathrm{m}$. A finite slab also reflects at its surfaces. Inferring bulk loss from transmitted power alone requires accounting for those reflections.

For a narrow packet, expand the real propagation constant around its carrier:

$$k'(\omega) \simeq k'_c + (\omega-\omega_c)\,dk'/d\omega,$$

where k' here denotes the real part of k. The carrier phase moves at ω_c/k'_c. If attenuation and curvature vary little across the packet bandwidth, the envelope delay over length L is $L\,dk'/d\omega$. Strong frequency-selective loss can reshape the peak so much that tracking it no longer measures the arrival of new information.

Link the two parts of susceptibility through causality

A causal linear material obeys $P(t) = \epsilon_0 \int_0^\infty h(\tau) E(t-\tau)\,d\tau$. Its susceptibility for our convention is $\int_0^\infty h(\tau) e^{i\omega\tau} d\tau$. For a decaying response this expression is analytic when $\operatorname{Im}\omega > 0$. Closing a contour in that half plane relates real and imaginary components by principal-value integrals. For a response decaying at large frequency one useful form is

$$\chi'(\omega) = \frac{2}{\pi}\,\mathcal{P}\int_0^\infty \frac{\omega'\chi''(\omega')}{\omega'^2-\omega^2}\,d\omega'.$$

Thus a narrow absorption band contributes to index variation at frequencies where the local absorption is small. A measurement over a short spectral window cannot uniquely evaluate the integral without

information about loss outside it. The damped-oscillator poles lie below the real axis, as the causal convention requires.

Change the damping of a bound charge

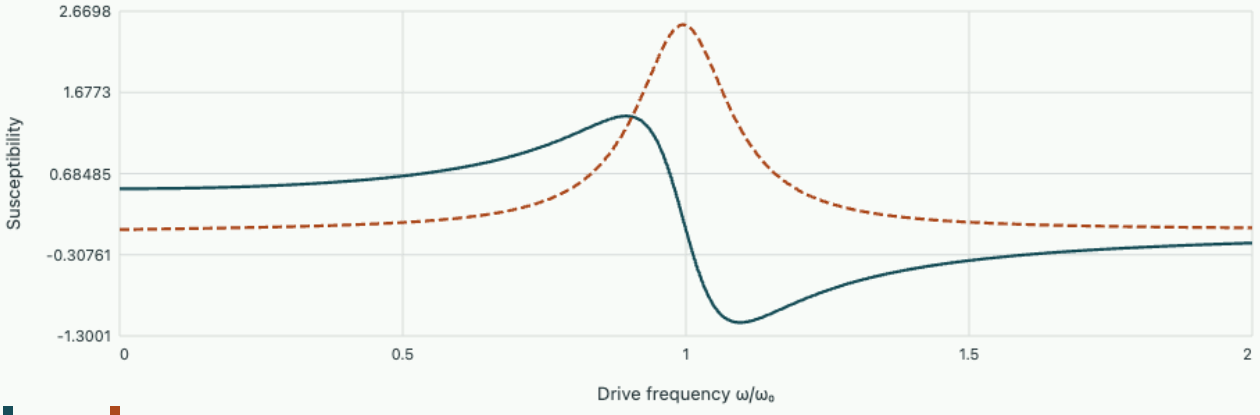
Track the in-phase and out-of-phase parts of one oscillator response. With strength 0.50 and damping 0.20 in the plotted units, the absorptive part at resonance is 2.50. Compare that point with the frequency of maximum absorption as damping increases.

Open this experiment online

Try the calculation

Damping γ/ω₀ **0.2**

Oscillator strength ω_p^2/ω_0^2 **0.5**



At ω = ω₀ the susceptibility is purely imaginary with Im χ = 2.5. At zero frequency the real susceptibility is 0.5.

Use attenuation and loss as independent checks

Infer an extinction coefficient

After correcting for both surfaces, a liquid transmits 25% of its incident intensity in 0.50 mm at vacuum wavelength 1.00 µm. Find κ and the field-amplitude ratio.

▼ **Work through the solution**

With $T = \exp(-4\pi\kappa L/\lambda_0)$,

$$\kappa = -\lambda_0 \ln(0.25)/(4\pi L) = 2.21 \times 10^{-4}.$$

The amplitude ratio is

$$\sqrt{T} = 0.50.$$

Using the amplitude decay law on intensity would overestimate κ by a factor of two.

Recover the conductivity limit

Set $\omega_0 = 0$ in the charge model and express the response as $J = \sigma E$. Find the low-frequency real conductivity.

▼ **Work through the solution**

Using $J = -i\omega P$ yields

$$\sigma(\omega) = \epsilon_0 \omega_p^2 / (\gamma - i\omega).$$

At $\omega \ll \gamma$,

$$\sigma \simeq \epsilon_0 \omega_p^2 / \gamma = N q^2 / (m\gamma).$$

The absence of a restoring force lets charges drift steadily instead of settling to a finite static polarisation.

Further reading

- Griffiths, §9.4, absorption and dispersion
- Tong, Electromagnetism in Matter, dispersion, Kramers–Kronig and the Drude model

Study the full derivation

Causal material response and the propagation of a finite pulse · Waves, radiation and relativity

An optical pulse entering a transparent slab is delayed, broadened and sometimes reshaped. Derive these effects from the motion of bound charge, and learn what a measured group delay actually says about a pulse carrying information.

Chapter 23 · Potentials, gauge freedom and the arrival of a signal

When a source changes at noon, which part of its history can affect a detector at twelve o'clock and one nanosecond? The answer emerges from Maxwell's equations once we specify both the source and the boundary condition in time.

Why introduce two potentials?

A magnetic field with zero divergence can, on a contractible region, be written as $\mathbf{B} = \nabla \times \mathbf{A}$. Faraday's law then says that $\mathbf{E} + \partial_t \mathbf{A}$ has zero curl. Define its scalar potential as $-\nabla\phi$. These two definitions solve the homogeneous Maxwell equations identically. Two inhomogeneous equations still determine how charge and current generate the fields. The potentials contain more components than necessary. For any smooth function χ, put

$$\mathbf{A}' = \mathbf{A} + \nabla\chi,$$
$$\phi' = \phi - \partial_t \chi.$$

The curl of a gradient vanishes. The two extra terms in the electric field cancel because mixed derivatives commute. This freedom lets us choose which equations to solve while keeping every force measurement identical.

> **Background · Retarded time**
>
> $$t - t_r = |\mathbf{r} - \mathbf{r}_s(t_r)|/c$$
>
> The source is evaluated at emission. For a moving source its location on the right also depends on the unknown emission time.

Uncoupling the field equations

Insert the potentials into Gauss's law. This gives

$$\nabla^2\phi + \partial_t \nabla\cdot\mathbf{A} = -\rho/\epsilon_0.$$

Now substitute them in Ampère–Maxwell and use the curl-of-curl identity. Keeping the gradient terms together gives

$$(\nabla^2 - c^{-2}\partial_t^2)\mathbf{A} = -\mu_0\mathbf{J} + \nabla(\nabla\cdot\mathbf{A} + c^{-2}\partial_t\phi).$$

Choose the Lorenz condition $\nabla\cdot\mathbf{A} + c^{-2}\partial_t\phi = 0$. A gauge transformation can achieve it by solving the scalar wave equation for χ with the negative of the original condition as source. Both potentials then obey wave equations with the same propagation speed. Under this choice, charge continuity is essential. Applying the wave operator to the Lorenz condition returns $-\mu_0(\nabla\cdot\mathbf{J} + \partial_t\rho)$, which vanishes for a conserved source. Residual gauge functions obey the homogeneous wave equation.

Constructing the retarded kernel

In a source-free spherical shell a radial scalar field ψ obeys $\partial_r^2(r\psi) = c^{-2}\partial_t^2(r\psi)$. The one-dimensional equation has an outgoing branch $r\psi = f(t - r/c)$ and an incoming branch $g(t + r/c)$. An isolated source with no incident wave selects the outgoing branch. The response to an instantaneous impulse is supported at $t = r/c$. Its spatial normalisation follows by integrating the static kernel over a tiny sphere, where $-\nabla^2(1/4\pi r) = \delta^3(\mathbf{r})$. The causal kernel is therefore $\delta(t - r/c)/(4\pi r)$. Integrating the source over time removes the delta function:

$$\phi(\mathbf{x},t) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{x}', t - R/c)}{R}\, d^3x',$$

$$\mathbf{A}(\mathbf{x},t) = \frac{\mu_0}{4\pi}\int \frac{\mathbf{J}(\mathbf{x}', t - R/c)}{R}\, d^3x',$$

$$R = |\mathbf{x} - \mathbf{x}'|.$$

Every volume element contributes at its own source time. If an incident wave is already present, add its homogeneous solution. In Coulomb gauge the scalar potential instead satisfies an instantaneous Poisson equation. Its instantaneous contribution to $-\nabla\phi$ combines with $-\partial_t\mathbf{A}$ to give the same causal electric field.

Two calculations with potentials

A pulse recorded at two distances

Measure an isolated outgoing spherical wave component in a source-free exterior. At radius 0.30 m its field trace peaks at 10.0 ns. At 1.80 m its peak arrives $1.50/c = 5.00$ ns later. The peak amplitude is $0.30/1.80 = 1/6$ of the first. Multiplying each trace by its radius removes spreading. Shifting the second trace back by 5.00 ns then superposes them. For an electromagnetic dipole these statements describe its outgoing transverse radiation term along the same ray. Reactive near-field terms have different radial dependences.

Two vector potentials for a uniform field

For a constant magnetic field $B\hat{\mathbf{z}}$ one choice is $\mathbf{A} = Bx\hat{\mathbf{y}}$, $\phi = 0$. Choose $\chi = -Bxy/2$. Its gradient adds $-By\hat{\mathbf{x}}/2 - Bx\hat{\mathbf{y}}/2$, giving $\mathbf{A}' = (-By\hat{\mathbf{x}} + Bx\hat{\mathbf{y}})/2$. Its curl has z component $B/2 - (-B/2) = B$. The line integrals around any closed loop match because the integral of the added gradient is zero.

Compare the arrival of a compact pulse at two radii

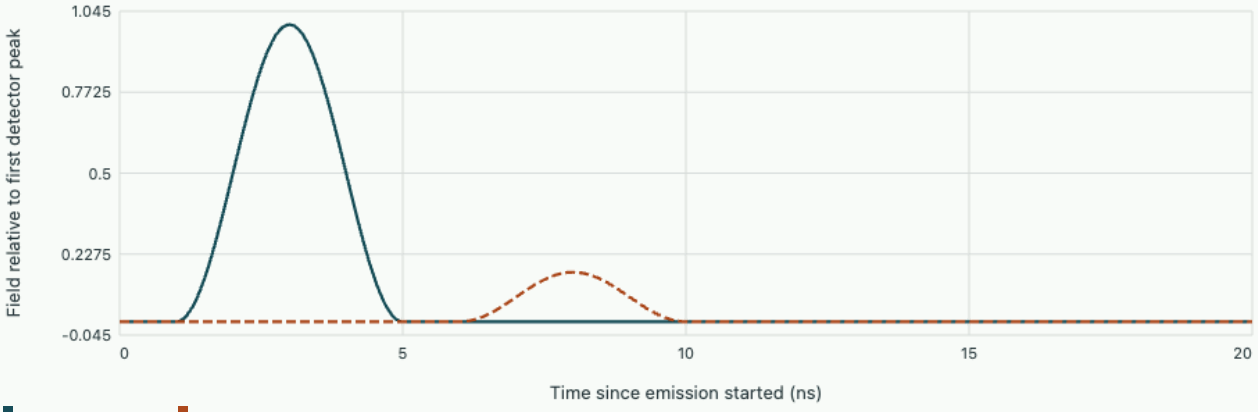
This source-free outgoing component has a finite start and finish. A 1.80 m detector records the front 5.00 ns after a 0.30 m detector and measures one sixth of its amplitude. Can changing the pulse duration change either observation?

Open this experiment online

Try the calculation

Second detector radius (m) **1.8**

Pulse duration (ns) **4**



At 0.30 m Second detector

The second trace starts 5 ns after the first. Its peak is 16.7% of the first peak.

Problem sheet · delays and a time-dependent gauge

A finite source

Two ends of a source are 12 cm apart along a detector's line of sight. Bound the difference in retarded source times. At what frequency does the corresponding maximum phase span reach one radian?

▼ **Work through the solution**

The triangle inequality gives $|R_1 - R_2| \le 0.12$ m. Thus $\Delta t \le 0.400$ ns. For a harmonic source the maximum phase difference is $2\pi f \Delta t$. One radian requires $f = 398$ MHz approximately. Well below that frequency one common retarded time can be accurate for this part of the phase. A dipole approximation also requires a source small compared with the observation distance.

A gauge with a time derivative

Start with zero potentials. Set $\chi = \alpha x t$, where α is constant. Determine both transformed potentials and check the fields.

▼ **Work through the solution**

The new potentials are $\mathbf{A}' = \alpha t \hat{\mathbf{x}}$ and $\phi' = -\alpha x$. Hence $-\nabla\phi' = +\alpha\hat{\mathbf{x}}$ and $-\partial_t \mathbf{A}' = -\alpha\hat{\mathbf{x}}$. Their sum and the curl of the spatially uniform vector potential are both zero. The example also satisfies the Lorenz condition, consistent with χ solving the homogeneous wave equation.

Further reading

- Griffiths, 5th ed., §§10.1–10.2 · potentials and retarded sources
- Fließbach, chapter 17 · general solution of Maxwell's equations
- Cambridge Electromagnetism, §5.1 · retarded potentials

Study the full derivation

Constructing the retarded Green function from an outgoing spherical shell · Waves, radiation and relativity

A short source pulse reaches two receivers at different times and amplitudes. Derive the signal from a localized spacetime impulse, see why the potential is evaluated at a past time, and identify which observables change when the source occupies an extended region.

Chapter 24 · A moving charge, its retarded position and its radiation

A moving source moves while its field travels. Finding its position on the past light cone supplies an extra factor in its potential and explains why an accelerated charge can carry energy to arbitrarily large distances.

Liénard–Wiechert potentials from the emission event

Let the charge follow the prescribed path $\mathbf{z}(s)$. At time t and detector position $\mathbf{x}$, define $R(s) = |\mathbf{x} - \mathbf{z}(s)|$ and $\mathbf{n} = (\mathbf{x} - \mathbf{z})/R$. The emission time s solves $h(s) = s + R(s)/c - t = 0$. Differentiate to obtain

$$h' = 1 - \mathbf{n} \cdot \boldsymbol{\beta} = \kappa,$$

with $\boldsymbol{\beta} = \dot{\mathbf{z}}/c$. For a subluminal charge $\kappa > 0$, so the root is unique when it exists. Two emissions a time ds apart arrive only κds apart. Looking forwards along a fast charge's path shortens this interval.

Insert $\rho = q\delta^3(\mathbf{x}' - \mathbf{z}(s))$ into the spacetime integral for the retarded potential. Integrate over $\mathbf{x}'$ first. The remaining integral is proportional to $\int ds\, \delta(h(s))/R(s)$. Changing variable from s to h introduces $1/\kappa$. With the current density $q\dot{\mathbf{z}}\delta^3(\mathbf{x}' - \mathbf{z})$, the same argument yields the vector potential:

$$\phi = \frac{q}{4\pi\epsilon_0}\left[\frac{1}{\kappa R}\right]_{\text{ret}},$$

$$\mathbf{A} = \left[\frac{\mathbf{v}}{c^2}\right]_{\text{ret}}\phi.$$

Simply replacing the position of a static charge with its retarded position misses the Jacobian. Both the geometry and the current vary when taking derivatives of these potentials.

Distinguish persistent and radiative parts

Implicit differentiation of $h = 0$ gives

$$\partial_t s = 1/\kappa,$$

$$\nabla s = -\mathbf{n}/(c\kappa).$$

Apply these identities in $-\nabla\phi - \partial_t\mathbf{A}$. Derivatives of the velocity supply an acceleration term, while the remainder can be expressed using the retarded velocity and geometry:

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0}\left[\frac{(1-\beta^2)(\mathbf{n}-\boldsymbol{\beta})}{\kappa^3 R^2} + \frac{\mathbf{n}\times[(\mathbf{n}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}]}{c\kappa^3 R}\right]_{\text{ret}},$$

$$\mathbf{B} = \frac{\mathbf{n}\times\mathbf{E}}{c}.$$

A detector subtends a fixed solid angle with area proportional to R^2. The square of the first field falls as R^{-4} and its flux through the detector vanishes at large radius. The second field is transverse and its

square falls as R^{-2}, so finite power travels to infinity. Uniform motion leaves only the first term.

Recover the nonrelativistic radiation law

For $\beta \ll 1$, $\kappa \simeq 1$. If θ is measured from the acceleration, the radiative electric-field magnitude is $qa\sin\theta/(4\pi\epsilon_0 c^2 R)$ in absolute value. Its magnetic partner has magnitude E/c. Therefore the outward Poynting flux is $\epsilon_0 c E^2$, and integration over a distant sphere gives

$$\frac{dP}{d\Omega} = \frac{q^2 a^2}{16\pi^2 \epsilon_0 c^3}\sin^2\theta,$$

$$P = \frac{q^2 a^2}{6\pi\epsilon_0 c^3}.$$

The angular factor integrates to $2\pi\int_0^\pi \sin^3\theta\, d\theta = 8\pi/3$. At finite speed with acceleration parallel to velocity, the vector triple product in the numerator has magnitude $(a/c)\sin\theta$. The flux per unit arrival time contains κ^{-6}. Converting to power per unit emission time multiplies by $dt/ds = \kappa$, leaving the angular dependence $\sin^2\theta/(1-\beta\cos\theta)^5$. The distinction matters whenever emissions bunch up at the detector.

Reading charge motion from a signal

A moving source changes its pulse spacing

A charge moving at $0.60c$ directly toward a distant observer emits two brief disturbances 2.0 ns apart in the lab frame. It travels 0.36 m between emissions. The second propagation time is shorter by 1.2 ns, leaving an arrival interval of 0.80 ns. Viewed from behind, the interval is 3.2 ns. This calculation concerns arrival times. An observer exactly along a parallel acceleration sees zero radiation in that direction.

How much power misses the polar caps?

For a slow charge, collect radiation at angles $60^\circ \le \theta \le 120^\circ$. Put $u = \cos\theta$. The collected fraction is

$$\int_{-1/2}^{1/2} (1-u^2)\, du/(4/3) = 11/16.$$

This equatorial belt spans half the solid angle but receives 68.75% of the power. Integrating with respect to θ alone would overcount the polar regions.

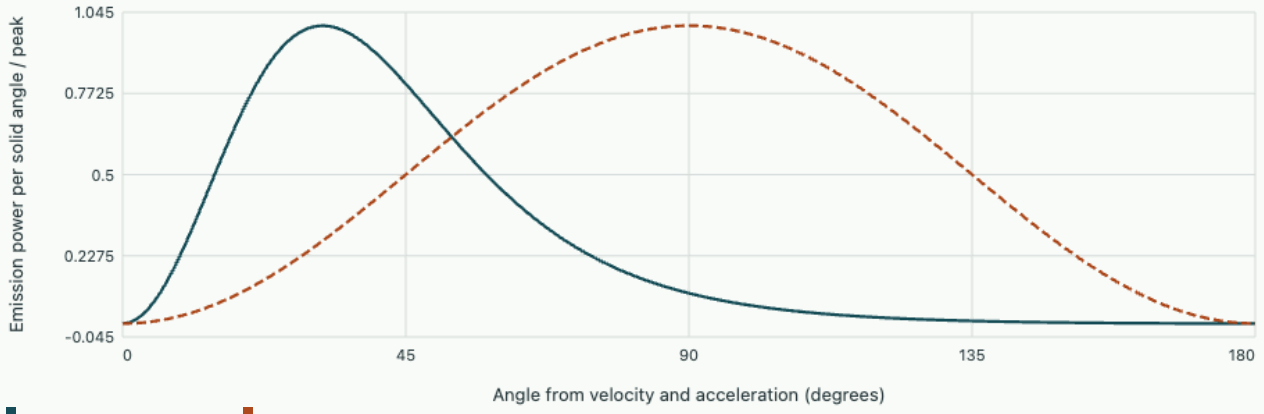
Track the brightest direction of an accelerating charge

The two angular curves are each normalised to their own maximum. For parallel acceleration and β = 0.60, the maximum moves from 90° to about 31.8°. Compare peak location and the separately reported unnormalised peak height.

Open this experiment online

Try the calculation

Source speed β 0.6



Peak emission is at 31.8°. Its unnormalised height is 9.82 times the slow-source maximum at the same instantaneous acceleration.

Problem sheet · inversion and beaming

Solve a retarded time

A charge moves along $z = 0.4cs$. At $t = 0$ a detector at $z = 3.0$ m receives its field. Find the emission time and the distance travelled by the signal.

▼ **Work through the solution**

Before arrival the charge is behind the origin. Here $R = 3.0 - 0.4cs$ in metres, so $0 = 0.6s + 3.0/c$. The emission time is -16.7 ns, at $z = -2.0$ m. The propagation distance is 5.0 m. Using the current separation would predict only 10.0 ns of travel.

Locate the angular maximum

For parallel acceleration differentiate $F(u) = (1 - u^2)/(1 - \beta u)^5$, $u = \cos\theta$, and predict the maximum for $\beta = 0.6$.

▼ **Work through the solution**

The numerator of F', after dropping positive factors, is $5\beta - 2u - 3\beta u^2$. The physical root is $u = 5\beta/(1 + \sqrt{1 + 15\beta^2})$. At 0.6 it equals 0.850, giving $\theta \simeq 31.8^\circ$. The other root lies outside the physical interval. At zero speed the limiting angle is 90°, matching the dipole pattern.

Practicum · distinguish a large near field from outgoing power

A neutral dipole of size much less than the wavelength oscillates with complex moment $p\,\hat{\mathbf{z}}$ at angular frequency ω. Two probes lie in its equatorial plane. An experimentalist proposes to estimate outgoing power as $\epsilon_0 c|E_\theta|^2/2$ times detector area at each radius. Is this plane-wave calibration valid close to the source? Compare field amplitudes, phases and a conserved flux before forming an opinion.

Recover the two equatorial fields from potentials

Use $e^{-i\omega t}$, $k=\omega/c$, $x=kr$, and the Lorenz-gauge dipole potentials $\phi = p\cos\theta\,(1-ikr)e^{ikr}/(4\pi\epsilon_0 r^2)$,

$$\mathbf{A} = -i\mu_0\omega p\, e^{ikr}\hat{\mathbf{z}}/(4\pi r).$$

Recover E_θ and B_ϕ at $\theta=\pi/2$.

▼ **Follow the calculation and check the result**

At the equator $\hat{\mathbf{z}} = -\hat{\boldsymbol{\theta}}$. Applying $-r^{-1}\partial_\theta$ to the scalar potential gives $p(1-ix)e^{ix}/(4\pi\epsilon_0 r^3)$. The vector part $i\omega A_\theta$ adds $-px^2e^{ix}/(4\pi\epsilon_0 r^3)$. Define the common complex scale $K = pe^{ix}/(4\pi\epsilon_0 r^3)$. Then

$$E_\theta = K(1-ix-x^2),$$

$$cB_\phi = K(-ix-x^2).$$

For the magnetic field use $\nabla\times[f(r)\hat{\mathbf{z}}] = f'(r)\hat{\mathbf{r}}\times\hat{\mathbf{z}}$ with $f=-i\mu_0\omega p\,e^{ikr}/(4\pi r)$. Both fields contain induction and radiation terms. The extra electrostatic term appears only in E_θ. It prevents the local field impedance from equalling the plane-wave value.

Find what two phase-sensitive probes would report

Multiply $E_\theta(cB_\phi)^*$, retaining both quadratures. Determine the mean outward energy flux. How quickly does the out-of-phase component disappear relative to it?

▼ **Follow the calculation and check the result**

The complex product is $|K|^2(1-ix-x^2)(ix-x^2) = |K|^2(x^4+ix)$. Radiation measurements average real instantaneous products, so the mean radial flux is $\epsilon_0 c|K|^2x^4/2$, proportional to r^{-2} at every radius outside the ideal source. The imaginary component of the complex Poynting vector measures quadrature exchange; its ratio to the real part is x^{-3} in this convention. It does not add a second time-averaged outward power.

At $kr=0.50$ the quadrature component is eight times the real part. At $kr=2.00$ it is one eighth. A spherical integral restores the angular factor $\sin^2\theta$ and area r^2, yielding $\omega^4|p|^2/(12\pi\epsilon_0 c^3)$, independent of radius. This checks the result against the dipole radiation formula while retaining the full near-zone amplitudes.

Quantify the bias of an electric-only reading

Compute the ratio of the proposed plane-wave estimate to the true mean radial flux for the two probes. Design a measurement that can separate reactive and radiative fields.

▼ **Follow the calculation and check the result**

The ratio is

$$|1 - ix - x^2|^2 / x^4 = 1 - x^{-2} + x^{-4}.$$

It is 13 at $x = 0.5$: a large electric field can coexist with comparatively little outward energy transport. At $x = 2$ the ratio is 0.8125; the same calibration now underestimates the flux by 18.75%. A monotonic intuition based on discarding all inverse powers except $1/r$ can even miss the sign of the error.

For a phase-sensitive measurement, record E_θ and B_ϕ against the same clock, form the real part of their cross-product and compare $r^2 \langle S_r \rangle$ at several radii. Simultaneously examine the imaginary-to-real ratio and check its r^{-3} trend. For a finite-sized source keep the closest probe outside the source and test the dipole approximation by changing the observation angle. Other multipoles or a reflected wave break the simple three-term model.

Further reading

- Griffiths, 5th ed., §§10.3, 11.2 · moving charges and power
- Fließbach, chapter 23 · accelerated charge
- Cambridge Electromagnetism, §§5.2–5.3 · radiation from charges

Study the full derivation

Constructing the retarded Green function from an outgoing spherical shell · Waves, radiation and relativity

A short source pulse reaches two receivers at different times and amplitudes. Derive the signal from a localized spacetime impulse, see why the potential is evaluated at a past time, and identify which observables change when the source occupies an extended region.

Chapter 25 · Changing observers and classifying electromagnetic fields

Can an observer move so that a measured electric field disappears? Two numbers constructed from the fields settle the question before a velocity is chosen. The moving observer's force measurements explain why the classification works.

Measure fields at the same event

Take two inertial laboratories with parallel axes. The primed laboratory moves with constant velocity $v\hat{\mathbf{x}}$ relative to the unprimed one. Compare fields at one spacetime event. The position labels and times assigned to that event differ by $x' = \gamma(x - vt)$ and $t' = \gamma(t - vx/c^2)$, where $\gamma = (1 - v^2/c^2)^{-1/2}$. Comparing both experiments at identical numerical position and time labels would usually compare different events.

How can a primed observer measure E'_y? Release a probe initially at rest in the primed frame. In the unprimed frame the probe has velocity $v\hat{\mathbf{x}}$, so its transverse force is $q(E_y - vB_z)$. The transverse momentum p_y is invariant under this boost, but along the probe's instantaneous trajectory $dt' = dt/\gamma$. Thus $dp'_y/dt' = \gamma dp_y/dt$, and $E'_y = \gamma(E_y - vB_z)$. The other transverse component is $E'_z = \gamma(E_z + vB_y)$. Longitudinal electric field is unchanged. A probe with adjustable primed velocity can also distinguish the magnetic force. The complete pair of rules is

$$\mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel},$$

$$\mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel},$$

$$\mathbf{E}'_{\perp} = \gamma(\mathbf{E}_{\perp} + \mathbf{v} \times \mathbf{B}),$$

$$\mathbf{B}'_{\perp} = \gamma(\mathbf{B}_{\perp} - \mathbf{v} \times \mathbf{E}/c^2).$$

Which differences survive a boost?

Examine perpendicular laboratory fields $E_y = E$, $cB_z = b$. In their common units the transformed pair is $e' = \gamma(E - \beta b)$,

$$b' = \gamma(b - \beta E).$$

Squaring and subtracting cancels the cross terms. The factor $\gamma^2(1 - \beta^2)$ is unity. In a general orientation the parallel parts remain unchanged and the other pair of transverse components cancels in the same manner. The scalar product also remains unchanged. The two invariants are

$$I_1 = |\mathbf{E}|^2 - c^2|\mathbf{B}|^2,$$

$$I_2 = \mathbf{E} \cdot \mathbf{B}.$$

If $I_2 \neq 0$, neither field can be transformed away. For nonzero orthogonal fields with $I_1 < 0$, choose $\mathbf{v} = \mathbf{E} \times \mathbf{B}/B^2$. Its speed is $E/B < c$, and it cancels $\mathbf{E}'$. If $I_1 > 0$ choose $\mathbf{v} = c^2\mathbf{E} \times \mathbf{B}/E^2$ to cancel $\mathbf{B}'$. For nonzero orthogonal fields of equal strength in common units the required speed would be c. A

plane wave illustrates this null case. A zero field also has both invariants zero, so both vanishing invariants alone do not prove that a wave is present.

Two changes of observer

An observer travelling with drifting charges

Let $E_y = 1.80 \times 10^5$ V/m and $B_z = 1.00$ mT. The cancelling velocity is $+1.80 \times 10^8$ m/s along x, approximately $0.600c$. Then $B_z' = B_z\sqrt{1 - E^2/(cB_z)^2} \simeq 0.800$ mT. A charge moving with that primed laboratory feels zero force at the chosen event. In the laboratory, the electric and magnetic terms cancel directly. The statements concern one probe velocity; charges moving differently in the primed frame can experience magnetic force.

Devices following a light pulse

A plane pulse travelling in $+x$ has $E_y = cB_z$. A detector moving along $+x$ measures

$$E_y'/E_y = \gamma(1 - \beta) = \sqrt{(1 - \beta)/(1 + \beta)}.$$

At $\beta = 0.60$ this is $1/2$, and the intensity ratio is $1/4$. Transform a phase $kx - \omega t$ using $k = \omega/c$. Its frequency receives the same factor $1/2$. Both detectors obtain $E'/B' = c$.

Why a neutral wire can have a charge density

For enclosed charge to agree between inertial descriptions, the charge and current densities must form the four-current $J^\mu = (c\rho, \mathbf{J})$. A boost gives

$$\rho' = \gamma(\rho - vJ_x/c^2),$$

$$J_x' = \gamma(J_x - v\rho).$$

For equal positive and negative charge densities producing a laboratory current, $\rho = 0$ but $J_x \neq 0$. In the moving laboratory the densities do not cancel. The change comes from how a simultaneous spatial slice intersects the worldlines of moving charge. Dividing a fixed list of particles by a contracted segment while keeping its endpoint events fixed would miscount particles crossing either endpoint.

Find a frame that cancels a field by varying its velocity

The axes are oriented as in the crossed-field derivation. For perpendicular fields E = 90 and cB = 150 kV/m, a boost at 0.60c cancels E and leaves cB′ = 120 kV/m. Exchange the two inputs and predict which field can now vanish.

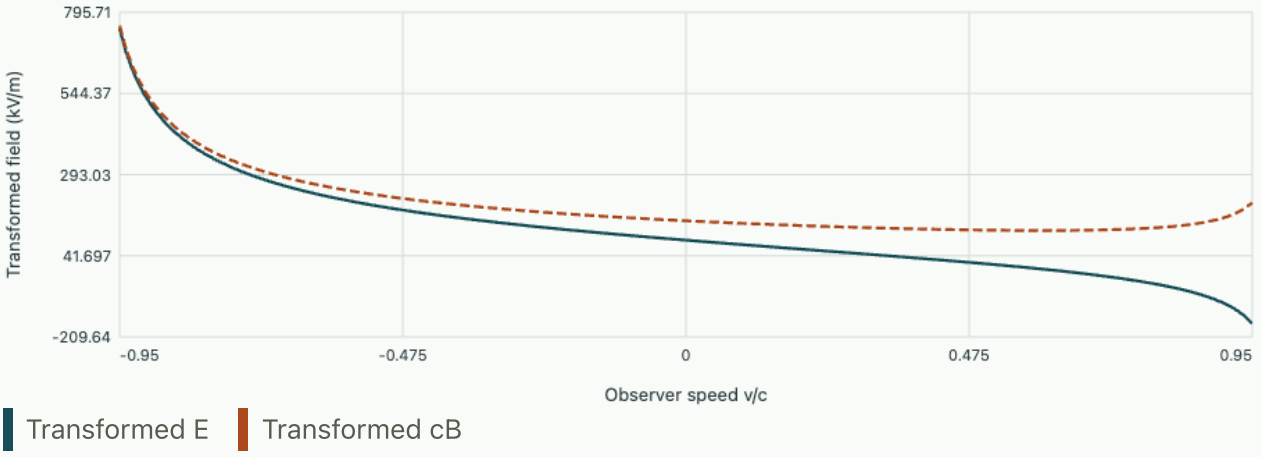
Open this experiment online

Try the calculation

Electric field E (kV/m) **90**

Magnetic field cB (kV/m) **150**

Observer speed β **0.6**



At the selected speed E′ = 0 kV/m and cB′ = 120 kV/m. E′² − c²B′² = -14400 (kV/m)².

Problem sheet · classification and measurements

Classify before calculating

Take a field with $\mathbf{E} = E_0\hat{\mathbf{y}}$ and $c\mathbf{B} = 2E_0\hat{\mathbf{y}} + E_0\hat{\mathbf{z}}$. Could either field disappear? Explain why consulting the difference of squares alone gives incomplete information.

▼ **Work through the solution**

The magnetic magnitude is $\sqrt{5}E_0/c$, making $I_1 = -4E_0^2$. However $I_2 = 2E_0^2/c \neq 0$. A frame with either field zero would have $I_2 = 0$, contradicting its observed invariant value.

A density reconstructed from a current

A laboratory measures $\rho = 0$ and $J_x = 9.0$ A/m². A second lab moves at $0.60c$. Determine its current density and charge density, including signs.

▼ **Work through the solution**

Here $\gamma = 1.25$. The transformed current is $+11.25$ A/m². The transformed density is $-1.25(0.60)(9.0)/c = -2.25 \times 10^{-8}$ C/m³. Inserting these into $c^2\rho'^2 - J_x'^2$ returns -81 in squared current-density units, a check independent of sign conventions in the voltage probes.

Practicum · reconstruct a crossed-field trajectory from two clocks

A positron initially rests in a laboratory with uniform crossed fields $\mathbf{B} = B\hat{\mathbf{z}}$,

$$\mathbf{E} = E\hat{\mathbf{y}},$$

with $B = 10.0$ mT and $E/(cB) = 0.600$. Ignore radiation losses and edge effects. Predict its maximum speed, arrival point after one cycle, and energy gain halfway through the cycle. Work by transforming whole events between observers. A velocity transformation alone cannot predict the travel time.

Choose the frame in which the trajectory is simplest

Find a frame with zero electric field, transform the initial particle velocity, and derive the orbit radius and frequency in that frame. Define what a “cycle” means for this measurement.

▼ **Follow the calculation and check the result**

Boost along $+x$ at $v_d = E/B = 0.600c$ with $\gamma_d = 1.25$. Then $E'_y = 0$ and $B'_z = B/\gamma_d = 8.00$ mT. Initially the particle has $v'_x = -v_d$, so its primed energy is constant at $\gamma_d mc^2$. With positive charge e, its initial magnetic force points towards $+y$. Defining $\Omega' = eB'/(\gamma_d m)$ and $R' = \gamma_d m v_d/(eB')$, one obtains

$$x' = -R'\sin(\Omega' t'),$$

$$y' = R'[1 - \cos(\Omega' t')].$$

Differentiating verifies the initial velocity and yields the rotating transverse acceleration. The primed radius is approximately 0.160 m and the primed period $T' = 2\pi/\Omega'$ is 5.58 ns. One cycle means two successive events at which the primed particle returns to its initial location and velocity. In the laboratory these events occur at different positions.

Measure time and the maximum lab speed

Transform the coordinates of the first return event to the laboratory. At what part of its orbit does the particle reach its maximum lab speed? Why is it incorrect to use a single constant γ_d to convert every subinterval along the curved path?

▼ **Follow the calculation and check the result**

For a full primed period $\Delta x' = 0$. The inverse Lorentz transformation gives

$$\Delta t = \gamma_d T' = 6.98\,\text{ns},$$

$$\Delta x = \gamma_d v_d T' = 1.26\,\text{m}.$$

Hence the average lab drift is $0.600c$ along $+x$. For an arbitrary subinterval the formula is instead $\Delta t = \gamma_d(\Delta t' + v_d \Delta x'/c^2)$. Away from rest or half-period endpoints its second term generally contributes.

The primed momentum has largest positive x' component at half a period, when $v'_x = v_d$. Then $v_x = 2v_d/(1 + v_d^2/c^2) = 0.88235c$, $v_y = 0$. The lab energy transformation $\mathcal{E} = \gamma_d(\mathcal{E}' + v_d p'_x)$ confirms that this is the energy maximum and hence the speed maximum. At the end of a period the particle returns to laboratory rest even though the apparatus has seen a net displacement.

Check the energy with an independent voltage measurement

At half a cycle compute the work performed by the actual laboratory electric field from the displacement alone. Compare with the relativistic kinetic energy and explain how it can be zero over a whole cycle.

▼ **Follow the calculation and check the result**

At the top of the primed orbit $y' = 2R'$. A boost along x leaves $y = 2R'$ unchanged. The electric work is $eE(2R') = 2\gamma_d^2\beta_d^2 mc^2 = 1.125mc^2$, or about 575 keV. The lab maximum speed gives $\gamma_{\max} = (1 + \beta_d^2)/(1 - \beta_d^2) = 2.125$, so its kinetic energy is the same $1.125mc^2$. The magnetic force performs zero work throughout. During the second half the particle moves back down in y, and the electric field removes exactly this kinetic energy.

At a fixed laboratory event a numerical trajectory can be checked against both $\mathcal{E} - mc^2 = eEy$ and the parametric orbit. Comparing the two paths at equal numerical values of t and t' would compare different events and could create apparent errors in both period and speed. At each successive cusp the lab velocity momentarily vanishes; the motion on either side reverses its transverse component. The cusp spacing supplies an additional check on the drift distance.

Further reading

- Griffiths, 5th ed., §§12.1–12.3 · relativity and field transformations
- Fließbach, chapter 22 · field transformations
- Cambridge Electromagnetism, chapter 4 · relativity

Study the full derivation

What different observers measure · Waves, radiation and relativity

A moving observer carries clocks, rulers and a calibrated test charge. A statement about the electric field must specify which observer held that test charge at rest. In this chapter we derive how different measurements can agree on the force law and yet assign different electric fields, magnetic fields and charge densities to the same physical apparatus. A current-carrying wire and a light pulse will provide independent checks.

Chapter 26 · From local symmetry to Maxwell's equations

Some electromagnetic equations describe the geometry of a potential. Others describe its response to charge. Tensor notation identifies the distinction precisely, and a variational argument shows why the source must obey a local conservation law.

Build an antisymmetric tensor from four potentials

Write $x^\mu = (ct, x, y, z)$, $\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$,

$$A^\mu = (\phi/c, \mathbf{A}),$$

$$J^\mu = (c\rho, \mathbf{J}).$$

Greek indices run from 0 to 3 and repeated upper and lower indices are summed. Lowering a spatial index changes its sign. Thus

$$\partial^\mu = (c^{-1}\partial_t, -\nabla).$$

This convention fixes every sign below.

Form $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. Then

$$F^{0i} = \partial_t A_i^{\mathrm{cart}}/c + \partial_i \phi/c = -E_i^{\mathrm{cart}}/c.$$

For spatial components $F^{12} = -\partial_x A_y + \partial_y A_x = -B_z$ with Cartesian subscripts on the right. Antisymmetry leaves six independent quantities. These are the three electric and three magnetic components. Adding a four-gradient to the potential leaves F unchanged because its mixed derivatives commute.

Unpack the two tensor equations

Consider $\partial_\mu F^{\mu\nu} = \mu_0 J^\nu$. For $\nu = 0$ the only terms are spatial:

$$\partial_i F^{i0} = \nabla \cdot \mathbf{E}/c = \mu_0 c\rho,$$

which is Gauss's law. For $\nu = 1$, $F^{01} = -E_x/c$, $F^{21} = B_z$, $F^{31} = -B_y$. The equation gives

$$(\nabla \times \mathbf{B})_x - c^{-2}\partial_t E_x = \mu_0 J_x.$$

Cyclic permutations give the remaining components. The other half of Maxwell's equations is an identity:

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0.$$

Insert the definition of F. Each second derivative cancels a partner of opposite sign. With three spatial indices the identity says $\nabla \cdot \mathbf{B} = 0$; with a time index it gives Faraday's law. Applying ∂_ν to the sourced equation also gives zero on the left by antisymmetry, enforcing $\partial_\nu J^\nu = 0$.

Vary the potential at fixed external current

We seek a Lorentz scalar built quadratically from the fields, linearly from the source and potential, and locally from their values at a spacetime event. In SI units choose

$$S = \frac{1}{c}\int d^4x \left[-\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu} - J^\mu A_\mu\right].$$

The factor $1/c$ converts dx^0 to seconds. The two antisymmetric terms in $\delta F_{\mu\nu}$ contribute equally. Thus the field part varies as $-F^{\mu\nu}\partial_\mu \delta A_\nu/\mu_0$. Integrate by parts, fixing $\delta A_\nu = 0$ on the boundary. The remaining arbitrary variation multiplies

$$\frac{1}{\mu_0}\partial_\mu F^{\mu\nu} - J^\nu = 0.$$

Under $A_\mu \mapsto A_\mu + \partial_\mu \chi$ the source term changes by $-J^\mu\partial_\mu\chi$. Integration by parts leaves a boundary term and $+\chi\partial_\mu J^\mu$. Gauge independence of the action for arbitrary interior χ therefore requires charge continuity. The same conservation law has appeared by a second route.

Read the action in familiar variables

Why a plane wave has zero field Lagrangian density

A direct contraction gives

$$F_{\mu\nu}F^{\mu\nu} = 2(B^2 - E^2/c^2).$$

The vacuum field Lagrangian density is $\epsilon_0 E^2/2 - B^2/(2\mu_0)$, whereas the physical energy density is the sum of those two positive terms. For a plane wave $E = cB$, the Lagrangian density vanishes; the energy density equals $\epsilon_0 E^2$. At $E = 100$ V/m instantaneous energy density is 8.85×10^{-8} J/m³. The cancellation in the action carries a distinct physical meaning from an empty field.

Test the stationary potential inside a capacitor

Between plane electrodes at $0, V_0$, take $\phi = V_0[x/L + \eta \sin(\pi x/L)]$. Each trial curve has the correct endpoints. For no space charge the static functional per area is $U = \epsilon_0 \int_0^L (\partial_x\phi)^2 dx/2$. Its cross term integrates to zero because the sine perturbation vanishes on the boundaries, and the squared cosine averages to $1/2$. Consequently

$$U/U_{\text{exact}} = 1 + \pi^2\eta^2/2.$$

A perturbation with $\eta = 0.10$ raises the functional by 4.93%. The flattening of the energy near $\eta = 0$ tests stationarity directly.

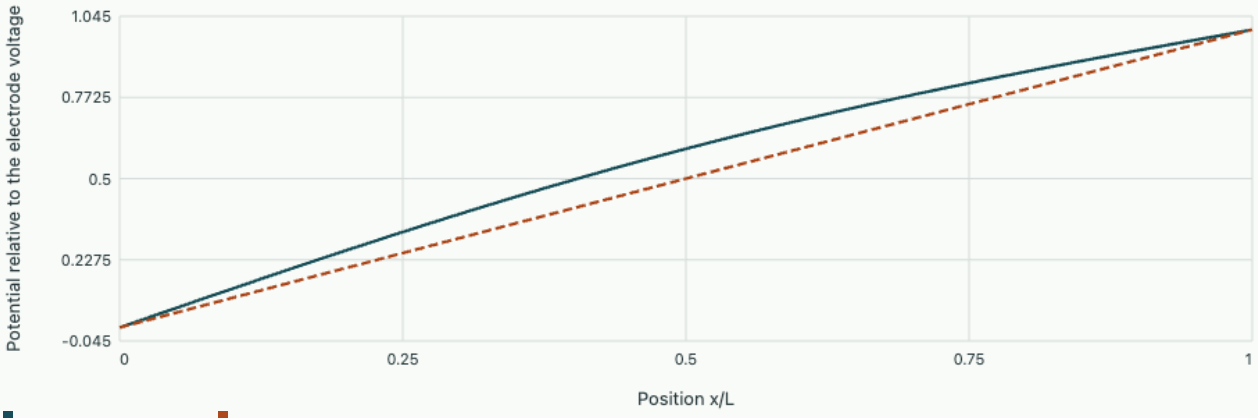
Perturb the potential while keeping both electrodes fixed

Every trial curve has the prescribed endpoints. At amplitude 0.10 the trial energy is 4.93% above the exact value. Compare positive and negative perturbations, then double the amplitude and predict how the energy penalty changes.

Open this experiment online

Try the calculation

Trial distortion η **0.1**



The trial field energy exceeds the stationary value by 4.93%. Reversing the distortion changes its direction while preserving this energy difference.

Problem sheet · gauge and variational checks

Introduce a known static charge

For fixed boundary potentials, vary $I[\phi] = \int (\epsilon_0 |\nabla\phi|^2/2 - \rho\phi)\, dV$. Which PDE follows? Why is it necessary to specify the sign of the linear term?

▼ **Work through the solution**

Integrating the gradient term by parts produces $\delta I = \int (-\epsilon_0 \nabla^2\phi - \rho)\delta\phi\, dV$, because $\delta\phi$ vanishes on the electrodes. Stationarity gives

$$-\epsilon_0 \nabla^2 \phi = \rho.$$

Reversing the source sign would reverse the curvature induced by positive charge. This functional includes a source coupling and generally differs from just the stored field energy.

Contract with a particle velocity

At an instant a particle is at rest. Insert $u^\mu = (c, 0, 0, 0)$ in $f^\mu = qF^{\mu\nu}u_\nu$ to recover its spatial force. Why is the time component zero?

▼ **Work through the solution**

Here $F^{i0} = E_i/c$ and $u_0 = c$, so $f^i = qE_i$. Antisymmetry gives $F^{00} = 0$. At that instant the velocity vanishes and hence so does the rate of work. More generally $u_\mu f^\mu = qu_\mu F^{\mu\nu}u_\nu = 0$, since the velocity product is symmetric. This preserves the particle's invariant rest mass.

Further reading

- Griffiths, 5th ed., §12.3 · relativistic electrodynamics
- Fließbach, chapter 19 · Lagrangian formalism
- Cambridge Electromagnetism, chapter 4 · covariant equations and Maxwell action

Study the full derivation

Relativistic particle dynamics and the field tensor · Waves, radiation and relativity

A circular track measures momentum per charge. A clock following the particle registers proper time, and a detector records how much kinetic energy it has gained. We will put these observations in one dynamical equation, solve that equation in uniform fields and exhibit a numerical pusher that can get the energy right while its orbit phase slowly slips.

Chapter 27 · Cavity eigenmodes, measured linewidth and ring-down

A cavity has both a pitch and a decay time. One is set predominantly by the field pattern and the boundaries. The other measures how much stored energy escapes per cycle. A spatial perturbation can reveal where a particular mode stores its electric energy.

Make a conducting box satisfy every boundary condition

Let the empty cavity occupy $0 < x < a,\ 0 < y < b,\ 0 < z < d$. Its ideal conducting walls require zero tangential electric field. Choose an electric field in the y direction independent of y. This automatically gives

$$\nabla \cdot \mathbf{E} = 0.$$

The field is normal to the two y-walls and must vanish at the four x, z-walls. Write $E_y = X(x)Z(z)e^{-i\omega t}$. The wave equation reduces to

$$X''/X + Z''/Z = -\omega^2/c^2.$$

Sines separately satisfy the vanishing boundary values. Thus a family of normal modes is

$$E_y = E_0 \sin(m\pi x/a)\sin(p\pi z/d)e^{-i\omega t},$$

$$f_{m0p} = \frac{c}{2}\sqrt{(m/a)^2 + (p/d)^2},$$

$$m, p = 1, 2, \ldots.$$

Faraday's law then gives

$$\mathbf{B} = (\nabla \times \mathbf{E})/(i\omega).$$

Its components are proportional to $-\partial_z E_y$ and $+\partial_x E_y$; its normal component vanishes at the appropriate walls. This family is transverse electric relative to the z axis. It does not exhaust the transverse magnetic modes or the modes with y-dependence. Whether a particular solution is excited depends on the overlap of the source with its field pattern.

Track energy between electric and magnetic form

Choose a real mode proportional to $\cos\omega t$. The associated magnetic field is proportional to $\sin\omega t$. At an electric maximum the stored energy is

$$U = \epsilon_0 \int E_{\text{peak}}^2 dV/2.$$

Both squared sines average to $1/2$, giving $U = \epsilon_0 E_0^2 abd/8$. A quarter-cycle later the magnetic field stores the same energy. For a weakly leaking resonator, define $Q = \omega U/P_{\text{lost}}$ using cycle-averaged lost power and slowly varying total energy. At fixed Q,

$$\frac{dU}{dt} = -\frac{\omega}{Q}U,$$

$$U(t) = U_0 e^{-\omega t/Q},$$

$$E_{\mathrm{env}}(t) = E_{\mathrm{env}}(0) e^{-\omega t/(2Q)}.$$

A field probe measuring voltage follows the amplitude envelope. A square-law power detector follows the energy envelope. A ring-down fit delivers $Q = \omega\tau_U = \omega\tau_E/2$, depending on which time constant is reported. Independent weak loss channels have additive powers and therefore additive inverse quality factors.

Connect a time trace to a frequency scan

The Fourier amplitude of a decaying positive-frequency signal is proportional to $\int_0^\infty e^{-\omega_0 t/(2Q)} e^{i(\omega-\omega_0)t} dt = [\omega_0/(2Q) - i(\omega-\omega_0)]^{-1}$. Its squared modulus is a Lorentzian with full width at half power $\Delta\omega = \omega_0/Q$. This derivation resolves a common experimental ambiguity: the two half-power points have opposite detunings, and the width spans both. The approximation assumes a single mode and damping small on the scale of the carrier frequency.

Measuring geometry and locating an antinode

A rectangular resonator

Set $a = 60.0$ mm, $d = 80.0$ mm, $b = 25.0$ mm. The 101 mode gives $f = 3.123$ GHz. At $Q = 6000$ the energy decay time is 306 ns and the field decay time is 612 ns. The full power linewidth is 520 kHz. Increasing b changes the stored energy at fixed peak field but leaves this ideal mode frequency unchanged. Other mode families can move as b changes.

A small dielectric probe

Write the electric eigenproblem as $\nabla \times \mu_0^{-1} \nabla \times \mathbf{E} = \omega^2 \epsilon \mathbf{E}$. Multiply by the complex conjugate field and integrate. The wall term vanishes and the numerator becomes $\int |\nabla \times \mathbf{E}|^2/\mu_0 \, dV$. First-order perturbation of the stationary quotient gives

$$\delta f / f = -\int \delta\epsilon |\mathbf{E}|^2 dV / (2\int \epsilon |\mathbf{E}|^2 dV).$$

A weak permittivity increase of 2% in a region containing 3% of the original electric-field integral shifts the mode down by 3×10^{-4}, approximately 0.94 MHz here. At an electric node the first-order shift vanishes. Large beads distort the field and require a new eigenproblem.

Read a cavity decay with an amplitude or power detector

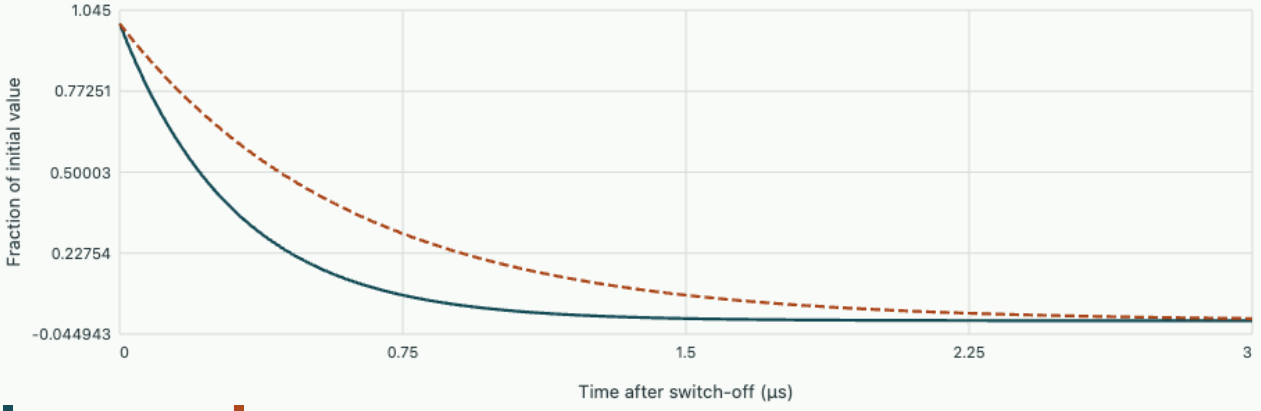
The cavity width is fixed at 60 mm. With 80 mm length and Q = 6000, the energy falls to 1/e in 0.306 µs. Read the field amplitude there and decide which trace a square-law detector would follow.

Open this experiment online

Try the calculation

Loaded quality factor Q **6000**

Cavity length d (mm) **80**



The mode frequency is 3.123 GHz with full power linewidth 520 kHz. The energy e-folding time is 0.306 µs.

Problem sheet · two ways to measure loss

Extract internal loss

A port-coupled resonator has measured $Q_L = 3000$. An independent calibration gives external $Q_e = 12000$. Find internal Q_i and the fraction of power leaving through the port.

▼ Work through the solution

Since $1/Q_L = 1/Q_i + 1/Q_e$, the internal quality factor is 4000. The port power fraction is

$$(1/Q_e)/(1/Q_L) = 1/4.$$

Closing the port in the idealised model leaves $Q = 4000$.

Compare scans and ring-downs

A single mode near 500 MHz has a half-power width of 100 kHz. After drive removal how long until the field amplitude halves?

▼ **Work through the solution**

The linewidth gives $Q = 5000$, so $\tau_E = Q/(\pi f) = 3.18$ µs. The field halves after $\tau_E \ln 2 = 2.21$ µs. At the same instant the energy has fallen to one quarter. This inference requires that frequency jitter has not broadened the measured line.

Further reading

- Griffiths, 5th ed., §9.5 · guided waves
- Fließbach, chapter 21 · cavity modes

Study the full derivation

A guided mode, a cavity line and the energy they carry · Waves, radiation and relativity

A microwave pulse sent through a rectangular guide arrives late even though the guide is evacuated. Terminating the guide converts the travelling pattern into a family of resonances. Use the same transverse eigenproblem to predict its cutoff, pulse delay, cavity frequency and measured linewidth.

Chapter 28 · Why a small transparent sphere scatters light

A particle can redistribute light without consuming it. The calculation starts with an electrostatic boundary-value problem and ends with power crossing a faraway sphere. Connecting the two limits predicts a wavelength dependence that can be tested with colour.

Find the induced dipole by matching two potentials

Take a sphere of radius a and relative permittivity ϵ_r in vacuum. Its size must be small compared with the wavelength everywhere, including inside the particle. Across its diameter, the incident wave then looks almost like a uniform electric field $E_0\hat{\mathbf{z}}$. Spherical symmetry and the applied field single out the angular factor $\cos\theta$. Choose $\phi_{\mathrm{in}} = -Cr\cos\theta$ and $\phi_{\mathrm{out}} = -E_0 r\cos\theta + D\cos\theta/r^2$. Continuity of tangential electric field is equivalent here to continuity of potential. With zero free surface charge, the normal displacement is also continuous. At $r = a$,

$$C = E_0 - D/a^3,$$

$$\epsilon_r C = E_0 + 2D/a^3.$$

Eliminate C to obtain

$$D = a^3 E_0(\epsilon_r - 1)/(\epsilon_r + 2).$$

The coefficient of the dipole potential is $p/(4\pi\epsilon_0)$. Hence $p = \alpha_0 E_0$ with $\alpha_0 = 4\pi\epsilon_0 a^3(\epsilon_r - 1)/(\epsilon_r + 2)$. The internal field $3E_0/(\epsilon_r + 2)$ explains the denominator. For $\epsilon_r = 1$ the disturbance vanishes. In the perfectly polarizable limit the internal field tends to zero.

Match a small dipole to an outgoing wave

Use complex amplitudes with time dependence $e^{-i\omega t}$. The far electric field is $(k^2/4\pi\epsilon_0 r)[(\hat{\mathbf{n}}\times\mathbf{p})\times\hat{\mathbf{n}}]e^{ikr}$, $k = \omega/c$. Squaring its amplitude and multiplying by $\epsilon_0 c/2$ gives the mean intensity. Divide by the incident intensity $\epsilon_0 c|E_0|^2/2$. The differential scattering cross section is then $k^4|\alpha|^2\sin^2\theta/(16\pi^2\epsilon_0^2)$, with θ measured from the dipole axis. Integrating the angular factor gives $8\pi/3$:

$$\sigma_{\mathrm{sca}} = \frac{k^4|\alpha|^2}{6\pi\epsilon_0^2} \simeq \frac{8\pi}{3}k^4 a^6\left|\frac{\epsilon_r - 1}{\epsilon_r + 2}\right|^2.$$

The sixth power originates in the coherent addition of dipoles throughout a volume proportional to a^3. The fourth power in frequency comes from the second time derivative of the dipole in the radiative electric field. A cross section has units of area and measures removed power divided by incident intensity.

Where the scattered energy comes from

The incident field does work at mean rate $\langle \mathbf{E}\cdot\dot{\mathbf{p}}\rangle = (\omega/2)\,\mathrm{Im}(\alpha)|E_0|^2$. Dividing by incident intensity yields

$$\sigma_{\mathrm{ext}} = k\,\mathrm{Im}(\alpha)/\epsilon_0.$$

A strictly real electrostatic α_0 would predict no extracted power while predicting positive scattered power. The static calculation captured the leading dipole amplitude but omitted its small radiation-induced phase delay. For a lossless electric dipole its leading radiative correction satisfies

$$\frac{1}{\alpha} = \frac{1}{\alpha_0} - i\frac{k^3}{6\pi\epsilon_0}.$$

Indeed $\operatorname{Im}\alpha = k^3|\alpha|^2/(6\pi\epsilon_0)$, making extinction equal to scattering. In an absorbing particle their difference is absorbed power per incident intensity. Finite-size dispersion and higher multipoles require additional terms once ka grows.

Size, colour and the limits of inference

A particle doubles in radius

For dilute spheres of the same weakly dispersive material, doubling a gives $2^6 = 64$ times the scattered power per particle at fixed incident intensity. If the total volume of dielectric is held fixed, the number of spheres falls by eight. The whole sample then scatters eight times the power, provided phases between different particles are effectively uncorrelated and multiple scattering is negligible. The two questions describe different samples.

A frequency-independent contrast

For $a = 20$ nm, $\epsilon_r = 2.25$ and vacuum wavelength 600 nm, $ka = 0.209$ and $(\epsilon_r - 1)/(\epsilon_r + 2) = 0.2941$. The leading cross section is 5.58×10^{-19} m², around 4.4×10^{-4} of the geometric area. At 450 nm the model predicts 3.16 times that cross section. A large discrepancy can indicate dispersion, additional multipoles, or a distribution of particle sizes.

Compare the scattering of several particle sizes

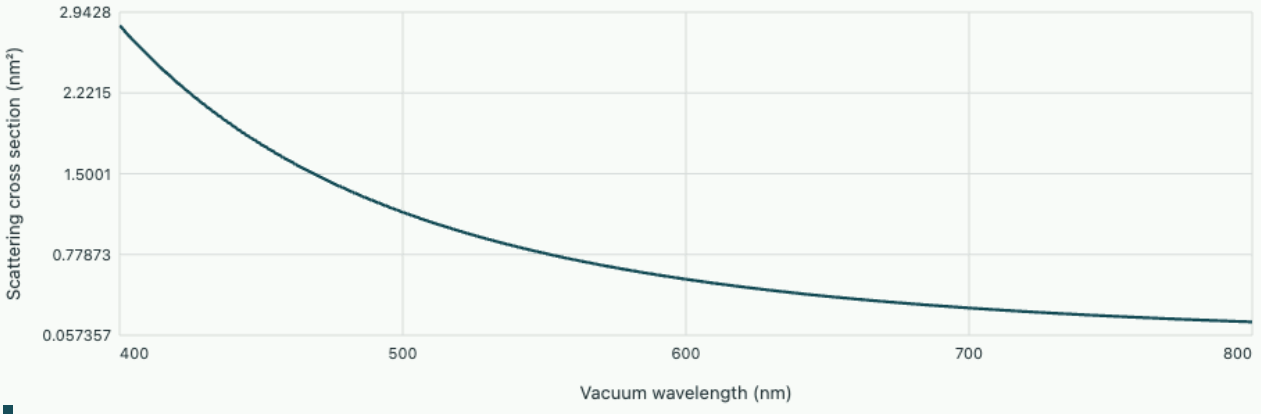
For radius 20 nm, relative permittivity 2.25 and wavelength 600 nm, the predicted cross section is about 0.56 nm². Halve the radius and estimate the drop before reading the plot. The curve uses a frequency-independent dielectric contrast.

Open this experiment online

Try the calculation

Sphere radius (nm) **20**

Relative permittivity **2.25**



Rayleigh sphere

The leading cross section at 600 nm is 0.558 nm². The shortest plotted wavelength gives ka = 0.314; finite-size corrections grow as this ratio grows.

Problem sheet · inverse measurements

Estimate a radius ratio

Two dilute suspensions have the same number density, composition and optical path length. In the single-scattering regime one produces twice the scattered intensity at the same angle. Infer the ratio of particle radii.

▼ Work through the solution

The proportionality to a^6 gives

$$a_2/a_1 = 2^{1/6} = 1.122.$$

A 12.2% change in radius doubles the signal. If counts or material contrasts are unknown, the same intensity change has multiple explanations.

Recover Thomson scattering

A free electron is driven by a weak field. Use $m\ddot{x} = -eE_0 e^{-i\omega t}$ to find its dipole and cross section. Why does the wavelength law change?

▼ Work through the solution

The displacement amplitude is $eE_0/(m\omega^2)$. Multiplying by electronic charge gives

$$\alpha = -e^2/(m\omega^2).$$

The ω^{-4} in $|\alpha|^2$ cancels k^4. The cross section becomes $(8\pi/3)[e^2/(4\pi\epsilon_0 mc^2)]^2$. This classical approximation assumes nonrelativistic electron motion and negligible recoil.

Further reading

- Fließbach, chapter 25 · scattering of light
- Cambridge Electromagnetism, chapter 5 · Rayleigh and Thomson scattering

Study the full derivation

Scattering, extinction and the phase of the forward wave · Waves, radiation and relativity

A suspension of transparent particles dims a narrow transmitted beam while absorbing almost none of its energy. Derive the scattered angular power, identify the interference term seen in transmission, and use energy conservation to put a bound on the response of a single small particle.

Chapter 29 · When a circuit element must be treated as a field

Can a short pair of conductors be called a capacitor? At low frequencies it stores energy almost electrostatically. As propagation time becomes important, the same open conductors support standing waves and their input current no longer grows proportionally to frequency.

Build a transmission line from small field volumes

Consider a uniform two-conductor line supporting a transverse electromagnetic wave. Let $V(z,t)$ measure the transverse potential difference and $I(z,t)$ the current toward $+z$. Denote its inductance and capacitance per unit length by L' and C'. For the moment assume ideal conductors, a linear lossless dielectric and negligible radiation. A segment dz stores charge $C'V\,dz$. Charge balance gives

$$I(z) - I(z+dz) = C' dz\, \partial_t V.$$

Its linked magnetic flux is $L' I\,dz$; Faraday's law gives the voltage drop $V(z) - V(z+dz) = L' dz\, \partial_t I$. Division by dz yields

$$\partial_z I = -C' \partial_t V,$$
$$\partial_z V = -L' \partial_t I.$$

Differentiate the second relation with respect to z and substitute the first. The result is

$$\partial_z^2 V = L'C' \partial_t^2 V,$$

with speed $u = 1/\sqrt{L'C'}$. An outgoing wave $F(t - z/u)$ satisfies

$$V/I = Z_0 = \sqrt{L'/C'}.$$

The electric and magnetic field energies per length are $C'V^2/2$ and $L'I^2/2$, and they are equal for this travelling wave. Multiplying the two equations by V and I also gives a one-dimensional conservation law:

$$\partial_z (VI) = -\partial_t (C'V^2/2 + L'I^2/2).$$

Derive the input admittance of an open end

Use $e^{-i\omega t}$ time dependence, and let the far end $z = \ell$ be open so $I(\ell) = 0$. Then $\partial_z V(\ell) = 0$. The solution is

$$V(z) = V_L \cos k(\ell - z),$$

where $k = \omega/u$. Since $\partial_z V = i\omega L' I$, the current is $-iV_L \sin k(\ell - z)/Z_0$. At the input,

$$Y_{\text{in}} = -\frac{i}{Z_0}\tan(k\ell),$$

$$Y_{\text{lumped}} = -i\omega C'\ell,$$

$$\frac{Y_{\text{in}}}{Y_{\text{lumped}}} = \frac{\tan x}{x},$$

$$x = \frac{\omega \ell}{u}.$$

At small x,

$$\tan x / x = 1 + x^2/3 + O(x^4).$$

Thus a circuit approximation can be assigned an error estimate. Near $x = \pi/2$ the input voltage approaches a node and the ideal input admittance diverges. Losses round the pole in a real line. Beyond this frequency even the sign can change, although the far end remains open.

Two decisions about model size

A 15 cm open lead

Take $u = 2.00 \times 10^8$ m/s and $Z_0 = 50.0\,\Omega$. Then $C' = 1/(Z_0 u) = 100$ pF/m and the lumped capacitance is 15.0 pF. At 20.0 MHz, $x = 0.09425$; the exact admittance exceeds the capacitive prediction by 0.297%. At 200 MHz, $x = 0.9425$, and the excess is about 46.0%. A component measured on an open fixture can therefore appear to acquire substantial extra capacitance at high frequency.

A time trace on the same lead

The one-way flight time is 0.75 ns. A step launched down the line first sees the characteristic impedance. Information about the open end returns only after 1.50 ns. At an open end an incident positive voltage wave is reflected with the same sign. The current waves cancel there, enforcing zero terminal current, while the terminal voltage doubles at their first overlap. For a source matched to Z_0, the returning wave is absorbed at the source. This time-domain argument provides a check on the origin of the standing-wave solution.

A second circuit limit from magnetic energy

In the slow-variation limit, linked flux in two fixed coils is linear in both currents. For a reciprocal medium,

$$\lambda_1 = L_1 I_1 + M I_2,$$

$$\lambda_2 = M I_1 + L_2 I_2.$$

Integrate the supplied power $I_1\dot{\lambda}_1 + I_2\dot{\lambda}_2$ to obtain

$$U = (L_1 I_1^2 + 2M I_1 I_2 + L_2 I_2^2)/2.$$

Completing the square gives

$$2U = L_1 (I_1 + M I_2 / L_1)^2 + (L_2 - M^2 / L_1) I_2^2.$$

A passive magnetic energy requires

$$|M| \le \sqrt{L_1 L_2}.$$

The sign of M depends on terminal orientations. This bound is an independent way to check a measured or numerically inferred circuit model.

Measure the range of a lumped capacitor model

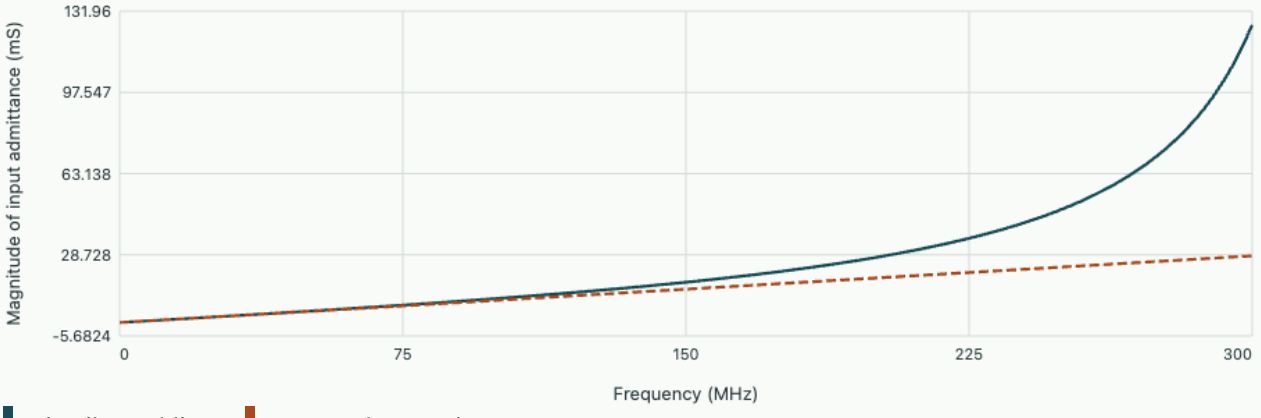
The line has speed 2.00 × 10⁸ m/s and impedance 50 Ω. For a 15 cm open section its first quarter-wave pole is 333 MHz. At 60% of that frequency the distributed admittance exceeds the lumped value by 46%. Move the open end and compare the two curves.

Open this experiment online

Try the calculation

Open line length (cm) **15**

Test frequency / first pole **0.6**



At 200 MHz distributed admittance exceeds lumped admittance by 46%. The first pole is at 333 MHz.

Problem sheet · estimating the cost of a circuit approximation

Set a one-percent frequency limit

For a 0.30 m open line with speed 1.8×10^{8} m/s, estimate the highest frequency for a 1% correction in input admittance. State the approximation used.

▼ **Work through the solution**

Set $x^2/3 = 0.01$, giving $x \simeq 0.173$ and $f = xu/(2\pi\ell) = 16.5$ MHz. The next term is $2x^4/15 \simeq 0.00012$, so the estimate slightly overstates the allowable frequency. Bisection on $\tan x/x = 1.01$ gives $x \simeq 0.1722$.

Check a claimed mutual inductance

A two-coil model has $L_1 = 4$ mH, $L_2 = 9$ mH and $M = 7$ mH. Find the current combination that reveals a problem.

▼ **Work through the solution**

Choose $I_1 = -MI_2/L_1 = -1.75I_2$. The surviving coefficient is $9 - 49/4 = -3.25$ mH, giving a negative stored energy for every nonzero I_2. A passive reciprocal model with these self-inductances can support at most 6 mH of mutual inductance in magnitude.

Further reading

- Fließbach, chapter 26 · oscillating circuits
- Griffiths, 5th ed., §§7.2, 9.5 · inductance and guided waves
- Cambridge Electromagnetism, chapter 3 · induction and inductance

Study the full derivation

Field momentum, momentum exchange and the quasistatic limits · Magnetism, induction and field dynamics

A system can have static electric and magnetic fields with nonzero local electromagnetic momentum density. A circuit can also be electrically small while either induction or displacement current remains indispensable. Both puzzles become clearer when one follows the relevant conservation law and compares the scales of the discarded terms.

Chapter 30 · From wavefronts to rays, focusing and caustics

A ray is a prediction extracted from the phase of a wave. Following its amplitude at the same time reveals the point at which a ray calculation loses its validity. A gently graded transparent material gives a solvable example of both effects.

Derive the phase and transport equations together

In a lossless isotropic dielectric with a slowly varying refractive index $n(\mathbf{r})$, consider one scalar polarisation component satisfying $\nabla^2\psi + k_0^2 n^2 \psi = 0$, where $k_0 = \omega/c$. This scalar description works when vector-polarisation coupling is negligible over the scales of interest. Set $\psi = A(\mathbf{r})e^{ik_0 S(\mathbf{r})}$ with a slowly varying real amplitude. Differentiation twice gives

$$0 = \nabla^2 A + ik_0[2\nabla A \cdot \nabla S + A\nabla^2 S] + k_0^2 A(n^2 - |\nabla S|^2).$$

If the material and amplitude vary little over a wavelength, first equate the leading k_0^2 term to zero. At the next order the k_0 term vanishes. The two results are the eikonal equation and transport equation:

$$|\nabla S|^2 = n^2,$$

$$\nabla \cdot (A^2 \nabla S) = 0.$$

The first fixes the optical length accumulated between neighbouring wavefronts. The second transports energy down a thin ray tube. Across a tube of area $\mathcal{A}$ perpendicular to the phase gradient, $A^2 n \mathcal{A}$ is constant to this order. In a uniform medium a spherical tube grows as r^2, giving $A \propto 1/r$, consistent with the outgoing wave.

Follow the normal to a wavefront

Choose arclength s along the ray and $\hat{\mathbf{t}} = \nabla S/n$. In the eikonal equation take a spatial derivative. Since ∇S is a gradient,

$$(\nabla S \cdot \nabla)\nabla S = \nabla(|\nabla S|^2)/2 = n\nabla n.$$

Divide by n. The ray equation is

$$\frac{d}{ds}(n\hat{\mathbf{t}}) = \nabla n.$$

A transverse increase of index bends the ray toward larger optical index. If n depends only on the coordinate normal to layers, the component of $n\hat{\mathbf{t}}$ parallel to them stays constant. Thus $n\sin\theta$ is conserved for an angle measured from the normal, recovering Snell's law. The same equation follows by requiring stationary optical length $\int n\,ds$ when varying the path with fixed endpoints.

Solve a graded-index profile

For paraxial rays travelling close to $+z$, take $n(y) \simeq n_0(1 - g^2 y^2/2)$. Assume $|gy| \ll 1$ and $|dy/dz| \ll 1$. Then $ds \simeq dz$, $\hat{t}_y \simeq y'$, and $d(ny')/dz \simeq n_0 y''$ to leading order. Since $\partial_y n = -n_0 g^2 y$, the trajectory solves $y'' + g^2 y = 0$. Initial height y_0 and slope s_0 give

$$y(z) = y_0 \cos(gz) + \frac{s_0}{g}\sin(gz),$$

$$y'(z) = -g y_0 \sin(gz) + s_0 \cos(gz).$$

For a collimated bundle, all $s_0 = 0$. The transverse magnification $\partial y/\partial y_0 = \cos(gz)$ vanishes simultaneously for all rays at $z_f = \pi/(2g)$. This is a caustic of the ray description. In this planar example the ray density grows as $1/|\cos(gz)|$. Its divergence signals that the neglected $\nabla^2 A$ becomes important. The underlying wave stays finite and develops diffraction near the focus.

A focus and a reconstructed gradient

Refocus an off-axis bundle

Take $g = 20\ \mathrm{m}^{-1}$, initial heights between -1 and $+1$ mm and zero slopes. The predicted first focus is 78.5 mm downstream. The extreme rays acquire slopes of magnitude 0.020 at the focus, well within the paraxial regime. Tilt every initial ray by 0.010 rad. At z_f the common height shifts to $s_0/g = 0.50$ mm. The focal length is insensitive to this common tilt in the parabolic approximation.

Estimate the index from a photograph

Two successive crossings of initially parallel symmetric rays are 120 mm apart. Crossings occur at $gz = \pi/2 + m\pi$, so

$$g = \pi/0.120 = 26.2\ \mathrm{m}^{-1}.$$

At a height of 2 mm, the predicted fractional index decrease from the central index is

$$g^2 y^2/2 = 1.37 \times 10^{-3}.$$

Resolving this small contrast through repeated bending can be easier than resolving it through a single refraction measurement.

Focus and tilt a bundle of paraxial rays

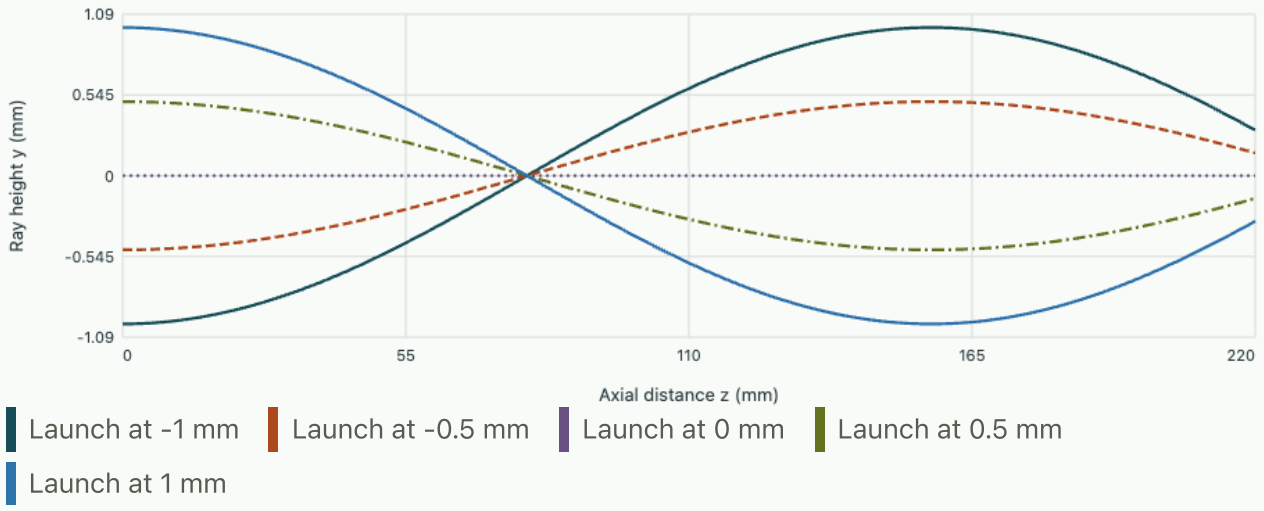
At gradient 20 m⁻¹, five parallel rays meet after 78.5 mm. A common tilt of 10 mrad displaces their meeting point by 0.50 mm. Follow the bundle through a crossing and decide why the geometric density ceases to predict a finite intensity there.

Open this experiment online

Try the calculation

Index gradient g (1/m) **20**

Common input slope (mrad) **0**



The first crossing occurs at z = 78.5 mm and y = 0 mm. Finite-wavelength diffraction limits the intensity at that crossing.

Problem sheet · optical path and diffraction

A travel-time check

A ray passes through two homogeneous layers of thicknesses a, b, indices n_1, n_2, with lateral endpoint separation D. Write travel time as a function of crossing coordinate x. Recover the interface condition.

▼ **Work through the solution**

The time is

$$T(x) = [n_1\sqrt{a^2 + x^2} + n_2\sqrt{b^2 + (D - x)^2}]/c.$$

The derivative vanishes when $n_1 x/\sqrt{a^2 + x^2} = n_2(D - x)/\sqrt{b^2 + (D - x)^2}$. The two fractions are the sines of the angles to the interface normal. Each segment's length has positive second derivative, making this stationary path a local minimum in this geometry.

Estimate the diffraction scale

A circular pupil of numerical aperture 0.030 focuses vacuum light of wavelength 600 nm. Estimate the first dark-ring radius and compare it with a ray focus.

▼ **Work through the solution**

For a uniformly illuminated circular aperture in the paraxial limit the first Airy minimum is $0.61\lambda_0/\mathrm{NA}$, giving 12.2 µm. Different pupil shapes change the numerical factor. The wavelength dependence is absent from the ray crossing, so a ray trace alone cannot distinguish two experiments with identical geometry but different wavelengths.

Further reading

- Fließbach, chapter 38 · geometrical optics
- Fließbach, chapter 36 · interference and diffraction

Study the full derivation

Diffraction from the near field to the focal plane · Waves, radiation and relativity

An illuminated aperture is a boundary condition for a wave. A screen a millimetre away, an intensity camera at a lens focus and a microscope measuring an evanescent field sample different consequences of the same condition. We will derive operators for propagating the field, estimate the approximations each instrument makes and use energy and sampling checks to tell numerical structure from a measurable fringe.

Chapter 31 · Radiation reaction and the energy stored near a charge

A charge can radiate at an instant when a proposed frictional force would do no work. To account for the energy one needs three measurements: mechanical work on the charge, power leaving in radiation and a reversible change in its nearby field.

A moving oscillator tests a friction hypothesis

For a point charge moving nonrelativistically in SI units define $\tau = q^2/(6\pi\epsilon_0 m c^3)$. The Larmor formula gives

$$P_{\mathrm{rad}} = m\tau a^2.$$

Suppose a charge reaches a turning point of a slow sinusoidal oscillation. Its velocity there is zero and its acceleration is maximal. Radiation carries positive power away, yet any finite force on the particle does zero mechanical work at that instant. Equating instantaneous mechanical loss to instantaneous radiated power cannot account for the turning point.

Integrate over a smooth acceleration history. The product rule gives

$$d(\mathbf{a}\cdot\mathbf{v})/dt = \dot{\mathbf{a}}\cdot\mathbf{v} + a^2.$$

Rearranging shows that a local force $\mathbf{F}_{\mathrm{self}} = m\tau\dot{\mathbf{a}}$ can recover the Larmor loss between endpoints where $\mathbf{a}\cdot\mathbf{v}$ agrees:

$$\int_{t_1}^{t_2} \mathbf{F}_{\mathrm{self}}\cdot\mathbf{v}\,dt = m\tau[\mathbf{a}\cdot\mathbf{v}]_{t_1}^{t_2} - \int_{t_1}^{t_2} P_{\mathrm{rad}}\,dt.$$

This identity motivates the nonrelativistic Abraham–Lorentz term. An integrated energy argument only fixes its work under the chosen endpoint conditions. Deriving the full local self-force requires analysing the near field, regularising the point-source singularity and identifying the measured particle mass. These assumptions delimit the classical effective equation.

Follow the energy that can return to the particle

Let $K = mv^2/2$, $E_S = -m\tau\,\mathbf{a}\cdot\mathbf{v}$. Write $m\mathbf{a} = \mathbf{F}_{\mathrm{ext}} + \mathbf{F}_{\mathrm{self}}$, dot with $\mathbf{v}$, and use the product rule once more. The result is

$$\mathbf{F}_{\mathrm{ext}}\cdot\mathbf{v} = \frac{d}{dt}(K + E_S) + P_{\mathrm{rad}}.$$

The acceleration-dependent quantity E_S, called the Schott energy, represents a reversible bound-field term after regularisation. Its sign depends on motion; it is a changeable contribution relative to an energy reference. For constant acceleration over a short nonrelativistic interval, the self-force vanishes, but $dE_S/dt = -m\tau a^2$. The decrease supplies the radiated energy. Switching the acceleration on and off

modifies the budget near the endpoints. Extending a constant-acceleration segment to infinite time would break the assumed slow-speed model.

Why taking the third-order equation literally fails

Take a constant mass and one Cartesian direction. With no external force the equation becomes $a - \tau\dot{a} = 0$, whose general solution includes $a = Ae^{t/\tau}$. Standard position and velocity data leave A undetermined. This runaway is incompatible with a passive free particle. The prescription $a(t) \to 0$ after an applied force has ended removes that branch, but integrating backward gives

$$a(t) = \frac{1}{m\tau}\int_t^\infty e^{-(s-t)/\tau} F_{\mathrm{ext}}(s)\, ds.$$

The integral samples future forces. A discontinuous idealised force can give apparent preacceleration over order τ. For an electron $\tau \simeq 6.27 \times 10^{-24}$ s. On such scales an elementary classical point-charge description is inadequate. For smooth forcing varying over $T \gg \tau$, instead use the leading equation $a = F_{\mathrm{ext}}/m + O(\tau/T)$ on the right. To first order in τ/T,

$$ma = F_{\mathrm{ext}} + \tau\,\frac{dF_{\mathrm{ext}}}{dt}.$$

For a force depending on position and velocity, evaluate its total derivative along the leading-order motion. This reduced-order equation selects the smooth, perturbative dynamics without adding independent initial acceleration data.

A spectral width and a phase-resolved budget

Radiative damping of a weakly bound oscillator

Let $F_{\mathrm{ext}} = -m\omega_0^2 x$. Then $\ddot{x} + \tau\omega_0^2\dot{x} + \omega_0^2 x = 0$ to first order. The amplitude falls with e-folding time $2/(\tau\omega_0^2)$. At $\omega_0 = 3.0 \times 10^{15}$ s⁻¹ the formal electron model gives 35.4 ns, with $\tau\omega_0 = 1.88 \times 10^{-8}$. The weak-damping approximation is internally consistent. Actual atomic linewidths additionally depend on quantum matrix elements and other decay channels. The calculation describes a classical bound oscillator with a chosen restoring force.

The same energy viewed at two phases

Impose $v = v_a \sin\omega t$, so $a = \omega v_a \cos\omega t$ and $\dot{a} = -\omega^2 v_a \sin\omega t$. In units $P_0 = m\tau\omega^2 v_a^2$, radiation power is $\cos^2\omega t$ while mechanical self-force power is $-\sin^2\omega t$. At phase zero all the instantaneous radiation is supplied by decreasing bound-field energy. A quarter period later the radiation vanishes at the corresponding retarded emission phase, but the self-force still removes mechanical energy. Over one period the integrated bound-field term vanishes, so the mean mechanical loss is $P_0/2$.

Close the energy accounts at a chosen emission phase

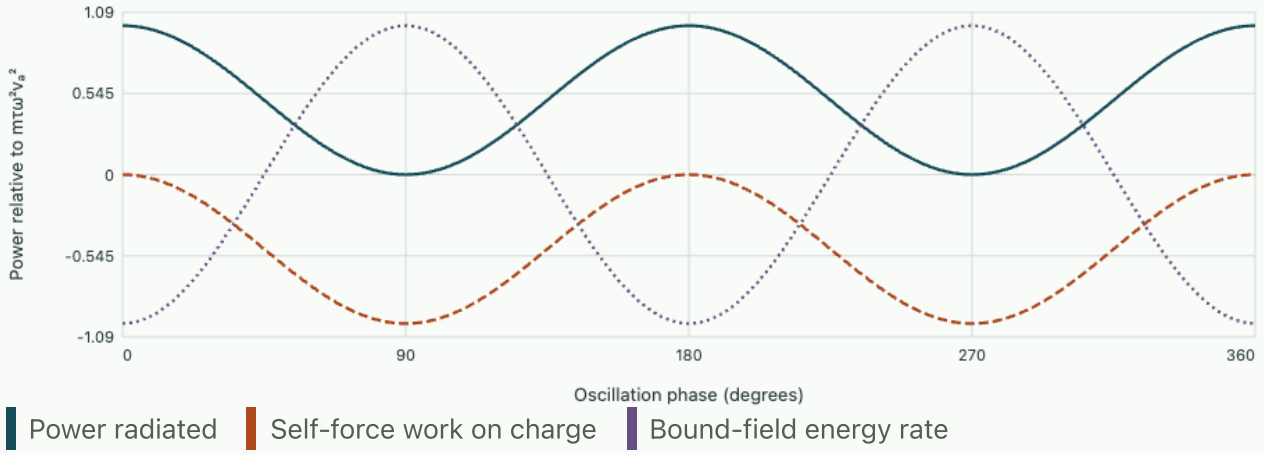
The radiated power, the self-force work on the particle and the rate of change of bound-field energy sum to zero at every emission phase. With zero mean speed and phase zero, their dimensionless values are +1, 0 and –1. Shift the mean speed and decide which of the three readings must stay the same.

Open this experiment online

Try the calculation

Mean velocity / oscillating speed **0**

Emission phase (degrees) **0**



At 0°, radiation, self-force work and bound-field rate are 1, 0, -1 in the plotted units. Their sum is zero.

Problem sheet · how local is radiative friction?

Calculate a complete pulse budget

Take $a(t) = a_0 \sin(\pi t/T)$ for $0 < t < T$, zero elsewhere, with initial velocity zero. Calculate final speed, total radiated energy and the ratio of radiation to final kinetic energy.

▼ **Work through the solution**

The velocity gain is $2a_0T/\pi$. Integrating a^2 yields

$$E_{\mathrm{rad}} = m\tau a_0^2 T/2.$$

The final kinetic energy is $2ma_0^2T^2/\pi^2$, so the ratio is $\pi^2\tau/(4T)$. Both endpoints have zero acceleration. The time-integrated external work must therefore equal the sum of the final kinetic energy and radiated energy. Choose $a_0T \ll c$ and $T \gg \tau$ for this calculation.

Diagnose a spurious trajectory

A computation with no force doubles its acceleration every few multiples of the electron radiation-reaction time. Would closer timesteps necessarily remove the instability?

▼ **Work through the solution**

The exact homogeneous third-order equation already has $a = Ae^{t/\tau}$. A smaller step can track that unphysical branch more accurately. A model restricted to smooth weak radiation reaction instead evolves the reduced second-order equation. In free motion this gives $a = 0$ identically. Compare physical timescales and equations before interpreting a converged numerical trajectory as an observable.

Further reading

- Griffiths, 5th ed., §§11.2–11.3 · radiated power and radiation reaction

Study the full derivation

Radiated power, mechanical work and energy limits · Waves, radiation and relativity

A charge can radiate while its speed is constant. A source can exchange energy with its nearby field before a pulse reaches a receiver. We will use the flux through a distant sphere, the work at the charge and the energy stored in a local field as three operationally distinct measurements. Agreement requires the relevant integration times and the range of validity of the particle model.

Chapter 32 · Diffraction from an aperture of finite width

Geometrical optics predicts a sharp edge and an infinitely narrow ideal focus. A scalar diffraction integral describes how contributions with different phase interfere. A slit gives an exact far-field reference against which to test an oscillatory numerical integral.

Accumulate phase across an aperture

Let coherent monochromatic light pass through a slit of full width a in the plane $z = 0$. The slit is long in its other direction and illuminated by a normally incident uniform plane wave. Work with a scalar component in free space. Suppose observation distance z is large compared with a, $|x|$ and the wavelength λ. Contributions from aperture point ξ to detector position x carry a phase kR, $k = 2\pi/\lambda$. The path expansion is

$$R = \sqrt{z^2 + (x-\xi)^2} = z + \frac{(x-\xi)^2}{2z} - \frac{(x-\xi)^4}{8z^3} + \cdots .$$

Approximating slow amplitude and obliquity factors by their axial values while retaining the quadratic phase gives the Fresnel integral. Check the dropped phase $k|x-\xi|^4/(8z^3) \ll 1$ over both the aperture and detector range. Its smallness is more restrictive than simply calling the angle small in a high-frequency problem. The scalar aperture model also assumes that edge polarisation and finite screen thickness are negligible.

Reduce the integral to two dimensionless variables

Set $s = \xi/a$, $u = ax/(\lambda z)$ and $N = a^2/(\lambda z)$. Expanding the square splits off a common phase depending only on x, z. Normalise away that phase and the common geometrical amplitude, leaving the shape integral

$$F(u, N) = \int_{-1/2}^{1/2} \exp[i\pi N s^2 - 2\pi i u s]\, ds,$$

$$I_{\mathrm{rel}} = |F|^2.$$

Here a denotes the full slit width. Definitions using the half-width produce Fresnel numbers four times smaller. The maximum phase of the s^2 term is $\pi N/4$. This sets a quantitative criterion for dropping it. In the far-field limit $N \to 0$, direct integration of the remaining exponential gives $F(u, 0) = \sin(\pi u)/(\pi u)$ and $F(0, 0) = 1$ by continuity. The first zeros occur at $u = \pm 1$. Each slit half has an opposite-phase partner at this observation angle.

Discretise complex amplitude before forming intensity

Divide the slit into M equal cells. For their midpoint coordinates $s_j = (j + 1/2)/M - 1/2$, evaluate $\phi_j = \pi N s_j^2 - 2\pi u s_j$ and accumulate the real and imaginary parts:

$$F_M = \frac{1}{M}\sum_{j=0}^{M-1}(\cos\phi_j + i\sin\phi_j),$$

$$I_M = (\operatorname{Re} F_M)^2 + (\operatorname{Im} F_M)^2.$$

Summing $|e^{i\phi_j}|^2$ first gives a flat result and loses all interference. An estimate for the largest phase step is $\pi(N + 2|u|)/M$. Reduce it while checking convergence at specified points. For a smooth integrand, composite midpoint quadrature has error of order M^{-2} in complex amplitude at fixed u, N. Relative error in intensity near a zero can be arbitrarily large. There compare absolute intensities or complex amplitudes. Agreement between a coarse grid and the far-field formula at large N is not a valid convergence test.

Design a detector scan and identify a false minimum

Choose a distance that resolves the first minimum

A slit has full width $a = 100$ µm in light of $\lambda = 500$ nm. At $z = 1.0$ m the Fresnel number is 0.020 and the dropped aperture phase is at most 0.0157 rad. The first far-field minimum lies at $x = \lambda z/a = 5.0$ mm. At $z = 20$ mm the same slit gives $N = 1$. Dropping 0.785 rad at the aperture edge can materially change the measured distribution. Moving the detector closer also changes the physical width associated with one unit of u.

How much light fills the nominal dark point?

At $u = 1$ the Fraunhofer amplitude vanishes. For small N, expand the chirp to first order. The remaining integral is $i\pi N\int_{-1/2}^{1/2} s^2 e^{-2\pi i s}ds$. The imaginary odd integrand vanishes; two integrations by parts give the real integral $-1/(2\pi^2)$. Hence $F(1, N) \simeq -iN/(2\pi)$ and $I_{\mathrm{rel}}(1, N) \simeq N^2/(4\pi^2)$. At $N = 0.10$ the predicted residual intensity is 2.53×10^{-4}. A correctly converged computation can have a nonzero value at a nominal far-field dark point.

Integrate the slit and refine the aperture mesh

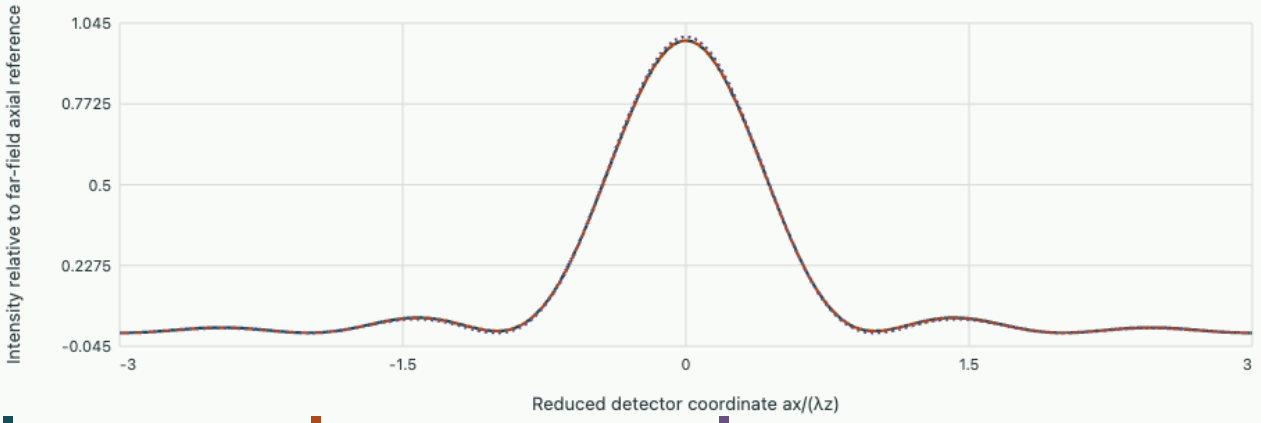
Here the slit width is the entire opening. At Fresnel number 0.10 the first far-field dark point has residual intensity about 0.000253 on the converged scale. Change the mesh and the Fresnel number independently to distinguish numerical error from a physical filling of the minimum.

Open this experiment online

Try the calculation

Fresnel number (full slit width) **0.5**

Midpoint quadrature cells **16**



Selected quadrature | Refined, twice as many cells | Fraunhofer limit

At u = 1 the two quadratures predict intensities 0.006203 and 0.006262. A persistent nonzero value for finite Fresnel number measures near-field interference.

Problem sheet · reconstruction and convergence

Use two measured minima

For wavelength 630 nm and distance 1.20 m, the separation of the first minima is 9.0 mm. Recover the slit width and check the implied far-field assumption.

▼ **Work through the solution**

The distance from the centre to either first zero is 4.5 mm. Thus $a = \lambda z / x = 168$ µm. Then $N = a^2/(\lambda z) = 0.0373$, and the maximum neglected quadratic phase is 0.0293 rad. With the stated precision the use of the far-field formula is plausible. The relevant length in the inverse calculation was half the measured separation.

Choose and validate a quadrature

For $|u| \le 3$ and $N = 4$, require the estimated phase change between adjacent source nodes to remain below 0.25 rad. What grid size is suggested? How would you estimate the remaining error?

▼ **Work through the solution**

The bound gives $M > \pi(4+6)/0.25 \simeq 126$, so 128 cells meet the criterion. Compare 128 and 256 cells at the centre and near the first fringes. In the second-order convergence regime the fine-grid amplitude

error is approximately one third of the coarse-to-fine difference. Verify the trend with 512 cells if high accuracy is needed. Refining the detector sampling alone cannot fix inadequate sampling of the slit.

Further reading

- Fließbach, chapter 35 · Huygens principle
- Fließbach, chapter 36 · interference and diffraction

Study the full derivation

Diffraction from the near field to the focal plane · Waves, radiation and relativity

An illuminated aperture is a boundary condition for a wave. A screen a millimetre away, an intensity camera at a lens focus and a microscope measuring an evanescent field sample different consequences of the same condition. We will derive operators for propagating the field, estimate the approximations each instrument makes and use energy and sampling checks to tell numerical structure from a measurable fringe.

Chapter 33 · Energy, momentum and a charged particle's worldline

A field may turn the momentum, change its magnitude, or do both. Beginning with a measurable clock carried by the particle gives a force law that remains well behaved even when increasing momentum produces little visible change in speed.

Use ticks of a comoving clock

A massive particle has a timelike worldline. In a laboratory with spacetime coordinates $x^\mu = (ct, \mathbf{x})$, the invariant interval along its path gives $c^2 d\tau^2 = c^2 dt^2 - d\mathbf{x}^2$, so $d\tau = dt/\gamma$,

$$\gamma = (1 - v^2/c^2)^{-1/2}.$$

Its four-velocity $u^\mu = dx^\mu/d\tau = \gamma(c, \mathbf{v})$ has fixed squared norm c^2. Multiplication by the measured rest mass gives four-momentum $p^\mu = (\mathcal{E}/c, \mathbf{p})$, with $\mathcal{E} = \gamma m c^2$,

$$\mathbf{p} = \gamma m \mathbf{v}.$$

Squaring and differentiating at fixed rest mass gives

$$\mathcal{E}^2 = m^2 c^4 + p^2 c^2,$$

$$d\mathcal{E} = \frac{c^2 \mathbf{p}}{\mathcal{E}} \cdot d\mathbf{p} = \mathbf{v} \cdot d\mathbf{p}.$$

Use $\dot{\mathbf{p}} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$. Taking its scalar product with $\mathbf{v}$ shows $d\mathcal{E}/dt = q\mathbf{E} \cdot \mathbf{v}$. A uniform magnetic field changes momentum direction while conserving its magnitude and energy. Resolving a general force along and across the velocity gives

$$F_\parallel = \gamma^3 m a_\parallel,$$

$$F_\perp = \gamma m a_\perp.$$

These factors follow from differentiating $\gamma\mathbf{v}$, without assigning the particle a changing rest mass.

Integrate motion in a constant electric field

Choose a positive charge initially at rest with E along $+x$, and neglect other forces. A negative charge follows the reflected trajectory. Define $s = qEt/(mc)$. Lab momentum grows linearly, $p = qEt$. Substitute it into the energy relation and use $v = c^2 p/\mathcal{E}$. Integration with $x(0) = 0$ gives

$$\gamma = \sqrt{1 + s^2},$$

$$\frac{v}{c} = \frac{s}{\sqrt{1 + s^2}},$$

$$x = \frac{mc^2}{qE}\left(\sqrt{1 + s^2} - 1\right).$$

Multiplying the distance by qE recovers kinetic energy $\mathcal{E} - mc^2$. A second check comes from the clock. Integrating $d\tau = dt/\sqrt{1+s^2}$ gives

$$\tau = (mc/qE)\,\operatorname{arsinh} s.$$

With rapidity $\eta = qE\tau/(mc)$,

$$s = \sinh\eta,\ \gamma = \cosh\eta.$$

The comoving proper acceleration is qE/m even while the laboratory acceleration decreases as γ^{-3}. Expanding at early times yields

$$v = qEt/m + O(t^3),$$
$$x = qEt^2/(2m) + O(t^4).$$

Read an energy from the curvature of a track

In a uniform magnetic field along z, longitudinal momentum and energy remain fixed. For transverse motion the change in momentum direction has magnitude $p_\perp v_\perp/r$. Equating it with $|q|v_\perp B$ gives

$$r = p_\perp/(|q|B).$$

The angular frequency is $\omega_c = |q|B/(\gamma m)$, and the longitudinal distance travelled per turn is $2\pi p_\parallel/(|q|B)$. The sign of the charge determines the sense of rotation for a specified field direction. Radius alone measures transverse momentum. To reconstruct total momentum from a helical path one must also measure the pitch.

Reconstruct two ideal laboratory observations

When does the classical speed fail?

For an electron starting at rest in a uniform field of magnitude 100 kV/m, the momentum magnitude reaches mc after $mc/(eE) = 17.0$ ns. A Newtonian extrapolation then predicts speed c. The relativistic result is $c/\sqrt{2} = 0.707c$, kinetic energy $(\sqrt{2}-1)mc^2 = 212$ keV and travel distance 2.12 m. The potential difference over this distance is 212 kV, giving the same kinetic energy independently. These are ideal constant-field values for a prescribed test particle.

A helical track with a missing component

In $B = 0.20$ T a singly charged particle has circular projection of radius 5.0 mm and advances 20 mm along the field per revolution. Using elementary charge e, one obtains $p_\perp = 1.60 \times 10^{-22}$ kg m/s and $p_\parallel = 1.02 \times 10^{-22}$ kg m/s. For an electron $p/(mc) = 0.696$, $\gamma \simeq 1.22$, and the kinetic energy is 112 keV. Assuming the track lay entirely in a transverse plane would underestimate its kinetic energy. A heavier singly charged particle with the same track has the same momentum but a different kinetic energy.

Compare a particle clock with a lab timer

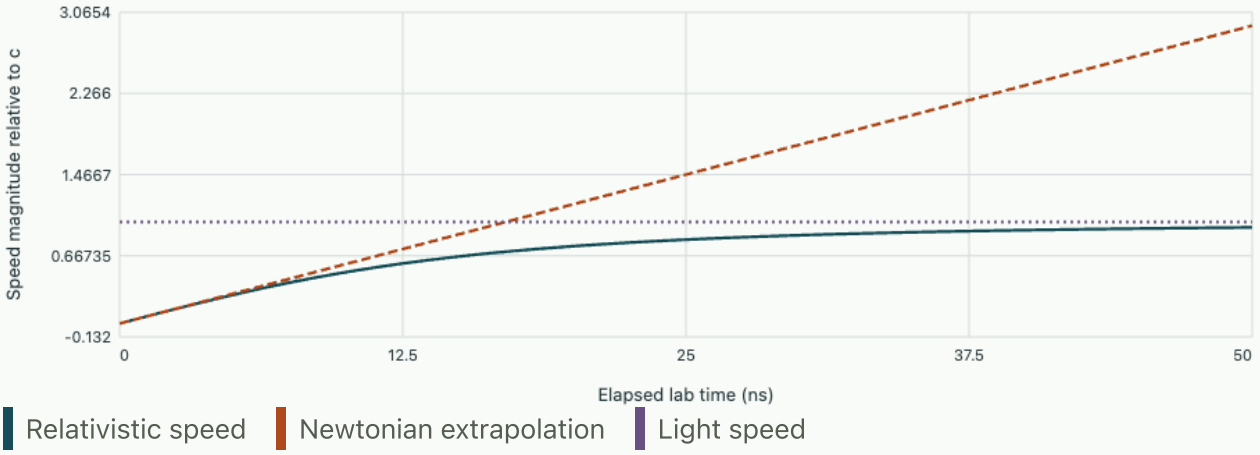
In a 100 kV/m uniform field, an initially stationary electron reaches about 0.706c after 17 ns lab time, while its own clock advances roughly 15 ns. Identify where the Newtonian extrapolation crosses the light-speed line and inspect the kinetic energy at that time.

Open this experiment online

Try the calculation

Electric field magnitude (kV/m) **100**

Lab time for a reading (ns) **17**



At 17 ns laboratory time the electron speed is 0.706 c. Its clock has advanced 15 ns and it has gained 211 keV.

Problem sheet · frame and trajectory checks

Measure the elapsed proper time

For the electric trajectory starting from rest, let the final momentum be $2mc$. Find the ratio of elapsed proper time to lab time, and compare it with the instantaneous final $1/\gamma$.

▼ **Work through the solution**

Since $s = 2$, the elapsed-time ratio is

$$\operatorname{arsinh}(2)/2 = 0.722.$$

At the endpoint $1/\gamma = 1/\sqrt{5} = 0.447$. The particle was slower earlier. Applying its final clock rate to the entire trajectory would discard its acceleration history.

Predict an orbit in crossed fields

Take perpendicular fields with $E/(cB) = 0.50$ and a particle initially at rest in the laboratory. Describe its motion in the frame moving at $\mathbf{E} \times \mathbf{B}/B^2$, and recover its average transverse drift direction.

▼ **Work through the solution**

The boost speed is $0.50c$. The primed electric field vanishes, the magnetic field is B/γ_d and the initial particle velocity is opposite the boost. Its primed motion is circular with constant speed. After every complete primed period its primed position returns to its starting value. Transforming the displacement between successive returns gives $\Delta x/\Delta t = v_d$ along the drift direction, independently of the sign of charge. Opposite signs reverse the sense of the primed circle.

Further reading

- Griffiths, 5th ed., §12.2 · relativistic mechanics
- Cambridge Electromagnetism, chapter 4 · relativistic particle motion

Study the full derivation

Relativistic particle dynamics and the field tensor · Waves, radiation and relativity

A circular track measures momentum per charge. A clock following the particle registers proper time, and a detector records how much kinetic energy it has gained. We will put these observations in one dynamical equation, solve that equation in uniform fields and exhibit a numerical pusher that can get the energy right while its orbit phase slowly slips.

Chapter 34 · Capacitance as a matrix of electrostatic energies

Three conductors occupy a grounded metal enclosure. Charging one can induce charge on another even when its voltage stays fixed. To predict a measurement we must specify which conductors have fixed voltage, which are isolated and what defines the reference potential.

Construct unit-voltage solutions

Consider a connected dielectric region with stationary geometry, linear reciprocal permittivity tensor $\boldsymbol{\epsilon}(\boldsymbol{r})$, no free volume charge, and perfectly conducting boundaries. For conductor j, define u_j as the solution of $\nabla \cdot (\boldsymbol{\epsilon} \nabla u_j) = 0$ which equals one on conductor j and zero on all other boundaries. When a grounded enclosing wall is used, each of these basis solutions also vanishes on that wall. Linearity and uniqueness give

$$\phi(\boldsymbol{r}) = \sum_j V_j u_j(\boldsymbol{r}).$$

Setting just one conductor to unit voltage determines an entire column of responses, measurable by holding the other potentials fixed.

Take $\boldsymbol{n}$ to be the unit normal pointing out of the dielectric, into the conductor. Since $\boldsymbol{D} = -\boldsymbol{\epsilon}\nabla\phi$, the conductor's charge facing the dielectric is $Q_i = \int_{S_i} \boldsymbol{n} \cdot \boldsymbol{\epsilon} \nabla \phi \, dS$. The two sign reversals, one from the electric field and one from the conductor's outward normal, cancel. The coefficients $C_{ij} = \partial Q_i / \partial V_j$ consequently integrate the flux from each basis solution.

$$Q_i = \sum_j C_{ij} V_j,$$

$$C_{ij} = \int_\Omega \nabla u_i \cdot \boldsymbol{\epsilon} \nabla u_j \, d^3 r.$$

To obtain the volume formula, apply the divergence theorem to $u_i \boldsymbol{\epsilon} \nabla u_j$. Only the gradient product survives in the volume. Only boundary i survives on the surface because it is the sole boundary where $u_i = 1$. Reciprocal permittivity makes the volume formula symmetric under $i \leftrightarrow j$; capacitance readings must therefore agree when drive and pickup electrodes are exchanged under the same conditions.

Background · Maxwell capacitance coefficient

$$Q_i = \sum_j C_{ij} V_j$$

Potentials are defined relative to a stated reference. A floating electrode has a prescribed total charge and a potential determined by the other electrodes.

Identify the reference and the energy

The electrostatic energy is

$$U = \frac{1}{2}\int \boldsymbol{E}\cdot\boldsymbol{D}\, d^3r = \frac{1}{2}\boldsymbol{V}^T C \boldsymbol{V}.$$

Its derivative with respect to V_i equals Q_i at fixed geometry. When the enclosing wall is included as an additional voltage degree of freedom, adding the same constant to all voltages leaves every gradient unchanged. The basis solutions including the wall sum to one. The full matrix then annihilates the vector of ones, its row sums vanish, and energy depends only on potential differences. Fixing the wall voltage removes that null mode. The reduced matrix acting on inner voltages can have nonzero row sums and is positive definite for a connected enclosure with positive permittivity.

For example, two electrodes have pair capacitance $c = 4$ pF, with additional $g_1 = 2$ pF and $g_2 = 3$ pF from each to the wall. The energy is $[g_1 V_1^2 + g_2 V_2^2 + c(V_1 - V_2)^2]/2$ in pJ for volt units. Differentiate to obtain reduced $C_{11} = 6,\ C_{22} = 7,\ C_{12} = C_{21} = -4$ pF. At $V_1 = 2, V_2 = -1$ V the charges are $16, -15$ pC and the energy is 23.5 pJ. Including the wall restores total charge zero for the charges facing the dielectric; the wall carries -1 pC.

Float one terminal and read a different capacitance

Disconnect the second electrode with initial charge zero. Its potential adjusts to satisfy $0 = -cV_1 + (g_2 + c)V_2$, giving $V_2 = 4V_1/7$. Substituting into Q_1 yields

$$C_{\mathrm{in}} = g_1 + c - c^2/(g_2 + c) = 26/7\,\mathrm{pF}.$$

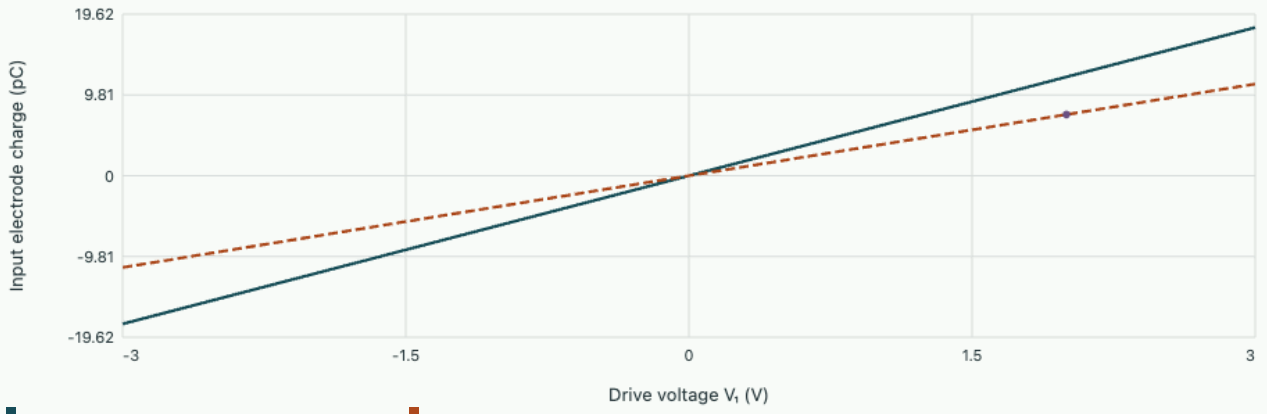
With electrode two grounded instead, $C_{\mathrm{in}} = 6$ pF. Both results came from the same electrostatic matrix; a voltmeter or guard changes the boundary conditions. The plot compares the charge delivered as the driven voltage changes. Vary the mutual coupling, and check that eliminating an isolated electrode agrees with differentiating the energy after minimising over its voltage for zero charge.

Open this experiment online

Try the calculation

Mutual capacitance (pF) 4

Drive voltage V_1 (V) 2



Second electrode grounded
Second electrode neutral and floating
Selected floating charge on electrode 1

Neutral electrode 2 floats to 1.143 V. Grounded input capacitance 6 pF, floating input capacitance 3.714 pF. Floating-case field energy 7.429 pJ.

Tutorial on calibrated terminal measurements

A measured reduced matrix has diagonal entries 8 and 9 pF and off-diagonal entries -3 pF. Recover the three pair capacitances to one another and to a reference enclosure. Predict the wall charge at $V_1 = 1, V_2 = 2$ V.

▼ **Work through the solution**

The mutual pair is 3 pF. Subtracting it from the diagonals gives 5 and 6 pF to the wall. The driven charges are $8 - 6 = 2$ pC and $-3 + 18 = 15$ pC. The enclosure-facing surface supplies the remaining -17 pC.

In the worked example electrode two is isolated with a trapped $Q_2 = 7$ pC. Determine the voltage induced when $V_1 = 0$, and the small-signal input capacitance measured by varying V_1.

▼ **Work through the solution**

Solve $7 = -4V_1 + 7V_2$ to find $V_2 = 1 + 4V_1/7$ V. The input charge is then $-4 + (26/7)V_1$ pC. The trapped charge shifts the intercept while leaving the incremental capacitance $26/7$ pF in this linear dielectric model.

Why can an inferred matrix with a negative eigenvalue fail even if its two diagonal entries are positive?

▼ **Work through the solution**

Choosing potentials proportional to that eigenvector would make the quadratic field energy negative. In a passive static dielectric with positive permittivity the field integral is nonnegative for every potential pattern. Positivity constrains the off-diagonal terms as well, for two terminals requiring $C_{11}C_{22} \geq C_{12}^2$.

Further reading · MIT 8.07, conductors and the capacitance matrix.

Study the full derivation

Capacitance matrices, floating nodes and controlled approximation bounds · Electrostatic field theory and measurement

A multi-electrode device generally has no single capacitance until the other terminals have been connected or isolated. An energy identity determines the matrix that relates voltages to charges, and a complementary flux identity can certify estimates even when the exact field is unknown.

Chapter 35 · How a conductor admits a field step and a sinusoid

A time-varying magnetic field enters a conducting slab from its surface. A suddenly imposed constant field continues to penetrate with elapsed time. A sinusoid establishes a persistent depth profile. Both behaviours follow from the same initial–boundary value equation.

Reduce Maxwell's equations to a diffusion equation

Use a uniform stationary Ohmic medium with scalar conductivity σ and permeability μ. Work on length and time scales for which displacement current is small compared with conduction current, $\omega\epsilon/\sigma \ll 1$ for each important Fourier component. Then

$$\nabla \times \boldsymbol{B} = \mu\sigma\boldsymbol{E}.$$

Taking a curl and using Faraday's law gives

$$\nabla \times \nabla \times \boldsymbol{B} = -\mu\sigma\partial_t \boldsymbol{B}.$$

The identity $\nabla \times \nabla \times \boldsymbol{B} = \nabla(\nabla \cdot \boldsymbol{B}) - \nabla^2\boldsymbol{B}$, together with absence of magnetic monopoles, leaves

$$\partial_t \boldsymbol{B} = D_m \nabla^2 \boldsymbol{B},$$

$$D_m = \frac{1}{\mu\sigma}.$$

D_m has units of area per time. Larger conductivity slows penetration: a small electric field can drive larger currents that oppose the changing applied field. An ideal mathematical step contains arbitrarily high frequencies outside this reduction. The result describes observations made after the brief electromagnetic start-up, on distances large enough for a continuum Ohmic description.

Background · Skin depth

$$\delta = \sqrt{2/(\mu\sigma\omega)}$$

In a good homogeneous conductor under sinusoidal excitation this is the amplitude decay length. Intensity decays twice as fast in the exponent.

Solve the initial–boundary value problem for a step

Fill $x > 0$ with conductor and impose a tangential field $B(0,t) = B_0$ for $t > 0$, with the interior initially field-free. At each finite time require $B(x,t) \to 0$ as $x \to \infty$. There is no imposed length scale, so try $B/B_0 = F(\eta)$ with $\eta = x/(2\sqrt{D_m t})$. The time derivative produces $-\eta F'/(2t)$; the second space derivative times D_m gives $F''/(4t)$. Therefore

$$F'' + 2\eta F' = 0.$$

First integrate $F' = Ae^{-\eta^2}$, then impose $F(0) = 1, F(\infty) = 0$. The normalisation $\int_0^\infty e^{-s^2} ds = \sqrt{\pi}/2$ gives

$$\frac{B(x,t)}{B_0} = \operatorname{erfc}\left(\frac{x}{2\sqrt{D_m t}}\right).$$

Any fixed fractional field level moves with $\sqrt{t}$. The depth at half amplitude is

$$x_{1/2} = 2\,\operatorname{erfc}^{-1}(0.5)\sqrt{D_m t} = 0.9539\sqrt{D_m t}.$$

For copper represented by $\sigma = 5.8 \times 10^7\,\mathrm{S/m}$, $\mu = \mu_0$, $D_m = 0.01372\,\mathrm{m^2/s}$. At $0.10\,\mathrm{ms}$ the half-depth is $1.12\,\mathrm{mm}$. Reaching twice that depth requires $0.40\,\mathrm{ms}$.

Solve for a periodic steady state

At the same surface impose $B(0,t) = \operatorname{Re} B_0 e^{i\omega t}$ for long enough that the transient has passed. Seek $B = \operatorname{Re} B_0 e^{i\omega t - kx}$. Substitution gives

$$D_m k^2 = i\omega.$$

Select the root with positive real part, $k = (1+i)/\delta$ with $\delta = \sqrt{2D_m/\omega}$. Thus

$$B(x,t) = B_0 e^{-x/\delta}\cos(\omega t - x/\delta).$$

At one skin depth the amplitude is e^{-1} of the surface amplitude and the oscillation lags by one radian. For the same model copper at 1 kHz, $\delta = 2.09\,\mathrm{mm}$, while the sinusoidal half-amplitude depth is $\delta \ln 2 = 1.45\,\mathrm{mm}$. A half-step depth and an exponential e-fold depth are different measurement definitions. Frequency and elapsed time are independent controls in the two experiments. In a slab comparable in thickness to either penetration scale the second interface supplies another boundary condition.

Read the two profiles at a shared probe

The graph plots the step response and the sinusoidal envelope against physical depth. Place the probe at 2 mm, vary the step duration and check the square-root law using the half-field depth. Hold that time fixed while quadrupling frequency. Only the periodic envelope shrinks by a factor of two in length. Time-dependent field and envelope both use normalised field-amplitude units.

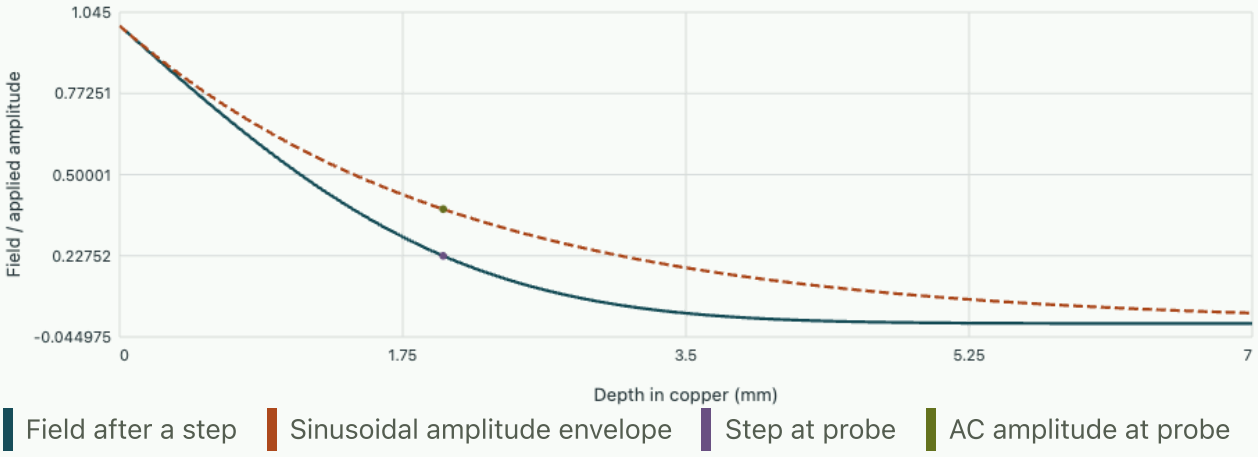
Open this experiment online

Try the calculation

Time after field step (ms) **0.1**

Sinusoidal drive frequency (kHz) **1**

Probe depth (mm) **2**



Half-step depth 1.117 mm. AC skin depth 2.09 mm. At the probe the step fraction is 0.2273, AC amplitude fraction 0.384 and AC phase lag 0.957 rad.

Tutorial on inversion and currents

A step profile has half-field depth $2.0\,\mathrm{mm}$ after $0.30\,\mathrm{ms}$. Infer D_m and σ if $\mu = \mu_0$.

▼ **Work through the solution**

$$D_m = (0.002/0.9539)^2/0.00030 = 0.01465\,\mathrm{m^2/s}.$$

Conductivity is

$$1/(\mu_0 D_m) = 5.43 \times 10^7\,\mathrm{S/m}.$$

A single profile measures the product $\mu\sigma$ unless permeability is independently known.

If $B = B_z(x,t)$, find the sign and surface magnitude of J_y for a positive ideal step. Compare the current integrated over depth with the imposed field.

▼ **Work through the solution**

Ampère gives

$$J_y = -\partial_x B_z/\mu.$$

The surface derivative is $-B_0/\sqrt{\pi D_m t}$, hence current is positive and its magnitude falls as $t^{-1/2}$. Its integral over all $x > 0$ is B_0/μ. The current spreads over a growing region while maintaining the total sheet current dictated by the field difference.

At a probe the AC envelope is 0.10 of the surface amplitude. What is the phase lag, and what is its time-delay equivalent at 1 kHz?

▼ **Work through the solution**

Solve $e^{-x/\delta} = 0.1$, so

$$x/\delta = \ln 10 = 2.303\,\mathrm{rad}.$$

At 1 kHz this equals

$$2.303/(2\pi\,1000) = 0.367\,\mathrm{ms}.$$

This is the phase delay of a single sinusoid. A pulse contains many frequencies with different attenuation and phase shifts.

Further reading · MIT Electromechanical Dynamics, magnetic diffusion and sinusoidal skin effect.

Study the full derivation

How a magnetic field relaxes into a finite conductor · Magnetism, induction and field dynamics

A conducting sheet initially shields its interior from an applied magnetic step. Given time, a normal resistive metal admits a static tangential magnetic field. The same equations give a finite slab a discrete relaxation spectrum and a periodic drive a complex penetration depth. These are different initial and boundary conditions for one diffusion operator.

Chapter 36 · Static screening and collective charge oscillations

In a neutral ion–electron gas a small positive probe draws an excess of electrons. A displaced electron fluid also experiences a restoring electric field and can oscillate. The equilibrium screening length and the collective frequency involve the same density but describe different experiments.

Count the electrons in a weak electrostatic potential

Assume fixed singly charged positive ions of uniform density n_0, and classical electrons in local thermal equilibrium at temperature T_e. Electron potential energy in an electric potential ϕ is $-e\phi$, where $e > 0$. The Boltzmann distribution relative to distant neutral plasma is $n_e = n_0 \exp[e\phi/(k_B T_e)]$. The net plasma charge density $e(n_0 - n_e)$ is negative near a small positive potential. Expand the exponential when $|e\phi| \ll k_B T_e$. Including a prescribed weak external charge density gives

$$(\nabla^2 - \lambda_D^{-2})\phi = -\rho_{\text{ext}}/\epsilon_0,$$

$$\lambda_D^2 = \frac{\epsilon_0 k_B T_e}{n_0 e^2}.$$

For a point source in an unbounded uniform plasma the source-free spherical equation at $r > 0$ is $(r\phi)'' = (r\phi)/\lambda_D^2$. Discard the growing exponential at infinity. The flux through a sufficiently small sphere fixes the coefficient of the $1/r$ singularity, yielding $\phi = Qe^{-r/\lambda_D}/(4\pi\epsilon_0 r)$ as the ideal linear Green function. A literal point charge makes ϕ large arbitrarily close to it. Applying this linear expression as a microscopic prediction there would violate the assumption that produced it.

At $r = \lambda_D$ the potential is e^{-1} times the bare Coulomb value at that same radius. The electric field is its derivative, so its ratio to the unscreened field is $(1 + r/\lambda_D)e^{-r/\lambda_D}$. At one Debye length this is $2/e = 0.736$. Equating the field ratio to the potential ratio would incorrectly remove a derivative term. If several mobile thermal species can equilibrate, the linear charge responses add, giving inverse squared screening lengths that add.

Displace the electron fluid before the ions move

Now consider a rapid, small longitudinal displacement $\xi(x, t)$ of a cold electron fluid against the uniform positive background. Particle conservation gives density perturbation $\delta n = -n_0 \partial_x \xi$. The resulting charge perturbation is $e n_0 \partial_x \xi$. For a nonzero spatial Fourier mode Gauss's law gives

$$E = e n_0 \xi/\epsilon_0,$$

with any uniform externally imposed field excluded. Newton's law $m_e \ddot{\xi} = -eE$ then gives

$$\ddot{\xi} + \omega_p^2 \xi = 0,$$

$$\omega_p^2 = \frac{n_0 e^2}{m_e \epsilon_0},$$

$$\omega_p \lambda_D = \sqrt{\frac{k_B T_e}{m_e}}.$$

This frequency is in radians per second. The oscillation cycles at $f_p = \omega_p/(2\pi)$. The length–frequency product equals a one-dimensional thermal-speed scale. The derivations used different closures: the static density had time to approach thermal equilibrium, while the oscillation neglected pressure gradients, collisions and ion motion. At appreciable $k\lambda_D$, pressure and kinetic effects change the longitudinal frequency and damping. A fully kinetic model is required to predict Landau damping.

Infer density and temperature from two measurements

At $n_0 = 10^{16}\,\mathrm{m}^{-3}$ and $k_B T_e = 2\,\mathrm{eV}$, the expressions give $\lambda_D = 105\,\mu\mathrm{m}$ and $f_p = 898\,\mathrm{MHz}$. Doubling temperature alone increases screening length by $\sqrt{2}$ and leaves this cold oscillation frequency fixed. Quadrupling density at fixed temperature halves the length and doubles the frequency. With a measured frequency $1.80\,\mathrm{GHz}$, density would be about four times the above value. An independently measured λ_D then determines $k_B T_e = m_e (2\pi f_p \lambda_D)^2$.

Set density and temperature in the graph and compare screened potential and field ratios at a selected physical distance. Use the readout to turn an oscillation period into a density estimate. Check the continuum hypothesis by estimating $N_D = 4\pi n_0 \lambda_D^3/3$, the number of particles in a Debye sphere. It is approximately 4.9×10^4 at the worked point, making a collective description plausible. A small N_D, degeneracy or strong coupling would undermine the thermal mean-field treatment.

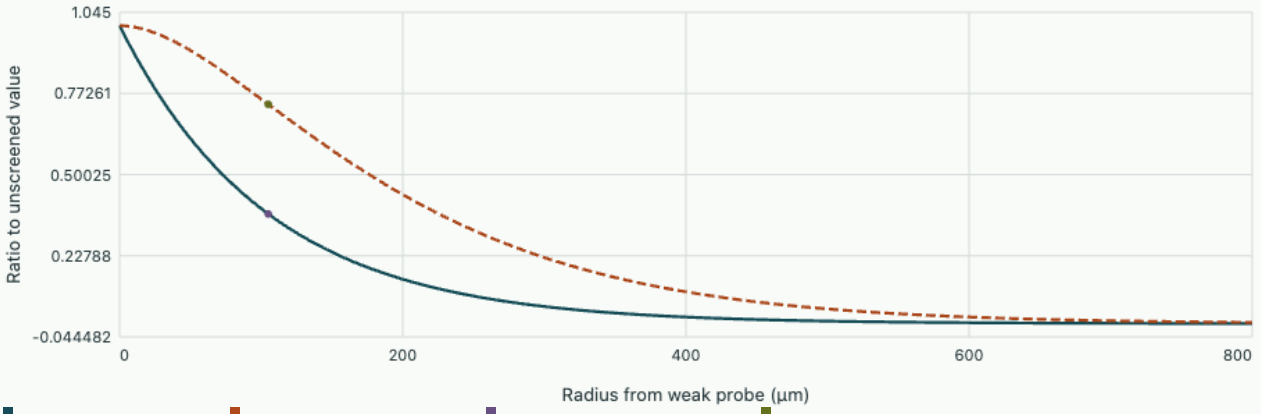
Open this experiment online

Try the calculation

Electron density (10^{16} m^{-3}) 1

Electron thermal energy (eV) 2

Probe radius (µm) 105



Debye length 105.1 µm, plasma frequency 897.9 MHz, period 1.114 ns. The Debye sphere contains approximately 48670 electrons.

Tutorial on independent spatial and temporal tests

A probe measures $f_p = 0.90\,\mathrm{GHz}$ and $\lambda_D = 210\,\mu\mathrm{m}$. Infer approximate density and thermal energy relative to the worked example.

▼ **Work through the solution**

The frequency is essentially unchanged, so density is approximately $10^{16}\,\mathrm{m}^{-3}$. The doubled screening length at fixed density requires four times the thermal energy, about 8 eV. The independent product check gives a thermal speed of roughly $1.19 \times 10^{6}\,\mathrm{m/s}$.

How far from a weak source does the screened potential fall to 1 percent of the bare value? Calculate the field ratio at that point.

▼ **Work through the solution**

The distance is

$$r = \lambda_D \ln 100 = 4.605 \lambda_D.$$

The field ratio is

$$(1 + 4.605)/100 = 0.0561.$$

These are ratios to the appropriate unscreened quantities at the same physical position.

If both electrons and singly ionised positive ions with equal temperature and density can rearrange in the static experiment, compare the new screening length with the electron-only value. Does this change the electron-frequency formula on a time scale where ions are stationary?

▼ **Work through the solution**

Both species respond to weak potential with charge-density changes proportional to $-q_s^2 n_s \phi/(k_B T_s)$. The two identical inverse squared lengths add, yielding $\lambda_D/\sqrt{2}$. On the much shorter oscillation time scale the ions remain the fixed background in the cold electron derivation.

Further reading · MIT Introduction to Plasma Physics, shielding and plasma oscillations.

Study the full derivation

From mobile ions to screening lengths and differential double-layer capacitance · Electrostatic field theory and measurement

A charged planar wall in an electrolyte can be studied without assuming a capacitance in advance. Diffusion and electrostatic drift first establish an ion cloud. The charge–voltage slope of that cloud is an experimental output of the model.

Chapter 37 · Reconstructing a partially polarised electromagnetic wave

A detector averages enormous numbers of optical cycles. Its horizontal and vertical readings may be equal for a circularly polarised wave, a diagonal linearly polarised wave, or unpolarised light. Which additional intensities distinguish these preparations?

Start from what a polarising analyser measures

Take a quasi-monochromatic beam propagating in a fixed direction, resolved in transverse axes x, y. Use complex field amplitudes with time factor $e^{-i\omega t}$. Absorb common power and impedance factors into their units. The time- or ensemble-averaged coherency matrix $J = \langle \boldsymbol{E}\boldsymbol{E}^{\dagger} \rangle$ has four real degrees of freedom because it is Hermitian. An analyser selecting the normalised complex state $\boldsymbol{a}$ passes amplitude $\boldsymbol{a}^{\dagger}\boldsymbol{E}$ and measures $I_a = \boldsymbol{a}^{\dagger} J \boldsymbol{a}$. Thus every analyser projects the same matrix onto a distinct polarisation.

Choose $S_0 = J_{xx} + J_{yy}$, $S_1 = J_{xx} - J_{yy}$,

$$S_2 = 2\,\mathrm{Re}\,J_{xy},$$

$$S_3 = -2\,\mathrm{Im}\,J_{xy}.$$

For this convention the state $(1, i)/\sqrt{2}$ has positive S_3. Naming it by its explicit Jones vector avoids a handedness ambiguity between observer viewing directions. Inverting those definitions gives

$$J = \frac{1}{2}\begin{pmatrix} S_0 + S_1 & S_2 - iS_3 \\ S_2 + iS_3 & S_0 - S_1 \end{pmatrix}.$$

Horizontal and vertical linear analysers therefore give $(S_0 \pm S_1)/2$. The $+45^{\circ}, -45^{\circ}$ pair gives $(S_0 \pm S_2)/2$. The circular pair $(1, \pm i)/\sqrt{2}$ gives $(S_0 \pm S_3)/2$. Every pair sum recovers total intensity, providing an internal calibration check. Subtracting the complementary measurements cancels a common additive detector offset in each Stokes difference, though S_0 still needs a dark correction.

Derive the physically allowed domain

Because $I_a \geq 0$ for every analyser, both eigenvalues of J are nonnegative. Solve its quadratic characteristic equation using the trace S_0 and determinant $(S_0^2 - S_1^2 - S_2^2 - S_3^2)/4$. The eigenvalues are $(S_0 \pm |\boldsymbol{S}|)/2$, where $\boldsymbol{S} = (S_1, S_2, S_3)$. Consequently $P = |\boldsymbol{S}|/S_0$ lies between zero and one for a beam with nonzero intensity. A single fully polarised Jones state gives rank one and $P = 1$. An incoherent mixture of two orthogonal states of equal intensity has $J = (S_0/2) I$ and $P = 0$.

Suppose a corrected data set gives $I_x, I_y = 70, 30$; $I_{+45}, I_{-45} = 65, 35$; $I_{+}, I_{-} = 60, 40$ in the same units. The pair sums agree at 100. Reconstructed parameters are $(100, 40, 30, 20)$, giving $P = \sqrt{2900}/100 = 0.539$. The eigenvalues are 76.93 and 23.07. Subtracting the smaller times the identity leaves a rank-one contribution of intensity 53.85 and an unpolarised remainder of intensity 46.15. This decomposition describes the observed second moments; several microscopic preparations can produce the same J.

Predict a rotating analyser and identify its blind spot

For $\boldsymbol{a} = (\cos\alpha, \sin\alpha)$, expanding the projection and using double-angle identities gives

$$I_\alpha = [S_0 + S_1 \cos 2\alpha + S_2 \sin 2\alpha]/2.$$

A linear analyser alone is blind to S_3, even when sampled at hundreds of angles. In the example the extrema are 50 ± 25, with the maximum at $\alpha = \frac{1}{2}\,\mathrm{atan}\,2(30, 40) = 18.43°$. The linear fringe visibility 0.50 is slightly below the total degree 0.539. Inserting a quarter-wave plate with Jones matrix $\mathrm{diag}(1, -i)$ before the same analyser transforms S_2 into S_3 in the intensity formula. One can now obtain the previously hidden signed correlation from readings at $45°$ and $135°$.

The experiment prepares an adjustable polarised fraction with azimuth ψ and ellipticity angle χ, mixed with an unpolarised background. Its normalised Stokes vector is $P(\cos 2\chi \cos 2\psi, \cos 2\chi \sin 2\psi, \sin 2\chi)$. Make the unassisted intensity scan flat by choosing circular polarisation, then insert the fixed retarder virtually and predict the recovered difference. Setting $P = 0$ makes both scans flat. This comparison shows which measurement discriminates the two cases.

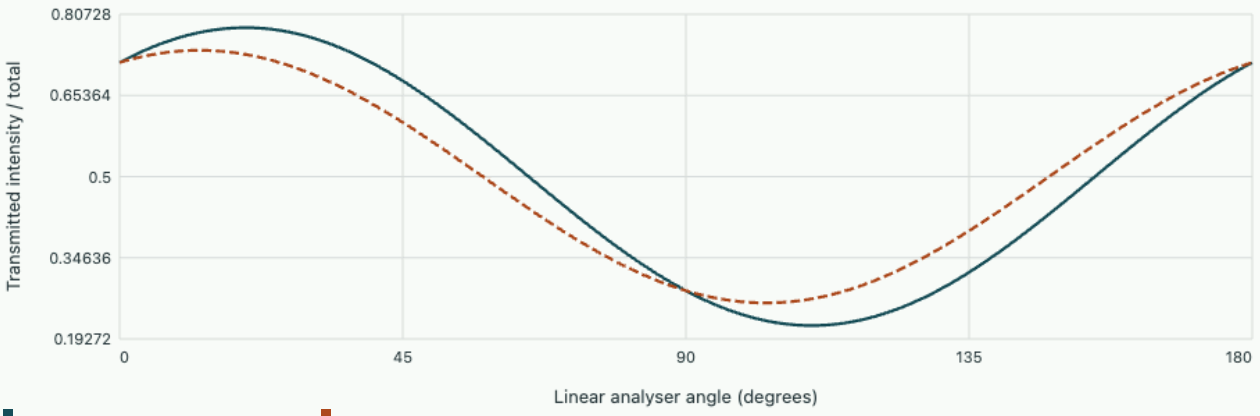
Open this experiment online

Try the calculation

Polarised fraction P **0.6**

Azimuth ψ (degrees) **20**

Ellipticity χ (degrees) **10**



Measured complementary pairs (horizontal, diagonal, circular): 0.716 / 0.284 · 0.6812 / 0.3188 · 0.6026 / 0.3974. Reconstructed $(S_1,S_2,S_3)/S_0$ = (0.4319, 0.3624, 0.2052); degree 0.6.

Tutorial on instrument design

A linear scan alone has $I_{\max} = 75, I_{\min} = 25$. Give the range of total degrees of polarisation allowed and the maximum possible magnitude of S_3.

▼ **Work through the solution**

The linear magnitude is 50 and $S_0 = 100$. Thus $0.5 \le P \le 1$, and the positivity bound gives

$$|S_3| \le \sqrt{100^2 - 50^2} = 86.60.$$

A circular analyser measurement fixes the remaining magnitude and sign.

An ideal beam contains 60 units of pure horizontal light and 40 units of unpolarised light. Compare its Stokes vector with a coherent superposition of horizontal and vertical components with intensities 80 and 20 and zero relative phase.

▼ **Work through the solution**

The incoherent mixture has $(100, 60, 0, 0)$ and $P = 0.6$. The coherent preparation has $(100, 60, 80, 0)$ since $2\sqrt{80 \times 20} = 80$, and $P = 1$. Their horizontal and vertical readings coincide. A diagonal analyser gives 50 for the mixture and 90 for the coherent preparation.

All three measured component differences equal 70 at measured total 100. Can a physical beam have these parameters?

▼ **Work through the solution**

They imply $|\boldsymbol{S}| = 70\sqrt{3} = 121.2 > S_0$ and a negative matrix eigenvalue. Check dark subtraction, gain calibration, alignment and whether the source changed between readings. A physical fit with stated uncertainties is required before interpreting the data as an optical state.

Further reading · ETH Zürich, Stenflo on coherency and Stokes parameters.

Study the full derivation

What an analyser measures about polarisation and coherence · Waves, radiation and relativity

Two red beams produce the same reading through every rotating linear polarizer. One might be circularly polarized and the other an incoherent mixture. Build the measurements that distinguish them, derive the visibility of their interference fringes, and decide which information is lost on a slow detector.

Chapter 38 · Recovering a vector field from divergence, curl and boundary data

A numerical field solver returns three components at every point. Divergence measures sources and curl measures circulation, yet neither measurement detects a spatially uniform offset. Helmholtz decomposition makes precise what can be recovered, how to compute it, and what the boundary must supply.

Invert the vector Laplacian in open space

Let $\boldsymbol{F}$ be smooth on all of space and rapidly decreasing, together with its needed derivatives, so integration by parts has no surface term at infinity. The identity $\nabla^2 \boldsymbol{F} = \nabla(\nabla \cdot \boldsymbol{F}) - \nabla \times (\nabla \times \boldsymbol{F})$ reduces reconstruction to three scalar Poisson inversions. Define $G(\boldsymbol{r}) = 1/(4\pi|\boldsymbol{r}|)$; distributionally $-\nabla^2 G = \delta^3$. For $d = \nabla \cdot \boldsymbol{F}$ and $\boldsymbol{c} = \nabla \times \boldsymbol{F}$, convolution with G gives

$$\Phi = G * d,$$

$$\boldsymbol{A} = G * \boldsymbol{c},$$

$$\boldsymbol{F} = -\nabla \Phi + \nabla \times \boldsymbol{A}.$$

To verify, first note that $\nabla \cdot \boldsymbol{c} = 0$; a supplied curl field must satisfy this compatibility condition. Moving the derivative through convolution gives

$$\nabla \cdot \boldsymbol{A} = G * (\nabla \cdot \boldsymbol{c}) = 0.$$

The divergence of the proposed $\boldsymbol{F}$ is $-\nabla^2 \Phi = d$, and its curl is

$$\nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A} = \boldsymbol{c}.$$

Any difference from the original field has zero divergence and curl, hence componentwise zero Laplacian. Smoothness and decay on all space force this harmonic difference to vanish. The reconstructed vector pieces are unique in this setting, while adding a gradient to an unrestricted vector potential still leaves its curl unchanged.

Derive the Fourier projectors

On a periodic grid separate the mean mode first. For a nonzero Fourier wavevector $\boldsymbol{k}$ with convention $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$, differentiation multiplies by $i\boldsymbol{k}$. A longitudinal component is parallel to $\boldsymbol{k}$, so its amplitude must reproduce the scalar product $\boldsymbol{k} \cdot \widehat{\boldsymbol{F}}$. Subtraction leaves a perpendicular component:

$$\widehat{\boldsymbol{F}}_L = \frac{\boldsymbol{k}(\boldsymbol{k} \cdot \widehat{\boldsymbol{F}})}{|\boldsymbol{k}|^2},$$

$$\widehat{\boldsymbol{F}}_T = \widehat{\boldsymbol{F}} - \widehat{\boldsymbol{F}}_L.$$

The first projector squared equals itself, and the two pieces are orthogonal. Thus their squared magnitudes add mode by mode, even for complex amplitudes with the Hermitian inner product. Equivalently the longitudinal piece has zero curl and the transverse piece has zero divergence. At $\boldsymbol{k} = 0$,

the denominator vanishes and both derivative measurements are zero. A known mean, applied uniform field or boundary condition supplies that information. Setting the mean to zero silently would discard an actual applied bias.

Recover a patterned measurement

Take a two-dimensional periodic field $\boldsymbol{F} = \boldsymbol{b} + \boldsymbol{a}\cos(2x + y)$ in dimensionless periodic coordinates with $\boldsymbol{a} = (3, 1)$, $\boldsymbol{b} = (1, 0)$. Its measured divergence is $-7\sin(2x + y)$. Its out-of-plane curl is $+\sin(2x + y)$, from $\partial_x F_y - \partial_y F_x$. Projection with $\boldsymbol{k} = (2, 1)$ gives $\boldsymbol{a}_L = (2.8, 1.4)$, $\boldsymbol{a}_T = (0.2, -0.4)$. The squared amplitudes are 9.8 and 0.2, summing to the original 10; only 2 percent of the nonconstant modal power is transverse. Neither differentiated pattern contains the bias. At phase zero the observed x-component is four, of which 2.8 is longitudinal, 0.2 transverse and one is uniform.

The graph scans phase, showing the x-component of the measured field, the two recovered varying pieces and the mean. Change the input amplitudes until $\boldsymbol{k} \cdot \boldsymbol{a} = 0$, then until $\boldsymbol{k} \times \boldsymbol{a} = 0$. Changing the mean should move only the total and constant terms. In a discretised solver the differentiation symbol used in a projection must match the numerical divergence being tested; the exact spectral $i\boldsymbol{k}$ and a finite difference need not represent the same grid operator.

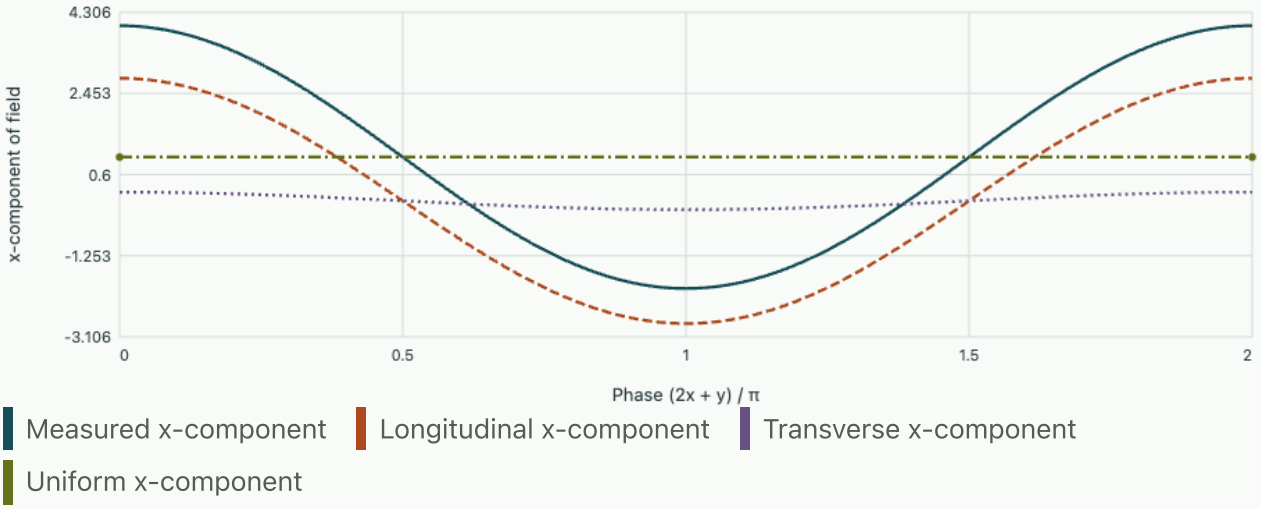
Open this experiment online

Try the calculation

Cosine amplitude a_x **3**

Cosine amplitude a_y **1**

Uniform component b_x **1**



Longitudinal vector (2.8, 1.4), transverse (0.2, -0.4). Squared amplitudes add to 9.8 + 0.2 = 10. The constant x-component is 1.

Specify a boundary on a finite domain

One convenient decomposition writes $\boldsymbol{F} = \nabla u + \boldsymbol{T}$ with $\nabla \cdot \boldsymbol{T} = 0$ and $\boldsymbol{T} \cdot \boldsymbol{n} = 0$ on the boundary. Solve the Neumann problem $\nabla^2 u = \nabla \cdot \boldsymbol{F},\ \partial_n u = \boldsymbol{F} \cdot \boldsymbol{n}$. The divergence theorem makes these data compatible. The potential is determined up to an additive constant in a connected domain, which does not affect its gradient. Orthogonality follows from integrating $\nabla u \cdot \boldsymbol{T}$ by parts: both the volume divergence and boundary normal term vanish. Holes can admit nontrivial fields with zero curl and divergence, so reconstructing a full vector field from its derivatives alone can require additional circulation data.

Tutorial on recovering missing information

With $\boldsymbol{k} = (2, 1)$, a sensor returns divergence $-5\sin(2x + y)$ and curl $-5\sin(2x + y)$. Infer the cosine amplitudes a_x, a_y.

▼ **Work through the solution**

From the divergence, $2a_x + a_y = 5$. From the curl, $2a_y - a_x = 5$. Solving gives

$$a_x = 1, a_y = 3.$$

The projected amplitudes are $(2, 1)$ and $(-1, 2)$, with equal squared magnitude. The mean remains independent of both readings.

On an annulus $a < r < b$, consider $\boldsymbol{H} = \boldsymbol{e}_\theta / r$. Which derivative and boundary data vanish? Propose one measurement that distinguishes it from zero.

▼ **Work through the solution**

Its planar divergence vanishes and its out-of-plane curl is $r^{-1}\partial_r(rH_\theta) = 0$ on the annulus. Its normal component vanishes at both circular boundaries. Nevertheless circulation around any circle surrounding the hole is 2π. The axis where the field is singular is excluded, so the all-space uniqueness argument does not apply.

A computed projection claims longitudinal power 6, transverse power 4, and original power 13 in one nonzero Fourier mode. Is this consistent with exact orthogonal projection in the same norm?

▼ **Work through the solution**

No. The two projected vectors are perpendicular, so their powers must sum to the original 13. Check whether powers include the same normalisation, whether the mean was included, and whether the solver and diagnostic use the same differentiation symbol and inner product.

Further reading · University of Virginia, Helmholtz's theorem and uniqueness · Rochester electromagnetic theory, transverse and longitudinal projections.

Study the full derivation

Vector calculus through flux, circulation and singular sources · Electrostatic field theory and measurement

A closed-surface flux and a closed-path circulation answer different experimental questions. Starting from small physical collecting surfaces shows why the associated differential operators look different in Cartesian, cylindrical and spherical coordinates.

Chapter 39 · A ring, a disk and an electric-field maximum in empty space

Gauss's law gives the flux through a closed surface. For a finite circular source the field varies across that surface. Adding the fields of actual charge elements predicts both the near field and a surprising axial maximum when the centre is removed.

Decide what symmetry eliminates

A uniformly charged ring of radius s lies in the xy plane. A detector on the z axis is a distance $R = \sqrt{s^2 + z^2}$ from every charge element. Contributions transverse to the axis cancel in diametrically opposite pairs. The potential adds as a scalar, giving $dV = dq/(4\pi\epsilon_0 R)$ for the whole ring charge dq. Choose zero potential at infinity. Differentiating with respect to detector position gives

$$dE_z = z\,dq/(4\pi\epsilon_0 R^3).$$

The same factor z/R appears if each Coulomb force is projected directly onto the axis. At the ring centre the potential is finite and the field is zero.

Construct an annulus with uniform surface density σ, inner radius a and outer radius b. A radial interval ds contributes $dq = 2\pi\sigma s\,ds$; forgetting s would assign equal charge to equal widths at the centre and rim. Integrating and differentiating yields

$$V(z) = \frac{\sigma}{2\epsilon_0}\left[\sqrt{z^2 + b^2} - \sqrt{z^2 + a^2}\right],$$

$$E_z(z) = \frac{\sigma z}{2\epsilon_0}\left[\frac{1}{\sqrt{z^2 + a^2}} - \frac{1}{\sqrt{z^2 + b^2}}\right].$$

The potential is even in z and the field odd. When $a > 0$ every point on the central axis is in empty space and both functions are smooth there. This smoothness changes in the disk limit.

Investigate two noncommuting limits

Set $a = 0$ before moving the probe to the plane. The potential then contains $-|z|$ and the field contains $\operatorname{sgn}(z)$. Thus $E_z(0^+) - E_z(0^-) = \sigma/\epsilon_0$, just the charge-sheet jump. By contrast, taking $z \to 0$ first at any positive hole radius gives zero field. The order of shrinking the hole and approaching its centre determines whether the probe sits on material. A grid that places its probe exactly at $z = 0$ for both geometries can conceal the difference.

At large positive z, Taylor-expand the two square roots before subtracting. The z terms cancel. Since $Q = \pi\sigma(b^2 - a^2)$,

$$V(z) = \frac{Q}{4\pi\epsilon_0 z}\left[1 - \frac{a^2 + b^2}{4z^2} + O(b^4/z^4)\right],$$

$$E_z = \frac{Q}{4\pi\epsilon_0 z^2}\left[1 - \frac{3(a^2 + b^2)}{4z^2} + \cdots\right].$$

The axial dipole term is absent by reflection symmetry. For very large z a numerical difference between almost identical square roots loses precision. Multiply by the conjugate to evaluate $V = \sigma(b^2 - a^2)/[2\epsilon_0(\sqrt{z^2 + b^2} + \sqrt{z^2 + a^2})]$.

Find an extremum without violating the vacuum field equations

For a nonzero hole the field grows linearly near the origin and decreases as an inverse square in the far field. Differentiate $z/\sqrt{z^2 + s^2}$ to get $s^2/(z^2 + s^2)^{3/2}$. Equating the two edge contributions gives

$$z_{\max}^2 = \frac{b^2 a^{4/3} - a^2 b^{4/3}}{b^{4/3} - a^{4/3}}.$$

This is a maximum of one component along one line. No interior maximum of the electrostatic potential is implied. A lateral displacement would break the ring-pair cancellation and requires integration over the azimuthal coordinate. On axis the field also equals the difference between the fields of two disks of common density.

Plan two field measurements

Locate a peak in a centimetre-scale annulus

Choose $a = 10$ mm, $b = 20$ mm and $\sigma = 1.0$ nC/m². The maximum occurs at $z = 0.4934b = 9.87$ mm. At this point $2\epsilon_0 E_z/\sigma = 0.25992$, so $E_z = 14.7$ V/m. A point charge approximation would incorrectly predict that the field increases indefinitely as the detector approaches the hole. The missing charge near the centre matters even though the whole source has nonzero net charge.

Turn a desired error into a measurement distance

For the same radius ratio the leading relative correction to the monopole field is $0.9375(b/z)^2$. To keep this correction near 1%, choose $z/b \simeq 9.68$, about 0.194 m here. Substituting this distance in the exact expression checks the term neglected in the estimate. For a radial charge distribution with changing sign the monopole could cancel altogether, invalidating the inference.

Integrate an annulus ring by ring

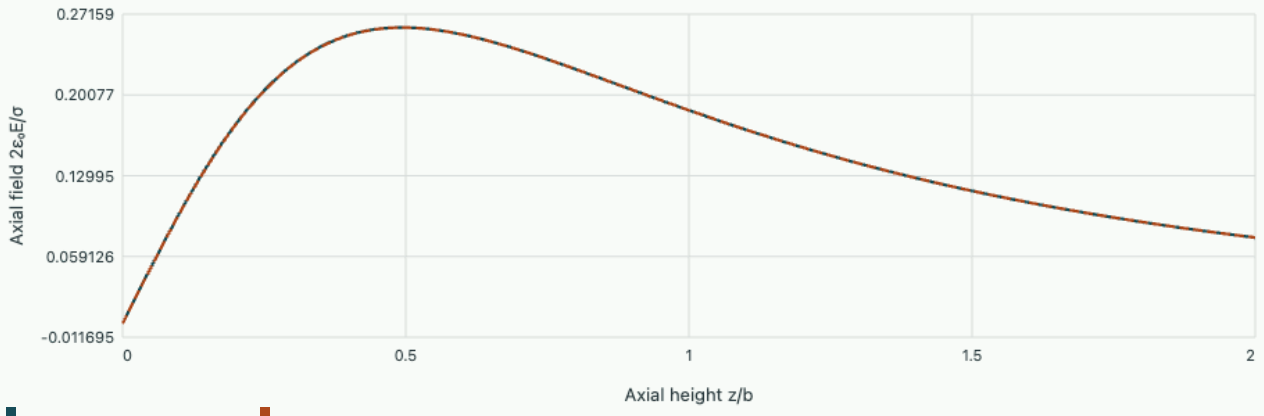
The graph compares midpoint integration over physical ring areas against the analytic axial field. A hole radius of half the outer radius puts the maximum at z/b = 0.4934. Move the inner edge and refine the source cells separately. A changing peak location and a shrinking numerical error have different causes.

Open this experiment online

Try the calculation

Hole / outer radius 0.5

Radial integration cells 8



Analytic integral Ring quadrature

The analytic peak is at z/b = 0.4934 with field 0.2599. At that location the quadrature differs by -0.01507%.

Problem sheet · what can an axial probe tell you?

Locate the peak of a narrow ring

Shrink the annulus to a narrow ring of fixed total charge and radius a. Find the field maximum directly and compare the answer with the picture above.

▼ **Work through the solution**

The ring field is proportional to $z/(z^2+a^2)^{3/2}$. Its derivative is proportional to $(a^2-2z^2)/(z^2+a^2)^{5/2}$. Thus the maximum on the positive axis occurs at $a/\sqrt{2}$. Keeping total charge fixed requires the surface density to grow as the width shrinks.

Compute with a nonuniform surface density

Suppose the density increases linearly from zero at the centre of a disk to σ_0 at its outer radius b. Describe an integration rule and give two independent checks that do not require an analytic field formula.

▼ **Work through the solution**

Divide the radius into cells with midpoint s_j. Assign each cell its exact area $\pi(s_{j,+}^2 - s_{j,-}^2)$, multiply by $\sigma_0 s_j/b$, and add the axial ring field. The summed source charge must converge to $2\pi\sigma_0 b^2/3$. At large z the computed field times $4\pi\epsilon_0 z^2$ approaches this same charge. At each fixed positive z, double the cell count and check convergence separately from the far-field comparison.

Further reading

- Griffiths, 5th ed., §§2.1.4, 2.3.4 and 3.4 · integration and distant fields
- Fließbach, chapter 5 · Coulomb law

Study the full derivation

From a force reading to Poisson’s equation · Electrostatic field theory and measurement

Imagine trying to infer an invisible charge distribution with a force probe and a voltmeter. Coulomb’s law, Gauss’s law and the electrostatic potential describe the same experiment in three different ways. Deriving the connections also exposes the measurements a single probe cannot make.

Chapter 40 · The same metal sphere connected to ground or electrically isolated

A grounded sphere can draw charge through a wire. An isolated one can only rearrange what is already present. Both are equipotentials, but the second problem needs a different image system and gives a different force at long range.

Find the image by matching distances on a sphere

A positive test charge q sits at $z = d > a$ outside a conducting sphere of radius a centred at the origin. In the exterior vacuum region the potential has the known Coulomb singularity at q and is otherwise harmonic. Look for an auxiliary point $s\hat{\mathbf{z}}$ inside the metal whose distance from each surface point is proportional to the distance from the real charge. The squared distances at polar angle θ are $a^2 + s^2 - 2as\cos\theta$ and $a^2 + d^2 - 2ad\cos\theta$. Matching their constant and cosine terms yields

$$s = a^2/d,$$

$$R_s = (a/d)R_q.$$

A charge $q' = -qa/d$ at the auxiliary point then cancels the real potential on the sphere. The superposition vanishes at infinity. Uniqueness for prescribed potentials makes this the actual exterior solution. The mathematical source is outside the physical domain of the solution. A surface enclosing just the conductor can be evaluated using the equivalent exterior field, so the total induced charge is q'. It arrived through the grounding connection.

Change the constraint by adding one image

For an isolated sphere of prescribed total charge Q_c, its surface potential is a constant that must be determined. Add a central charge $q_0 = Q_c - q'$. Its surface contribution $q_0/(4\pi\epsilon_0 a)$ is uniform, preserving equipotentiality, while the total equivalent enclosed charge becomes Q_c. The resulting sphere potential is $V_s = Q_c/(4\pi\epsilon_0 a) + q/(4\pi\epsilon_0 d)$. The grounded and neutral isolated solutions agree only when the geometry or external charge makes their net-charge constraints coincide.

Use Gauss's boundary condition $\sigma = -\epsilon_0 \partial_r V|_{a^+}$. The grounded result and its modification are

$$\sigma_g(\theta) = -\frac{q}{4\pi a}\,\frac{d^2 - a^2}{(d^2 + a^2 - 2ad\cos\theta)^{3/2}},$$

$$\sigma_c(\theta) = \sigma_g(\theta) + \frac{q_0}{4\pi a^2}.$$

For a neutral sphere the charge layer facing the source is negative, while a sufficiently distant part of the surface becomes positive. Integrating over the sphere involves the area element $2\pi a^2 \sin\theta\, d\theta$. A uniform angular average without that weight miscounts the sharply concentrated polar charge.

Take a force at the real charge, then take a large-distance limit

Evaluate the electric field of the image charges only at the real charge. The grounded force F_z and its neutral-isolated counterpart are

$$F_g = -\frac{q^2}{4\pi\epsilon_0}\frac{ad}{(d^2-a^2)^2},$$

$$F_{\text{neutral}} = F_g + \frac{q^2}{4\pi\epsilon_0}\frac{a}{d^3}.$$

At $d \gg a$, the grounded force begins at order d^{-3} because the grounded sphere acquires charge proportional to $1/d$. In the neutral case that leading term cancels, leaving $-2q^2a^3/(4\pi\epsilon_0 d^5)$. The conductor responds mainly as an induced dipole. This cancellation offers a demanding numerical check at large separation. For the grounded case the energy relative to infinite separation follows from integrating $-F_g$ with respect to d, yielding $U = -q^2a/[8\pi\epsilon_0(d^2-a^2)]$. It also equals $qV_{\text{image}}(d)/2$, with the half arising as the conductor polarises continuously during assembly.

Compare two preparations with identical geometry

Unplug a sphere before approaching it

Set $a = 20$ mm, $d = 40$ mm and $q = 1.0$ nC. The auxiliary image lies at 10 mm and has charge -0.50 nC. Grounding gives $F_g = -4.99$ µN. For a neutral isolated sphere the added central $+0.50$ nC exerts $+2.81$ µN and the total is -2.18 µN. Its surface potential is $q/(4\pi\epsilon_0 d) \simeq 225$ V. In both experiments the metal still has no interior electric field in static equilibrium.

Disconnect after the charge is already present

If the wire is removed at $d = 2a$, the sphere keeps its acquired $-q/2$. With the real charge still at $2a$ the exterior solution is unchanged. Move it to $4a$. The new off-centre image is $-q/4$ and the central image is also $-q/4$. The net conductor charge remains $-q/2$, and its potential becomes negative. A calculation that continually sets the sphere voltage to zero after disconnection would model an ongoing exchange of charge with an external reservoir.

Measure the charge density of two spheres with different connections

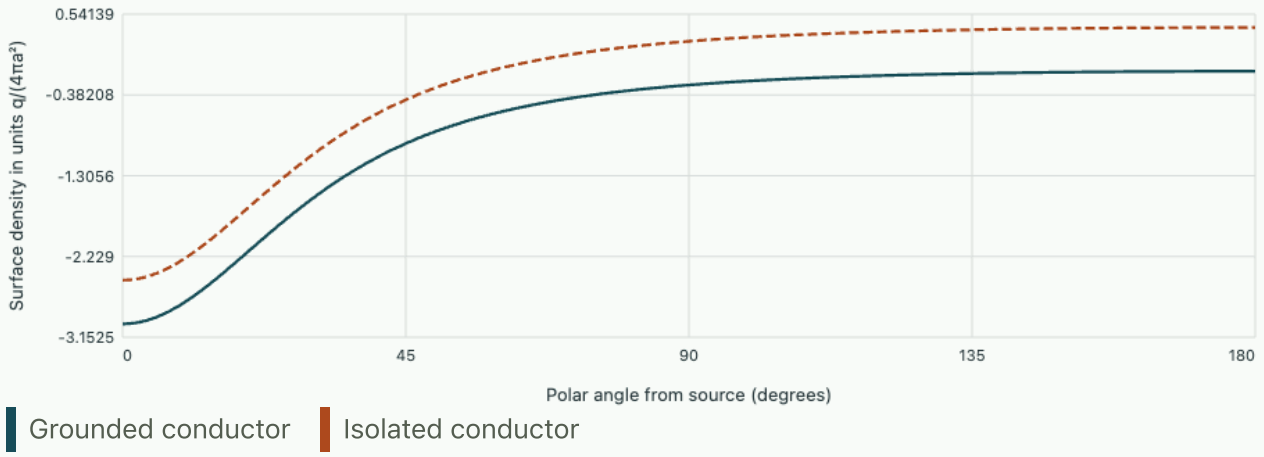
With the source at two sphere radii, the grounded surface draws a total charge –q/2; the neutral isolated sphere retains zero total charge. Change the distance and the trapped initial conductor charge. Inspect the source-facing pole and the far side before deciding whether a net force can be repulsive.

Open this experiment online

Try the calculation

Source distance d/a **2**

Isolated conductor charge Qc/q **0**



Grounding draws net charge -0.5 q. Forces in units q²/(4πε₀a²) are -0.2222 grounded and -0.09722 isolated; positive means away from the sphere.

Problem sheet · charge balance and a near-surface limit

Recover a plane image from a curved surface

Let $d = a + h$ with $h \ll a$. Show that the near-surface force on the charge approaches the grounded plane result and determine when the central neutralising image matters.

▼ **Work through the solution**

Use $d^2 - a^2 = 2ah + h^2$. The leading grounded force is $-q^2/(16\pi\epsilon_0 h^2)$, identical to attraction to an image $-q$ a distance $2h$ away. The central neutralising field is of order $q/(4\pi\epsilon_0 a^2)$, smaller by order $(h/a)^2$ in the force budget.

Where does a neutral sphere change surface charge sign?

For $d = 2a$ and $Q_c = 0$, solve $\sigma_c = 0$ for the polar angle. How can a sphere still attract the external charge with zero net charge?

▼ **Work through the solution**

In units $q/(4\pi a^2)$ the equation is

$$-3/(5 - 4\cos\theta)^{3/2} + 1/2 = 0.$$

Therefore $\cos\theta = (5 - 6^{2/3})/4$ and the sign change occurs near 65°. Negative surface charge is closer to the positive real charge; the inverse-square field weights its attraction more strongly than repulsion by distant positive charge.

Further reading

- Griffiths, 5th ed., §§2.5, 3.1.6, 3.2.4 · conductors and image methods
- Fließbach, chapter 7 · boundary-value problems

Study the full derivation

Images in wedges, cylinders and dielectric interfaces · Electrostatic field theory and measurement

An image source is a construction for a particular physical domain and boundary condition. Three examples show how geometry determines the image position, how material response sets its amplitude, and why an apparent success on part of a boundary is insufficient.

Chapter 41 · How a circular wall selects Bessel modes

A voltage pattern is imposed on the end of a long grounded tube. Its penetration length depends on the tube radius through a number that looks unfamiliar in a Cartesian Fourier series. Deriving that number also explains the proper radial weight for projecting data.

Separate Laplace’s equation without flattening the circular wall

Take an empty cylinder of radius a extending from $z = 0$ to large positive z. Its side is grounded. A segmented end electrode imposes the continuous axisymmetric potential $V(r, 0) = V_0[1 - (r/a)^2]$, matching zero at the rim. In cylindrical coordinates $\nabla^2 V = r^{-1}\partial_r(r\partial_r V) + \partial_z^2 V = 0$. Set $V = R(r)Z(z)$. A decaying solution requires $Z = e^{-kz}$ with $k > 0$. Then

$$(rR')' + k^2 rR = 0.$$

Using $\xi = kr$ gives Bessel's equation of order zero. The regular solution $J_0(\xi)$ starts at one with zero derivative. Its series follows by substituting $R = \sum c_m \xi^{2m}$: the recurrence is $c_{m+1} = -c_m/[4(m+1)^2]$. The independent solution diverges logarithmically at the origin and would introduce an unmodelled line source. The side boundary requires $J_0(ka) = 0$, so $k_n = j_n/a$ with j_n the positive zeros of J_0. The first three are approximately $2.40483, 5.52008, 8.65373$.

Derive the radial weight and coefficients

Multiply the equation for R_m by R_n and subtract the reversed equation. Integration gives $(k_m^2 - k_n^2)\int_0^a r R_m R_n \, dr$ equal to a boundary term. At a both functions vanish; at zero their derivatives are regular and the factor r vanishes. Hence distinct modes are orthogonal with the radial weight r. That factor represents the area of a circular ring. The squared norm follows from the derivative identities $J_0' = -J_1$, $(tJ_1)' = tJ_0$, using the known zero at a:

$$\int_0^1 x J_0(j_n x)^2 \, dx = \frac{J_1(j_n)^2}{2},$$

$$b_n = \frac{2\int_0^1 x(1-x^2)J_0(j_n x)\,dx}{J_1(j_n)^2}.$$

Use $(xJ_1(jx))' = jxJ_0(jx)$ to integrate the numerator by parts. The endpoint factor $1 - x^2$ vanishes at $x = 1$; the remaining integral is

$$2j^{-1}\int_0^1 x^2 J_1(jx)\,dx = 2J_2(j)/j^2.$$

At a zero of J_0, $J_2(j) = 2J_1(j)/j$. Thus $b_n = 8/[j_n^3 J_1(j_n)]$, with alternating signs. The full solution is

$$V(r, z) = V_0 \sum_{n=1}^{\infty} b_n J_0(j_n r/a) e^{-j_n z/a}.$$

Each partial sum exactly grounds the side wall and decays toward infinity. The representation of the end electrode improves as modes are added. For a finite cylinder closed at $z = L$, replace each exponential by $\sinh[k_n(L - z)]/\sinh(k_n L)$. This satisfies both end conditions and tends to the long-cylinder result at fixed z as L grows.

Estimate a leakage length and a modal error

How far does the slowest mode survive?

For $a = 12$ mm the first modal decay length is

$$a/j_1 = 4.99\,\mathrm{mm}.$$

Far from the end, two axis probes separated by 5.0 mm should read voltages in ratio about $e^{-2.40483(5/12)} = 0.367$. The first coefficient $b_1 \simeq 1.108$ is larger than unity because the higher modes bring the prescribed central boundary value back to exactly V_0. A fit using $V_0 e^{-j_1 z/a}$ would get the correct distant slope but the wrong amplitude.

Predict the accuracy of a one-mode model

The second coefficient is about -0.140. On axis the magnitude of its contribution relative to the first is approximately $0.126\,e^{-(5.520-2.405)z/a}$. At $z = a$ this is 0.56%. At $z = 0$ higher modes cannot be neglected at the same tolerance. Near a radial node of a chosen reference mode a relative error becomes unstable; use absolute potential error measured in units of V_0 there.

Reconstruct a circular electrode pattern with Bessel modes

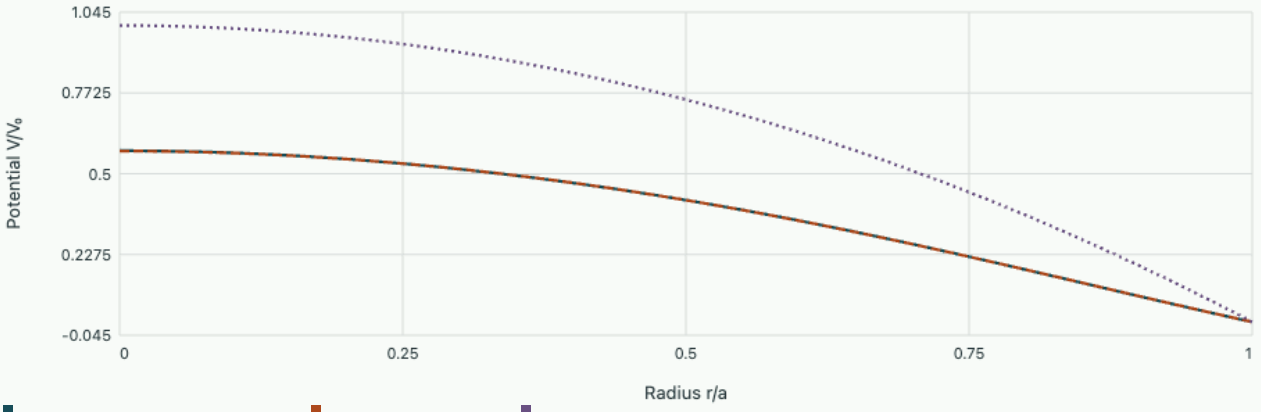
Each retained mode exactly satisfies the side-wall condition. At the end, one mode alone overestimates the central voltage by about 10.8%. Move the detector deeper into the tube and compare three and eight modes until their interior traces coincide to the precision you need.

Open this experiment online

Try the calculation

Height z/a **0.25**

Retained modes **3**



Selected modal sum Eight modes Imposed end voltage

The on-axis potential is 0.5774 V_0 with the selected modes, versus 0.5765 V_0 using eight. The slowest mode has e-folding length 0.4158 a.

Problem sheet · source weighting and a radial inverse problem

Explain a nonzero axis slope in a file of radial data

Someone fits the central potential by $V = A + Br$ near the cylinder axis. What does a nonzero B imply in the continuum problem?

▼ Work through the solution

A smooth axisymmetric Cartesian scalar must have zero radial derivative at the axis. The term Br would give $r^{-1}\partial_r(r\partial_r V) = B/r$, incompatible with a smooth source-free interior and finite homogeneous axial curvature. An interpolation can manufacture this feature if it treats the axis as an arbitrary material boundary.

Infer tube radius using only distant measurements

Two distant on-axis probes separated by 3.00 mm have a voltage ratio of 0.500. Infer a under first-mode dominance, then design a check against residual higher modes.

▼ Work through the solution

From $e^{-j_1\Delta z/a} = 1/2$, $a = j_1\Delta z/\ln 2 = 10.4$ mm. Move both probes farther from the end by the same amount. A persistent ratio supports the one-mode approximation. If higher modes remain, their changing

relative weight changes the ratio. Very small voltages may instead expose the noise floor and lead to an apparently shallower decay.

Further reading

- Fließbach, chapter 10 · cylindrical boundary-value problems
- Griffiths, 5th ed., §3.3 · separation of variables

Study the full derivation

Cylindrical geometry, coaxial boundaries and Bessel spectra · Electrostatic field theory and measurement

The radial coordinate of a cylinder leads to logarithms in one experiment and to Bessel functions in another. The difference is set by the variation along the axis and around the circumference. Recovering the radial weight from flux clarifies both.

Chapter 42 · Forces and torques from a matrix of flux linkages

Two measurements made while holding current fixed and flux fixed can produce different torque curves. They agree on the torque of the same physical state. A derivation from electrical work identifies precisely what must be held fixed when differentiating.

Keep track of each electrical and mechanical port

Let θ be the mechanical angle of a pair of coils, i_j the currents entering their positive terminals and λ_j their linked fluxes, including all turns. Use ideal windings, a reversible magnetic medium and negligible radiation. Faraday's terminal relation is

$$v_j = d\lambda_j/dt.$$

With positive torque T doing work on the moving rotor, energy balance reads

$$dW_m = \sum_j i_j \, d\lambda_j - T \, d\theta.$$

The natural variables of magnetic energy W_m are fluxes and position. Consequently $i_j = (\partial W_m/\partial \lambda_j)_\theta$ and $T = -(\partial W_m/\partial \theta)_\lambda$. A current-controlled experiment supplies or removes electrical work during motion. Define the coenergy $W'_m = \sum i_j \lambda_j - W_m$. Its differential is $\sum_j \lambda_j \, di_j + T \, d\theta$, giving $T = (\partial W'_m/\partial \theta)_i$. Coenergy is a Legendre transform. Its numerical value equals stored energy for linear zero-bias materials but can differ for nonlinear response.

Derive reciprocity, positivity and torque

Take $\boldsymbol{\lambda} = L(\theta)\mathbf{i}$ with a two-port inductance matrix. From the existence of W'_m, mixed derivatives at fixed angle commute, so

$$\partial \lambda_1/\partial i_2 = \partial \lambda_2/\partial i_1.$$

This gives $L_{12} = L_{21} = M$ for the given terminal orientation. With positive self-inductances, positive stored energy requires $L_1 L_2 - M^2 > 0$ for independent currents in a nonsingular model. Integrating coenergy at fixed angle gives

$$W'_m = \mathbf{i}^T L \mathbf{i}/2.$$

The energy in its own natural variables is $W_m = \boldsymbol{\lambda}^T L^{-1} \boldsymbol{\lambda}/2$. Differentiate $L^{-1}L = I$ to obtain

$$(L^{-1})' = -L^{-1} L' L^{-1}.$$

Thus both torque formulas at the same state give

$$T = \tfrac{1}{2} \mathbf{i}^T \frac{\partial L}{\partial \theta} \mathbf{i}.$$

For constant self-inductances and $M = M_0 \cos\theta$,

$$T = -M_0 i_1 i_2 \sin\theta.$$

Reversing one winding reverses the coordinate definitions of its current and mutual inductance together, preserving the physical result. Under prescribed fluxes the same formula holds with currents $\mathbf{i} = L(\theta)^{-1}\boldsymbol{\lambda}$ recalculated at every angle.

Compare a current supply with two flux-conserving loops

Start at the same state, then turn the coils

Use $L_1 = 4.0$ mH, $L_2 = 9.0$ mH and $M_0 = 3.0$ mH. At $\theta_0 = 60^\circ$, let currents be 2.00 A and 1.00 A. The flux linkages are 9.50 and 12.0 mWb. At this configuration $T = -5.20$ mN m and the energy is 15.5 mJ. The units follow from an inductance times current squared, which is energy, while angle in radians is dimensionless.

Move to 90°. Keeping both currents fixed gives $T = -6.00$ mN m. Keeping both fluxes fixed instead gives $i_1 = 9.50/4.0 = 2.375$ A and $i_2 = 12.0/9.0 = 1.333$ A because the mutual term has vanished. The new torque is -9.50 mN m. The two paths through state space now represent genuinely different experiments. For the fixed-current path the field energy falls by 3.00 mJ. The sources absorb 6.00 mJ and the external mechanical agent supplies 3.00 mJ, consistent with the first law.

A nonlinear constitutive curve

For one winding model a reversible flux linkage $\lambda = L_\ell i + \Lambda \tanh(i/I_s)$, where positive L_ℓ, Λ, I_s may depend on geometry. Integration gives

$$W_m' = L_\ell i^2/2 + \Lambda I_s \ln\cosh(i/I_s);$$

$$W_m = i\lambda - W_m'.$$

Only in the linear weak-current limit do they coincide. A measured hysteresis loop would make this single-valued construction invalid without additional state variables describing material history.

Compare turning a rotor at fixed current and at fixed flux

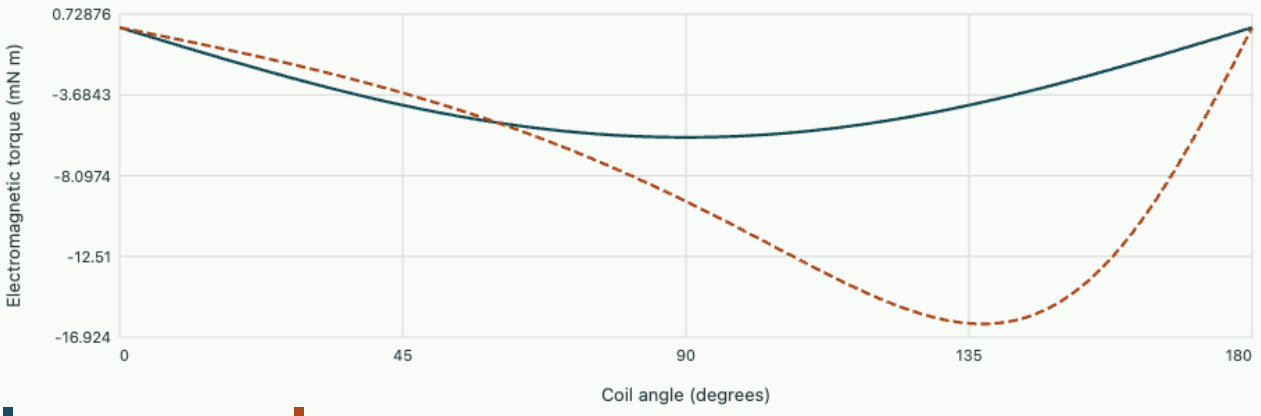
Both protocols start at 60°. With initial currents 2 A and 1 A, the torques there coincide at –5.20 mN m but reach –6.00 and –9.50 mN m at 90°. Reverse one initial current and identify the terms changed by the corresponding constraints.

Open this experiment online

Try the calculation

First current at preparation (A) 2

Second current at preparation (A) 1



At 60° both preparations give -5.196 mN m. At 90° the current-controlled and flux-controlled torques are -6 and -9.5 mN m.

Problem sheet · independent torque measurements

Test a computed torque numerically

How would a finite-difference experiment recover torque in the preceding fixed-flux state without differentiating any matrix entries symbolically?

▼ **Work through the solution**

Take the known $\boldsymbol{\lambda}$, invert L at $\theta \pm h$ separately and compute $W_{\pm} = \boldsymbol{\lambda}^T L_{\pm}^{-1} \boldsymbol{\lambda}/2$. Estimate $T = -(W_+ - W_-)/(2h)$ using h in radians. At 90° the result converges to -9.50 mN m. Reusing the same currents at both perturbations silently changes what the finite difference measures.

Infer the size of the conservative part of coupling

At fixed angle a small current cycle on the i_1, i_2 plane produces $\oint (\lambda_1 di_1 + \lambda_2 di_2) \neq 0$ after correcting lead resistance. What does this say about a proposed coenergy function?

▼ **Work through the solution**

An exact differential has zero integral around every closed path. A repeatable nonzero result means that a single-valued coenergy of the two currents alone cannot describe these measurements. Check frequency-dependent losses, hysteresis and whether an unmeasured mechanical coordinate moved. Making the inductance matrix symmetric by averaging two readings alone cannot explain the residual work.

Further reading

- Griffiths, 5th ed., §§7.2.3–7.2.4 · mutual inductance and magnetic energy
- Cambridge Electromagnetism, chapter 3 · induction and magnetic energy

Study the full derivation

Flux linkage, coupled coils and magnetic coenergy · Magnetism, induction and field dynamics

A winding links flux through each turn. In two windings the current in either one can change the voltage in the other. A matrix of flux linkages describes this coupling, but predicting power or mechanical force requires specifying the current directions and the control conditions.

Chapter 43 · Where current and energy go in a charging capacitor

A wire carries current to the centre of a plate, yet charge appears across its face and the energy increases in the gap. Following that current over differently chosen surfaces gives a spatial test of charge conservation and an independent power budget.

A circulation must give one answer for every spanning surface

Surround the charging lead by a loop. A flat surface inside the lead intercepts conduction current; a surface stretched into the insulating gap can avoid the lead altogether. Applying the magnetostatic law to both would assign different circulation to the same loop. Charge continuity and Gauss's law instead give

$$\nabla \cdot (\mathbf{J} + \epsilon_0 \partial_t \mathbf{E}) = -\partial_t \rho + \partial_t \rho = 0.$$

Consequently the flux of this sum through any two surfaces with the same boundary agrees. Ampère–Maxwell equates the circulation of $\mathbf{B}/\mu_0$ to that common enclosed current in vacuum. In a material with identified free current the macroscopic balance uses $\mathbf{H}, \mathbf{J}_f, \partial_t \mathbf{D}$, incorporating polarisation. The divergence identity verifies consistency; the magnetic field generated by the additional term is a physical prediction of Maxwell's equation.

Resolve the charging current across the plate

Take circular metal plates of radius a, separation $d \ll a$, connected at their centres. Work at times long compared with a/c and neglect small edge and lead corrections. In the gap $E_z = V/d$ is approximately uniform with sign chosen from positive to negative plate. $Q = \epsilon_0 \pi a^2 E_z$, so $\partial_t D_z = I/(\pi a^2)$ where $I = \dot{Q}$. A circle of radius $r < a$ spanning the gap intercepts displacement current $I r^2/a^2$. Azimuthal symmetry gives

$$B_\phi(r) = \frac{\mu_0 I r}{2\pi a^2} \quad (r < a),$$

$$B_\phi(r) \simeq \frac{\mu_0 I}{2\pi r} \quad (r > a).$$

The exterior expression assumes the return conductor lies farther away and retains the leading large-plate approximation. Magnetic field is continuous at the idealised rim. Now cut a ring of radius r into the positive plate. The portion of supplied current needed to charge its central disk is $I r^2/a^2$; the rest crosses that ring outwards as conduction current. Thus $2\pi r K_r = I(1 - r^2/a^2)$. Its apparent $1/r$ singularity at the centre is rounded off by the finite contact patch. Distributed charging is compatible with current entering through one point.

Locate the power flow in the gap

The Poynting vector in the gap is

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} = -E_z H_\phi \hat{\mathbf{r}}.$$

For increasing field it points radially inward. Integrate the inward flux through the cylindrical side of a smaller pillbox of radius $r < a$ and thickness d:

$$P_{\mathrm{in}}(r) = 2\pi r d E_z H_\phi(r) = VI\,\frac{r^2}{a^2}$$

$$\frac{d}{dt}\left(\frac{\epsilon_0}{2}E_z^2 \pi r^2 d\right).$$

Thus energy crosses the side before being stored throughout that subvolume. At $r = a$ the result equals the terminal power VI. This derivation uses the leading quasistatic fields. A rapidly changing current also produces induction fields and time-varying magnetic energy; in that regime the full Poynting theorem includes them. During discharging, $\dot{Q}$ reverses and the direction of radial power flow reverses.

Predict two independent instrument readings

Sample midway to the rim

Let $a = 20$ mm, $d = 0.50$ mm and $I = 4.0$ mA. At $r = 10$ mm a gap circle encloses 1.0 mA displacement current. Its magnetic field is 20.0 nT. The outward conduction current crossing the same radius in the positive metal is 3.0 mA, corresponding to $K_r = 0.0477$ A/m. The gap reading and the plate reading describe different surfaces; their sum recovers the lead current in this construction.

Record a finite charge ramp

The ideal capacitance here is 22.3 pF. A constant $I = 4.0$ mA for 20.0 ns raises the voltage by 3.60 V from zero. The source supplies $\int VI\,dt = CV_{\mathrm{final}}^2/2 = 144$ pJ. At the half-radius pillbox only one quarter of this energy appears, 36.0 pJ. A real pulse has finite rise and fall times. The simple uniform-field solution is accurate away from abrupt switching transients when the spectral content is well below c/a.

Measure a displaced current and a radial conduction current separately

At 20 mm plate radius, 4 mA lead current and a 10 mm probe radius, 1 mA crosses the gap disk and 3 mA crosses the metal ring. The graph compares both currents over the entire plate. Reverse the current or enlarge the plate and compare the magnetic reading at the selected fractional radius.

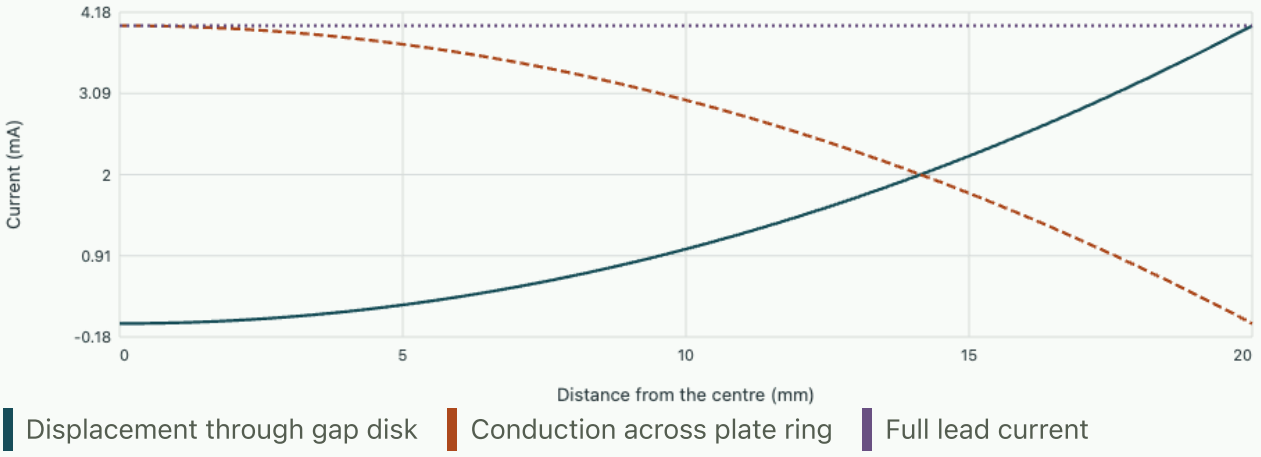
Open this experiment online

Try the calculation

Plate radius (mm) **20**

Charging current (mA) **4**

Measurement radius / plate radius **0.5**



At r = 10 mm the gap encloses 1 mA and the radial plate current is 3 mA. The local azimuthal magnetic field is 20 nT.

Problem sheet · challenge a misleading measurement

Determine the sign during reversal

The capacitor field still points from the positive to the negative plate, but its magnitude is decreasing. Which of B_ϕ, S_r, E_z changes sign relative to an increasing field with the same instantaneous voltage?

▼ **Work through the solution**

The voltage and E_z keep their signs. $I = C\dot{V}$ reverses, so the azimuthal magnetic field reverses. Hence radial energy flow reverses and leaves the pillbox. A measurement of voltage alone does not determine the direction of energy transport.

What changes if the centre contact is replaced by a rim feed?

Assume a uniform annular feed deposits charge symmetrically at the outer edge. Predict the signed metal current crossing a circle of radius $r < a$ if positive is radially outward.

▼ **Work through the solution**

All charge accumulating inside the test circle must now enter from larger radii. The inward flux is Ir^2/a^2, so $2\pi r K_r = -Ir^2/a^2$. The gap displacement current for the same uniform charging history is

unchanged. The external lead and return geometry determines the remainder of the global magnetic field configuration.

Further reading

- Griffiths, 5th ed., §§7.3.1–7.3.3, 8.1 · displacement current and energy
- Cambridge Electromagnetism, chapter 3 · displacement current

Study the full derivation

Displacement current and the charging capacitor · Magnetism, induction and field dynamics

A wire feeds a capacitor whose dielectric blocks the transport of mobile charges between its plates. A magnetic probe near the gap still responds to charging. Resolving which surface a circulation law should use leads to displacement current, an electric field correction during a ramp, and an energy-flow test around the rim.

Chapter 44 · A lossy half-space seen through its surface impedance

An absorbing medium transmits a complex field and converts electrical work into heat. Matching one complex number at its surface predicts reflection. Integrating the heating through its depth provides an independent check that is especially revealing near the crossover between dielectric and conductor.

Choose a branch that decays into the material

Let a stationary plane wave enter the homogeneous half-space $z > 0$ at normal incidence. Use $e^{-i\omega t}$ and fields proportional to e^{ikz}, real $\epsilon, \mu > 0$ and Ohmic $\mathbf{J} = \sigma\mathbf{E}$, $\sigma \geq 0$. Ampère's law combines the two currents into $-i\omega(\epsilon + i\sigma/\omega)\mathbf{E}$. Faraday's law yields

$$H_y = kE_x/(\omega\mu).$$

Eliminating H gives

$$k^2 = \omega^2\mu\epsilon(1 + it),$$

$$t = \frac{\sigma}{\omega\epsilon},$$

$$k = \beta + i\alpha.$$

Choose $\beta > 0, \alpha \geq 0$ so the interior field weakens with depth. Let $k_d = \omega\sqrt{\mu\epsilon}$. Equating real and imaginary parts gives $\beta^2 - \alpha^2 = k_d^2$, $2\beta\alpha = k_d^2 t$, and hence

$$\frac{\beta}{k_d} = \sqrt{\frac{\sqrt{1 + t^2} + 1}{2}},$$

$$\frac{\alpha}{k_d} = \frac{t}{2(\beta/k_d)},$$

$$Z = \frac{E_x}{H_y} = \frac{\omega\mu}{k}.$$

The rational expression for α avoids subtracting nearly equal numbers for small t. As $t \to 0$,

$$\alpha/k_d \simeq t/2.$$

For $t \gg 1$, $\beta \simeq \alpha \simeq \sqrt{\omega\mu\sigma/2}$: skin depth $1/\alpha$ measures field-amplitude decay and the power has the shorter e-folding depth $1/(2\alpha)$.

Read the complex reflection and the power entering the surface

Let the incident medium be lossless with impedance Z_1. Continuity of tangential E, H gives

$$1 + r = h,\ (1 - r)/Z_1 = h/Z,$$

where $h = E(0^+)/E_{\mathrm{inc}}$. Solving,

$$r = \frac{Z - Z_1}{Z + Z_1},$$

$$h = \frac{2Z}{Z + Z_1},$$

$$1 - |r|^2 = Z_1 \operatorname{Re}(1/Z^*)|h|^2.$$

The flux relation follows by dividing the real part of $EH^*/2$ by the incident intensity. It includes a real admittance and is generally different from $|h|^2$. The imaginary part of the complex Poynting vector measures exchange in quadrature with the drive. In the good-conductor limit $Z \simeq (1-i)\sqrt{\omega\mu/(2\sigma)}$. Its positive real part is essential for passive heating, even where the reflected voltage is almost the negative of the incident one.

Recover the entire inward power from a volume measurement

Ohmic heating per volume is $\langle p_J \rangle = \sigma |E(0)|^2 e^{-2\alpha z}/2$. Integrating from zero to infinity gives

$$P_J/A = \sigma |E(0)|^2/(4\alpha).$$

Use $2\beta\alpha = \mu\omega\sigma$ to rewrite it as $|E(0)|^2\beta/(2\omega\mu)$. This equals the inward mean surface flux from $EH^*/2$. No separate efficiency factor is needed. The fraction of surface power dissipated before depth D is $1 - e^{-2\alpha D}$, so 90% of the heat is deposited by $D = \ln 10/(2\alpha)$, about 1.15 field skin depths.

Compare two frequency regimes

A crossover with comparable currents

Consider material whose lossless impedance equals the incident medium's and whose loss ratio is $t = 2$. Then $\beta/k_d = 1.272$, $\alpha/k_d = 0.786$, and $Z/Z_1 = 0.569 - 0.352i$. Direct complex division gives a reflected power fraction of about 0.120. The balance enters the half-space and ultimately heats it. One lossless phase length k_d^{-1} into the material, the remaining forward power is $e^{-1.572} = 0.208$ of its surface value. This illustrates substantial heating even though most incident power enters the material.

An amplitude probe and a calorimeter

Suppose a field-amplitude probe reads 1/e of its surface value at depth 0.80 mm. Then $\alpha = 1250$ m⁻¹, the power half-depth is $\ln 2/(2\alpha) = 0.277$ mm and 90% of the absorbed energy is deposited by 0.921 mm. These predictions do not require knowing the incident intensity. Disagreement between the two types of measurement can reveal spatially varying σ or a reflected wave from a finite back surface.

Close a reflection and heating budget at every depth

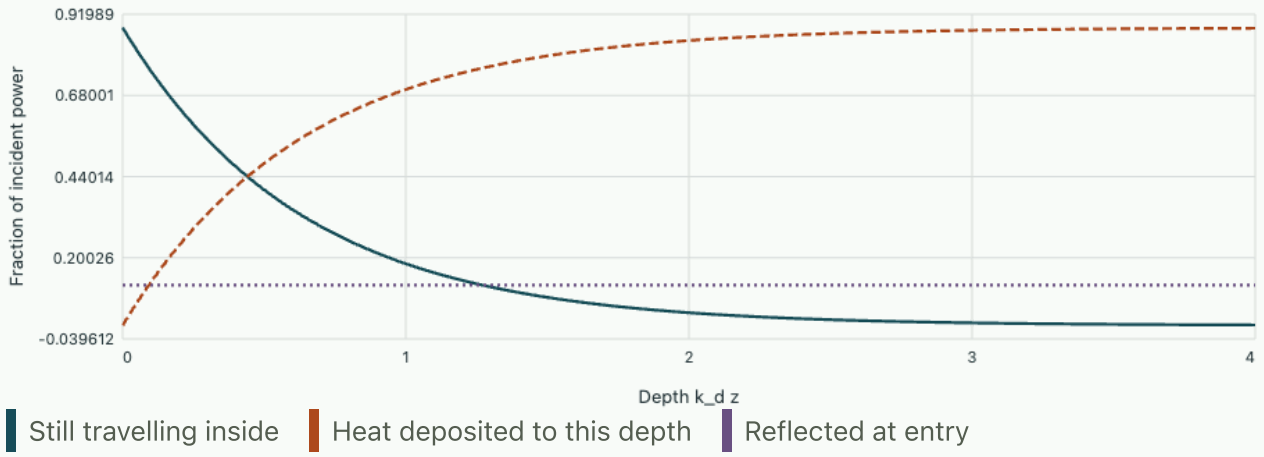
At loss ratio 2 and matched lossless impedances, about 12% of the incident power is reflected. The remaining power gradually becomes heat. Change loss and impedance independently and check that the surface reflection, transmitted flux at depth and integrated heat always add to one.

Open this experiment online

Try the calculation

Loss ratio σ/(ωε) **2**

Lossless impedance / incident impedance **1**



The face reflects 11.97% of incident power. The field attenuation α/k_d is 0.7862 and impedance relative to the incident medium is 0.5689 − 0.3516i.

Problem sheet · determine a material from two probes

Recover material parameters

At frequency f in a nonmagnetic half-space, phase measurements give β and amplitude measurements give α. Recover σ, ϵ, and specify the consistency condition for the dielectric model used here.

▼ **Work through the solution**

Solve for $\sigma = 2\beta\alpha/(\mu_0\omega)$ and $\epsilon = (\beta^2 - \alpha^2)/(\mu_0\omega^2)$. A medium with positive real permittivity in this simple Ohmic model requires $\beta > \alpha$. In the good-conductor limit their difference is small, making the inferred permittivity sensitive to measurement errors.

When does an infinite depth assumption fail?

For a slab of thickness D, the back-face field reflection magnitude is 0.50. Bound the return amplitude at the entrance relative to the forward wave inside, ignoring subsequent round trips. How thick keeps it

below 1%?

▼ **Work through the solution**

The round trip attenuates the field by $e^{-2\alpha D}$. Require $0.50e^{-2\alpha D} < 0.01$, giving $D > \ln 50/(2\alpha) = 1.96/\alpha$. Phase may suppress the measured amplitude at selected frequencies, but the magnitude bound does not depend on that cancellation.

Further reading

- Griffiths, 5th ed., §§9.4.1–9.4.2 · conducting media and reflection
- Fließbach, chapter 34 · dispersion and absorption

Study the full derivation

Causal material response and the propagation of a finite pulse · Waves, radiation and relativity

An optical pulse entering a transparent slab is delayed, broadened and sometimes reshaped. Derive these effects from the motion of bound charge, and learn what a measured group delay actually says about a pulse carrying information.

Chapter 45 · Two polarisations moving through a uniaxial crystal

A crystal can send two pulses in different directions even when their wavefronts have the same normal. The distinction emerges by applying the divergence condition to electric displacement and measuring energy flow with the Poynting vector.

A field transverse to the wavevector need not be the electric field

Let a transparent, nonmagnetic uniaxial medium have $\boldsymbol{\epsilon}/\epsilon_0 = \mathrm{diag}(n_o^2, n_o^2, n_e^2)$. The optic axis is z; dispersion over the bandwidth under study is negligible. For a harmonic plane wave with wavevector in the xz plane, write $\mathbf{k} = (k_x, 0, k_z)$. Source-free Maxwell equations require $\mathbf{k}\cdot\mathbf{D} = 0$, $\mathbf{k}\cdot\mathbf{B} = 0$,

$$\mathbf{k}\times\mathbf{E} = \omega\mathbf{B},$$
$$\mathbf{k}\times\mathbf{H} = -\omega\mathbf{D}.$$

Eliminating the magnetic field gives

$$\mathbf{k}\times(\mathbf{k}\times\mathbf{E}) + k_0^2(\boldsymbol{\epsilon}/\epsilon_0)\mathbf{E} = 0,$$
$$k_0 = \omega/c.$$

An electric field along y satisfies this system with $k^2 = k_0^2 n_o^2$. This ordinary mode has transverse $\mathbf{E}, \mathbf{D}$ and energy travels parallel to $\mathbf{k}$. A second mode has $\mathbf{E}$ in the xz plane; $\mathbf{D}$ is transverse but generally not parallel to $\mathbf{E}$.

Find the phase index and the direction of energy separately

The remaining two Cartesian equations form a matrix with diagonal $k_0^2 n_o^2 - k_z^2,\ k_0^2 n_e^2 - k_x^2$ and off-diagonal elements $k_x k_z$. Its determinant vanishes for nontrivial fields. Expanding cancels the fourth-order cross term and gives

$$\frac{k_x^2}{n_e^2} + \frac{k_z^2}{n_o^2} = k_0^2,$$
$$\frac{1}{n_{\mathrm{ph}}(\theta)^2} = \frac{\sin^2\theta}{n_e^2} + \frac{\cos^2\theta}{n_o^2},$$

where θ is the angle from the optic axis to the wavevector. Along the axis both modes have the same index; perpendicular to it the extraordinary phase index is n_e. For the latter mode choose transverse $\mathbf{D} \propto (\cos\theta, 0, -\sin\theta)$. Applying the inverse permittivity tensor gives

$$\mathbf{E} \propto (\cos\theta/n_o^2, 0, -\sin\theta/n_e^2).$$

The magnetic field is along y, so the energy flow angle ψ from z satisfies

$$\tan\psi = -\frac{E_z}{E_x} = \frac{n_o^2}{n_e^2}\tan\theta.$$

The group velocity obtained by differentiating $\omega(\mathbf{k}) = c\sqrt{k_x^2/n_e^2 + k_z^2/n_o^2}$ has the same direction in this nondispersive model. This independent construction is normal to the constant-frequency ellipse. At an interface it is the tangential component of $\mathbf{k}$, fixed by phase matching, that selects the transmitted wavevector. The energy path can then diverge from the corresponding wave normal.

A crystal plate as a beam displacer and as a retarder

A normally incident beam with an oblique optic axis

Let $n_o = 1.50,\ n_e = 1.60$ and let the plate normal make 45° to the optic axis. At normal incidence the phase normal is the plate normal for both modes. For the extraordinary wave $\psi = \arctan(2.25/2.56) = 41.3^\circ$. Relative to its wave normal the energy travels about -3.7° toward the optic axis. Across a 10.0 mm thick plate the lateral displacement relative to the ordinary beam is approximately 0.65 mm. This describes propagation inside a parallel plate and neglects edge clipping.

A plate that turns linear input into circular output

Propagate perpendicular to the optic axis so both eigenmodes travel in the same direction. Their accumulated phase difference is $\Delta = k_0(n_e - n_o)d$. Linear incident polarisation at 45° to the eigenaxes has equal amplitudes. When $|\Delta| = \pi/2$ modulo 2π, those orthogonal components have a quarter-period phase difference and form circular polarisation, assuming equal transmitted amplitudes. At 600 nm with $n_e - n_o = 0.10$, the smallest such thickness is 1.50 µm. A larger multi-order plate can give the same retardance at this wavelength but responds more strongly to bandwidth and thickness errors.

Recover retardance from a pair of polarisers

Place the crystal axes at 45° to the input polariser and use a second polariser crossed with the first. Express the output field in the two crystal axes, then project onto the crossed analyser; its amplitude is proportional to $(e^{i\Delta} - 1)/2$. Thus the transmitted fraction is $\sin^2(\Delta/2)$. An arbitrary phase offset common to both modes disappears. A single intensity reading determines neither the sign of Δ nor its full order. Varying thickness or scanning wavelength can resolve the branch; the material dispersion then enters the interpretation.

Predict the transverse displacement of a crystal plate

Choose principal indices 1.50 and 1.60 and a 45° wave normal. The extraordinary beam is displaced by about –0.65 mm through a 10 mm plate; the ordinary beam follows the normal. Swap the principal indices and compare the sign and magnitude of walk-off.

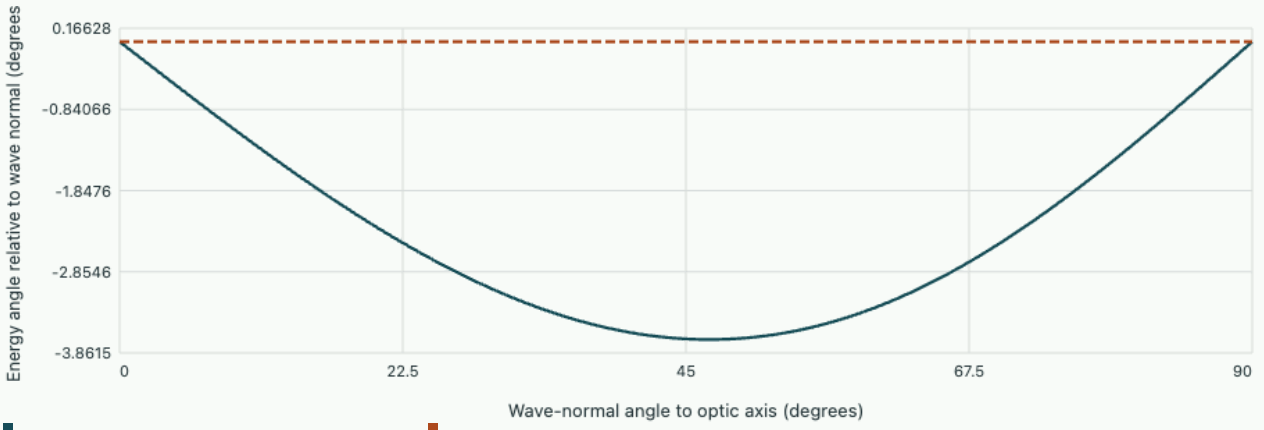
Open this experiment online

Try the calculation

Ordinary principal index 1.5

Extraordinary principal index 1.6

Probe wave-normal angle (degrees) 45



At the chosen angle the extraordinary phase index is 1.548. A 10 mm plate normal to the wavevector displaces its energy by -0.6445 mm relative to the ordinary ray.

Problem sheet · use energy and phase observations

Infer birefringence from a walk-off angle

At $\theta = 45^\circ$, an extraordinary energy ray makes $\psi = 40^\circ$ to the optic axis. Assuming the same model, recover n_e/n_o. Is the medium positive or negative uniaxial?

▼ **Work through the solution**

The ratio is

$$\sqrt{\tan 45^\circ / \tan 40^\circ} = 1.092.$$

The extraordinary principal index exceeds the ordinary one, a positive uniaxial material. A measurement referenced to the plate normal must first be converted to the angle from the optic axis used in the equation.

Calculate the contrast loss caused by angular misalignment

An ideal half-wave plate has its eigenaxis at angle α to the input polariser. Keep the analyser crossed and find the intensity transmitted when the angle is 40° instead of 45°.

▼ **Work through the solution**

Projection into the crystal axes and back out gives

$$I/I_0 = \sin^2(2\alpha)\sin^2(\Delta/2).$$

At half-wave retardance the factor from phase is unity, and the misaligned transmission is

$$\sin^2 80^\circ = 0.970.$$

The reduction alone cannot distinguish angular error from a small error in retardance.

Further reading

- Griffiths, 5th ed., §§4.3–4.4 and 9.1.4 · constitutive laws and polarisation
- Fließbach, chapter 33 · wave solutions in media

Study the full derivation

What an analyser measures about polarisation and coherence · Waves, radiation and relativity

Two red beams produce the same reading through every rotating linear polarizer. One might be circularly polarized and the other an incoherent mixture. Build the measurements that distinguish them, derive the visibility of their interference fringes, and decide which information is lost on a slow detector.

Chapter 46 · Radiation when the electric dipole cancels

Two optical dipoles can cancel in the long-wavelength electric-dipole approximation and still radiate. Expanding their source phase shows which moments survive, where the remaining light goes, and how a small asymmetry restores a different angular pattern.

Keep the source phase inside the radiation integral

For a localised harmonic current $\mathbf{J}(\mathbf{x}')e^{-i\omega t}$, take a detector far compared with source size and wavelength. In the outgoing potential set $1/R \simeq 1/r$, while retaining $kR \simeq kr - k\mathbf{n}\cdot\mathbf{x}'$, where $\mathbf{n} = \mathbf{r}/r$. This approximation requires the quadratic path phase across the source to be negligible. The radiative electric field is proportional to the transverse part of the current integral

$$\mathbf{C}(\mathbf{n}) = \int \mathbf{J}(\mathbf{x}')e^{-ik\mathbf{n}\cdot\mathbf{x}'}\,d^3x',$$

$$\mathbf{E}_{\mathrm{rad}} = \frac{i\mu_0\omega e^{ikr}}{4\pi r}[\mathbf{C} - \mathbf{n}(\mathbf{n}\cdot\mathbf{C})].$$

For small source size a with $ka \ll 1$, expand the exponential. Continuity and integration by parts give $\int J_i\,d^3x = \dot{p}_i$ for electric dipole $\mathbf{p} = \int \rho\mathbf{x}\,d^3x$. The first correction is $-ikn_j\int J_i x_j\,d^3x$. Split this integral into its symmetric and antisymmetric parts. Define $Q_{ij} = \int \rho x_i x_j\,d^3x$ and $\mathbf{m} = \frac{1}{2}\int \mathbf{x}\times\mathbf{J}\,d^3x$. Differentiating Q_{ij} and using charge conservation yields

$$\int J_i x_j\,d^3x = \tfrac{1}{2}\dot{Q}_{ij} - \epsilon_{ijk}m_k,$$

$$\mathbf{C} = \dot{\mathbf{p}} - \frac{ik}{2}\dot{Q}\,\mathbf{n} + ik\,\mathbf{n}\times\mathbf{m} + \cdots.$$

The trace of Q contributes a vector along $\mathbf{n}$; the far-field transverse projection removes it. Thus only its traceless electric quadrupole part can radiate. The antisymmetric contribution is the magnetic dipole. Both occur at the next order in this source-phase expansion. Their importance depends on source symmetry and on the actual sizes of their transition moments.

Construct an electric quadrupole with two displaced dipoles

Put narrow identical z-directed electric dipoles at $z = \pm d/2$ and drive them with moments $+p$ and $-p$, perfectly out of phase. Their total electric dipole is zero. The currents are collinear with their displacement, so this geometry also gives no net magnetic dipole. In direction θ to the axis the single-dipole field has factor $\sin\theta$. The additional two-source factor is $e^{-ikd\cos\theta/2} - e^{ikd\cos\theta/2}$. In units of the equatorial intensity of one dipole the complete pattern is

$$F(\theta) = 4\sin^2\theta\,\sin^2\!\left(\tfrac{1}{2}kd\cos\theta\right) \simeq (kd)^2\sin^2\theta\cos^2\theta.$$

This has nodes on the axis from transversality and in the equatorial plane from source cancellation. The leading lobes peak at 45° and 135°. Integrating over solid angle, $\int_{-1}^{1}(1-u^2)u^2\,du = 4/15$; the single-dipole integral is $4/3$. The pair-to-single total-power ratio is $(kd)^2/5$ at leading order. A fixed single-dipole amplitude is the reference; normalising each curve to its own maximum would hide this suppression.

Relate the symmetry argument to optical transitions

In the long-wavelength quantum interaction the electric-dipole term contains a matrix element of position, odd under spatial inversion. Between two states of the same definite parity, the integrand for an electric dipole moment changes sign on inversion and integrates to zero. The electric quadrupole contains two position factors and has even parity. The magnetic dipole is an axial vector and also has even parity. Either next-order channel can therefore satisfy parity between such states, subject to angular-momentum and other selection rules. Classical source symmetries illuminate the long-wavelength expansion; actual atomic emission rates require the relevant quantum matrix elements and state populations.

Estimate the strength and detect imperfect cancellation

An optical-scale separation

For $d = 20$ nm and $\lambda = 600$ nm, $kd = 0.2094$. The leading integrated pair-to-single power ratio is 0.00877, under 1%. In the 45° direction the dimensionless intensity is $4(1/2)\sin^2[kd/(2\sqrt{2})] \simeq 0.01095$. Tripling d at fixed wavelength should initially give roughly nine times the power; check the exact sine before applying that extrapolation.

An equatorial leak

Reduce the opposing dipole magnitude to $(1-\delta)p$. Expanding the squared two-source factor gives $\delta^2 + 4(1-\delta)\sin^2(kd\cos\theta/2)$. At the equator the second term is zero, leaving intensity fraction δ^2 relative to one dipole. A measured fraction 0.0025 corresponds to a 5% amplitude imbalance if this is the only imperfection. A phase mismatch can also fill the node, so the observation cannot distinguish both mechanisms without additional phase information.

Break the symmetry of an optical dipole pair

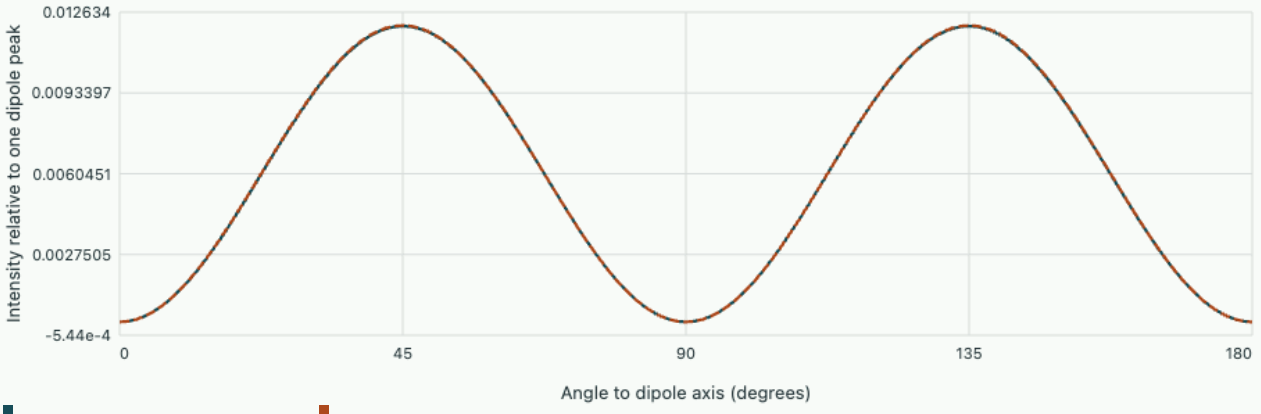
Compare a coherent two-source calculation against its first long-wavelength approximation. At d/λ = 0.035, equal and opposite dipoles emit about 1% of the power of a single dipole. Introduce a 5% amplitude mismatch and check the equatorial node. Then increase spacing to test the limit of the expansion.

Open this experiment online

Try the calculation

Separation d/λ **0.035**

Opposing amplitude deficit **0**



Full phase difference Long-wavelength expansion

Integrating over solid angle gives 0.9656% of the total power of one dipole. At 90° the intensity relative to its equatorial peak is 0.

Problem sheet · which moment radiates?

Separate the amplitude and power scaling

For the symmetric pair suppose the separation is halved with the two moments and frequency unchanged. Predict the leading power change at 45° and at 90°. Where is a ratio meaningful?

▼ **Work through the solution**

At 45° the leading field is proportional to kd, so the intensity drops by a factor of four. At 90° the exact two-source phase difference is zero for every spacing. Both intensities vanish there and their ratio is undefined. A finite measurement at that angle probes symmetry breaking or background light.

Compare with a closed current loop

A small circular loop carries uniform oscillating current I and has area A. Identify its leading moment and the directions in which its radiation vanishes.

▼ **Work through the solution**

The net electric dipole of this ideal divergence-free closed current is zero. Its magnetic dipole is $IA\hat{\mathbf{n}}_{\text{loop}}$. The radiation term contains $\mathbf{n} \times \mathbf{m}$, hence vanishes along the loop axis and peaks transverse to it. At leading order the loop can radiate strongly in its equatorial plane, unlike the collinear electric-quadrupole pair.

Further reading

- Griffiths, 5th ed., §§11.1.3–11.1.4 · magnetic dipoles and general sources
- Fließbach, chapter 24 · dipole radiation

Study the full derivation

Radiating multipoles and what a detector actually integrates · Waves, radiation and relativity

A small oscillating source looks point-like from a distant detector, yet its angular pattern reveals its structure. Derive electric and magnetic dipole radiation from a conserved current, identify what survives when electric dipoles cancel, and turn a detector aperture into a calibrated estimate of total power.

Chapter 47 · Four contacts on a conductor of finite thickness

A voltmeter scarcely draws any current, yet its reading depends on how the current from a separate pair of leads spreads inside the specimen. Derive the conversion from a four-terminal reading to resistivity. Decide from the geometry whether the result describes a bulk material or a conducting sheet.

What the two unused current terminals buy us

Place four point contacts on a straight line on the top of a uniform conducting sample. Drive current $+I$ into the first contact and remove it at the fourth. Attach an ideal high-impedance voltmeter to the second and third contacts. Equal probe spacing is called s; the sample thickness is h. Record $V = \phi_2 - \phi_3$ with the sign convention that positive conventional current flows toward lower electric potential. We seek the resistance of the actual measurement V/I, rather than the terminal resistance between the current contacts.

The voltage lost across a current contact can be large because all the drive current passes through it. That loss does not enter the ideal difference between the inner probe potentials. Likewise a contact resistance in series with an ideal voltmeter has zero IR drop. The separation does not eliminate all geometric uncertainties. A voltage contact with appreciable area can average the field or perturb it, a voltmeter with finite input impedance draws a small current, and an edge in the conducting specimen redirects current. Derive first what an ideal experiment would do before introducing corrections for a manufactured sample.

In the ohmic, low-frequency limit $\mathbf{J} = -\sigma \nabla \phi$ with constant scalar conductivity $\sigma = 1/\rho$. Conservation in the conducting interior means $\nabla \cdot \mathbf{J} = 0$. A point current lead on an otherwise insulating surface is a boundary flux. Unlike a prescribed voltage, it fixes the strength of a source without fixing its absolute potential. Only currents and differences should survive an arbitrary additive constant. This will be a useful test when we introduce infinitely many reflected current sources.

A spreading experiment in three dimensions

First make the sample so thick that the lower face is inaccessible to the current on the scale of the probes. Around a point input on the surface of a conducting half-space consider a hemisphere of radius r. All its outward current is I; rotational symmetry plus reflection across the ideal upper plane produces a radial field of magnitude $J = I/(2\pi r^2)$. Integrate $E_r = \rho J = -d\phi/dr$ inward from infinity. The surface potential relative to infinity is $\rho I/(2\pi r)$. Subtract a second copy centred on the sink.

At probe 2 the distances from source and sink are s and $2s$, respectively. Its potential is $+\rho I/(4\pi s)$; symmetry gives the opposite sign at probe 3. Therefore

$$R_{\text{bulk}} = \frac{V}{I} = \frac{\rho}{2\pi s}.$$

A bulk resistivity in ohm metres can be extracted by multiplying the transfer resistance by $2\pi s$ in metres. The inverse spacing is not a contact effect. It is the result of spreading current over an area that grows quadratically with radius. In particular, doubling s halves the measured resistance of a sufficiently thick sample. This scaling is a check independent of the calibration of a fixed current source as long as it remains stable across both measurements.

Confine the same current to a thin sheet

If current rapidly becomes uniform through a thin specimen, the spreading area is instead the circumference of a circle times h. Thus $J_r = I/(2\pi r h)$. The integral of $1/r$ is logarithmic and contains an arbitrary reference length. A single source in an infinite sheet cannot be assigned zero potential at infinity by a finite constant. The source–sink pair poses no difficulty because the reference cancels from the potential difference. With $R_s = \rho/h$ and source and sink distances r_+ and r_-, the sheet potential can be written as

$$\phi = \frac{I R_s}{2\pi} \ln \frac{r_-}{r_+},$$

$$R_{\mathrm{thin}} = \frac{R_s \ln 2}{\pi}.$$

At the two inner contacts the ratios are two and one half. The logarithms are equal and opposite, which supplies the extra factor of two. Doubling the probe spacing now leaves the measured transfer resistance unchanged. The name ohms per square refers to R_s, the resistance across opposite complete edges of a square with uniform flow; a four-point probe is a different arrangement and its resistance is smaller by $\ln 2/\pi$. Attaching the word square to a current probe reading without this conversion would produce a systematic error.

Reflect the current from an insulating lower face

A real film has an upper surface and a back surface. Assume both prevent normal current except at the upper contacts. Reflecting a source with the same sign across an insulating plane cancels the normal derivative there. Continuing the reflection between both parallel faces generates current images displaced vertically by $2nh$. A grounded conducting backing would impose a different boundary condition and cannot be represented by these signs. On the top surface the half-space spreading result therefore acquires a sequence of image distances. Define $z = s/h$. After subtracting the source and sink contributions at the inner probes one obtains

$$\frac{h}{\rho}\frac{V}{I} = \frac{1}{\pi}\left[\frac{1}{2z} + 2\sum_{n=1}^{\infty}\left(\frac{1}{\sqrt{z^2+4n^2}} - \frac{1}{\sqrt{4z^2+4n^2}}\right)\right].$$

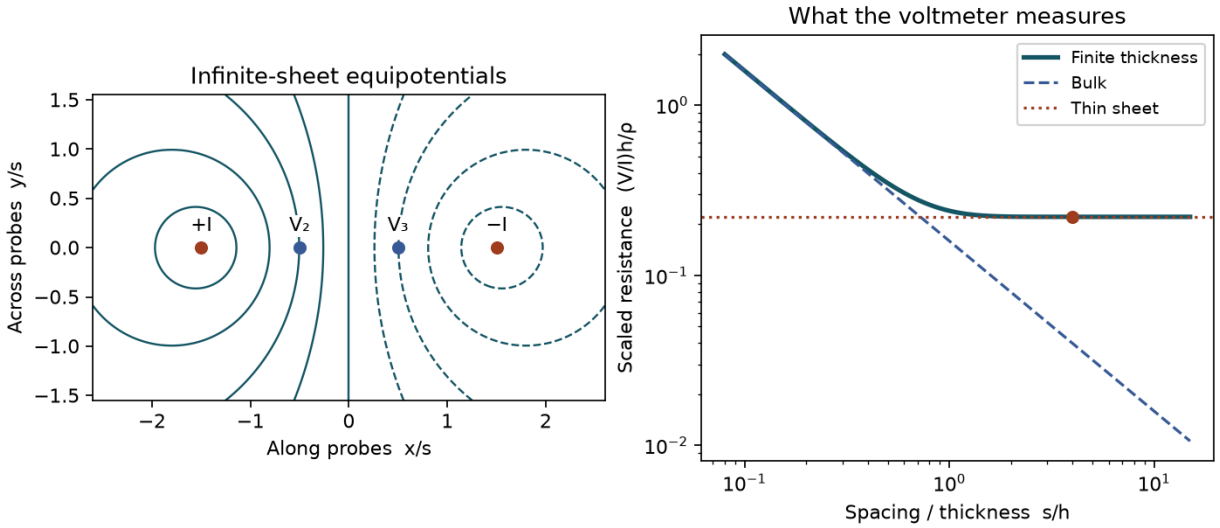
Every term in the remaining sum is positive: a nearby reflecting back increases the measured resistance relative to the infinite bulk expression. Individual absolute image potentials have slowly diverging harmonic tails; the paired differences converge. To see why, use $1/a - 1/b = (b^2 - a^2)/[ab(a+b)]$. With the two square roots above the numerator is exactly $3z^2$, independent of n, while the denominator grows as n^3. This form suppresses subtractive cancellation and makes an image-cutoff refinement test meaningful.

The thin-film limit can also be read from the sum without the original sheet argument. For large z the image spacing is small compared with the distances. Replace the sum over positive and negative images by an integral along the vertical coordinate with density $1/(2h)$. Integrating the difference of two inverse square roots gives the difference of two inverse hyperbolic sines. At a large upper limit only the ratio of the lateral distances survives. Their ratio is two, returning $\ln 2$. This second derivation identifies thickness

averaging as a long-distance limit of a three-dimensional field, rather than a separate electrical transport law.

Open editable notebook Download .ipynb

Use two spacings to test a material interpretation



The two inner probe positions lie on opposite sides of the zero-potential bisector. The solid curve in the right panel includes reflection of the current at the insulating back surface. Only the spacing on the horizontal scale was changed.

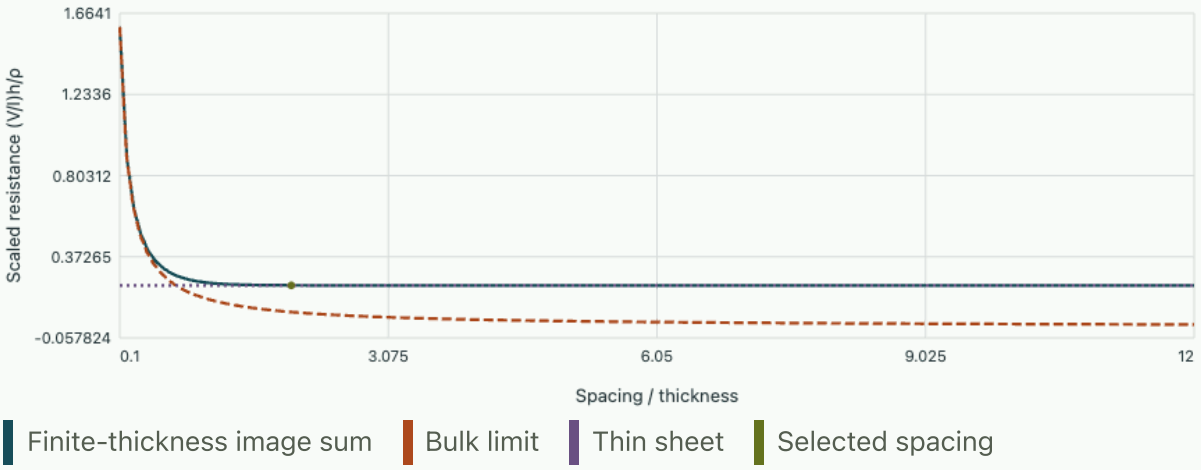
Try it · Four probe

Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Probe spacing / film thickness 2



Thin-sheet inference bias 0.0026442. Image sum uses 1500 pairs; the notebook checks its truncation with a longer sum.

Download the corresponding editable notebook. The browser recalculates the reduced equations directly; the notebook exposes the full calculation and its checks.

Take $\rho = 2.0 \times 10^{-5}$ ohm m, $h = 20$ micrometres and $s = 80$ micrometres. The sheet resistance is 1.00 ohm per square. For a 5.0 mA current the thin-sheet formula predicts $V = 1.1032$ mV. Summing the images changes this prediction by only a few parts per million at $s/h = 4$. A bulk interpretation would instead give $2\pi s V/I \simeq 1.11 \times 10^{-4}$ ohm m, more than five times the true resistivity. Both conversions have the right dimensions and both produce a positive, plausible laboratory number. Their different response to spacing supplies the experimental distinction.

The notebook lets you vary h, s, ρ and I separately. First keep s/h fixed while doubling both lengths. The scaled resistance stays fixed while $R_s = \rho/h$ halves. Then change only s. A real thickness change and an alternative probe head now follow different paths on the dimensionless graph. The optional OneSimulation fit adds a small constant thermoelectric voltage to synthetic observations at two current polarities. Its design matrix has a geometry-weighted current column and a constant column. Fitting both recovers the sheet resistance independently of offset, a useful distinction when the desired signal is only a millivolt.

Tutorial on offsets, symmetry and geometry

Current reversal at a fixed spacing gives +1.112 mV at +5.00 mA and −1.088 mV at −5.00 mA. Find the resistive signal and the current-independent offset. Assuming a thin sample, infer the sheet resistance.

▼ **Work through the solution**

Half the difference is 1.100 mV. Half the sum is the offset, 0.012 mV. The transfer resistance is 0.2200 ohm. Multiplication by π/ln 2 gives about 0.997 ohm per square. If the temperature changes during the reversal cycle, a drifting thermoelectric offset may survive this subtraction.

The current contacts are shifted outward so that the two inner probes have distances a, b from the left source and b, a from the right sink, with $b > a$. Derive the thin-sheet conversion.

▼ **Work through the solution**

The two inner potentials are opposite and have magnitude $IR_s \ln(b/a)/(2\pi)$. The differential resistance is $R_s \ln(b/a)/\pi$. Equal neighbouring spacings give $b/a = 2$. The sign of the measured voltage reverses if the two potential leads are exchanged, while the physical sheet resistance remains positive.

Why can current and voltage terminal pairs be interchanged in an ideal ohmic specimen, and what is a useful check if their readings differ?

▼ **Work through the solution**

The linear scalar conduction operator with insulating external walls is reciprocal. Integrating the two field equations against the potential of the other experiment, then subtracting, leaves only terminal terms; symmetry of the conductivity cancels the interior contribution. This equates the two transfer resistances using consistent orientations. Check damping of instrument transients, finite voltage input current and polarity reversal before assigning a difference to the material. A magnetic-field-induced antisymmetric conductivity requires a different reciprocity comparison.

Further reading · F. M. Smits · Measurement of Sheet Resistivities with the Four-Point Probe

Problem sheet 7 · Project · Infer a charge distribution from a grounded tube

A voltage probe sits on the centreline of a long square metal tube. Derive a reference field, decide what the probe can identify and test a numerical estimate of the charge density with OneSimulation.

Construct a forward problem that checks its own units

The walls of a square of side $L = 1$ m are held at zero. Along the tube nothing changes. Assume a vacuum and $\rho = \rho_0 \sin(\pi x/L) \sin(\pi y/L)$. Find the potential throughout the cross-section, explain the sign of the voltage and identify the dimensionless shape parameter if L changes.

▼ **Work through the solution**

Gauss's law gives

$$-\nabla^2 V = \rho/\epsilon_0.$$

Each second derivative on the sine product returns $-(\pi/L)^2$ times that product. Thus $V = \rho_0 L^2 \sin(\pi x/L) \sin(\pi y/L)/(2\pi^2\epsilon_0)$. The factor $\rho_0 L^2/\epsilon_0$ has units V, and the coordinates are $x/L, y/L$. Positive charge makes the interior positive relative to the grounded walls. The grounded boundary and Poisson uniqueness fix the solution.

Infer the charge before opening the solver

At the centre an instrument measures 5.72 ± 0.05 V. Determine the inferred peak charge density with uncertainty, assuming the sine shape is independently known. If a second probe is placed at $(L/4, L/2)$, what should it measure? Which error affects both estimates in the same way?

▼ **Work through the solution**

Divide the reading by $L^2/(2\pi^2\epsilon_0)$, or about 5.72165 V per nC/m³, to obtain 1.000 ± 0.009 nC/m³ after rounding. The second voltage is the first times $\sin(\pi/4)$, about 4.04 V. A gain error common to both instruments biases the inferred charge while leaving their voltage ratio unchanged.

Verify the mesh and distinguish forward from inverse error

Use the grounded-square experiment at 1 nC/m³. Compare 6, 12 and 24 cells across the tube. Run the same PDE in OneSimulation and check its centre row. Derive the sign of the grid error and the corresponding direction of bias if that grid were used to infer the charge.

▼ **Work through the solution**

The discrete Laplacian eigenvalue is $8\sin^2(\pi/(2n))/h^2$, $h = L/n$. Because $\sin t < t$ for positive t, it is smaller than the continuum eigenvalue $2(\pi/L)^2$. The discrete voltage is therefore high for the same density. Centre values for the three grids are about 5.8542, 5.7544 and 5.7298 V. Fitting the observed voltage with this grid would lower the charge estimate. The leading error decreases like n^{-2}.

Show precisely what one probe misses

Add an unknown volume charge proportional to $\sin(2\pi x/L)\sin(\pi y/L)$. Predict its centre signature and design a measurement to isolate it even in the presence of the original mode.

▼ **Work through the solution**

The extra potential is proportional to the same shape divided by $5\pi^2/L^2$. It vanishes at $x = L/2$. The original centre probe cannot detect it at any amplitude. Place probes at $x = L/4$ and $3L/4$, both at $y = L/2$. Their difference cancels the symmetric first mode and doubles the antisymmetric second. This is an identifiability gain from a change of measurement, independently of mesh refinement.

Problem sheet 8 · Project · Trace energy from a wave to the material supporting it

A plane wave illuminates a transparent dielectric and then a metal surface. Use field boundary conditions, flux, mechanical reaction and dissipation to account for the same energy in different descriptions.

Start with field amplitudes

A normally incident 300 V/m peak plane wave arrives from vacuum at a lossless nonmagnetic glass of refractive index 1.5. Assume a semi-infinite medium. Derive the reflected and transmitted field amplitudes by matching tangential electric and magnetic fields. Check the net cycle-averaged power.

▼ **Work through the solution**

Tangential continuity gives $1 + r = t$ and $1 - r = 1.5t$, where the minus sign follows from the reversed magnetic field of the reflected wave. Hence $r = -0.2$, $t = 0.8$. The reflected electric amplitude is -60 V/m and the transmitted amplitude 240 V/m. Reflectance is $r^2 = 0.04$ and power transmittance is

$$1.5t^2 = 0.96.$$

Squaring the transmitted electric amplitude ratio alone would omit the wave impedance.

Locate stored energy on the incident side

Use the incident and reflected energy experiment. Predict the two boundary contributions at phase zero with $r = -0.2$. Explain how standing structure is compatible with a constant time-averaged net flux.

▼ **Work through the solution**

At the boundary the normalised electric part is $(1 + r)^2 = 0.64$ and the magnetic part $(1 - r)^2 = 1.44$. The product EH has cancelling mixed terms between opposite propagation directions. Averaging over a period leaves the incident flux times $1 - |r|^2$, independent of position in the homogeneous incident region.

Replace the glass by a good conductor

For $\sigma = 5.8 \times 10^7$ S/m, $\mu = \mu_0$ and 100 kHz, derive the skin depth and compare with the penetration experiment. At what depth is only one per cent of the surface Joule power density left?

▼ **Work through the solution**

$\delta = 1/\sqrt{\pi f \mu \sigma} \approx 0.209$ mm. Field amplitude decays as $e^{-z/\delta}$, so time-averaged local Joule heating proportional to $|E|^2$ decays as $e^{-2z/\delta}$. Setting this to 0.01 gives $z = \delta \ln(100)/2 \approx 0.481$ mm. Confusing intensity with amplitude would double the inferred depth.

Check the global heat budget and a mechanical limit

If the known internal surface field has peak amplitude E_s, integrate the Joule heating over a thick conductor. Separately consider an ideal lossless mirror with incident time-averaged intensity I. How can its heat absorption vanish while it experiences a force?

▼ **Work through the solution**

The heat per unit area and time is

$$(\sigma |E_s|^2/2) \int_0^\infty e^{-2z/\delta} dz = \sigma |E_s|^2 \delta/4.$$

This equals the inward mean Poynting flux in the good-conductor limit, using $1/Z = (1-i)\sigma\delta/2$ for the stated phasor convention. At a perfect mirror the momentum flux reverses, giving pressure $2I/c$. A stationary mirror can exchange momentum without doing mechanical work.

Problem sheet 9 · Project · Measure a cavity and identify two different loss mechanisms

A narrow optical resonance can identify a cavity's round-trip phase and total decay. Use a time trace, a spectrum and a change of coupling to distinguish information that a single measurement cannot supply.

Establish what resonates

In a one-dimensional ideal cavity of length L in vacuum the tangential electric field vanishes at two perfectly conducting end planes. Derive the allowed wavelengths and the frequency spacing. For $L = 0.15$ m calculate the spacing. Which assumption lets the spacing remain constant across the spectrum?

▼ **Work through the solution**

A standing mode is proportional to $\sin(kz)$ with $kL = m\pi$. Thus $\lambda_m = 2L/m$ and $f_m = mc/(2L)$. Adjacent modes stay about 0.9993 GHz apart. A frequency-independent phase speed and frequency-independent boundary phases have been assumed. Mirror phase dispersion or a dispersive fill changes the spacing.

Reconcile a ring-down trace with the spectral line

For a mode at $f_0 = 10$ GHz the stored energy falls by $1/e$ in 0.2 µs after the drive is removed. Compute the energy decay rate, quality factor and expected full power linewidth. Compare energy and field-amplitude time constants in the cavity experiment.

▼ **Work through the solution**

Write $U = U_0 e^{-\kappa t}$. The rate is 5.0×10^6 s⁻¹. The field amplitude falls as $e^{-\kappa t/2}$, with time constant 0.4 µs. The driven response has denominator $(\kappa/2)^2 + (\omega - \omega_0)^2$. The full width at half maximum in ordinary frequency is $\kappa/(2\pi) = 0.796$ MHz, giving $Q \simeq 12566$. These expressions agree with $Q = \omega_0 U / P_{\text{loss}}$.

Infer loss rates from a second measurement

The initial linewidth includes an intrinsic energy decay rate κ_i and output-coupling rate κ_e. A calibrated modification halves κ_e while keeping κ_i fixed, lengthening the energy decay time to 0.25 µs. Find each rate and the fraction of stored energy that could leave through the port before the change.

▼ **Work through the solution**

Originally $\kappa_i + \kappa_e = 5 \times 10^6$ s⁻¹. After the change $\kappa_i + \kappa_e/2 = 4 \times 10^6$ s⁻¹. Subtraction gives $\kappa_e = 2 \times 10^6$ and $\kappa_i = 3 \times 10^6$ s⁻¹. The initial out-coupled fraction is

$$\kappa_e / (\kappa_i + \kappa_e) = 0.4.$$

A lone linewidth constrains only the sum. The second experiment supplies independent information.

Challenge the assumption of a single stored mode

An observed photodetector signal oscillates on top of an exponential envelope. Describe a cause consistent with a linear cavity and an experiment that distinguishes it from a defective exponential integrator.

▼ **Work through the solution**

If two modes are populated, the detector records $|a_1 e^{-i\omega_1 t} + a_2 e^{-i\omega_2 t}|^2$. The cross term oscillates at their difference frequency. Narrow the excitation band or spatially select one mode and repeat. A numerical time-step test can identify solver error, but changing the prepared physical field tests the single-mode assumption directly.

Notebook practice · Four contacts on a conductor of finite thickness

Electrostatic field theory and measurement

Learn to infer a field from charges, potentials or boundary measurements. Build the mathematics of flux, uniqueness and Green functions as each problem requires it.

1. ## From a force reading to Poisson's equation

   Imagine trying to infer an invisible charge distribution with a force probe and a voltmeter. Coulomb's law, Gauss's law and the electrostatic potential describe the same experiment in three different ways. Deriving the connections also exposes the measurements a single probe cannot make.

2. ## Vector calculus through flux, circulation and singular sources

   A closed-surface flux and a closed-path circulation answer different experimental questions. Starting from small physical collecting surfaces shows why the associated differential operators look different in Cartesian, cylindrical and spherical coordinates.

3. ## When do boundary measurements determine a unique electrostatic field?

   A boundary-value solver may converge while solving the wrong electrical experiment. Establishing which potentials, charges and fluxes are fixed gives a uniqueness proof and a way to turn an approximate field into an upper bound.

4. ## Images in wedges, cylinders and dielectric interfaces

   An image source is a construction for a particular physical domain and boundary condition. Three examples show how geometry determines the image position, how material response sets its amplitude, and why an apparent success on part of a boundary is insufficient.

5. ## Green functions, the Poisson kernel and what a small electrode can influence

   A Green function is the response of an entire boundary-value problem to a unit source. Once that response is known, a continuous charge distribution and a patterned boundary voltage can both be treated by integration. The two integrations require different kernels.

6. ## Rectangular geometries, orthogonality and sources that drive individual modes

   Fourier modes are solutions selected by their electrode walls. On a rectangle they turn a boundary-voltage problem into independent vertical equations. When a charge fills the rectangle the same eigenfunctions give a different denominator and a different physical response.

7. ## Cylindrical geometry, coaxial boundaries and Bessel spectra

   The radial coordinate of a cylinder leads to logarithms in one experiment and to Bessel functions in another. The difference is set by the variation along the axis and around the circumference. Recovering the radial weight from flux clarifies both.

8. Normalised spherical harmonics and the field of an arbitrary three-dimensional source

An axial voltage scan can determine Legendre coefficients only if the source is rotationally symmetric. A general spherical electrode or molecule needs angular modes with two indices. Their normalisation fixes the conversion from a measured multipole coefficient to a physical voltage.

9. Capacitance matrices, floating nodes and controlled approximation bounds

A multi-electrode device generally has no single capacitance until the other terminals have been connected or isolated. An energy identity determines the matrix that relates voltages to charges, and a complementary flux identity can certify estimates even when the exact field is unknown.

10. Anisotropic and nonlinear dielectrics from local measurements to global fields

A single dielectric constant cannot describe a crystal whose displacement tilts away from its electric field, or a reversible material whose incremental response changes with bias. Starting with the constitutive measurement clarifies which parts of ordinary electrostatics survive.

11. From mobile ions to screening lengths and differential double-layer capacitance

A charged planar wall in an electrolyte can be studied without assuming a capacitance in advance. Diffusion and electrostatic drift first establish an ion cloud. The charge–voltage slope of that cloud is an experimental output of the model.

12. Reconstructing a source from noisy voltage measurements

A forward electrostatic calculation can be unique and stable while the inverse problem constructed from it is unstable or nonunique. A scanned electrode, a finite collecting area and an explicitly chosen noise level make these statements experimentally testable.

13. Which electrode sees a charge moving above a detector?

A point charge moves without touching the electrodes. Derive the induced charge on one strip from a second, source-free boundary problem. Use the two-dimensional map to distinguish motion across the detector from motion away from it.

14. Build the field operators in curved coordinates

Derive the geometric factors in gradient, divergence, curl and Laplacian, then use dimensional conversion to compare electromagnetic conventions.

Electrostatic field theory and measurement · 1. From a force reading to Poisson's equation

Imagine trying to infer an invisible charge distribution with a force probe and a voltmeter. Coulomb's law, Gauss's law and the electrostatic potential describe the same experiment in three different ways. Deriving the connections also exposes the measurements a single probe cannot make.

Calibrate a probe before calling its reading a field

Place a small charged bead on an insulating support, sufficiently far from other objects for the effect of surrounding matter to be negligible. At a known point nearby, hold a much smaller probe charge q_t on a force balance. If the only force changing when the source is introduced is electrical, the difference in force readings is $q_t\mathbf{E}$. Operationally one should repeat the measurement with decreasing probe charge and extrapolate the ratio $\mathbf{F}/q_t$. Otherwise a mobile source or a nearby conductor might rearrange in response to the probe itself. A field at a point means the force per unit positive test charge in this limiting preparation.

In SI units the vacuum force between stationary point charges is parallel to their separation. A convenient convention places the source at $\mathbf{y}$, the detector at $\mathbf{x}$, and defines $\mathbf{R} = \mathbf{x} - \mathbf{y}$. Then $\mathbf{E} = q\mathbf{R}/(4\pi\epsilon_0 R^3)$. The factor $\mathbf{R}/R^3$ contains a unit direction vector and an inverse square in magnitude. Reversing source sign reverses the vector. Doubling distance reduces its magnitude to one quarter. Doubling source and test charge together increases the force fourfold while doubling the field only twice relative to the original experiment.

For $q = 40$ pC at distance 60 mm the magnitude is 99.9 V/m. A positive 2 pC test bead experiences 0.200 nN away from the source. A negative probe gives the opposite force at the same field. This example also establishes why reporting a force in newtons and a field in volts per metre are physically different measurements. Making the test charge twice as large can improve the force signal but can increase disturbance of the original experiment.

Replace a sum by an integral without losing the direction

For two sources superposition adds forces on the same probe, so the electric fields add as vectors. In a continuous charge distribution divide space into cells small on the scale of the observation distance and assign each cell $dq = \rho(\mathbf{y})d^3y$. Refining the partition gives

$$\mathbf{E}(\mathbf{x}) = \frac{1}{4\pi\epsilon_0}\int \frac{\mathbf{x}-\mathbf{y}}{|\mathbf{x}-\mathbf{y}|^3}\rho(\mathbf{y})\,d^3y.$$

Surface or filament models replace dq by $\sigma\,dA$ or $\lambda\,ds$, respectively. Their densities have different units. A wire of radius a may be modelled as a line at observation distances large compared with a, while the field right at the ideal line diverges. Replacing a physical distribution with a lower-dimensional source is an approximation about distances. It is not a claim that metal contains literal curves with zero area.

Take a straight filament extending along z from $-L$ to $+L$, with constant λ. On the perpendicular bisector at $x = s > 0$, terms from z and $-z$ cancel their axial components. A slice contributes $dE_x = \lambda s\,dz/[4\pi\epsilon_0(s^2+z^2)^{3/2}]$. Using $d[z/\sqrt{s^2+z^2}]/dz = s^2/(s^2+z^2)^{3/2}$ yields

$$E_x = \frac{\lambda}{2\pi\epsilon_0 s}\frac{L}{\sqrt{s^2+L^2}}.$$

At $s \ll L$ this is the infinite-line expression. At $s \gg L$ it becomes $2L\lambda/(4\pi\epsilon_0 s^2)$, exactly the point-charge field for the full segment. The same model changes apparent radial power law between two limits. For $s = L$ the long-line formula would overpredict by a factor $\sqrt{2}$. Fitting an observed power law over a limited range does not identify the microscopic source shape.

Discover why inverse-square strength produces a flux law

Choose a patch of surface at distance R from a point charge. Its oriented area is $\hat{\mathbf{n}}\,dA$. Only the component projected perpendicular to the ray contributes to radial-field flux. That projected area is $\hat{\mathbf{n}}\cdot\hat{\mathbf{R}}\,dA = R^2 d\Omega$, where $d\Omega$ is an oriented solid angle. Consequently $d\Phi_E = q\,d\Omega/(4\pi\epsilon_0)$. The decrease with R^2 exactly cancels the spread of a fixed cone. An enclosing sphere covers total solid angle 4π, and any smooth closed surface surrounding the point gives the same oriented total because each outward ray leaves once more than it re-enters. For a charge lying outside, entering and leaving contributions cancel.

Superposition extends this argument to arbitrary bounded sources. For a closed surface with no point source on the surface,

$$\oint \mathbf{E}\cdot d\mathbf{A} = Q_{\text{inside}}/\epsilon_0.$$

This equation returns one scalar from an entire field map. It determines pointwise E only where symmetry fixes directions and makes the relevant component constant over the chosen pieces. Draw a cube around an off-centre charge. The six face fluxes still sum to q/ϵ_0, but symmetry cannot divide that sum by six. Draw the same cube centred on that charge and all six face integrals become equal even though the field is not uniform across an individual face.

Gauss's law also explains the field of a thick spherical distribution without integrating the Coulomb kernel. For $\rho(r)$ in a sphere centred at the origin, invariance under all rotations forces $\mathbf{E} = E_r(r)\hat{\mathbf{r}}$. A spherical collecting surface gives

$$E_r = Q(<r)/(4\pi\epsilon_0 r^2),$$

with $Q(<r) = 4\pi\int_0^r \rho(s)s^2 ds$. A hollow shell in vacuum therefore has zero interior field; its potential there may still be large relative to infinity. Changing the reference of the voltmeter affects its number but cannot change the force probe.

Why a line integral can be reduced to its endpoints

For a static point charge,

$$d(1/R) = -\hat{\mathbf{R}} \cdot d\mathbf{x}/R^2.$$

The work per unit test charge along a path from A to B is $\int_A^B \mathbf{E} \cdot d\mathbf{x} = q[1/R_A - 1/R_B]/(4\pi\epsilon_0)$. It depends only on endpoints. Superposing static sources preserves this property. We define $V(B) - V(A) = -\int_A^B \mathbf{E} \cdot d\mathbf{x}$. With a reference at infinity for a compact distribution, direct integration gives

$$V(\mathbf{x}) = \int \rho(\mathbf{y})/(4\pi\epsilon_0|\mathbf{x} - \mathbf{y}|)\, d^3y.$$

A numerical evaluation adds scalar potentials; a directional derivative recovers the desired force component. Differentiating local data can amplify noise, so a voltage map and a force map are equivalent descriptions mathematically but different experiments.

An infinitesimal displacement yields

$$\mathbf{E} = -\nabla V.$$

To see the sign, move a positive probe in the field direction. The field does positive work and the potential decreases. Surfaces of constant V are perpendicular to $\mathbf{E}$ wherever its magnitude is nonzero. Uniformly spaced contour values on a map denote equally spaced voltages, not equal physical intervals. Where those contours crowd together the component normal to them has greater magnitude.

For the finite filament, integrating $dz/\sqrt{s^2 + z^2}$ on the perpendicular bisector gives

$$V(s) = \lambda \operatorname{arsinh}(L/s)/(2\pi\epsilon_0).$$

Its negative derivative recovers the field above. If L tends to infinity at fixed s, the field reaches a finite limit but V diverges logarithmically when referenced to infinity. A measured difference $V(s_2) - V(s_1)$ stays finite, tending to $-\lambda \ln(s_2/s_1)/(2\pi\epsilon_0)$. The failure is in the proposed reference for the infinite model. Real finite conductors and return paths supply physically well-defined voltage references.

From integral measurements to a boundary problem

Apply the divergence theorem to Gauss's law in a smooth region: $\int_\Omega (\nabla \cdot \mathbf{E} - \rho/\epsilon_0) d^3x = 0$. It holds for arbitrary small volumes, so $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$ there. Combining it with the potential gives Poisson's equation $-\epsilon_0 \nabla^2 V = \rho$, reducing to Laplace's equation in empty space. At point or sheet sources this statement has an integral, or distributional, interpretation. A single derivative at the singularity is not an ordinary finite number.

The equation alone is insufficient to compute a laboratory voltage. Adding any harmonic function changes the field while preserving the same charge density inside the region. Installing external electrodes changes those harmonic contributions. To predict a unique measurement, supply boundary potentials, conductor charges, flux conditions or an appropriate condition at infinity. For example, $V(x) = \rho x(L - x)/(2\epsilon_0)$ solves a one-dimensional region of uniform charge between grounded planes at 0 and L. Raising the right

plane to V_b adds $V_b x/L$. The curvature and inferred density remain identical, while the force throughout the gap acquires a uniform extra component.

How does this fit time-dependent Maxwell theory? A static source satisfies a zero-curl electric field and the Gauss constraint. With a varying magnetic flux, Faraday changes the circulation to $-d\Phi_B/dt$, so a purely scalar, path-independent description of the complete electric field can fail. A slowly driven experiment can remain close to an electrostatic sequence when induction corrections and propagation delays are negligible. Compare the system flight time L/c with its driving timescale, but also check any imposed magnetic flux. A small device placed around a deliberately changing magnetic circuit can have appreciable induced circulation even when its own dimensions are small.

A finite charged sheet and the solid angle it subtends

An insulating rectangle spans $-a \le x' \le a$, $-b \le y' \le b$ in the horizontal plane, carrying prescribed uniform surface charge σ. The sides need not be equal. Measure the field directly above its centre at $(0, 0, h)$, $h > 0$. Reflection suppresses the two lateral components but there is no sufficiently symmetric closed Gaussian surface that supplies the remaining field from flux alone. Instead each patch contributes $dE_z = \sigma h\, dx'dy'/[4\pi\epsilon_0(x'^2 + y'^2 + h^2)^{3/2}]$. The factor $h\,dx'dy'/R^3$ is the positive solid angle of the patch as seen from the detector. Thus $E_z = \sigma\Omega/(4\pi\epsilon_0)$, where Ω is the solid angle of the whole rectangle. This reverses the usual flux picture: the detector is the point from which the source surface is viewed.

The integral can be evaluated without a solid-angle table. By symmetry integrate one quadrant and multiply by four. First integrate along x' using the antiderivative $x'/[(h^2 + y'^2)\sqrt{x'^2 + h^2 + y'^2}]$. This leaves $\int_0^b ah\, dy'/[(h^2 + y'^2)\sqrt{a^2 + h^2 + y'^2}]$. Differentiating $\arctan[ay'/(h\sqrt{a^2 + h^2 + y'^2})]$ recovers precisely this integrand. Consequently

$$E_z = \frac{\sigma}{\pi\epsilon_0}\arctan\left(\frac{ab}{h\sqrt{a^2 + b^2 + h^2}}\right).$$

Before inserting numbers, examine both extreme geometries. When $h \to 0^+$ at fixed a, b, the arctangent tends to $\pi/2$. A small probe above the interior sees the infinite-sheet field $\sigma/(2\epsilon_0)$. Reverse h to obtain its opposite below the sheet, giving the correct normal jump. If $h \gg a, b$, the arctangent argument is small and $E_z \simeq \sigma ab/(\pi\epsilon_0 h^2) = Q/(4\pi\epsilon_0 h^2)$ for $Q = 4ab\sigma$. A finite source with sheet-like near field therefore regains a point-like field without needing a second physical law.

For a square of half-width 4.0 cm, at $h = 4.0$ cm its solid angle is

$$4\arctan(1/\sqrt{3}) = 2\pi/3.$$

The field is only one third of the infinite-sheet estimate. If $\sigma = 2.0$ nC/m², that is 37.6 V/m instead of 113 V/m. A cube of side 8.0 cm drawn around a point at its centre has exactly the same solid angle on each face. Six copies of $2\pi/3$ close to 4π; this checks the integral's normalisation without referring back to the same antiderivative. At the same h, replacing the entire charged square by a point would instead give 71.9 V/m. Its error has the other sign because the approximation concentrates all charge at the closest radial location.

The same formula can be used for a point above a rectangular sensor away from its centre. Break the rectangle into four axis-aligned pieces whose common corner lies at the projected point. Add their signed quadrant solid angles with their actual edge distances, taking care of sign when the projection lies outside the detector. A useful numerical strategy is to integrate the near patch analytically and the smooth distant patches by quadrature. A uniform mesh chosen only to match the physical electrode widths may entirely miss the sharply peaked angular kernel at small h. Refine around the projection while checking the total field against the measured large-distance and small-distance limits.

What a differential voltage measurement rejects

A voltmeter spanning two nearby points measures an integral of the field rather than a pointwise derivative. Put two identical probes at $x = \pm s/2$ on an axis and report $\Delta V = V(-s/2) - V(+s/2)$. Taylor expansion around the origin gives

$$\Delta V / s = -V'(0) - s^2 V'''(0)/24 + O(s^4).$$

It cancels a common voltage offset and the quadratic term. The estimate approaches the centre field with an error set by curvature of the field itself, not by the local voltage magnitude. A common reference shift can be huge and still disappear exactly, while spatially varying disturbances remain.

Take a small source charge q on the same axis at $x = d > s/2$. The exact potential is $kq/(d - x)$,

$$k = 1/(4\pi\epsilon_0).$$

Substitution gives a signed differential estimate $\Delta V / s = -kq/[d^2 - (s/2)^2]$. The true field at the midpoint is $-kq/d^2$, pointing away from a positive charge to the left. A probe of finite separation overestimates its magnitude by $1/[1 - (s/2d)^2]$. For $s = 10$ mm, $d = 25$ mm the bias is 4.17%. Reducing s to 5 mm gives 1.01%. This systematic error does not disappear by repeating the same pair of readings many times.

Measurement noise produces the opposite trend. If each voltage reading has independent standard deviation δV, their difference divided by s has standard deviation $\sqrt{2}\,\delta V / s$. Halving probe separation doubles the field uncertainty even as it quarters the leading geometrical bias. With $\delta V = 0.5$ mV, a 10 mm baseline gives 0.071 V/m uncertainty. At a true field of 10 V/m and the preceding geometry, the systematic 0.417 V/m is larger; at 5 mm it falls to 0.101 V/m while random uncertainty grows to 0.141 V/m. Comparing both scales helps choose a practical probe spacing. Correlated noise from a common reference changes the variance calculation and may cancel more completely than independent probe noise.

This distinction persists when measurements are replaced by a grid of computed potentials. A finite-difference derivative measures change over a cell, while the Coulomb integral can give the field directly at a point. Agreement of the raw potential to many digits does not guarantee a comparably accurate local field if neighbouring numerical errors vary rapidly. Use a known analytic source, vary the sampling baseline separately from the underlying field mesh, and compare the measured scaling with the predicted s^2 truncation and $1/s$ noise behaviour. That procedure checks an actual numerical reading of the physical observable rather than simply refining every parameter together.

Can a bead of nonzero size measure the field at one point?

Consider a rigid insulating bead of radius a_t, carrying a fixed spherically symmetric probe charge density $\rho_t(r)$ and total charge q_t. The external sources all lie outside the bead and do not move when the bead is introduced. These assumptions exclude rearrangement of either the probe's own charge or the external conductors. Its net electric force is the volume integral $\int \rho_t(\mathbf{s})\mathbf{E}_{\text{ext}}(\mathbf{x}_0+\mathbf{s})d^3s$. It might appear necessary to treat every nonzero bead radius as a source of measurement bias. In this particular experiment the result equals $q_t\mathbf{E}_{\text{ext}}(\mathbf{x}_0)$ exactly.

To see why, the external potential is harmonic throughout the enclosing ball. Each Cartesian component of its field is also harmonic there, because a Cartesian derivative commutes with the Laplacian in a smooth uniform vacuum. The mean-value property says that the angular average of a harmonic field component on each spherical shell equals its value at the centre. Pull this value through the remaining radial integral of $\rho_t(r)\,4\pi r^2 dr$ and the result follows. The field can change strongly across the bead. What matters is the combination of a source-free exterior solution within its volume and a rigid radially symmetric charge weight. A probe that averages along a straight baseline, as above, does not possess this spherical mean-value property.

What changes if the source charge distribution lies partly inside the sampling ball? In a charged region $\nabla^2 E_i = \partial_i \rho/\epsilon_0$, so a density gradient can produce a finite-size correction. For a uniform spherical probe a Taylor expansion gives $\langle s_i s_j\rangle = (a_t^2/5)\delta_{ij}$, hence leading force per total charge $E_i(\mathbf{x}_0) + (a_t^2/10)\nabla^2 E_i(\mathbf{x}_0)$. In a locally uniform volume charge even this correction vanishes, since an electric field linear in position averages to its central value under any centred symmetric probe. Material interfaces or singular sources require separate treatment; the smooth expansion cannot be extended through them by taking more terms.

A conducting sensor is a different preparation. A neutral conducting bead carries an induced dipole in an external field and can feel a force in a field gradient even with lab-calibrated net charge zero. One experimental check is to reverse the deliberately deposited test charge while keeping its magnitude fixed. In a weakly perturbing regime the desired force changes sign with q_t; a charge-independent polarisation force contributes to the even part. Dividing a single force measurement by q_t without checking such an offset may produce an apparent field that changes with probe charge. A physically meaningful field measurement begins by deciding what the actual sensor averages and which parts of its signal follow the calibration law.

Measure the field of a finite source

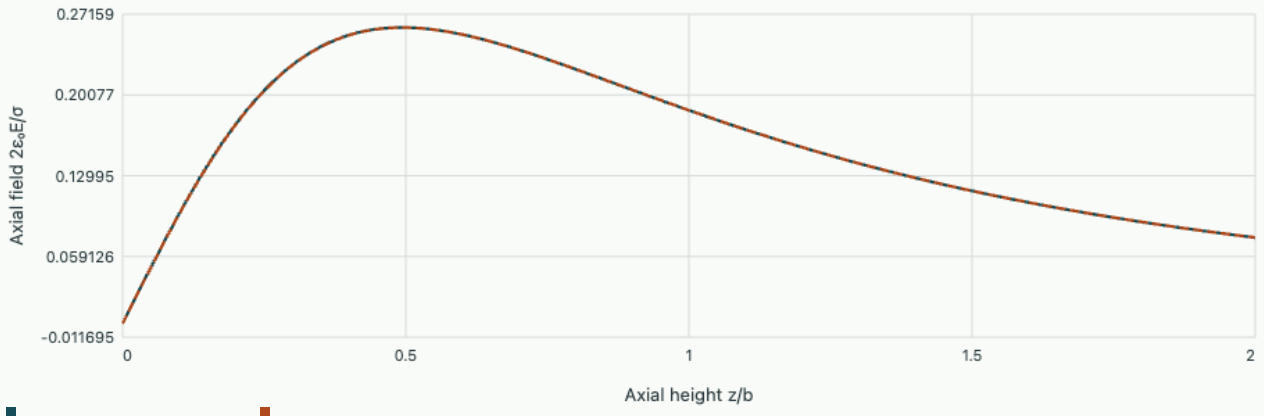
Move along the axis of an annulus and compare the integrated source with its distant point-charge limit. Identify the length scale that controls the difference.

Open this experiment online

Try the calculation

Hole / outer radius 0.5

Radial integration cells 8



Analytic integral Ring quadrature

The analytic peak is at z/b = 0.4934 with field 0.2599. At that location the quadrature differs by -0.01507%.

Tutorial · design measurements that distinguish sources

A finite line mistaken for an infinite one

A filament extends from −40 to +40 mm. A probe is placed 10 mm from its midpoint on the perpendicular bisector. Find the fractional error from using the infinite-line formula. How close must the probe be for that approximation to overstate the exact field by less than 1%?

▼ **Work through the solution**

The ratio of the approximation to the exact field is $\sqrt{1+(s/L)^2}$. At $s/L = 1/4$ this is 1.0308, so the overestimate is 3.08%. Requiring the ratio below 1.01 gives $s/L < \sqrt{1.01^2 - 1} = 0.1418$, or $s < 5.67$ mm. The proximity criterion must coexist with the condition that the physical filament radius be much smaller than s. No universal sensor distance satisfies both criteria for every wire.

A second check is to take two radial readings. Their ratio cancels any unknown absolute calibration of λ. The infinite-line prediction is simply the inverse distance ratio, whereas the exact result contains the two end correction factors. With finite signal noise a reading very near the source may not be necessary if the probe is stable enough to resolve the predicted ratio difference.

Can an external cage identify a hollow charge distribution?

A spherical shell at radius a and a solid sphere of radius a carry the same spherically symmetric total charge. Compare external voltage, centre voltage, and electrostatic work on a positive probe from the

centre to the surface for uniform charge in each model.

▼ **Work through the solution**

Outside both, the field is $q/(4\pi\epsilon_0 r^2)$ and the potential with zero at infinity is $q/(4\pi\epsilon_0 r)$. No exterior electrostatic measurement can distinguish these sources within the spherical model. For the shell the interior field vanishes, so centre and surface potentials are equal and the journey does no electrical work on the probe. For a solid uniform sphere $E_r = qr/(4\pi\epsilon_0 a^3)$, so $V(0) - V(a) = q/(8\pi\epsilon_0 a)$. The centre potential becomes $3q/(8\pi\epsilon_0 a)$ and the outward electrical work on test charge q_t is $q_t q/(8\pi\epsilon_0 a)$.

At fixed radius the solid's higher centre voltage coexists with zero centre field, because the potential slope vanishes there by symmetry. An internal probe may itself disturb the charges in a liquid or conductor. The comparison assumes the specified source profiles stay fixed as measurements are taken.

Further reading

- Integrate a finite annulus and compare its field with the analytic integral
- Reconstruct a screened charge from two voltage readings
- MIT, Electromagnetic Fields and Energy, chapter 4 · source integrals
- Griffiths, 5th ed., chapters 1–2 · vector analysis and electrostatics

Electrostatic field theory and measurement · 2. Vector calculus through flux, circulation and singular sources

A closed-surface flux and a closed-path circulation answer different experimental questions. Starting from small physical collecting surfaces shows why the associated differential operators look different in Cartesian, cylindrical and spherical coordinates.

How a short displacement detects a derivative

A voltmeter measures a difference between endpoints. If a small displacement $d\boldsymbol{\ell}$ changes a smooth potential by dV, the gradient is the unique vector satisfying $dV = \nabla V \cdot d\boldsymbol{\ell}$ for every displacement. Its component along an arbitrary unit vector is the slope in that direction. Its largest rate of increase occurs when the displacement is parallel to the gradient. A unit step in a coordinate label does not necessarily mean a unit physical displacement. This distinction is the source of the metric factors in curvilinear coordinates.

At position (s, ϕ, z) in cylindrical coordinates, the three physical components of displacement are ds, $s\,d\phi$, dz. Equating the general dot product to $\partial_s V\,ds + \partial_\phi V\,d\phi + \partial_z V\,dz$ gives

$$\nabla V = \hat{\mathbf{s}}\partial_s V + \hat{\boldsymbol{\phi}} s^{-1}\partial_\phi V + \hat{\mathbf{z}}\partial_z V.$$

The azimuthal factor converts a voltage change per radian into a voltage change per metre around a circle. For $V = V_0(s/a)\cos\phi$ the radial and azimuthal components of $-\nabla V$ are $-V_0\cos\phi/a$ and $+V_0\sin\phi/a$. Together they equal the constant Cartesian field $-V_0\hat{\mathbf{x}}/a$, since $s\cos\phi = x$. The apparent angular variation comes only from the rotating local unit vectors. Leaving out $1/s$ would predict a nonuniform field and components with incompatible dimensions.

Count what leaves the sides of a shrinking cell

Consider a vector $\mathbf{F}$, perhaps electric displacement measured in C/m². Define its divergence at a smooth point as the limiting outward flux per cell volume. In a Cartesian box $\Delta x\Delta y\Delta z$, the opposing x faces give $[F_x(x+\Delta x) - F_x(x)]\Delta y\Delta z$. Dividing by volume and adding the other pairs gives $\partial_x F_x + \partial_y F_y + \partial_z F_z$. The signed flux measures the local excess of outflow over inflow. A field can be large and uniform while its divergence is zero.

Repeat the count in a thin cylindrical wedge between s and $s + ds$, angular width $d\phi$, height dz. The outer curved face has larger area than the inner one. Their net flux is $\partial_s(sF_s)ds\,d\phi\,dz$. The two cuts at fixed angle each have physical area $ds\,dz$ and differ by $\partial_\phi F_\phi d\phi$. The axial faces contribute $\partial_z F_z\,s\,ds\,d\phi\,dz$. Division by the wedge volume gives

$$\nabla\cdot\mathbf{F} = \frac{1}{s}\partial_s(sF_s) + \frac{1}{s}\partial_\phi F_\phi + \partial_z F_z.$$

A radial flow $F_s = A/s$ has zero divergence at positive s because the same flux crosses every cylindrical side of fixed height. A constant radial magnitude instead has divergence A/s: larger cylinders carry greater total flux and therefore enclose additional source. Specifying that a magnitude is constant says nothing about the divergence of its direction field.

In a spherical sector the radial areas grow as r^2. For a spherically symmetric radial field $F_r(r)$, the same count gives $r^{-2}d(r^2F_r)/dr$. A sphere of radius r then has flux $4\pi r^2F_r(r)$. Choosing $F_r = Ar$ produces $3A$ everywhere instead of A; the two missing contributions in a one-dimensional derivative account for growth of transverse area. For a uniform charge density ρ, setting $3A = \rho/\epsilon_0$ immediately reconstructs the interior electrostatic field of a sphere.

Inspect a small loop before asserting that a potential exists

The component of curl normal to a surface is the limiting circulation of $\mathbf{F}$ around an oriented small loop divided by its area. In an xy rectangle the upper side is traversed oppositely to the lower one. Opposite x-directed sides leave $-\partial_y F_x$; opposite y-directed sides leave $+\partial_x F_y$. Thus

$$(\nabla \times \mathbf{F})_z = \partial_x F_y - \partial_y F_x.$$

For a planar polar loop the curved segments have different lengths, giving $(\nabla \times \mathbf{F})_z = s^{-1}[\partial_s(sF_\phi) - \partial_\phi F_s]$. Curl and divergence use different faces and signs even when their cylindrical formulas share a factor $1/s$.

Join many neighboring loops. Circulations on common edges cancel in pairs, leaving only the outer boundary; the summed area integrals of the curl give Stokes's theorem. A similar cancellation of internal face fluxes gives the divergence theorem. These are physical bookkeeping identities, and they motivate discretisations that attach oriented fluxes to shared faces and voltages to shared edges. Assigning two independent approximations to one shared face can spoil exact cancellation in a computed conservation budget.

Take $\mathbf{F} = A\hat{\boldsymbol{\phi}}/s$ outside a thin excluded cylinder. Its ordinary curl is zero there. Any circle about the cylinder has circulation $2\pi A$. A disk spanning this circle would pass through the excluded region, so a surface lying entirely inside the stated domain cannot supply the usual zero-curl conclusion. Locally $\mathbf{F} = \nabla(A\phi)$ on any small patch avoiding a branch cut, while the putative potential changes by $2\pi A$ after a full turn. This example explains the role of domain topology. A genuine static electrostatic source in ordinary free space has a globally consistent potential; a changing flux through an excluded region can instead generate precisely such an exterior circulation.

Measure curvature through neighboring voltages

Taking the divergence of a gradient defines $\nabla^2 V$. A Taylor expansion at a point gives

$$[V(x+h) - 2V(x) + V(x-h)]/h^2 = \partial_x^2 V + O(h^2).$$

Adding the three Cartesian expressions relates the six neighbouring voltage readings to the local charge density. Equivalently, averaging a potential over directions at radius h removes all terms odd in the displacement. The angular mean of $n_i n_j$ is $\delta_{ij}/3$, so for smooth V

$$\langle V(\mathbf{x} + h\mathbf{n})\rangle_\Omega = V(\mathbf{x}) + \frac{h^2}{6}\nabla^2 V(\mathbf{x}) + O(h^4).$$

In a region with positive charge, Poisson's equation makes the spherical average smaller than the centre value. An uncharged saddle can have strong curvature along each of two axes that cancels in the sum. A local maximum along a single measurement line therefore does not prove that an interior source exists. To make such an inference one needs curvature in every independent spatial direction or independent symmetry information.

Near a boundary or dielectric interface this smooth Taylor analysis may break down. Enforce the integral flux balance on a cell straddling the interface instead. In media of different permittivity the normal derivative of V can jump even where there is no free charge. The physical differential operator is then $\nabla \cdot (\epsilon \nabla V)$; simply inserting a pointwise averaged permittivity in front of a smooth Laplacian may erase the correct jump condition.

Use a test volume to give singular expressions meaning

Outside the origin the divergence of $\hat{\mathbf{r}}/r^2$ is zero, yet every sphere centred there has flux 4π. To characterize its source measure the field against a smooth test function f of compact support. Excise a ball of radius η, integrate by parts in the remaining exterior and note that its inward-facing inner boundary normal is $-\hat{\mathbf{r}}$. The zero bulk divergence eliminates the interior term, leaving $-\int (\hat{\mathbf{r}}/r^2) \cdot \nabla f \, dV \to 4\pi f(0)$ as the ball shrinks. This is the defining action of the three-dimensional delta distribution:

$$\nabla \cdot (\hat{\mathbf{r}}/r^2) = 4\pi \delta^3(\mathbf{r}),$$

$$\nabla^2 (1/r) = -4\pi \delta^3(\mathbf{r}).$$

The factors and signs follow from an ordinary finite flux calculation. A delta distribution has no finite peak height to compare with a mesh value. A convergent field computation should reproduce its integral and its action on smooth functions. In spherical coordinates a radially concentrated charge can be represented with a radial delta only when the $4\pi r^2$ integration weight and the convention for the endpoint are treated consistently. It is often safer to specify the enclosed charge directly.

For a finite regularisation take $V_a = q/(4\pi\epsilon_0\sqrt{r^2 + a^2})$. Differentiation gives

$$E_r = qr/[4\pi\epsilon_0 (r^2 + a^2)^{3/2}].$$

A spherical flux measurement encloses $Q/q = r^3/(r^2 + a^2)^{3/2}$. Its derivative divided by $4\pi r^2$ gives

$$\rho_a = 3qa^2/[4\pi (r^2 + a^2)^{5/2}].$$

For every $a > 0$ the density is smooth; its peak scales as a^{-3}, preserving a finite total charge. For fixed $r > 0$, $Q/q \to 1$ as $a \to 0$. Shrinking the collecting radius and the source core simultaneously at fixed r/a instead retains a nontrivial fraction. The resolution of the probe is part of the limiting experiment.

A sheet of dipoles produces a voltage jump without a net sheet

charge

Replace an ideal neutral polarised layer by two parallel charge sheets, one at $z = -d/2$ with density $-\sigma$, another at $z = +d/2$ with density $+\sigma$. Assume both planes extend far beyond the detectors so their fields can be treated as infinite-plane fields. The field outside cancels. Between them it points from positive to negative,

$$E_z = -\sigma/\epsilon_0.$$

A voltmeter moved from below to above the pair rises by $+\sigma d/\epsilon_0$. The result can stay finite when d shrinks provided that the dipole moment per area $p_s = \sigma d$ is fixed. The fields away from the layer then converge to zero on both sides while the voltage difference persists. This model is useful for a surface dipole layer, for example when positive and negative charge distributions are displaced by molecular-scale distances.

It is essential not to infer the voltage jump from the net charge enclosed in a pillbox, since the net is zero for every d. Write the signed density as $\rho_d = \sigma[\delta(z - d/2) - \delta(z + d/2)]$. For any smooth test voltage $f(z)$, integration gives $\sigma[f(d/2) - f(-d/2)] \to p_s f'(0)$. The distribution with action $f \mapsto p_s f'(0)$ is $-p_s \delta'(z)$. In one dimension Poisson's equation is $-\epsilon_0 V'' = \rho$, so a solution is

$$V = (p_s/\epsilon_0) H(z) + \text{constant},$$

where H is a step. Its derivative is a delta, its second derivative a delta derivative. Reversing the layer orientation reverses both the jump and the sign of this source distribution.

A conventional monopole sheet instead has $\rho = \sigma_f \delta(z)$. Its potential is continuous and its normal derivative jumps. The two types of ideal source thus leave different signatures in electrode boundary models. For a voltage jump of 20 mV, the required $p_s = \epsilon_0 (0.020) = 1.77 \times 10^{-13}$ C/m in vacuum. Spreading it across 1.0 nm requires $\sigma = 1.77 \times 10^{-4}$ C/m² and an internal field of 20 MV/m in that thin-layer model. The large internal field is consistent with a small voltage step because it acts only over one nanometre. At fixed p_s the field energy per area grows as $p_s^2/(2\epsilon_0 d)$. Thus taking a perfect zero-thickness dipole layer gives useful exterior fields and voltage jumps but not a finite internal energy. A physical model of finite layer thickness is necessary for that observable.

Moving a numerical grid across the sheet illustrates how to recognise the two cases. If its potential samples resolve a continuous cusp, integrating the inferred charge across one symmetric layer gives finite sheet charge. If instead the voltage contains a step, two neighbouring density bins can acquire large opposite entries whose sum tends to zero. Their first moment stays finite. A useful verification uses both $\int \rho \, dz$ and $\int z\rho \, dz$ and compares potential readings on each side. A code that preserves charge but smooths a dipole step into an incorrect zero voltage would pass the first test but fail the others.

An explicit flux audit on a nonuniform mesh

To separate geometrical signs from a particular physical equation, take a rectangular box $0 < x < L_x,\ 0 < y < L_y,\ 0 < z < L_z$ and an exactly known displacement $\mathbf{D} = Axy\,\hat{\mathbf{x}} + By^2\,\hat{\mathbf{y}} + Cz\,\hat{\mathbf{z}}$. The constants have units making each term C/m². Its divergence is $(A + 2B)y + C$. Integrating over the volume gives total free charge $L_x L_z[(A + 2B)L_y^2/2 + CL_y]$. At $x = 0, y = 0, z = 0$ the relevant normal components vanish. At $x = L_x$, integrating over the face gives

$AL_xL_zL_y^2/2$. The $y = L_y$ face gives $BL_xL_zL_y^2$, and the top gives $CL_xL_yL_z$. Addition reproduces the volume calculation, including its factors of two.

Choose numerical constants by specifying $L_x = L_y = L_z = 1$ m, $A = 2$ nC/m⁴, $B = 3$ nC/m⁴ and $C = 4$ nC/m³. The three positive-face contributions are 1, 3 and 4 nC, for total 8 nC. Split the volume at $y = 0.30$ m. The bottom subvolume has volume charge $\int_0^{0.3}(8y + 4)\,dy = 1.56$ nC. Its new upper face carries $B(0.30)^2 = 0.27$ nC. The bottom section of the $x = L_x$ face carries 0.090 nC and that of the top carries 1.20 nC. Their sum is 1.56 nC. For the upper subvolume the same internal face must be counted with the opposite sign. Physical flux is single-valued on that shared face even if the cells on its sides have very different thicknesses.

Suppose a program evaluates the shared face once using a derivative from the lower cell and independently using another from the upper cell. A mismatch of 0.03 nC in those estimates generates a global charge defect when summing cells, although each side may display a locally plausible gradient. A conservative scheme stores one oriented face flux and uses the same number with opposite signs in both cells. Flux conservation can then hold to floating-point accuracy on any partition, without guaranteeing the shared flux is physically accurate. In our manufactured example a shared wrong value would move the two cell errors in opposite directions and leave the total exact. Compare the two individual subvolume charges as well as the global one.

Is this manufactured $\mathbf{D}$ an electrostatic vacuum solution? In a homogeneous scalar dielectric, $\nabla \times \mathbf{D}$ would have to vanish for a purely electrostatic field. Its z component here is $-Ax$. Thus the displacement tests the divergence theorem and a face-flux data structure perfectly, but it does not simultaneously test a potential-gradient solver with constant permittivity. For the latter choose $\mathbf{D} = -\epsilon\nabla V$ from an explicit potential before manufacturing the charge. This example separates validation of a mathematical operation from validation of a whole physical model, an important distinction when a beautifully closed flux budget is the only available diagnostic.

Reconstruct a field from its divergence and curl in a periodic measurement cell

Suppose a three-dimensional field map is tabulated on a periodic cube and you know two derived maps: its divergence d and its curl $\mathbf{w}$. Can these determine the original field? Fourier analysis gives a practical construction and reveals the missing information. For a nonzero spatial wavevector $\mathbf{k}$, use the convention $\mathbf{F}(\mathbf{x}) = \sum_{\mathbf{k}} \mathbf{F}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}$. Then $d_{\mathbf{k}} = i\mathbf{k} \cdot \mathbf{F}_{\mathbf{k}}$ and $\mathbf{w}_{\mathbf{k}} = i\mathbf{k} \times \mathbf{F}_{\mathbf{k}}$. Decompose $\mathbf{F}_{\mathbf{k}}$ into its component parallel to $\mathbf{k}$ and the remaining transverse component. Elementary vector identities give

$$\mathbf{F}_{\mathbf{k}} = -\frac{i\mathbf{k}}{k^2} d_{\mathbf{k}} + \frac{i}{k^2}\mathbf{k} \times \mathbf{w}_{\mathbf{k}} \qquad (\mathbf{k} \neq 0).$$

The first term supplies the requested divergence and zero curl. The second supplies the requested curl when $\mathbf{k} \cdot \mathbf{w}_{\mathbf{k}} = 0$, a compatibility requirement because the divergence of a curl vanishes. At $\mathbf{k} = 0$ both operations annihilate the constant field. Its mean must be measured independently. In a nonperiodic region suitable boundary data enter as well. This is a constructive version of the statement that divergence and curl determine a field only after specifying its domain, boundary information and any harmonic freedom.

For a concrete field take $\mathbf{F} = (A\cos kx, B\sin kx, C)$. Its divergence is $-kA\sin kx$, its only curl component is $w_z = kB\cos kx$, and the constant offset $C\hat{\mathbf{z}}$ drops out of both maps. The first input alone reconstructs the x component and the second the y component. A purely electrostatic solver may legitimately require $B = 0$ but will not recover C from the local charge distribution alone. A uniform applied field can be imposed by suitable distant electrodes. A strictly periodic scalar potential cannot represent a nonzero mean field without separately adding a linear background.

A grid-based reconstruction needs its own consistency check. A centred difference on a periodic grid maps e^{ikx} to $i\sin(kh)e^{ikx}/h$, not exactly ike^{ikx}. A source plane wave at the highest alternating grid mode $k = \pi/h$ has zero centred derivative at every node despite alternating sign. Treating discrete divergence measurements as exact continuum coefficients near that grid scale would infer falsely that an oscillatory field has vanished. Use a consistent discrete operator, stagger quantities if appropriate, and restrict quantitative interpretation to adequately sampled modes. The continuous Helmholtz identity is exact; the data acquisition and differentiation scheme decides which components of that identity are actually observable.

Enclose different amounts of a regularised source

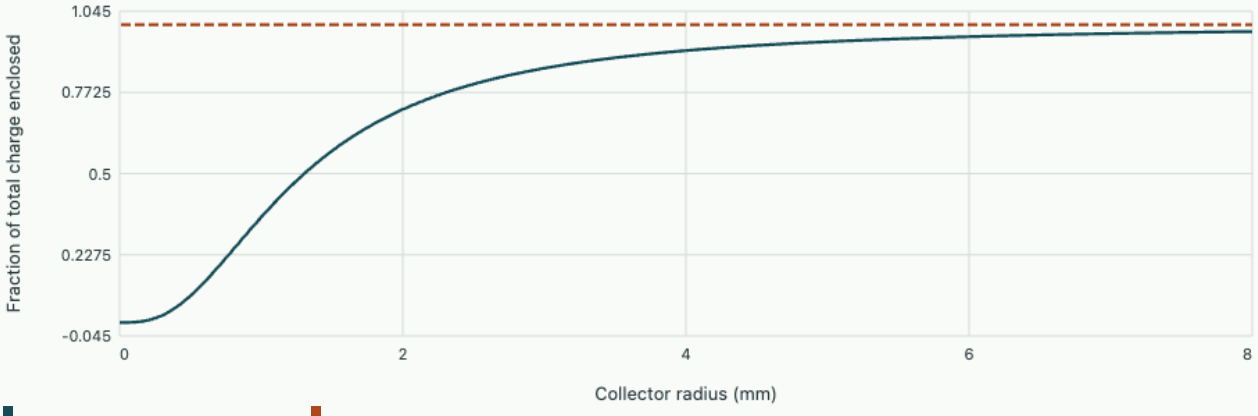
Predict how the flux of a small sphere changes as it crosses the charge cloud. Compare the enclosed fraction with the local source density.

Open this experiment online

Try the calculation

Core radius a (mm) 1

Collector radius R (mm) 2



At 2 mm the collector encloses 71.6% of the charge. Doubling its radius would enclose 91.3%.

Tutorial · turn field values into reliable local measurements

Reconstruct a sheet from a voltage cusp

Two electrodes at $z = \pm L$ are grounded. Inside the gap a measured potential is $V_0(1 - |z|/L)$. Find both fields and reconstruct the source, then explain what a centred finite difference reports at the cusp.

▼ **Work through the solution**

For $z < 0$, $E_z = -V_0/L$; for $z > 0$, $E_z = +V_0/L$. The jump gives a sheet $\sigma = 2\epsilon_0 V_0/L$ at zero. A centred estimate of the first derivative there returns zero by symmetry, reflecting the arithmetic mean of the two one-sided fields. It cannot be interpreted as either physical one-sided limit. The centred second difference is $-2V_0/(Lh)$, so the inferred cell density grows as $1/h$. Multiplying by a cell width h returns the finite sheet charge. This is the correct quantity to compare when refining the grid.

With $V_0 = 2$ V and $L = 2$ mm the fields have magnitudes 1000 V/m and the sheet charge is 17.7 nC/m². Reversing V_0 reverses both field vectors and the sheet sign without moving the cusp.

Can a voltage map distinguish a saddle from a charge?

Near the origin a section shows $V(x, 0, 0) = V_c - Ax^2$, $A > 0$. Give a source-free three-dimensional extension and a second extension consistent with uniform positive charge. Propose a minimum additional measurement given no prior symmetry.

▼ **Work through the solution**

The first is $V_c - Ax^2 + Ay^2$, independent of z. Its Laplacian vanishes. The second is $V_c - A(x^2 + y^2 + z^2)$, whose Laplacian is $-6A$, giving $\rho = 6\epsilon_0 A$. Both match the entire original line. Curvatures along y, z together with the given x-curvature determine the Laplacian. A small enclosing surface measuring normal electric flux provides the same information integrated over a volume and does not require numerical second differentiation. Neither a second derivative along one line nor a visibly crowded pair of voltage contours is a universal charge detector.

Further reading

- Vary the width and collecting radius of a regularised charge
- Fließbach, chapter 3 · distributions
- Griffiths, 5th ed., chapter 1 · vector analysis and the delta function

Electrostatic field theory and measurement · 3. When do boundary measurements determine a unique electrostatic field?

A boundary-value solver may converge while solving the wrong electrical experiment. Establishing which potentials, charges and fluxes are fixed gives a uniqueness proof and a way to turn an approximate field into an upper bound.

Specify the experiment before specifying the equation

An insulating region occupies a bounded volume Ω with smooth conducting or insulating boundaries. A known free charge density ρ_f is embedded in the region. For a positive scalar linear permittivity $\epsilon(\mathbf{x})$, electrostatics reduces to

$$-\nabla \cdot (\epsilon \nabla V) = \rho_f.$$

To turn this into an experiment one must decide how each boundary is prepared. A conductor attached to a source supplies a fixed voltage; an isolated conductor keeps a fixed total charge while becoming an equipotential; an insulating surface with known displacement flux prescribes a normal component. These preparations require different mathematical data. The same electrode shape can participate in any of them.

At an interface between two ordinary dielectrics the potential is continuous in the absence of a prescribed dipole sheet. Integrating Gauss's law across the interface gives

$$(\mathbf{D}_2 - \mathbf{D}_1) \cdot \mathbf{n} = \sigma_f,$$

where $\mathbf{n}$ points from medium 1 to 2. Thus the appropriate field problem can be posed either in a weak integral form across both materials or as two equations joined by voltage and flux conditions. The latter view is especially helpful when a measured charge appears on just one internal sheet. A jump in permittivity alone does not permit an arbitrary jump in voltage.

Prove uniqueness by testing the difference, not the solution

Suppose two solutions V_1, V_2 fit identical prescribed volume sources, material data and boundary voltages. Define $w = V_1 - V_2$. It satisfies $\nabla \cdot (\epsilon \nabla w) = 0$ and vanishes at the boundary. Multiply by w and apply the product rule to $\nabla \cdot (w \epsilon \nabla w)$. Integration yields

$$\int_\Omega \epsilon |\nabla w|^2 \, dV = \oint_{\partial\Omega} w \epsilon \, \partial_n w \, dA.$$

The boundary term is zero. The integrand on the left is nonnegative and $\epsilon > 0$, so the gradient vanishes wherever the material is regular. The function must be constant on each connected component; its fixed boundary value makes it zero. This proves that at most one field solves the stated problem. It does not itself construct a solution. It converts a difficult task of checking every possible competitor into checking the PDE and boundary data for one candidate.

The same reasoning works for a symmetric positive-definite permittivity tensor: replace $\epsilon|\nabla w|^2$ by $(\nabla w)^T \boldsymbol{\epsilon} \nabla w$. It also applies with internal material interfaces. Integrate separately on either side. Their face terms cancel because w is continuous and the prescribed free sheet makes the jump in the difference flux zero. An unaccounted interface charge would supply precisely the term that invalidates this cancellation.

If only the normal flux is fixed everywhere, the boundary difference flux vanishes and the left integral again gives a constant w. Constants are now indistinguishable by the boundary data; the electric field remains unique. Gauss's law also imposes compatibility $\oint_{\partial\Omega} \mathbf{D} \cdot \mathbf{n}\, dA = \int_\Omega \rho_f\, dV$. Choosing flux values whose sum disagrees with the enclosed charge produces no solution, even after choosing a voltage reference. Numerically this can appear as a residual that never falls below a threshold, or as a solver modifying the source silently to make the equations compatible.

Prove uniqueness when only total conductor charge is known

Take a region surrounding several isolated conductors. The value on each metal is constant but initially unknown. For the difference of two candidate solutions denote these constants by w_i. Then each conductor's contribution to the right-hand side of the energy identity factors as $w_i \int_{S_i} \epsilon \partial_n w\, dA$. With outward normal from the dielectric into the metal this integral is the difference in the conductor's free charge. Prescribing total charge for each metal makes every such term vanish individually. Provided a reference conductor or condition at infinity removes an overall constant, the potential is unique.

Knowing only the sum of charges over several electrically disconnected metals can be insufficient. For two isolated spheres, allow one to take charge $+\delta q$ and the other $-\delta q$ while fixing their total at zero. Their equipotential values can change relative to each other, and an external field may appear. If instead they are joined by an ideal wire they must share one voltage, so the common boundary contribution factors against their total charge. The proof itself identifies the physical distinction between strings of electrically connected metal and separate conductors.

A functional whose variation is Gauss’s law

For fixed boundary voltage consider

$$\mathcal{I}[V] = \int_\Omega \left[\frac{1}{2} \nabla V \cdot \epsilon \nabla V - \rho_f V \right] dV.$$

The two meanings of V in this notation are distinguishable from context: $V(\mathbf{x})$ is potential and dV is a volume element. Add a trial perturbation η that vanishes on the fixed-voltage boundary. The first-order change is $\int [\nabla \eta \cdot \epsilon \nabla V - \rho_f \eta] dV$. Integrate the first term by parts. Requiring zero for every admissible η gives

$$-\nabla \cdot (\epsilon \nabla V) = \rho_f.$$

On a boundary portion where flux rather than voltage is prescribed, an additional linear boundary contribution accounts for the specified normal derivative. It is not generally correct to hold every trial voltage zero there.

For an exact fixed-voltage solution V_* expand $\mathcal{I}[V_* + \eta]$. The first-order terms vanish, leaving

$$\mathcal{I}[V_* + \eta] - \mathcal{I}[V_*] = \frac{1}{2}\int_\Omega \nabla\eta \cdot \epsilon \nabla\eta \, dV.$$

The exact solution minimises the functional, and the gap measures the field error in an energy-weighted norm. If $\rho_f = 0$ the functional is the stored field energy. With embedded sources the linear term matters; the value of the functional is generally not just field energy. Restrict the trial fields to a finite space and minimize there. This is the variational origin of a potential-based finite-element method. A small change in the energy between two grids can coexist with a noticeably changing local peak field because the integral weights errors by volume.

A complete numerical comparison without a black box

Take a unit area filled by two dielectric slabs each of thickness $d = 1$ mm and relative permittivities 2 and 8. Set one electrode at 10 V, the other at zero, with no free charge inside. A trial straight line through the whole gap obeys both electrode voltages. Its energy per area is $\epsilon_0(2+8)d(5000)^2/2 = 1.107\ \mu\mathrm{J/m^2}$ after unit conversion. The exact interface problem requires constant normal displacement. It gives voltage drops of 8 and 2 V, fields of 8000 and 2000 V/m and energy per area $0.708\ \mu\mathrm{J/m^2}$. A single continuous slope produced too much energy by about 56%. The trial is admissible despite its incorrect interface flux, because stationarity, rather than a kinematic constraint, enforces that flux.

Recover the exact answer as a one-variable minimisation. Let u be the voltage at the interface measured relative to the grounded end, with the first material next to the 10 V electrode. The energy per area is $\epsilon_0[2(10-u)^2 + 8u^2]/(2d)$. Differentiation gives $-2(10-u) + 8u = 0$, hence $u = 2$ V. The derivative with respect to an interior trial voltage measures the flux imbalance between neighbouring elements. This makes the algebraic finite-element equation a local charge balance. Conversely a trial that fails to equal an electrode's prescribed voltage is inadmissible, even if its energy appears lower than the computed answer.

A maximum principle as a diagnostic independent of energy

In a charge-free region with uniform permittivity the potential at the centre of a sphere equals its mean over the sphere. To see this, differentiate the angular average with respect to its radius. The result is proportional to the total normal derivative, which is zero by Gauss's law for an empty sphere. Shrink the sphere to its centre to identify the constant average. Therefore a nonconstant harmonic function cannot have a strict interior maximum or minimum. If the centre assumes the maximum value, all sufficiently nearby directions must assume that same value to keep the average equal; continuing through the connected region forces constancy.

For source-free Poisson solvers an interior voltage above all fixed outer values indicates an issue with boundary data, source terms or discretisation. But the same observation is entirely possible with positive internal charge, as the one-dimensional grounded-plane parabola demonstrates. The maximum principle concerns potentials; it does not say that a field magnitude, an individual component in an inhomogeneous medium, or the potential restricted to a line must remain between the corresponding boundary values on that line.

A thin dielectric coating turns a boundary voltage into a Robin condition

Suppose a grounded metal enclosure is lined with an insulating film of thickness d_f and permittivity ϵ_f. We want the electrostatic potential in an interior region large compared with the film thickness. Simply declaring the film's exposed face grounded neglects its voltage drop. If the tangential scale of the interior field and the radius of curvature of the surface are both much larger than d_f, the film behaves locally like a plane capacitor with capacitance per area $c_f = \epsilon_f / d_f$. Let $\mathbf{n}$ point from the interior into the film, let the backing metal be at V_b, and denote the interior-side surface potential by v. Without free charge at the contact, the outward displacement from the interior equals $c_f (v - V_b)$. Since $\mathbf{D} = -\epsilon \nabla V$ inside, the effective boundary condition is

$$\epsilon \, \partial_n V + c_f (V - V_b) = 0.$$

This is a Robin boundary condition: both the potential and its derivative enter. The coefficient c_f has units F/m², so both terms are surface charges. For $c_f \to \infty$ a bounded normal field forces V to the backing voltage. For $c_f \to 0$ a finite voltage difference can sustain almost no normal flux. The two familiar boundary ideals are limiting cases of one measurable material preparation. The derivation requires the film to contain no embedded free charge and to respond linearly; measured leakage over long times adds a current-balance problem.

The variational proof above extends naturally. Add $\frac{1}{2}\int_{S_f} c_f (V - V_b)^2 dA$ to the interior functional. Its first variation at the exposed face gives $\epsilon \partial_n V + c_f (V - V_b)$ times the arbitrary boundary variation. The added quadratic term is exactly the electrostatic energy stored in the eliminated thin layer under the one-dimensional approximation. For the difference of two solutions the energy identity becomes

$$\int \epsilon |\nabla w|^2 dV + \int c_f w^2 dA = 0.$$

Either a prescribed reference voltage or a positive-area Robin patch with $c_f > 0$ removes the constant null mode. This follows without pretending that the exposed potential itself was prescribed.

Consider a planar interior region of thickness 2.0 mm filled with $\epsilon_i = 4\epsilon_0$, separated from its 0 V backing by a 50 µm film with $\epsilon_f = 2\epsilon_0$. Set the opposite electrode to 10 V. The exact one-dimensional displacement is

$$D = 10 / [L/\epsilon_i + d_f/\epsilon_f].$$

The effective electric length in vacuum units is $0.50 + 0.025 = 0.525$ mm, hence $D = 169$ nC/m² and the film takes $10(0.025/0.525) = 0.476$ V. Imposing a direct 0 V boundary at the exposed film face would overpredict displacement by 5%. Solving with the Robin condition instead reproduces the series result exactly in the flat one-dimensional model. For a patterned electrode with wavelength comparable to film thickness, tangential field in the film cannot be ignored and the effective boundary response depends on spatial frequency.

What a voltage error bound does and does not control

If two source-free solutions have slightly different Dirichlet data, their difference is harmonic, not necessarily zero. Let their boundary discrepancy have maximum magnitude δ. The maximum principle applied to the difference and its negative guarantees $|V_1 - V_2| \le \delta$ throughout the region. This gives a stability estimate for voltage even where the exact solution is unknown. It requires an error bound on the whole boundary and identical sources inside. Collocation residuals evaluated only at a few points do not establish the premise between those points.

Does the same number bound the error in local electric field? Consider the upper half-space with boundary error $\delta\cos(kx)$. The harmonic difference at height h is $w = \delta e^{-kh}\cos(kx)$. Immediately at the boundary the field-error amplitude is $k\delta$, which can be arbitrarily large as the spatial wavelength shrinks even though the voltage error stays below δ. At nonzero h the field-error amplitude for this single mode is $k\delta e^{-kh}$. For fixed h it is maximal at $k = 1/h$, with value $\delta/(eh)$. Depth regularises the sensitivity to ever finer ripples. A global bound on arbitrary superpositions requires accounting for how their amplitudes combine; maximising one sinusoid does not prove that every boundary function attains the same bound.

Suppose a nominally grounded instrument panel has a sinusoidal ripple of peak voltage 5 mV and wavelength 1.0 mm. At the surface it creates an electric-field amplitude of 31.4 V/m. Just 1.0 mm away its amplitude is multiplied by $e^{-2\pi} = 0.00187$, leaving 0.0587 V/m. A sensor directly at the surface and one millimetre inside may therefore make very different accuracy demands on the boundary model, despite the same nominal voltage tolerance. A field solver claiming agreement of all electrode voltages to 5 mV still needs a spatial smoothness or resolution statement before claiming local field accuracy.

A useful related identity controls energy error when an approximate potential has an interior residual. Let V_h satisfy the exact boundary values, and define $r = \rho + \nabla\cdot(\epsilon\nabla V_h)$. With $e = V_* - V_h$ zero on the boundary,

$$-\nabla\cdot(\epsilon\nabla e) = r.$$

Multiplying by e gives $\int \epsilon|\nabla e|^2 dV = \int re\, dV$ for a smooth admissible approximation. If a piecewise field has jumps at internal element faces their residual contributions must be included. A tiny volume residual computed inside each piece can coexist with incorrect flux matching at every face. In the two-layer example, a straight-line trial has zero second derivative in both layers, yet its displacement mismatch at the interface is exactly what forces a different physical solution.

Why a conductor becomes equipotential when charge can move

So far conductor voltages or charges entered as boundary conditions. A related variational argument explains the equipotential condition itself. Imagine charges on one connected conductor that may redistribute while their total stays fixed. Treat the final material state as ideal and static, with no ongoing currents or dissipative constraints. The change in electrostatic field energy for an infinitesimal change in free surface charge is $\delta U = \int_S V(\mathbf{x})\delta\sigma(\mathbf{x})\, dA$. This follows either by slowly bringing in each incremental charge at its current potential or by varying the symmetric double integral $\frac{1}{2}\int\int \rho(\mathbf{x})G(\mathbf{x},\mathbf{y})\rho(\mathbf{y})d^3x\, d^3y$. The half in the energy disappears upon variation because either source factor may vary.

Move a small positive charge dq from a patch at potential V_A to another at V_B, preserving total conductor charge. The change is $(V_B - V_A)dq$. If the two voltages differ, one direction of transfer lowers the energy. At equilibrium every such redistribution must have zero first-order energy change, hence the potential is constant over the connected conductor. For disconnected conductors whose individual totals are fixed the same reasoning gives a different allowed constant on each. If an ideal wire joins them, transfers across the wire become admissible and their two constants equalise. This is a statement about electrical connectivity and constraints, not merely the smoothness or shape of the conducting material.

The second-order change for a neutral redistribution is its own nonnegative field energy. For a fixed-charge isolated conductor it follows that the electrostatic equilibrium distribution minimises energy over all allowed placements of the same surface charge. This is a different trial principle from prescribing electrode voltages and varying a potential throughout the space. A trial charge that obeys a total-charge constraint need not produce an exactly equipotential conductor; its excess energy measures a failure of equilibrium. Confusing these two variational principles can switch the direction of an inferred capacitance bound.

As a measurable limit, take two small spheres far apart, with approximate capacitances $C_1 = 4\pi\epsilon_0 a_1$ and $C_2 = 4\pi\epsilon_0 a_2$ to infinity. Temporarily neglect mutual coupling. With total charge Q, write $U(q) = q^2/(2C_1) + (Q-q)^2/(2C_2)$. Differentiating gives $q/C_1 = (Q-q)/C_2$, the equipotential solution after wiring. For radii 10 and 20 mm and $Q = 30$ pC this gives shares 10 and 20 pC at a common voltage about 9.0 V. Putting all the charge initially on the smaller sphere would give energy $Q^2/(2C_1)$. After connection the energy is $Q^2/[2(C_1 + C_2)]$, one third as large. The missing field energy leaves during transient currents through radiation or dissipation; static uniqueness alone describes neither time history nor dissipation partition. As sphere spacing is reduced, mutual capacitance must be included and the simple radius ratio becomes an approximation to test rather than an exact rule.

Check a potential independently of its residual

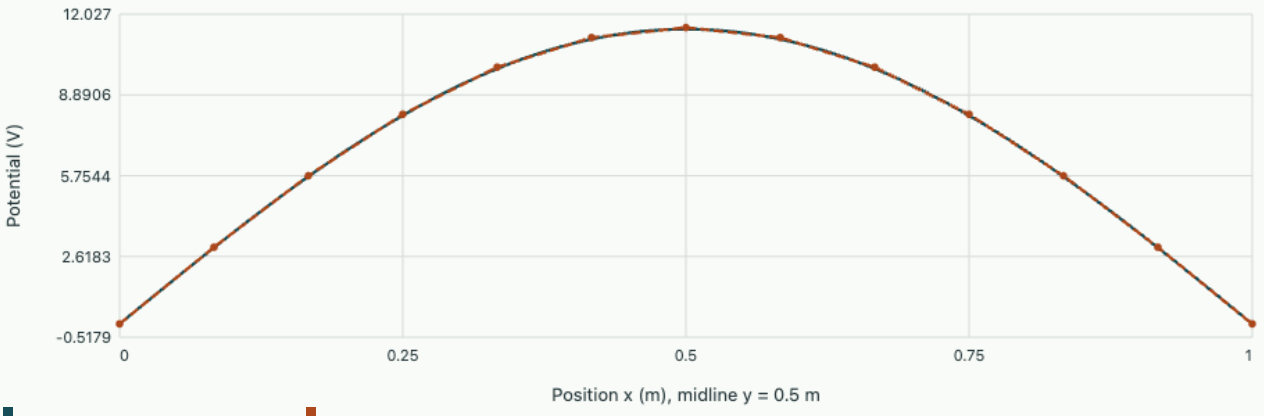
Refine a grounded square with known boundary data. Compare centre voltage with the series solution as well as the discrete residual. The accompanying OneSimulation model solves the same field problem.

Open this experiment online

Try the calculation

Cells per side 12

Charge amplitude ρ_0 (nC/m³) 2



Continuous solution Five-point grid solution

Centre potential 11.50889 V on the grid, 11.44331 V in the continuum. Nodal error 0.06558 V.

Download the OneSimulation model

In your OneSimulation environment, run the model with these inputs.

```
python grounded-square.py --cells 12 --charge 2
```

Tutorial · can the proposed computation represent this experiment?

Find a compatible flux and determine the remaining freedom

A slab $0 < x < L$ contains uniform positive free charge ρ. At $x = 0$ the outward displacement is measured as $D_n = -D_0$. Determine what flux must be prescribed at L, the electric field throughout, and one piece of information needed for the absolute potential.

▼ **Work through the solution**

Positive x displacement at the left is D_0 because that face's outward normal points towards negative x. Integrating $dD_x/dx = \rho$ gives

$$D_x = D_0 + \rho x.$$

Therefore the outward flux at L is $D_0 + \rho L$. Their sum per unit area equals ρL. In constant permittivity $E_x = (D_0 + \rho x)/\epsilon$ and $V(x) = V(0) - (D_0 x + \rho x^2/2)/\epsilon$. A reference voltage fixes $V(0)$; none can fix an incompatible pair of flux measurements. A negative right-hand outward reading with very large positive enclosed charge would contradict Gauss's law before one even solves for V.

Bound a capacitance with an intentionally poor trial

In a charge-free parallel-plate vacuum gap of thickness L, use $V(x) = V_0(x/L)^2$ to bound the capacitance per area. Compare with the exact linear voltage. How does improving the potential shape turn into a monotonic bound?

▼ **Work through the solution**

The trial satisfies the endpoints. It gives energy per area $\epsilon_0 \int_0^L (2V_0 x/L^2)^2 dx/2 = 2\epsilon_0 V_0^2/(3L)$, hence a capacitance estimate $4\epsilon_0/(3L)$. The exact value is ϵ_0/L, lower by the predicted variational inequality. Expanding a trial space to include both the old functions and new ones cannot raise its optimal energy because the old minimiser remains available. This monotonic property requires nested trial spaces and consistent integration; changing quadrature or replacing the physical geometry changes the comparison.

Further reading

- Perturb a potential and measure the quadratic energy penalty
- Recover an interface charge using independent field measurements
- Griffiths, 5th ed., §§3.1.5–3.1.6 · uniqueness theorems

Electrostatic field theory and measurement · 4. Images in wedges, cylinders and dielectric interfaces

An image source is a construction for a particular physical domain and boundary condition. Three examples show how geometry determines the image position, how material response sets its amplitude, and why an apparent success on part of a boundary is insufficient.

Start with a domain in which the fictitious sources are excluded

Suppose we have a point charge in an open region bounded by conducting metal. Inside the vacuum we know where all real sources lie. Near each real point source the potential must have its prescribed Coulomb singularity. Elsewhere it must be harmonic. A method of images extends an expression for V beyond the physical region using additional sources chosen to enforce boundary values on the metal. The added singularities must lie outside the field domain. If all boundary data and the behaviour at infinity agree, the uniqueness theorem guarantees the solution there. The extension says nothing about the actual field inside the metal.

There are three separate checks. First locate the singularities. Next test every boundary value as a function over the entire surface, not at one point of symmetry. Finally check the decay or far-field background. Using a few image charges to make the voltage zero at a finite list of collocation points gives an approximation unless a separate argument enforces the boundary everywhere. It can still be an effective numerical method when one monitors the untested parts of the surface.

A family of wedges that permits a finite reflection construction

Take two grounded half-planes meeting along z, with vacuum opening angle $\alpha = \pi/m$ for positive integer m. A real point charge sits in the wedge at cylindrical coordinates (s_0, ϕ_0, z_0), $0 < \phi_0 < \alpha$. Reflection across either grounded plane reverses the charge sign. Two successive reflections amount to rotation by 2α with no sign reversal. Hence a useful set consists of m charges $+q$ at $\phi_0 + 2j\alpha$ and m charges $-q$ at $-\phi_0 + 2j\alpha$, $j = 0, \ldots, m-1$, all at the same s_0, z_0. Only the original lies in the physical wedge. Each metal face pairs equal-distance charges of opposite signs, so the summed potential vanishes on both faces.

For $\alpha = \pi/3$ there are five images in addition to the real charge. If the real charge is on the bisector $\phi_0 = \pi/6$, the total configuration is six alternating charges on a circle. All forces tangential to the bisector cancel by symmetry. Consider any other charge q_j displaced through azimuthal angle Δ from the real charge. Its separation is $2s_0|\sin(\Delta/2)|$, and the radial projection of the displacement vector from that image to the real source is $s_0(1-\cos\Delta)$. Its radial force on q is therefore $qq_j/[16\pi\epsilon_0 s_0^2|\sin(\Delta/2)|]$. Summing the five alternating terms gives

$$F_s = \frac{q^2}{16\pi\epsilon_0 s_0^2}\left(-5 + \frac{4}{\sqrt{3}}\right).$$

The sign points toward the wedge vertex. The nearest negative images at $\pm 60^\circ$ each contribute -2 in the bracket, the two positive ones at $\pm 120^\circ$ give $+2/\sqrt{3}$ each, and the opposite negative image gives -1. The radial scaling $1/s_0^2$ follows independently from the geometric similarity of the wedge. For arbitrary opening angles the repeated images can fail to terminate, or introduce spurious singularities in the vacuum

wedge. Separation in wedge coordinates then offers a better construction. A valid result for one rational angle cannot simply be extrapolated to a neighboring angle.

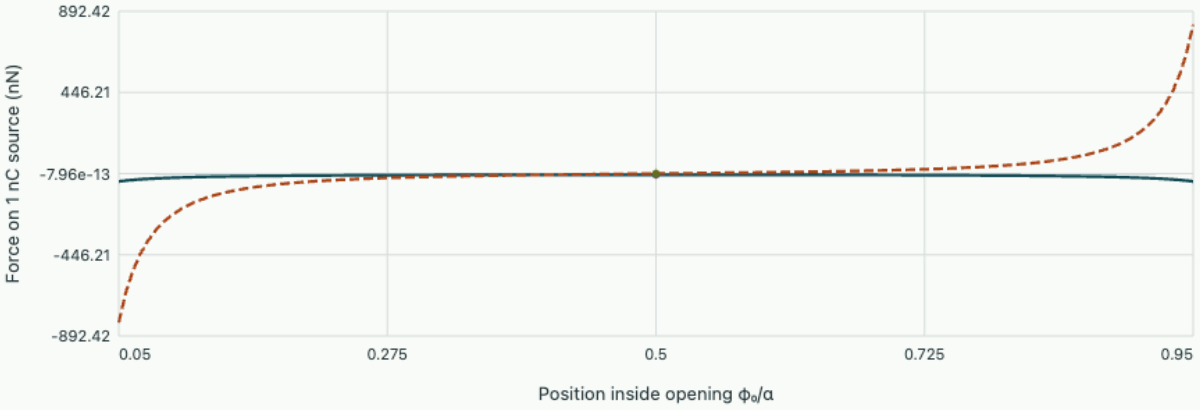
Try it · Change the opening and move the charge

Keep m integer, as required by this finite image construction. Move the source off the bisector and identify the nearer attracting face. Double the radial distance and check the inverse-square force scaling.

Open this boundary experiment

Try the calculation

Wedge denominator m	3
Radial source distance (m)	1
Selected fraction of opening	0.5



Opening 60°; 5 fictitious images. At ϕ₀/α=0.5, radial force -6.0455 nN and azimuthal force -2.6645e-15 nN. On the bisector the azimuthal component vanishes.

Notebook: reproduce the exact boundary constructions and the 3D collocation model

Find an image line from a constant distance ratio

An infinite metal cylinder of radius a is grounded. Outside it a parallel positive line charge λ is located a transverse distance $d > a$ from the centre, on the x-axis. Choose an interior image $-\lambda$ at $s = a^2/d$. For a surface point at polar angle ϕ, its distance from the real line is $R_q = \sqrt{a^2 + d^2 - 2ad\cos\phi}$. Its distance from the image satisfies

$$R_i = (a/d)R_q.$$

A line potential varies as a logarithm, so the ratio of distances, rather than a weighted sum of inverse distances, must be constant on the metal. With zero on the cylinder the exterior expression is

$$V = \frac{\lambda}{2\pi\epsilon_0}\ln\left(\frac{dR_i}{aR_q}\right).$$

At large transverse radius the logarithm approaches $\ln(d/a)$, yielding a constant nonzero distant voltage. This is consistent with the effectively two-dimensional infinite geometry and the chosen grounding convention. Insisting on zero potential both on the grounded cylinder and at infinity without an external return changes the problem. In the physical exterior the electric field nonetheless decays.

At the real line the image field attracts with magnitude per length $\lambda^2/[2\pi\epsilon_0(d - a^2/d)] = \lambda^2 d/[2\pi\epsilon_0(d^2 - a^2)]$. For $d = a + h$ and $h \ll a$, the image separation tends to $2h$, yielding the plane result for a line. For $a = 10$ mm, $d = 25$ mm and $\lambda = 1.0$ nC/m the attraction is 0.856 µN/m. A conducting sphere near a point charge uses the same geometric inversion $s = a^2/d$ but the image amplitude there is $-qa/d$ because its Green function is $1/R$. Dimensions of the source and its fundamental solution determine which rules transfer between the two calculations.

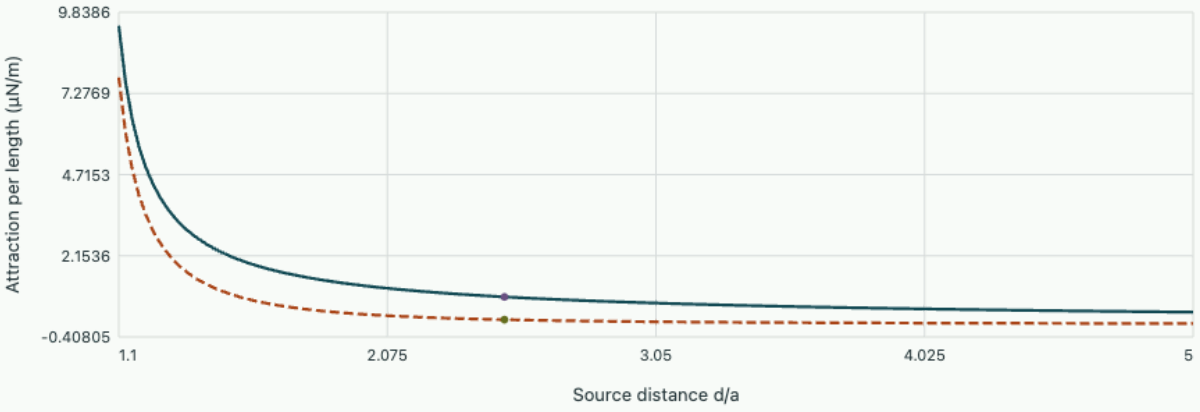
Try it · Compare grounded and isolated cylinders

Set a=10 mm, λ=1 nC/m and d/a=2.5 to recover the example above. Move farther away and compare the two asymptotic force laws. The two curves differ in the total conductor charge constraint.

Open this boundary experiment

Try the calculation

Cylinder radius a (mm)	10
Line charge (nC/m)	1
Selected d/a	2.5



At d=25 mm: grounded attraction 0.85596 μN/m, isolated attraction 0.13695 μN/m. The image lies 4 mm from the axis; electrical preparation changes the distant force law.

Notebook: reproduce the exact boundary constructions and the 3D collocation model

Try it · Compare grounded and isolated image spheres

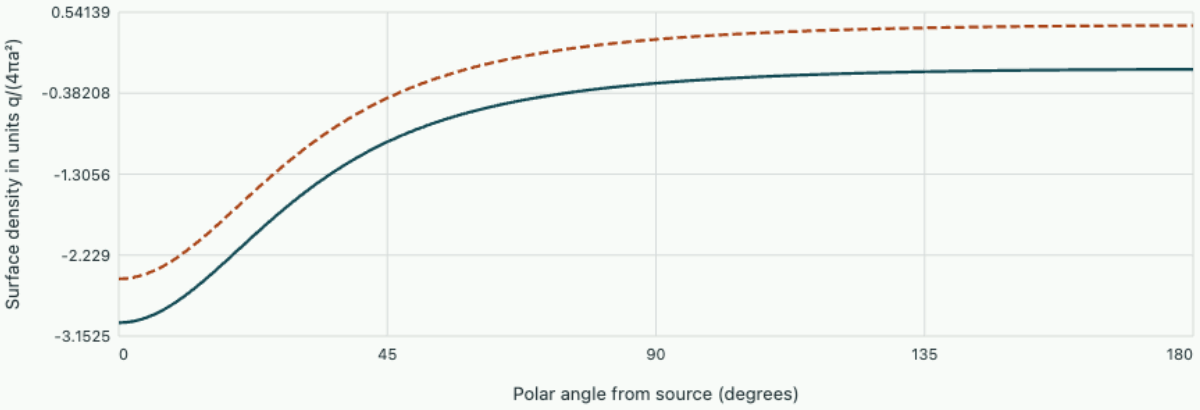
Change the source position and compare the induced-charge constraints before trying the general collocation system below.

Open the exact sphere comparison

Try the calculation

Source distance d/a **2**

Isolated conductor charge Qc/q **0**



Grounding draws net charge -0.5 q. Forces in units q²/(4πε₀a²) are -0.2222 grounded and -0.09722 isolated; positive means away from the sphere.

An interface that supports bound charge but does not fix voltage

Let a point charge q occupy $z = d > 0$ in a medium of permittivity ϵ_1. The lower half-space has permittivity ϵ_2. Both extend to infinity and there is no free sheet on the interface. An ansatz for $z > 0$ contains the real charge plus an image q' at $z = -d$, using $1/(4\pi\epsilon_1)$ as its common prefactor. For $z < 0$, put an effective charge q'' at $z = d$ with prefactor $1/(4\pi\epsilon_2)$. That effective charge also lies outside its own solution domain. At the plane, the two distances are equal. Potential continuity gives

$$(q + q')/\epsilon_1 = q''/\epsilon_2.$$

Taking the oriented normal derivative gives $q - q' = q''$ because original and reflected source distances have opposite z-derivatives there. Solving,

$$q' = q\frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2},$$
$$q'' = \frac{2\epsilon_2 q}{\epsilon_1 + \epsilon_2}.$$

Check the equal-material limit: $q' = 0, q'' = q$. For an increasingly polarizable lower region $q' \to -q$, reproducing the grounded-plane exterior field, while the lower electric field becomes weak. For a lower-permittivity lower region the image has the same sign and the charge is repelled. Positive permittivity materials can therefore produce either direction of interface force; it is the comparison of the two constitutive laws that decides.

For $\epsilon_1 = 4\epsilon_0, \epsilon_2 = 2\epsilon_0$, the image is $q/3$. At $d = 1$ mm and $q = 10$ pC the repulsive force is $q^2/[4\pi(4\epsilon_0)12d^2] = 18.7$ nN. Moving the charge quasistatically from infinity to that position requires positive work. The interaction contribution to energy is

$$qV_{\text{image}}/2 = qq'/(16\pi\epsilon_1 d).$$

Minus its derivative with respect to d gives the same force as the real charge times the image field. The half accounts for inducing the bound charge. Treating the image as a second independently prescribed charge and differentiating its pair energy with both positions moving doubles the force.

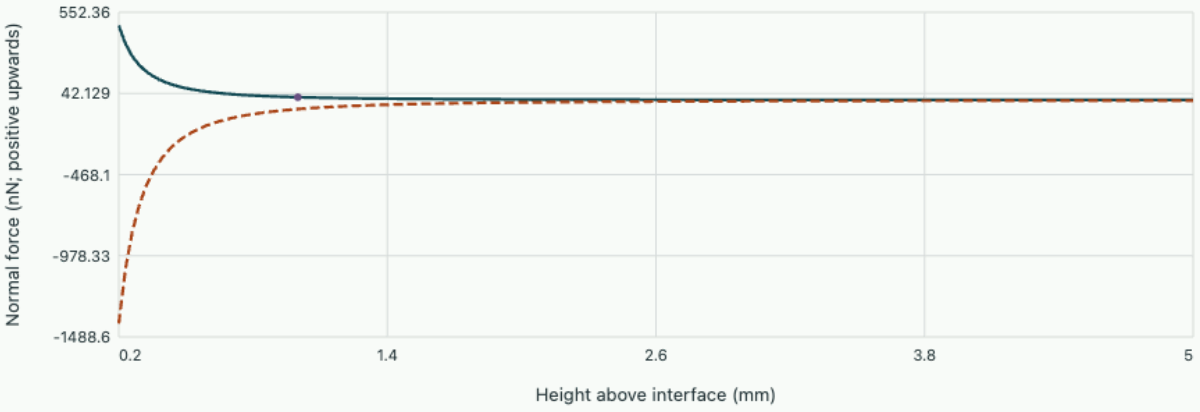
Try it · Switch attraction into repulsion

Move $\varepsilon_2/\varepsilon_1$ through one. Check the equal-material limit and compare with the grounded-plane limit. The sign denotes force along increasing z; a negative value attracts the source toward the interface.

Open this boundary experiment

Try the calculation

Permittivity ratio $\varepsilon_2/\varepsilon_1$	0.5
Charge q (pC)	10
Selected height d (mm)	1



With $\varepsilon_1=4\varepsilon_0$, $q'/q=0.33333$ and $q''/q=0.66667$. Force at 1 mm is 18.724 nN: repulsive.

Notebook: reproduce the exact boundary constructions and the 3D collocation model

How to test a numerical image representation

For a smooth conducting enclosure of less convenient shape, place hypothetical source points inside the metal, impose the desired boundary voltages at accessible sample points, and solve linearly for strengths. The physical solution must remain smooth throughout the vacuum except at genuine specified sources. Choosing numerical source positions too close to the boundary may create narrow voltage oscillations between sampled surface points. Placing them too far can make influence columns almost indistinguishable, producing enormous opposite coefficients and poor numerical conditioning.

Test the candidate on surface points deliberately withheld from the fit. In a source-free region the difference between the approximation and the exact potential is harmonic if both use the same physical singularities. Its maximum voltage error cannot exceed the largest absolute boundary discrepancy when the far boundary is also controlled. This does not directly bound the normal-field error at the conductor, which can be sensitive to small short-wavelength ripples in voltage. Check flux and force separately, for

example by integrating the Maxwell stress over a vacuum surface displaced away from difficult corners. A voltage map, a force and a conductor's total charge impose complementary demands on accuracy.

OneSimulation in 3D · Test a fitted source system between its boundary points

A grounded sphere with a=1 m is placed beside q=1 nC at z=d. The exact image charge is –qa/d at z=a²/d. In the numerical construction, choose N fictitious points sj on an interior sphere of radius αa. Require zero potential at N points bi on the physical surface:

$$\sum_{j=1}^{N} \frac{w_j}{|\mathbf{b}_i - \mathbf{s}_j|} = -\frac{1}{|\mathbf{b}_i - d\mathbf{e}_z|},$$

$$V_h(\mathbf{x}) = \frac{q}{4\pi\epsilon_0}\left[\frac{1}{|\mathbf{x} - d\mathbf{e}_z|} + \sum_j \frac{w_j}{|\mathbf{x} - \mathbf{s}_j|}\right].$$

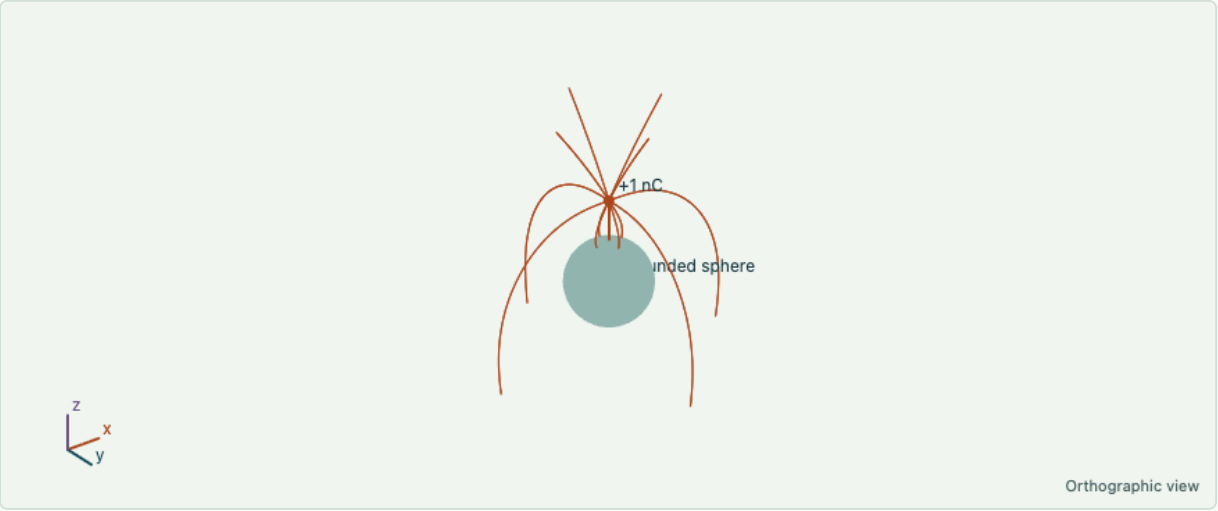
OneSimulation solves this dense collocation system in the downloadable notebook. The browser implements the same system with pivoted Gaussian elimination. All sources are inside the excluded metal; the field is evaluated outside. Its sum of numerical source strengths predicts total induced charge, independently checked against –qa/d. A tiny residual at fitted points can coexist with a substantial voltage error between them.

Refine N at fixed d and α. Then move the fictitious sources toward the boundary and observe the error between fitted points. Rotate the 3D field lines and notice that small azimuthal differences disappear with refinement. Field-line density here is chosen for readability and does not encode flux.

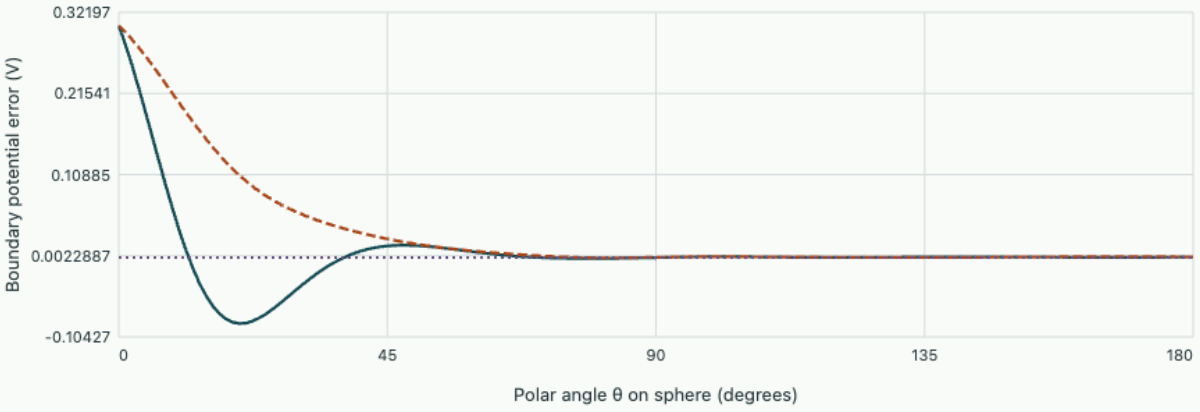
Open the 3D sphere experiment

Try the calculation

Number of source points	48
Charge distance d/a	2
Source radius / a	0.65



Numerically traced electric-field lines outside a grounded sphere (a=1 m). The computed field uses fitted interior sources; slight loss of axial symmetry is approximation error. The fictitious interior charges are not physical charges in the metal. Axes: x = x/a; y = y/a; z = z/a.



Computed surface, ϕ=0° Computed surface, ϕ=90° Exact image construction

Linear residual 5.55e-16; maximum error on 1201 withheld surface points 0.41847 V. Induced charge -0.49919 nC; exact -0.5 nC.

Download the OneSimulation model

In your OneSimulation environment, run the model with these inputs.

```
python complete-spatial-sphere.py --sources 48 --distance 2 --source-radius 0.65
```

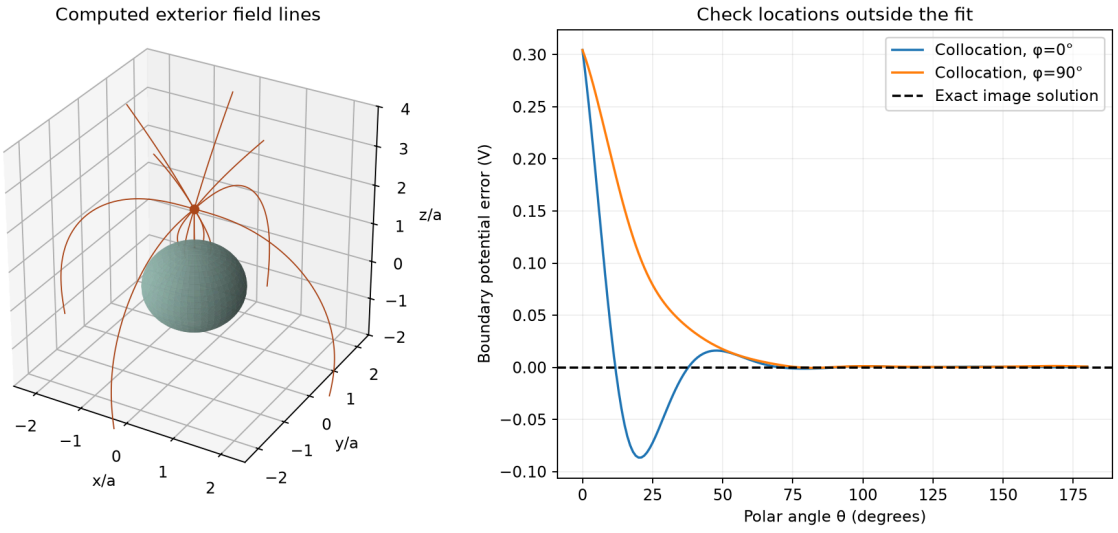
Download the executed 3D OneSimulation notebook · Python model

```
"""Grounded sphere: a 3D source-collocation solve and an independent image solution."""
import argparse
import numpy as np
import matplotlib.pyplot as plt
from onesim.numerics.linear.matrices.dense_linear_algebra import solve_dense

parser = argparse.ArgumentParser(description=__doc__)
parser.add_argument('--sources', type=int, default=48)
parser.add_argument('--distance', type=float, default=2.0, help='Real-charge distance / sphere
radius')
parser.add_argument('--source-radius', type=float, default=.65, help='Fictitious-source radius /
sphere radius')
args = parser.parse_args()
if not 24 <= args.sources <= 96 or not 1.5 <= args.distance <= 3 or not .45 <= args.source_radius <=
.85:
    parser.error('Use 24–96 sources, distance 1.5–3, and source radius 0.45–0.85')

def fibonacci_sphere(n, offset=0.0):
    j = np.arange(n)
    z = 1 - 2*(j+.5)/n
    phi = j*np.pi*(3-np.sqrt(5))+offset
    return np.column_stack([np.sqrt(1-z*z)*np.cos(phi),np.sqrt(1-z*z)*np.sin(phi),z])

boundary = fibonacci_sphere(args.sources)
sources = args.source_radius*boundary
charge = np.array([0.,0.,args.distance])
influence = 1/np.linalg.norm(boundary[:,None,:]-sources[None,:,:],axis=2)
rhs = -1/np.linalg.norm(boundary-charge,axis=1)
weights = np.asarray(solve_dense(influence,rhs))
factor = 8.9875517923  # q/(4*pi*eps0), for q=1 nC and radius=1 m

def potential(points):
    points=np.atleast_2d(points)
    return factor*(1/np.linalg.norm(points-charge,axis=1)
                   + (1/np.linalg.norm(points[:,None,:]-sources[None,:,:],axis=2))@weights)

def exact(points):
    points=np.atleast_2d(points)
    image=np.array([0.,0.,1/args.distance])
    return factor*(1/np.linalg.norm(points-charge,axis=1)
                   -1/args.distance/np.linalg.norm(points-image,axis=1))

def field(point):
    real=point-charge
    delta=point-sources
    return real/np.linalg.norm(real)**3 + (weights[:,None]*delta/np.linalg.norm(delta,axis=1)
[:,None]**3).sum(axis=0)

check = fibonacci_sphere(1201,.37)
boundary_error=float(np.max(np.abs(potential(check))))
algebraic_error=float(np.max(np.abs(influence@weights-rhs)))
print(f'Sources: {args.sources}; d/a={args.distance:g}; source radius/a={args.source_radius:g}')
print(f'Linear residual: {algebraic_error:.3e}; withheld boundary error: {boundary_error:.6e} V')
print(f'Induced charge: {weights.sum():.8f} nC; exact: {-1/args.distance:.8f} nC')
assert algebraic_error < 1e-9

fig = plt.figure(figsize=(11,4.8),layout='constrained')
ax=fig.add_subplot(121,projection='3d')
theta=np.linspace(0,np.pi,25);phi=np.linspace(0,2*np.pi,49)
T,P=np.meshgrid(theta,phi)
ax.plot_surface(np.sin(T)*np.cos(P),np.sin(T)*np.sin(P),np.cos(T),color='#94b8af',alpha=.85,linewidth=
```

```
0)
for angle in [np.pi/6,np.pi/2,5*np.pi/6,np.pi]:
    for azimuth in [0,np.pi/2,np.pi,3*np.pi/2]:

p=charge+.08*np.array([np.sin(angle)*np.cos(azimuth),np.sin(angle)*np.sin(azimuth),np.cos(angle)])
        path=[p.copy()]
        for _ in range(160):
            e=field(p);p=p+.035*e/np.linalg.norm(e);path.append(p.copy())
            if np.linalg.norm(p)<1.015 or np.linalg.norm(p)>4.4:break
        curve=np.array(path);ax.plot(*curve.T,color='#ad491d',linewidth=.8)
ax.scatter(*charge,color='#ad491d',s=30)
ax.set(xlabel='x/a',ylabel='y/a',zlabel='z/a',title='Computed exterior field lines')
ax.set_box_aspect([1,1,1]);ax.set(xlim=(-2.5,2.5),ylim=(-2.5,2.5),zlim=(-2,4))
ax=fig.add_subplot(122)
angles=np.linspace(0,np.pi,181)
for azimuth in [0,np.pi/2]:

points=np.column_stack([np.sin(angles)*np.cos(azimuth),np.sin(angles)*np.sin(azimuth),np.cos(angles)])
    ax.plot(np.rad2deg(angles),potential(points),label=f'Collocation, ϕ={np.rad2deg(azimuth):g}°')
ax.axhline(0,color='black',linestyle='--',label='Exact image solution')
ax.set(xlabel='Polar angle θ (degrees)',ylabel='Boundary potential error (V)',title='Check locations
outside the fit')
ax.legend();ax.grid(alpha=.2)
fig.savefig('complete-spatial-sphere.png',dpi=170)
```

Actual output from the default OneSimulation solve:

```
Sources: 48; d/a=2; source radius/a=0.65
Linear residual: 2.220e-16; withheld boundary error: 4.184726e-01 V
Induced charge: -0.49918756 nC; exact: -0.50000000 nC
```



Full three-coordinate collocation model. The exact grounded-sphere solution has zero boundary potential; the plotted discrepancy measures approximation error.

Derive the external capacitance of a two-wire pair

Two long parallel metal cylinders of equal radius a have centre-to-centre separation $D > 2a$. They carry equal and opposite line charges and are maintained at opposite voltages. If the surface separation is

comparable to the wire radius, putting equivalent lines on the conductor axes fails to make each circumference equipotential. Instead exploit the geometry of a potential difference of two line sources. Place $+\lambda$ at $x = s$ and $-\lambda$ at $x = -s$, and write $R_\pm$ for their distances to a field point. The potential with distant zero reference is $\lambda \ln(R_-/R_+)/(2\pi\epsilon)$. Every contour with a constant distance ratio is a circle or a straight line.

For a positive conductor the ratio is

$$R_-/R_+ = e^\eta > 1.$$

Square the distances and complete the square in x, y. The resulting circle has centre $x_c = s\coth\eta$ and radius $s/\sinh\eta$. Its negative twin has centre $-x_c$. Requiring the physical radius and separation fixes $\cosh\eta = D/(2a)$ and $s = \sqrt{(D/2)^2 - a^2}$. Both effective lines lie inside their respective conductors. Since the cylinder voltages are $\pm\lambda\eta/(2\pi\epsilon)$, their voltage difference is $\lambda\eta/(\pi\epsilon)$. The pair capacitance per axial length is therefore

$$C' = \frac{\pi\epsilon}{\operatorname{arcosh}(D/(2a))}.$$

Here charge per length on one conductor is divided by the voltage difference between both. Dividing by the voltage of just the positive conductor relative to the symmetry plane would give twice this value, appropriate to a different terminal convention. Integrating field energy over the whole exterior also yields $\lambda^2/(2C')$, so the factor of two can be checked by comparing charge and energy definitions using the same two terminals.

For $a = 1.0$ mm and $D = 5.0$ mm in vacuum, $\eta = \operatorname{arcosh} 2.5 = 1.567$ and $C' = 17.8$ pF/m. In the far-separated limit $\operatorname{arcosh}(D/2a) \simeq \ln(D/a)$, explaining the familiar logarithmic dependence. In the opposite limit $D = 2a + g$ with $g \ll a$,

$$\operatorname{arcosh}(1 + g/2a) \simeq \sqrt{g/a}.$$

The capacitance then grows as $\pi\epsilon\sqrt{a/g}$, slower than the $1/g$ divergence of broad parallel plates. Why? Near the closest points the local gap increases approximately as $g + y^2/a$. Treating narrow strips as locally parallel gives

$$dC'/dy \simeq \epsilon/(g + y^2/a).$$

Integrating over all y produces $\pi\epsilon\sqrt{a/g}$, independently recovering the leading result. The active area shrinks laterally as the wires approach.

One can predict the mechanical consequence without integrating a surface force directly. For fixed $\pm\lambda$, energy per length is

$$U' = \lambda^2\eta/(2\pi\epsilon).$$

The force tending to increase separation is

$$-dU'/dD = -\lambda^2/[2\pi\epsilon\sqrt{D^2 - 4a^2}].$$

It is attractive. With fixed voltage the charges vary, but evaluating the appropriate coenergy derivative at the same $\lambda = C'\Delta V$ gives the same instantaneous force. This is a strong consistency check because a seemingly small geometric error in the image positions can preserve a reasonable far-field capacitance while badly mispredicting the near-gap force.

Try it · Resolve the two-wire capacitance

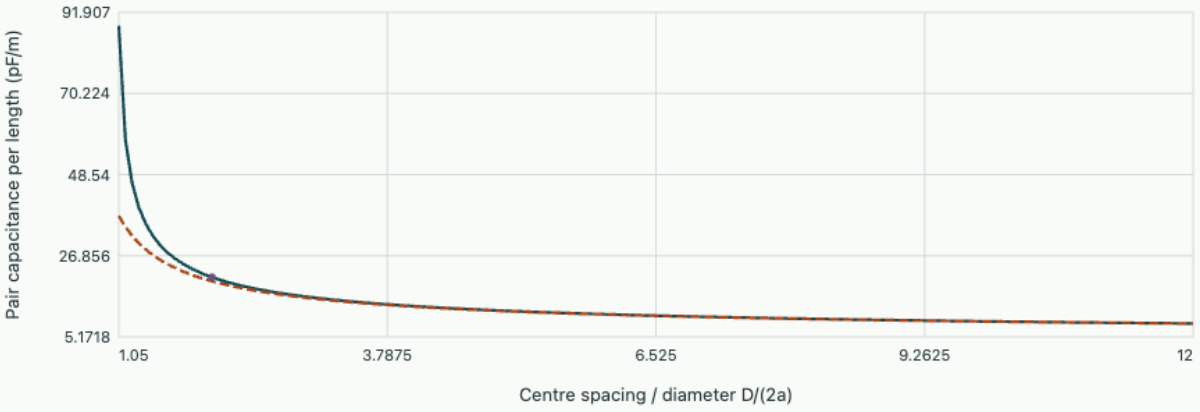
Reduce the gap and compare the exact result with its large-separation logarithm. Change permittivity and predict the scaling before moving the slider. C′ uses the voltage between both wires.

Open this boundary experiment

Try the calculation

Relative permittivity 1

Selected D/(2a) 2



Exact bipolar-coordinate result | Widely separated approximation | Selected spacing

At D/(2a)=2, C′=21.122 pF/m. For a homogeneous nonmagnetic dielectric the corresponding TEM impedance is 157.93 Ω.

Notebook: reproduce the exact boundary constructions and the 3D collocation model

An infinite image series for two grounded planes

Put a charge q at $z = s$ between infinite grounded planes at $z = 0, L$. Reflecting in the bottom plane creates a charge $-q$ at $-s$; now the top plane no longer has zero voltage. Repeated reflections generate charges $+q$ at $z = 2nL + s$ and $-q$ at $z = 2nL - s$, for every integer n. At the bottom, equal-distance terms cancel pairwise by reflection across zero. At the top, reflection $z \mapsto 2L - z$ again exchanges the

two families. The original charge is the $n = 0$ positive one. Each other singularity lies outside the open strip $0 < z < L$.

For axial separation from the charge line $r_\perp$, the potential is the sum of $q/(4\pi\epsilon)$ times the differences of reciprocal distances to these pairs. Group each positive and negative image into a neutral pair before summing. Individual series of absolute monopole magnitudes diverge like a harmonic series, while the paired differences decay at least as fast as $1/n^2$ at fixed interior observation point. A numerical calculation should specify this paired convention and test voltage on both planes as more pairs are included. Summing a large collection of positives first and negatives afterward can produce large finite truncation offsets.

At $s = L/2$ the charge is equidistant from both grounded planes. Reflection about the midplane makes the field due to all images vanish at the source, so it feels no net axial force. Each conductor nevertheless receives induced charge $-q/2$. Reciprocity gives the split without summing any images: the weighting potential of the upper plane is z/L, so it receives $-qs/L$, with the rest on the lower plane. As s approaches zero the nearest negative image at $-s$ dominates the force, reproducing the single-plane attraction $q^2/(16\pi\epsilon s^2)$. A zero net central force therefore does not imply weak local conductor fields or no induced charge.

A useful numerical experiment places the source at several small offsets from the midpoint. The two conductors' integrated charge measurements change linearly with s by the exact reciprocity rule, while the force changes nonlinearly and diverges near either face. These independent analytic targets expose different errors. A coarse mesh may resolve smooth electrode weighting potentials and hence the charge split long before it resolves the closely spaced source-image pair needed for a trustworthy force. Replacing all images by the nearest one succeeds only near a chosen plate and fails both far-boundary conditions and midpoint force cancellation.

Try it · Test both grounded faces at once

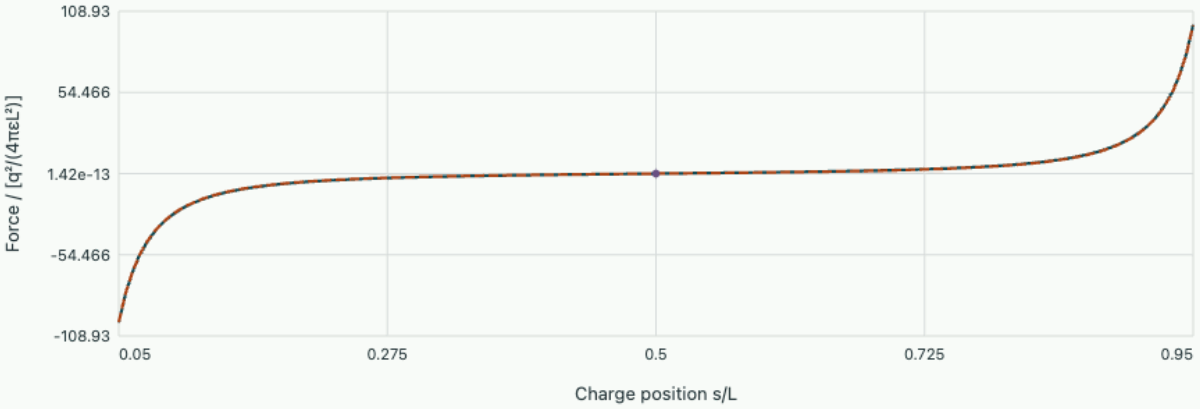
Move the source around the midpoint and refine the image sum. Check antisymmetry of force and compare the nearest-image approximation. The induced electrode charges are also reported from reciprocity.

Open this boundary experiment

Try the calculation

Included image-pair order **12**

Selected charge position s/L **0.5**



At s/L=0.5, normalised force=6.0715e-18. Induced charge fractions: lower plate -0.5, upper plate -0.5. At the midpoint force cancels, while both plates carry –q/2.

Notebook: reproduce the exact boundary constructions and the 3D collocation model

A buried grounded plane produces a scale-dependent image response

A charge is placed in medium 1 above a film of medium 2 with thickness t. The film spans $-t < z < 0$ and is backed by a grounded metal at $z = -t$. Both permittivities are positive and there is no free charge on the material interface. An intuitively appealing model uses one image representing the dielectric interface and another representing the metal. Their two boundary effects actually interact: the dielectric layer changes how strongly the field of the backing is seen at the upper surface. A sinusoidal lateral Fourier component exposes that interaction without constructing an infinite image series.

Let $k > 0$ be the magnitude of the wavevector parallel to the plane. In the portion of medium 1 below the source write the incident contribution near the interface as Ae^{kz} and the reflected contribution as Be^{-kz},

so their values there are A, B. In the film the grounded backing selects $C\sinh[k(z+t)]$. Potential continuity gives

$$A + B = C\sinh(kt).$$

Equality of oriented displacement at $z = 0$ gives

$$\epsilon_1 k(A - B) = \epsilon_2 kC\cosh(kt).$$

Dividing these two equations and solving for the ratio yields

$$r(k) = \frac{B}{A} = \frac{\epsilon_1 - \epsilon_2\coth(kt)}{\epsilon_1 + \epsilon_2\coth(kt)}.$$

At $kt \gg 1$ lateral variations decay within a small fraction of the film thickness, so the result tends to the familiar dielectric image factor. At $kt \ll 1$, $\coth(kt) \sim 1/(kt)$ and the ratio tends to -1: long-range variations see an effectively grounded boundary. When the two dielectric materials are identical the result is exactly $-e^{-2kt}$, the propagation factor to a grounded plane at depth t and back. These three limits independently test the sign convention for the decaying and growing terms in the local upper-region decomposition.

Take $\epsilon_1 = 4\epsilon_0, \epsilon_2 = 2\epsilon_0$. For a very short wavelength the real image coefficient tends to $+1/3$, while for very long wavelengths the reflected coefficient is negative. It crosses zero when $\coth(kt) = 2$, or $kt = \frac{1}{2}\ln 3 = 0.5493$. With a 1.0 mm film the crossover wavelength is

$$2\pi t/0.5493 = 11.4\,\mathrm{mm}.$$

A single image charge with a constant amplitude cannot represent all wavelengths. A point charge excites a broad spectrum; the force it experiences is an integral over those modes weighted by propagation between charge and interface. The sign change of one Fourier coefficient alone does not locate a force reversal at the same distance.

An experiment with a sinusoidal patterned source and a movable probe can estimate this reflection coefficient mode by mode, comparing to a calibration without the backing. Measuring at several wavelengths tests the film model more stringently than fitting one effective permittivity to one force reading. If the backing is ungrounded but maintained at one unknown uniform potential, nonzero k components can still use the same zero-variation boundary there while the $k = 0$ component must enforce the actual conductor charge. This subtle distinction is another reason to state both electrical reference and spatial mode whenever an image coefficient is used.

Try it · Find which wavelengths see the backing

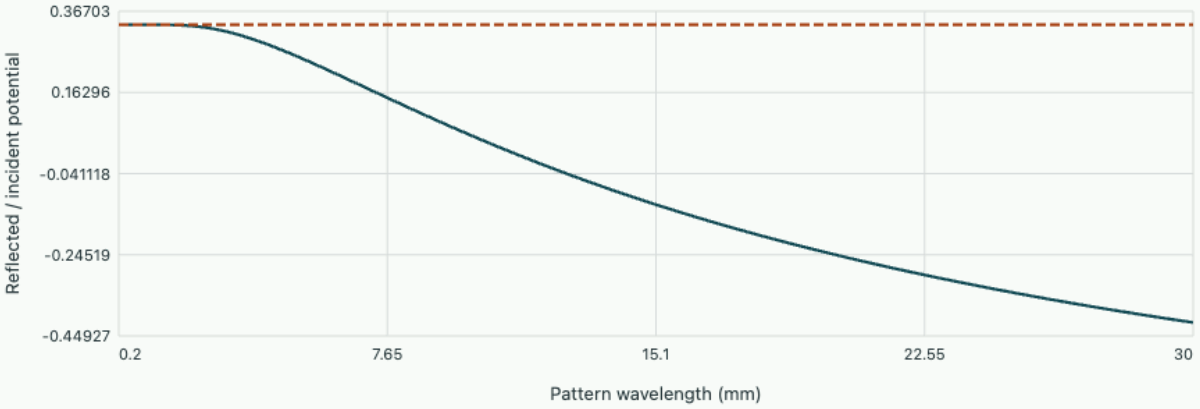
Change film thickness and relative permittivity. Check the identical-material limit and locate the sign crossover when ε_2 is smaller than ε_1. This is a mode reflection coefficient; the force on a point charge requires integrating over its spectrum.

Open this boundary experiment

Try the calculation

Permittivity ratio $\varepsilon_2/\varepsilon_1$ **0.5**

Film thickness (mm) **1**



The reflection changes sign at wavelength 11.438 mm. Long wavelengths see the buried ground; short wavelengths see the dielectric interface.

Notebook: reproduce the exact boundary constructions and the 3D collocation model

Tutorial · apply and challenge an image system

What does an outside line induce on an isolated cylinder?

An initially neutral isolated conducting cylinder replaces the grounded one. Find the change needed in the image system, and calculate the leading far-separation force per length. State the observable that distinguishes the two preparations.

▼ **Work through the solution**

Retain the off-centre negative image to enforce equipotentiality, and add a central $+\lambda$. Its logarithmic potential has constant value on the cylinder. The equivalent enclosed charge per length is now zero. The radial force per length is

$$\lambda^2[1/d - d/(d^2 - a^2)]/(2\pi\epsilon_0) = -\lambda^2 a^2/[2\pi\epsilon_0 d(d^2 - a^2)].$$

At $d \gg a$ this varies as $-d^{-3}$, whereas the grounded-line problem gives a leading $-d^{-1}$ force. A total-charge measurement or a force-vs-distance measurement distinguishes the preparations. The total source including the real infinite line is no longer neutral, so potential differences must be referred to a finite reference or a specified return electrode.

Can a single medium describe the potential on both sides?

For the dielectric interface let $\epsilon_2 = 3\epsilon_1$. A voltmeter scanned along the plane finds a potential that looks like a charge of half the expected strength in medium 1. Predict the force direction, the transmitted field below, and why one apparent charge cannot describe the whole space.

▼ **Work through the solution**

Here $q' = -q/2$, so on the plane $q + q' = q/2$. The force is attractive toward the higher-permittivity medium. The lower-region effective charge is $q'' = 3q/2$, producing potential prefactor $q''/\epsilon_2 = (q/2)/\epsilon_1$, so the surface voltage agrees. Away from the plane the upper field contains two distances while the lower field contains only the continued source distance. Normal electric fields at the surface differ by the factor $\epsilon_2/\epsilon_1 = 3$ required by continuous displacement. Identical voltage along the interface does not imply equal three-dimensional fields near it.

Further reading

- Compare grounded and isolated image spheres numerically
- Integrate the induced charge on the two faces of a corner
- Griffiths, 5th ed., §§3.2, 4.4.2 · images and dielectric boundaries

Electrostatic field theory and measurement · 5. Green functions, the Poisson kernel and what a small electrode can influence

A Green function is the response of an entire boundary-value problem to a unit source. Once that response is known, a continuous charge distribution and a patterned boundary voltage can both be treated by integration. The two integrations require different kernels.

Define a point response with its boundary conditions

Let Ω be a vacuum domain with boundary on which the voltage is given. Use the convention $-\nabla_{\mathbf{y}}^2 G(\mathbf{x},\mathbf{y}) = \delta^3(\mathbf{y}-\mathbf{x})$ and $G = 0$ on the boundary as a function of $\mathbf{y}$. In free space $G = 1/(4\pi|\mathbf{x}-\mathbf{y}|)$ if decay is required at infinity. This definition places no permittivity in G; physical charge enters as ρ/ϵ_0. A different frequent convention puts 4π in the defining equation, changing all subsequent prefactors. Comparing the near-source flux to unity is a straightforward way to fix which convention is being used.

Begin with two differentiable functions u, v. The difference $\nabla\cdot(u\nabla v - v\nabla u) = u\nabla^2 v - v\nabla^2 u$ has no mixed-gradient term. Integrate this identity with $u = V(\mathbf{y})$,

$$v = G(\mathbf{x},\mathbf{y}).$$

The source equation for G samples V at $\mathbf{x}$, while Poisson's equation for V supplies a volume density. With the outward normal taken from Ω,

$$V(\mathbf{x}) = \int_\Omega G(\mathbf{x},\mathbf{y})\,\frac{\rho(\mathbf{y})}{\epsilon_0}\,d^3y - \oint_{\partial\Omega} V(\mathbf{y})\,\partial_{n_{\mathbf{y}}} G(\mathbf{x},\mathbf{y})\,dA_{\mathbf{y}}.$$

The $G\partial_n V$ boundary term vanishes because $G = 0$ there. If the singularity causes concern, exclude a small sphere about $\mathbf{x}$, apply the ordinary identity, and shrink the sphere. The inward normal on the excised sphere produces the same $V(\mathbf{x})$ term. The representation separates charges in the volume from voltages maintained by electrode sources; it does not combine them into one fictitious material distribution.

Solve the half-space and differentiate with respect to source position

Consider $\mathbf{x} = (\mathbf{X}, h)$ with $h > 0$ above the plane $z = 0$. For an interior source $\mathbf{y} = (\mathbf{Y}, z)$, denote its reflection by $\mathbf{y}^* = (\mathbf{Y}, -z)$. The grounded half-space Green function is

$$G = \frac{1}{4\pi}\left[\frac{1}{\sqrt{|\mathbf{X}-\mathbf{Y}|^2 + (h-z)^2}} - \frac{1}{\sqrt{|\mathbf{X}-\mathbf{Y}|^2 + (h+z)^2}}\right].$$

It vanishes as the source coordinate approaches the plane. The outward domain normal there is $-\hat{\mathbf{z}}$; therefore the coefficient of a prescribed boundary voltage is

$$P_h = -\partial_n G = \partial_z G|_{z=0}.$$

Differentiating the two terms gives equal positive contributions:

$$P_h(\mathbf{X}-\mathbf{Y}) = \frac{h}{2\pi[|\mathbf{X}-\mathbf{Y}|^2+h^2]^{3/2}}.$$

Its units are inverse area, while G has inverse length in this three-dimensional convention. For no interior charge $V(\mathbf{X},h) = \int P_h(\mathbf{X}-\mathbf{Y})f(\mathbf{Y})\,d^2Y$. The kernel is positive. Integrate over the whole plane using $2\pi s\,ds$; the result is one. Hence a constant boundary voltage is reproduced, and an arbitrary voltage bounded between two numbers produces values between the same bounds inside. The integral solves the bounded half-space problem with the far-field condition appropriate to the data. Demanding decay for a nonzero constant boundary value would be inconsistent.

A detector at height h takes a weighted average over a neighbourhood of width comparable with h. The kernel has algebraic tails, so its support is not a sharp circular footprint. The contribution from radii smaller than b is $1 - h/\sqrt{b^2+h^2}$. Half the total weight lies within $b = \sqrt{3}h$. Ninety per cent requires

$$b = \sqrt{99}h \simeq 9.95h.$$

These quite different radii show why a full-width-at-half-maximum measurement of the kernel cannot by itself define the region that contains most of its influence.

A finite driven electrode and a reciprocity experiment

Fix a circular disk of radius b in the plane at voltage V_d while grounding the rest of the plane, with ideal thin insulating seams. Directly over its centre,

$$V(0,h) = V_d\left(1 - \frac{h}{\sqrt{h^2+b^2}}\right),$$

$$E_z(0,h) = V_d\frac{b^2}{(h^2+b^2)^{3/2}}.$$

At $h \to 0^+$ the electrode voltage is recovered at its centre. For $h \gg b$ the voltage scales as $V_d b^2/(2h^2)$. A small electrode surrounded by ground does not behave at long distance as an isolated charged electrode with a $1/h$ potential. Its surrounding return charge makes a combined disturbance whose far field has a different order. At the sharp ideal edge the boundary voltage jumps, and local fields there require an actual seam geometry if one wants a bounded peak or a breakdown prediction.

For $b = 5.0$ mm, $V_d = 4.0$ V and $h = 5.0$ mm the axial voltage is 1.17 V and field is 283 V/m. A separate experiment grounds every electrode and places a point charge q at that observation point. Green reciprocity relates the net charge on the disk to the previous weighting potential $w = V/V_d$: $Q_d = -q\,w$. With $q = 20$ pC the disk collects -5.86 pC. Direct integration of the image-plane surface charge $\sigma(s) = -qh/[2\pi(s^2+h^2)^{3/2}]$ gives the same result. It is precisely $-qP_h$. Thus the boundary kernel has both a voltage-producing interpretation and a charge-collection interpretation measurable with different instrumentation.

Recover the kernel of an electrode strip

A boundary pattern invariant in y should not require a two-dimensional source integral. Integrate P_h over the unused coordinate. With lateral separation $s = X - X'$, the result is

$$p_h(s) = h/[\pi(s^2 + h^2)].$$

It integrates to one over the real line. A strip between $-b, b$ at V_d gives on its central normal $V(0, h) = 2V_d \arctan(b/h)/\pi$. At $h = b$ the value is half the applied voltage, distinct from the circular electrode's $1 - 1/\sqrt{2}$. The strip extends in one more direction and occupies a larger angular range of boundary influence.

For sinusoidal boundary data $f(X) = \cos kX$ there is an independent way to determine the result without evaluating a difficult oscillatory integral. Translational symmetry and linearity preserve the sinusoid. Solving Laplace's equation with vertical factor $Z(h)$ gives

$$Z'' = k^2 Z.$$

The bounded decaying solution with unit boundary amplitude is $e^{-|k|h}$. Thus the Fourier transform of the strip kernel is $e^{-|k|h}$. Convolving twice with heights h_1, h_2 is equivalent to convolution once with height $h_1 + h_2$. This composition property checks a numerical propagation code without requiring separate analytic solutions for many different boundary patterns.

Which symmetries does an influence table possess?

Exchange source and detector in the defining equations for $G(\mathbf{x}, \cdot)$ and $G(\mathbf{z}, \cdot)$. Their difference identity gives $G(\mathbf{x}, \mathbf{z}) - G(\mathbf{z}, \mathbf{x}) = 0$, since both vanish on the same boundary. Symmetry extends to scalar or symmetric-tensor positive permittivity when the operator is defined as $-\nabla \cdot \epsilon \nabla$. It refers to exactly the same material, boundary preparation and domain in both experiments. Replacing one experiment's grounded conductor by a floating conductor changes the Green function.

On a nonuniform mesh it matters whether influence entries correspond to equal point charges or equal charge densities in unequal cells. If a finite-volume solver stores sources as densities, a column's total injected charge scales with its cell volume. An unweighted table of sampled voltages can look nonsymmetric even when the discrete physics is reciprocal. Form the matrix with unit total source and consistent detector averaging, or retain the correct volume-weighted inner product. Test positivity for positive sources as a second check. Symmetric influence entries by themselves cannot expose a common incorrect normalisation or a misplaced outer boundary.

A kernel that reproduces every harmonic around a circle

In a two-dimensional circular domain of radius a, prescribe a voltage $f(\phi')$ around the wall. For a detector at (r, ϕ) let $t = r/a < 1$. The angular Fourier component with order m is multiplied by $t^{|m|}$ on continuation into the disk. This follows directly from the regular radial power solution of Laplace's

equation. Substitute the Fourier coefficient integral into the series and interchange the integral and the absolutely convergent series at fixed $t < 1$. The boundary response per unit angle is then

$$P_t(\delta) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} t^{|m|} e^{im\delta} = \frac{1 - t^2}{2\pi(1 - 2t\cos\delta + t^2)},$$

$$\delta = \phi - \phi'.$$

To obtain the rational form, separate the $m = 0$ term and sum the two conjugate geometric series $te^{\pm i\delta}/(1 - te^{\pm i\delta})$. The result is positive for every interior point. Integrating it over angle leaves only the constant Fourier mode, so it is normalised to unity. At the centre the kernel is simply $1/(2\pi)$, recovering the exact mean-value theorem for harmonic functions. Near the wall it becomes very sharply concentrated around the detector's azimuth while retaining a small nonlocal tail. Using prescribed boundary data as a flat average at every interior point would reproduce only the centre reading.

For a concrete experiment put a short arc of half-angle α centred on $\phi = 0$ at 1 V and ground the remaining circular boundary. Along $\phi = 0$ the potential equals the integral of P_t between $-\alpha, \alpha$. With $u = \tan(\phi'/2)$ one has $\cos\phi' = (1 - u^2)/(1 + u^2)$, $d\phi' = 2du/(1 + u^2)$. Substitution gives the weight

$$w(t) = \frac{2}{\pi} \arctan\left[\frac{1 + t}{1 - t} \tan(\alpha/2)\right].$$

At $t = 0$ it gives α/π, exactly the fraction of the circumference occupied by the arc. For a 60°-wide arc $\alpha = \pi/6$, hence the centre reads $1/6$ V. At $t = 0.6$ on the centreline it reads about 0.522 V. Placing a small positive line charge there while grounding the whole cylindrical wall induces -0.522 times the line charge on that arc by the same reciprocity argument used on the flat disk. The angular share depends on position inside the circle even though every point on the conductor has the same actual grounded voltage.

An independent check takes the limit of a very large circle with the probe only a distance $h = a - r$ from the nearest boundary point. For small angular displacement $\delta = s/a$, $1 - 2t\cos\delta + t^2 \simeq (h/a)^2 + (s/a)^2$ and $1 - t^2 \simeq 2h/a$. The angular probability element $P_t(\delta)d\delta$ therefore tends to $h\,ds/[\pi(h^2 + s^2)]$, precisely the translation-invariant strip kernel obtained by integrating the half-space formula over one coordinate. Curvature corrections can be estimated by increasing a/h at fixed physical arc width and comparing both predictions.

Why an insulated Green function needs a zero-mode subtraction

Consider the simpler interval $0 < x < L$ with prescribed zero flux at both ends. We would like a point response satisfying $-G'' = \delta(x - s)$, $G'(0) = G'(L) = 0$. Integrate this proposed equation over the interval: the left-hand side is $-G'(L) + G'(0) = 0$, whereas the right-hand side is 1. The attempted kernel violates compatibility. Physically a sealed insulating region cannot take a net free test charge without changing its prescribed outer flux. Adding an arbitrary constant to G cannot fix this contradiction.

A useful alternative removes the constant source mode and fixes the remaining voltage reference. Define $-G_N'' = \delta(x - s) - 1/L$, $G_N'(0) = G_N'(L) = 0$, and $\int_0^L G_N(x, s)\,dx = 0$. Away from the point the curvature is $+1/L$. The left derivative must be x/L to vanish at zero, the right derivative $x/L - 1$ to

vanish at L. They jump by -1 at $x = s$. Integrating, imposing continuity and subtracting the mean gives the compact symmetric expression

$$G_N(x, s) = \frac{x^2 + s^2}{2L} - \max(x, s) + \frac{L}{3}.$$

Differentiating verifies both end conditions and the jump. Integrating the quadratic part and the piecewise maximum over x verifies the mean condition. Its sign can be negative because the response includes an opposite uniform return source and uses a zero-mean reference. Unlike the positive grounded Dirichlet kernel, it is not a positive voltage weight to use for a maximum-principle test.

Apply a neutral pair of sheets $+\sigma$ at $s_1 = 0.2L$ and $-\sigma$ at $s_2 = 0.7L$. Their voltage is $\sigma[G_N(x, s_1) - G_N(x, s_2)]/\epsilon$. The two artificial uniform return sources cancel exactly. To the left of both sheets the field vanishes, between them it equals $+\sigma/\epsilon$, and to the right it vanishes again. The potential drops by $0.5L\sigma/\epsilon$ across the two sheets. This matches direct Gauss-law jumps, and the zero-mean potential fixes the constant that Gauss's law could not determine. For more complex closed domains numerical solvers implement the same idea by projecting out the constant eigenvector and checking the compatible total charge before inverting. An arbitrary tiny diagonal regularisation of a singular matrix can hide a physically impossible net-source input.

Construct a point response when the medium itself screens charge

A linear ionic or phenomenological screening model between two grounded parallel electrodes leads in one dimension to $[-\partial_x^2 + \kappa^2]V = \rho_{\text{imposed}}/\epsilon$, $0 < x < L$. Here ρ_{imposed} excludes the mobile linear response already represented by $\kappa^2 V$. We seek $G_\kappa(x, s)$ with zero voltage at both ends. Away from its unit source the equation has hyperbolic solutions. Use $A\sinh(\kappa x)$ on the left to vanish at zero and $B\sinh[\kappa(L - x)]$ on the right to vanish at L. Continuity at $x = s$ removes one constant. Integrating across the source gives $G'_\kappa(s^+) - G'_\kappa(s^-) = -1$; the finite mass term integrates to zero in a shrinking source interval.

To solve transparently use the functions $u(x) = \sinh(\kappa x)$,

$$v(x) = \sinh[\kappa(L - x)].$$

Their Wronskian $u'v - uv'$ is $\kappa\sinh(\kappa L)$, independent of x. Then continuity, the derivative jump, and symmetry under $s \leftrightarrow x$ are manifest in

$$G_\kappa(x, s) = \frac{\sinh(\kappa x_<)\sinh[\kappa(L - x_>)]}{\kappa\sinh(\kappa L)},$$

$$x_< = \min(x, s),$$

$$x_> = \max(x, s).$$

For $\kappa \to 0$, expand each hyperbolic sine to first order. The result is $x_<(L - x_>)/L$, the grounded unscreened response. When both observation and source lie many screening lengths from every boundary, the kernel tends to $e^{-\kappa|x-s|}/(2\kappa)$, the infinite-line screened solution. Thus one formula contains both geometric screening by grounded conductors and material screening by mobile charges.

At the midpoint source and detector coincide. The finite kernel reduces to $\tanh(\kappa L/2)/(2\kappa)$. Compared with the unscreened midpoint value $L/4$, its ratio is $2\tanh(\kappa L/2)/(\kappa L)$. At $\kappa L = 2$ a fixed inserted sheet therefore creates only 76.2% of its unscreened midpoint voltage. A naïve replacement of plate separation by the screening length would give a constant ratio and miss the smooth small-κ crossover. For long gaps the measured voltage saturates at fixed sheet charge, because the countercharge response near the sheet now sets the local scale.

Check the convolution using a uniform positive imposed charge density ρ_0. Without integrating the kernel directly, seek a constant particular solution plus a symmetric homogeneous correction. The result is

$$V(x) = [\rho_0/(\epsilon\kappa^2)]\{1 - \cosh[\kappa(x - L/2)]/\cosh(\kappa L/2)\}.$$

Twice differentiation reconstructs the source and zero endpoints. In the small-κL limit it tends to $\rho_0 x(L - x)/(2\epsilon)$. A numerical integration of the point response against a uniform source should match this independent closed solution throughout the interval, including the near-electrode regions. It is a stronger test than checking the point kernel at the centre alone.

Integrate the original equation across the gap. The sum of outward displacement fluxes equals imposed charge minus $\epsilon\kappa^2 \int V\,dx$, the magnitude of the compensating mobile charge. For the uniform source the fraction of imposed charge balanced by electrode rather than mobile charge is $2\tanh(\kappa L/2)/(\kappa L)$. For a long gap this fraction falls like twice the screening length divided by gap width, because only the two boundary regions communicate appreciably with the electrodes. A current meter attached to the electrodes and a chemical measurement of accumulated mobile ions can therefore close the same source budget independently.

Inspect the response to a localised source

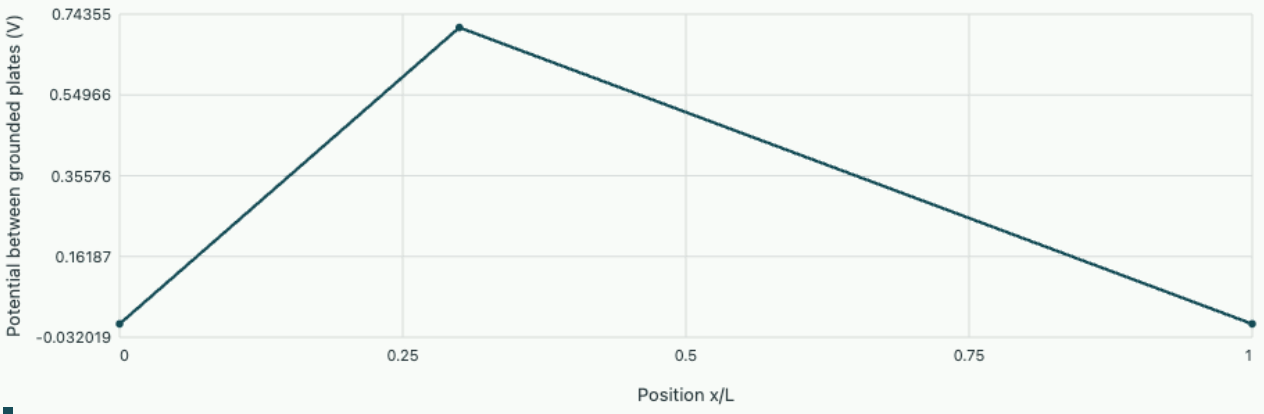
Move the source while holding the boundaries fixed. Test the symmetry between source and observation point predicted by reciprocity.

Open this experiment online

Try the calculation

Sheet position s/L **0.3**

Sheet charge density (nC/m²) **3**



Grounded 10 mm gap

Induced densities are -2.1 nC/m² on the left and -0.9 nC/m² on the right. Potential at the sheet is 0.712 V.

Design a differential collecting electrode from its weighting functions

Two concentric electrodes supply one signed position signal

A grounded plane is divided into a central disk of radius a, a surrounding annulus from a to $2a$, and the remainder. A small charge moves along the central normal. Two charge amplifiers measure the disk and annulus. Construct an output that vanishes at $h = a$, then find its local height sensitivity. Why can that output alone not distinguish uncertain charge from height away from the null?

▼ **Derive and check the result**

Let $w_a = 1 - h/\sqrt{h^2 + a^2}$ and $w_{2a} = 1 - h/\sqrt{h^2 + 4a^2}$. The disk reads $Q_d = -qw_a$; the annulus reads $Q_r = -q(w_{2a} - w_a)$. An output $S = Q_d - \gamma Q_r$ vanishes at $h = a$ if $\gamma = (1 - 1/\sqrt{2})/(1/\sqrt{2} - 1/\sqrt{5}) = 1.127$ approximately. Both readings are negative for positive q, so it is their difference, not their sum, that rejects the reference position. The ideal zero does not depend on the magnitude of any nonzero q.

Differentiate with respect to height using $w_b' = -b^2/(h^2 + b^2)^{3/2}$. At the null $a\,w_a' = -1/(2\sqrt{2})$ and $a\,w_{2a}' = -4/(5\sqrt{5})$. Hence

$$a\,d(S/q)/dh = -(1 + \gamma)aw_a' + \gamma a w_{2a}' \simeq 0.349.$$

For $q = 10$ pC, $a = 2.0$ mm, a small upward displacement of 0.10 mm produces approximately $+0.175$ pC. Finite differences of the full two-disk formula at heights on both sides of a check the linear slope. For displacements that are no longer small compared with a, use the full expression rather than extrapolating that single sensitivity.

An additional sum channel $Q_d + Q_r = -qw_{2a}$ permits a dimensionless output $S/(Q_d + Q_r)$ that cancels the magnitude and sign of q while retaining height dependence. This ratio becomes noise sensitive far from the electrodes, where collected charge is small. A null by itself can also mean that there is no charge at all, so a usable position detector checks the sum channel at the same time. The exercise combines electrostatic reciprocity, electronic weighting and a finite signal-to-noise constraint while requiring only two precomputed boundary responses.

Tutorial · predict before numerically integrating the kernel

How much charge belongs to an annular collector?

An annular electrode with inner radius h and outer radius $3h$ lies in a grounded plane. A charge q sits at height h over its centre. Derive its induced charge. If a numerical integration omits all radii beyond $3h$, where does the missing charge go?

▼ Work through the solution

Subtract two circular weights. The annulus receives $Q/q = -[1/\sqrt{2} - 1/\sqrt{10}] = -0.3909$. The central disk receives $-(1 - 1/\sqrt{2})q = -0.2929q$. The whole plane should sum to $-q$, so the part beyond the outer edge supplies

$$-q/\sqrt{10} = -0.3162q.$$

Nothing has been lost by the physical system; a finite integration patch excludes a considerable piece of the long-tailed kernel. Radial extrapolation with the analytic far tail or an expanding integration domain distinguishes missing area from poor quadrature within the patch.

Use resistance to a modelling mistake

A student proposes G itself as the weight for setting electrode voltages on the plane. Another replaces it by $2P_h$, arguing that an image doubles the response. Predict two different tests that diagnose these errors.

▼ Work through the solution

The first proposal gives zero contribution since the Dirichlet Green function is zero at the boundary. It confuses volume source response with a boundary normal derivative. The second produces twice the prescribed constant voltage because the kernel would integrate to two. Independent electrostatic reciprocity would assign $-2q$ to the whole grounded plane around a charge q, violating the charge sum

derived from the real and image fields. For a numerical test, set successively a constant, a sinusoid, and a bounded disk on the same plane. Those cases examine normalisation, spectral attenuation, and spatial tails separately.

Further reading

- Move a source between two grounded electrodes
- Compare vertical attenuation for two boundary harmonics
- Fließbach, chapter 7 · electrostatic boundary problems

Electrostatic field theory and measurement · 6. Rectangular geometries, orthogonality and sources that drive individual modes

Fourier modes are solutions selected by their electrode walls. On a rectangle they turn a boundary-voltage problem into independent vertical equations. When a charge fills the rectangle the same eigenfunctions give a different denominator and a different physical response.

Determine which sign of a separation constant can fit the walls

Take a vacuum rectangle $0 < x < a,\ 0 < y < b$, invariant along z. Ground the left, right and bottom edges and impose $V(x, b) = f(x)$ at the top. If the boundary data disagree at a corner, the continuum problem can have large edge fields; we first choose smooth f vanishing at both endpoints. In the charge-free interior, a product solution $X(x)Y(y)$ gives $X''/X = -Y''/Y$, equal to a constant wherever the product is nonzero. If $X'' = k^2 X$, the only combination of hyperbolic functions zero at both $x = 0, a$ is zero. The same is true for the linear case $k = 0$. Thus $X'' = -k^2 X$. The left boundary picks a sine and the right boundary enforces $k = n\pi/a,\ n = 1, 2, \ldots$.

Then $Y'' = k^2 Y$, with the bottom condition choosing $\sinh ky$. Choosing its top value as unity gives the transfer function $Y_n = \sinh(k_n y)/\sinh(k_n b)$. Summing products,

$$V(x, y) = \sum_{n \geq 1} f_n \sin(k_n x) \frac{\sinh(k_n y)}{\sinh(k_n b)},$$

$$f_n = \frac{2}{a} \int_0^a f(x) \sin(k_n x)\, dx.$$

These coefficients follow by multiplying the top boundary by $\sin(k_m x)$ and integrating. For $n \neq m$ the product reduces to cosines with integer half-wave multiples and integral zero; for equal indices its integral is $a/2$. This orthogonality incorporates the edge boundary values. Periodically repeating a voltage function without regard to its endpoint conditions can give cosine or constant terms that violate the grounded side walls.

Track how geometry filters electrode detail

A mode's vertical derivative is $k_n \cosh(k_n y)/\sinh(k_n b)$, so the field normal to the electrode weights high modes differently from potential. At the top the magnitude multiplier is $k_n \coth(k_n b)$. For a half-space, high-frequency field content can be large very near a patterned electrode even if it disappears rapidly a short distance inside. For large $k_n b$, direct numerical evaluation of hyperbolic sines may overflow. The algebraically equivalent ratio $e^{-k_n(b-y)}[1 - e^{-2k_n y}]/[1 - e^{-2k_n b}]$ stays bounded; for very small arguments evaluate $1 - e^{-t}$ with a numerically stable exponential-difference routine.

Suppose the rectangle is square and the electrode has $f = 4\sin(\pi x/a) - 2\sin(3\pi x/a)$ V. At mid-height $y = a/2$, the fundamental transfer factor is $1/[2\cosh(\pi/2)] = 0.1993$, while the third is

$$1/[2\cosh(3\pi/2)] = 0.00898.$$

At $x = a/2$ the first sine is 1 and the third is -1, so the potential is 0.815 V approximately, compared with the 6 V drive at the same x on the electrode. Taking the simple arithmetic average of the top and bottom voltages at that x would predict 3 V and miss leakage to the grounded side walls.

If a second boundary carries a pattern, superpose the source-free solution produced by it with the other boundaries grounded. For example a nonzero bottom voltage pattern uses $\sinh[k_n(b-y)]/\sinh(k_n b)$. This works because the field equation and boundary conditions in fixed-voltage form are linear. A floating conductor imposes a different constraint coupling many modal coefficients, not the prescription of one extra arbitrary voltage profile.

Integrate a discontinuous electrode and distinguish two kinds of convergence

Let a short strip centred on the top boundary span $a/2 - w/2 < x < a/2 + w/2$ with voltage V_d, the remaining top grounded. Integrating the sine over that interval gives

$$f_n = 4V_d \sin(n\pi/2)\sin(n\pi w/(2a))/(n\pi).$$

Even modes vanish by mirror symmetry. Individual allowed coefficients decrease as $1/n$ at most because the boundary has jumps. On the boundary partial sums overshoot close to a jump. Adding terms narrows that region rather than removing its limiting fractional overshoot. At fixed interior distance the exponential transfer renders the series much better behaved.

For $w = a/3$, the first and third coefficients are $2V_d/\pi$ and $-4V_d/(3\pi)$. Along the centreline both contribute positively because their sine factors have opposite signs. A finite-width electrode cannot be captured exactly with a single spatial mode even when its interior signal far away is mainly fundamental. Recovering the electrode width from a distant scan by dividing attenuated high modes by their transfer factors will amplify measurement errors. Forward propagation and inverse reconstruction need different treatments of frequency content.

Let the volume rather than the electrode excite the modes

Ground all four edges and embed a known charge density $\rho(x, y)$. Choose double sines $\sin(m\pi x/a)\sin(n\pi y/b)$. They already vanish on all edges and are eigenfunctions of the negative Laplacian with eigenvalues $\lambda_{mn} = (m\pi/a)^2 + (n\pi/b)^2$. Project the charge using $\rho_{mn} = 4(ab)^{-1}\int_0^a \int_0^b \rho \sin(m\pi x/a)\sin(n\pi y/b)\,dy\,dx$. Applying Poisson's equation coefficient by coefficient yields

$$V_{mn} = \frac{\rho_{mn}}{\epsilon_0 \lambda_{mn}}.$$

High spatial frequencies of a volume charge produce weaker potentials by an inverse-square spectral factor. The corresponding normal field on a boundary involves an additional derivative. This polynomial filtering differs from the exponential decay of a boundary pattern measured away from its source. Both follow from the same PDE but answer different experiments.

For a manufactured test choose a square of side 10 mm with $\rho = \rho_0 \sin(\pi x/a)\sin(2\pi y/a)$. The voltage magnitude peaks at $\rho_0 a^2/(5\pi^2\epsilon_0)$. With $\rho_0 = 1.0\ \mu\mathrm{C/m^3}$ it is 0.229 V. The total charge integrates to zero, since the second sine is odd about the mid-height. The field and stored energy are nonzero. At $x = a/2, y = a/4$ the voltage is a positive maximum, consistent with a positive local source. At $x = a/2, y = 3a/4$ the same amplitude is negative.

Use an exact discrete eigenvalue for a rigorous code check

On a uniform square grid of spacing $h = a/N$, replace $-\partial_x^2$ by its centred three-point formula. Acting on the sampled sine gives eigenvalue $4h^{-2}\sin^2(m\pi h/(2a))$. The two-dimensional sum of the corresponding factors is therefore the exact discrete eigenvalue for this interior mode with zero edges. A computed solution can be compared first to $\rho_{mn}/(\epsilon_0\lambda_{mn,h})$, which tests the solver algebra, and then to the continuous $\rho_{mn}/(\epsilon_0\lambda_{mn})$, which tests spatial discretisation. Reducing the iterative residual alone cannot remove the difference between these two physical amplitudes.

At $N = 10$, modes $(1,2)$ have $a^2\lambda_h = 4N^2[\sin^2(\pi/20) + \sin^2(\pi/10)] = 47.99$, versus $5\pi^2 = 49.35$. The discrete amplitude is about 2.8% high despite solving its linear equations exactly. For fixed mode under grid halving, expansion of the sine gives an eigenvalue error proportional to h^2. At wavelengths only a few cells wide this expansion is unreliable. Try a sequence of grids with the same physical mode and compare actual amplitude ratios. Holding instead a fixed number of cells per wavelength while changing the physical wavelength tests a different convergence question.

Two transverse dimensions select a different attenuation length

Extend the rectangle into a rectangular conducting shaft $0 < x < a,\ 0 < z < c$, running to $y > 0$. Ground all four side walls, require bounded potential far down the shaft and prescribe a voltage on its mouth at $y = 0$. Separation in the cross section now gives products $\sin(m\pi x/a)\sin(n\pi z/c)$ with both $m, n \geq 1$. The axial equation is

$$Y'' = [(m\pi/a)^2 + (n\pi/c)^2]Y.$$

Each boundary coefficient therefore propagates as $e^{-\gamma_{mn} y}$,

$$\gamma_{mn} = \pi\sqrt{m^2/a^2 + n^2/c^2}.$$

A narrow second dimension suppresses penetration even when the first dimension is very broad. Approximating the shaft by a two-dimensional strip removes an entire contribution to the decay constant.

If the aperture is driven uniformly at V_0, projecting gives $f_{mn} = 16V_0/(\pi^2 mn)$ for odd m, n, zero otherwise. The ideal mouth voltage jumps at the grounded rim. Deep inside, only the lowest odd pair survives and its amplitude exceeds V_0 before the exponential is applied; the higher terms restore the actual boundary value at $y = 0$. With $a = 20$ mm and $c = 10$ mm the slowest decay length is

$$1/\gamma_{11} = 2.85\,\mathrm{mm}.$$

A two-dimensional strip ignoring the 10 mm side would instead give

$$a/\pi = 6.37\,\mathrm{mm}.$$

After 10 mm penetration the first-mode attenuation factors are 0.0298 and 0.208, almost a factor of seven apart.

A simple numerical test samples the centreline voltage at several heights and calculates adjacent log slopes. Close to the mouth many odd modes matter, while deep inside the ratio between two positions Δy apart approaches $e^{-\gamma_{11}\Delta y}$. Compare with a second geometry that swaps a, c. The two lowest-order decay lengths must agree because the eigenvalue is symmetric in the two dimensions; an implementation that uses only the nominal 'width' may fail this check. A second observation path away from the centre can sense different mode weights, even though every mode's exponential is unchanged.

Suppose the mouth instead has opposite voltages on the left and right halves, antisymmetric about $x = a/2$. Its coefficients vanish for odd m, so the $(1,1)$ mode is absent by symmetry. The leading allowed mode can be $(2,1)$, shortening the decay length. A centreline probe would then read zero everywhere by antisymmetry and could not fit any length at all. Move the probe to $x = a/4$ while keeping it at mid-height of the cross section. This illustrates how an excited spectrum, a geometry's complete set of eigenvalues, and an instrument's observed spectrum may differ.

Why the energy of an ideal electrode voltage step depends on edge resolution

Return to a two-dimensional strip or half-plane. To focus on a single boundary Fourier mode take $V = A\sin(kx)e^{-ky}$ in $0 < x < a,\ y > 0$, with $ka = n\pi$. Its gradients in x and y have equal prefactor kAe^{-ky}. The sum of their squared sine and cosine factors is one, so the energy per perpendicular length is

$$U_n' = \epsilon a k A^2/4.$$

For many orthogonal sine modes cross terms vanish after integration. Thus the stored field energy involves the weighted sum $\sum_n n|f_n|^2$, while the mean-square boundary voltage itself involves simply $\sum_n |f_n|^2$. The extra factor of n comes from the steep near-electrode field of short-wavelength structure.

For a finite electrode patch abruptly stepping from 0 to V_d, the allowed Fourier amplitudes typically decay as $1/n$ along nonzero subsequences. The voltage series converges in mean square, but its energy sum can contain a logarithmically diverging $\sum 1/n$. This is the local singularity of the zero-width voltage discontinuity. It does not prevent a well-defined voltage at any interior distance, but it prevents assigning a finite total energy or a cut-off-independent peak electric field to that exact sharp-edge idealisation. Mesh refinement can produce progressively higher energy estimates even as interior voltage curves appear perfectly stable.

There is a local way to see the same conclusion without a Fourier series. Consider a small half-disk of vacuum around a junction where an otherwise flat boundary jumps between voltages zero and V_d. In polar coordinates about the junction, the leading harmonic solution depends only on angle: $V = V_d\phi/\pi$ with

angle origin chosen on the grounded side. Its field magnitude is $V_d/(\pi r)$. Integrating $\epsilon E^2/2$ over a half-annulus of inner radius δ and outer radius R gives

$$U' = \epsilon V_d^2 \ln(R/\delta)/(2\pi).$$

Halving the microscopic length changes this single-edge contribution by $\epsilon V_d^2 \ln 2/(2\pi)$. The independence of R in that change is a practical test of the logarithmic asymptotic regime.

A finite insulation seam between real electrodes supplies a different boundary problem. For a deliberately smoothed transition of width δ, the spectrum falls faster above $k \sim 1/\delta$; the local singularity is cut off. One can compare a family of specified physical seam widths while separately refining each numerical grid until it resolves its own seam. Shrinking seam width along with mesh spacing instead changes the physical geometry at each refinement. Report the interior voltage, near-edge energy within a named observation region and the physical seam model separately. Each can converge at a different rate even in a technically correct Laplace solver.

A discrete residual can fall quickly while the wrong spatial scale remains

The separated sine basis makes it possible to analyse the convergence of a concrete numerical Laplace solver as well as the continuum solution. On a square grid with equal spacing, the five-point discretisation in a charge-free cell says its voltage equals the mean of its four neighbours. A Jacobi sweep sets all new values to this mean of the old grid, preserving fixed boundary values meanwhile. Subtract the exact discrete solution from the current approximation: the error evolves by the same averaging step with zero boundary values, independently of the inhomogeneous source that generated the original solution.

Insert an error pattern $e_{ij} = \sin(m\pi i/N)\sin(n\pi j/N)$. Elementary sine addition on the two x-neighbours produces $2\cos(m\pi/N)e_{ij}$; the y-neighbours produce the analogous factor. Hence one sweep multiplies this error mode by $\mu_{mn} = [\cos(m\pi/N) + \cos(n\pi/N)]/2$. For the longest $(1,1)$ mode, $\mu \simeq 1 - \pi^2/(2N^2)$. Its amplitude takes roughly $2N^2/\pi^2$ sweeps to drop by $1/e$. Doubling resolution therefore makes this relaxation about four times slower per physical longest mode even before counting the four times as many cells per sweep.

Choose $N = 64$. Then $\mu_{11} = 0.99880$. After 100 sweeps about 88.6% of that initial smooth error still remains, although the update per cell has become small. At the opposite corner of the spectrum $(N-1, N-1)$ gives nearly -1, so a checkerboard mode also decays slowly while changing sign on successive sweeps. Weighted Jacobi replaces an old value by $(1-\omega)$ times itself plus ω times the neighbour mean. At $\omega = 2/3$ that checkerboard amplification becomes approximately $-1/3$, removing short-wavelength oscillations quickly. The long smooth mode still has amplification near one. An attractive smooth colour map after a few sweeps can thus be misleading.

A multigrid correction exploits these distinct errors. First relax high-frequency errors on the fine grid. Compute the residual of the original discrete equation and restrict it to a coarser grid. A long fine-grid mode becomes easier to change there. Solve the coarse error equation, prolong its correction back to the fine grid, and smooth again. The correction must solve for error with the appropriate homogeneous boundary constraints, not repeat the original boundary excitation a second time. The residual and

prolongation should respect the same physical units and grid volumes as the original source. Otherwise a visually plausible coarse correction can inject net charge at each level.

For a manufactured check initialise the fine-grid error to a known sine mode on top of an exact discrete solution. After one weighted sweep measure the inner product of the new error with the original mode and divide by its squared norm. The observed ratio must equal $1 - \omega + \omega\mu_{mn}$ before roundoff or boundary effects. Repeat with $(1, 1)$, a mixed $(1, N - 1)$ mode, and the checkerboard-like $(N - 1, N - 1)$ mode. This directly tests the implementation's indexing, old-versus-new updates and boundary handling; it does not require the continuum equation to be discretisation exact.

Finally compare iterative and truncation errors in measurable volts. For a discrete mode, if the algebraic residual coefficient is r_{mn} in the form $-\nabla_h^2 e = r$, its error coefficient is $r_{mn}/\lambda_{mn,h}$. The same residual amplitude in a long-wavelength mode produces a larger voltage error than in a high-frequency mode. An absolute solver tolerance with no physical scaling is therefore insufficient to certify a voltage reading. Stop iterating when the solver error in the actual observable is small compared with discretisation and experimental uncertainty, then refine the spatial representation separately.

Resolve a short-wavelength electrode pattern

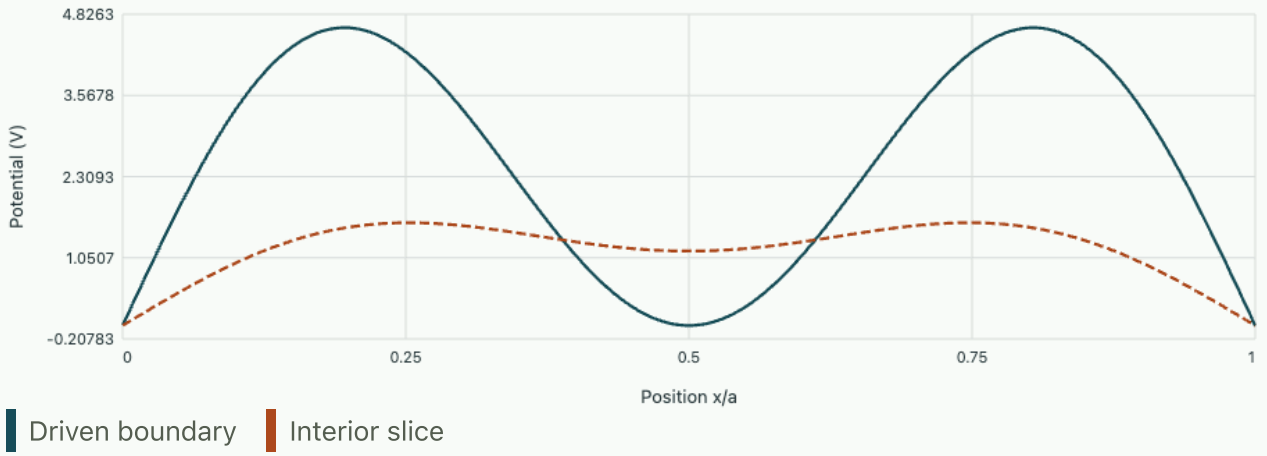
Increase the transverse mode number. Predict how far the pattern reaches into the strip before inspecting the voltage profile.

Open this experiment online

Try the calculation

Height y/a **0.18**

Third harmonic at the boundary (V) **3**



At the middle of the strip the potential is 1.15 V. The third harmonic retains 18.3% of its boundary amplitude.

Build a resistor-network analogue of the rectangular boundary problem

Use Kirchhoff measurements to check a field solver

A square lattice of nodes is connected by identical resistors R between nearest neighbours. Outer nodes are held at prescribed voltages. Derive the interior equation from charge conservation at a node. Which measured electrical quantity can test a normal field at the boundary? What changes on an anisotropic grid?

▼ **Derive and check the result**

For a steady interior node with no supplied current, the sum of currents $(V_{ij} - V_{\text{neighbour}})/R$ is zero. Rearrangement gives the same mean-of-four equation as the square-grid discrete Laplacian. Thus the resistor array reproduces the *discrete* Dirichlet solution for the same pinned boundary voltages. Its Joule dissipation $\sum_{\text{edges}} (V_i - V_j)^2/R$ is stationary against variations of the unfixed voltages, giving the same equations a second way. The numerical voltage maps can be compared node by node; the physical power and the electrostatic field energy are different dimensional quantities even though their quadratic forms are analogous.

To estimate the normal field near a flat boundary, measure the signed current from a boundary node along the edge connecting it to its nearest interior neighbour. This equals the observed voltage difference divided by R; multiplying by R/h converts it to the corresponding one-sided field magnitude in volts per metre. At sharp corners the one-sided difference has large discretisation error and the continuum physical field may be singular. Summing all currents supplied to a terminal edge gives a more robust integrated flux analogue. Account for any links between nodes at different prescribed boundary voltages if those links contribute separate current to the laboratory power supplies.

For horizontal spacing h_x and vertical spacing h_y, replacing every resistor by the same value no longer matches the isotropic Cartesian Laplacian. A finite-volume cell requires conductance proportional to face length divided by edge length: horizontal links have conductance proportional to h_y/h_x, vertical links to h_x/h_y. Their ratio is h_y^2/h_x^2. This restores a discrete conservation law with the correct physical area factors. A measured profile agreeing exactly with an equal-resistor lattice on a stretched geometry could therefore validate the wrong continuum material or geometry unless those weights were checked first.

Tutorial · design a boundary measurement and a solver benchmark

Identify a forbidden mode

A symmetric positive top electrode has its centre at $x = a/2$. A numerical sine transform reports a large $n = 2$ coefficient. Is it evidence of physical high-frequency content? Which controlled change could verify the sign of the defect?

▼ **Work through the solution**

Mirror substitution $x \mapsto a - x$ preserves the boundary voltage but changes $\sin(2\pi x/a)$ to its negative. Its exact projection is zero. Sharp electrode edges may require many odd modes; they cannot explain an even coefficient under exact symmetry. Check coordinate origin, edge alignment on the grid and measurement offsets. Shift the actual patterned patch by a small amount to one side, then the same amount to the opposite side. A genuine even coefficient changes sign under these mirrored displacements, while odd-mode amplitudes agree. The response at interior height has the same parity because the transfer factor depends on mode number alone.

Measure where electrode charge resides

For a single sinusoidal top voltage $V_d \sin(\pi x/a)$, derive the total free charge per perpendicular length on the upper electrode face. Why does it depend on the bottom distance even though the top voltage profile is the same?

▼ **Work through the solution**

At the upper metal the normal from metal into the dielectric is $-\hat{\mathbf{y}}$. With $E_y = -\partial_y V$, the metal charge density is

$$+\epsilon_0 \partial_y V = \epsilon_0 k V_d \coth(kb) \sin(kx),$$

$$k = \pi/a.$$

Integrating along $0 < x < a$ yields $2\epsilon_0 V_d \coth(\pi b/a)$ per length. Bringing the bottom closer increases $\coth(kb)$. For $kb \ll 1$ the expression tends to $\epsilon_0 b^{-1} \int_0^a f(x)dx$, the local parallel-plate limit. A segmented electrode supplies a nonuniform fixed profile; the result counts its integrated face charge and assumes the ideal voltage boundary remains prescribed in each segment.

Further reading

- Move the measurement plane above a patterned electrode
- Evaluate a grounded-enclosure boundary problem
- Griffiths, 5th ed., §3.3.1 · Cartesian separation

Electrostatic field theory and measurement · 7. Cylindrical geometry, coaxial boundaries and Bessel spectra

The radial coordinate of a cylinder leads to logarithms in one experiment and to Bessel functions in another. The difference is set by the variation along the axis and around the circumference. Recovering the radial weight from flux clarifies both.

A conducting annulus selects a logarithmic radial potential

Put a long cylindrical metal electrode of radius a inside a coaxial grounded tube of radius b, with homogeneous dielectric of permittivity ϵ in the gap. Long means that probes are far enough from both ends for axial variation to be negligible. With no free gap charge and no angular variation, Laplace's equation reduces to $\partial_r(r\partial_r V) = 0$. One integration gives $rV' = C$, a statement of conserved radial flux through cylinders; the second gives a logarithm. With inner electrode at V_0,

$$V(r) = V_0 \frac{\ln(b/r)}{\ln(b/a)},$$

$$E_r = \frac{V_0}{r \ln(b/a)},$$

$$C' = \frac{2\pi\epsilon}{\ln(b/a)}.$$

Here C' is capacitance per length. To obtain it, integrate D_r over the inner metal face per unit axial length and divide by V_0. Another route integrates the energy density $\epsilon E_r^2/2$ times $2\pi r\, dr$, giving $C'V_0^2/2$. For $a = 2$ mm, $b = 8$ mm and $\epsilon = 2\epsilon_0$, $C' = 80.3$ pF/m. At $r = \sqrt{ab} = 4$ mm the voltage is $V_0/2$; the half-voltage radius is the geometric mean. An equally spaced radial ruler does not sample equal voltage steps.

The logarithm is not allowed at the axis of a solid charge-free cylinder because it diverges and carries line-source flux. In the annular domain the singular axis is outside the field region. The same radial function can therefore be admissible or inadmissible in two problems with very similar geometry. Before discarding a mathematical solution, check which radii actually lie in the domain and whether singular sources were explicitly prescribed.

Nonuniform wall voltages in a two-dimensional disk

Next take a full disk of radius a with prescribed side voltage $f(\phi)$, invariant in z. Seek $R(r)e^{im\phi}$. Single-valuedness after $\phi \to \phi + 2\pi$ forces integer m. The radial equation $r^2R'' + rR' - m^2R = 0$ is solved by $r^{|m|}$ and $r^{-|m|}$ for $m \neq 0$; regularity at the centre chooses the growing power. The zero mode is constant, discarding its logarithmic partner. Thus an angular Fourier component on the wall penetrates as $(r/a)^{|m|}$. At $r = a/2$, a third harmonic retains one eighth while a sixth harmonic retains only one sixty-fourth. Both can coexist on the wall without being equally visible in the interior.

For $f(\phi) = 2 + 3\cos(2\phi)$ V, the interior is

$$2 + 3(r/a)^2 \cos(2\phi) = 2 + 3(x^2 - y^2)/a^2.$$

The potential has a saddle at the centre, with zero field there and opposite Cartesian curvatures. A scan along a single diameter could show a conspicuous maximum without detecting any charge. The two-dimensional Laplacian is zero because the curvatures cancel. This is also a useful finite-difference benchmark: a centred Cartesian Laplacian differentiates this quadratic exactly in the interior when the circular boundary has been represented correctly.

End effects produce a Sturm–Liouville problem

Now ground the side of a circular tube, excite its end at $z = 0$ and look for fields decaying into $z > 0$. For an axisymmetric source-free potential $V = R(r)Z(z)$, choose $Z = e^{-kz}$ for decay. Laplace's equation becomes

$$(rR')' + k^2 rR = 0.$$

The solutions are $J_0(kr)$ and $Y_0(kr)$. The former is regular with value one at zero; the latter diverges logarithmically. Grounding at $r = a$ requires $J_0(ka) = 0$, so $k_n = j_n/a$ with positive roots $j_1 = 2.40483,\ j_2 = 5.52008, \ldots$. Unlike sine wavenumbers they are not evenly spaced at low index.

Multiplying two such eigen-equations by the opposite radial function, subtracting and integrating gives $(k_m^2 - k_n^2) \int r R_m R_n \, dr$ equal to the endpoint term $r(R_n' R_m - R_m' R_n)$. It vanishes at a grounded wall and at a smooth axis. Hence modes at distinct eigenvalues are orthogonal with weight r. This weight arises physically from the circumference of a radial ring. Normalize dimensionless radius $x = r/a$. For a zero j_n, the squared modal norm is $J_1(j_n)^2/2$. An end profile $f(x)$ therefore gives

$$V(r, z) = \sum_n c_n J_0(j_n r/a) e^{-j_n z/a},$$

$$c_n = \frac{2}{J_1(j_n)^2} \int_0^1 x f(x) J_0(j_n x) \, dx.$$

For $f = V_0$ across the cap, $\int_0^1 x J_0(jx) dx = J_1(j)/j$ by differentiating $xJ_1(jx)$. Then $c_n = 2V_0/[j_n J_1(j_n)]$, with alternating signs. The cap voltage jumps at the wall corner, making its near-rim reconstruction slow. A smooth end pattern that already vanishes at a yields faster coefficient decay. In either case a narrow radial strip at the axis should not be given as much integration weight as an equal width near the circumference.

Two ways to see which axial scale survives

Far enough from the cap the first excited nonzero coefficient dominates. Its length is

$$a/j_1 = 0.416a.$$

Ratios of distant on-axis voltages can then identify tube radius without absolute calibration of the cap voltage. For $a = 15.0$ mm, an added 5.0 mm of distance should give factor $e^{-2.40483/3} = 0.449$. Repeat this estimate at several starting positions. The ratio only approaches that value once competing modes

have decayed. A datum deep enough to suppress all higher modes may also be close to the instrument noise level, which biases a fitted decay length.

Instead impose an axial sinusoid e^{ikz} on a very long cylinder's side wall. The radial equation now has the opposite sign, $(rR')' - k^2 rR = 0$, whose regular solution is the modified Bessel function $I_0(kr)$. The potential is $V_0 I_0(kr)\cos(kz)/I_0(ka)$. The experiment selects a real oscillation along z, not an exponential attenuation away from an end. Near the axis $I_0(\xi) = 1 + \xi^2/4 + \cdots$; the field generated by a long-wavelength side modulation is nearly uniform across a narrow cross section. At large ka, $I_0(ka)$ grows exponentially and the axis scarcely senses a short-wavelength wall pattern. Calling every radial solution simply a Bessel mode without tracking the separation sign can exchange these two physical predictions.

Construct a grid that treats the axis as a regular point

A conservative radial discretisation starts from $\int_{r_-}^{r_+} r\,\nabla_r^2 V\,dr = [rV']_{r_-}^{r_+}$. Use annular cells and fluxes at shared ring faces. Near $r = 0$ the left face has zero circumference, requiring no unknown point beyond the axis. For a nodal grid with Taylor expansion $V(r) = V(0) + Ar^2 + O(r^4)$, the radial operator at zero is $4A$, so $\nabla_r^2 V(0) \simeq 4[V(h) - V(0)]/h^2$. Simply setting $1/r = 0$ and using the Cartesian one-dimensional second derivative misses a factor of two.

Compare a numerical response to the analytic J_0 mode at a fixed z/a. Test the wall value, axis derivative, radial energy norm and distant decay rate. These are independent observables. A discretisation can reproduce the slowest length while misrepresenting an edge profile composed of many faster modes. Conversely, matching a discontinuous cap perfectly at mesh nodes does not demonstrate that high-order interior amplitudes are accurate.

Can a coaxial dielectric be graded to produce a uniform field?

The homogeneous coaxial field grows as $1/r$ and reaches its highest magnitude at the inner wire. A designer proposes grading the dielectric radially so the full annular gap experiences one common radial field. Assume both metal faces are equipotentials, the scalar permittivity $\epsilon(r)$ may vary continuously, and there is no free volume charge. Cylindrical symmetry still gives $D_r = \lambda/(2\pi r)$, independent of how the material is graded. A constant electric-field magnitude E_g requires $\epsilon(r) = \lambda/(2\pi E_g r)$, that is, permittivity proportional to $1/r$. This construction replaces a uniform material by a known spatial distribution; applying $\epsilon(r)\nabla^2 V = 0$ would miss the material gradient and give the wrong field.

Write $\epsilon(r) = \epsilon_a a/r$. At applied voltage V_0, the constant field is $E_g = V_0/(b - a)$, and charge per length is

$$\lambda = 2\pi\epsilon_a a V_0/(b - a).$$

Thus $C' = 2\pi\epsilon_a a/(b - a)$. For $a = 2.0$ mm, $b = 6.0$ mm, $\epsilon_a = 6\epsilon_0$, the outer relative permittivity is 2, and $C' = 167$ pF/m. At 12 V every radius experiences 3000 V/m. For a homogeneous dielectric at the same voltage the inner surface field would be $12/[0.002\ln 3] = 5461$ V/m regardless of its constant permittivity. The gradient reduces peak field at fixed voltage. Whether a real material benefits also depends on its local field limit, losses and interfaces.

Where is the polarisation charge that makes this possible? In the dielectric the true total charge satisfies

$$\rho_{\mathrm{tot}} = \epsilon_0 \nabla \cdot \mathbf{E} = \epsilon_0 E_g / r.$$

All of it is bound since $\rho_f = 0$. Direct evaluation from $\mathbf{P} = (\epsilon(r) - \epsilon_0) E_g \hat{\mathbf{r}}$ gives $-r^{-1}\partial_r(rP_r) = +\epsilon_0 E_g / r$, the same result. Material surfaces add bound charges $P_r(b)$ on the outer radial face and $-P_r(a)$ on the inner face. Integrating their surface contributions and the volume charge over a full annulus returns zero net bound charge, as required for a polarised neutral body. A constant radial electric magnitude can therefore coexist with zero free charge and nonzero total charge.

A manufactured graded material is often approximated by concentric homogeneous shells. For shell j from r_j to r_{j+1}, its voltage drop at fixed λ is $\lambda \ln(r_{j+1}/r_j)/(2\pi\epsilon_j)$. To reproduce the exact voltage drop of the continuously graded design over each shell, choose $\epsilon_j = \epsilon_a a \ln(r_{j+1}/r_j)/(r_{j+1} - r_j)$. This matches electrode capacitance even with relatively few shells. Yet the field inside each discrete shell still falls as $1/r$ and jumps at its interfaces. Matching terminal capacitance alone does not verify the target uniform-field profile. Compute the field just inside each shell edge and refine the radial layering against a local peak-field tolerance.

What a small displacement of the end electrode excites

An axisymmetric tube calculation retains only angular order $m = 0$. Now impose at the tube entrance a voltage profile proportional to $r\cos\phi$ near the axis and taper it to zero at the wall. This is a dipolar lateral pattern with positive voltage on one side and negative on the other. Separating $R(r)\cos(m\phi)e^{-kz}$ into Laplace's equation yields

$$(rR')' + \left(k^2 r - \frac{m^2}{r}\right) R = 0.$$

The regular solutions are $J_m(kr)$. Near the axis their leading powers are $r^{|m|}$, so an $m = 1$ potential vanishes on the centreline while its transverse gradient generally does not. The grounded wall selects zeros of J_m. The first zero of J_1 is 3.83171, larger than the first axisymmetric zero, so the slowest lateral disturbance decays faster than the slowest axisymmetric voltage. Probe design therefore changes which decay laws are experimentally accessible.

For a cap shaped exactly as $f = V_1 J_1(jr/a)\cos\phi$, $j = 3.83171$, the whole solution is a single mode. Near the axis $J_1(jr/a) \simeq jr/(2a)$; using $r\cos\phi = x$,

$$V \simeq V_1 j x e^{-jz/a}/(2a).$$

Hence $E_x(0,z) = -V_1 j e^{-jz/a}/(2a)$. With $V_1 = 2.0$ V and $a = 10$ mm, the lateral field starts at 383 V/m and falls to 8.30 V/m after an axial distance equal to one radius. A centreline voltmeter sees zero at both positions and gives no hint of the finite transverse force. A pair of voltmeters at small opposite x or a field probe oriented along x measures the relevant derivative.

For a realistic shifted electrode the input can contain both $m = 0$ and $m = 1$. Place detectors at $(r, 0, z)$ and (r, π, z). Their average selects even azimuthal orders; their difference selects odd ones. The ratio of

the difference at two heights tests $m = 1$ decay after higher odd modes subside. The average also includes $m = 2, 4, \ldots$, so attributing it to $m = 0$ requires symmetry information or more angular sensors. Reversing the applied electrode displacement changes the sign of an odd response but leaves even corrections unchanged to leading order. This experiment links angular Fourier selection and radial eigenvalues to a measurable alignment tolerance without requiring an entire three-dimensional voltage map.

Project a ring source without calling an axis cell a point detector

Deposit a thin circular ring of total free charge q at radius s in the plane $z = z_0$ inside a long grounded cylindrical tube. Its volume density in an axisymmetric model is $q\,\delta(r - s)\delta(z - z_0)/(2\pi s)$, for $s > 0$. The $2\pi s$ removes the circumference introduced by the volume element. A radial code that deposits q directly as a density in one mesh bin, without dividing by annular volume, changes the total charge each time the radius or mesh spacing changes. In this calculation one can track that normalisation all the way to a voltage coefficient.

Take the tube infinite in both axial directions for simplicity. Expand $V = \sum_n v_n(z) J_0(j_n r/a)$ in the grounded radial eigenfunctions. Project Poisson's equation using weight $r\,dr$ and the norm $a^2 J_1(j_n)^2/2$. The ring produces

$$[-d_z^2 + (j_n/a)^2] v_n = \frac{q J_0(j_n s/a)}{\pi \epsilon a^2 J_1(j_n)^2}\,\delta(z - z_0).$$

The infinite-line Green response for $-d_z^2 + k^2$ is $e^{-k|z-z_0|}/(2k)$, as can be checked by integrating its derivative across z_0. Thus each modal voltage has a finite amplitude and two decaying tails, and the complete field is obtained by summing radial modes. At a detector separated axially from the ring the exponential factors ensure rapid convergence. Exactly on the mathematical ring the local potential diverges logarithmically like that of a thin filament; comparing finite-element peak voltages there requires a finite wire cross section or a specified detector average.

An unexpected and testable consequence concerns the source radius. The coupling to mode n is proportional to $J_0(j_n s/a)$. It becomes exactly zero when the source sits on a radial node of that mode. The second mode has an interior node at $s/a = j_1/j_2 \simeq 0.4357$. A ring there still carries charge and excites many other modes but contributes no second modal coefficient. As an experimental analogy, a weighting potential that changes sign across part of a sensor can receive cancelling contributions from an extended excitation. Here the cancellation is encoded exactly in orthogonality and the pointwise coupling value.

A check independent of the ring-charge normalisation uses reciprocity. Imagine two narrow coaxial source/detector rings of radii s, r, each with the same definition of total charge and voltage measured around its circumference. The modal response contains the symmetric product $J_0(j_n s/a) J_0(j_n r/a)$ multiplied by a factor that depends only on axial separation. Swapping radius and axial position therefore leaves voltage per total charge unchanged. If a numerical table instead looks symmetric per charge density but not per total ring charge, check whether the radius-dependent integration weights were applied on both deposition and sampling.

Another useful test fixes total deposited charge but reduces a resolved annular thickness around s. First compare with an integral of the analytic ring solution across the same finite distribution. Then take the thin-ring limit at a detector away from the deposition region. Taking grid thickness and source thickness to zero together while comparing a self-voltage at the source conflates numerical refinement with a mathematically singular observable. Stating detector location, normalisation and order of limits turns a superficially straightforward ring source into a robust benchmark for cylindrical Poisson solvers.

Measure a radial cylindrical mode

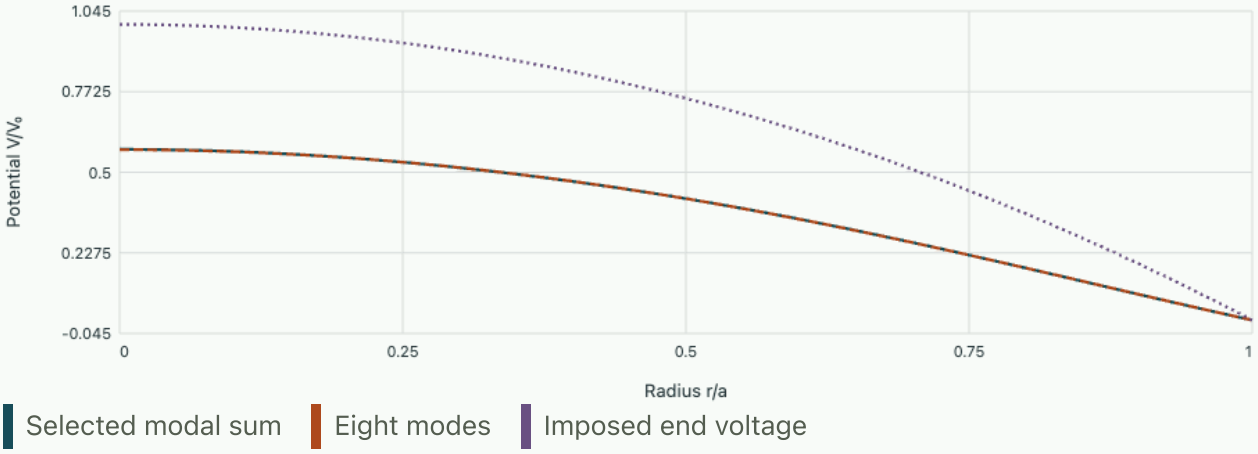
Change the electrode mode and compare radial nodes with axial decay. A cylindrical boundary chooses a different eigenbasis from a rectangular one.

Open this experiment online

Try the calculation

Height z/a **0.25**

Retained modes **3**



The on-axis potential is 0.5774 V_0 with the selected modes, versus 0.5765 V_0 using eight. The slowest mode has e-folding length 0.4158 a.

A cylindrical interface can carry induced bound charge with an entirely radial field

Separate the two charges detected at a coaxial material boundary

A long coaxial gap contains two homogeneous radial layers with interface at s, no free sheet there and fixed inner line charge $+\lambda$. Find the bound interface charge per length, verify it from the field jump, and determine the interface radius at which the layers take equal voltages. Evaluate the result for $a = 2$ mm, $b = 18$ mm and relative permittivities 2 and 4.

▼ **Derive and check the result**

For both regions $D_r = \lambda/(2\pi r)$. In material i the polarisation is $P_i = D_r(1 - \epsilon_0/\epsilon_i)$ along the radial direction. At the interface the outer face of material 1 has bound density $+P_1$, the inner face of material 2 has $-P_2$. Their sum times the circumference gives bound line charge $\lambda_b = \lambda(\epsilon_0/\epsilon_2 - \epsilon_0/\epsilon_1)$. It equals $-\lambda/4$ for the given ordering; the value per length is independent of s because the $1/s$ field cancels the interface circumference. A local density meter would see a different value at a different radius even though the integrated result stays the same.

Direct use of Gauss's law for the true electric field gives $\epsilon_0\, 2\pi s(E_{2r} - E_{1r}) = \lambda[\epsilon_0/\epsilon_2 - \epsilon_0/\epsilon_1]$, confirming the sign and magnitude. A check with displacement alone gives zero jump and thus correctly reports no *free* charge at that same surface. The two field laws answer different questions. Extending the shell calculation to include both exterior material faces returns zero total bound charge per material, so the internal negative interface charge never represents a net creation of charge.

The voltage drops are proportional to $\ln(s/a)/\epsilon_1$ and $\ln(b/s)/\epsilon_2$. Equality gives $s^{\epsilon_1+\epsilon_2} = a^{\epsilon_2} b^{\epsilon_1}$, or $s = a(b/a)^{\epsilon_1/(\epsilon_1+\epsilon_2)}$, a dimensionally transparent form. Numerically $s = 2(9)^{1/3} = 4.16$ mm. A half-voltage location at 4.16 mm is far from the arithmetic centre of this gap. It is also distinct from the 6.0 mm geometric mean for a homogeneous annulus, because the inner lower-permittivity layer needs less geometric thickness to take the same voltage.

Tutorial · radial modes with different physical boundaries

Replace the grounded wall by a reflecting boundary

A long tube has $\partial_r V = 0$ on its side instead of $V = 0$, and its end voltage profile is nonuniform. What happens to the radial eigenvalues and to the assertion that every end disturbance dies away?

▼ **Work through the solution**

The wall now requires

$$J_0'(ka) = -J_1(ka) = 0.$$

In addition a constant radial mode with $k = 0$ satisfies both side and axis conditions. Its axial equation is $Z'' = 0$, so boundedness allows a constant component. The end profile's weighted cross-sectional mean supplies that persistent offset. Only the components orthogonal to the constant decay. Requiring the potential to tend to zero at infinity while prescribing a nonzero mean cap voltage would be incompatible for this insulated side model unless other boundaries or sources are introduced.

Infer a radial charge density from a coaxial reading

Suppose the measured potential in a coaxial gap is $V = A\ln(b/r) + B(b^2 - r^2)$ with constant permittivity. Determine the volume charge and the signs of its contribution to the two face fluxes. Which

part can be changed by externally altering the applied voltage at fixed charge density?

▼ **Work through the solution**

The logarithmic term is harmonic away from the excluded axis. For the quadratic term $\nabla_r^2[-Br^2] = -4B$, giving uniform $\rho = 4\epsilon B$. The radial field is $A/r + 2Br$. The net outward displacement flux through the gap per length is $2\pi\epsilon[bE(b) - aE(a)] = 4\pi\epsilon B(b^2 - a^2)$, exactly $\rho\pi(b^2 - a^2)$. At positive B the quadratic term contributes positive outer flux and negative inward-face flux, with the larger outer contribution from both area and field. Changing only the source-free boundary excitation adds a logarithmic term and, if the reference permits it, a constant. It alters both field readings without altering the inferred uniform volume density.

Further reading

- Reconstruct a radial electrode profile with one to eight Bessel modes
- Fließbach, chapter 10 · cylindrical problems

Electrostatic field theory and measurement · 8. Normalised spherical harmonics and the field of an arbitrary three-dimensional source

An axial voltage scan can determine Legendre coefficients only if the source is rotationally symmetric. A general spherical electrode or molecule needs angular modes with two indices. Their normalisation fixes the conversion from a measured multipole coefficient to a physical voltage.

Separate three coordinates without imposing axial symmetry

In spherical coordinates the source-free Laplace equation is $r^{-2}\partial_r(r^2\partial_r V) + r^{-2}\Delta_\Omega V = 0$, where $\Delta_\Omega = (\sin\theta)^{-1}\partial_\theta(\sin\theta\,\partial_\theta) + (\sin^2\theta)^{-1}\partial_\phi^2$. Assume $V = R(r)Y(\theta,\phi)$. Separation requires

$$-\Delta_\Omega Y = \lambda Y.$$

Multiply by Y^* and integrate over the unit sphere. Integration by parts has no actual surface boundary for globally smooth fields and gives

$$\lambda \int |Y|^2 d\Omega = \int |\nabla_\Omega Y|^2 d\Omega \geq 0.$$

The constant mode has eigenvalue zero.

Separate azimuth as $e^{im\phi}$. Continuity after a full rotation forces integer m. With $u = \cos\theta$, the remaining equation becomes

$$[(1-u^2)P']' + [\lambda - m^2/(1-u^2)]P = 0.$$

Requiring regular behaviour at both poles yields $\lambda = \ell(\ell+1)$ with integers $\ell \geq |m|$. The corresponding angular functions are the associated Legendre functions. For each ℓ there are $2\ell+1$ independent m's, all with identical radial behaviour. A rotation mixes coefficients within this subspace but cannot convert a dipole into a quadrupole.

Fix one convention and verify its first modes

Choose $P_\ell(1) = 1$ and for $m \geq 0$ include the Condon–Shortley phase in $P_\ell^m(u) = (-1)^m(1-u^2)^{m/2}d^m P_\ell/du^m$. Define

$$Y_{\ell m} = \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}\,P_\ell^m(\cos\theta)e^{im\phi},$$

$$Y_{\ell,-m} = (-1)^m Y_{\ell m}^*.$$

With this convention $\int Y_{\ell m}^* Y_{\ell' m'}\,d\Omega = \delta_{\ell\ell'}\delta_{mm'}$, using $d\Omega = \sin\theta\,d\theta\,d\phi$. In particular $Y_{00} = 1/\sqrt{4\pi}$,

$$Y_{10} = \sqrt{3/(4\pi)}\cos\theta,$$

$$Y_{11} = -\sqrt{3/(8\pi)}\sin\theta e^{i\phi}.$$

To check the dipole normalisation independently,

$$\int \cos^2\theta \, d\Omega = 4\pi/3.$$

For Y_{11},

$$\int \sin^2\theta \, d\Omega = 8\pi/3.$$

The complex basis is an accounting device. A real potential obeys $v_{\ell,-m} = (-1)^m v^*_{\ell m}$, leaving $2\ell+1$ real angular degrees of freedom at each order.

On a sphere of radius a suppose segmented electrodes prescribe $f(\theta,\phi)$. Its coefficients are $v_{\ell m} = \int Y^*_{\ell m} f \, d\Omega$. These have units of volts because the angular functions and solid angle are dimensionless. The radial equation $(r^2R')' - \ell(\ell+1)R = 0$ has solutions $r^\ell, r^{-\ell-1}$. For an empty interior, regularity selects $V_{\text{in}} = \sum v_{\ell m}(r/a)^\ell Y_{\ell m}$. For empty exterior decaying at infinity,

$$V_{\text{out}} = \sum v_{\ell m}(a/r)^{\ell+1} Y_{\ell m}.$$

Nonuniform prescribed surface voltage requires electrically distinct electrode patches or a specified ideal boundary; one continuous metal conductor could not sustain that static variation.

Represent a tilted dipole in the laboratory frame

Impose $f = V_0 \hat{\mathbf{p}}\cdot\hat{\mathbf{r}}$ with $\hat{\mathbf{p}} = (\sin\alpha, 0, \cos\alpha)$. In the rotated frame with polar axis along $\hat{\mathbf{p}}$ this is a pure $m=0, \ell=1$ pattern. In the laboratory it is $V_0[\cos\alpha\cos\theta + \sin\alpha\sin\theta\cos\phi]$. Using the explicit harmonics,

$$v_{10} = V_0\sqrt{4\pi/3}\cos\alpha,$$

$$v_{11} = -V_0\sqrt{2\pi/3}\sin\alpha,$$

$$v_{1,-1} = +V_0\sqrt{2\pi/3}\sin\alpha.$$

The sum of their squared magnitudes is $4\pi V_0^2/3$, independent of tilt. The interior solution simplifies to $V_0\hat{\mathbf{p}}\cdot\mathbf{r}/a$, giving uniform $\mathbf{E} = -V_0\hat{\mathbf{p}}/a$. At $a = 40$ mm, $V_0 = 2.0$ V and $\alpha = 30^\circ$, $E_x = -25.0$ V/m, $E_z = -43.3$ V/m. An axis-only fit would infer the correct axial component while losing the transverse one. A full vector field measurement at one interior point distinguishes the rotated dipole from a weaker aligned one.

Tie angular modes to an actual charge distribution

The addition theorem expresses $P_\ell(\hat{\mathbf{r}}\cdot\hat{\mathbf{s}}) = [4\pi/(2\ell+1)]\sum_m Y_{\ell m}(\hat{\mathbf{r}})Y^*_{\ell m}(\hat{\mathbf{s}})$. Rotational invariance explains its structure: the sum is unchanged by simultaneous rotation of both directions, so it depends only on their angle. Put $\hat{\mathbf{s}}$ at the north pole, where only $m=0$ survives, and the proportionality constant follows from $Y_{\ell 0}$ and $P_\ell(1)=1$. Checking $\ell=1$ recovers the dot product of the two unit vectors.

Use the generating series $1/|\mathbf{r}-\mathbf{s}| = \sum_{\ell\ge 0}(r_<^\ell/r_>^{\ell+1})P_\ell(\hat{\mathbf{r}}\cdot\hat{\mathbf{s}})$. For an observation radius outside every source point, define $q_{\ell m} = \int \rho(\mathbf{s}) s^\ell Y^*_{\ell m}(\hat{\mathbf{s}})\, d^3s$. Inserting the addition theorem into the Coulomb integral gives

$$V(\mathbf{r}) = \frac{1}{\epsilon_0}\sum_{\ell,m}\frac{q_{\ell m}}{(2\ell+1)r^{\ell+1}}Y_{\ell m}(\hat{\mathbf{r}}).$$

Unlike boundary voltage coefficients, $q_{\ell m}$ carries units C m$^\ell$. At $\ell=0$, $q_{00} = Q/\sqrt{4\pi}$ and multiplication by Y_{00} returns $Q/(4\pi\epsilon_0 r)$, fixing the prefactor. At $\ell=1, m=0$,

$$q_{10} = \sqrt{3/(4\pi)}\,p_z.$$

The two transverse coefficients reconstruct p_x, p_y. Measuring the voltage coefficient at two radii can distinguish a decaying exterior multipole from a regular potential imposed by distant equipment, whose local expansion instead grows as r^ℓ.

Design an angular quadrature that measures the claimed coefficient

Uniformly spaced latitude rings do not have equal area. In angular integrals substitute $u=\cos\theta$, giving $d\Omega = du\,d\phi$. Equal spacing in u and ϕ already avoids overweighting the polar caps. More accurate quadrature uses Gauss–Legendre nodes in u with their weights, combined with equispaced azimuthal samples for the Fourier factors. If a known angular pattern contains $|m|\le M$, at least $2M+1$ azimuthal samples are needed to distinguish these signed orders on a fixed ring. A higher order present in an unknown pattern can alias into the fitted low orders, so adding more sample angles and reserving independent test directions is informative.

A robust reconstruction checks the constant mode on $f=1$, the total power $\sum|v_{\ell m}|^2 = \int |f|^2 d\Omega$ for an exactly represented pattern, and the predicted angular field at withheld directions. With noisy overdetermined data use a weighted linear least-squares fit to the modes, including weights for sampling geometry or measurement uncertainty as appropriate. The two choices are physically different: one approximates an angular integral, the other quantifies independent noisy observations. State which objective your fit minimises. An angular reconstruction can predict a voltage while still having very poor sensitivity to high-order exterior source moments, whose signals decay rapidly with radius.

A displaced charge supplies an exact all-orders convergence experiment

Place a point charge q at $s\hat{\mathbf{z}}$ within an otherwise empty sphere used as a measurement surface. At radius $R > s$ the potential is $q/[4\pi\epsilon_0\sqrt{R^2 + s^2 - 2Rs\cos\theta}]$. The axial symmetry makes only $m = 0$ coefficients nonzero, giving the exact Legendre series $[q/(4\pi\epsilon_0 R)]\sum_{\ell\geq 0}(s/R)^\ell P_\ell(\cos\theta)$. This example exhibits the convergence parameter directly. It is the radius of the enclosing source relative to the measurement radius, not the magnitude of the charge or the number of sampled angles.

On the ray nearest the point charge, $\theta = 0$, every Legendre polynomial is unity. Truncation at order L there gives a geometric partial sum and a relative error exactly $(s/R)^{L+1}$ when compared with the exact voltage at that point. At $s/R = 0.40$, six retained orders $\ell = 0, \ldots, 5$ leave 0.410% relative error. At $s/R = 0.85$ the same approximation leaves 37.7%. Achieving 1% in the latter case requires $L + 1 > \ln(0.01)/\ln(0.85) = 28.3$, so at least 29 orders. A claim that retaining through quadrupole always suffices has no physical meaning without an observation radius and a desired error.

On the opposite ray $P_\ell(-1) = (-1)^\ell$, making the terms alternate. The absolute voltage error there has a smaller prefactor $1/(1 + s/R)$, in contrast to $1/(1 - s/R)$ on the near side. For this particular geometric series its relative error magnitude is still $(s/R)^{L+1}$, despite the alternating signs. At other angles the angular polynomial oscillations can produce much smaller error but not a uniformly improved bound. The simple inequality $|P_\ell(u)| \leq 1$ gives an absolute bound $|q|(s/R)^{L+1}/[4\pi\epsilon_0 R(1 - s/R)]$ for the entire angular surface. For an arbitrary signed source replace $|q|$ by integrated absolute charge and s by a radius enclosing all source points. A nearly neutral distribution can have a small exact voltage at some angles, making any comparable relative bound much less informative.

Now rotate the point charge through polar and azimuthal angles θ_s, ϕ_s without changing its radius. The coefficient is exactly $q_{\ell m} = q s^\ell Y^*_{\ell m}(\theta_s, \phi_s)$. Summing $|q_{\ell m}|^2$ over m gives $q^2 s^{2\ell}(2\ell + 1)/(4\pi)$, by setting the two angular arguments equal in the addition theorem. This rotational invariant is a strong test of a numerical angular projection: a rotation changes individual complex phases, but the total power at each angular order remains fixed. Rotating a sensor array rather than the source should produce the equivalent opposite rotation of the fitted description.

A spherical split electrode drives an interior field without net charge

Construct a spherical boundary of radius a whose northern half is held at $+V_0$ and southern half at $-V_0$, separated by a thin insulating equatorial seam. In the mathematical limit take $f(u) = V_0 \operatorname{sgn} u$, $u = \cos\theta$. The problem is axisymmetric and changes sign under reflection through the equator, so only odd ℓ occur. Using the Legendre (rather than unit-normalised harmonic) coefficients, the series inside is $V = \sum b_\ell (r/a)^\ell P_\ell(u)$, where $b_\ell = (2\ell + 1)V_0 \int_0^1 P_\ell(u)\, du$ for odd ℓ. There is no monopole, hence no change to the net charge when the two voltages are applied symmetrically in a symmetric environment.

The first two nonzero coefficients can be found without a general table. $P_1 = u$ gives

$$b_1 = 3V_0/2.$$

$P_3 = (5u^3 - 3u)/2$ integrates from zero to one to $-1/8$, so $b_3 = -7V_0/8$. At the centre all terms with $\ell \geq 2$ have vanishing gradient, leaving an exact central field $E_z(0) = -3V_0/(2a)$. Even though the electrode voltages jump abruptly at the equator, a single low-order coefficient determines this point measurement exactly. A finely resolved electrode charge density near the seam requires a very different number of modes and a model of the seam width.

Set $a = 30$ mm and $V_0 = 2.0$ V. The central field is -100 V/m. At $z = 0.20a$ on the northern axis, the $\ell = 1$ voltage contributes 0.600 V and $\ell = 3$ adds -0.0140 V. The first omitted odd term scales as $(0.2)^5$ times its angular coefficient. At $z = 0.9a$ the same rapidly convergent two-term approximation is no longer trustworthy. A student can compare the actual surface voltages estimated from a truncated expansion to predictions at selected interior radii. A converged central field and a conspicuously rounded numerical step can coexist in the same exact analytic model.

For a finite insulating angular seam suppose instead the boundary interpolates continuously and antisymmetrically near $u = 0$. Only a thin band changes relative to the sign function. The dipole coefficient depends on $\int f(u) u \, du$; the extra weight u suppresses the influence of perturbations concentrated around the equator. If the band spans $|u| < \delta$ and the change in potential is bounded by V_0, the dipole coefficient error is at most $3V_0\delta^2/2$. High-order edge charge can nevertheless change markedly. This estimate shows how a real imperfection can have a small impact on one observable and a large impact on another. It also provides a quantitative tolerance when using the ideal split boundary to predict a centre-field calibration.

Separate material inside a measurement sphere from fields imposed outside it

A laboratory sometimes measures voltages on two concentric source-free spherical surfaces, without knowing whether the disturbance comes from an object inside both surfaces or from distant electrodes outside both. A single angular map determines the coefficient of each measured harmonic at that radius but cannot in general distinguish these two origins. In an annulus lacking sources, each angular component takes the radial form $v_{\ell m}(r) = A_{\ell m} r^{\ell} + B_{\ell m} r^{-\ell-1}$. The regular term is compatible with external forcing, and the decaying term with sources inside the inner radius. Both are harmonic in the measurement shell.

For a single $\ell = 1$ component, denote its coefficient of an ordinary $\cos\theta$ pattern by $v(r) = a_1 (r/R_1) + b_1 (R_1/r)^2$, so a_1, b_1 are voltages at the first radius. If two scans give $v(R_1) = 2.00$ V and $v(2R_1) = 1.30$ V, the equations are

$$a_1 + b_1 = 2.00, \; 2a_1 + b_1/4 = 1.30.$$

Multiply the first by one quarter and subtract, giving $a_1 = 0.4571$ V and $b_1 = 1.5429$ V. At a third radius $1.5R_1$ the predicted value is 1.371 V. A pure internal dipole would have returned 0.50 V on the second sphere, while a pure uniform external field would have returned 4.00 V. Fitting a single dipole to the first reading alone would grossly misdescribe the second.

For general order the determinant of the two normalised equations is $t^{-\ell-1} - t^{\ell}$ with $t = R_2/R_1$. As the two radii approach each other, this determinant tends to zero and the coefficients cannot be separated

reliably. Increasing separation improves discrimination of the two radial powers up to the point where one component falls below the measurement noise or additional sources enter the annulus. The monopole case deserves separate attention: its growing branch is just a constant voltage offset, while its decaying branch reveals total charge. A voltage reference common to both scans is therefore indispensable when using their difference to infer monopole charge.

Suppose the field was measured rather than voltage. Radial field coefficients are $-\ell A_{\ell m} r^{\ell-1} + (\ell+1) B_{\ell m} r^{-\ell-2}$. Measuring both radial field and voltage on one sphere can, in principle, supply the two independent equations needed to separate regular and decaying parts. For $\ell = 0$ the constant reference disappears from the radial field. An independently calibrated force probe then constrains total internal charge even when the voltmeter has a common offset. Numerically estimated radial derivatives from two very close voltage scans, however, share measurement errors with those voltages; treating them as two independent precise observations overstates the information gained.

A final issue is the coordinate origin. Moving it can create higher moments from lower ones for a nonneutral distribution. If the actual source is a displaced point charge and the coordinate origin shifts toward it, the series converges faster at a given observation sphere. If the charge distribution is unknown, choosing the origin solely to minimise apparent high-order coefficients can hide genuine structure or amplify noise. Record the physical origin of the measuring geometry, report withheld predictions, and compare the integrated monopole flux, which is independent of the choice of origin as long as the source remains enclosed. This separates a useful coordinate convention from an inferred physical symmetry.

Propagate measured boundary harmonics

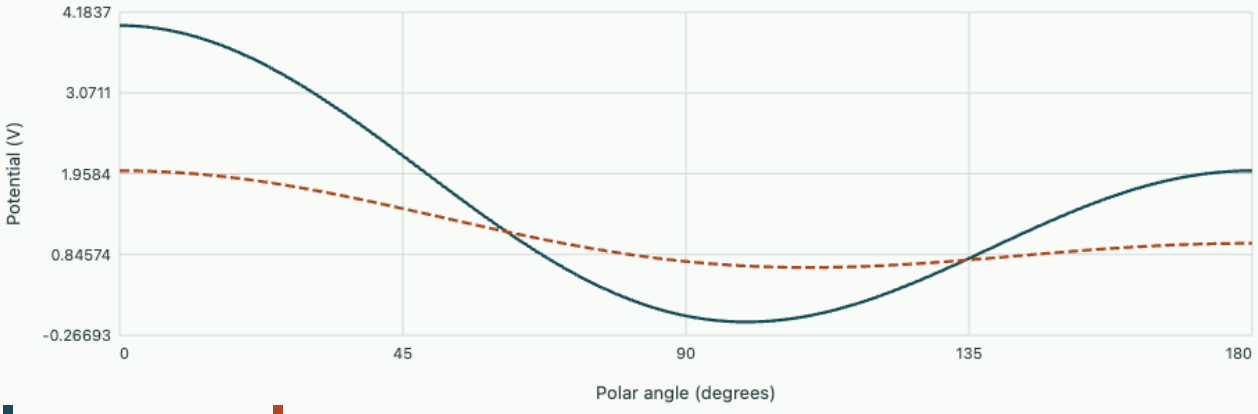
Choose angular boundary data and examine its attenuation away from the spherical surface. Relate the relative loss of detail to harmonic degree.

Open this experiment online

Try the calculation

Observation radius r/a **0.5**

Quadrupole on the sphere (V) **2**



The north pole and equator read 2 V and 0.75 V at this radius. Both branches coincide on the boundary.

Predict an angular sampling requirement before buying more probes

How many angular measurements are needed for a declared band limit?

A real boundary pattern is known, independently of the current measurement, to contain only harmonics with $\ell \le 3$. Count its real degrees of freedom. Design a separable quadrature that exactly projects such a pattern onto the same harmonics in exact arithmetic. Explain what can still go wrong when the true band limit was assumed incorrectly.

▼ **Derive and check the result**

There are $\sum_{\ell=0}^{3}(2\ell+1) = 16$ independent real degrees of freedom. Sixteen arbitrary readings need not determine them: their locations must give a full-rank observation matrix. Equatorial-only data, for example, miss every axisymmetric odd pattern there. A straightforward overdetermined grid uses uniform azimuths and Gauss–Legendre nodes in $u = \cos\theta$. For $L = 3$ use at least $2L+1 = 7$ equally spaced azimuths so that differences $m - m'$ up to 6 cannot alias a nonzero azimuthal difference into the constant component of a projection.

After that exact azimuthal sum, only equal m terms survive. Their products in u are proportional to $P_\ell^m(u)P_{\ell'}^m(u)$. Each contains $(1-u^2)^{|m|}$ times a polynomial and is itself a polynomial of degree at most $\ell + \ell' \le 6$. A four-point Gauss–Legendre rule integrates polynomials through degree 7 exactly. Thus four latitudes times seven azimuths, 28 weighted samples, suffice to reconstruct all 16 coefficients for the

declared band limit. Both polar caps are represented by quadrature weights; no point must lie precisely on either pole.

If the true field contains order $m = 7$, then $e^{i7\phi} = 1$ at the seven uniformly spaced azimuths starting at zero. Its azimuthal variation can masquerade as a constant on each sampled circle. A second scan shifted by half an azimuthal spacing flips the sign of this harmonic but preserves a true constant, exposing the alias. Unmodelled high latitudinal orders can similarly corrupt low coefficients through the finite radial quadrature. With noisy measurements evaluate withheld off-grid directions and report both the sensitivity to additional modes and the precision assumed in the protected low orders. An exactly integrating grid under a band-limit hypothesis is not a proof that the physical source obeyed that hypothesis.

Tutorial · derive a surface charge and test identifiability

Two spheres agree on their voltage patterns

The same prescribed $v_{\ell m}$ is realised on an infinitesimally thin spherical electrode array between identical dielectric media of permittivity ϵ. Find the coefficient of its free charge density. Which integrated measurement isolates the monopole?

▼ Work through the solution

The interior and exterior radial derivatives on $r = a$ are $\ell v_{\ell m} Y_{\ell m}/a$ and $-(\ell+1) v_{\ell m} Y_{\ell m}/a$. With outward normal pointing from interior to exterior, $\sigma_f = \epsilon(E_{\mathrm{out},r} - E_{\mathrm{in},r})$, giving coefficient $\epsilon(2\ell+1)v_{\ell m}/a$. A whole-surface integral of $\sigma_f a^2 d\Omega$ retains only the monopole since all nonconstant harmonics are orthogonal to Y_{00}. For uniform boundary voltage V_0, $v_{00} = \sqrt{4\pi} V_0$ and total charge $4\pi\epsilon a V_0$, the capacitance of an isolated sphere times voltage. High-order voltage components can instead demand strong local charge densities despite zero net added charge.

A ring of detectors misses an entire pattern

Twelve probes placed uniformly on the equator read zero voltage. Can this exclude a substantial $\ell = 1$ component? Give an explicit counterexample and a second configuration still invisible after adding a north-pole probe.

▼ Work through the solution

A pure Y_{10} vanishes all along the equator, whatever its amplitude. A north-pole reading would detect it. But the real quadrupole proportional to $\sin\theta\cos\theta\cos\phi$, corresponding to xz/r^2, vanishes on both the equator and at either pole. An intermediate ring or a set of oblique directions is needed. Increasing the number of equatorial samples alone cannot resolve a lack of checks at other latitudes. Under a declared band limit compute the rank and singular values of the probe-by-harmonic matrix before interpreting a null reading as an absent multipole.

Further reading

- Compare interior and exterior harmonics with a polar scan
- Infer a tilted quadrupole from six independent probe readings
- Fließbach, chapter 11 · spherical harmonics
- Griffiths, 5th ed., §§3.3–3.4 · spherical separation and multipoles

Electrostatic field theory and measurement · 9. Capacitance matrices, floating nodes and controlled approximation bounds

A multi-electrode device generally has no single capacitance until the other terminals have been connected or isolated. An energy identity determines the matrix that relates voltages to charges, and a complementary flux identity can certify estimates even when the exact field is unknown.

Build the matrix from unit-voltage experiments

Consider n isolated conductor surfaces embedded in a connected linear dielectric, plus a grounded reference conductor or decay relative to a reference at infinity. There is no free charge in the dielectric. Define the weighting function $w_j(\mathbf{x})$ by putting conductor j at 1 V, all other conductors and the reference at zero, then dividing by 1 V. For arbitrary voltages V_j, linearity and uniqueness give

$$\phi(\mathbf{x}) = \sum_j V_j w_j(\mathbf{x}).$$

The charge on conductor i is an integral of displacement leaving its metal into the dielectric. Using a normal $\mathbf{n}$ pointing out of the dielectric into the metal gives

$$Q_i = \int_{S_i} \epsilon \partial_n \phi \, dA.$$

Therefore

$$Q_i = \sum_j C_{ij} V_j,$$

$$C_{ij} = \int_{S_i} \epsilon \, \partial_n w_j \, dA = \int_\Omega \nabla w_i \cdot \epsilon \nabla w_j \, dV.$$

The last equality follows by multiplying the charge-free field equation for w_j by w_i and integrating by parts. The boundary integral selects just conductor i because its weighting function is unity there and zero on every other conductor. The volume expression proves reciprocity $C_{ij} = C_{ji}$ for a symmetric dielectric tensor. The diagonal terms are positive, since they integrate the squared fields of each independent one-volt experiment. In an ordinary connected isotropic configuration a positive voltage on j with $i \neq j$ grounded puts negative induced charge on i; the corresponding mutual coefficient C_{ij} is nonpositive.

The stored energy is

$$U = \tfrac{1}{2} \sum_{ij} V_i C_{ij} V_j.$$

Positive material response makes the reduced matrix with an effective reference positive definite under usual connected boundary conditions. If all conductors including the external closing conductor are retained without fixing a reference, changing all voltages by a common constant produces no field. The full matrix then has a zero common-voltage mode and each row sums to zero. Removing the grounded

reference from the matrix removes this gauge null direction; its remaining rows need not sum to zero. Treating both statements as universal would create a spurious constraint on measured terminal data.

Convert a field matrix to a circuit measurement

A two-port device might have reduced matrix in picofarads $C = \begin{pmatrix} 12 & -3 \\ -3 & 8 \end{pmatrix}$ relative to its grounded enclosure. With conductor 1 at 2 V and conductor 2 at 0 V, charges are 24 pC and -6 pC. The grounded enclosure takes the compensating -18 pC. Reciprocity predicts that holding conductor 1 at 0 V and conductor 2 at 2 V puts -6 pC on conductor 1. The diagonal 12 pF includes both the coupling to conductor 2 and the path to the enclosure. In an equivalent lumped model the positive pair capacitance between 1 and 2 is 3 pF, while the two separate couplings to ground are 9 and 5 pF. The sign of a Maxwell matrix off-diagonal is opposite to a positive capacitor drawn between two nodes.

Now disconnect conductor 2 after neutralising it. Its total charge stays zero while its voltage becomes an unknown. The second matrix row demands $-3V_1 + 8V_2 = 0$. Thus $V_2 = 3V_1/8$, giving apparent capacitance seen at terminal 1 equal to $12 - 9/8 = 10.875$ pF. Grounding conductor 2 would instead yield 12 pF. The change is a measurable effect of the prepared state, not a change in the dielectric constant of the material. If conductor 2 already carries charge Q_2^0, its voltage acquires a constant offset Q_2^0/C_{22} in addition to the response to V_1.

For several floating nodes write their block equation as $Q_f = C_{fa}V_a + C_{ff}V_f$. If the reference fixes the potential freedom, solve $V_f = C_{ff}^{-1}(Q_f - C_{fa}V_a)$ and substitute into the active-node block. The differential active capacitance is the Schur complement $C_{aa} - C_{af}C_{ff}^{-1}C_{fa}$. It is symmetric and positive for a positive full reduced matrix. With fixed floating charges the constant term shifts active charges but leaves this differential matrix unchanged in a linear medium. An actual cable measurement that results from leaving a neighbouring shield unconnected should be compared to this elimination, not to the original diagonal.

Obtain an upper capacitance bound by trying a voltage

For a two-terminal charge-free dielectric at prescribed voltage difference V, every admissible continuous trial potential satisfying the electrode voltages gives an energy at least as large as the exact electrostatic energy. This follows by writing the trial as exact plus a perturbation vanishing on fixed-voltage electrodes and expanding the energy. The mixed term integrates to zero because the exact displacement is divergence free. If exterior walls are insulated, the exact normal displacement vanishes there and the trial need not be assigned a fixed potential on them. Thus $C \le 2U_{\mathrm{trial}}/V^2$. The fields must genuinely arise as a gradient of one trial voltage. Guessing unrelated components of an electric field does not provide this upper bound.

For a capacitor of plate area A, separation d, and scalar permittivity varying in the gap, choose a linear voltage along the plate normal. The trial energy gives

$$C/A \le \langle \epsilon \rangle / d,$$

where the average is over the gap volume divided by Ad. For dielectric columns that run parallel to the applied field with no change of material along the field direction, the linear potential can be exact. For layers stacked normal to the field it usually overestimates capacitance. The same arithmetic average thus serves as the actual result or a strict upper bound depending on geometry, even when the material fractions are identical.

A complementary trial with a divergence-free displacement

Fix total plate charges $\pm Q$ instead, taking no volume free charge. An admissible displacement trial must have the prescribed flux into each electrode, no flux through insulating side walls, continuous normal component across free-charge-free interfaces, and zero divergence inside. It need not correspond to an exact potential-gradient electric field. Consider $U_D[\mathbf{D}] = \frac{1}{2}\int \mathbf{D}\cdot\epsilon^{-1}\mathbf{D}\,dV$. Write a trial as $\mathbf{D}_* + \delta\mathbf{D}$. The cross term is $\int(-\nabla\phi_*)\cdot\delta\mathbf{D}\,dV = 0$: integration by parts gives zero bulk divergence and each conductor has constant voltage multiplying zero net flux change. Thus $U_D \geq Q^2/(2C)$ and

$$\frac{Q^2}{2U_D^{\mathrm{trial}}} \leq C \leq \frac{2U_V^{\mathrm{trial}}}{V^2}.$$

Unlike two competing estimates, these inequalities identify which side contains the exact answer. For the same parallel-plate geometry choose constant displacement $D_z = Q/A$. It yields $C/A \geq 1/[d\langle 1/\epsilon\rangle]$, the harmonic-mean lower bound. For layers stacked normal to the plates that trial is exact. For columns parallel to the field the upper bound is exact. An arbitrarily shaped two-phase composite satisfies both under the stated plate and insulating-side conditions. This gives a strong physical check before solving detailed microstructure on a mesh.

Use measured numbers to rule out a material model

Take equal volume fractions of relative permittivities 2 and 8 in a cell of area 100 mm² and thickness 1.0 mm. The harmonic mean is 3.2, arithmetic mean 5.0. Since $\epsilon_0 A/d = 0.8854$ pF, the admissible capacitance interval is 2.83–4.43 pF. A measured 3.3 pF is consistent with many internal geometries and does not determine how the two materials are arranged. A calibrated value of 5.2 pF would require examining fringing, area and thickness, parallel cable capacitance, dispersive or conducting response, and the assumed constituent data. Merely refining an interior mesh cannot produce a static value outside the proven interval if it faithfully solves the same ideal model.

For a field computation use the matrix reciprocity relation as one diagnostic and the two variational bounds as another. An electrostatic finite-element potential may produce a reliable energy estimate while numerical differentiation at a sharp electrode edge produces poor charge density. Integrating boundary fluxes gives each terminal charge, but reconstructing C_{ij} by the mixed energy integral often has a different error sensitivity. Comparing both over a grid sequence can indicate whether the error is concentrated at edges or in the field volume. Symmetrising a noisy measured matrix by averaging its transpose removes an antisymmetric discrepancy but cannot establish positivity or explain an inaccurate common scale.

Construct a complete measurable matrix without numerical differentiation

Three thin spherical conductors have radii $a < b < c$. The outer one is tied to the reference potential. Homogeneous material of permittivity ϵ_1 fills the inner annulus $a < r < b$ and another material ϵ_2 fills $b < r < c$. Spherical symmetry gives each annulus its own independent radial displacement field fixed by its electrode voltage difference. Define the ordinary two-sphere gap capacitances $K_1 = 4\pi\epsilon_1 ab/(b-a)$,

$$K_2 = 4\pi\epsilon_2 bc/(c-b).$$

The inner conductor takes $K_1(V_1 - V_2)$. The middle one has two faces, carrying the opposite inner gap charge and an outer gap charge $K_2 V_2$. Thus $Q_1 = K_1 V_1 - K_1 V_2$ and $Q_2 = -K_1 V_1 + (K_1 + K_2)V_2$. Every sign and every term can be read from a physical face.

In vacuum choose $a = 10$ mm, $b = 20$ mm, $c = 40$ mm. Then $K_1 = 2.225$ pF and $K_2 = 4.451$ pF. Driven at 3.0 V with the middle conductor grounded, the inner sphere carries 6.68 pC and the outer gap has zero field. With the middle conductor floating and neutral, $V_2 = K_1 V_1/(K_1 + K_2) = 1.0$ V. The inner sphere now carries 4.45 pC. The inner and outer faces of the floating sphere carry exactly opposite charges, although the conductor has a nonzero voltage. Integrating a current meter attached to the conductor during charging need not tell you each face's charge separately.

Now electrically connect the inner two conductors. Their common voltage sets the inner field to zero. The combined conductor has capacitance K_2 to the exterior, independent of a. This third preparation is distinct from grounding or floating the middle shell. Before calculating an experimental 'device capacitance', draw the electrical connections as well as the geometric cross section. In a larger matrix a driven conductor and a driven guard at the same voltage should have their charges combined only if the experimental ammeter actually measures their total current.

One can also recover the two constituent permittivities from just two carefully selected scalar measurements. Ground the middle shell to determine K_1 directly. Then float it neutral and measure $K_{\mathrm{ser}} = K_1 K_2/(K_1 + K_2)$, giving $K_2 = K_{\mathrm{ser}} K_1/(K_1 - K_{\mathrm{ser}})$. If K_2 is very large, the floating and grounded readings nearly coincide and their difference becomes noise sensitive. An alternative is to drive inner and middle together and measure the outer-gap current directly. The algebra selects a better experimental protocol instead of hiding ill-conditioning in a formula.

Tighten a variational bound without solving the entire field

Take the stacked two-layer capacitor with no space charge, equal layer thickness and isotropic permittivities $\epsilon_1 = 2\epsilon_0, \epsilon_2 = 8\epsilon_0$. A single linear trial produces the arithmetic-mean upper bound. Insert one interior voltage degree of freedom at a point that coincides with the material interface. The resulting piecewise-linear trial space contains the exact solution. Minimising its energy makes the normal displacements equal, reproducing the harmonic-mean capacitance exactly. A global quadratic trial with the same number of extra parameters will usually perform worse because its derivative is continuous while the exact normal field jumps.

To see this quantitatively write $x = z/d$ on the unit interval and $V(x) = V_0[x + \eta x(1-x)]$, with the lower half at relative permittivity 2 and upper half at 8. This trial fixes both endpoints. Its dimensionless

capacitance upper estimate is

$$c(\eta) = \int_0^1 \epsilon_r(x)[1 + \eta(1 - 2x)]^2 dx.$$

Integrating gives

$$c(\eta) = 5 - 3\eta + (5/3)\eta^2.$$

The best quadratic trial has $\eta = 0.9$ and $c = 3.65$, compared with 5.00 for the straight line and 3.20 for the exact layered solution. The unrefined constant-displacement lower trial already supplies 3.20, so the two-field comparison certifies an interval before any field-map visualisation.

Why did the smooth trial still overestimate by 14%? At its endpoints it predicts slopes 1.9 and 0.1 times the mean field magnitude, whereas the exact field has slopes 1.6 and 0.4 in the two halves. Optimising a volume energy distributes a global smooth approximation across both layers. It has no mechanism to jump at the known interface, so an error persists even with its one coefficient optimised perfectly. The example explains why fitting material interfaces in a mesh, or using basis functions that admit correct normal derivative jumps, can matter more than merely increasing polynomial degree.

In an irregular two-dimensional composite an exactly divergence-free complementary trial can be constructed from a stream function Ψ: take $\mathbf{D} = (\partial_y \Psi, -\partial_x \Psi)$ in Cartesian coordinates. Its divergence vanishes identically. Total flux through an oriented path is a difference of Ψ at its endpoints, so carefully chosen boundary values enforce required electrode charges and zero leakage on insulating sides. Optimising $\int \mathbf{D} \cdot \epsilon^{-1} \mathbf{D}/2$ over such trial functions tightens the lower capacitance bound. Optimising the independent voltage trial tightens the upper one. The difference between their resulting bounds quantifies what the chosen approximation spaces still fail to resolve.

What an AC bridge measures when a nominally floating terminal leaks

A real high-impedance lead is neither perfectly floating nor an ideal ground at every frequency. Add a leakage resistance R from conductor 2 of the two-port example to its grounded enclosure. Keep the reduced static capacitance matrix C fixed, linear and quasistatic at the frequencies of interest. Let terminal 1 be driven sinusoidally and terminal 2 have no other external drive. Using $e^{i\omega t}$ here for circuit phasors, charge continuity at terminal 2 gives

$$i\omega(C_{21} V_1 + C_{22} V_2) + V_2/R = 0.$$

The outward conductive current plus the rate of charge stored in the electrode must vanish. Solving for the active terminal admittance gives

$$Y_{\mathrm{in}} = i\omega C_{11} + \frac{\omega^2 C_{12} C_{21}}{1/R + i\omega C_{22}}.$$

At low frequency the leakage holds terminal 2 close to ground, so the leading measured capacitance is C_{11}. At high frequency the floating condition is recovered and $Y_{\mathrm{in}}/(i\omega) \to C_{11} - C_{12}C_{21}/C_{22}$. The effective admittance also has a positive real part, since the driven voltage on terminal 2 supplies Joule heat in the leakage resistance. This term persists even if the dielectric itself is lossless. An AC capacitance meter reporting a frequency-dependent loss tangent might therefore be detecting a boundary preparation rather than intrinsic material loss.

Take $C_{22} = 8$ pF and $R = 100$ MΩ. The crossover $1/(2\pi R C_{22})$ is 199 Hz. At this frequency the quadrature capacitance is halfway between 12.0 and 10.875 pF, namely 11.4375 pF. The corresponding loss conductance divided by ω is 0.5625 pF. An independent verification uses the port voltage ratio at crossover,

$$|V_2/V_1| = |C_{21}|/(\sqrt{2}C_{22}) = 0.265.$$

Compute dissipated power as $|V_2|^2/(2R)$ for peak phasor amplitudes and compare it with $\mathrm{Re}\, Y_{\mathrm{in}} |V_1|^2/2$; the two results must coincide.

The crossover shifts inversely with the leakage resistance while its two limiting capacitances stay fixed in this simple model. Measuring several shunt resistances or adding a known external one provides a direct way to test this interpretation. If a neighbouring shield is connected to a voltage follower instead, its response includes the follower's own gain and phase; static floating-node elimination is insufficient. Extending the field-derived capacitance matrix by experimentally known terminal circuits makes it possible to distinguish changes in electrostatic geometry from changes in what the surrounding measurement wiring enforces.

Close the charge balance at a floating conductor

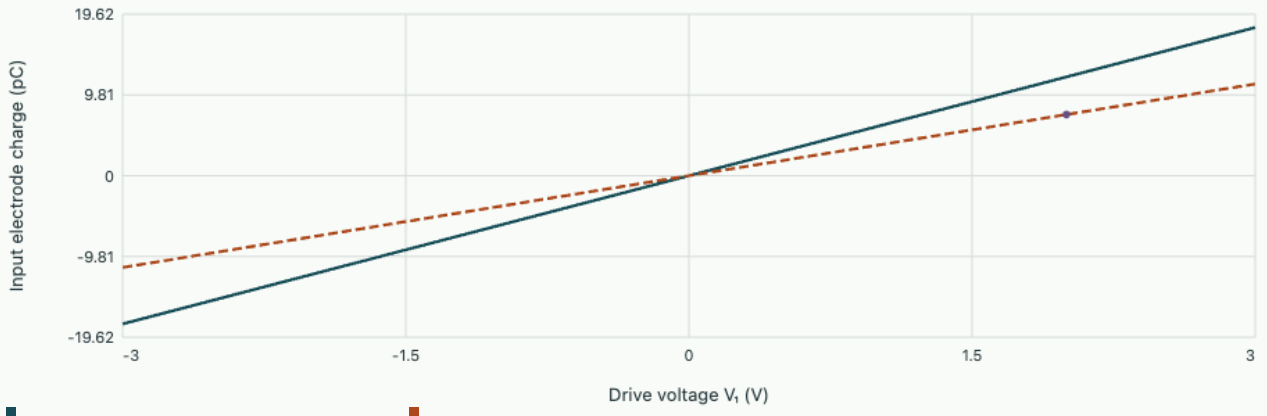
Compare voltages obtained by fixing a conductor potential and by prescribing its net charge. Check the sign of the off-diagonal capacitance terms.

Open this experiment online

Try the calculation

Mutual capacitance (pF) 4

Drive voltage V_1 (V) 2



Second electrode grounded Second electrode neutral and floating

Selected floating charge on electrode 1

Neutral electrode 2 floats to 1.143 V. Grounded input capacitance 6 pF, floating input capacitance 3.714 pF. Floating-case field energy 7.429 pJ.

Tutorial · distinguish a floating port and an inadmissible trial

What changes after disconnecting a biased guard?

Using the 12, 8 and –3 pF coefficients, start with terminal 1 at 0 V and terminal 2 at 1 V. Disconnect terminal 2, then raise terminal 1 to 2 V. Find the new floating voltage, both charges and the change of active charge.

▼ **Work through the solution**

Initially $Q_2 = 8$ pC, which stays fixed. The new condition is $8 = -3(2) + 8V_2$, giving $V_2 = 1.75$ V. Terminal 1 initially held -3 pC; finally it holds $12(2) - 3(1.75) = 18.75$ pC. Its change is 21.75 pC, or 10.875 pF times the 2 V increment. This reproduces the neutral-floating differential result but a different absolute charge. A charge amplifier reset at the initial state observes the change, whereas a device measuring absolute node potential distinguishes the stored preparation.

Can a lower voltage-trial estimate be trusted?

A potential approximation for the two-phase cell returns 2.70 pF while a constant-displacement trial gives 2.83 pF. The implementer reports the smaller number because its potential looked smoother. Diagnose the inconsistent pair.

▼ **Work through the solution**

For admissible fields on the same problem the flux trial gives a lower capacitance bound and the potential trial an upper one. The reversed interval cannot result from arbitrary coarseness alone. Check that the voltage trial satisfies the exact terminal difference and that all physical material volumes and energy factors are included. Check that the displacement trial has the correct area-normalised total charge and satisfies flux continuity. Finally examine numerical quadrature and unit conversion. Comparing the plots of the two fields does not fix an inadmissible constraint. If both are corrected, their gap gives an actual a posteriori uncertainty interval even without an analytic solution.

Further reading

- Measure a variational energy gap
- Change the geometry and permittivity of a layered capacitor
- Griffiths, 5th ed., §§2.4–2.5 and 4.4 · capacitance and field energy

Electrostatic field theory and measurement · 10. Anisotropic and nonlinear dielectrics from local measurements to global fields

A single dielectric constant cannot describe a crystal whose displacement tilts away from its electric field, or a reversible material whose incremental response changes with bias. Starting with the constitutive measurement clarifies which parts of ordinary electrostatics survive.

Measure a tensor rather than assuming collinearity

In a homogeneous linear anisotropic material,

$$\mathbf{D} = \boldsymbol{\epsilon}\mathbf{E}.$$

The entry ϵ_{ij} measures displacement component i in response to electric-field component j. A reversible reciprocal dielectric has a symmetric positive-definite tensor. Its principal axes are orthogonal and its positive eigenvalues are the principal permittivities. Field-energy density $u = \mathbf{E} \cdot \boldsymbol{\epsilon}\mathbf{E}/2$ is then positive for a nonzero field. Positive definiteness concerns all directions; having three positive diagonal entries in one basis alone is insufficient.

Choose principal relative permittivities $2, 4, 6$ along x, y, z. With a field $(1000, 0, 1000)$ V/m the displacement is $\epsilon_0(2000, 0, 6000)$. The field makes 45° to z, the displacement only 18.4°; their directions differ. Energy density is 3.54×10^{-5} J/m³. A flat free-charge sensor normal to z responds to D_z; an electrostatic force probe constrained along x measures E_x. Their readings cannot be interpreted as two component measurements of one vector differing by a single scalar.

Derive the field of a point charge in an anisotropic bulk

For a constant symmetric tensor write $\boldsymbol{\epsilon} = \epsilon_0 K$, with positive dimensionless matrix K. Poisson's equation is

$$-\epsilon_0 \nabla_{\mathbf{x}} \cdot K \nabla_{\mathbf{x}} V = q\delta^3(\mathbf{x}).$$

Change coordinates to $\mathbf{y} = K^{-1/2}\mathbf{x}$. The gradient transforms by the same symmetric matrix and the operator becomes the ordinary isotropic Laplacian in $\mathbf{y}$. The source transforms too: physical volume is $d^3x = \sqrt{\det K}\, d^3y$, so

$$\delta^3(\mathbf{x}) = \delta^3(\mathbf{y})/\sqrt{\det K}.$$

The decaying solution is

$$V(\mathbf{x}) = \frac{q}{4\pi\epsilon_0\sqrt{\det K}} \frac{1}{\sqrt{\mathbf{x}^T K^{-1}\mathbf{x}}},$$

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0\sqrt{\det K}} \frac{K^{-1}\mathbf{x}}{(\mathbf{x}^T K^{-1}\mathbf{x})^{3/2}}.$$

A check is the scalar case $K = kI$. The determinant contributes $k^{3/2}$ and the distance contributes $1/\sqrt{k}$, returning $q/(4\pi\epsilon_0 kr)$. The equipotentials in laboratory coordinates are ellipsoids, while $\mathbf{D} = \epsilon_0 K\mathbf{E}$ happens to point along $\mathbf{x}$ for this infinite homogeneous point-source problem. Its magnitude varies with direction on a physical sphere. Gauss's law still gives total free flux q; equating that flux to $4\pi r^2$ times a constant magnitude would be invalid. Conducting or material boundaries also transform under this coordinate map and generally cease to be simple physical spheres.

Along any principal axis i, at distance r, the potential is proportional to $\sqrt{K_i}/\sqrt{\det K}$. With principal values $2, 4, 6$, two probes equally far away on the z, x axes measure a voltage ratio $\sqrt{6/2} = \sqrt{3}$. One might have expected the largest permittivity to reduce the voltage most strongly along that axis. The observed result reflects how the full anisotropic medium distributes the displacement of a point source; it cannot be derived by inserting a direction-specific permittivity into the isotropic radial formula.

Which permittivity does a slab capacitance measure?

Place broad metal plates normal to a unit vector $\mathbf{n}$ and neglect edge effects. Constant electrode potentials force a purely normal electric field $V\mathbf{n}/d$, independent of the off-diagonal dielectric entries. The displacement generally has a tangential component, but the metal's free-charge reading is D_n. Thus the measured normal capacitance per area is

$$C/A = (\mathbf{n}^T \boldsymbol{\epsilon} \mathbf{n})/d.$$

For axes with relative values 2 and 6, cut the normal at 45° between them. The relative plate permittivity is 4.

A different preparation constrains displacement to be purely normal, for instance in an effectively uniform bulk state with known flux vector. Then $\mathbf{E} = \boldsymbol{\epsilon}^{-1} D_n \mathbf{n}$, so

$$D_n/E_n = 1/(\mathbf{n}^T \boldsymbol{\epsilon}^{-1} \mathbf{n}).$$

For the same cut the result is $1/[(1/2 + 1/6)/2] = 3$ in relative units. Both are correct for their own field constraints. Flat equipotential plates do not generally enforce tangential displacement equal to zero. A boundary-value calculation using that extra condition could report the wrong dielectric constant from an otherwise perfect experiment.

With an interface between two tensors, integrate Faraday's law around a small static rectangular loop to enforce continuous tangential E and Gauss's law over a pillbox to enforce the specified normal D jump. Write the first material's measured tangential field into the second's constitutive law and solve for its normal electric field from D_n. No scalar Snell-like rule involving only a ratio of two relative permittivities can represent arbitrary crystal orientations. Some field components may vanish by symmetry in one experiment and reappear after rotating the sample.

Reversible response distinguishes stored energy and coenergy

For a slowly and reversibly assembled homogeneous dielectric, increment of stored electrical energy per volume is $du = \mathbf{E} \cdot d\mathbf{D}$. A complementary density $w(\mathbf{E}) = \mathbf{E} \cdot \mathbf{D} - u$ has $dw = \mathbf{D} \cdot d\mathbf{E}$. In a scalar nonlinear law $D(E) = \epsilon E + \alpha E^3$ with $\epsilon, \alpha > 0$, integration gives

$$w(E) = \tfrac{1}{2}\epsilon E^2 + \tfrac{1}{4}\alpha E^4,$$

$$u(E) = ED(E) - w(E) = \tfrac{1}{2}\epsilon E^2 + \tfrac{3}{4}\alpha E^4.$$

The familiar $ED/2$ holds only in the linear case. For a small perturbation superposed on bias E_b, $dD/dE = \epsilon + 3\alpha E_b^2$ is the differential permittivity. A dc charge-to-voltage ratio measures the secant value $\epsilon + \alpha E_b^2$. At $\alpha E_b^2 = 0.20\epsilon$ these are 1.60ϵ and 1.20ϵ, a substantial difference even though the constitutive curve remains smooth.

The field equation becomes $\nabla \cdot \mathbf{D}(-\nabla V) = \rho_f$ and superposition generally fails. For a differentiable energy-based vector law, $\partial D_i/\partial E_j = \partial^2 w/\partial E_i \partial E_j$ is symmetric. A positive-definite differential tensor gives local stability and helps preserve ellipticity of the electrostatic boundary problem. A hysteretic material can depend on its prior history and need extra state variables; an area in a closed constitutive cycle then measures dissipated work, not reversible recoverable energy.

A nonlinear field calculation that needs only a scalar root

Place a sphere with charge q in an infinite isotropic nonlinear medium of the preceding type. Spherical symmetry and Gauss's law still determine $D_r = q/(4\pi r^2)$ outside. Invert $\epsilon E_r + \alpha E_r^3 = D_r$ point by point. For $\alpha > 0$ the derivative is positive everywhere, ensuring one physical real root for each D_r. A bisection bracket for $q > 0$ is $0 \le E_r \le D_r/\epsilon$; evaluating the residual at the upper endpoint gives positive cubic surplus. Compute voltage relative to infinity by integrating this field, rather than replacing D/E by an arbitrary constant taken at the surface.

For weak nonlinearity set $E = D/\epsilon + \delta$ and keep the leading cubic correction. One obtains

$$E \simeq D/\epsilon - \alpha D^3/\epsilon^4.$$

At large radii a $1/r^2$ displacement therefore acquires a negative $1/r^6$ correction in the electric field, giving a $1/r^5$ potential correction. A probe can misinterpret that extra radial scale as a change of source shape if material nonlinearity is not independently controlled. Double the source charge and compare the field after division by q. In a linear medium those curves coincide; the perturbative deviation here scales as q^2.

Resolve an oblique crystal interface component by component

A lateral detail or an oblique electrode can produce a field with both normal and tangential components at an interface. Consider a locally planar interface with tangent $\hat{\mathbf{x}}$ and normal $\hat{\mathbf{z}}$ from isotropic medium 1 into a crystal. In the xz plane let the crystal's relative tensor be $\begin{pmatrix} 4 & 1 \\ 1 & 3 \end{pmatrix}$, with entries in the order x, z. Medium

1 has relative permittivity 2 and measured field $E_{1x} = 120$ V/m, $E_{1z} = 240$ V/m. There is no free sheet at the interface. Neither input field is assumed normal to the boundary.

From a very short rectangular loop crossing the interface, the tangential electrostatic field is continuous, so $E_{2x} = 120$ V/m. The normal displacement on side 1 is $D_{1z}/\epsilon_0 = 2(240) = 480$ V/m. Equating it to $E_{2x} + 3E_{2z}$ yields $E_{2z} = 120$ V/m. The tangential displacement can jump. In the crystal it is $D_{2x}/\epsilon_0 = 4(120) + 120 = 600$ V/m, compared with $D_{1x}/\epsilon_0 = 240$ V/m. Neither tangent D nor normal E is required to be continuous. A field solver that enforces continuity of all components can silently erase the very anisotropy the experiment was intended to measure.

Find the bound surface charge using two routes. The true charge jump is $\epsilon_0(E_{2z} - E_{1z}) = -120\epsilon_0 = -1.06$ nC/m². With no free sheet this must be bound. Directly, $P_{1z}/\epsilon_0 = D_{1z}/\epsilon_0 - E_{1z} = 240$ V/m and $P_{2z}/\epsilon_0 = 480 - 120 = 360$ V/m. The net bound charge from the two facing polarisation surfaces is $P_{1z} - P_{2z} = -120\epsilon_0$, giving the same result. At a rotated interface it can be safer to perform these component checks than to infer charge from one visual change in field-line angle.

A controlled reversal can expose the off-diagonal tensor entry. Keep E_{1z} fixed and reverse E_{1x} using independent side electrodes. Then $E_{2x} = -120$ V/m and $E_{2z} = (480 + 120)/3 = 200$ V/m. The normal response therefore changes by 80 V/m even though the imposed normal displacement is unchanged. A scalar model, or a crystal whose tensor had zero xz entry in this coordinate system, would predict no such change. In practice the side electrodes must be far enough from the chosen interface patch for their intended approximately uniform fields to be realised locally.

Derive pressure for a reversible nonlinear capacitor under both electrical controls

Take a parallel-plate gap d, area A, completely filled by a reversible material with specified constitutive law $D(E)$. For this idealised calculation assume that changing d changes only the volume occupied by otherwise identical material at fixed thermodynamic properties. Exclude any additional mechanical elastic stress, electrostriction or chemical work needed to preserve that preparation. Keep the field homogeneous and neglect fringing. Under charge control the displacement $D = Q/A$ is fixed, so the stored electric energy is $U = Ad\,u(D)$ with $u = \int_0^D E(D')\,dD'$. The force tending to increase the gap is

$$F = -\partial_d U|_Q = -Au(D).$$

Thus the attractive field contribution has magnitude Au.

Now hold voltage fixed. The electric field is $E = V/d$; decreasing the gap also changes the charge the source must supply. Legendre transforming the energy by subtracting VQ gives the appropriate potential $\mathcal{G} = U - VQ = -Ad\,w(E)$. Its derivative at fixed V is $-Aw(E) + AE\,dw/dE = -Aw + AED = Au$. Hence $F = -\partial_d \mathcal{G}|_V = -Au$, agreeing for the same instantaneous state. Differentiating U at fixed voltage without the source term would produce a different answer. The calculation identifies the source work excluded by that incorrect derivative.

For $D = \epsilon E + \alpha E^3$ with $\alpha E_b^2 = 0.20\epsilon$,

$$u(E_b) = 0.65\epsilon E_b^2.$$

The often-used linear formula $E_b D_b/2$ would give $0.60\epsilon E_b^2$, underestimating the stated ideal field contribution. With $\epsilon = 4\epsilon_0,\ E_b = 1.0 \times 10^5$ V/m, the predicted magnitude per area is 0.230 Pa. This number follows from a particular reversible constitutive model and mechanical variation. A real force between electrodes compressing a solid sample also contains material mechanical stresses, and the permittivity itself can change with density. The correct experiment must specify which matter moves when the gap changes.

An independent numerical test computes $U(d \pm \delta d, Q)$ at fixed charge or $\mathcal{G}(d \pm \delta d, V)$ at fixed voltage, forms centred derivatives and chooses the charges or voltages that describe the same midpoint state. If the two forces converge to opposite signs, check the mechanical displacement convention and the source-work term. Agreement of the pressures at a single bias does not prove the constitutive approximation is valid at other fields. A separate differential capacitance scan can verify the slope required by the energy model across the displacement range.

Solve a layered nonlinear capacitor by displacement continuation

A useful intermediate problem between a uniform nonlinear slab and a full finite-element solve is a pair of different nonlinear layers stacked normal to large plates. With no free interface charge, the same normal displacement D passes through both. Each has a reversible monotone law $D_i(E_i) = \epsilon_i E_i + \alpha_i E_i^3$, $\epsilon_i > 0, \alpha_i \geq 0$. At fixed terminal charge $D = Q/A$ is known. Invert each material separately and add the voltages,

$$V(D) = d_1 E_1(D) + d_2 E_2(D).$$

At fixed voltage solve this last scalar equation for D. Its derivative is positive:

$$\frac{dV}{dD} = \frac{d_1}{\epsilon_1 + 3\alpha_1 E_1^2} + \frac{d_2}{\epsilon_2 + 3\alpha_2 E_2^2},$$

$$\frac{C_{\mathrm{diff}}}{A} = \left(\frac{dV}{dD}\right)^{-1}.$$

Thus it is the differential permittivities, evaluated at the two actual and generally different local fields, that combine as a series small-signal capacitance. Using a differential response measured at the overall average field can give an inconsistent result even when the series formula itself is correct. The secant ratio D/V has another interpretation and should be compared only with a dc charge reading.

Choose a field scale $E_* = 1000$ V/m and $D_* = \epsilon_0 E_*$. In the first material let $D/D_* = 2e_1 + e_1^3$, $e_1 = E_1/E_*$, while the second is linear $D/D_* = 6e_2$. Give both layers thickness 1 mm. At $D = 3D_*$, the first equation is solved by $e_1 = 1$; the second gives $e_2 = 0.50$. The total voltage is 1.50 V, and $D = 26.56$ nC/m². The dc capacitance per area is 17.71 nF/m². At this same operating point the differential relative permittivities are 5 and 6, giving $C_{\mathrm{diff}}/A = \epsilon_0/[0.001/5 + 0.001/6] = 24.15$ nF/m². A small voltage ripple measures a capacitance about 36% larger than the secant value.

Check the bias state from energy without using terminal current. The stored energy per volume in layer 1 is $\epsilon_0 E_*^2[e_1^2 + 3e_1^4/4]$ and in layer 2 is $3\epsilon_0 E_*^2 e_2^2$. Multiplying by the respective thicknesses at the chosen state gives total energy per area $0.00250\epsilon_0 E_*^2 = 22.14$ nJ/m². Integrating the voltage as a function of displacement from zero to $3D_*$ must give the same result, since $d(U/A) = V\,dD$. The shortcut $VD/2$ would instead give 19.92 nJ/m². It lies between neither two guaranteed bounds nor two voltage controls; it is simply unjustified for this nonlinear path.

For voltage-controlled computation use a continuation in applied voltage. Start at low bias with the known linear two-layer solution and reuse its displacement as a first guess for the next monotone scalar solve. A bracketing method with increasing positive endpoints always finds the root for positive applied voltage in these specific laws. Once converged, evaluate both layers' displacement residuals, their voltage sum and the independent energy integral. A lower iteration error alone cannot correct a model that accidentally multiplied the cubic coefficient by field rather than field cubed, or used a common E across both materials. Direct field checks at a selected bias make the numerical experiment physically interpretable.

Rotate the field relative to a microstructure

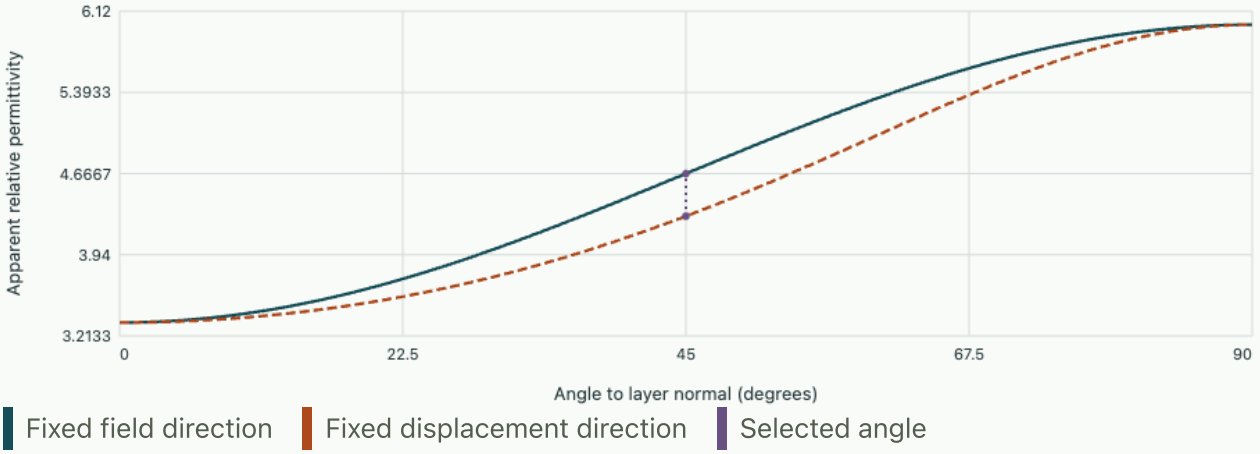
Study a layered material with calculable normal and tangential responses. Change its composition before deciding whether one scalar permittivity can describe both measurements.

Open this experiment online

Try the calculation

Fraction of low-permittivity phase **0.5**

Measurement angle (degrees) **45**



Principal relative permittivities: normal 3.3333, tangential 6. At 45° the two responses are 4.6667 and 4.2857.

Recover dielectric alignment torque from electrode work

Will positive or negative rotation increase capacitance?

A slab completely filling an ideal parallel-plate cell can rotate its uniaxial permittivity tensor reversibly without changing plate geometry. The angle θ between plate normal and the larger principal-permittivity axis is adjustable. Derive the torque on this ideal material coordinate, compare voltage and charge control, and calculate the maximum for a stated geometry.

▼ **Derive and check the result**

With area A, separation d, principal absolute permittivities $\epsilon_{\parallel} > \epsilon_{\perp}$, the normal capacitance is $C(\theta) = (A/d)(\epsilon_{\parallel}\cos^2\theta + \epsilon_{\perp}\sin^2\theta)$. For fixed voltage electrical coenergy is $CV^2/2$. The torque doing work in the direction of increasing θ is its derivative,

$$T = (A/d)(\epsilon_{\perp} - \epsilon_{\parallel})V^2\sin(2\theta)/2.$$

For $0 < \theta < \pi/2$ it drives the larger-permittivity principal axis towards the plate normal, increasing capacitance. The configurations at 0° and 90° both have zero torque but only the former minimises the fixed-voltage mechanical potential $-CV^2/2$.

Under fixed charge the stored energy is $Q^2/(2C)$, so the torque is $-\partial_\theta U = Q^2C'/(2C^2)$. At the same instantaneous voltage Q/C it agrees with the voltage-controlled result. Following two entire rotation curves with different preparations generally yields different magnitudes because the isolated cell voltage varies with angle. This equality at the same state is an independent sign check and distinguishes the field torque from the external holding torque of opposite sign.

Take $A = 1.0$ cm², $d = 1.0$ mm, principal relative permittivities 6 and 2, and $V = 20$ V. At 45°, the magnitude is $0.5(A/d)4\epsilon_0V^2 = 7.08\times10^{-10}$ N m. A small-angle torque measurement close to the aligned state would also reveal a positive angular restoring stiffness. In a real sample the crystal lattice, liquid-crystal elasticity or mount supplies additional energy and boundary conditions for rotation. The electrostatic contribution just calculated can be isolated by varying V^2 and checking the $\sin 2\theta$ angular dependence at sufficiently slow drive.

Tutorial · diagnose a tensor and a bias dependence

Recover a rotated tensor from three plate measurements

In a material plane, cuts with normals $x, y, (x+y)/\sqrt{2}$ yield relative plate permittivities $3, 5, 5$, respectively. Reconstruct the 2×2 tensor, check positivity and predict the result for the other bisector.

▼ **Work through the solution**

The first two values are

$$K_{xx} = 3, K_{yy} = 5.$$

The third is $(3 + 5)/2 + K_{xy} = 5$, so $K_{xy} = 1$. The symmetric matrix has determinant 14 and positive first diagonal entry, hence is positive definite. Its principal values are $4 \pm \sqrt{2}$. The other bisector gives $4 - 1 = 3$. This extra reading can test the assumed orientation and symmetry independently. All four inferences assume the plate experiment imposes electric field normal to each cut and fringing is negligible.

Distinguish charge measurements before integrating electrical work

At a bias E_b, a sample following the cubic constitutive law gives secant permittivity $1.30\epsilon_*$ and differential permittivity $1.70\epsilon_*$, where ϵ_* is a calibration permittivity. Infer its linear coefficient, cubic contribution at bias and true stored energy density.

▼ **Work through the solution**

Subtract the two measurements. The difference is $2\alpha E_b^2 = 0.40\epsilon_*$, so $\alpha E_b^2 = 0.20\epsilon_*$,

$$\epsilon = 1.10\epsilon_*.$$

The energy at bias is

$$[0.55 + 0.15]\epsilon_* E_b^2 = 0.70\epsilon_* E_b^2.$$

Using $ED/2$ would give $0.65\epsilon_* E_b^2$. To test the law rather than just fit two parameters, obtain a second set of secant and differential readings at a different bias and check the predicted quadratic change with bias field.

Further reading

- Inspect normal displacement across a dielectric interface
- Follow the optical consequences of a permittivity tensor
- Griffiths, 5th ed., §§4.3–4.4 · constitutive laws and dielectric work

Electrostatic field theory and measurement · 11. From mobile ions to screening lengths and differential double-layer capacitance

A charged planar wall in an electrolyte can be studied without assuming a capacitance in advance. Diffusion and electrostatic drift first establish an ion cloud. The charge–voltage slope of that cloud is an experimental output of the model.

An electric field competes with entropy of mixing

Take a bulk reservoir containing mobile ions with charges $+e$, $-e$, each at number density n_b. The solvent is a uniform continuum of permittivity ϵ, the bath holds absolute temperature T, and no reactions or Faradaic currents occur at the wall. At thermal equilibrium the electrochemical potential of each species is spatially constant. For an ideal dilute gas of particles it contains $k_B T \ln(n_\pm / n_b)$ from mixing plus electrostatic energy $\pm eV$. Set the bulk potential to zero. Equating the local and bulk electrochemical potentials yields

$$n_\pm = n_b e^{\mp eV/k_B T}.$$

At positive V the negative counterions concentrate and the positive coions are depleted. Total mobile charge is

$$\rho = e(n_+ - n_-) = -2en_b \sinh(eV/k_B T).$$

Substituting into Poisson's equation and defining dimensionless $\psi = eV/(k_B T)$ gives

$$\nabla^2 \psi = \kappa^2 \sinh \psi,$$

$$\kappa^2 = \frac{2e^2 n_b}{\epsilon k_B T}.$$

The parameter n_b is a density of one species, not both together. For a 1:1 salt specified in mol/litre convert to formula units per cubic metre using $1000 N_A$ before substitution. For a mixture the linear screening coefficient generalises to $e^2 \sum_i z_i^2 n_i / (\epsilon k_B T)$, whereas the full nonlinear right-hand side remains a sum of exponentials. An externally maintained reservoir fixes chemical potentials; a sealed very small cell can instead require separate number conservation for every species.

What the small-potential limit predicts and what it neglects

For $|\psi| \ll 1$, replace $\sinh \psi$ by ψ. Above a planar wall at $z = 0$ with prescribed potential V_s and $V \to 0$ far away, the bounded solution is

$$V(z) = V_s e^{-\kappa z}.$$

With the normal into liquid $+\hat{\mathbf{z}}$, the wall displacement is

$$D_z(0^+) = -\epsilon V'(0) = \epsilon \kappa V_s.$$

For a conducting wall with no interior field this equals its free surface charge. Therefore $C_{\mathrm{diff}}/A = d\sigma/dV_s = \epsilon\kappa$, dimensionally a dielectric layer of thickness κ^{-1}. The ionic concentration actually varies continuously; it is not a plate of opposite charge located at that distance.

Integrating charge density $-\epsilon\kappa^2 V(z)$ over $z > 0$ gives

$$-\epsilon\kappa V_s = -\sigma.$$

The screening cloud and wall together are neutral per area. The fraction of compensating charge within distance h is $1 - e^{-\kappa h}$. At one screening length the potential has fallen to $1/e$ and 63.2% of the countercharge has accumulated. For a spherical bead curvature changes this simple relationship between potential ratio and enclosed charge. The ionic constitutive relation is the same, but the spatial Laplacian and collecting area differ.

With $\epsilon_r = 78.5,\ T = 298$ K and 1.00 mM of a fully dissociated symmetric monovalent salt, $n_b \simeq 6.02 \times 10^{23}$ m⁻³ and $\kappa^{-1} \simeq 9.62$ nm. The linear diffuse capacitance is about 0.072 F/m². Increasing concentration fourfold halves the screening length and doubles this capacitance at fixed solvent response. Close to molecular length scales the continuum assumptions and concentration-independent permittivity become progressively questionable, so the square-root law should be tested in the intended concentration range.

Background · How small is a small electrostatic potential?

In a symmetric electrolyte the mean charge term is proportional to $\sinh u$, with $u = ze\psi/(k_B T)$ measured relative to the neutral reservoir. A potential being small in volts gives no absolute criterion: the valence and temperature set the relevant scale. Taylor expansion yields the fractional error in replacing the ionic source by its linearised value,

$$\frac{\sinh u - u}{u} = \frac{u^2}{6} + O(u^4).$$

For roughly 1% local source accuracy this estimate asks for $|u| \lesssim 0.245$. At 298 K it gives about 6.3 mV for monovalent ions and 3.1 mV for divalent ions. This is a criterion on the local differential equation. An integrated charge or differential capacitance can have a different relative error and merits its own check.

Solve the nonlinear planar boundary problem

With planar symmetry $\psi'' = \kappa^2 \sinh\psi$. Multiply both sides by ψ' and integrate from the remote bulk where $\psi, \psi' \to 0$. This gives

$$\tfrac{1}{2}(\psi')^2 = \kappa^2(\cosh\psi - 1) = 2\kappa^2 \sinh^2(\psi/2).$$

Choose the branch $\psi' = -2\kappa \sinh(\psi/2)$, which decays to zero for either sign of surface potential. The identity $d\ln|\tanh(\psi/4)|/d\psi = [2\sinh(\psi/2)]^{-1}$ turns the separated equation into

$$\tanh\frac{\psi(z)}{4} = \tanh\frac{\psi_s}{4}\, e^{-\kappa z}.$$

The physical potential itself is not a simple exponential at finite bias. Its distant tail has the same decay length but extrapolates back to a different amplitude $4\tanh(\psi_s/4)$ in dimensionless units. A far-field slope can therefore recover κ, while a linear extrapolation of its intercept underestimates large surface potentials. The nonlinear solution reduces continuously to the linear one because $\tanh x \sim x$.

At $z = 0$, using $D_z = -\epsilon k_B T \psi'/e$ gives

$$\sigma(V_s) = \frac{2\epsilon\kappa k_B T}{e}\sinh\left(\frac{eV_s}{2k_B T}\right),$$

$$\frac{C_{\text{diff}}}{A} = \epsilon\kappa\cosh\left(\frac{eV_s}{2k_B T}\right).$$

The first relation is odd in surface voltage and the differential capacitance is even for this symmetric salt model. Increasing voltage concentrates counterions and allows more charge per further voltage increment. For $V_s = 50$ mV at 298 K, $V_s/(2k_BT/e) \simeq 0.973$. The differential capacitance is about 1.51 times its zero-bias value, while the secant ratio σ/V_s is only about 1.17 times that value. A small AC perturbation superposed slowly about a DC bias measures the differential slope if ionic transport remains equilibrated.

An independent mechanical check of the nonlinear field

The ideal mobile ions have an osmotic pressure excess relative to the reservoir $\Delta p = k_B T(n_+ + n_- - 2n_b) = 2n_b k_B T(\cosh\psi - 1)$. From the first integral the electrical normal stress magnitude $\epsilon E^2/2$ equals this excess. Thus $\Delta p - \epsilon E^2/2$ is spatially constant and vanishes in the bulk of the isolated-wall problem. One can derive the same result by differentiating the pressure: $d\Delta p/dz = \rho E$ at equilibrium. The cancellation provides a sensitive joint test of the ion distribution and the potential gradient.

Confine electrolyte between two equal positively biased parallel walls instead. By symmetry $E = 0$ at the midplane but ψ_m generally remains positive. The constant excess normal pressure is then $2n_b k_B T(\cosh\psi_m - 1) > 0$. Overlap of the diffuse layers produces a repulsive osmotic force per area in this mean-field model. With large separation the midplane potential tends to zero and so does that force. A solution built by naively superposing two high-voltage nonlinear wall profiles will generally violate the constant-pressure test.

Separate a near-wall compact layer from the diffuse response

A finite exclusion distance for ion centres near a real electrode can be modelled at first by a compact dielectric layer of capacitance per area c_s. Its voltage drop is σ/c_s; the diffuse layer starts at potential V_d. For applied bias measured relative to the bulk and a chosen neutral reference,

$$V_{\text{app}} = V_d + \sigma(V_d)/c_s.$$

Differentiate with respect to surface charge:

$$\frac{1}{c_{\mathrm{meas}}} = \frac{1}{c_s} + \frac{1}{c_d(V_d)}.$$

The diffuse constitutive relation must use V_d, not the entire applied voltage. For $c_s = 0.20$ F/m² and low-bias $c_d = 0.072$ F/m², the measured capacitance is about 0.053 F/m². Even as c_d grows, this simple series model cannot exceed c_s. It offers a physically distinct explanation for a measured saturation, alongside finite ion size, adsorption, dielectric variation and other model extensions. A static capacitance measurement alone cannot uniquely distinguish all of them.

Use layer overlap to infer a mechanical force independently of capacitance

Consider identical parallel electrodes at $z = \pm H$, held at the same small positive potential V_s relative to a salt reservoir in chemical equilibrium with the gap. Neither electrode is an ordinary reference for the other: a voltage meter connected only between them reads zero while each diffuse layer experiences a nonzero potential relative to the reservoir. Under the linear screening approximation the potential obeys $V'' = \kappa^2 V$. Reflection symmetry enforces $V'(0) = 0$. Matching both wall potentials gives $V(z) = V_s \cosh(\kappa z)/\cosh(\kappa H)$, so the midplane voltage is

$$V_m = V_s/\cosh(\kappa H).$$

The magnitude of free charge per area on each inner electrode face is $\epsilon\kappa V_s \tanh(\kappa H)$. At large H each plate recovers its isolated-wall charge. At small H the whole gap is nearly at V_s, so very little normal field is required to maintain it and each facing surface carries less charge. This behaviour differs from two oppositely biased plates, whose nearly uniform internal field grows as their separation decreases. The reservoir reference and voltage symmetry decide which experiment is being performed.

To leading nonzero order in V_s, the osmotic excess force per area between these identical walls can be found at the field-free midplane. There it is $2n_b k_B T(\cosh\psi_m - 1) \simeq n_b k_B T \psi_m^2 = \epsilon\kappa^2 V_s^2/[2\cosh^2(\kappa H)]$. Why use a term quadratic in V_s when the potential was found linearly? The pressure's zeroth-order excess vanishes and the first-order ionic density changes cancel between opposite species. Its first measurable correction is quadratic. The cubic error in the linear potential changes the computed pressure only at fourth order in V_s, making the mixed expansion consistent.

For $\kappa^{-1} = 10$ nm, $\epsilon = 78.5\epsilon_0$ and $V_s = 5.0$ mV, at half-gap $H = 10$ nm the predicted repulsive pressure is about 36.5 Pa. At $H = 20$ nm it drops to roughly 6.14 Pa. The potential rose by only a few millivolts yet its ionic osmotic consequence extends across the entire overlap region. As H becomes large, $\cosh^{-2}(\kappa H) \simeq 4e^{-2\kappa H}$, giving exponential decay with the full separation $2H$. This is a force measurement of the same screening scale measured independently by a voltage probe. Surface roughness, salt exclusion and adsorption can influence a real nanometre-scale force curve and should be distinguished from deviations in the bulk screening length.

If the electrode charges rather than their potentials are fixed, the wall derivative is prescribed. The even solution then has $V_m = \sigma/[\epsilon\kappa \sinh(\kappa H)]$ in the linear regime, and pressure $\sigma^2/[2\epsilon \sinh^2(\kappa H)]$. As H shrinks the linear approximation will eventually fail for fixed nonzero σ, because V_m grows. Comparing the constant-potential and constant-charge formulas without adjusting their boundary conditions could suggest contradictory predictions where each applies to a different experiment.

Estimate how quickly the ion cloud can follow a capacitance measurement

Static Poisson–Boltzmann theory assumes ions are in equilibrium at each measured voltage. A time-dependent measurement can test that assumption. For dilute symmetric monovalent ions with equal diffusivity D_i, small density perturbations obey the linearised number fluxes $\mathbf{j}_\pm = -D_i[\nabla \delta n_\pm \pm e n_b \nabla V/(k_B T)]$, combining diffusion with electric drift. Assume no bulk reactions. Subtract the two continuity equations and multiply by e, yielding a charge-current density $\mathbf{J} = -D_i[\nabla \rho + 2e^2 n_b \nabla V/(k_B T)]$. Use $\nabla^2 V = -\rho/\epsilon$ in the bulk to obtain

$$\partial_t \rho = D_i(\nabla^2 \rho - \kappa^2 \rho).$$

A spatial harmonic of wavenumber k decays at rate $D_i(k^2 + \kappa^2)$ in the unbounded linear bulk model. The long-wavelength Debye charge-relaxation time is $\tau_D = 1/(D_i \kappa^2)$, also recovered by dividing permittivity by the dilute ionic conductivity $\sigma_{\text{ion}} = 2e^2 n_b D_i/(k_B T)$. The total salt-density perturbation instead obeys ordinary diffusion because the leading positive and negative drift terms cancel. Fast local charge relaxation and slower redistribution of neutral salt are different observables.

For $D_i = 1.0 \times 10^{-9}$ m²/s and $\kappa^{-1} = 10$ nm, this local time is 100 ns, corresponding to an angular rollover scale $1/\tau_D = 10^7$ s⁻¹ or ordinary frequency about 1.6 MHz. Blocking electrodes in a cell of thickness $L \gg \kappa^{-1}$ can display a slower charging time. Estimate its series resistance $R = L/(\sigma_{\text{ion}} A)$ and diffuse capacitance $C \sim \epsilon\kappa A$ for one interface. Their product is of order $L/(D_i \kappa)$. For $L = 100$ µm with the preceding parameters this is about 1 ms, orders of magnitude above the local bulk time. Two identical interfaces in series supply an extra numerical factor; finite reservoirs and unequal mobilities can introduce further scales.

A low-frequency impedance experiment interpreted as the equilibrium $d\sigma/dV$ must therefore allow the actual circuit and cell to charge, not merely wait several local Debye times. Varying cell spacing helps distinguish a geometry-dependent charging time from an intrinsic screening-length change. This diffusion calculation and the static pressure calculation share the same κ, giving independent checks on a physical fit to voltage, force and time-resolved measurements.

Check reservoir and continuum assumptions using actual particle numbers

The grand-canonical ion profiles implicitly allow the diffuse layers to draw particles from a large reservoir without changing the reservoir's chemical potentials. Whether this is reasonable can be checked with a dimensioned inventory. For a 1 mM monovalent salt, each species has about 6.02×10^{23} particles per

cubic metre. A cell of area 1 mm² and depth 100 µm therefore starts with 6.02×10^{13} ions of each species. A surface charge density of 1.0 mC/m² on one plate amounts to 6.24×10^{9} elementary charges on that area. Even a conservative estimate in which each net countercharge requires one ion from the bulk uses only about 10^{-4} of the available counterions.

Keep the same area and salt concentration but reduce the gap to 20 nm. Now only 1.20×10^{10} ions of each species occupy the cell initially. The same surface charge is comparable with half the inventory of one species. The remote-neutral-reservoir approximation is no longer automatically credible, especially if both walls attract counterions of the same sign. One must specify how ions and compensating charge can enter or leave, what the electrodes' chemical or electrical preparations enforce, and how much fluid actually belongs to the reservoir. An applied electrostatic boundary voltage fixes no ionic chemical potential by itself.

For a sealed cell write the equilibrium densities instead as $n_{\pm}(z) = A_{\pm} \exp[\mp eV(z)/(k_B T)]$. The constants $A_{\pm}$ must be chosen so that integrals of $n_{\pm}$ over the physical volume equal the known numbers of each species. Substitution into Poisson's equation gives a coupled boundary problem with two additional global constraints. Adding a constant to V and appropriately rescaling $A_{\pm}$ leaves the densities unchanged; electrical reference and chemical normalisation must be chosen consistently. A calculation holding both fixed bulk densities and exact sealed inventories can be overconstrained if diffuse layers hold appreciable salt.

The continuum approximation also makes a statement about scale. A cube of side one screening length near 10 nm at 1 mM holds roughly 0.60 ions of either species on average. This does not by itself invalidate a mean density for a large planar electrode, which averages over many lateral regions and thermal configurations. It does warn against interpreting the smooth one-dimensional density as a snapshot of a single nanometre-sized volume with a deterministic fractional particle. Force or voltage measurements with different lateral averaging areas can show quite different fluctuations even when their mean profiles follow the same equation.

As another self-check compare the predicted contact concentration with available molecular packing. The ideal theory gives

$$n_{\mathrm{counter}}(0) = n_b e^{|\psi_s|}.$$

Starting from 1 mM, a surface potential near 0.25 V at 298 K would multiply concentration by roughly 1.7×10^{4}, corresponding to tens of molar even before allowing room for solvent. A molecular model or a measured finite-size constitutive relation is then more appropriate for quantitative contact density. The failure is in extrapolating the dilute ideal entropy, not in Gauss's law or the electrostatic boundary condition. Source, solvent, reservoir and detector scales determine where the otherwise elegant analytic solution becomes an approximation to test.

Resolve a screening length

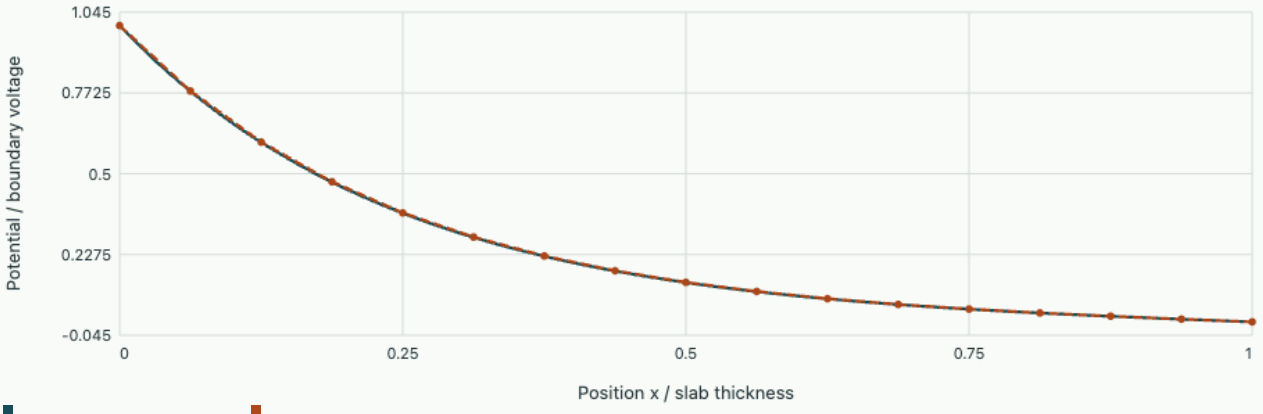
Compare the linear screened profile with a grid solution as the ratio of thickness to screening length grows. Separate failure to resolve the layer from failure of the linear material law.

Open this experiment online

Try the calculation

Cells across slab **16**

Gap / screening length **4**



Continuum slab Solved grid values

Midpoint 0.13357, exact 0.1329. Discrete decay constant × thickness = 3.9897.

Download the OneSimulation model

In your OneSimulation environment, run the model with these inputs.

```
python field-laboratory.py screen --cells 16 --contrast 4
```

How many excess ions, not just how much net charge?

Count the salt adsorbed in a planar diffuse layer

The total countercharge per area equals the negative surface charge. Separately calculate the excess number of ions of both signs over the neutral bulk value for a symmetric electrolyte at prescribed surface potential. Which quantity changes sign if the wall voltage reverses?

▼ **Derive and check the result**

Define $\Gamma = \int_0^\infty [n_+(z) + n_-(z) - 2n_b]\,dz = 2n_b \int_0^\infty (\cosh\psi - 1)dz$. The integrand is nonnegative for either sign of wall voltage. For $\psi_s > 0$, substitute $dz = -d\psi/[2\kappa \sinh(\psi/2)]$ from the first integral. Then $\cosh\psi - 1 = 2\sinh^2(\psi/2)$ makes the resulting integral elementary. The result, also valid at negative bias by symmetry, is $\Gamma = (4n_b/\kappa)[\cosh(\psi_s/2) - 1]$. It has units particles per area. The signed excess charge divided by elementary charge, in contrast, is

$$-\sigma/e = -(4n_b/\kappa)\sinh(\psi_s/2).$$

For positive surface charge the negative counterion excess and positive coion depletion follow by taking half the sum and difference of these two measurements. At small ψ_s, the net charge excess is linear in

voltage while Γ is quadratic: the two ion species redistribute in opposite directions before appreciable neutral salt adsorption appears. Reversing the bias swaps which species is enriched but leaves the total ion-number excess unchanged. Thus a mass-sensitive and a charge-sensitive measurement give complementary responses even in the same equilibrium model.

At $\psi_s = 2$,

$$\Gamma/[|\sigma|/e] = [\cosh 1 - 1]/\sinh 1 = \tanh(1/2) = 0.462.$$

The total excess count can be smaller than the magnitude of countercharge measured in elementary charges because some compensation occurs through removing coions rather than adding counterions. An optical concentration measurement that tracks only total dissolved ion count cannot be converted into net countercharge by multiplying by e. Sampling both quantities, or calibrating the signal against the appropriate species, is needed to test the adsorption prediction. The derivation assumes the same dilute reservoir as the planar Poisson–Boltzmann solution and should not be applied unmodified to a cell with fixed small ion inventories.

Tutorial · infer a screening length and check the model

Use two potential measurements at finite bias

In a 298 K symmetric electrolyte, $\tanh[eV(z)/(4k_BT)]$ decreases from 0.30 to 0.15 over 8.0 nm. Infer the screening length, and reconstruct the two physical voltages. Why is taking a raw voltage ratio slightly different?

▼ Work through the solution

The transformed profile decays exponentially, so

$$\kappa^{-1} = 8.0/\ln 2 = 11.54\,\text{nm}.$$

Since $4k_BT/e \simeq 102.7$ mV, the voltages are $102.7\,\text{artanh}(0.30) = 31.8$ mV and $102.7\,\text{artanh}(0.15) = 15.5$ mV. Their physical ratio is about 0.488 rather than 0.500. Inferring a linear-log decay using only these points gives a biased length. Repeating at smaller voltages would test convergence towards the same result without assuming the nonlinear transformation.

Which capacitance was measured?

At low bias a diffuse model independently predicts 0.080 F/m² but a low-frequency measurement gives 0.050 F/m². Infer the compact capacitance of the ideal series model. What would make this inversion impossible with positive capacitances?

▼ Work through the solution

From the reciprocal difference $1/c_s = 20 - 12.5 = 7.5$ m²/F we obtain $c_s = 0.133$ F/m². A measured capacitance larger than the predicted diffuse one would imply a negative reciprocal compact capacitance in this two-positive-series model. Check the normalisation area, frequency response, adsorption or additional parallel charging processes before assigning a negative parameter. To distinguish the two capacitances experimentally, vary reservoir concentration at fixed temperature and compare the predicted $\sqrt{n_b}$ trend of the diffuse contribution with the remaining offset in reciprocal capacitance.

Further reading

- Compare planar screening to a spherical bead
- Cambridge Electromagnetism, chapter 6 · screening and electrostatics of mobile charge

Electrostatic field theory and measurement · 12. Reconstructing a source from noisy voltage measurements

A forward electrostatic calculation can be unique and stable while the inverse problem constructed from it is unstable or nonunique. A scanned electrode, a finite collecting area and an explicitly chosen noise level make these statements experimentally testable.

Write down what the instrument actually measures

Pattern a long electrode at $z = 0$ with an unknown voltage $f(x)$, periodic over length L for this model. The region above is homogeneous, charge free, and contains a sufficiently weak probe at height h. For a Fourier component $f_m e^{ik_m x}$, $k_m = 2\pi m/L$, the bounded harmonic solution at height h is $f_m e^{-|k_m|h} e^{ik_m x}$. Treat the constant component according to the actual reference. Suppose the probe reports an average over a rectangular window of width w centred at x_j, plus noise e_j. For uniform probe sensitivity the forward relation is

$$y_j = \sum_m H_m f_m e^{ik_m x_j} + e_j,$$

$$H_m = e^{-|k_m|h}\,\mathrm{sinc}(k_m w/2),$$

$$\mathrm{sinc}\, t = \frac{\sin t}{t}.$$

The exponential is physical propagation from boundary to probe; the sinc factor is averaging by the detector. The latter has exact zeros in this idealised model. For modes at those zeros no noise-free reading at the same width and height can recover the lost amplitude. More closely spaced scan positions increase the sampling bandwidth but cannot undo either such instrumental nulls or the exponential decrease of the true high-frequency signal.

A nonuniform probe shape replaces the sinc by its Fourier transform. Unknown height adds an additional calibration problem: increasing h and increasing high-frequency boundary amplitudes can yield the same readings if the amplitudes are unrestricted. Before inferring boundary detail, calibrate the probe width and decide whether height is known, measured with uncertainty or jointly estimated using an independent constraint.

A direct inverse amplifies uncertainty in predictable bands

In an ideal complete periodic scan the discrete Fourier transform separates resolvable modes. Let $\hat{y}_m = H_m f_m + \eta_m$. Inverting one coefficient gives $f_m^{\mathrm{dir}} = \hat{y}_m / H_m$ whenever $H_m \neq 0$. A perturbation $\delta\hat{y}_m$ becomes

$$\delta f_m = \delta\hat{y}_m / H_m.$$

With $H_m = e^{-|k|h}$, the amplification grows exponentially with spatial frequency. The mode values can be uniquely determined in ideal noiseless data of infinite precision, yet the inverse is unbounded as increasingly fine modes are admitted. Finite data can additionally produce exact nonuniqueness.

For $h = 2.0$ mm and negligible aperture, boundary wavelengths 20, 10 and 4 mm have transfers $0.533, 0.285, 0.0432$, respectively. With 1 mV noise in the measured amplitude of each Fourier component, inverse noise levels are $1.87, 3.51, 23.1$ mV. A genuine 10 mV boundary ripple at wavelength 4 mm contributes only 0.432 mV at the probe, smaller than that noise. Plotting its raw inverted coefficient with a smooth line does not improve its information content. Reducing height to 1 mm raises that ripple's signal to 2.08 mV and changes the experiment's resolving power before any reconstruction choice is made.

Derive a regularised inverse and expose its bias

Assume coefficients are inferred separately and use a quadratic penalty. For each m, minimise $J_m = |H_m f - \hat{y}_m|^2 + \alpha^2 w_m^2 |f|^2$ with positive weights w_m and positive regularisation parameter α. Splitting real and imaginary parts, or differentiating with respect to f^*, gives

$$\widehat{f}_m = \frac{H_m^* \hat{y}_m}{|H_m|^2 + \alpha^2 w_m^2},$$

$$W_m = \frac{|H_m|^2}{|H_m|^2 + \alpha^2 w_m^2},$$

$$\mathbb{E}[\widehat{f}_m] = W_m f_m$$

for zero-mean measurement noise. The reconstruction suppresses noise near a small transfer by introducing deterministic bias $(W_m - 1) f_m$. Its variance for known mode-noise variance s_m^2 is $|H_m|^2 s_m^2 / (|H_m|^2 + \alpha^2 w_m^2)^2$. Neither low variance alone nor a small data residual is evidence that a spatial ripple was recovered. Setting α very large produces a smooth, nearly zero estimate regardless of the data. Setting it very small can reproduce noise almost exactly.

For $w_m = 1, \alpha = 0.10$ and wavelength 4 mm in the example above, $W_m \simeq 0.157$. A true 10 mV ripple is estimated at only 1.57 mV on average. The inversion gain is about 3.64 rather than 23.1. At 20 mm, the same penalty retains 96.6% of the true amplitude. If the purpose is to quantify a short ripple, improving the experiment can be preferable to changing the penalty. A derivative penalty $w_m \propto |k_m|$ encodes a stronger preference against fine structure; a constant mode must be handled separately if the voltage reference or the derivative penalty leaves it unconstrained.

What survives away from a uniform grid or simple geometry?

For arbitrary electrode shapes and detector positions, represent source or boundary data in a chosen basis $f = \sum_j a_j b_j$. Solve the forward field problem for each unit coefficient and sample it exactly as the real probe would, giving matrix entries A_{ij}. A noise covariance Σ weights the misfit through $\|\Sigma^{-1/2}(A\mathbf{a} - \mathbf{y})\|^2$. For an identity penalty apply the singular-value decomposition to the whitened matrix,

$$\Sigma^{-1/2} A = U S V^*.$$

Each right singular vector is a source pattern; its singular value tells how strongly it appears in the whitened measurement space. Multiplication by U^* diagonalises the data misfit and reduces the problem to the scalar filters above.

A zero singular value means the chosen sensors are insensitive to an entire source combination. No choice of numerical precision or regularisation changes that observational fact. A small nonzero value means one needs sufficiently small noise and modelling errors to identify the corresponding component. A practical resolution matrix $R = (A^*\Sigma^{-1}A + \alpha^2 I)^{-1}A^*\Sigma^{-1}A$ maps a noiseless true coefficient vector to the regularised estimate. A column of R shows which neighbouring coefficients respond when one ideal source coefficient is present. Testing such impulse responses describes the reconstruction's spatial smoothing more honestly than quoting the spacing of its output grid.

Use withheld measurements and identifiable physical quantities

Reserve some independently measured positions, heights or probe widths. Fit the rest and use the recovered candidates to predict the withheld measurements by applying the forward operator. The test is more informative when the withheld observations have different sensitivity patterns. A second distant scan can verify low modes while remaining insensitive to every high-frequency ambiguity. A second close scan or a different probe width can fill modes attenuated or nulled by the first. Correlated calibration errors, such as one common voltage offset, should be modelled as such, rather than being counted as many independent observations.

A distinct nonuniqueness occurs when trying to infer an arbitrary three-dimensional charge density solely from exterior voltages. Take any smooth compactly supported potential $g(\mathbf{x})$ inside the source region whose value and derivatives vanish outside. Add $\delta\rho = -\epsilon_0\nabla^2 g$ to a candidate charge distribution. Its induced potential is g and therefore vanishes at every external detector. Unless additional sign, support or constitutive assumptions constrain the source, infinitely many internal distributions remain consistent with exact exterior data. This is stronger than amplification of measurement noise. Total charge from a closed-surface flux or low multipole moments can still be identifiable when the internal density is not.

See an ill-conditioned inverse in a two-source instrument

Two possible source patches sit symmetrically beneath two voltage probes. A calibrated forward response, in volts of probe signal per chosen source unit, is $A = \begin{pmatrix} 1 & 0.98 \\ 0.98 & 1 \end{pmatrix}$. Each patch affects both probes almost equally. The sum pattern $(1, 1)/\sqrt{2}$ and difference pattern $(1, -1)/\sqrt{2}$ are orthonormal eigenvectors as well as singular vectors, with gains 1.98 and 0.02. The condition number is 99. The instrument localises total source strength much more reliably than it locates its division between the two patches.

Suppose the actual source vector is $(1, 0)$. The predictions are $(1.00, 0.98)$ V. A measurement perturbation $(+0.01, -0.01)$ V, though small compared with the signal, lies in the weak difference mode. The exact inverse interprets it as source error $(+0.50, -0.50)$, giving a reconstructed distribution $(1.50, -0.50)$. A common perturbation $(+0.01, +0.01)$ instead produces almost negligible source

changes $+0.00505$ each. A single rms error statistic for the voltage readings misses how their spatial correlation changes uncertainty in the reconstruction.

Use an identity Tikhonov penalty $\alpha = 0.10$. The response factors for the two true source directions become

$$W_{+} = 1.98^2/(1.98^2 + 0.10^2) = 0.9975,$$
$$W_{-} = 0.02^2/(0.02^2 + 0.10^2) = 0.0385.$$

In the absence of noise the estimated source from $(1, 0)$ is approximately $(0.518, 0.480)$. Both patches light up almost equally. This is the resolution matrix in action: the model predicts the observed voltages well yet cannot confidently identify which of the two patches generated them. Under the same opposite noise perturbation the additional reconstructed difference is only 0.0192 on each patch in opposite directions. A visually smooth result is reproducible but biased towards nearly equal strengths.

Add a third probe designed to respond as $(0.6, -0.6)$. This linear combination may be implemented physically by a differential electrode, so negative coefficients are meaningful. The squared singular gain of the old difference direction increases from 0.0004 to $0.0004 + 2(0.6)^2 = 0.7204$, while the common gain is unchanged. The new difference singular value is 0.849, reducing the ratio of the largest to smallest singular values to about 2.33 if the three channels have comparable independent noise in the specified units. Adding another copy of the first common-sensitive probe would have been much less effective. Designing measurements against the weak singular vector can improve source resolution far more than collecting many redundant readings.

Separate numerical error from height uncertainty and unknown backgrounds

Return to a sinusoidal source with negligible aperture. If its true height is h but one divides by an assumed transfer computed at $h + \delta h$, then even exact noiseless data return $\widehat{f}/f = e^{|k|\delta h}$. For a 4 mm wavelength and positive height error 0.10 mm, the fractional amplitude bias is

$$e^{0.157} - 1 = 17.0\%.$$

Unlike random noise this bias does not shrink by averaging repeated scans at the same miscalibrated height. A 20 mm calibration pattern at that same error suffers only about 3.2% bias, so testing one long-wavelength standard need not guarantee short-wavelength accuracy.

For an unknown but stable boundary pattern recorded at two known relative probe heights, $\hat{y}_m(h_2)/\hat{y}_m(h_1) = e^{-|k_m|(h_2 - h_1)}$ for modes above noise when the aperture is unchanged. The source amplitude and common aperture factor cancel. This measures a relative vertical displacement but not the absolute height if the true boundary amplitudes are unknown. Add a calibration electrode of known voltage and wavelength, or an independently measured mechanical height, to determine the absolute scale. Near a sensor's aperture zero the ratio is exceptionally noise sensitive and should not be used for the height fit.

Consider also a slowly varying stray background produced by distant conductors. In a source-free layer of finite thickness a Fourier mode can contain both $Ae^{-|k|z}$ and $Be^{+|k|z}$. The growing term is excluded in a

bounded half-space continued to infinity, but can represent a field launched from the upper boundary of a finite domain. For a single spatial mode measured at two heights one can solve a 2×2 system for A, B. Its determinant is $2\sinh[|k|(h_2 - h_1)]$, up to a sign fixed by the column order. As the two heights coalesce their rows become nearly dependent. A measurement showing unexpected height scaling may thus indicate the wrong outer boundary in the forward model rather than a faulty inverse algorithm.

A reproducible inverse analysis can be audited along three independent axes. First manufacture noise-free readings from a source whose exact forward voltage is known, then refine only the forward discretisation. Second add noise with a specified amplitude and covariance and repeat enough draws to estimate recovered variation in identified modes. Third perturb physical nuisance parameters such as probe width, height and background independently of the noise draws. For each stage state what quantity is being recovered: total charge, a few spatial coefficients, or a real-space image biased by a chosen penalty. These stages expose different error mechanisms even when they yield similar changes in a plotted voltage trace.

Build an exactly invisible internal charge perturbation

The general construction of a silent source can be made explicit for a spherical support of radius a. Pick a smooth interior potential perturbation $g(r) = V_g[1 - (r/a)^2]^3$ for $0 \le r < a$, and set $g = 0$ outside. At $r = a$ both the potential and its first radial derivative vanish; its second radial derivative also approaches zero, avoiding unmodelled surface sheets. It is regular at the origin. Therefore g itself is the potential relative to infinity of the compact charge distribution $\delta\rho = -\epsilon_0 r^{-2} d(r^2 g')/dr$. With $x = r/a$, expansion gives

$$\delta\rho = \frac{\epsilon_0 V_g}{a^2}(18 - 60x^2 + 42x^4), \qquad 0 \le x < 1.$$

The perturbation is positive near the centre and negative in an outer band if $V_g > 0$. Its enclosed charge is

$$-4\pi\epsilon_0 r^2 g' = 24\pi\epsilon_0 V_g a\, x^3(1 - x^2)^2.$$

It vanishes at both ends of the radius interval and is positive at intermediate radii. Gauss's law outside then gives exactly zero electric field and the chosen reference gives zero exterior voltage. A numerical inverse recovering only zero perturbation from exterior data is observationally consistent with both the absence of charge and this particular structured distribution. Integrating the polynomial with its correct $4\pi r^2 dr$ weight confirms the exact zero total charge independently.

For $a = 20$ mm, $V_g = 0.10$ V, the internal centre voltage differs by 100 mV while every ideal exterior detector, no matter how accurately calibrated, records no difference whatsoever. At $r = a/2$, $g = 0.0422$ V and the enclosed perturbation charge is 9.39×10^{-14} C. An internal measurement can therefore break this particular ambiguity. Moving external probes closer while keeping them outside the sphere cannot. Positivity restrictions on the total density can reduce ambiguity, but adding a sufficiently small signed perturbation to a sufficiently positive background can preserve nonnegativity of both candidate densities and still leave their exterior data identical.

This example suggests a practical way to report results when full source reconstruction is impossible. Ask whether the quantity of interest changes along a silent perturbation. Total charge and all external moments stay unchanged here, whereas the central voltage and internal charge profile change. A claim about total charge can be appropriate even when detailed shell-by-shell densities are not identifiable. Conversely a fitted compact central peak may depend on the chosen smoothness penalty or positivity constraints even when every reported exterior measurement is reproduced.

A multiscale numerical exercise can demonstrate the effect reproducibly. Discretise both g and its associated $\delta\rho$, solve the independent forward Poisson problem, and evaluate internal and external probes. Coarse grids may show a small exterior signal if source deposition fails to preserve its neutral moments. Refine the source representation, plot the exterior residual separately from the interior voltage difference and compare with the analytic polynomial. A residual that tends to zero while the interior difference remains finite is the expected physical outcome. Using exactly the same discrete operator to manufacture data and to recover them can mask representation errors, so a second mesh, an analytic source integral or a differently implemented forward solver is especially valuable.

See which part of a specimen electrodes can resolve

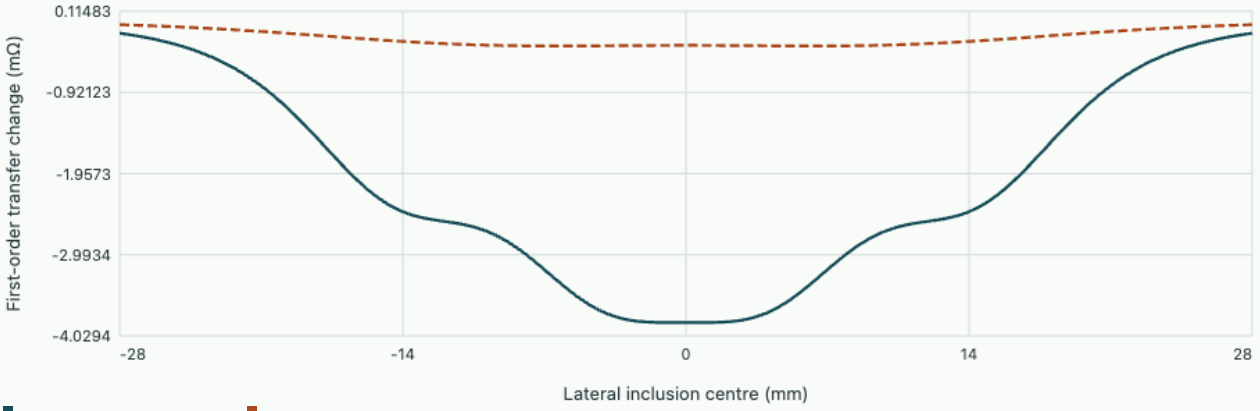
Compare two contact configurations and their sensitivity to a local material change. An accurate potential at the electrodes does not necessarily resolve the interior.

Open this experiment online

Try the calculation

Inclusion depth (mm) 8

Sensing half-spacing (mm) 5



Centre perturbation -3.858 mΩ; unperturbed transfer 79.58 Ω. Sphere radius 1 mm, conductivity increase 0.020 S/m.

Can two independent estimates agree for the wrong reason?

Separate common uncertainty from independent noise in two height scans

A probe measures a 10 mm wavelength mode at heights 1.0 and 2.0 mm with negligible aperture. The two Fourier amplitudes are 5.34 and 2.85 mV, each with independent 0.10 mV random uncertainty, but they share an unknown common voltage gain. Estimate the boundary amplitude and its random error under unit gain. Decide what the agreement can establish about calibration.

▼ **Derive and check the result**

The two transfers are

$$H_1 = e^{-0.2\pi} = 0.5335,$$
$$H_2 = e^{-0.4\pi} = 0.2846.$$

Each individually predicts about 10.0 mV at the boundary. For independent equal-variance amplitude noise, minimise $\sum (y_i - H_i f)^2$. Its solution is $f = (H_1 y_1 + H_2 y_2)/(H_1^2 + H_2^2)$, and random standard deviation $0.10/\sqrt{H_1^2 + H_2^2} = 0.165$ mV. Simply averaging the two directly inverted amplitudes gives the noisier second inversion equal weight and hence a larger variance.

The measured amplitude ratio is about 0.534, as predicted by the additional 1 mm of propagation. A common multiplicative calibration factor cancels from this ratio. If the true voltage gain were 1.10, both unit-gain inversions would agree at about 10 mV even though the physical amplitude would be 9.1 mV. Repeated scans reduce the random uncertainty but cannot remove this systematic scale. A known reference electrode or independent calibrated voltage source is needed.

If instead both probes share a small additive voltage offset in every spatial sample, a full ideal Fourier analysis would place that constant entirely in the zero mode. Incomplete sampling, finite windows or background leakage can mix the offset into the fitted sinusoid. Including a constant nuisance term in a weighted least-squares fit is one way to test that effect directly. Identifying the form of a shared error is as important as quoting its magnitude: common gain, common offset and relative height uncertainty enter the inverse problem in different directions and respond to different calibration experiments.

Tutorial · compare an experimental redesign with a stronger prior

Resolve a window null without changing sample spacing

A voltage camera averages over $w = 4$ mm. A suspected boundary pattern has wavelength 4 mm. Find its measured transfer and propose a second window width that recovers sensitivity. Include the effect of a 1 mm measurement height.

▼ **Work through the solution**

At equal aperture and wavelength the sinc argument is π, giving zero. With a 2 mm window the aperture transfer becomes

$$\sin(\pi/2)/(\pi/2) = 2/\pi.$$

Height attenuation is $e^{-2\pi(1/4)} = 0.2079$, so the combined transfer is 0.132. A 20 mV boundary amplitude therefore gives 2.65 mV in the second scan and exactly zero in the first idealised scan. Collecting more centres with the original width would still give zero. Calibrating the second aperture and its orientation is part of the forward model, especially near a sinc null where small geometry errors strongly change inference.

Choose a regularisation level by its permitted bias

For a mode with transfer 0.20 and unit penalty weight, require the mean reconstructed amplitude to retain at least 80% of the true value. Find the allowed maximum α. For 2 mV noise in the measured coefficient, compare inverse standard deviations with and without this penalty.

▼ **Work through the solution**

Set $W = 0.04/(0.04 + \alpha^2) \geq 0.80$. This gives $\alpha \leq 0.10$. At the limiting value the noise gain is $0.20/(0.04 + 0.01) = 4$, so the reconstructed noise standard deviation is 8 mV. Without regularisation the inverse noise is 10 mV. This modest improvement costs a 20% amplitude bias. The calculated tradeoff may or may not be acceptable for the physical question. Stating both numbers allows a reader to decide whether this experiment can support a claimed amplitude difference.

Further reading

- Measure the transfer of two spatial frequencies
- Reconstruct a tensor from finite electrostatic observations
- MIT, Smith · electric-field inversion and models of interacting electrodes
- Christian Clason · Regularization of Inverse Problems, spectral filters and Tikhonov regularisation

Electrostatic field theory and measurement · 13. Which electrode sees a charge moving above a detector?

A point charge moves without touching the electrodes. Derive the induced charge on one strip from a second, source-free boundary problem. Use the two-dimensional map to distinguish motion across the detector from motion away from it.

Exchange the source and the measuring electrode

An infinite conducting plane is split into long parallel strips. The gaps have negligible width and every strip is held at zero potential during the measurement. Put one charge q at (x, z) above the plane. Computing the field separately for every position would be expensive. Instead consider an auxiliary problem with the measuring strip at unit potential, all other strips grounded and no charge above the plane. Call its dimensionless solution w.

Green's reciprocity pairs the electrode potentials of one problem with the charges of the other. The grounded real electrodes make one boundary contribution vanish. The test electrode gives Q_{strip} and the real point source gives $qw(x, z)$. Their sum vanishes, hence $Q_{\mathrm{strip}} = -qw$. For a strip occupying $-a < u < a$, the Poisson kernel of the upper half-plane integrates the boundary value into the interior:

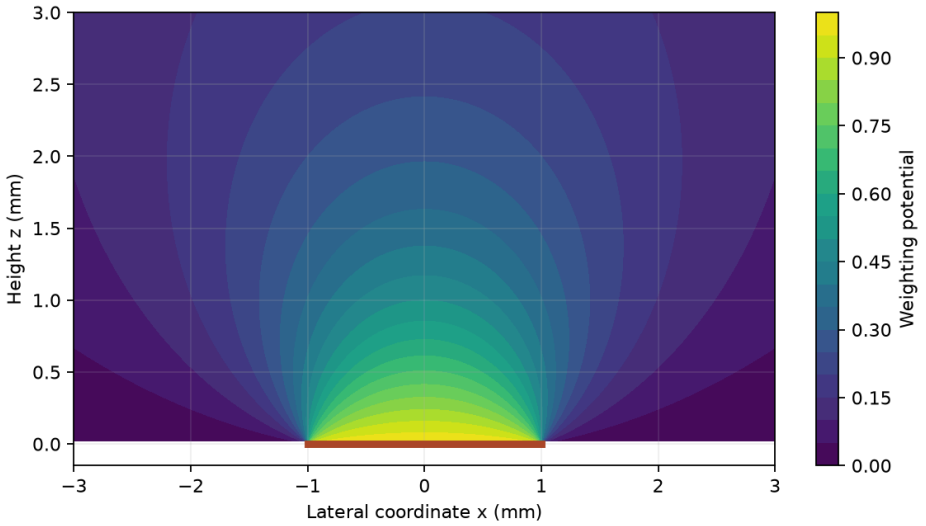
$$w(x, z) = \frac{1}{\pi} \int_{-a}^{a} \frac{z\, du}{(u - x)^2 + z^2} = \frac{\arctan[(a - x)/z] + \arctan[(a + x)/z]}{\pi}.$$

This weighting potential is computed for the electrode geometry. It is independent of the test charge. Its gradient predicts the readout current as a charge travels through different weightings. Infinite strip length and an effectively infinite return plane are the assumptions that permit this two-dimensional construction.

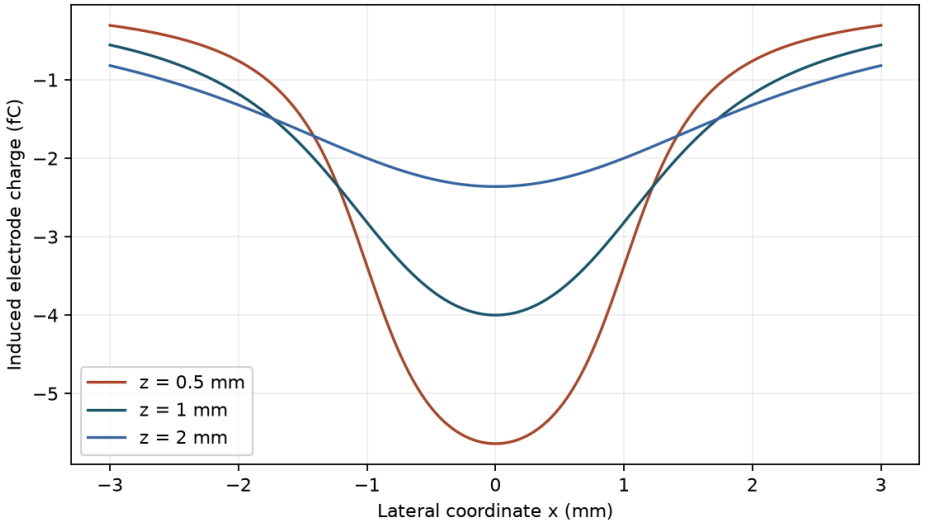
Read a charge and a displacement separately

At the centre and at height $z = a$, the strip subtends a right angle, giving $w = 1/2$. A positive charge of 8 fC induces −4 fC on this electrode. At $x = a, z = a$ the weight is $\arctan(2)/\pi$, about 0.352. The change in measured charge can occur with the same charge and the same height. The peak and width of a lateral scan together carry more positional information than one sample. A finite insulating gap or a grounded back plane would change the boundary problem. A second measurement illustrates why electrode geometry matters for a current signal. Differentiate the central response $w(0, z) = 2\arctan(a/z)/\pi$, giving $\partial_z w = -2a/[\pi(a^2 + z^2)]$. At $z = a = 1$ mm, an 8 fC charge moving upward at 1 mm/µs changes the induced charge at $+8/\pi$ fC/µs, or $+2.55$ nA. The sign here describes the time derivative of charge on the strip. The sign of an external current reading also depends on the chosen lead direction.

Summing the strip weights over an entire tiled plane gives one because the Poisson kernel integrates to one. Total induced charge on that plane is therefore $-q$, even as charge redistributes between individual electrodes. A finite numerical map can miss the long tails of the weighting function. Enlarge the lateral range and check the sum before interpreting a shortfall as an uncollected signal.



Contours show equal sensitivity to a point charge. The steepest change occurs close to an electrode edge.



Move the charge at constant height and record the induced charge. The scan width can help distinguish a small nearby charge from a larger distant one.

Exercise · close the charge balance

If the entire plane is read out, how much induced charge do all electrodes collect together?

▼ Work through the solution

The full boundary integral runs from minus infinity to plus infinity. Its arctangent increases by π, so the weights sum to one and the total induced charge is $-q$. In the notebook, add the weights from neighbouring strips and watch the deficit shrink as the readout array grows.

Exercise · infer a height

A known charge above the centre produces a weight of one third. Find the height in units of the half-width.

▼ **Work through the solution**

At the centre $w = 2\arctan(a/z)/\pi$. Thus $a/z = \tan(\pi/6) = 1/\sqrt{3}$ and $z = \sqrt{3}a$. An unknown lateral offset would spoil this one-parameter inversion.

Try it · Strip charge

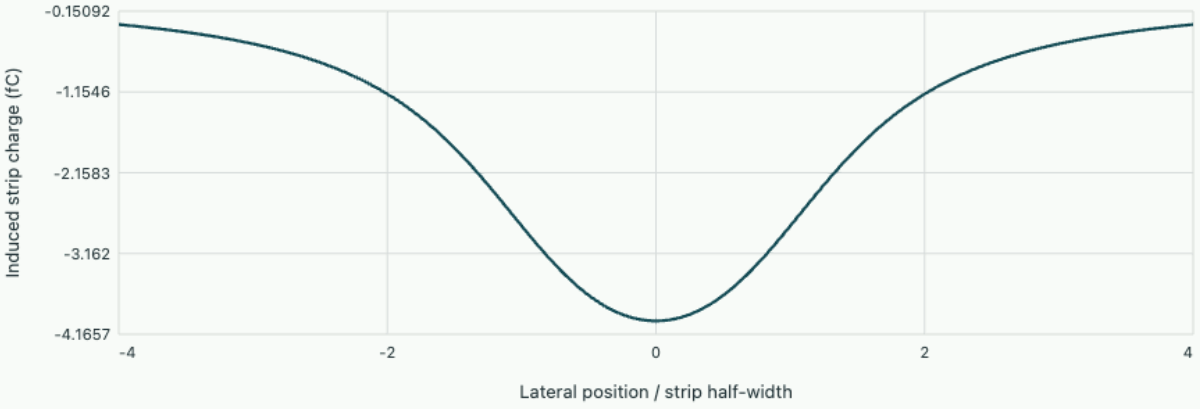
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Charge height / strip half-width **1**

Point charge (fC) **8**



Grounded strip charge

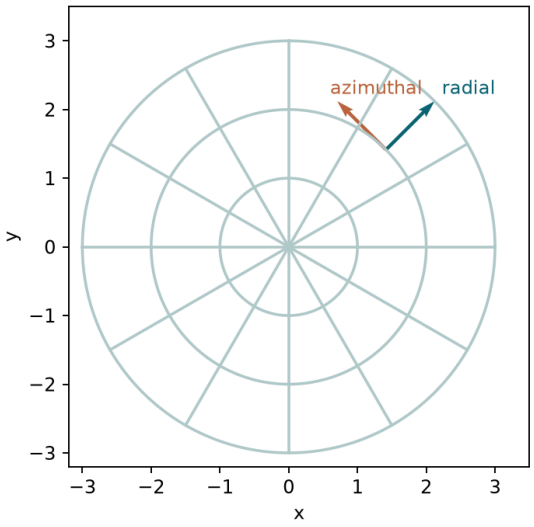
At the centre the electrode receives -4 fC. Reciprocity converts the strip weighting potential into induced charge.

Download the corresponding editable notebook. The browser recalculates the reduced equations directly; the notebook exposes the full calculation and its checks.

Open editable notebook Download .ipynb

Electrostatic field theory and measurement · 14. Build the field operators in curved coordinates

Derive the geometric factors in gradient, divergence, curl and Laplacian, then use dimensional conversion to compare electromagnetic conventions.



Cylindrical coordinate curves and local radial and azimuthal basis vectors. Equal angular increments span longer arcs at larger radius.

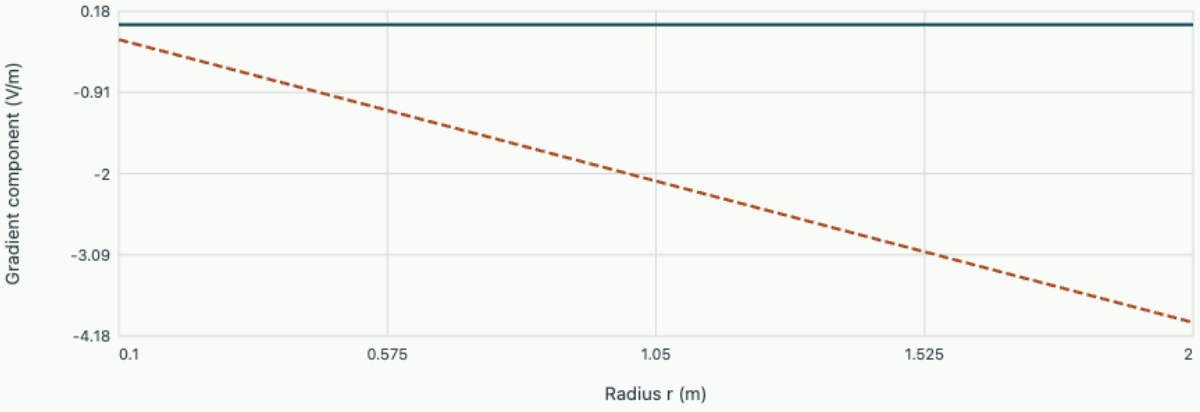
Try it · Cylindrical

Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Fixed angle ϕ (degrees) 45



For V = (1 V/m²) r² cos(2ϕ), the Laplacian is zero: radial 4 cos(2ϕ) and angular −4 cos(2ϕ) cancel. The azimuthal derivative needs the metric factor 1/r.

Attach a length to a coordinate increment

Coordinates label points, while differential operators describe spatial change. In a right-handed orthogonal system q_1,q_2,q_3 write the position as r(q). The coordinate tangent vectors are mutually perpendicular. Define $h_i=|\partial r/\partial q_i|$ and unit vectors $e_i=(\partial r/\partial q_i)/h_i$. A small displacement is $\Sigma e_i h_i dq_i$. Matching $dV=\Sigma(\partial V/\partial q_i)dq_i$ with $dV=\nabla V\cdot dr$ gives the gradient.

$$\nabla V = \sum_i \frac{\mathbf{e}_i}{h_i}\frac{\partial V}{\partial q_i},$$

$$d\mathcal{V} = h_1 h_2 h_3\, dq_1 dq_2 dq_3.$$

For cylindrical coordinates (s,ϕ,z), r=(s cosϕ,s sinϕ,z). Differentiation gives h=(1,s,1), e_s=(cosϕ,sinϕ,0) and e_ϕ=(−sinϕ,cosϕ,0). Thus ∂ϕ e_s=e_ϕ and ∂ϕ e_ϕ=−e_s. For spherical coordinates (r,θ,ϕ), direct differentiation gives h=(1,r,r sinθ). The coordinate basis changes direction even if the components of a field stay constant.

Recall · Angle increments are dimensionless

An angular change dϕ corresponds to arc length s dϕ in a cylinder. Accordingly the azimuthal derivative of potential has physical component (1/s)∂ϕV. This has the same units, volts per metre, as ∂sV. Omitting the geometric factor changes both dimensions and the electric field.

Derive divergence by subtracting opposite face fluxes

A small face normal to e_1 has area $h_2h_3dq_2dq_3$. The two opposite faces have different field and potentially different area. Subtract their outward fluxes using a first-order expansion, sum over the three coordinate directions and divide by cell volume.

$$\nabla \cdot \mathbf{A} = \frac{1}{h_1 h_2 h_3}\left[\partial_1(h_2 h_3 A_1) + \partial_2(h_3 h_1 A_2) + \partial_3(h_1 h_2 A_3)\right].$$

The cylindrical radial term is (1/s)∂s(sA_s). The spherical radial term is $(1/r^2)\partial r(r^2A_r)$. These follow from circumference and surface area respectively. For A=Cr e_r in a sphere, the divergence is 3C. A sphere of radius R then encloses flux $4\pi CR^3$, matching the volume integral of 3C.

Derive curl from the oriented boundary of one face

Choose the q_2–q_3 face with positive normal e_1. Circulation along its four edges gives $\partial_2(h_3A_3)-\partial_3(h_2A_2)$, multiplied by dq_2dq_3. Divide by $h_2h_3dq_2dq_3$. Cyclic permutation supplies all three components.

$$(\nabla \times \mathbf{A})_1 = \frac{\partial_2(h_3 A_3) - \partial_3(h_2 A_2)}{h_2 h_3}.$$

In cylindrical symmetry A=Aϕ(s)eϕ yields $(\text{curl } A)z=s^{-1}d(sA\phi)/ds$. Aϕ=Bs/2 has curl B ez, the familiar vector potential for a uniform axial magnetic field. Aϕ=C/s has zero local curl away from the axis but a nonzero circulation 2πC around it. The excluded axis supplies the topological or singular information.

Remember · Zero local curl needs a domain condition

Stokes' theorem requires a spanning surface inside the region where the field is regular. A loop encircling an excluded axis need not have such a surface. To infer a single-valued potential from zero curl, check the domain as well as the local differential equation.

Compose the two operators

Insert the gradient components into the divergence formula. The scalar Laplacian is therefore

$$\nabla^2 V = \frac{1}{h_1 h_2 h_3} \sum_i \partial_i \left(\frac{h_1 h_2 h_3}{h_i^2} \partial_i V \right).$$

In spherical coordinates this becomes $r^{-2}\partial r(r^2\partial rV)+(r^2\sin\theta)^{-1}\partial\theta(\sin\theta\partial\theta V)+(r^2\sin^2\theta)^{-1}\partial\phi^2V$. For $V=r^2P_2(\cos\theta)$, the radial contribution is $6P_2$ and the angular contribution is $-6P_2$. The potential is harmonic even though neither contribution vanishes separately. This cancellation underlies separation into radial powers and angular eigenfunctions.

Convert a physical experiment between conventions

In SI vacuum electrostatics the point-source law is $E=q/(4\pi\varepsilon_0 r^2)$, with force qE. Gaussian conventions write the corresponding laws $E=q/r^2$ and $F=qE$, while magnetic force contains $v\times B/c$. The numerical values and dimensions of both charge and field change. To convert, equate a measurable force in the same physical experiment first. One statcoulomb is $3.335640952\times10^{-10}$ C, one statvolt/cm is 29979.2458 V/m, and one gauss is 10^{-4} T.

$$\nabla \cdot \mathbf{E}_{\mathrm{G}} = 4\pi\rho_{\mathrm{G}},$$

$$\nabla \times \mathbf{B}_{\mathrm{G}} = \frac{1}{c}\partial_t \mathbf{E}_{\mathrm{G}} + \frac{4\pi}{c}\mathbf{J}_{\mathrm{G}}.$$

Gaussian electric and magnetic fields share dimensions. Rationalised Heaviside–Lorentz conventions absorb the source factors 4π into the field and charge normalisation. A conversion must state whether c is retained or set to unity. The route from an equation to a voltage or force requires restoring that convention. A ratio E/B in SI is a velocity, whereas E_G/B_G for a plane vacuum wave equals one.

Exercises on geometry and conversion

1. Derive the cylindrical Laplacian and apply it to $V=s^2\cos2\phi$. What Cartesian polynomial does this represent?

▼ **Work through the solution**

The radial term is $4\cos2\phi$ and the angular term is $-4\cos2\phi$. Their sum vanishes. Since $s^2\cos2\phi=x^2-y^2$, the Cartesian Laplacian is $2-2=0$.

---

2. A plane wave has E=300 V/m. Determine B in tesla and gauss, then check the Gaussian electric-field number.

▼ **Work through the solution**

$B=E/c\approx1.00069$ μT$=0.0100069$ G. Dividing 300 V/m by 29979.2458 V/m per statvolt/cm gives 0.0100069 statvolt/cm, the same Gaussian number.

---

3. Evaluate the outward flux of C e_s/s through a cylindrical shell and explain why local zero divergence does not remove an axial source.

▼ Work through the solution

The curved outer face gives 2πCL for length L, independent of radius. A shell with both inner and outer faces has net zero flux. A solid cylinder includes the singular axis, where the distributional source supplies 2πC per unit length.

Magnetism, induction and field dynamics

Begin with a steady current and follow what changes when a source, a conductor or a field moves. Track energy and momentum with explicit system boundaries.

1. Building a magnetic field from a measured current

A Hall probe beside a bent wire reports a direction and a magnitude. The current flowing past it is easy to measure, but a single current reading cannot specify the field unless the return path and geometry are also known. Starting with the local field equations, we will derive the integral law that makes those geometrical details calculable.

2. The vector potential as an instrument for computing flux

An experimenter can measure a flux change with a coil without measuring the magnetic field over every square millimetre of the coil area. The vector potential gives a corresponding computational tool. Its line integral determines flux, while the freedom to choose its gauge can turn a difficult boundary problem into a manageable one.

3. Which features of a current distribution survive at a distance?

A current loop just fits behind a probe mount. At a nearby point its size and placement matter. Far away the same winding behaves as though all that geometry were a single vector. Predicting when that replacement is accurate, and what remains when the vector cancels, requires an expansion of the source rather than an appeal to a field-line sketch.

4. From microscopic moments to measured magnetisation

A magnetised sample can make a compass turn without a power supply attached. The macroscopic theory assigns it bound currents in its bulk and on its surface. These are calculated sources for the magnetic field. We will reconstruct them from magnetic moments, predict signals from finite samples, and ask what an exterior field measurement can identify.

5. Interfaces, demagnetising fields and the work of magnetisation

Two samples made of the same material can respond very differently to an externally applied field if one is a long rod and the other is a thin disc. The induced moment depends on the internal field, which the magnetised sample helps create. The same distinction reappears when one tries to infer a force from stored magnetic energy.

6. Motional emf and transformer emf in moving circuits

Two laboratory demonstrations produce the same voltmeter trace. In one, a stationary loop sees a changing magnetic field. In the other, a conductor moves through a stationary field. The charges experience different local forces, yet under identifiable conditions both readings follow from the rate of change of magnetic flux.

7. Flux linkage, coupled coils and magnetic coenergy

A winding links flux through each turn. In two windings the current in either one can change the voltage in the other. A matrix of flux linkages describes this coupling, but predicting power or

mechanical force requires specifying the current directions and the control conditions.

8. How a magnetic field relaxes into a finite conductor

A conducting sheet initially shields its interior from an applied magnetic step. Given time, a normal resistive metal admits a static tangential magnetic field. The same equations give a finite slab a discrete relaxation spectrum and a periodic drive a complex penetration depth. These are different initial and boundary conditions for one diffusion operator.

9. Displacement current and the charging capacitor

A wire feeds a capacitor whose dielectric blocks the transport of mobile charges between its plates. A magnetic probe near the gap still responds to charging. Resolving which surface a circulation law should use leads to displacement current, an electric field correction during a ramp, and an energy-flow test around the rim.

10. Energy flow and mechanical force from vacuum fields

An external support holds a charged capacitor plate or a current-carrying coil in place. One can infer the support force from fields just outside the body, without summing forces on every microscopic charge. A second surface integral tracks the energy feeding the same apparatus. Both methods start with Maxwell's equations, but they transport different conserved quantities.

11. Field momentum, momentum exchange and the quasistatic limits

A system can have static electric and magnetic fields with nonzero local electromagnetic momentum density. A circuit can also be electrically small while either induction or displacement current remains indispensable. Both puzzles become clearer when one follows the relevant conservation law and compares the scales of the discarded terms.

12. Posing an electromagnetic evolution problem

A solver cannot evolve 'a pulse in a box' from geometry and a time step alone. It needs material laws, compatible initial fields, a charge-conserving source, and enough boundary data to specify what enters the domain. The same bookkeeping is required when interpreting a real transient measurement.

13. Why does a Helmholtz pair have a useful uniform region?

A magnetometer averages the field over a finite sample. Derive the coil spacing that suppresses curvature and predict how rapidly the remaining calibration error grows away from the midpoint.

14. From atomic motion to magnetic susceptibility

Calculate the response of a bound orbit and of an ensemble of existing moments, then follow how interactions permit spontaneous magnetisation.

15. The electric field of a rotating magnet

Use charge balance and a surface matching calculation to predict a voltage around a rotating magnetic body, then close the mechanical and electrical power accounts.

Magnetism, induction and field dynamics · 1. Building a magnetic field from a measured current

A Hall probe beside a bent wire reports a direction and a magnitude. The current flowing past it is easy to measure, but a single current reading cannot specify the field unless the return path and geometry are also known. Starting with the local field equations, we will derive the integral law that makes those geometrical details calculable.

Decide what is steady

Imagine making two maps at successive times of the current density in a conducting ring. Charges move, but if the maps agree, the source can be stationary. The current through an oriented cross section is $I = \int \boldsymbol{J} \cdot d\boldsymbol{a}$. Charge balance in any fixed volume requires

$$\partial_t \rho + \nabla \cdot \boldsymbol{J} = 0.$$

A genuinely steady system has $\nabla \cdot \boldsymbol{J} = 0$, including across contacts where current changes direction. A single finite wire element carrying a constant current cannot be an entire isolated source: continuous current would have to appear at one end and disappear at the other. It can be a contribution to the field of an explicitly closed arrangement.

In vacuum the magnetostatic equations are

$$\nabla \cdot \boldsymbol{B} = 0,$$
$$\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J}.$$

Stokes's theorem turns the second into $\oint \boldsymbol{B} \cdot d\boldsymbol{\ell} = \mu_0 I_{\mathrm{enc}}$. It gives one number for the whole closed path. For an infinite straight wire rotational and translational symmetry establish in advance that the field is azimuthal and constant along a coaxial circle. Only then does the integral determine a local field $B = \mu_0 I/(2\pi s)$. For a bent wire the same circulation equality holds. Extracting a pointwise magnitude from it is a separate inverse problem. A linear superposition integral resolves that problem directly.

Invert the local curl equation

Take the curl of Ampère's equation. In Cartesian coordinates the identity $\nabla \times \nabla \times \boldsymbol{B} = \nabla(\nabla \cdot \boldsymbol{B}) - \nabla^2 \boldsymbol{B}$ applies componentwise. Since its divergence vanishes,

$$-\nabla^2 \boldsymbol{B} = \mu_0 \nabla \times \boldsymbol{J}.$$

Assume first that the sources occupy a bounded region of space and the field decays at large distance. For each component the inverse of $-\nabla^2$ is the Newton kernel $G(\boldsymbol{r} - \boldsymbol{r}') = 1/(4\pi|\boldsymbol{r} - \boldsymbol{r}'|)$. Its normalisation follows from integrating $-\nabla^2 G$ over a little sphere. The outward derivative contributes precisely one.

Thus the field is $\mu_0 \int G\, \nabla' \times \boldsymbol{J}\, d^3 r'$. Move the source derivative off the current by integrating the product $\nabla' \times (G\boldsymbol{J})$. Its surface integral at infinity vanishes for localised currents. The remaining term is $-\mu_0 \int (\nabla' G) \times \boldsymbol{J}\, d^3 r'$. Writing $\boldsymbol{R} = \boldsymbol{r} - \boldsymbol{r}'$,

$$\nabla' G = +\boldsymbol{R}/(4\pi R^3).$$

Interchanging the two factors in the cross product supplies another minus sign.

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \int \frac{\boldsymbol{J}(\boldsymbol{r}') \times (\boldsymbol{r} - \boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|^3}\, d^3 r'.$$

The integrand tells us where direction comes from. It is perpendicular both to the local current and to the vector pointing from source to probe. A segment aimed straight at the probe contributes zero there. A very distant transverse segment contributes as its length divided by squared distance. The special $1/s$ field of an infinite wire results from summing contributions along an extended length. A free-space field could still contain an additional homogeneous contribution if large coils elsewhere impose a background; the assumed boundary condition fixed that possibility. In a cavity bounded by magnetic material one needs the corresponding boundary solution or all equivalent sources.

Collapse a slender wire into a curve

Within a narrow conducting filament take arc length u, unit tangent $\boldsymbol{t}(u)$ in the current direction, and transverse section $S(u)$. If the observation distance and the local radius of curvature both greatly exceed the wire radius, the kernel changes little across a section. The volume element contributes its area times du, while the current density integrated over that area gives I. Their product becomes

$$I \boldsymbol{t}\, du = I\, d\boldsymbol{\ell}'.$$

The resulting Biot–Savart integral is a geometric idealisation with units $\mu_0 I/\mathrm{m} = \mathrm{T}$.

For a straight section on the z-axis from $-a$ to $+a$, observe on the perpendicular bisector at $(s, 0, 0)$, $s > 0$. A current directed towards $+z$ gives

$$\boldsymbol{e}_z \times (s\boldsymbol{e}_x - z'\boldsymbol{e}_z) = s\boldsymbol{e}_y.$$

The only surviving component requires the elementary integral $\int s(s^2 + z'^2)^{-3/2} dz' = z'/[s\sqrt{s^2 + z'^2}]$, hence

$$B_y = \frac{\mu_0 I}{2\pi s} \frac{a}{\sqrt{a^2 + s^2}}.$$

With $a = 0.10\,\mathrm{m},\ s = 0.040\,\mathrm{m},\ I = 3.0\,\mathrm{A}$, the infinite-wire value is $15.0\,\mu\mathrm{T}$ and the finite-section contribution is $13.93\,\mu\mathrm{T}$. A detector with a one-percent calibration would resolve the difference. Moving the probe twice as far changes the finite-length factor as well as $1/s$, so simply dividing the field by two is insufficient. To predict the experiment one must add the field from the return conductor. Tilting that return to point approximately towards the probe can reduce its contribution even when its nearest distance stays similar.

Sum a loop before invoking its symmetry

Put a circular loop of radius a in the xy-plane with counterclockwise current as seen from $+z$. A source element $a\,d\varphi\,\boldsymbol{e}_{\varphi}$ sees the axial point $z\boldsymbol{e}_z$ along $z\boldsymbol{e}_z - a\boldsymbol{e}_s$. Its cross product is $a\,d\varphi(z\boldsymbol{e}_s + a\boldsymbol{e}_z)$. The radial terms integrate to zero because their directions rotate around the ring. Every axial term has the same direction and distance factor. Integrating the factor around the circle gives

$$B_z(z) = \frac{\mu_0 I a^2}{2(a^2 + z^2)^{3/2}}.$$

At the centre the field is $\mu_0 I/(2a)$. For $a = 0.050\,\mathrm{m}$ and $I = 2.0\,\mathrm{A}$, this is $25.13\,\mu\mathrm{T}$, comparable to ordinary ambient geomagnetic field scales. At $z = a$ it is only $2^{-3/2} = 0.3536$ of the central reading, or $8.89\,\mu\mathrm{T}$. A practical isolation procedure measures once with positive current and once with negative current. Half the reading difference removes a background stable between the two runs. An intensity-only sensor that records $|\boldsymbol{B}|$ instead of a signed axial component needs a separate procedure, since magnitudes of vectors do not add linearly.

The loop also supplies an internal test of rough field approximations. At $|z| \gg a$, expanding $(1 + a^2/z^2)^{-3/2}$ gives

$$B_z = \mu_0 I a^2/(2|z|^3)[1 - 3a^2/(2z^2) + \cdots].$$

In the median plane a dipole moment points opposite to the local external field, whereas on either extension of its axis the field points in the moment direction. A scalar drawing that simply marks a plus sign for $z > 0$ and a minus sign for $z < 0$ would confuse the direction of the observation position with the direction of the field.

Give a singular source a physical cross section

The filament field diverges as the mathematical distance to an ideal line tends to zero. For a long cylindrical conductor of radius a carrying uniform steady current I, the field is regular. An Amperian circle at $s < a$ encloses Is^2/a^2. Its circulation yields $B_{\varphi}(s) = \mu_0 I s/(2\pi a^2)$, tending to zero at the centre. Outside it joins the filament result continuously. If all the current instead flowed in an infinitesimal surface sheet, the interior field would vanish and a tangential jump would appear at the surface. The outside measurement of a perfectly symmetric straight wire would be identical in these two ideal models.

For a $2.0\,\mathrm{mm}$ radius wire with $I = 4.0\,\mathrm{A}$, the surface value is $0.400\,\mathrm{mT}$. Halfway out, uniform bulk flow gives $0.200\,\mathrm{mT}$; a thin outer sheet gives zero there. The same reasoning underlies an inverse estimate of current density from measured interior field:

$$J_z = (1/\mu_0 s) d(sB_{\varphi})/ds.$$

Differentiating noisy measurements can amplify high spatial frequencies, so one can instead compare flux integrals over annuli. The difference $2\pi[s_2 B(s_2) - s_1 B(s_1)]/\mu_0$ gives the current through an annulus without numerical differentiation.

Set up a finite-current computation

For a polygonal coil, sum each oriented segment contribution as a three-component vector. Split a curved path into short sections, evaluate geometry at each midpoint, and compare with a second run using twice as many sections. For a smooth integrand at a fixed point away from the winding one expects second-order midpoint convergence. Near a filament the kernel varies on the source-to-probe separation. Refining until the segment length is small relative to that separation can require much finer sampling than refining the plotted contour itself. At a point inside the conductor, improving a line approximation cannot replace a cross-sectional current model.

An independent quality test is symmetry. The two transverse axial components of a centred circular coil must cancel separately to roundoff for a symmetric discretisation. An axial field may nonetheless be inaccurate if all segment lengths systematically approximate the wrong radius. Check the absolute axial benchmark as well as the zero components. A second useful test reverses every edge orientation while keeping the probe and axes fixed; all three field components must reverse. These checks detect mistakes in geometry, integration and sign that a visually plausible field-line plot can hide.

Compare source geometry in the steady-current experiment

Design a nearly uniform calibration field

A single loop has a local maximum of axial field at its centre, with appreciable curvature. A sensor of finite size averages across positions and can read a different value from a nominal point probe. Two identical coaxial loops with equal current permit the curvature to be removed at a chosen point. Put their centres at $z = \pm d/2$. Let $b(z) = K(a^2 + z^2)^{-3/2}$, $K = \mu_0 N I a^2/2$, be the field from one N-turn loop centred at the origin. By superposition,

$$B(z) = b(z - d/2) + b(z + d/2).$$

This is even in z for every spacing, so its first derivative at the midpoint vanishes. A flat-looking midpoint on a coarse graph consequently gives no information about whether the separation was set correctly.

Differentiate once to obtain

$$b' = -3Kz(a^2 + z^2)^{-5/2}.$$

The next derivative is

$$b'' = 3K(4z^2 - a^2)(a^2 + z^2)^{-7/2}.$$

The pair therefore has $B''(0) = 0$ when $4(d/2)^2 = a^2$: the separation equals the radius. At that setting the central value is

$$B(0) = (4/5)^{3/2}\mu_0 N I/a.$$

With $a = d = 0.10\,\mathrm{m}$, 50 turns on each winding and 0.20 A, it is $89.9\,\mu\mathrm{T}$. A thermal expansion increasing both radius and separation by one percent preserves their ratio yet lowers the field calibration nearly one

percent at fixed current. Geometrical uniformity and absolute calibration are distinct tolerances.

There is a useful way to estimate the size of the flat region without trusting the plot scale. Write $u = z/a$ and expand each factor $[1 + (u \pm 1/2)^2]^{-3/2}$ in the exact two-coil formula. Odd terms cancel in the sum, and the quadratic term was eliminated by design. The first surviving correction is

$$B(z)/B(0) = 1 - (144/125)u^4 + O(u^6).$$

At one tenth of a radius it predicts a fractional axial fall of 1.15×10^{-4}, roughly 0.012 percent. At one quarter radius the quartic estimate is 0.45 percent; comparison with the unexpanded expression is more appropriate if a tight tolerance is specified. With a single loop, by contrast, $B(z)/B(0) = 1 - 3u^2/2 + O(u^4)$, already losing about 1.5 percent at the first of those positions.

Is the field also uniform slightly off the axis? In a current-free axisymmetric region, smooth fields obey $\nabla \times \boldsymbol{B} = 0$ and $\nabla \cdot \boldsymbol{B} = 0$. Write $B_z(s, z) = B_0(z) + c(z)s^2 + \cdots$ and $B_s = d(z)s + \cdots$; odd radial axial terms would fail smoothness at the axis. From the divergence $2d + B_0' = 0$. The azimuthal curl gives $d' = 2c$, hence $B_s = -sB_0'/2 + \cdots$ and $B_z = B_0 - s^2 B_0''/4 + \cdots$. Eliminating axial curvature removes the leading radial variation too. It does not remove the field of an asymmetrically routed supply lead, which breaks the expansion's assumed symmetry. Twisting a distant supply pair and measuring a reversed-current map test that separate contribution.

Separate the feed from the return in a field measurement

Consider two long straight parallel conductors separated by $2a$. The right conductor carries $+I$ along z; the left carries $-I$. On the line of centres at position $x > a$, their azimuthal fields oppose. Adding signed magnitudes gives

$$B_y = \frac{\mu_0 I}{2\pi}\left(\frac{1}{x-a} - \frac{1}{x+a}\right) = \frac{\mu_0 I a}{\pi(x^2 - a^2)}.$$

At $a = 5.0\,\mathrm{mm}$, $x = 50\,\mathrm{mm}$, $I = 1.0\,\mathrm{A}$, the field is $0.808\,\mu\mathrm{T}$. The nearer conductor alone would contribute $4.44\,\mu\mathrm{T}$. Omitting the return because it is 'the other lead' would overpredict the signal by more than a factor of five. Well outside the wire spacing but far from the ends, the leading transverse field falls as $1/x^2$; at still greater distances than the full length of a finite closed rectangular winding, the complete assembly ultimately behaves as a three-dimensional dipole and falls as $1/r^3$. The apparent power law therefore depends on which distances are large relative to which parts of the apparatus.

Now move the probe between the conductors to $x = 0$. Both fields point in the same transverse direction there, with magnitude $\mu_0 I/(\pi a) = 80.0\,\mu\mathrm{T}$ for the same parameters. A probe displaced laterally by a fraction of the spacing can therefore see a large change even at fixed current. Reversing both lead currents swaps every sign. Reversing just the nominal return current would create parallel currents and describe a new unseen return elsewhere, illustrating why 'individual wire fields' require their circuit context. If the readings are used to infer a small current, fit the two known conductor positions and the background rather than using the distance to the nearest visible piece of copper as the only geometry input.

The two examples provide complementary ways to verify a coil calculation experimentally. For a calibration pair, motion over a known traverse tests a deliberately small spatial derivative, while reversing current rejects an additive background. For opposed leads, the sharp transition between addition and cancellation tests probe alignment and whether the return has been modelled. A single calibration at the midpoint could succeed in either geometry with the wrong sign convention or an erroneously rescaled current. A signed scan over a measured range of positions exposes those degeneracies.

Recover a magnetic field from a steady current

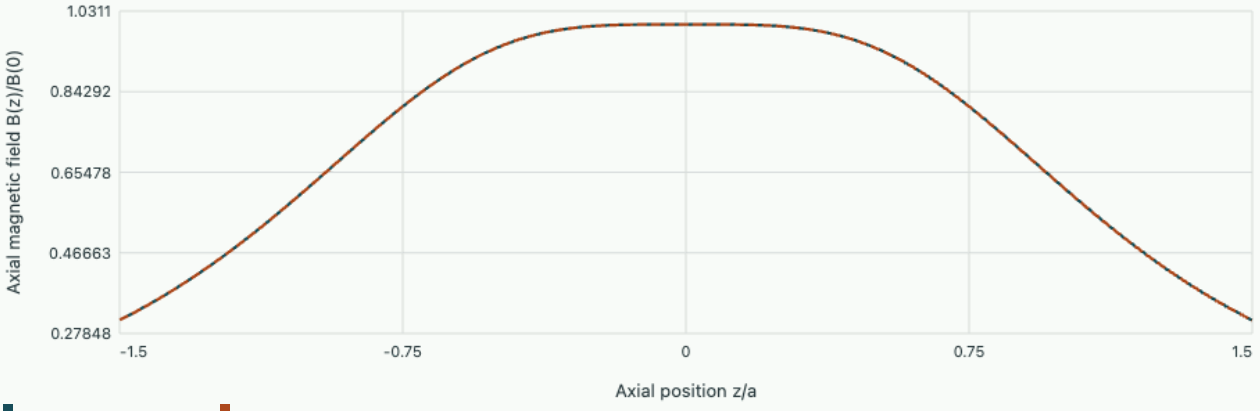
Compare the on-axis contributions of current loops before and after changing their spacing. Check the sign and spatial uniformity near the centre.

Open this experiment online

Try the calculation

Coil separation / radius **1**

Axial test position / radius **0.25**



At z/a = 0.25 the selected field differs from the central field by -0.42%. The sign distinguishes a central peak from a central dip close to the origin.

Resolve a current strip from its measured field

A long conducting strip occupies the region between $x = -a$ and $x = a$ in the plane $z = 0$. Conventional current flows in positive y. Suppose leads far from the observation region distribute the current uniformly across the width. Calling the surface current K fixes the current in the narrow ribbon between x' and $x' + dx'$ as $K\,dx'$. The total current is therefore $I = 2aK$. A wide trace on a board and a narrow round wire carrying the same current look similar at a distance. An observer close enough to resolve the width should be able to tell them apart. We can calculate exactly how close the observer has to come in this ideal geometry.

Each ribbon is an infinitely long straight source. Its contribution lies on a circle around the y axis through x'. At an observation point $(x, 0, z)$, its direction is proportional to $\hat{y} \times [(x - x')\hat{x} + z\hat{z}]$. The field vector is the transverse displacement rotated through a right angle and divided by its squared length. The familiar inverse distance of the straight wire is still present because the numerator has the magnitude of one distance. On the centre line the contributions to B_z cancel in pairs. All contributions to B_x add with the same sign above the plane. For positive z,

$$B_x(0, z) = \frac{\mu_0 K}{2\pi} \int_{-a}^{a} \frac{z\, dx'}{x'^2 + z^2} = \frac{\mu_0 K}{\pi} \arctan \frac{a}{z}.$$

At heights much smaller than the half-width the arctangent approaches $\pi/2$. A locally infinite sheet gives the same field, $\mu_0 K/2$. At heights much greater than the half-width, replacing the arctangent by its argument gives $\mu_0 I/(2\pi z)$. The first correction to this wire prediction has relative size $-a^2/(3z^2)$. A single formula thus shows when an idealisation loses the length scale that distinguishes different sources. At height $z = a$ the strip gives exactly one half of its value immediately above the centre of the sheet.

Take a strip of width $40\,\mathrm{mm}$ with current $8.0\,\mathrm{A}$. The surface current is $200\,\mathrm{A/m}$, and the limiting centre field is $125.7\,\mu\mathrm{T}$. At $z = 5\,\mathrm{mm}$, the exact field is about $106.1\,\mu\mathrm{T}$. At $20\,\mathrm{mm}$, it is $62.8\,\mu\mathrm{T}$. A line-current estimate at this second position gives $80.0\,\mu\mathrm{T}$. This discrepancy occurs even when the source is perfectly long. Extending the source in the y direction cannot cure an approximation that discarded the width in the x direction.

A scan at fixed height reveals another part of the vector field. Integrating the signed z component gives

$$B_z(x, z) = \frac{\mu_0 K}{4\pi} \ln \frac{(x - a)^2 + z^2}{(x + a)^2 + z^2}.$$

The logarithm vanishes at the centre, changes sign across it and is negative on the positive x side. Just above an edge the ideal formula grows in magnitude as the height decreases. The singular geometry assumed zero thickness and terminated a uniformly flowing sheet abruptly at a perfectly sharp boundary. A probe of finite active size measures a spatial average. A model of a real edge also needs its thickness and actual current distribution. Neither complication changes the centre-line limiting fields at heights well above the thickness and well inside the long region of the strip.

A student measures a short-distance plateau and records a centre field equal to half the plateau at a nominal height of $12\,\mathrm{mm}$. What width follows, and how could a sensor-position offset corrupt the inference?

▼ Work through the solution

The half-height condition requires an arctangent equal to $\pi/4$, so the actual separation equals the half-width. The inferred full width is $24\,\mathrm{mm}$ if the sensitive element lies at the point used to define the nominal height. Suppose instead that the sensitive element sits $2\,\mathrm{mm}$ behind the face of its package and the nominal reading gives the face height above the strip. The field probe was then $14\,\mathrm{mm}$ from the current sheet, and the inferred full width becomes $28\,\mathrm{mm}$. A fit at several heights can estimate an unknown offset together with the width. A remote return wire adds an approximately constant background over a short

scan, so reverse the source current and measure the field of the return geometry separately when the desired precision demands it.

What happens to a far-field and a near-field reading if the same total current is spread over twice the width?

▼ **Work through the solution**

In the far field the leading term depends only on total current, so it is insensitive to the change provided the observation height is still much larger than the new width. On the near-field centre plateau the field halves because $K = I/(2a)$ halves. At an intermediate height use the arctangent with the new half-width and the new value of K. Doubling the width while keeping the surface current fixed describes a different experiment because the supply then delivers twice as much total current.

David Tong, Magnetostatics, on deriving a field from its current distribution.

Tutorial on recovering fields from geometry

A square loop has side $2a$ and current I. Find the field at its centre by using the finite-section result. For $a = 0.050\,\mathrm{m},\ I = 2.0\,\mathrm{A}$, compare it with a circular loop of equal enclosed area.

▼ **Work through the solution**

Each straight side is at perpendicular distance a, and its half-length is a, giving $\mu_0 I/(2\pi a\sqrt{2})$ per side. All four point along the same normal. The result is

$$\sqrt{2}\mu_0 I/(\pi a) = 22.63\,\mu\mathrm{T}.$$

Equal area requires a circular radius $2a/\sqrt{\pi} = 0.0564\,\mathrm{m}$; its field is $22.27\,\mu\mathrm{T}$. Equal enclosed areas imply equal dipole moments far away, but a central probe distinguishes the two windings.

A cylindrical wire with azimuthal field $B_\varphi = bs^3$ inside radius a carries an unknown axial current distribution. Reconstruct it and compare the current inside $a/2$ with the current inferred at a.

▼ **Work through the solution**

Ampère gives

$$I_{\mathrm{enc}} = 2\pi b s^4/\mu_0.$$

The differential current through an annulus is $2\pi s J_z\,ds$, yielding $J_z = 4bs^2/\mu_0$. The half-radius encloses $1/16$ of the total. A uniform-density wire would enclose $1/4$, so a single exterior reading cannot determine the distribution while several interior readings could do so.

At what ratio a/s does the infinite-wire approximation overestimate the finite straight section on its perpendicular bisector by less than one percent, measured against the exact contribution?

▼ **Work through the solution**

The ratio of approximate to exact is $\sqrt{1+(s/a)^2}$. Requiring it below 1.01 gives

$$a/s > 1/\sqrt{1.01^2 - 1} = 7.05.$$

The full length must exceed roughly 14 times the perpendicular distance. Even a section meeting this condition still requires an estimate of the return path field.

Magnetism, induction and field dynamics · 2. The vector potential as an instrument for computing flux

An experimenter can measure a flux change with a coil without measuring the magnetic field over every square millimetre of the coil area. The vector potential gives a corresponding computational tool. Its line integral determines flux, while the freedom to choose its gauge can turn a difficult boundary problem into a manageable one.

Separate local representation from global information

In a small simply connected patch with no holes a sufficiently smooth divergence-free magnetic field has a potential $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. Taking a divergence gives zero automatically because mixed Cartesian derivatives commute. For a closed curve bounding an oriented surface within the patch, Stokes's theorem gives

$$\oint \boldsymbol{A} \cdot d\boldsymbol{\ell} = \int \boldsymbol{B} \cdot d\boldsymbol{a}.$$

The dimensions are useful: A has units of tesla metres or webers per metre; A integrated along a curve has units of webers, whereas B is flux per area. Neither a single component nor the pointwise magnitude of $\boldsymbol{A}$ equals a force per charge.

Add the gradient of any smooth single-valued scalar χ to $\boldsymbol{A}$. Its curl vanishes and its integral around any closed loop is zero, so this changes neither flux nor local force. In a time-dependent problem one must at the same time change ϕ by $-\partial_t \chi$, since $\boldsymbol{E} = -\nabla\phi - \partial_t \boldsymbol{A}$. This coupled transformation cancels the two derivatives of χ. It permits moving an electric field contribution between the scalar and vector formulas while leaving their sum intact. A gauge choice simplifies the equations for the potentials; grounding a conductor instead imposes a physical condition on a voltage measurement.

Derive the free-space Coulomb-gauge potential

Substitute $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ into Ampère's equation. The double curl gives

$$\nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A} = \mu_0 \boldsymbol{J}.$$

Choose $\nabla \cdot \boldsymbol{A} = 0$, called the Coulomb gauge. Starting from any initial $\boldsymbol{A}$ one can impose this condition by solving $\nabla^2 \chi = -\nabla \cdot \boldsymbol{A}$ with appropriate boundary behaviour. A further gauge addition is possible if χ is harmonic and preserves those conditions. With localised steady currents and decay at infinity the resulting solution is

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \int \frac{\boldsymbol{J}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\, d^3 r'.$$

The scalar-looking kernel acts separately on the Cartesian components of the vector current. Verify the gauge explicitly by taking the divergence under the integral. Because $\nabla G = -\nabla' G$, integrating by parts leaves a term proportional to $\int G\, \nabla' \cdot \boldsymbol{J}\, d^3 r'$ and a surface current term at infinity. Both vanish for the specified steady localised source. Taking the curl instead gives $\nabla G \times \boldsymbol{J}$, which is precisely the Biot–

Savart integrand after reversing the cross-product order. For a curved wire, integrating its directed elements with the scalar kernel can be simpler than integrating three different projected field components.

Begin with a uniform field and two different potentials

For $\boldsymbol{B} = B_0 \hat{\boldsymbol{z}}$, the choice $\boldsymbol{A}_1 = B_0 x \hat{\boldsymbol{y}}$ immediately yields

$$B_z = \partial_x A_y.$$

A symmetric construction uses $\boldsymbol{A}_2 = \boldsymbol{B} \times \boldsymbol{r}/2 = B_0(-y\hat{\boldsymbol{x}} + x\hat{\boldsymbol{y}})/2$. Both have zero divergence, so the Coulomb condition alone has not fixed them. Their difference is $\nabla(-B_0 x y/2)$. A uniform imposed field does not decay at infinity, so no contradiction with the free-space uniqueness statement arises.

A rectangle has vertices $(0,0), (a,0), (a,b), (0,b)$. In the first gauge only the right edge contributes $B_0 a b$ to a counterclockwise line integral. In the symmetric gauge that edge contributes half and the upper edge, traversed leftward, contributes the other half. For $a = 0.20\,\mathrm{m}, b = 0.080\,\mathrm{m}, B_0 = 0.030\,\mathrm{T}$, both return $4.80 \times 10^{-4}\,\mathrm{Wb}$. If the field doubles smoothly over 0.040 s, average emf around the stationary loop is $-0.0120\,\mathrm{V}$. Trying to assign the measured emf exclusively to the right edge just because one static gauge makes its line integral special would confuse a chosen potential representation with a physical local electric field.

Work around the axis of a long solenoid

Consider an ideal long solenoid of radius a, with axial field B_0 inside and negligible external field. Rotational symmetry suggests $\boldsymbol{A} = A_\varphi(s)\hat{\boldsymbol{\varphi}}$. A circle of radius $s < a$ encloses flux $\pi s^2 B_0$; a circle outside encloses the entire $\pi a^2 B_0$. Dividing by its circumference gives

$$A_\varphi(s) = \begin{cases} B_0 s/2, & s < a, \\ B_0 a^2/(2s), & s > a. \end{cases}$$

The two expressions agree at $s = a$. In cylindrical coordinates $B_z = s^{-1}\partial_s(s A_\varphi)$; it gives the desired constant inside and zero outside. Nevertheless every exterior loop encircling the solenoid has a nonzero integral. One cannot remove this exterior potential by a globally single-valued scalar in that punctured domain. Locally A_φ equals a constant times the gradient of the polar angle, but the angle changes by 2π upon one circuit. A small loop entirely outside and not surrounding the axis has zero integral. These two experiments distinguish a local zero curl from a global zero circulation.

At $B_0 = 12\,\mathrm{mT}$, $a = 10\,\mathrm{mm}$, the enclosed flux is $3.77\,\mu\mathrm{Wb}$. At $s = 20\,\mathrm{mm}$, $A_\varphi = 3.00 \times 10^{-5}\,\mathrm{T\,m}$; at twice the radius it is half as large but its loop length is twice as large. If B_0 changes in time, the exterior loop can carry a nonzero induced emf even where its wire sits in nearly zero magnetic field. The change of enclosed flux determines it. For a finite solenoid the returning flux eventually changes the large-radius behaviour, so extending the ideal $1/s$ result to all distances would describe the wrong source.

Calculate the coupling of two separate circuits

Drive a slender closed loop 1 with steady current I_1. Its vector potential at loop 2 is $\mu_0 I_1/(4\pi)\oint_1 d\boldsymbol{\ell}_1/R$. Integrating along the second loop yields the Neumann integral $M = \Phi_{21}/I_1 = [\mu_0/(4\pi)]\oint_1\oint_2 d\boldsymbol{\ell}_1\cdot d\boldsymbol{\ell}_2/R$. The dot product and distance are symmetric under interchange, so $M_{21} = M_{12}$ in this reciprocal vacuum setting. Reversing either chosen winding orientation changes the sign; reversing both restores it. The two loops must stay distinct for this simple filament integral. Sending loop 2 onto loop 1 creates a singular self-term. A finite wire radius and the interior current distribution enter a physical self-inductance.

Approximate a small pickup as sampling a uniform axial field from a large drive loop. Let $a_1 = 50\,\mathrm{mm}$, pickup radius $a_2 = 5\,\mathrm{mm}$, and separation $z = a_1$. The drive produces

$$B_z/I_1 = \mu_0/(2a_1\,2^{3/2}) = 4.44\,\mu\mathrm{T/A}.$$

The flux per ampere through a single pickup turn is area πa_2^2 times this field, or $0.349\,\mathrm{nH}$. Fifty tightly wound pickup turns give about $17.45\,\mathrm{nH}$. A changing drive of $2.0\times10^4\,\mathrm{A/s}$ induces about $0.349\,\mathrm{mV}$ in magnitude. The approximation uses the small pickup radius relative to the scale on which the drive field varies. Increasing the pickup radius tenfold calls for actual area integration.

Compare linkage, reciprocity and stored energy for coupled coils

Calculate internal and external inductance in a coaxial return

Flux linkage can be subtle even for the simplest straight conductor once the current has a finite cross section. Consider a long coaxial pair with inner radius a and an ideal thin return shell at b. Work well away from the ends. For equal and opposite currents the exterior field vanishes. In the annulus $a < s < b$,

$$B_\varphi = \mu_0 I/(2\pi s).$$

A longitudinal rectangular surface spanning the two conductors and length ℓ has flux $\Phi = \ell\int_a^b B_\varphi\,ds$. All the inner current is enclosed at every radius in that annulus, so its external flux linkage per unit length is $(\mu_0 I/2\pi)\ln(b/a)$. The corresponding inductance per unit length has units H/m and is finite because the return sets b.

In terms of the potential $B_\varphi = -dA_z/ds$, so the same answer is $\ell[A_z(a) - A_z(b)]$ for the oriented rectangle. A constant added to A_z cancels between the two longitudinal edges. There are no end contributions for a purely longitudinal potential along radial closing segments. Only the potential difference matters. This example displays which pieces of the earlier infinite-wire result are measurable. Extending the return to ever larger radius makes the linkage grow logarithmically and eventually invalidates the approximation of infinitely long straight sections, since real end connections cannot be neglected at arbitrarily large radii.

If the inner conductor carries uniform steady current, $B_\varphi(s) = \mu_0 I s/(2\pi a^2)$ inside it. The internal magnetic energy per length follows directly by integrating over rings:

$$U_{\mathrm{int}}/\ell = \int_0^a \frac{B_\varphi^2}{2\mu_0}\, 2\pi s\, ds = \frac{\mu_0 I^2}{16\pi},$$

$$L_{\mathrm{int}}/\ell = \frac{\mu_0}{8\pi}.$$

The unexpectedly radius-independent internal result is an instructive scaling check. At fixed current a smaller radius raises the typical internal magnetic-field magnitude inversely with radius, but the area in which energy is stored falls as the square of radius. Those factors cancel. This coefficient applies to a uniformly conducting solid inner wire and an ideal sheet return. A solid outer conductor with finite thickness has its own internal energy, and at high frequency skin effect progressively removes the approximately uniform low-frequency current distribution.

A naive internal flux integral $\int_0^a B_\varphi ds$ would give twice the correct linkage. The reason appears by following current filaments: flux in a shell at radius s links only the fraction $I_{\mathrm{enc}}/I = s^2/a^2$ of the inner current on the relevant side. Weighting by that fraction yields

$$\int_0^a (s^2/a^2) B_\varphi ds = \mu_0 I/(8\pi).$$

This is also the answer from the energy derivative and gives an independent interpretation of partial internal linkages. For $a = 1.0\,\mathrm{mm}, b = 4.0\,\mathrm{mm}$, the external contribution is $0.2773\,\mu\mathrm{H/m}$ and the inner contribution $0.0500\,\mu\mathrm{H/m}$. In a 2.0 m low-frequency section carrying 1.0 A, about $0.327\,\mu\mathrm{J}$ of magnetic energy is stored with these approximations. Inferring the geometry from a measured value without subtracting the internal part would bias the inferred logarithmic radius ratio.

Recover the transverse part of a computed potential

A numerical solution may represent an otherwise exact magnetic field by a potential containing substantial gradient contributions. An error estimate based on comparing two raw vector-potential arrays can therefore be misleading. For a periodic calculation choose a nonzero Fourier wavevector $\boldsymbol{k}$. The curl depends only on the component of $\widehat{\boldsymbol{A}}$ perpendicular to $\boldsymbol{k}$, since $i\boldsymbol{k} \times \widehat{\boldsymbol{A}}_{\parallel} = 0$. A Coulomb-gauge representative is obtained by subtracting the projection $\boldsymbol{k}(\boldsymbol{k} \cdot \widehat{\boldsymbol{A}})/k^2$. This is precisely the gradient that the scalar Poisson gauge-fixing problem removes. A constant zero mode needs its own convention since both the derivative and this formula vanish or become undefined at $\boldsymbol{k} = 0$.

Take coordinates in metres, $\boldsymbol{k} = (2, 1, 0)\,\mathrm{m}^{-1}$, and a real potential $\boldsymbol{A} = (3, 1, 0)\cos(2x + y)\,\mathrm{mT\,m}$. The projection amplitude is $(2.8, 1.4, 0)\,\mathrm{mT\,m}$. Removing it leaves $(0.2, -0.4, 0)\,\mathrm{mT\,m}$. Directly differentiating in either representation gives

$$B_z = \partial_x A_y - \partial_y A_x = +\sin(2x + y)\,\mathrm{mT}.$$

The potential arrays differ by a component more than six times as large in norm as the retained one, but the predicted magnetic field agrees. A solver that tracks flux through a physical pickup instead of one gauge component would recognise this equivalence.

This Fourier argument does not turn every longitudinal component of the physical electric field into a gauge artifact. Physical $\boldsymbol{E}$ combines both potentials and its longitudinal projection is constrained by charge. A shift of $\boldsymbol{A}$ in time must be accompanied by the scalar shift derived earlier. An independent check on a computed transient is to form both $-\nabla\phi - \partial_t \boldsymbol{A}$ and $\nabla \times \boldsymbol{A}$, then apply exactly the same discrete differentiation conventions as were used in the field solver. Discrete gradient, curl and divergence operators built on incompatible grids need not cancel as their continuum identities do. If two gauges yield slightly different measured voltages after discretisation, refining the mesh and comparing closed-loop fluxes helps separate differentiation error from genuinely different imposed boundary data.

Follow a charged particle through a time-dependent change of gauge

An electric field in the central region between large closely spaced electrodes can be approximately uniform over a particle trajectory. Let its measured value be $E_x(t) = E_0 \cos \omega t$, and neglect magnetic deflection over the short central path. A scalar-potential description takes $\phi = -x E_0 \cos \omega t$ and $\mathbf{A} = 0$. The electric force then appears to arise entirely from a spatial derivative. Can the same experiment be described with a vanishing scalar potential? The gauge rule answers this question constructively. To make $\phi' = \phi - \partial_t \chi = 0$, choose a function whose time derivative is the given scalar potential. Integrating at fixed position gives

$$\chi(x,t) = -\frac{x E_0}{\omega} \sin \omega t,$$

$$A'_x = -\frac{E_0}{\omega} \sin \omega t,$$

$$\phi' = 0.$$

The curl of this spatially uniform vector potential vanishes. Its negative time derivative gives precisely the measured electric field. A spatially uniform oscillating electric field with no magnetic field everywhere in unbounded source-free space would violate the Maxwell–Ampère law. The electrode example uses a central-region approximation whose accuracy depends on dimensions and frequency. Changing gauge inside that approximation leaves its range of validity exactly where it was. Retaining the small induction fields of a real device also allows gauge transformations of the more complete potentials.

A particle of charge q and mass m, released from rest at $t = 0$, has kinetic momentum $m v_x = (q E_0/\omega) \sin \omega t$. Obtain this result first by integrating the physical force. In the new potentials the canonical momentum is $p_x = m v_x + q A'_x$, which vanishes at all times for this initial condition. The Lagrangian $L = m v^2/2 + q \mathbf{v} \cdot \mathbf{A} - q\phi$ in these variables has no explicit x dependence. Its Euler–Lagrange equation indeed says that p_x is conserved. In the original potentials $p_x = m v_x$ changes because the scalar potential depends on x. Both descriptions assign the same speed, trajectory and work to the particle, even though the conserved canonical expression looks different.

This agreement follows for any path, without assuming the particle followed the physical trajectory. Substitute $\mathbf{A} + \nabla\chi$ and $\phi - \partial_t \chi$ into the Lagrangian. Its change is $q(\mathbf{v} \cdot \nabla\chi + \partial_t \chi) = q\, d\chi/dt$. Integrated between fixed spacetime endpoints it contributes a fixed boundary term to the action. Varying the path between those endpoints leaves the equations of motion unchanged. The extra canonical momentum $q\nabla\chi$ is the derivative of that boundary term with respect to the endpoint position. Keeping the

canonical momentum numerically fixed when comparing two gauges would accidentally assign different physical initial velocities.

For a singly charged ion of mass $6.64 \times 10^{-27}\,\mathrm{kg}$, take $E_0 = 2.0\,\mathrm{V/m}$ and $f = 100\,\mathrm{kHz}$. At one quarter cycle the impulse is $5.10 \times 10^{-25}\,\mathrm{kg\,m/s}$ and the speed is about $76.8\,\mathrm{m/s}$. The displacement follows by a second integration,

$$x(t) - x(0) = qE_0[1 - \cos(\omega t)]/(m\omega^2).$$

It is about $0.122\,\mathrm{mm}$ at the same instant. The geometry ought therefore to supply an approximately uniform region comfortably longer than a few tenths of a millimetre. At half a cycle the velocity is again zero. The particle has still moved because its velocity remained positive throughout that interval.

Change a vector potential while holding the magnetic field fixed

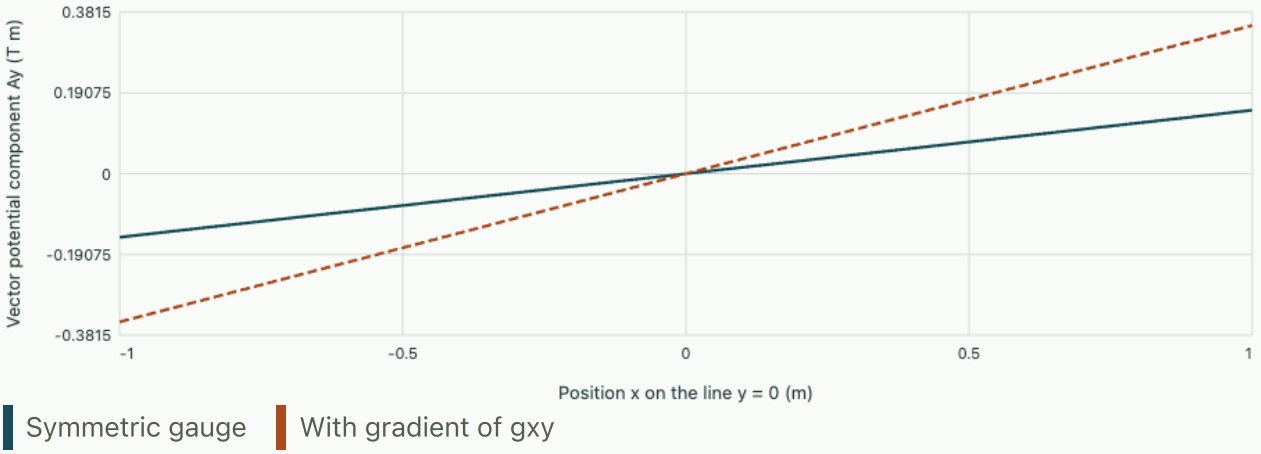
Add a gradient to the potential of a uniform magnetic field. Compare the change of a local component with the circulation around a closed rectangle.

Open this experiment online

Try the calculation

Uniform magnetic field B (T) **0.3**

Gauge coefficient g (T) **0.2**



∂xAy = 0.35 T, ∂yAx = 0.05 T. Curl = 0.3 T. Both loop integrals around 0.30 × 0.20 m are 0.018 Wb.

Preserve gauge invariance in a numerical circulation

One can check gauge invariance without differentiating noisy data twice. Store each oriented edge integral $a_{ij} = \int_i^j \mathbf{A} \cdot d\mathbf{l}$ on a mesh. Adding a scalar gauge function specified at the vertices changes an edge integral by precisely $\chi_j - \chi_i$. Summing around a face gives four such differences on a rectangular grid.

Each vertex occurs once as an end and once as a beginning. The terms cancel exactly, including on irregularly shaped faces, provided that the last edge returns to the first vertex and shared edges use opposite signs for adjacent faces. The result is a simple test for a discrete curl operator that should survive grid refinement unchanged.

A rectangular face has consecutive gauge values $0, 2, 5, -1$ in units of $\mu\mathrm{Wb}$. A program stores initially zero edge integrals. What should it write after the gauge transformation, and what does the sum measure?

▼ **Work through the solution**

Walking around the oriented boundary yields changes of $+2, +3, -6, +1\,\mu\mathrm{Wb}$. Each value is allowed to change independently, but their signed sum is zero. That sum is the magnetic flux through the face by Stokes' theorem. Reversing the boundary should negate every contribution and still give zero. If a neighbouring face uses the same stored value with the same sign for their shared edge, interior contributions will survive in the total flux of a patch. Such a result indicates a mismatch of edge orientation even if some perfectly symmetric test configurations accidentally give zero.

Why does division by the edge length matter when using vector components exported by another code?

▼ **Work through the solution**

The discrete cancellation applies directly to line integrals. A stored component tangential to an edge has units of potential per unit length in this representation and must first be multiplied by its edge length, or integrated with a suitable quadrature if it varies. Adding raw component values around a stretched cell weights physically different lengths equally. A gauge test on a uniform mesh can conceal this error. Repeat it on cells of unequal lengths and compare fluxes in webers.

David Tong, Electromagnetism and Relativity, on potentials and gauge transformations.

Tutorial on gauge and flux

An infinite axial wire yields

$$B_\varphi = \mu_0 I/(2\pi s).$$

Choose a vector potential of the form $A_z(s)\hat{\boldsymbol{z}}$. Determine the difference $A_z(2s) - A_z(s)$, and explain why $A_z(\infty) = 0$ cannot be imposed.

▼ **Work through the solution**

The cylindrical curl gives $B_\varphi = -dA_z/ds$, so $A_z(s) = -(\mu_0 I/2\pi)\ln(s/s_0)$. The difference is $-(\mu_0 I/2\pi)\ln 2$, independent of s_0. An infinite wire is not a localised current with a remote return. Its logarithm cannot both fit the field and tend to zero at infinity. A coaxial return supplies a physically meaningful outer length.

A closed circle outside the ideal solenoid is deformed into an elongated loop without crossing the solenoid. The computed circulation changes by 8 percent. What three physical or numerical hypotheses

would you test?

▼ **Work through the solution**

Under the ideal conditions the enclosed flux stays fixed for deformations entirely through field-free space. Check whether the solenoid has significant end fields, whether the computed geometry changes the winding number, and whether the line quadrature resolves the parts passing closest to the axis. A change of a smooth single-valued gauge cannot change that closed integral.

Take $\chi = \alpha x t$ in a field-free configuration initially represented by $\phi = 0, \boldsymbol{A} = 0$. Recover the fields after transformation.

▼ **Work through the solution**

The new potentials are

$$\phi' = -\alpha x, \ \boldsymbol{A}' = \alpha t \hat{\boldsymbol{x}}.$$

The static-looking scalar gradient supplies $+\alpha\hat{\boldsymbol{x}}$ and the negative vector time derivative supplies $-\alpha\hat{\boldsymbol{x}}$, cancelling. The curl of the spatially uniform vector potential vanishes. A scalar voltage difference alone need not be the complete physical electric field in a time-dependent gauge.

Magnetism, induction and field dynamics · 3. Which features of a current distribution survive at a distance?

A current loop just fits behind a probe mount. At a nearby point its size and placement matter. Far away the same winding behaves as though all that geometry were a single vector. Predicting when that replacement is accurate, and what remains when the vector cancels, requires an expansion of the source rather than an appeal to a field-line sketch.

Expand the kernel of a localised source

Take a bounded steady current distribution of size a near the origin, with observation distance $r \gg a$. The Coulomb-gauge potential is $\boldsymbol{A} = (\mu_0/4\pi)\int \boldsymbol{J}(\boldsymbol{r}')/|\boldsymbol{r}-\boldsymbol{r}'|\,dV'$. Taylor expand the scalar kernel about the observation point while holding the source coordinates fixed. Since $\partial_i(1/r) = -r_i/r^3$, the first two terms are $1/r + \boldsymbol{r}'\cdot\boldsymbol{r}/r^3$. Terms one order farther involve quadratic products $r'_i r'_j$ and second derivatives of $1/r$. For a source entirely inside a sphere of radius a the expansion is systematically useful beyond that sphere; a probe near a wire can violate the required scale separation even at an apparently large distance from the centre of the apparatus.

The leading vector coefficient would be $\int J_i\,dV'$. But $J_i = \partial'_j(r'_i J_j)$ for a divergence-free steady source. Integrating converts it to an outer surface term that vanishes for localised currents. For a line loop the same fact reads $\oint d\boldsymbol{\ell}' = 0$. A finite source with a remote return omitted can give a spurious nonzero leading term. The disappearance of the $1/r$ magnetostatic potential here follows from charge conservation and closed currents. A radiating time-dependent source is governed by a retarded integral with a different far-field expansion.

Identify the magnetic moment without assuming a circular loop

Define $T_{ij} = \int r'_i J_j\,dV'$. The integral of $\partial'_k(r'_i r'_j J_k)$ is zero at large distance, giving $T_{ij} + T_{ji} = 0$. Thus only its antisymmetric part survives, and three independent components remain. Define $\boldsymbol{m} = \frac{1}{2}\int \boldsymbol{r}'\times\boldsymbol{J}\,dV'$. Contracting its components with the observation vector converts the first surviving potential into

$$\boldsymbol{A}_{\mathrm{dip}} = \frac{\mu_0}{4\pi}\frac{\boldsymbol{m}\times\boldsymbol{r}}{r^3},$$

$$\boldsymbol{B}_{\mathrm{dip}} = \frac{\mu_0}{4\pi r^3}[3(\boldsymbol{m}\cdot\hat{\boldsymbol{r}})\hat{\boldsymbol{r}} - \boldsymbol{m}] \quad (r>0).$$

To take this curl, apply the product rule to $r^{-3}(\boldsymbol{m}\times\boldsymbol{r})$. The curl of the second factor is $2\boldsymbol{m}$, while $\nabla r^{-3} = -3\boldsymbol{r}/r^5$. The vector triple product in the other term leaves the radial projection and subtracts three copies of the moment term, giving the expression above. It describes exterior points of the actual source; replacing the source by an ideal point dipole at its own location needs distributional terms. The moment is origin-independent for a closed current because shifting the origin adds a constant vector crossed with $\int \boldsymbol{J}\,dV' = 0$.

For a planar loop $\boldsymbol{m} = (I/2)\oint \boldsymbol{r}'\times d\boldsymbol{\ell}' = I\boldsymbol{a}$, where the oriented area follows from projecting onto each Cartesian plane. For a polygon one may sum oriented triangles from any chosen origin; the edges to the origin cancel between adjacent triangles. A nonplanar closed wire still has an oriented projected area

vector, although it has no unique scalar enclosed area. Stacking N identical turns multiplies the vector by N. This definition makes a measured moment a current times a geometrical area, with units $\mathrm{A\,m^2} = \mathrm{J/T}$.

Use distance and orientation to determine a moment

A 40-turn circular winding of radius $15\,\mathrm{mm}$ carries $0.20\,\mathrm{A}$. Its moment is

$$40(0.20)\pi(0.015)^2 = 5.65 \times 10^{-3}\,\mathrm{A\,m^2}.$$

At $r = 0.15\,\mathrm{m}$ the axial point-dipole field is

$$2(\mu_0/4\pi)m/r^3 = 0.335\,\mu\mathrm{T};$$

at an equatorial point the same distance away it is $0.168\,\mu\mathrm{T}$ directed opposite the moment. A scalar probe looking only at magnitudes would find a two-to-one change. A three-axis probe can use the sign reversal of the axial component to distinguish a dipole from a residual uniform background.

The exact axial field of the circular loop is $B_z = (\mu_0 N I a^2/2)(a^2+z^2)^{-3/2}$. Divide it by the dipole value to obtain $[1+(a/z)^2]^{-3/2}$. At ten radii it is 0.9852 of the dipole prediction, giving an error near 1.5 percent despite the large separation. Reducing that error below one percent relative to the exact value requires $[1+(a/z)^2]^{3/2} - 1 < 0.01$, or $z/a > 12.3$. A far-field fit to the moment using several sensor distances can examine how the fitted $z^3 B_z$ approaches a constant. It also exposes uncertainty in the assumed source centre: a small offset in z produces a relative distance error amplified roughly threefold in the inferred moment.

Arrange two loops to cancel the leading term

Place two identical coaxial loops at $z = \pm d/2$ with moments $+m\hat{\boldsymbol{z}}$ on the positive side and $-m\hat{\boldsymbol{z}}$ on the negative side. At a far positive point each small loop contributes its axial dipole field with its own signed moment. Define $K = \mu_0/(2\pi)$. Their sum, $Km[(z-d/2)^{-3} - (z+d/2)^{-3}]$, has no z^{-3} term. Taylor expansion to first order in d/z gives $3Kmd/z^4$. This quadrupolar decay is steeper because probing from farther away makes the spatial separation of the opposed moments harder to resolve. The result applies when both d/z and individual loop radius divided by z are small.

Near the midpoint a different approximation is needed. The exact axial fields of the two finite loops cancel at $z = 0$ but their derivatives add. For single-turn loops of radius a, differentiating the exact expression gives $B_z'(0) = 3\mu_0 I a^2 d/[2(a^2 + d^2/4)^{5/2}]$ in the stated orientation. With $a = d = 0.040\,\mathrm{m}$, $I = 1.0\,\mathrm{A}$, this is $6.74 \times 10^{-4}\,\mathrm{T/m}$. A 2.0 mm displacement from the centre then produces approximately $1.35\,\mu\mathrm{T}$. A gradiometer uses differences like these to suppress an approximately uniform background. A slight current imbalance gives a residual ordinary dipole, which eventually dominates at sufficiently large distance because it falls more slowly.

Measure torque and work as an independent calibration

A small loop in a uniform external field has forces $I\,d\boldsymbol{\ell}\times\boldsymbol{B}$ on its sides. Their net force vanishes since $\oint d\boldsymbol{\ell}=0$. The net torque survives: reducing $\oint \boldsymbol{r}\times(I\,d\boldsymbol{\ell}\times\boldsymbol{B})$ with the same antisymmetric geometrical tensor yields

$$\boldsymbol{\tau}=\boldsymbol{m}\times\boldsymbol{B}.$$

With a fixed permanent moment or a maintained current in an imposed field, the mechanical potential for orientation is $-\boldsymbol{m}\cdot\boldsymbol{B}$ up to a constant. If an external source maintains the current its electrical work matters in the total energy balance.

The winding above placed in $B=0.030\,\mathrm{T}$ at 30° experiences $8.48\times10^{-5}\,\mathrm{N\,m}$. Turning it quasistatically from alignment to perpendicular increases this orientation energy by $mB=1.70\times10^{-4}\,\mathrm{J}$. A force measurement in a slowly varying field samples the gradient: along an axis for a fixed aligned moment $F_z\simeq m\,dB_z/dz$. In a truly uniform field the dipole has torque but no translational force. A convenient test of an inferred field gradient is therefore to compare the force at the same point with moment parallel and antiparallel. Their difference $2mB_z'$ rejects a constant support force. Magnetic susceptibility, moment reorientation or a large source spanning appreciable changes in gradient invalidate the rigid-point formula.

Calibrate a gradiometer without assuming the background vanishes

A paired pickup consists of two equal small sensing loops connected in opposition. Put their centres at $z=z_0\pm l/2$, both normal to the same axial field component $B(z)$, and use N_p turns of area A_p for each. In the small-loop limit their net linkage along the series path is $N_pA_p[B(z_0+l/2)-B(z_0-l/2)]$. The spatially uniform part cancels algebraically. Expanding about the midpoint leaves $N_pA_pl[B'(z_0)+l^2B'''(z_0)/24+\cdots]$. A finite separation therefore estimates the derivative averaged over the baseline. The cubic derivative, rather than the second derivative, sets its first truncation error because the measurement is centred.

Suppose a calibration winding produces a sinusoidal gradient of peak amplitude $2.0\,\mathrm{mT/m}$ at 100 Hz across that baseline. With $N_p=50,\ A_p=2.0\times10^{-4}\,\mathrm{m^2}, l=0.030\,\mathrm{m}$, the peak linkage is $6.0\times10^{-7}\,\mathrm{Wb}$. Its time derivative gives a peak emf $0.377\,\mathrm{mV}$. Changing frequency to 200 Hz while keeping gradient amplitude unchanged should double this emf if pickup impedance and input instrumentation do not distort the measurement. A DC gradient alone produces no voltage in a stationary inductive pickup; a calibrated DC Hall pair or a deliberately moved sample implements a different measurement.

A small fractional area-turn imbalance ε couples a uniform field back into the differenced signal. For example take the positive pickup to have factor $N_pA_p(1+\varepsilon/2)$ and the negative one $N_pA_p(1-\varepsilon/2)$. Expansion now contains both lB' and εB. If a sinusoidal background of amplitude $20\,\mu\mathrm{T}$ shares the drive frequency and $\varepsilon=0.01$, its apparent gradient is

$$\varepsilon B/l=6.67\,\mu\mathrm{T/m}.$$

This is only a third of one percent of the calibration gradient but could dominate a much smaller sample signal. Reversing the sample while leaving the pickup fixed tests which contribution reverses. Rotating the whole assembly might also change its coupling to external wiring and would implement a less selective test.

For an axial dipole field $B(z) = C/z^3$ sufficiently far from the loop, $B' = -3C/z^4$ and $B''' = -60C/z^6$. The relative leading baseline error is therefore $5l^2/(6z_0^2)$. At $l/z_0 = 0.30$, the differential reading overestimates the pointwise derivative magnitude by about 7.5 percent before accounting for finite sensing-loop area. A gradient device may be well calibrated in a nearly linear reference field and still give a biased point-gradient interpretation near a sample. Comparing measurements at two baselines tests this spatial averaging while retaining rejection of the common background.

Distinguish finite pickup area from multipole structure

Even a single coil usually records surface flux rather than the field at one mathematical point. For a dipole $m\hat{\boldsymbol{z}}$ on the symmetry axis of a coaxial circular pickup of radius b, one can calculate the flux exactly within the point-dipole approximation by integrating the vector potential along its boundary. The azimuthal potential is $A_\varphi = \mu_0 m s/[4\pi(s^2+z^2)^{3/2}]$. Multiplication by $2\pi b$ gives

$$\Phi(z) = \frac{\mu_0 m b^2}{2(b^2+z^2)^{3/2}}.$$

Compare with the point-field estimate $\pi b^2 B_z(0,z) = \mu_0 m b^2/(2|z|^3)$. The exact flux is smaller by the factor $[1+(b/z)^2]^{-3/2}$. This has exactly the same mathematical form as the correction for a finite drive loop and point probe. There is a physical reason: the mutual inductance is reciprocal. A measured deviation from a dipole fit can result from the source size, the pickup size, or both. A measurement at one distance without knowledge of the pickup geometry cannot assign the correction uniquely.

For $m = 1.0 \times 10^{-3}\,\mathrm{A\,m^2}$, $b = 10\,\mathrm{mm}$, $z = 30\,\mathrm{mm}$, the point-field estimate is $2.33\,\mathrm{nWb}$ per turn. The actual dipole flux through that circle is $1.99\,\mathrm{nWb}$, about 15 percent lower. If the pickup is doubled in radius, it does not simply collect four times as much flux at the same distance. In the formal limit $b \to \infty$ the explicit expression falls as $1/b$, tending to zero: the very distant return flux of the dipolar field then cancels the central flux in the whole plane. For fixed nonzero z, differentiating the expression with respect to b gives an extremum at $b = \sqrt{2}|z|$; collecting flux with an ever wider loop eventually reduces the signal.

One can turn that observation into a geometry choice for a weak-signal experiment. First choose a drive frequency slow enough for the moment to remain quasistatic and for the pickup circuit response to be calibrated. Select a pickup radius near the theoretical optimum for the fixed stand-off, then compare an intentionally smaller pickup whose field is more local. Agreement of the source moments inferred with each correct area formula provides evidence for the dipole model over the distances used. Disagreement that persists after correcting alignment and finite pickup radii motivates resolving higher moments or modelling the actual source current, rather than assigning a different magnetic moment to the same sample at each probe distance.

Build a current dipole from a rotating charged shell

Consider a thin spherical shell of radius a carrying charge Q uniformly painted onto its insulating surface. A motor rotates the shell with angular speed Ω about z. Every little patch moves, yet the charge density observed at a fixed point in space stays constant. The steady current density is a surface current $\mathbf{K} = \sigma \Omega a \sin\theta \, \hat{\boldsymbol{\phi}}$, where $\sigma = Q/(4\pi a^2)$. This model assumes that charge remains attached to the shell and that the rim speed is small compared with the speed of light. A conducting shell allowed to rearrange its charges introduces another physical question.

Slice the sphere into narrow latitude bands. The distance across a band perpendicular to its current is $a\,d\theta$, so the circulating current associated with the band is $dI = K_\phi a \, d\theta$. Its loop area is $\pi a^2 \sin^2\theta$. All the band moments point along the rotation axis for positive charge. Summing them gives

$$m_z = \pi \sigma \Omega a^4 \int_0^\pi \sin^3\theta \, d\theta = \frac{4\pi\sigma\Omega a^4}{3} = \frac{Q\Omega a^2}{3}.$$

The integral can be checked by substituting $u = \cos\theta$, which changes $\sin^3\theta\,d\theta$ into $-(1-u^2)\,du$. The equatorial bands dominate much more strongly than their charge alone suggests. Each patch there travels faster and encloses a larger area. Using the full charge as though it were on an equatorial ring would give $Q\Omega a^2/2$, an overestimate by a factor of three halves. Both models have the same radius and total charge, so the difference tests precisely what information a magnetic dipole moment contains.

There is a useful local check. For the latitude ring at angle θ, its radius squared is $a^2\sin^2\theta$ and the distance squared from every source point to the sphere centre is a^2. Insert these into the exact axial field of a loop. The result is

$$dB_z = \mu_0 dI \sin^2\theta / (2a).$$

Integrating gives

$$B_z(0) = 2\mu_0 \sigma \Omega a / 3 = \mu_0 Q \Omega / (6\pi a).$$

This finite answer at the origin cannot be obtained by extrapolating a point-dipole field inwards. A uniformly magnetised sphere with $\mathbf{M} = \sigma\Omega a\,\hat{z}$ has exactly this surface current. Its solution gives a uniform interior field and an exact dipolar exterior because its angular dependence is a single degree-one harmonic.

For $a = 50\,\mathrm{mm}$, $Q = 1.0\,\mu\mathrm{C}$ and $\Omega = 100\,\mathrm{rad/s}$, the dipole moment is $8.33 \times 10^{-8}\,\mathrm{A\,m^2}$. The central field is about $0.133\,\mathrm{nT}$. On axis at $200\,\mathrm{mm}$, the field is only $2.08\,\mathrm{pT}$. Separating a field that weak from an ambient steady field requires a mechanism such as reversing the rotation or the charge repeatedly and averaging synchronous differences. Rotor vibration in a spatially varying background can generate a signal correlated with rotation, too. An uncharged control run tests that contribution without changing the motor geometry.

Distinguish a local current from its far-field moment

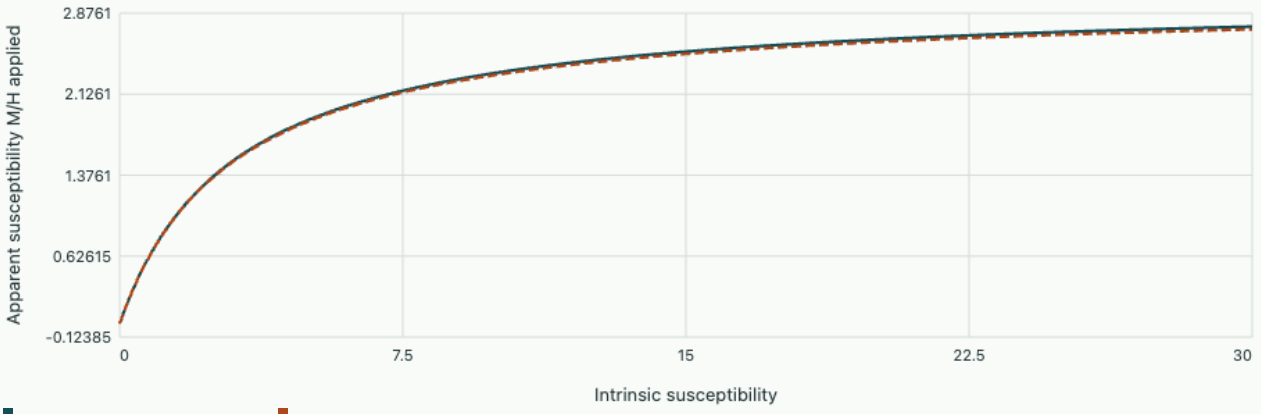
Explore the field generated by a magnetised body. Identify which exterior measurements are still sensitive to the detailed current distribution.

Open this experiment online

Try the calculation

Demagnetising factor N **0.33**

Intrinsic susceptibility **9**



Selected ellipsoid Sphere N = 1/3

At the selected susceptibility the apparent value is 2.27; the internal H field is 25.2% of the applied H field.

Compare magnetic moment and angular momentum

Suppose the mass M is distributed over the same shell in the same proportion as its charge. A mass element dm contributes axial angular momentum $dL_z = \Omega a^2 \sin^2\theta\, dm$. A transported charge element dq returns after the period $2\pi/\Omega$, so its contribution to the current moment is

$$dm_z = \Omega a^2 \sin^2\theta\, dq/2.$$

If every element has $dq/dm = Q/M$, these relations give $m_z/L_z = Q/(2M)$ on any rigid axisymmetric body, independent of whether the matter fills a solid ball or only a shell. The simple proportionality is classical and assumes the same spatial distributions. It supplies no derivation of the intrinsic spin moment of an electron.

A solid insulating sphere of radius a carries a uniformly distributed volume charge instead. Recover its moment by adding thin spherical shells and compare the angular momentum ratio.

▼ **Work through the solution**

The charge inside a radial shell is $dQ = 3Qr^2dr/a^3$. Each shell contributes $\Omega r^2 dQ/3$ from the latitude-band result. Integration from zero to a yields $Q\Omega a^2/5$. More charge now occupies small cylindrical radii, which reduces the moment at fixed total charge. If mass is also uniform in volume, the moment of inertia is $2Ma^2/5$. Their ratio is again $Q/(2M)$. If instead essentially all the mass resides in a solid core while the charge remains on an insulating outer shell, inserting the shell charge result and solid inertia gives $5Q/(6M)$, illustrating why the distribution hypothesis matters.

What mechanical measurement independently determines the moment of the shell without resolving its field in space?

▼ **Work through the solution**

Place the rotor in a calibrated approximately uniform external magnetic field at right angles to its axis. The torque has magnitude mB, obtained by summing the loop torques. With the numerical shell above and $B = 0.10\,\mathrm{T}$ it is $8.33 \times 10^{-9}\,\mathrm{N\,m}$. Reverse the charge and repeat the torque measurement to isolate the contribution odd in charge. Maintaining the axis fixed requires a compensating support torque. A nonuniform imposed field would also exert a net force and could couple a spurious lever-arm torque into the measuring fixture.

David Tong, Magnetostatics, on dipole moments, fields and forces.

Tutorial on testing a multipole approximation

Two nominally opposed loops have moment difference $0.01m$ and separation d. Estimate the positive axial distance at which their residual dipole equals their leading quadrupole in magnitude.

▼ **Work through the solution**

The respective magnitudes are $K(0.01m)/z^3$ and $3Kmd/z^4$. Equating them gives $z = 300d$. An imbalance of one percent can therefore determine the most distant field even though a closer scan shows an excellent quadrupolar law.

A nonplanar wire follows three sides of the unit square in the xy-plane, then returns to its start via a raised corner. Can its dipole moment have an in-plane component? Give a construction that determines it without surface integration.

▼ **Work through the solution**

Sum $\frac{I}{2}\sum_j \boldsymbol{r}_j \times \boldsymbol{r}_{j+1}$ over its vertices in traversal order, including the closing edge. Its projections onto the vertical coordinate planes can span nonzero signed area, generating in-plane moment components. A rigid translation of every vertex adds a telescoping zero, providing a check.

Opposed windings produce fields that scale nearly as r^{-4}. By what factor does a small transverse pickup's induced voltage amplitude change if source-to-pickup distance doubles while the whole source current is driven sinusoidally at the same sufficiently low frequency?

▼ **Work through the solution**

If the pickup area and orientation remain fixed, its linked flux scales as its normal field. The time derivative supplies the same factor ω in both runs. Its voltage amplitude therefore falls by about 16, provided the magnetostatic spatial approximation and quadrupolar regime remain valid. A dipole-dominated residual would instead give about eight.

Compare a static cancellation with cancellation of the leading radiating multipole

Magnetism, induction and field dynamics · 4. From microscopic moments to measured magnetisation

A magnetised sample can make a compass turn without a power supply attached. The macroscopic theory assigns it bound currents in its bulk and on its surface. These are calculated sources for the magnetic field. We will reconstruct them from magnetic moments, predict signals from finite samples, and ask what an exterior field measurement can identify.

Define a material field at an intermediate scale

Choose a volume large enough to average over microscopic moments and small enough to resolve the sample's macroscopic variation. Let $\boldsymbol{M}(\boldsymbol{r})$ be the vector sum of moments within it divided by its volume. Its units are $\mathrm{A/m}$. A permanent aligned collection may maintain $\boldsymbol{M}$ without an applied field. An induced collection may point with or against the applied field. Electric charges need not traverse the entire sample in a closed macroscopic wire for either mechanism. Electron orbital moments and spin contribute to the measured moment; for a classical exterior field calculation any small magnetic dipole has the same leading field once its moment is specified.

A bar of volume $8.0\,\mathrm{cm}^3$ with uniform $M = 2.0 \times 10^4\,\mathrm{A/m}$ has total moment $m = 0.160\,\mathrm{A\,m}^2$. A torsion balance in a uniform $10\,\mathrm{mT}$ transverse field would measure a maximum torque $1.60 \times 10^{-3}\,\mathrm{N\,m}$ if its moment stayed fixed. Dividing a measured torque by field and sample volume infers the component normal to the external field. Alignment of a previously soft material could instead alter the very moment being measured. Distinguishing fixed-source and induced-response experiments will matter when we calculate mechanical energy.

Derive the bound sources by integration by parts

Sum the dipole vector potentials over a material region V. With observation point outside the matter, a volume element contributes $(\mu_0/4\pi)\boldsymbol{M}(\boldsymbol{r}') \times (\boldsymbol{r} - \boldsymbol{r}')dV'/R^3$, where $R = |\boldsymbol{r} - \boldsymbol{r}'|$. Note that $\nabla'(1/R) = (\boldsymbol{r} - \boldsymbol{r}')/R^3$. Apply the identity $\nabla' \times (\boldsymbol{M}/R) = (\nabla' 1/R) \times \boldsymbol{M} + (\nabla' \times \boldsymbol{M})/R$. The desired first cross product is minus the first term on the right. A volume integral of a curl is the surface integral of outward normal crossed with the vector. This yields

$$\boldsymbol{A} = \frac{\mu_0}{4\pi}\left[\int_V \frac{\nabla' \times \boldsymbol{M}}{R}\,dV' + \oint_{\partial V} \frac{\boldsymbol{M} \times \boldsymbol{n}'}{R}\,dA'\right].$$

Comparing with the ordinary current potential identifies $\boldsymbol{J}_b = \nabla \times \boldsymbol{M}$, in A/m², and $\boldsymbol{K}_b = \boldsymbol{M} \times \boldsymbol{n}$, in A/m. The first can vanish throughout a uniformly magnetised object, while the second closes around its exterior. A literal set of microscopic loops illustrates the sign. Neighbouring loop-edge currents cancel between equal adjacent moments, whereas their edges on the outer boundary remain uncompensated. When the moments vary slowly, interior cancellation is incomplete and yields the curl.

An alternative treatment extends $\boldsymbol{M}$ by zero outside. Its distributional curl already includes the surface sheet at the discontinuity. Use either this full derivative or the separate bulk and surface integrals; including both descriptions simultaneously doubles the boundary source. At an internal interface with chosen normal from material 1 to 2 the two surface contributions combine to $(\boldsymbol{M}_1 - \boldsymbol{M}_2) \times \boldsymbol{n}$. This expression is unchanged by interchange of the two labels accompanied by reversal of the normal.

Work through a uniformly magnetised cylinder

Let a finite circular cylinder of radius a and length $2l$ have $\boldsymbol{M} = M\hat{\boldsymbol{z}}$. On the curved side $\hat{\boldsymbol{z}} \times \hat{\boldsymbol{s}} = \hat{\boldsymbol{\varphi}}$, hence $K_\varphi = M$; on its end faces the cross product vanishes. The cylinder is equivalent, for its magnetic field, to a finite solenoid having turn-current per unit length M. A strip of sidewall of height dz' carries ring current $M\,dz'$. Insert this into the axial finite-loop field from the first chapter and integrate from $-l$ to l.

$$B_z(z) = \frac{\mu_0 M}{2}\left[\frac{z+l}{\sqrt{a^2+(z+l)^2}} - \frac{z-l}{\sqrt{a^2+(z-l)^2}}\right].$$

At the centre $B_z = \mu_0 M l/\sqrt{a^2+l^2}$. For $a = 10\,\mathrm{mm}$, $l = 20\,\mathrm{mm}$ and $M = 1.0\times10^4\,\mathrm{A/m}$, the value is $11.24\,\mathrm{mT}$, below the long-cylinder limit $12.57\,\mathrm{mT}$. At the end centre $z = l$, only the first term contributes, giving $6.10\,\mathrm{mT}$. A probe just outside the end and one just inside see the same normal $\boldsymbol{B}$ in the limit, even though the auxiliary normal $\boldsymbol{H}$ can jump because $\boldsymbol{M}$ ends abruptly. The field near an edge is not uniform and can vary strongly with the sensor's position. Stating the axis and position is therefore essential when inferring magnetisation from a reading.

Use a nonuniform sample to separate bulk and surface effects

For an infinitely long cylinder set $\boldsymbol{M} = M_0(s/a)\hat{\boldsymbol{\varphi}}$ for $s < a$, zero outside. The cylindrical curl gives $J_{b,z} = s^{-1}\partial_s(sM_\varphi) = 2M_0/a$ in the bulk. The side normal is $\hat{\boldsymbol{s}}$, so $K_{b,z} = -M_0$. Integrated over a cross section the bulk current is $2\pi a M_0$; the axial sheet supplies $-2\pi a M_0$. Their cancellation means an exterior Amperian circle sees zero enclosed current. Its zero field follows from symmetry and the absence of externally imposed fields.

Inside, a circle at s encloses $2M_0\pi s^2/a$. Dividing its circulation by $2\pi s$ gives

$$B_\varphi = \mu_0 M_0 s/a.$$

Thus $\boldsymbol{H} = \boldsymbol{B}/\mu_0 - \boldsymbol{M} = 0$ both inside and outside here, despite nonzero interior $\boldsymbol{B}$. For $a = 20\,\mathrm{mm}$, $M_0 = 500\,\mathrm{A/m}$, the bulk current density is $5.0\times10^4\,\mathrm{A/m^2}$, total bulk current $62.8\,\mathrm{A}$, and field just inside the rim $0.628\,\mathrm{mT}$. A magnetometer outside reports zero within this ideal geometry. Such a null measurement establishes neither zero $\boldsymbol{M}$ nor absence of all microscopic moments.

Compare a field map and a pickup-coil measurement

A stationary inductive pickup near a stationary permanent magnet has no emf from that static flux alone. An experiment can move the sample, rotate it or modulate a known magnetising field to produce $-N\,d\Phi/dt$. A Hall sensor or calibrated magnetometer can instead respond to a static field. To compare methods, place a small pickup near the axis of the preceding finite cylinder and oscillate its axial position as $z = z_0 + \zeta\cos\omega t$. For ζ small, the flux per turn is approximately $A_p[B_z(z_0) + \zeta B_z'(z_0)\cos\omega t]$. The voltage amplitude is $NA_p\zeta\omega|B_z'(z_0)|$. It measures a gradient multiplied by displacement amplitude, whereas a static Hall value measures the field itself.

At the symmetric cylinder centre $B_z'(0) = 0$, so the first harmonic vanishes even though the field is strong. Expanding one more order gives a constant shift and a second harmonic proportional to $B_z''(0)\zeta^2$, since $\cos^2 \omega t = (1 + \cos 2\omega t)/2$. Finding a very small first harmonic at the centre is therefore a check on positioning and symmetry. It is not a calibration of zero magnetisation. Far from the ends, a measurement of the whole sample's dipole field instead yields $\int \boldsymbol{M}\, dV$. Different internal distributions can share that integral while producing different nearer fields. Report the geometry, the actual observable and the assumed uniformity whenever quoting a single recovered value of M.

Examine bound currents and field measurements in a magnetised sample

Use fictitious poles to compute a measurable field

There is a second useful representation of a given magnetisation. In a simply connected region without free current, $\nabla \times \boldsymbol{H} = 0$, so choose $\boldsymbol{H} = -\nabla\psi$. Substitution into $\nabla \cdot (\boldsymbol{H} + \boldsymbol{M}) = 0$ gives

$$\nabla^2\psi = \nabla \cdot \boldsymbol{M}.$$

The potential therefore behaves mathematically as if it were produced by volume density $\rho_m = -\nabla \cdot \boldsymbol{M}$ and surface density $\sigma_m = \boldsymbol{M} \cdot \boldsymbol{n}$. The surface term follows by extending the magnetisation by zero across the boundary, or by requiring continuity of $B_n = \mu_0(H_n + M_n)$. Integrate these densities with kernel $1/(4\pi R)$ to obtain ψ; there is no factor of ϵ_0^{-1} because these conventions assign ψ units of amperes. The resulting $\boldsymbol{H}$ must be converted back to $\boldsymbol{B}$, adding $\mu_0\boldsymbol{M}$ inside matter.

For the uniformly magnetised finite cylinder the pole representation puts densities $+M$ and $-M$ on the two end faces, none on the sides. Earlier the bound-current representation placed current on the curved side and none on the ends. They predict the same physical field. A single uniformly loaded disk of pole density σ_m at $z = 0$ has on-axis scalar potential $\psi = \sigma_m[\sqrt{z^2 + a^2} - |z|]/2$, obtained by integrating $2\pi s\, ds/[4\pi\sqrt{s^2 + z^2}]$. Its derivative gives

$$H_z = (\sigma_m/2)[\operatorname{sgn}(z) - z/\sqrt{z^2 + a^2}].$$

At the disk plane the two one-sided fields jump by σ_m, precisely the normal-H jump demanded by the magnetisation.

At the midpoint of the cylinder add the two oppositely charged disks. Their axial fields both point opposite the magnetisation and total $H_z = -M[1 - l/\sqrt{a^2 + l^2}]$. Adding M yields the same B_z as the side-current integral. At $l = 2a$, this internal demagnetising field is $-0.1056M$; for a very short cylinder $l \ll a$ it approaches $-M$. A uniformly specified magnetisation in a thin disc therefore produces only a small normal B at its centre despite the large M and H terms that nearly cancel. In practice, a soft magnetic material in such a shape may choose a nonuniform magnetisation or domains. The calculation describes the field of the imposed distribution; it does not establish that the distribution is a stable material state.

The 'poles' are convenient mathematical source terms, not isolated magnetic charges: adding the effective volume and surface poles over a finite bounded sample gives

$$-\int \nabla \cdot \boldsymbol{M}\, dV + \oint \boldsymbol{M} \cdot \boldsymbol{n}\, dA = 0.$$

North and south effective poles may be far apart, but a surrounding closed Gaussian surface still has zero net magnetic flux. If a model predicts a $1/r^2$ radial field of one sign through an entire distant sphere around a fully included bar magnet, test whether the opposite face or volume contribution has been accidentally dropped.

Infer the field of a spherical void by subtracting a known sample

Superposition can calculate a local cavity field before introducing a detailed microscopic model. Imagine a large body with uniform imposed magnetisation $\boldsymbol{M}$ near an interior point, and remove a small spherical piece. The magnetic field after removal is the original macroscopic field minus the contribution of the excised sphere, provided the surrounding magnetisation stays unchanged. A uniformly magnetised sphere in otherwise empty space alone has uniform internal $H = -M/3$, as can be derived from its pole distribution $\sigma_m = M\cos\theta$. At the sphere centre $\boldsymbol{H}$ follows by integrating the field from surface elements $-\sigma_m \hat{\boldsymbol{r}}\, dA/(4\pi a^2)$. The z-component involves the spherical average $\langle \cos^2\theta \rangle = 1/3$; transverse contributions vanish by symmetry. The unique dipolar harmonic extends that central value uniformly throughout the sphere.

The sphere's own internal B is therefore $2\mu_0 M/3$. Immediately after its removal the field in the empty cavity is

$$B_{\mathrm{cav}} = B_{\mathrm{bulk}} - 2\mu_0 M/3 = \mu_0 (H_{\mathrm{bulk}} + M/3).$$

This assumes the sphere is small relative to the variation scale of the pre-existing magnetisation and far from the sample's external boundary on its own size scale. A different cavity shape changes the subtraction factor, and a physical bore can rearrange the surrounding magnetisation or admit demagnetising fields from its ends. Local-field interpretations depend on probe geometry as well as material properties.

As a numerical illustration, suppose a permanent or pinned surrounding material has local $M = 1500\,\mathrm{A/m}$ and $H_{\mathrm{bulk}} = 400\,\mathrm{A/m}$, aligned. The macroscopic bulk field is $\mu_0(1900) = 2.39\,\mathrm{mT}$. A small spherical empty cavity under the stated frozen-distribution approximation has $B_{\mathrm{cav}} = \mu_0(900) = 1.13\,\mathrm{mT}$. Both readings can be correct for the same initial material. A Hall element inserted into a drilled hole measures the field in that hole; labelling it 'the internal material field' without modelling the probe cavity can cause a substantial discrepancy. Small hole volume does not by itself make the local cavity field correction small, since area and inverse-square field factors cancel in the spherical surface integral.

For a final comparison, cut a very thin flat slit perpendicular to $\boldsymbol{M}$ instead of a sphere, idealising its lateral dimensions as large relative to its thickness. The facing effective pole sheets produce a strong additional field within the slit. Normal B stays continuous across its faces, so the vacuum reading tends towards the original normal B_{bulk} when edge effects and magnetisation changes are negligible. A long narrow cavity parallel to the magnetisation instead tends towards continuity of tangential H on its long sides. Comparing the sphere and slit is an operational way to understand why B and H are different macroscopic variables rather than two symbols for one ambiguous 'magnetic intensity'.

Read the current implied by a magnetisation

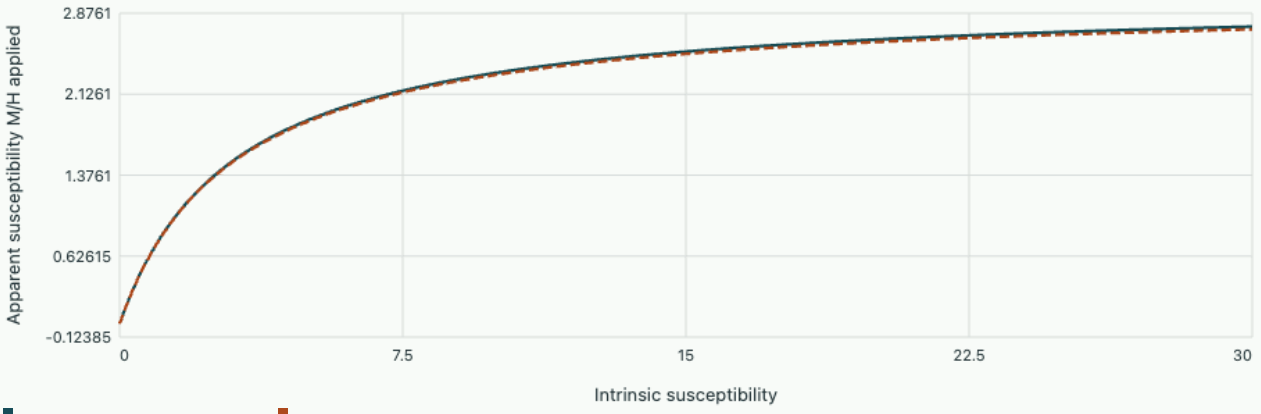
Change the magnetisation and compare interior and exterior fields at the surface. Check which tangential or normal field condition is relevant.

Open this experiment online

Try the calculation

Demagnetising factor N **0.33**

Intrinsic susceptibility **9**



Selected ellipsoid Sphere N = 1/3

At the selected susceptibility the apparent value is 2.27; the internal H field is 25.2% of the applied H field.

Reconstruct an invisible current distribution in a magnetised film

An extended magnetic film fills the slab $-a < z < a$. A fabrication process leaves a tangential magnetisation $M_x(z)\hat{x}$. Let the magnetisation vanish outside. There is no free current in this calculation, and there is no applied field at infinity. A field measurement above the centre of a very wide film can then give nearly zero even when the magnetisation inside varies substantially. Work out both the bound currents and the local magnetic field before deciding what an external zero reading means.

Inside the material the bound volume current is

$$J_{\mathrm{b},y} = dM_x/dz.$$

The upper face has outward normal $+\hat{z}$, so its bound sheet current is $-M_x(a^-)\hat{y}$. On the lower face the outward normal reverses and the sheet current is $+M_x(-a^+)\hat{y}$. Integrating the volume current through the thickness gives $M_x(a^-) - M_x(-a^+)$. Thus the combined sheet and volume currents sum to zero per unit length across the entire film. This cancellation is independent of how much structure the magnetisation has inside the slab. An alternative derivation extends M_x by zero beyond the faces and

differentiates that piecewise function as a distribution. The jumps generate both surface terms with their proper signs.

For an explicit depth profile take $M_x = M_0(1 + z/a)$. There is a uniform bulk current M_0/a, no lower surface current, and an upper surface current $-2M_0$. Planar symmetry allows only $B_x(z)$. Ampère's law gives

$$dB_x/dz = \mu_0 J_y.$$

Integrating from below the film, where the field vanishes, gives $B_x = \mu_0 M_0(1 + z/a)$ inside. At the upper face the field jumps back to zero. The result also follows directly from $H_x = B_x/\mu_0 - M_x$: its tangential component is continuous wherever there is no free surface current, and there is no volume source for its variation. In this geometry $H_x = 0$ throughout.

Take $a = 0.20\,\mathrm{mm}$ and $M_0 = 2.0 \times 10^4\,\mathrm{A/m}$. The interior current density is $1.0 \times 10^8\,\mathrm{A/m^2}$. That large number represents microscopic circulating moments after coarse graining. It does not imply a transport current passing through external electrical contacts. The peak interior field is approximately $50.3\,\mathrm{mT}$ just below the upper face, and the average is about $25.1\,\mathrm{mT}$. An exterior sensor over the central region of the infinite idealisation reads zero at any height. With a finite lateral extent, currents must close around the film edges. The resulting fringe signal contains shape information and becomes measurable as the probe approaches the rim.

Rotate the magnetisation through ninety degrees so that it now points in the z direction while its magnitude keeps the same depth profile. Each local curl vanishes. Each face has parallel magnetisation and normal, giving zero bound sheet current. Consequently the solution has zero B everywhere in the ideal infinite geometry. Within the film $H_z = -M_z$. In the magnetic-pole construction the two surface pole densities and the bulk density $-dM_z/dz$ provide precisely that auxiliary field. These two slab examples illustrate a real inverse problem: very different internal magnetic states can produce the same exterior measurement. Knowing the material orientation or imposing a changed external field supplies additional information.

Can two depth profiles with the same average magnetisation be distinguished by an ideal measurement of the tangential magnetic flux through the film thickness?

▼ Work through the solution

Choose a rectangular sensing surface with its normal along x, its height equal to the film thickness, and its length w along y. The ideal internal flux is $w\mu_0 \int_{-a}^{a} M_x(z)\,dz$. This measurement determines one weighted integral. A constant M_0 and the linear example above give identical results even though the constant profile has only surface currents. A depth-selective measurement would weight the profiles differently. For example, integrating over just the upper half gives $\mu_0 w a M_0$ in the constant case and $3\mu_0 w a M_0/2$ in the linear case. Actual pickup loops must close outside the sample, so include their full geometry when inferring an internal flux.

Suppose the linear profile relaxes uniformly in amplitude after an aligning field is removed. What information would the integrated voltage of a calibrated pickup provide?

▼ **Work through the solution**

If the pickup links N turns with the preceding full-thickness flux and M_0 decays from its initial value to zero, the time-integrated emf has magnitude $2N\mu_0 waM_0(0)$. Faraday's law supplies this integral independently of whether the decay is exponential or has several relaxation times. A finite integration window that ends before zero magnetisation underestimates the initial amplitude unless the residual flux is measured. Leakage from the changing applied alignment field can also drive the loop directly. A reference winding of the same area, measured without the specimen, allows that part of the transient to be estimated.

David Tong, Electromagnetism in Matter, on macroscopic magnetisation and bound currents.

Tutorial on reconstructing material sources

In the nonuniform cylinder let $M_\varphi = M_0(s/a)^3$ instead. Obtain the bulk current density, the side sheet, the total bound current through the full cross section and the interior field.

▼ **Work through the solution**

$$J_{b,z} = s^{-1}\partial_s(M_0 s^4/a^3) = 4M_0 s^2/a^3.$$

At the wall $K_{b,z} = -M_0$. Integration gives $2\pi aM_0$ in the bulk and its opposite on the wall. For an interior circle the enclosed bulk current is $2\pi M_0 s^4/a^3$, giving $B_\varphi = \mu_0 M_0 (s/a)^3$. Both total cancellation and dimensional consistency check the result.

Two adjoining media have tangential magnetisations $400\hat{\boldsymbol{x}}$ and $100\hat{\boldsymbol{x}}\ \mathrm{A/m}$. The normal from 1 to 2 is $+\hat{\boldsymbol{z}}$. Find the net bound sheet current and the associated tangential jump of $\boldsymbol{B}$ if no free surface current is present.

▼ **Work through the solution**

The equivalent sheet is

$$300\hat{\boldsymbol{x}} \times \hat{\boldsymbol{z}} = -300\hat{\boldsymbol{y}}\,\mathrm{A/m}.$$

Tangential H_x is continuous, so

$$B_{2x} - B_{1x} = \mu_0(M_{2x} - M_{1x}) = -0.377\,\mathrm{mT}.$$

Crossing this difference with the chosen normal gives $\mu_0 \boldsymbol{K}_b$, matching the microscopic Ampère law.

A long solenoidal sample has known M only to 5 percent. Can a single Hall reading at the centre establish that its magnetisation is uniform over the entire volume to that accuracy?

▼ **Work through the solution**

No. The reading is one geometry-weighted integral of the source. Reversing magnetisation can remove a stable external offset, and measurements at several axial positions can test the uniform-cylinder shape, but the single central datum cannot uniquely determine a spatially resolved source. Agreement of a fitted constant at one point supplies no independent uniformity test.

Magnetism, induction and field dynamics · 5. Interfaces, demagnetising fields and the work of magnetisation

Two samples made of the same material can respond very differently to an externally applied field if one is a long rod and the other is a thin disc. The induced moment depends on the internal field, which the magnetised sample helps create. The same distinction reappears when one tries to infer a force from stored magnetic energy.

Introduce the auxiliary field by removing a known current

With magnetisation present the total steady current is $\boldsymbol{J}_f + \nabla \times \boldsymbol{M}$, where the curl includes any sheet contribution if $\boldsymbol{M}$ is extended distributionally. Define $\boldsymbol{H} = \boldsymbol{B}/\mu_0 - \boldsymbol{M}$, giving $\nabla \times \boldsymbol{H} = \boldsymbol{J}_f$. A small ribbon crossing an interface with normal $\boldsymbol{n}$ directed from 1 to 2 gives $\boldsymbol{n} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{K}_f$, and a thin pillbox gives

$$\boldsymbol{n} \cdot (\boldsymbol{B}_2 - \boldsymbol{B}_1) = 0.$$

The boundary conditions distinguish the field whose circulation counts free current from the field whose flux has no sources.

If both media are linear, isotropic and local,

$$\boldsymbol{B}_i = \mu_i \boldsymbol{H}_i.$$

In the absence of free surface current H_t and B_n are continuous. For angles measured from the normal, $\tan\theta = B_t/B_n$ therefore gives $\tan\theta_2/\tan\theta_1 = \mu_2/\mu_1$ when the nonzero components permit the ratio. For instance, going from vacuum into a medium with $\mu_r = 5$, let $B_{1n} = 2\,\mathrm{mT}$ and $B_{1t} = 1\,\mathrm{mT}$. The new components are 2 and $5\,\mathrm{mT}$, tilting the local field from 26.6° to 68.2°. A mounted probe sensitive only to the normal component would register no jump. This local argument says which components can jump; it does not solve the complete field of a finite piece of material.

Solve a sphere to find its internal field

Put a uniform isotropic sphere of radius a and permeability $\mu_r\mu_0$ in an originally uniform external $\boldsymbol{H}_0 = H_0\hat{\boldsymbol{z}}$. There is no free current in or near it, so locally $\boldsymbol{H} = -\nabla\psi$. In each uniform domain $\nabla \cdot \boldsymbol{B} = 0$ implies Laplace's equation for ψ. Spherical symmetry of the geometry and the linear far-field potential select the angular harmonic $\cos\theta$. The finite interior solution is $-Cr\cos\theta$; the exterior is $-H_0 r\cos\theta + D\cos\theta/r^2$. Tangential H agrees at the surface when ψ is continuous up to a chosen common constant. Normal B enforces the second equation:

$$C = H_0 - D/a^3,$$

$$\mu_r C = H_0 + 2D/a^3,$$

$$C = \frac{3H_0}{\mu_r + 2}.$$

Inside, $H_{\mathrm{int}} = C$ is uniform. Multiplying by susceptibility $\chi = \mu_r - 1$ gives

$$M = \chi C.$$

Rearranging yields the memorable self-consistency relation $H_{\mathrm{int}} = H_0 - M/3$. The factor $1/3$ came from solving both boundary conditions for a sphere and is not a universal material property. The dipolar exterior carries moment $m = 4\pi D = (4\pi a^3/3)M$, as expected from summing the uniform magnetisation. An ellipsoid with a principal axis parallel to the drive can likewise have uniform magnetisation and $H_{\mathrm{int}} = H_0 - N_d M$ with a geometric demagnetising factor; an arbitrary block usually has spatial variations near edges.

Take $\mu_r = 10$ and $H_0 = 200\,\mathrm{A/m}$. The sphere has $H_{\mathrm{int}} = 50\,\mathrm{A/m}$, $M = 450\,\mathrm{A/m}$ and $B_{\mathrm{int}} = 0.628\,\mathrm{mT}$. A naive use of $M = 9H_0$ would give four times the actual moment. A long narrow specimen parallel to the drive, with very small N_d, comes much closer to that latter value if the material remains linear. For the sphere $\mu_r \to \infty$ gives $B_{\mathrm{int}} \to 3\mu_0 H_0$ rather than infinity; meanwhile $H_{\mathrm{int}} \to 0$. These two trends are compatible because B/H is growing. Calling either simply a 'field enhancement' without specifying the measured field obscures the limiting result.

Derive the incremental energy for a reversible medium

Imagine increasing the current in a winding slowly while holding all material boundaries fixed. Faraday's law makes the terminal voltage, apart from resistance, $d\lambda/dt$, where λ is the flux linkage. The source delivers incremental energy $I\,d\lambda$. In a uniform toroidal core of mean path length ℓ, area A, and N turns,

$$H = NI/\ell,\; \lambda = NBA.$$

Therefore $I\,d\lambda = (A\ell)H\,dB$. In a lossless quasistatic medium the magnetic energy density accumulated along a reversible constitutive path is $w(B) = \int_0^B H(B')\,dB'$. When $B = \mu H$ with constant μ, this reduces to $BH/2$. The corresponding coenergy $w'(H) = \int_0^H B(H')\,dH'$ satisfies $w + w' = BH$ for a single-valued curve through the origin.

As a nonlinear example, use $H = \alpha B + \beta B^3$ with $\alpha = 10^3\,\mathrm{A/(m\,T)}$ and $\beta = 4.0 \times 10^3\,\mathrm{A/(m\,T^3)}$ over the measured range. At $B = 0.50$ T, $H = 1000\,\mathrm{A/m}$. The integrated energy is

$$w = \alpha B^2/2 + \beta B^4/4 = 187.5\,\mathrm{J/m^3};$$

half the endpoint product gives $250\,\mathrm{J/m^3}$ and misses the curvature. Coenergy is $312.5\,\mathrm{J/m^3}$, compensating the difference. The differential permeability is $dB/dH = (\alpha + 3\beta B^2)^{-1} = 2.50 \times 10^{-4}\,\mathrm{H/m}$, whereas the ratio $B/H = 5.00 \times 10^{-4}\,\mathrm{H/m}$. A small alternating signal about a DC operating point samples the differential value. These numbers are properties of the model curve and should not be mistaken for a universal ferromagnetic law.

Add an air gap and localise the energy

Give a toroidal magnetic circuit a thin gap of length g, neglecting fringing, leakage and flux crowding. For equal cross sections the same normal B passes through core and gap. Ampère's law gives

$$NI = H_c \ell_c + B g/\mu_0.$$

For linear μ_r, this yields

$$B = \mu_0 NI/(\ell_c/\mu_r + g).$$

With $\ell_c = 0.20\,\mathrm{m},\ g = 1.0\,\mathrm{mm},\ \mu_r = 1000,\ NI = 100$ A-turns the result is 0.1047 T. The gap consumes 83.3 A-turns; the much longer core consumes 16.7. Vary the two lengths separately to see why high permeability cannot remove the gap contribution.

Their energy ratio follows from $B^2/(2\mu)$ times volume, giving $U_g/U_c = g\mu_r/\ell_c = 5.0$. For area $A = 1.0\,\mathrm{cm}^2$, the gap stores $0.436\,\mathrm{mJ}$ and the core $0.0873\,\mathrm{mJ}$. The simple force scale at one nearly uniform normal gap face is

$$B^2 A/(2\mu_0) = 0.436\,\mathrm{N}.$$

Decreasing g at fixed current changes the total flux, so differentiating only the instantaneous gap energy while holding everything else implicitly fixed need not reproduce the mechanical force. One may use a vacuum stress surface or a coenergy derivative at prescribed current. Both methods require a clear statement about what is held fixed during the imagined displacement.

Vary the gap and the permeability independently

Distinguish reversible storage from a measured loop

Real ferromagnetic samples can have history-dependent magnetisation. If B and H trace a closed hysteresis loop, the integral $\oint H\,dB$ measures energy per volume delivered in one quasistatic cycle, with a positive sign for a dissipative cycle under passive conventions. A single-valued energy density of B alone cannot reproduce this nonzero path integral. At finite frequency conductive eddy losses and relaxation can broaden a measured loop, so a loop measured at one rate alone cannot isolate a purely quasistatic hysteresis loss.

For an approximately parallelogram-shaped cycle spanning switching fields $\pm 120\,\mathrm{A/m}$ and switching induction swing $0.80\,\mathrm{T}$ from lower to upper branch, its approximate enclosed area is $240 \times 0.80 = 192\,\mathrm{J/m^3}$ per cycle. At 50 Hz a $12\,\mathrm{cm}^3$ sample dissipates approximately $0.115\,\mathrm{W}$. The stated swing here is the full separation, eliminating a common factor-of-two ambiguity. A power meter on the terminals also sees the winding's I^2R and any energy leaving the apparatus; a separate resistance measurement and a frequency sweep help isolate the material contribution.

Reconstruct an anisotropic interface from boundary data

At an interface with a single crystal the relation between $\boldsymbol{H}$ and $\boldsymbol{B}$ may be tensorial. The formulas for angle refraction derived earlier assume scalar permeability; the underlying boundary conditions do not. Take a planar interface with normal $+\hat{\boldsymbol{z}}$, vacuum below and a linear reciprocal crystal above. Choose units such that its relative permeability in the x, z plane is the symmetric matrix with entries

$\mu_{xx} = 4, \mu_{xz} = \mu_{zx} = 1, \mu_{zz} = 2$. Its positive determinant and positive diagonal make the associated magnetic energy positive. Suppose the vacuum drive near the boundary has $H_{1x} = 100\,\mathrm{A/m}$ and $B_{1z}/\mu_0 = 200\,\mathrm{A/m}$, with no y component and no free sheet current. These are local interface data, not a complete boundary condition for a finite crystal.

Continuity of tangential H gives $H_{2x} = 100$. Continuity of normal B gives $100 + 2H_{2z} = 200$, hence $H_{2z} = 50\,\mathrm{A/m}$. The crystal's tangential flux density then follows from $B_{2x}/\mu_0 = 4(100) + 50 = 450\,\mathrm{A/m}$; $B_{2z}/\mu_0 = 200$ by construction. In the crystal $\boldsymbol{H}$ and $\boldsymbol{B}$ point at different angles to the normal, with tangent ratios 2 and 2.25 respectively. The local reversible magnetic energy density is

$$\boldsymbol{H} \cdot \boldsymbol{B}/2 = \mu_0[100(450) + 50(200)]/2 = 0.0346\,\mathrm{J/m^3}.$$

Replacing the tensor by a scalar fitted to the normal component would give the wrong tangent field and energy even though that fitted scalar reproduces one measurement by design.

For a general symmetric matrix, the cross energy in $w(\boldsymbol{H}) = \boldsymbol{H} \cdot \boldsymbol{\mu}\boldsymbol{H}/2$ involves both off-diagonal elements. Its derivative with respect to one component yields the corresponding component of $\boldsymbol{B}$. One can test reciprocity by preparing two small field increments about the same reversible operating point and interchanging their directions. The mixed incremental energy should agree. Large hysteresis loops, saturation along an easy axis or externally biased nonreciprocal media require more information before equating the two cross responses. The scalar material tables often used for a gapped iron core therefore encode a particular direction and operating regime.

Determine a demagnetising factor or an intrinsic susceptibility

For an ellipsoidal sample along a principal axis, combine $H_{\mathrm{int}} = H_0 - N_d M$ with $M = \chi H_{\mathrm{int}}$. Define apparent susceptibility $\chi_{\mathrm{app}} = M/H_0$. Rearrangement gives

$$1/\chi_{\mathrm{app}} = 1/\chi + N_d.$$

This linear relation suggests a comparative measurement that can separate material and shape without having a probe inside either sample. Prepare geometrically distinct ellipsoids from one linear isotropic material, measure their moments along specified axes in the same calibrated uniform field, and divide by their respective volumes. If their demagnetising factors are independently known and the material history has not changed, the quantities $1/\chi_{\mathrm{app}} - N_d$ should agree.

Take intrinsic $\chi = 20$. A sphere with $N_d = 1/3$ gives $\chi_{\mathrm{app}} = 2.609$; a slender ellipsoid with $N_d = 0.050$ gives $\chi_{\mathrm{app}} = 10.0$. One might attribute their nearly fourfold moment ratio to different compositions if the shapes were omitted. Suppose the measured slender-sample value is 9.0 instead. Its inferred intrinsic susceptibility would be $1/(1/9 - 0.05) = 16.36$. Comparing the sphere's inferred value of 20 tests whether the discrepancy exceeds shape, volume and magnetometer uncertainties. The inversion is especially ill-conditioned near the large-χ apparent saturation $1/N_d$; small errors in a heavily demagnetised sample then produce large errors in the recovered intrinsic response.

The sensitivity can be quantified without introducing a fitting convention. Differentiate $\chi = \chi_{\mathrm{app}}/(1 - N_d\chi_{\mathrm{app}})$, giving $d\chi/d\chi_{\mathrm{app}} = 1/(1 - N_d\chi_{\mathrm{app}})^2$. At the preceding spherical point this

derivative is nearly 59, so an absolute error 0.02 in the apparent value shifts the recovered intrinsic result by about 1.18. For the slender sample the derivative is only four at the stated point. The large spherical amplification follows from subtraction of nearly equal quantities, not poor electrical calibration. An elongated specimen may therefore yield a more informative intrinsic measurement even if a sphere is easier to machine and align.

Check a material explanation against mechanical and electrical work

Imagine measuring a single coil fitted with an adjustable magnetic insert. At a particular position the current ramp up from zero to 0.50 A produces linkage $\lambda(I) = aI + bI^3$ with $a = 0.040\,\mathrm{H}$, $b = -0.020\,\mathrm{H/A^2}$ over this restricted range. Differential inductance is $a + 3bI^2$ and stays positive here. Stored energy is $\int I\,d\lambda = aI^2/2 + 3bI^4/4$, whereas coenergy is $\int \lambda dI = aI^2/2 + bI^4/4$. At the stated end current they are 4.06 and 4.69 mJ. Their sum 8.75 mJ equals $I\lambda$. A measured terminal waveform can integrate the first expression directly by accumulating $I(v - RI)dt$; independently plotting λ versus I allows the second by planimetric integration. These two areas have different shapes on a nonlinear constitutive curve.

If the insert moves, coenergy at each position can be obtained by repeating the reversible $\lambda(I)$ scan. Its spatial difference divided by travel distance estimates the average force for fixed end current. Simply using half the product of the end current and linkage would predict neither energy correctly in this example. A hysteretic insert follows different curves on raising and lowering the current, so a measured mean force or cycle loss must state the prepared branch. A force predicted by one reversible branch can be checked with a load cell at the same position and current; an electrical scan at another bias is not enough to test the mechanical boundary condition.

Put the magnetic energy on the correct side of a gap

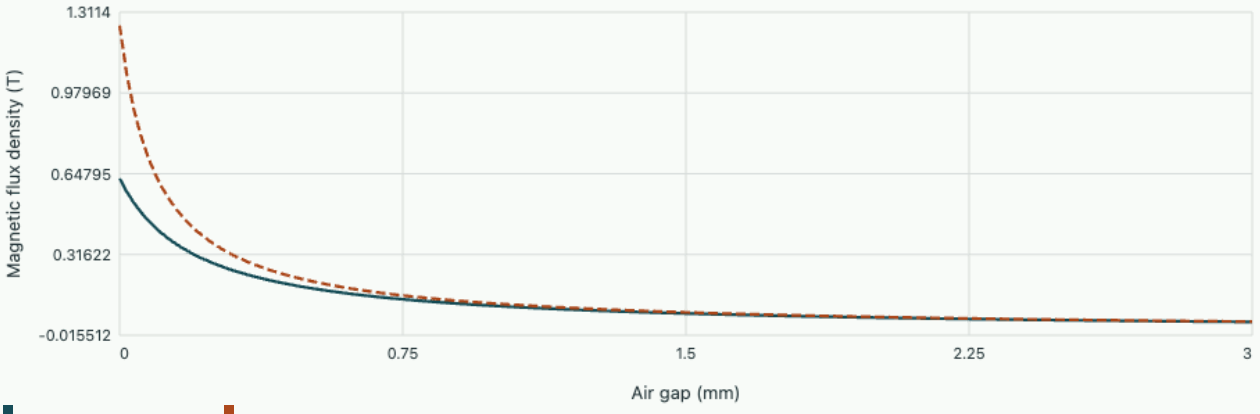
Change the reluctance of a narrow air gap and inspect which section controls flux. Check the assumptions needed before replacing the complete field with a circuit.

Open this experiment online

Try the calculation

Relative permeability **1000**

Magnetomotive force NI (A turns) **100**



At 1 mm the field is 0.1047 T. Equal core and gap reluctances occur at 0.200 mm.

Derive the directional permeability of a stack of thin layers

A material assembled from alternating sheets can respond differently to fields along and across the sheets even when each constituent is isotropic. Consider many parallel layers of two linear, reversible magnetic materials with absolute permeabilities μ_1 and μ_2. Their volume fractions in a representative period are f and $1 - f$. Start far from the outer edges. Suppose the imposed field varies only on lengths much longer than a period, and take the frequency low enough that eddy currents do not change the field distribution. These conditions make a comparison with homogeneous bulk measurements meaningful.

First apply a field tangent to the layers. The tangential field H_t is continuous at each interface because no free sheet current flows there. Each material has a different magnetic induction $B_{it} = \mu_i H_t$. Averaging the induction over one period and dividing by the shared field gives

$$\mu_{\parallel} = f\mu_1 + (1 - f)\mu_2.$$

This is an arithmetic mean. A heavily permeable layer can carry a large fraction of the tangential flux through only part of the period. No condition requires equal tangential induction on the two sides of the internal interface.

Next drive the magnetic flux through the stack in the normal direction. Normal induction is continuous because magnetic flux lines cannot end at a layer. The normal field in each layer is B_n/μ_i. Averaging the field at fixed common induction now gives a harmonic mean,

$$\frac{1}{\mu_{\perp}} = \frac{f}{\mu_1} + \frac{1 - f}{\mu_2}.$$

A small fraction of a weakly permeable spacer can account for most of the magnetomotive drop across a stack. One can derive the same result by adding the magnetic reluctances of a column of common area. That circuit calculation works here because the interface conditions guarantee a common induction. Tilting the applied field creates both normal and tangential components and requires both averages at once. In axes attached to the layers,

$$\langle B_t \rangle = \mu_{\parallel} \langle H_t \rangle,$$
$$\langle B_n \rangle = \mu_{\perp} \langle H_n \rangle.$$

The averaged field vectors are generally not parallel.

For a stack containing 90% by volume of a material with relative permeability 1000 and 10% of a spacer with relative permeability one, the two effective relative permeabilities are 900.1 and approximately 9.91. Suppose a fixture produces equal average components $H_t = H_n = 10\,\mathrm{A/m}$. The induction components are then about $11.31\,\mathrm{mT}$ and $0.125\,\mathrm{mT}$. Their ratio corresponds to an induction angle only 0.63° above the layers while the average field points at 45°. A probe sensitive to a single component could therefore return an apparently inconsistent permeability when the orientation of an anisotropic specimen is ignored.

Verify the energy before using the homogenised material law. For a purely normal induction the microscopic average energy density is

$$f B_n^2/(2\mu_1) + (1-f) B_n^2/(2\mu_2) = B_n^2/(2\mu_{\perp}).$$

For tangential field the microscopic average is

$$H_t^2 [f\mu_1 + (1-f)\mu_2]/2 = \mu_{\parallel} H_t^2/2.$$

Adding both contributions gives $\langle \mathbf{B} \rangle \cdot \langle \mathbf{H} \rangle / 2$ for this linear model. Directly replacing each constituent field by the spatially averaged field before evaluating its energy would give an incorrect normal contribution. The field concentration in the spacer is a physical effect that survives passage to an effective continuum.

Estimate how precisely the spacer fraction must be known if a low-frequency normal-induction calibration is to be reproducible.

▼ Work through the solution

Write $s = 1 - f$ and $\mu_{\perp,r} = [(1-s)/1000 + s]^{-1}$. Differentiating its logarithm gives $d\ln \mu_{\perp,r}/ds = -0.999/[0.001 + 0.999 s]$, about -9.90 at $s = 0.10$. Increasing the spacer volume fraction by one absolute percentage point therefore reduces normal permeability by roughly ten percent to first order. Direct substitution at $s = 0.11$ gives 9.02, or a reduction of about nine percent, showing the size of the linearisation error for that relatively large change. The same change reduces the tangential relative permeability from 900.1 to 890.1, only about one percent.

How does local saturation change the normal and tangential experiments?

▼ Work through the solution

In the normal test the induction is shared. A high-permeability layer reaches a specified saturation induction at approximately the same common induction one measures outside a suitably shaped test column. In the tangential test the field is shared and the high-permeability phase can have a much larger local induction than the reported average. For an infinitesimal variation around a bias point, repeat the interface argument with the local differential permeabilities dB_i/dH_i. The differential effective response obeys the same arithmetic or reciprocal averaging rules under the one-dimensional hypotheses. A finite path that traverses a hysteresis loop needs its history and dissipation as well; no single constant differential permeability specifies its loop area.

David Tong, Electromagnetism in Matter, on constitutive laws and material interfaces.

Tutorial on geometry and constitutive response

A spherical sample has apparent susceptibility $M/H_0 = 2.4$. Infer its intrinsic linear susceptibility and the internal field fraction. Can an apparent value of 3.5 arise in this model with positive intrinsic susceptibility?

▼ **Work through the solution**

For a sphere $M/H_0 = \chi/(1+\chi/3)$. Inversion gives

$$\chi = 2.4/(1-0.8) = 12,$$
$$H_{\mathrm{int}}/H_0 = 1/5.$$

Positive finite χ gives apparent values less than three. A measured larger result would call for checking shape, units, background or the linear homogeneous assumption.

An unknown material is placed in a long gapped circuit. At a given operating point twice the current gives only 1.6 times the old induction. Can a fixed linear permeability with unchanged geometry explain it?

▼ **Work through the solution**

A fixed linear reluctance at fixed geometry gives exactly proportional induction. One can infer $H_c = (NI - Bg/\mu_0)/\ell_c$ at each point and compare the results with a measured constitutive curve. Field-dependent permeability, leakage, fringing or a changing physical gap would each need evaluation before calling the effect saturation.

An interface has no sheet current, the normal field happens to vanish on both sides, and the tangential magnetic flux density differs by a factor of eight. What material ratio follows if both media are scalar and linear? Does this require a discontinuity in H_t?

▼ **Work through the solution**

Continuity of H_t gives $B_{2t}/B_{1t} = \mu_2/\mu_1 = 8$ if the tangential drive is nonzero. No tangential H jump is needed. The magnetisation jump can supply an equivalent bound surface current producing the jump in B_t.

Magnetism, induction and field dynamics · 6. Motional emf and transformer emf in moving circuits

Two laboratory demonstrations produce the same voltmeter trace. In one, a stationary loop sees a changing magnetic field. In the other, a conductor moves through a stationary field. The charges experience different local forces, yet under identifiable conditions both readings follow from the rate of change of magnetic flux.

Define the measured emf through a material loop

Consider a thin conducting loop that can translate or deform. A point of the wire moves with velocity $\boldsymbol{u}(\boldsymbol{r},t)$. In the laboratory a conducting charge also has drift velocity along the wire, and the magnetic force associated with that parallel drift has zero projection along a slender wire element. The tangential Lorentz force per charge, integrated in the direction of current convention, is therefore

$$\mathcal{E}=\oint_{C(t)}(\boldsymbol{E}+\boldsymbol{u}\times\boldsymbol{B})\cdot d\boldsymbol{\ell}.$$

It is a closed-path force per charge at one laboratory time. A voltmeter attached across an opening samples a specific pair of terminals and lead path. To compare it with a predicted closed-loop emf, include the instrumentation leads and consider their field exposure.

For a stationary contour, Faraday's law applies directly:

$$\oint_{C}\boldsymbol{E}\cdot d\boldsymbol{\ell}=-\int_{S}\partial_t\boldsymbol{B}\cdot d\boldsymbol{a}.$$

For a moving contour one must also differentiate the surface. Over a short time each boundary segment sweeps an oriented strip whose area vector is $\boldsymbol{u}\,dt\times d\boldsymbol{\ell}$ for the increment to the enclosed surface in the compatible orientation. Contracting with $\boldsymbol{B}$ gives $-(\boldsymbol{u}\times\boldsymbol{B})\cdot d\boldsymbol{\ell}\,dt$. Interior choices of how to move the spanning surface change nothing when $\nabla\cdot\boldsymbol{B}=0$ and all surfaces stay in the physical domain. Thus

$$\frac{d}{dt}\int_{S(t)}\boldsymbol{B}\cdot d\boldsymbol{a}=\int_{S(t)}\partial_t\boldsymbol{B}\cdot d\boldsymbol{a}-\oint_{C(t)}(\boldsymbol{u}\times\boldsymbol{B})\cdot d\boldsymbol{\ell},$$

$$\mathcal{E}=-\frac{d\Phi}{dt}.$$

The derivative is of the flux through the actual circuit. The partial derivative samples the field change at a fixed spatial point. Confusing the two would predict that moving a loop into a static field can never generate a signal. Conversely, inserting just $B\ell u$ whenever a conductor moves would incorrectly predict a signal for a rigid entire loop translating through an exactly uniform static field.

Follow a sliding bar and the energy feeding the current

Take parallel rails extending along x separated by length L in the y-direction, joined at their left ends, with a conducting bar making the right edge at $x(t)$. A uniform $\boldsymbol{B}=B\hat{\boldsymbol{z}}$ permeates the rails. Choose counterclockwise circuit orientation as viewed from $+z$; on the right edge $d\boldsymbol{\ell}$ points $+y$. For $\dot{x}=u>0$,

$\boldsymbol{u}\times\boldsymbol{B}=-uB\hat{\boldsymbol{y}}$, giving $\mathcal{E}=-BLu$. The area flux is BLx with the chosen orientation, yielding the same result. In a purely resistive loop $I=\mathcal{E}/R$ with this signed convention.

With $B=0.50\,\mathrm{T},\ L=0.20\,\mathrm{m},\ u=2.0\,\mathrm{m/s}, R=0.25\,\Omega$, the emf magnitude is $0.200\,\mathrm{V}$, current $0.800\,\mathrm{A}$, and magnetic drag magnitude $BLI=0.0800\,\mathrm{N}$. An external pull supplying constant speed delivers $0.160\,\mathrm{W}$, exactly I^2R. The magnetic Lorentz force on each individual charge is perpendicular to its total velocity and does no direct work on that charge. The moving conductor, its constraints and the electric field transfer mechanical energy from the pull to the circuit. Treating the magnetic force on the bar as directly heating a stationary isolated charge would miss the mechanism.

Allow a $40\,\mathrm{g}$ bar to coast after release. Neglecting mechanical friction and electrical inductance gives

$$m\dot{u}=-(BL)^2u/R.$$

Its decay time is

$$mR/(BL)^2=1.00\,\mathrm{s};$$

after one second the speed falls to $0.736\,\mathrm{m/s}$. Integrating heat from launch to that moment gives

$$mu_0^2(1-e^{-2})/2=0.0692\,\mathrm{J}.$$

An open bar still develops a charge-separation voltage if its terminals are measured with a high-impedance meter, but without sustained closed current the ideal damping disappears.

Compare the open and closed moving-bar traces

Change the field while changing the area

Use the same rails in a slowly varying uniform field $B(t)$. Differentiating flux gives

$$\mathcal{E}=-L[x\dot{B}+B\dot{x}].$$

The first contribution is transformer emf from a circulating electric field and the second is the motional contribution. If $L=0.20\,\mathrm{m},\ x=0.40\,\mathrm{m}, \dot{x}=0.10\,\mathrm{m/s}, B=0.50\,\mathrm{T}$, the motional term is -0.0100 V. A field ramp of $-0.125\,\mathrm{T/s}$ contributes $+0.0100$ V so net emf vanishes at that instant. Yet the local electric and motional forces need not vanish on every part of the loop. This null experiment tests whether a purported expression for the emf includes both mechanisms with the correct sign.

Keep the area fixed instead, but rotate a rigid N-turn coil of area A in a uniform field such that its normal makes angle ωt with $\boldsymbol{B}$. The flux linkage is $NBA\cos(\omega t)$, giving $\mathcal{E}=NBA\omega\sin(\omega t)$. For $N=100, A=4.0\times10^{-4}\,\mathrm{m}^2, B=0.050\,\mathrm{T}$, a rotation at 10 cycles/s gives peak 0.1257 V and rms 0.0889 V. The flux itself is largest when emf is zero and crosses zero when emf magnitude is largest. If the wiring rotates with the loop until brushes interrupt or reverse a contact, the terminal waveform depends on those connections as well as this geometrical flux. A slip ring and a commutator do not measure the same oriented winding throughout a mechanical cycle.

Calculate a rotating disc without inventing a moving loop

A conducting disc of radius a spins about its axis with angular speed Ω in a steady uniform axial field B. A high-resistance stationary voltmeter contacts the centre and the rim with sliding brushes. A point of the metal has $\boldsymbol{u} = \Omega s\hat{\boldsymbol{\varphi}}$, so

$$\boldsymbol{u} \times \boldsymbol{B} = \Omega s B \hat{\boldsymbol{s}}.$$

Positive charge is pushed radially outward until an electrostatic field counterbalances this force in open circuit. Thus $E_s = -\Omega s B$, and the rim potential relative to the centre is $+\Omega B a^2/2$.

With $a = 0.080\,\mathrm{m}$, $B = 0.20\,\mathrm{T}$ and 300 revolutions/min,

$$\Omega = 10\pi\,\mathrm{s}^{-1}.$$

The voltage is $20.1\,\mathrm{mV}$. The lab shape traced by a radial line and a stationary external lead can appear fixed; taking the time derivative of flux through that fixed geometrical contour alone would give zero. A radial line drawn on the disc rotates while the brushes continue to contact different material points, so the chosen fixed path is not one closed material contour carried along with the conductor velocity. Evaluating the local Lorentz force inside each moving and stationary element is the unambiguous starting point for sliding-contact arrangements. The terminals' open-circuit potential follows from the balancing electric field, which is compatible with zero electrostatic circulation on a fixed spatial loop.

If a resistive load is connected, radial currents in the disc experience an azimuthal magnetic force opposite the drive. For a current element carrying total radial current I at radius s, the torque increment has magnitude $sIB\,ds$ after integrating around the annulus. The whole torque is $IBa^2/2$; multiplying by Ω gives $I\mathcal{E}$. This independent energy calculation recovers the same half, which would be lost by using the rim speed for every radial point. In a real disc both contact resistances and the spreading resistance near a small central contact influence the load current.

Retain the electrical transient in the sliding-bar experiment

The purely resistive moving-bar calculation assumed that current adjusts on a time short compared with any change in speed. To test this assumption connect a fixed external inductor L_c in the return branch, large relative to the small variation of the rails' own inductance over the travel of interest. Let i denote the positive current that produces a force opposing positive bar velocity u, and put $k = BL$ for the field–length product. With the circuit closed at $t = 0$, the coupled equations become

$$m\dot{u} = -ki,\; L_c\dot{i} + Ri = ku.$$

These signs make the mechanical loss term and electrical source term cancel when their equations are multiplied by u and i. Thus $d(mu^2/2 + L_c i^2/2)/dt = -Ri^2$. At the instant of closure with $i = 0$, the bar starts with zero magnetic braking force even though it has nonzero speed and induced emf.

Eliminate $i = -m\dot{u}/k$ to obtain

$$mL_c\ddot{u} + mR\dot{u} + k^2 u = 0.$$

The two roots are $[-R/L_c \pm \sqrt{(R/L_c)^2 - 4k^2/(mL_c)}]/2$. In the limit L_c/R very short, one fast root fills in the initial current and the slow root approaches $-k^2/(mR)$, recovering the original decay. An electrical inductance can also make the idealised system underdamped. In that case magnetic energy temporarily stored in L_c can cause a velocity reversal, despite the total coupled energy decreasing monotonically through resistance. Observing a reversal alone would not imply that the resistor supplied energy.

Choose $m = 0.040\,\mathrm{kg}$, $k = 0.10\,\mathrm{T\,m}$, $R = 0.10\,\Omega$, $L_c = 0.040\,\mathrm{H}$. The characteristic equation is $s^2 + 2.50 s + 6.25 = 0$, giving roots $-1.25 \pm 2.165 i\,\mathrm{s^{-1}}$. With $u(0) = 2.0\,\mathrm{m/s}$, $i(0) = 0$, one obtains

$$u(t) = 2 e^{-1.25 t}[\cos(2.165 t) + (1.25/2.165)\sin(2.165 t)].$$

Initially $\dot{i}(0) = k u(0)/L_c = 5.0\,\mathrm{A/s}$ and $\ddot{u}(0) = -k\dot{i}(0)/m = -12.5\,\mathrm{m/s^3}$. The first zero of u satisfies $\tan(2.165 t) = -2.165/1.25$ on the second-quadrant branch, giving $t \simeq 0.967\,\mathrm{s}$. The approximation of fixed k and nearly fixed total inductance requires enough rail length and a sufficiently uniform field for this motion.

A practical experiment need not wait for a reversal. Record both current and speed at matching times and integrate $R i^2$ over the record. The decrease of mechanical energy alone will temporarily exceed the accumulated heat while the inductor is charging. Subtracting the measured increase $L_c i^2/2$ should close the budget. If it does not, evaluate mechanical friction, position-dependent inductance or an external drive before ascribing the discrepancy to a violation of Lenz's law. A lightweight probe loop placed next to the return inductor would register a changing magnetic field during this transfer, independently verifying storage between motion and heat.

Measure flux across a moving boundary without assuming a uniform field

Take a rectangular loop of height H spanning the interval $[x_0, x_0 + w]$ in a static field $B_z(x) = B_p \exp[-x^2/(2d^2)]$. Translate it along $+x$ at speed u. The flux is $H \int_{x_0}^{x_0+w} B_z(x) dx$; differentiating the moving endpoints gives $\mathcal{E} = H u [B_z(x_0) - B_z(x_0 + w)]$ for positive z orientation. This formula directly illustrates what a translating induction pickup measures. It returns a difference of the field at its two transverse conductor edges rather than the field at its centre. At the instant the loop is centred on the symmetric peak the flux is largest, the two edges see equal fields, and the measured voltage is zero.

For $w = d = 20\,\mathrm{mm}$, $H = 15\,\mathrm{mm}$, $u = 0.30\,\mathrm{m/s}$, $B_p = 0.040\,\mathrm{T}$, suppose the trailing edge is at $x_0 = 0$. Then the leading edge sees $B_p e^{-1/2}$; the emf is

$$H u B_p (1 - e^{-1/2}) = 70.8\,\mu\mathrm{V}.$$

When the leading edge had just reached $x = 0$, the same magnitude occurred with the opposite sign. Integrating the entire voltage pulse from far left to far right gives minus the change in initial and final flux, hence zero despite a nonzero positive lobe and a nonzero negative lobe. A DC offset in acquisition would spoil that cancellation and can imitate a change of background flux.

When $w \ll d$, a Taylor expansion about the centre $x_c = x_0 + w/2$ gives

$$\mathcal{E} \simeq -Huw\, dB_z/dx|_{x_c},$$

with the first odd correction containing $w^3 B_z'''/24$. When $w \gg d$, the two edges traverse the field at well-separated times and give separate pulses approximating its profile with opposite signs. These two regimes can be distinguished by changing only the loop width. A finite voltmeter bandwidth can also smooth the pulses and push their apparent extrema, so a timing fit must include the measurement response if it is comparable with the transit time d/u. An isolated open bar, by contrast, measures the difference in radial charge equilibrium and lacks the cancelling second edge. This accounts for apparently conflicting interpretations of 'measuring the field by moving a wire' in demonstrations with different connections.

Close the power balance of a moving conductor

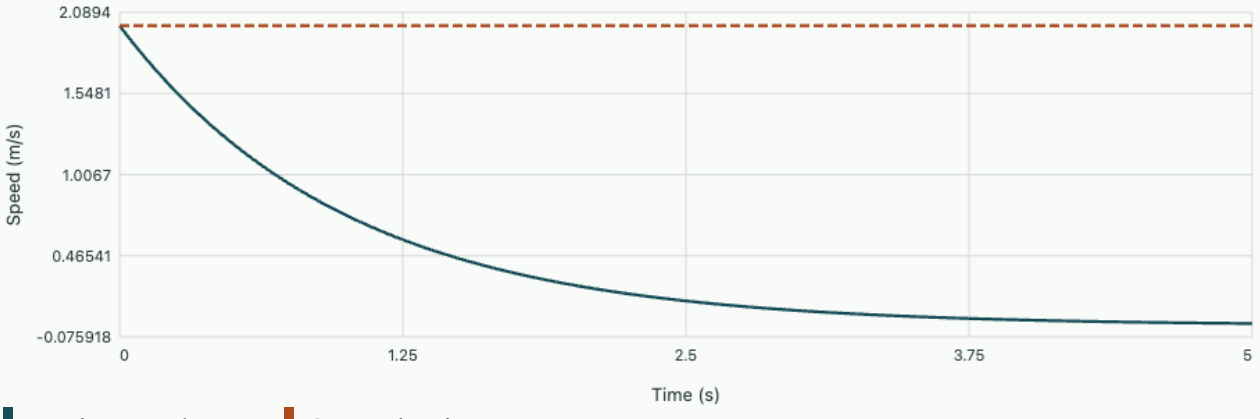
Change circuit resistance and magnetic field. Predict how quickly the bar slows, then relate the lost kinetic energy to the resistive power.

Open this experiment online

Try the calculation

Magnetic field (T) **0.5**

Circuit resistance (Ω) **0.25**



Decay time 1.000 s. At one decay time the speed is 0.736 m/s and 86.5% of the original kinetic energy has gone into Joule heat.

Let a freely moving loop brake itself at a field boundary

A rigid rectangular wire loop of extent w along x and height h along y slides on frictionless guides through a region of uniform vertical magnetic induction $B\hat{z}$. The magnetic region begins at $x = 0$ and ends at $x = L$, with $L > w$. Take its edges sufficiently sharp compared with the loop width, and suppose the

electrical relaxation time of the loop is short compared with the time to cross an edge. The ring then acts like a resistance R. Its mass m and initial speed u_0 determine whether it can cross each boundary at all. Holding it at constant speed would describe a separate experiment in which a drive supplies mechanical work.

While the front edge is inside and the rear edge is outside, the enclosed flux grows as Bhx, where x is the distance the front edge has entered. Faraday's law gives an emf of magnitude Bhu, and the circulating current has magnitude Bhu/R. The horizontal force on the immersed vertical side opposes the motion. The two horizontal sides carry equal and opposite vertical force components and supply no longitudinal drive. The braking force therefore has magnitude $(Bh)^2u/R$. Its direction can also be obtained without a sign convention for current by requiring the mechanical power loss to equal I^2R.

Write $k = (Bh)^2/R$ and $\tau = m/k$. Newton's equation reads $m\,du/dt = -ku$, so $u(t) = u_0 e^{-t/\tau}$. At a known position, a particularly simple expression follows by using $du/dt = u\,du/dx$. As long as the speed is positive one factor of u cancels and $du/dx = -1/\tau$. The loop loses the fixed speed increment w/τ in traversing the first edge, regardless of its initial speed, provided it reaches the end of that traversal. If $u_0\tau < w$, the penetration tends to $u_0\tau$ while the speed tends to zero. It approaches its stopping position asymptotically in this ideal model.

Once both vertical sides are in the same uniform field, changing the position leaves the flux unchanged. The loop coasts through this region at the speed it retained after entry. When it reaches the far boundary the enclosed flux decreases and the induced current reverses. The remaining immersed vertical side changes too. These two sign changes leave the net longitudinal force pointing backwards. A loop that reaches the far edge and crosses it completely loses a second speed increment w/τ. The signs of the current at the two edges therefore distinguish the direction of flux change, while both edges remove kinetic energy.

For $h = 0.10\,\mathrm{m}$, $w = 0.20\,\mathrm{m}$, $B = 0.50\,\mathrm{T}$, $R = 0.10\,\Omega$ and $m = 0.020\,\mathrm{kg}$, the damping coefficient is $0.025\,\mathrm{kg/s}$ and the time constant $0.80\,\mathrm{s}$. Released at $0.80\,\mathrm{m/s}$, the loop drops to $0.55\,\mathrm{m/s}$ after entry and $0.30\,\mathrm{m/s}$ after exit. The durations of the two braking intervals are respectively $0.80\ln(0.80/0.55) = 0.300\,\mathrm{s}$ and $0.80\ln(0.55/0.30) = 0.485\,\mathrm{s}$. They differ because the second identical field boundary is crossed more slowly. The total heat is

$$m(0.80^2 - 0.30^2)/2 = 5.50\,\mathrm{mJ}.$$

A laboratory replaces the free loop by a motor that drags it at $0.80\,\mathrm{m/s}$ throughout. What mechanical signature distinguishes the two experiments?

▼ Work through the solution

A constant-speed drive must balance $ku = 0.020\,\mathrm{N}$ during each partially immersed interval. The associated drive power is $0.016\,\mathrm{W}$. Each interval covers $0.20\,\mathrm{m}$, giving $4.0\,\mathrm{mJ}$ of heat and mechanical work per edge. The two intervals take equally long and the two current pulses have the same magnitude. In the freely moving test the current drops during each pulse and the second pulse starts at a lower magnitude. Agreeing values of B, R and geometry do not fix the energy loss unless the mechanical boundary conditions are also specified.

Under what conditions would a voltmeter or current sensor register a significant signal in the nominally uniform middle section?

▼ **Work through the solution**

With a finite loop inductance the current present at the end of entry cannot vanish instantaneously. It decays on the electrical time $\mathcal{L}/R$ while dissipating previously stored magnetic energy. A nonuniform field also produces an emf whenever the leading and trailing sides sample different inductions, even if both are inside the geometric footprint of the magnet. Finally, a change of the applied field in time drives a transformer emf over the whole loop. Varying the speed, holding the loop stationary, and measuring the field along the track discriminate these mechanisms because their time and position dependence differs.

David Tong, Electrodynamics, on induction and the energy of magnetic circuits.

Tutorial on signs, loops and measurement leads

A rectangular closed loop of width w in the motion direction and height h translates with speed u in $B_z(x) = B_0 + \alpha x$. Predict its signed emf for positive orientation and evaluate at $w = 0.10\,\mathrm{m},\ h = 0.040\,\mathrm{m}, \alpha = 0.30\,\mathrm{T/m}, u = 1.5\,\mathrm{m/s}$.

▼ **Work through the solution**

The flux is $h\int_{x_0}^{x_0+w}(B_0+\alpha x)dx$. Differentiating the two moving endpoints gives

$$hu[B(x_0+w) - B(x_0)] = hw\alpha u.$$

The signed emf is $-1.80\,\mathrm{mV}$. Uniform B_0 cancels between equal opposite moving edges, verifying the uniform-field translation limit.

Two stationary voltmeters are attached to the same pair of terminals but their leads enclose additional area $0.015\,\mathrm{m}^2$ between the two measurement paths. A normal field grows at $0.40\,\mathrm{T/s}$. What discrepancy can occur?

▼ **Work through the solution**

Concatenate the first measurement path with the reverse of the second. It forms a closed stationary loop with induced emf of magnitude $A|\dot{B}| = 6.0\,\mathrm{mV}$. The discrepancy's sign is fixed by the chosen paths and orientation. Keeping outgoing and returning meter leads together reduces the area and makes a lumped terminal voltage a better approximation.

A conducting ring expands radially so $a(t)$ grows while the applied uniform axial field falls as $B \propto a^{-2}$. Does the vanishing net emf imply zero local induced electric field?

▼ **Work through the solution**

The flux $B\pi a^2$ is constant, so a material ring has zero net Lorentz-force circulation. Faraday's law on a stationary circle at its instantaneous radius still gives

$$E_\varphi = -a\dot{B}/2.$$

The moving metal supplies $(\boldsymbol{u}\times\boldsymbol{B})_\varphi = -\dot{a}B$; these terms cancel for the stated time dependence. A stationary charge and a charge bound to the moving ring therefore probe different local forces.

Magnetism, induction and field dynamics · 7. Flux linkage, coupled coils and magnetic coenergy

A winding links flux through each turn. In two windings the current in either one can change the voltage in the other. A matrix of flux linkages describes this coupling, but predicting power or mechanical force requires specifying the current directions and the control conditions.

Measure a linkage rather than a field at one point

Choose a positive traversal direction for every turn of winding i. Define $\lambda_i = \sum_{\text{turns}} \int \boldsymbol{B} \cdot d\boldsymbol{a}$ with normals consistent with traversal. If every turn sees equivalent flux, this is $N_i \Phi_i$; a distributed winding can have significantly different flux per turn. A passive terminal reference takes current entering the positive terminal. For a stationary winding with series resistance R_i, its terminal voltage then obeys $v_i = R_i I_i + d\lambda_i/dt$. The opposing internal induced emf is $-d\lambda_i/dt$. These statements use different definitions of voltage direction, so the opposite signs are consistent.

To measure a differential inductance at fixed geometry, one may superimpose a small sinusoidal current on a DC bias, measure its voltage quadrature, and infer $d\lambda/dI$. A large slow ramp instead measures finite changes in linkage by integrating $v - RI$ over time. If the winding warms during the ramp, using a cold resistance biases the flux integral. At sufficiently high frequency winding capacitance, frequency-dependent conductor resistance and nonuniform flux can invalidate the simple two-terminal model. Entering the source's physical boundary and frequency regime is part of interpreting any inferred inductance.

Derive reciprocity and energy from the field equations

For a linear reciprocal material with fixed geometry and no remanent background, write $\lambda_i = \sum_j L_{ij} I_j$. In free space the Neumann double line integral for distinct slender loops is unchanged when loop indices are exchanged, giving $L_{12} = L_{21}$. For a broader proof take field solutions $(\boldsymbol{H}_i, \boldsymbol{B}_i, \boldsymbol{A}_i)$ for unit currents in windings i. The linked flux in winding j is $\int \boldsymbol{A}_i \cdot \boldsymbol{J}_j \, dV$, where $\boldsymbol{J}_j$ is its unit-current density. Integrate $\boldsymbol{A}_i \cdot (\nabla \times \boldsymbol{H}_j)$ by parts. With appropriate enclosing conditions the surface term vanishes and the integral is $\int \boldsymbol{B}_i \cdot \boldsymbol{H}_j dV$. Since the permeability tensor is symmetric, exchanging i, j gives the same value. Reciprocity depends on the constitutive law and boundary conditions.

At fixed mechanical coordinates the electrical work put into the field is $dW = \sum I_i \, d\lambda_i$ after subtracting losses. Ramp all currents up in fixed proportion from zero and use linearity. The instantaneous integrand scales linearly with fractional ramp amplitude, giving half its endpoint product:

$$W = \tfrac{1}{2} \sum_i I_i \lambda_i = \tfrac{1}{2} L_1 I_1^2 + M I_1 I_2 + \tfrac{1}{2} L_2 I_2^2, \qquad |M| \le \sqrt{L_1 L_2}.$$

The inequality follows from requiring nonnegative energy for every current vector. At fixed I_2, minimise the quadratic with respect to I_1; at $I_1 = -M I_2 / L_1$ energy is $(L_2 - M^2/L_1) I_2^2/2$. Thus no passive linear pair can have coupling magnitude greater than the geometric mean of its self-inductances. Equality corresponds to a combination of currents producing zero linked-field energy in the ideal model, as for perfect coupling without leakage. In an actual device leakage, parasitics and measurement uncertainty matter close to that limit.

Background · Can zero mutual flux coincide with maximum torque?

Consider a small coil that rotates in the nearly uniform field of a fixed coil. For a suitable angle origin the linked flux is proportional to the projected area, so $M(\theta) = M_0 \cos\theta$. Hold both currents fixed with ideal supplies. In a linear medium the angle-dependent part of the magnetic coenergy is $M(\theta) I_1 I_2$. Its derivative gives the torque.

$$\tau_\theta = I_1 I_2 \frac{dM}{d\theta} = -M_0 I_1 I_2 \sin\theta.$$

Precisely at a right angle the instantaneous mutual linkage is zero while its angular derivative has maximum magnitude. A freely rotating current loop therefore feels its largest aligning torque there. Measuring linkage at one orientation cannot supply this force. Measuring it at two nearby angles estimates the derivative that enters the mechanical response.

Reverse a connection and include a shorted secondary

Consider $L_1 = 9.0\,\mathrm{mH}$, $L_2 = 4.0\,\mathrm{mH}$, $M = 4.8\,\mathrm{mH}$, so coupling coefficient $k = M/\sqrt{L_1 L_2} = 0.80$. With currents $I_1 = 0.40$ A and $I_2 = -0.20$ A, the flux linkages are $2.64\,\mathrm{mWb}$ and $1.12\,\mathrm{mWb}$ using the signed references. The energy is

$$(0.40 \times 2.64 - 0.20 \times 1.12)/2 = 0.416\,\mathrm{mJ}.$$

Reversing only the second physical winding while keeping the named terminal references changes the mutual sign but leaves the diagonal self terms positive.

Connect the coils in series. When the same signed current flows in both, the total linkage along the series path gives

$$L_{\mathrm{aid}} = L_1 + L_2 + 2M = 22.6\,\mathrm{mH}.$$

With opposite reference directions the result is

$$L_{\mathrm{oppose}} = L_1 + L_2 - 2M = 3.4\,\mathrm{mH}.$$

The difference of the measured series inductances divided by four estimates the signed mutual coupling for the chosen assignment. Both inferred inductances are nonnegative. A negative value obtained from a model suggests incorrect signs, units or incompatible measurements.

Now close winding two through an ideal lossless short while holding geometry fixed. With zero initial secondary linkage the voltage condition $d\lambda_2/dt = 0$ implies $I_2 = -(M/L_2) I_1$, hence $\lambda_1 = (L_1 - M^2/L_2) I_1$. In the example the input inductance becomes $3.24\,\mathrm{mH}$, smaller than the open-secondary value of 9.0. Its reduction is not the same calculation as series opposition; the secondary current is determined by a flux constraint rather than imposed as $-I_1$. With finite secondary resistance the DC steady secondary current eventually returns to zero. The notion of a perfectly shorted, flux-conserving secondary is an ideal time-scale limit.

See the secondary's losses from the primary

For sinusoidal current use peak phasors with $e^{i\omega t}$. Let the full secondary resistance be R_2 and close it without an applied voltage. Solve $0 = (R_2 + i\omega L_2)\widetilde{I}_2 + i\omega M \widetilde{I}_1$, then substitute into the primary equation. The input impedance becomes

$$Z_{\mathrm{in}} = R_1 + i\omega L_1 + \frac{\omega^2 M^2}{R_2 + i\omega L_2}.$$

The additional real part $\omega^2 M^2 R_2/[R_2^2 + (\omega L_2)^2]$ is positive and represents power dissipated in winding two. Its imaginary part is negative and reduces the primary reactive voltage. In the high-frequency limit within the lumped model the reactive term tends towards $i\omega(L_1 - M^2/L_2)$. In the low-frequency limit the induced coupling vanishes quadratically in the real part. A sense of physical scale is obtained by choosing $\omega = 250\,\mathrm{s}^{-1}$ and $R_2 = 1.0\,\Omega$. Then

$$\omega L_2 = 1.0\,\Omega, \omega M = 1.20\,\Omega.$$

The added impedance is $1.44/(1+i) = 0.72 - 0.72i\,\Omega$, so a primary with negligible resistance sees $0.72 + 1.53i\,\Omega$.

With primary amplitude 1.0 A, the secondary amplitude is

$$1.20/\sqrt{2} = 0.849\,\mathrm{A}.$$

Its mean Joule loss is

$$0.849^2 (1.0)/2 = 0.360\,\mathrm{W}.$$

Independently the primary's reflected resistance draws $1.0^2(0.72)/2 = 0.360$ W. This agreement tests the sign of the induced secondary current and the correct use of peak versus rms amplitudes. Reversing the secondary orientation flips its voltage reference and its current phase by π, but the reflected impedance depends on M^2 and is unchanged.

Let geometry move while current is maintained

If a mechanical coordinate x changes, terminal electrical work can go into both field energy and mechanical output. For a reversible device $dW_{\mathrm{field}} = \sum I_i d\lambda_i - F_x dx$. Introduce coenergy $W' = \sum I_i \lambda_i - W_{\mathrm{field}}$; the product rule gives

$$dW' = \sum \lambda_i dI_i + F_x dx.$$

Therefore force at fixed currents is $F_x = (\partial W'/\partial x)_{\boldsymbol{I}}$. At fixed linkages instead,

$$F_x = -(\partial W_{\mathrm{field}}/\partial x)_{\boldsymbol{\lambda}}.$$

A linear reversible system has $W' = W_{\text{field}}$ at each state, but their natural independent variables differ. Differentiating the same algebraic expression under an inappropriate constraint reverses a force prediction.

Suppose translating one coil increases M at $0.20\,\mathrm{H/m}$, while its self-inductances are approximately constant. For maintained $I_1 = 0.40\,\mathrm{A}$, $I_2 = 0.30\,\mathrm{A}$, $F_x = I_1 I_2\, dM/dx = 0.024\,\mathrm{N}$ in the direction of increasing coupling. If x changes by 1.0 mm slowly, the field energy rises $24\,\mu\mathrm{J}$. Source work from both windings is twice this value because both linkages change. The remaining $24\,\mu\mathrm{J}$ goes to mechanical output. With currents of opposite sign the force reverses. The mechanism is the same as that behind an electromechanical transducer, but the sign is easier to see in a reciprocal two-port than in a one-coil moving-core drawing.

Check force by differentiating magnetic coenergy

Split the resonances of two coupled electrical oscillators

Close each winding through its own identical capacitor C, remove resistance initially, and take identical self-inductances L. Let q_1, q_2 be capacitor charges in orientations for which $I_i = \dot{q}_i$ and the passive capacitor voltage drop is q_i/C. The two source-free loop equations become

$$L\ddot{q}_1 + M\ddot{q}_2 + q_1/C = 0,$$
$$M\ddot{q}_1 + L\ddot{q}_2 + q_2/C = 0.$$

Adding and subtracting decouple the symmetric coordinate $q_1 + q_2$ and antisymmetric coordinate $q_1 - q_2$. Their angular frequencies are $1/\sqrt{(L+M)C}$ and $1/\sqrt{(L-M)C}$ respectively. Thus the pair measures $L + M$ and $L - M$ independently of any assumption that current remains confined to the driven winding.

For $L = 10.0\,\mathrm{mH}$, $M = 2.0\,\mathrm{mH}$ and $C = 1.0\,\mu\mathrm{F}$, the uncoupled frequency is 1592 Hz. The symmetric mode lies at 1453 Hz, the antisymmetric at 1779 Hz. A sensor observing one coil after an initial excitation may see both: with $q_1(0) = q_0, q_2(0) = 0$ and both currents zero, $q_1 = q_0[\cos(\omega_s t) + \cos(\omega_a t)]/2$, while $q_2 = q_0[\cos(\omega_s t) - \cos(\omega_a t)]/2$. The envelope of the first can almost vanish as energy appears in the second. Energy has not been externally driven during this interval. The coupled magnetic cross term is part of the complete energy, so simply adding the two uncoupled winding energies at arbitrary instants can give an apparent oscillation of the total.

Inverse measurement is possible. If the two frequencies are f_s, f_a, eliminate unknown C to get $M/L = (f_a^2 - f_s^2)/(f_a^2 + f_s^2)$ for the stated same-sign winding labels. If frequencies alone are known but one cannot tell which mode carries same-sign current, one obtains the magnitude of the coupling. Measuring the relative current phase identifies the assignment. Unequal capacitances, unequal self-inductances or appreciable damping change the eigenvalue problem, so a frequency splitting should not be identified with $2M$ or with a material coupling independent of the circuit model. A loss measurement can differentiate genuine normal-mode splitting from two overlapping modes broadened by resistance.

Recover transformer relations and identify the leakage mode

An ideal high-permeability closed core links all turns by a common flux Φ, so

$$\lambda_1 = N_1\Phi,\ \lambda_2 = \pm N_2\Phi.$$

Its terminal induced voltage ratio follows on differentiation, $v_2/v_1 = \pm N_2/N_1$ when resistive drops and parasitic voltages are negligible. The relation is between oriented time derivatives, not a rule about the magnitude of any static potential across a disconnected winding. A DC voltage held across a resistanceless winding would ramp the core flux indefinitely in this ideal model; a real core saturates or winding resistance establishes a different state. The harmonic transformer approximation operates within a finite flux and frequency range.

For a linear core with reluctance $\mathcal{R}$, $\Phi = (N_1 I_1 + N_2 I_2)/\mathcal{R}$ with consistent orientations. Thus $L_1 = N_1^2/\mathcal{R},\ L_2 = N_2^2/\mathcal{R},\ M = N_1 N_2/\mathcal{R}$. The inductance matrix has a zero eigenvalue since the current combination $N_1 I_1 + N_2 I_2 = 0$ creates no core flux. Real leakage breaks this exact cancellation. Adding small series leakage inductances corresponds to storing additional positive energy outside the common flux region. The shorted-secondary input test is particularly sensitive to this residual storage because the coupled part largely cancels.

Take an 80-turn primary and 20-turn secondary on the same linear core. With 4.0 V rms across the magnetising branch at 2.0 kHz, sinusoidal flux amplitude follows from $V_{\mathrm{rms}} = N_1\omega\Phi_{\mathrm{peak}}/\sqrt{2}$, yielding $5.63\,\mu\mathrm{Wb}$. A $2.0\,\mathrm{cm}^2$ core then sees 0.0281 T peak. The open secondary gives 1.0 V rms. A $5.0\,\Omega$ resistive secondary load would draw 0.20 A rms and dissipate 0.20 W in the ideal lossless-transformer limit. The load-related primary current is then 0.050 A rms, corresponding to a reflected resistance of $4.0/0.050 = 80\,\Omega$, as given by $(N_1/N_2)^2 R_L$. An additional reactive magnetising current is needed for finite $\mathcal{R}$; a primary ammeter can consequently read more than the reflected load component even with negligible winding power loss.

Halving drive frequency at unchanged voltage doubles peak flux. This can increase saturation and harmonic distortion even though a naive turns-ratio calculation still predicts the same rms open-circuit voltage. Repeating the experiment with unchanged voltage-to-frequency ratio instead holds the ideal peak flux fixed. These different protocols distinguish a failure of flux linearity from simple winding loss. They also show why mutual inductance as a constant scalar is a model tied to a current, geometry and frequency range. Its measurement should specify the secondary termination, DC bias and whether the observed winding responds to essentially all the changing flux.

Infer force from linked flux

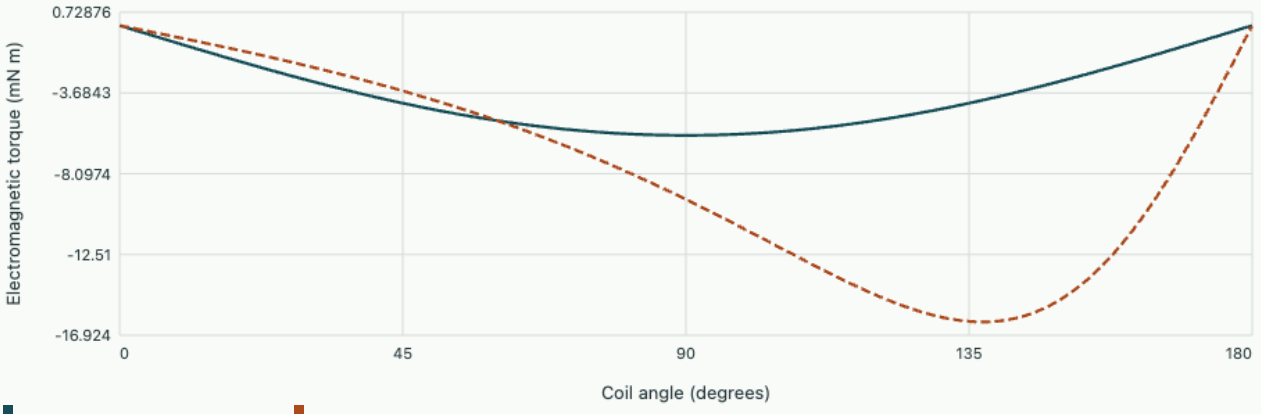
Change the coil geometry and current while checking which electrical variable is held fixed in the mechanical derivative.

Open this experiment online

Try the calculation

First current at preparation (A) 2

Second current at preparation (A) 1



Currents held fixed Fluxes held fixed after 60°

At 60° both preparations give -5.196 mN m. At 90° the current-controlled and flux-controlled torques are -6 and -9.5 mN m.

Compute the reaction of a ring that conserves its flux linkage

A primary coil is held at current I_1 by an ideal supply. Along its axis a second closed coil can move with coordinate x. The two self-inductances L_1, L_2 may be treated as constants, while their mutual inductance $M(x)$ increases as the centres approach. Choose both winding directions so that aiding positive currents produce positive M. Imagine first that resistance in the secondary is negligible over the entire experiment. Faraday's circuit equation then conserves its total linkage $\lambda_2 = L_2 I_2 + M(x) I_1$. The conserved quantity is fixed by the preparation of the ring, so an unreported initial field can change even the sign of the force.

Begin with zero secondary linkage. As the two coils approach at fixed current I_1, positive linkage from the primary grows and the secondary current must become negative. Its value is

$$I_2 = -M I_1 / L_2.$$

For two linear windings the magnetic coenergy with the instantaneous currents independently specified is $W' = (L_1 I_1^2 + 2 M I_1 I_2 + L_2 I_2^2)/2$. The mechanical force is the derivative with respect to x at fixed winding currents. Take that partial derivative first to get $F_x = I_1 I_2 M'(x)$, and then insert the actual secondary current. At zero trapped linkage this gives

$$F_x = -\frac{I_1^2}{L_2} M(x) M'(x).$$

For an approach coordinate that increases toward stronger coupling, the sign is negative. Magnetic induction supplies a restoring tendency against entry into the high-coupling region. The measured primary linkage is $[L_1 - M^2/L_2] I_1$, which gives the same terminal inductance previously found from a short

circuit at sufficiently high frequency in the linear model. With zero trapped linkage the stored energy also reduces to $I_1^2[L_1 - M^2/L_2]/2$. For fixed current and moving coils the supply exchanges energy. Its work $I_1\, d\lambda_1$ must be retained when using a field-energy calculation to deduce mechanical force.

For example, let $L_1 = 20\,\mathrm{mH}$, $L_2 = 5.0\,\mathrm{mH}$, $M = 3.0\,\mathrm{mH}$, $M' = 0.12\,\mathrm{H/m}$, and $I_1 = 0.50\,\mathrm{A}$. The secondary current is $-0.30\,\mathrm{A}$, the effective primary inductance is $18.2\,\mathrm{mH}$, and the force is $-0.018\,\mathrm{N}$. The coupling coefficient $M/\sqrt{L_1 L_2} = 0.30$ satisfies positive energy. A small approach of $1.0\,\mathrm{mm}$, over which this local gradient varies negligibly, gives approximately $-18\,\mu\mathrm{J}$ of work by the magnetic force. The external apparatus must do positive work to maintain a slow approach.

Now prepare a positive trapped linkage $3.0\,\mathrm{mWb}$, with all the other numerical data unchanged. The secondary current is $(\lambda_2 - MI_1)/L_2 = +0.30\,\mathrm{A}$, so the same geometry attracts with $+0.018\,\mathrm{N}$. This reversal comes from the initial state of the secondary. The threshold $MI_1 = \lambda_2$ gives zero secondary current and hence zero force at that point, although the primary still stores field energy. For a permanently closed ideal conductor the trapped linkage must be specified separately. Applying a steady-field circuit equation with zero resistance and zero terminal voltage does not determine its current by itself.

What changes if the secondary has a small measured resistance and the motion stops immediately after the first force reading?

▼ **Work through the solution**

Once the coils and the primary current are stationary, mutual linkage from the primary stops changing. The secondary equation becomes

$$L_2\, dI_2/dt + R_2 I_2 = 0.$$

The current and the magnetic force decay together with time constant L_2/R_2. With $R_2 = 0.10\,\Omega$ that time is $50\,\mathrm{ms}$, and the initially measured force has fallen to about 37% after one time constant. A slowly moving ordinary ring can lose much of its induced current during motion. A fast passage can resemble the conserved-linkage limit even when a later static reading shows no induced force.

A second student substitutes a nonzero conserved linkage into the stored field energy and differentiates at fixed I_1. Why does this prescription miss the extra trapped-current force?

▼ **Work through the solution**

Substitution puts the field energy in the form $W = (L_1 - M^2/L_2)I_1^2/2 + \lambda_2^2/(2L_2)$. At fixed I_1, λ_2 it has no term linear in trapped linkage. The source work does. The primary linkage is $\lambda_1 = (L_1 - M^2/L_2)I_1 + M\lambda_2/L_2$. At fixed primary current, subtract dW from $I_1 d\lambda_1$ and divide by dx. The resulting mechanical force is $I_1(\lambda_2 - MI_1)M'/L_2$, precisely the independently evaluated mutual force. An energy derivative determines a force only after the controlled variables and any energy supplied at electrical terminals have been identified.

David Tong, Electrodynamics, on mutual inductance and field energy.

Tutorial on electrical and mechanical constraints

A winding pair measures 35 mH in series aiding and 11 mH in series opposing. The first self-inductance is 16 mH. Infer the second, mutual coupling, and primary inductance with the ideal secondary shorted.

▼ **Work through the solution**

Adding the readings gives $2(L_1 + L_2) = 46$ mH, so $L_2 = 7$ mH. Their difference gives $4M = 24$ mH, hence $M = 6$ mH and $|k| = 6/\sqrt{112} = 0.567$. The shorted-secondary primary value is

$$16 - 36/7 = 10.86\,\mathrm{mH}.$$

For a single ideal coil, L increases from 2.0 to 3.0 mH while a source holds current at 0.50 A. Find the change of magnetic energy, source work and mechanical output. Compare to a separate isolated lossless coil with the same initial current and conserved linkage.

▼ **Work through the solution**

At fixed current the stored energy rises by $(0.5)^2(0.001)/2 = 0.125$ mJ. Source work is $I^2\Delta L = 0.250$ mJ, leaving 0.125 mJ mechanical. For the isolated coil $\lambda = 1.00$ mWb stays fixed and energy falls from 0.250 to 0.167 mJ. The released 0.0833 mJ supplies mechanical output without source work.

A sinusoidal test infers a negative reflected resistance in a pair of ordinary passive stationary windings with a passive termination. Which steps would you check before interpreting it physically?

▼ **Work through the solution**

Verify primary and secondary voltage/current reference directions and the distinction between a passive port voltage and the internally induced emf. Then check calibration phase, instrument delay, any active load, and the subtracted winding resistance. The passive model's reflected real part is explicitly nonnegative for $R_2 \geq 0$, regardless of the sign of M.

Magnetism, induction and field dynamics · 8. How a magnetic field relaxes into a finite conductor

A conducting sheet initially shields its interior from an applied magnetic step. Given time, a normal resistive metal admits a static tangential magnetic field. The same equations give a finite slab a discrete relaxation spectrum and a periodic drive a complex penetration depth. These are different initial and boundary conditions for one diffusion operator.

Identify the time derivative lost in the approximation

Let the conductor be stationary, Ohmic, uniform and isotropic with conductivity σ, permeability μ and permittivity ϵ. In the full Maxwell equations both $\sigma \boldsymbol{E}$ and $\epsilon \partial_t \boldsymbol{E}$ contribute to the curl of $\boldsymbol{H}$. At a characteristic angular frequency ω, the ratio of the latter to the former is $\omega\epsilon/\sigma$. For a good conductor at low frequencies this is very small. Neglect displacement current only in that constitutive region; a vacuum gap or a capacitor elsewhere can still carry significant displacement current. Ampère then gives

$$\boldsymbol{E} = (\nabla \times \boldsymbol{B})/(\mu\sigma).$$

Substitution into Faraday and use of $\nabla \cdot \boldsymbol{B} = 0$ yield

$$\partial_t \boldsymbol{B} = D\nabla^2 \boldsymbol{B},$$
$$D = 1/(\mu\sigma).$$

The diffusion time for scale a is of order a^2/D. Increasing conductivity reduces D, because smaller electric fields can sustain stronger opposing currents. In an ideal mathematical conductor with $\sigma \to \infty$, an initial flux can be frozen in this model. That is a different preparation from a superconductor expelling a field in equilibrium. Magnetic diffusion alone contains neither a thermodynamic superconducting transition nor London screening. At sufficiently short times the dropped displacement current also matters: an ideal step is an approximation to a finite-risetime drive, and the diffusion equation's nonzero Gaussian tails at arbitrarily large distances should not be used to infer a superluminal signal.

Obtain a transient similarity profile

Take a half-space $x > 0$ with field $\boldsymbol{B} = B_z(x,t)\hat{\boldsymbol{z}}$; the surface is driven to B_s at $t = 0^+$, while initially the interior has zero field. Scaling x by $\sqrt{Dt}$ suggests $B_z/B_s = F(\eta)$, $\eta = x/(2\sqrt{Dt})$. Compute $\partial_t F = -\eta F'/(2t)$ and $D\partial_x^2 F = F''/(4t)$; hence

$$F'' + 2\eta F' = 0.$$

Its first integral is

$$F' = Ce^{-\eta^2}.$$

Set $F(0) = 1, F(\infty) = 0$ and use the Gaussian integral to obtain

$$F(\eta) = \operatorname{erfc}\eta.$$

The associated current $J_y = -(1/\mu)\partial_x B_z$ is positive for this sign of surface field and has a Gaussian depth profile.

$$J_y(x,t) = \frac{B_s}{\mu\sqrt{\pi D t}} e^{-x^2/(4Dt)},$$

$$\int_0^\infty J_y\,dx = \frac{B_s}{\mu}.$$

The current integral remains fixed as the profile spreads, even though the surface current density falls as $t^{-1/2}$. A finite conductivity does not imply uniform current at every instant. For $\sigma = 5.8 \times 10^7\,\mathrm{S/m}$ and $\mu = \mu_0$, $D = 0.01372\,\mathrm{m^2/s}$. The point reaching half the surface field is

$$x = 2\,\mathrm{erfc}^{-1}(1/2)\sqrt{Dt} = 0.9539\sqrt{Dt}.$$

At $0.10\,\mathrm{ms}$ it is $1.12\,\mathrm{mm}$ deep. At $0.40\,\mathrm{ms}$, four times the time, it is twice as deep. Measuring several half-amplitude positions against time lets one fit D without calibrating the absolute magnetic field strength.

Satisfy both faces of a finite slab

Let the conductor occupy $-a < x < a$. Impose the same sudden tangential field B_s at both faces. At long time the unique steady solution with these equal boundaries is uniform B_s. Subtract this final value by defining the deficit $u = 1 - B_z/B_s$. It satisfies homogeneous boundary values $u(\pm a, t) = 0$ and begins at one inside. Even symmetry permits cosine modes $\cos(k_m x)$ with $k_m a = (m + 1/2)\pi$. Separating time gives the multiplier $e^{-D k_m^2 t}$. Orthogonality on $(-a, a)$ gives squared norm a for each cosine and projection of the constant equal to $2\sin(k_m a)/(a k_m)$. Thus

$$\frac{B_z(x,t)}{B_s} = 1 - \frac{4}{\pi}\sum_{m=0}^{\infty} \frac{(-1)^m}{2m+1} \cos\frac{(2m+1)\pi x}{2a}\, \exp\!\left[-(2m+1)^2 t/\tau_0\right],$$

$$\tau_0 = \frac{4a^2}{\pi^2 D}.$$

The initial boundary jump is real in the idealisation; the series reproduces the interior initial value in the usual Fourier sense, while the imposed boundary value applies for $t > 0$. Later times damp the higher spatial harmonics most rapidly. At the centre and $t = \tau_0$, the first term predicts

$$B/B_s = 1 - (4/\pi)e^{-1} = 0.5316.$$

The next correction is $4e^{-9}/(3\pi) = 0.0000524$, so this is already accurate to within 6×10^{-5}. At very early times many terms are needed because their cancellations encode a nearly field-free interior. A half-space similarity expression near each face can be more efficient there.

For a $4.0\,\mathrm{mm}$-thick copper slab $a = 2.0\,\mathrm{mm}$, giving $\tau_0 = 0.118\,\mathrm{ms}$. At late times the centre deficit $1 - B/B_s$ approaches a single exponential with log slope $-1/\tau_0$. Reaching 90 percent at the centre requires

$$t \simeq \tau_0 \ln(4/(0.1\pi)) = 2.54\tau_0 = 0.300\,\mathrm{ms}.$$

The initial square-root penetration law cannot simply be extrapolated into this final regime, since the penetration fronts from both sides interact. Doubling full thickness at constant material quadruples the long relaxation time.

Compare a periodic solution and a surface impedance

Instead drive the half-space sinusoidally and adopt $e^{i\omega t}$ peak phasors. Solving $i\omega\widetilde{B} = D\,d^2\widetilde{B}/dx^2$ gives $\widetilde{B} = B_s e^{-(1+i)x/\delta}$ with the root $\delta = \sqrt{2D/\omega}$ that decays towards infinity. The envelope falls as $e^{-x/\delta}$, while phase lags the boundary by x/δ. The electrical surface field follows by differentiating, yielding $\widetilde{E}_y = (1+i)\widetilde{H}_z/(\sigma\delta)$. The surface impedance in these orientations is $Z_s = (1+i)/(\sigma\delta)$. Its positive real part measures dissipation, and the equal quadrature term records energy storage with this convention.

At $f = 1.0\,\mathrm{kHz}$ in the copper model, $\delta = 2.09\,\mathrm{mm}$ and $\mathrm{Re}\,Z_s = 8.25\,\mu\Omega$. A tangential surface peak field $H_s = 1000\,\mathrm{A/m}$ gives mean incoming power density $\mathrm{Re}\,Z_s|H_s|^2/2 = 4.13\,\mathrm{W/m^2}$. The same number follows from volume Joule heating. Integrate $\sigma|\widetilde{E}_y|^2/2$ over depth; the squared envelope $e^{-2x/\delta}$ integrates to $\delta/2$. An experiment monitoring just the in-phase field can find a zero at $x = \pi\delta/2$ while the full harmonic magnitude is still $e^{-\pi/2} = 0.208$ of the surface amplitude.

Compare a half-space step and a sinusoidal depth profile using the same material

Use the solution to design a convergence test

A mesh across the finite slab can step $B_j^{n+1} = B_j^n + r(B_{j+1}^n - 2B_j^n + B_{j-1}^n)$, with $r = D\Delta t/h^2$. A Fourier mode e^{ikjh} is multiplied by $1 - 4r\sin^2(kh/2)$, so forward Euler is stable for $0 \le r \le 1/2$. Above that bound a high-frequency grid mode can grow despite the physical diffusion being dissipative. For $h = 0.20\,\mathrm{mm}$ in the copper model, $\Delta t \le 1.46\,\mu\mathrm{s}$. Halving the mesh spacing demands quartering this explicit time step. An implicit scheme can overcome the same stability restriction but still needs small enough steps to resolve the measured transient.

Set both face values, initialise the interior once and record the centre at $t = \tau_0$. Compare first against the series, then against a second mesh and time step. A mistaken use of full thickness in τ_0 would shift the inferred time by a factor four even if the code has converged perfectly. Independent checks include reflection symmetry about the slab centre, zero gradient at that centre, a bounded response between zero and B_s for the continuous problem, and ohmic heating that is nonnegative at every point. The latter can catch sign mistakes in an energy budget even though squaring J hides the sign of a local measured current.

Compare surface and centre in a periodically driven finite slab

At a fixed angular frequency a finite slab with equal tangential field on both faces has a different profile from a half-space. Write the phasor equation $d^2\widetilde{B}/dx^2 = k^2\widetilde{B}$, $k = (1+i)/\delta$ for temporal dependence $e^{i\omega t}$, and impose $\widetilde{B}(\pm a) = B_s$. Even symmetry eliminates $\sinh kx$, giving $\widetilde{B}(x) = B_s\cosh(kx)/\cosh(ka)$. The centre amplitude is $1/|\cosh[(1+i)a/\delta]|$ of the applied field, and

its phase lag is the argument of that denominator. Both incoming penetration profiles must be combined as complex amplitudes satisfying boundary conditions; adding two decaying real envelopes is not a valid phasor solution.

The low-frequency limit has $ka \ll 1$, yielding $B(0)/B_s \simeq 1 - (ka)^2/2 = 1 - i(a/\delta)^2$. The leading correction is a phase lag, while the magnitude change starts at higher order. Thus a sensitive phase detector can pick up eddy-current effects before a large change of amplitude is visible. At $a/\delta = 1$, $\cosh(1+i) = \cosh 1 \cos 1 + i \sinh 1 \sin 1 \simeq 0.834 + 0.989i$, giving centre amplitude 0.773 and lag 0.870 radians. An isolated half-space viewed one skin depth in from a single face would have amplitude 0.368 and lag 1.00 radian. These are genuinely different field boundary conditions, not two definitions of skin depth.

The phasor of the current varies as $\sinh(kx)$ and is odd. Current on one half of the slab runs opposite to that on the other; the net integrated current across both halves vanishes for equal exterior fields. That cancellation does not suppress losses, because local Joule heating is proportional to $|J|^2$. In the thin-slab limit expand Faraday directly. If $B_z(t)$ is nearly uniform,

$$\partial_x E_y = -\dot{B}_z.$$

Symmetry about the centre gives

$$E_y \simeq -x\dot{B}_z.$$

For a sinusoidal peak field B_s, the mean volumetric loss averaged across thickness is $\sigma\omega^2 B_s^2 a^2/6$. The a^2 scaling at fixed total magnetic volume explains why dividing a core into electrically insulated thin laminations can reduce classical eddy losses when the applied field lies along the lamination faces. The thin-slab approximation, isolation between laminae and field orientation are essential parts of that conclusion.

For a material with $\sigma = 2.0 \times 10^6\,\mathrm{S/m}$, half-thickness $0.10\,\mathrm{mm}$, sinusoidal peak field 0.10 T and frequency 50 Hz, the formula gives $3.29\,\mathrm{W/m^3}$. If effective permeability were $100\mu_0$, the skin depth at that frequency would be about $5.03\,\mathrm{mm}$, so the thin-slab ordering is internally consistent for these model numbers. Hysteresis or anomalous dynamic losses can still dominate a measured sample. Halving full lamination thickness at fixed material volume should quarter only this classical contribution, not every loss mechanism in the test fixture.

A thin conducting cylindrical shell gives a single time constant

A different finite system exposes how geometry changes a diffusion transient. Consider a long cylindrical shell of radius a and wall thickness $w \ll a$ surrounding vacuum. Assume the skin depth in the wall exceeds w on the important time scales, so current is approximately uniform across the wall, and the solenoid producing an approximately uniform external axial field is sufficiently long. An azimuthal sheet current K_φ produces an additional axial field inside and negligible external field in the long-cylinder approximation. The boundary condition is

$$B_{\mathrm{in}} = B_{\mathrm{ext}} + \mu_0 K_\varphi.$$

Faraday's law around the shell gives

$$2\pi a E_\varphi = -\pi a^2 \dot{B}_{\mathrm{in}}.$$

The sheet obeys $K_\varphi = \sigma w E_\varphi$. Combining the three relations gives

$$B_{\mathrm{in}} + \tau_s \dot{B}_{\mathrm{in}} = B_{\mathrm{ext}},$$
$$\tau_s = \mu_0 \sigma w a / 2.$$

For a step from zero to B_s the field inside grows as $B_s(1 - e^{-t/\tau_s})$. For a sinusoid the interior phasor transfer is $1/(1 + i\omega\tau_s)$. This model has a single time constant instead of a slab's infinite relaxation spectrum because the wall thickness has been collapsed into one azimuthal sheet-current degree of freedom; magnetic energy is stored mainly throughout the bore. Take copper with $a = 0.030\,\mathrm{m}, w = 0.10\,\mathrm{mm}$ to obtain $\tau_s = 0.109\,\mathrm{ms}$. At its half-power frequency $f = 1/(2\pi\tau_s) \simeq 1.46\,\mathrm{kHz}$, the copper skin depth is about 1.73 mm, comfortably larger than the specified wall. At 100 times this frequency the uniform-wall assumption is much weaker.

This result also suggests a discriminating experiment. Compare shells made of the same material with fixed wall thickness and different bore radius. The leading screening time grows linearly with radius. Compare solid sheets of different normal penetration thickness in the earlier slab experiment and their time grows with thickness squared. Either system can display a late exponential, so an observed straight tail in a semilog trace does not alone reveal which diffusion path limits the response. The thickness and radius dependences, together with the phase response and the known current paths, provide independent evidence. A longitudinal insulating slit in the cylindrical shell interrupts its closed eddy-current path, substantially altering the ideal shielding despite leaving most of the metal present.

Distinguish propagation from diffusion

Change conductivity or drive frequency and inspect how the field attenuates with depth. Compare its amplitude decay and phase lag.

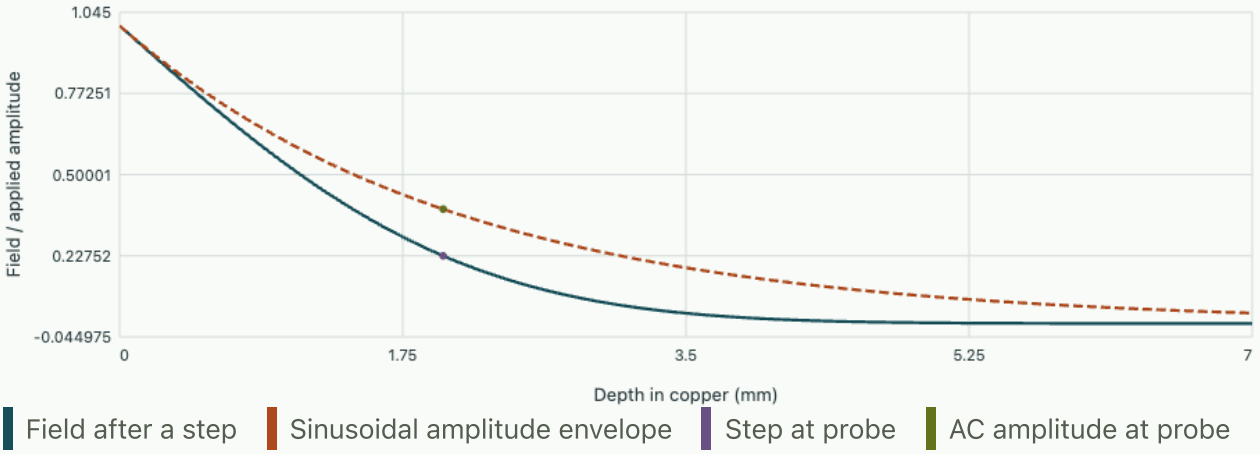
Open this experiment online

Try the calculation

Time after field step (ms) **0.1**

Sinusoidal drive frequency (kHz) **1**

Probe depth (mm) **2**



Half-step depth 1.117 mm. AC skin depth 2.09 mm. At the probe the step fraction is 0.2273, AC amplitude fraction 0.384 and AC phase lag 0.957 rad.

Account for every joule admitted by a magnetic step

A conducting half-space initially has zero internal field. At $t = 0$ a source clamps the surface induction to $B_s\hat{z}$. Let $x > 0$ measure inward distance and $D = 1/(\mu\sigma)$. The previously derived field profile is $B_z = B_s\,\mathrm{erfc}[x/(2\sqrt{Dt})]$. An observer sees an ever-deepening field front even though the surface induction subsequently remains constant. Someone must provide both the magnetic field energy and the Joule heat generated by the induced current. The appropriate bookkeeping is Poynting's theorem, with positive surface power defined into the conductor.

In this geometry Ampère's law gives

$$J_y = -(1/\mu)\partial_x B_z.$$

Differentiating the error function produces a Gaussian current density $J_y = B_s[\mu\sqrt{\pi Dt}]^{-1}\exp[-x^2/(4Dt)]$. Ohm's law sets $E_y = J_y/\sigma$. At the surface the Poynting vector $E_y H_z\hat{x}$ points into the material and has magnitude

$$\frac{P_{\mathrm{in}}}{A} = \frac{B_s^2}{\mu}\sqrt{\frac{D}{\pi t}}.$$

The initial mathematical step creates an integrable square-root singularity in the surface power. Holding B_s fixed still requires the source to deliver power for $t > 0$. Faraday's law relates the surface electric field to the changing flux below the surface. Its value falls with time because the depth derivative of the

induction spreads and weakens. A physical source with a finite rise time smooths the initial singularity and sets the earliest time for which the ideal step is a useful approximation.

To calculate heat production, square the full depth-dependent current and integrate J_y^2/σ from zero to infinity. The squared Gaussian has exponent $-x^2/(2Dt)$, and its integral on the half-line is $\sqrt{\pi Dt/2}$. The result is

$$P_{\mathrm{heat}}/A = (B_s^2/\mu)\sqrt{D/(2\pi t)}.$$

Consequently the heat fraction of instantaneous input is $1/\sqrt{2}$, about 70.7%. The rest, $1 - 1/\sqrt{2}$, supplies an increasing store of magnetic field energy. Both powers decay as $t^{-1/2}$; their integrals up to a finite time increase as $\sqrt{t}$. These fractions belong to the step-driven semi-infinite problem. They do not describe steady sinusoidal loss or a finite specimen after the diffusion front traverses it.

With $B_s = 10\,\mathrm{mT}$, copper conductivity $5.8 \times 10^7\,\mathrm{S/m}$, and permeability close to μ_0, the diffusion coefficient is $0.0137\,\mathrm{m^2/s}$. At $100\,\mu\mathrm{s}$, the incoming power is about $526\,\mathrm{W/m^2}$. Joule heating accounts for $372\,\mathrm{W/m^2}$, and field energy is increasing at about $154\,\mathrm{W/m^2}$. Integration from the ideal step to this time gives $2tP_{\mathrm{in}}(t)/A = 0.105\,\mathrm{J/m^2}$, split into approximately $0.0744\,\mathrm{J/m^2}$ of heat and $0.0308\,\mathrm{J/m^2}$ of stored field energy. At that time $2\sqrt{Dt}$ is about $2.34\,\mathrm{mm}$. Applying these numbers to a foil much thinner than a millimetre would use the half-space solution beyond its range.

A finite slab of thickness d is driven from both faces by the same step. How much magnetic energy is eventually stored, and why must a decaying input replace the half-space power law?

▼ **Work through the solution**

At long times, Ampère's law allows a uniform induction B_s with zero current and the prescribed values on both faces. Its stored magnetic energy is $dB_s^2/(2\mu)$ per unit footprint, independent of conductivity. With $d = 5.0\,\mathrm{mm}$ and the parameters above this is about $0.199\,\mathrm{J/m^2}$. At early times each face behaves approximately as an independent half-space. Once their diffusion profiles overlap, extended use of that approximation would predict unlimited energy accumulation in a finite volume with bounded field magnitude. The finite-slab modes replace the late algebraic law by exponential decay of the current and the surface electric field.

What is the total heat in a finite linear slab after full penetration if both surface inductions are clamped to the same ideal step?

▼ **Work through the solution**

Integrate the inward surface Poynting power over all time. With symmetry about the midplane, Faraday's law on the upper half relates the magnitude of the surface electric field to $d[\int_0^{d/2} B\,dx]/dt$. The surface magnetic field after the step is the constant B_s/μ. Including both faces gives total input energy per area $B_s^2 d/\mu$. Half remains stored, so half has been dissipated as heat, or about $0.199\,\mathrm{J/m^2}$ in the example. The integrated fraction one half over the complete finite transient differs from the early-time half-space

fraction. This energy identity provides a stringent check of a discretised slab calculation because it tests the surface flux, the volume heating and their time integration together.

Hermann Haus and James Melcher, Magnetoquasistatic Relaxation and Diffusion, including experimental demonstrations.

Tutorial on extracting material parameters

The late centre deficit of a slab of full thickness 6.0 mm falls by a factor e in $0.60\,\mathrm{ms}$. Find D, and estimate σ if $\mu = 2\mu_0$.

▼ **Work through the solution**

Using $a = 3.0\,\mathrm{mm}$ gives

$$D = 4a^2/(\pi^2 \tau_0) = 0.00608\,\mathrm{m^2/s}.$$

Then

$$\sigma = 1/(\mu D) = 6.55 \times 10^7\,\mathrm{S/m}.$$

A transient alone determines $\mu\sigma$; the independent permeability assumption is essential.

At fixed frequency the signal amplitude at one depth is 10 percent of the surface amplitude in a periodic half-space experiment. Find the phase lag there and predict the new amplitude when frequency drops to one quarter.

▼ **Work through the solution**

The depth is $x = \delta \ln 10$, so the lag is 2.303 radians. Quartering frequency doubles the skin depth; the ratio becomes

$$e^{-(\ln 10)/2} = 0.3162.$$

Treating the original 10 percent as an intensity ratio instead of a field ratio would produce a different inferred depth.

A step field applied at $t = 0$ is removed again at $t = T$. Obtain the field after removal and explain why it need not vanish immediately in the interior.

▼ **Work through the solution**

By linearity superpose a positive step and a delayed negative step. For $t > T$ the normalised half-space profile is $\mathrm{erfc}[x/(2\sqrt{Dt})] - \mathrm{erfc}[x/(2\sqrt{D(t-T)})]$. It vanishes at the driven boundary but remains positive at interior points where the later disturbance has not substantially diffused. The induced currents subsequently relax and dissipate energy.

Magnetism, induction and field dynamics · 9. Displacement current and the charging capacitor

A wire feeds a capacitor whose dielectric blocks the transport of mobile charges between its plates. A magnetic probe near the gap still responds to charging. Resolving which surface a circulation law should use leads to displacement current, an electric field correction during a ramp, and an energy-flow test around the rim.

Start from a contradiction involving two surfaces

Draw a circle around the conductor leading to one plate. One spanning surface crosses the metal wire and records free current I. Another bulges into the gap and misses the metal entirely. A single circulation $\oint \boldsymbol{H} \cdot d\boldsymbol{\ell}$ cannot equal two distinct free-current integrals for the same contour. This difficulty is already visible locally. If $\nabla \times \boldsymbol{H} = \boldsymbol{J}_f$ held at every time, its divergence would require $\nabla \cdot \boldsymbol{J}_f = 0$, but charge conservation near a plate gives

$$\nabla \cdot \boldsymbol{J}_f = -\partial_t \rho_f.$$

Now impose Gauss's law $\nabla \cdot \boldsymbol{D} = \rho_f$. Differentiating gives

$$\nabla \cdot (\partial_t \boldsymbol{D}) = \partial_t \rho_f.$$

Therefore $\boldsymbol{J}_f + \partial_t \boldsymbol{D}$ is divergence free. Choosing it in Maxwell's circulation law respects both the steady limit and surface independence. Conservation alone determines only the required divergence of the additional term; the full Maxwell law also specifies its transverse content and is tested by electromagnetic waves. A displacement current density has units $\mathrm{A/m^2}$. Its flux can exist in a vacuum region in which no conduction charge crosses the integration surface.

$$\oint_C \boldsymbol{H} \cdot d\boldsymbol{\ell} = \int_S (\boldsymbol{J}_f + \partial_t \boldsymbol{D}) \cdot d\boldsymbol{a}.$$

Subtract the fluxes through two surfaces with the same boundary. They form a closed surface when the second orientation is reversed. The difference vanishes by the divergence theorem applied to the sum. If one uses only the conduction term, the missing difference is the rate at which free charge accumulates in the closed volume. This explains why a DC magnetostatic approximation works in a settled resistor and fails for an insulating gap being charged through the same wire.

Resolve the magnetic field across the plate area

Model two circular conducting plates of radius a separated by $d \ll a$, with a linear dielectric filling the gap. Ignore edge fringing for interior points, use symmetric feeding, and take the charging time long compared with the propagation time across the plates. Then $D_z \simeq Q(t)/(\pi a^2)$. The total displacement current is $I = \dot{Q}$, but a centred Amperian circle at $s < a$ samples only fraction s^2/a^2 of that area. Cylindrical symmetry gives

$$H_\varphi(s) \simeq \frac{Is}{2\pi a^2} \quad (s < a),$$

$$H_\varphi(s) \simeq \frac{I}{2\pi s} \quad (s > a)$$

for a symmetric idealisation with negligible other return paths within the chosen circle. The second expression is a useful intermediate description near the capacitor before the distant return or fringing geometry intervenes. At the rim the formulas match. At $s = 0$ the field vanishes, so assigning the full terminal current to every small circle would produce a fictitious on-axis singularity.

Take $a = 0.050\,\mathrm{m}$, $d = 1.0\,\mathrm{mm}$, $I = 10\,\mathrm{mA}$, nonmagnetic dielectric. At $s = a/2$ the field is $H_\varphi = 0.0159\,\mathrm{A/m}$ and $B_\varphi = 20.0\,\mathrm{nT}$. At the rim $B_\varphi = 40.0\,\mathrm{nT}$. The enclosed displacement current at half radius is $2.50\,\mathrm{mA}$, a quarter of terminal current. Doubling current doubles both readings; doubling plate radius at the same fractional measurement radius halves them. For vacuum the voltage rises at $\dot V = Id/(\epsilon_0 \pi a^2) = 1.44 \times 10^8\,\mathrm{V/s}$, consistent with the ordinary lumped $I = C\dot V$.

Distinguish material and vacuum contributions

In matter $\partial_t \boldsymbol{D} = \epsilon_0 \partial_t \boldsymbol{E} + \partial_t \boldsymbol{P}$. In a transparent ideal linear dielectric with relative permittivity four, impose the same field ramp as in an empty capacitor. The required free current is four times as large, with three quarters of the total changing displacement supplied by polarisation current. If instead the external supply maintains the same current, $\dot D$ is unchanged, the field ramp is four times smaller, and the interior magnetic field calculated above is unchanged. These statements compare different control conditions. Supplying an equivalent charge-free background with the same prescribed voltage waveform is a different experiment from supplying it with the same current.

Add weak leakage $\boldsymbol{J}_f = \sigma \boldsymbol{E}$ in the dielectric. Under a harmonic spatially uniform drive the terminal current per area is $(\sigma + i\omega\epsilon)\widetilde{E}$ with $e^{i\omega t}$. In phase with the voltage there is real conduction, and in quadrature there is changing polarisation and vacuum electric displacement. For area $A = 0.010\,\mathrm{m}^2$, thickness $d = 0.002\,\mathrm{m}$, $\epsilon_r = 5$, $\sigma = 10^{-8}\,\mathrm{S/m}$, the static parallel model has $C = 221\,\mathrm{pF}$, conductance $G = 50.0\,\mathrm{nS}$, and crossover $\omega = G/C = \sigma/\epsilon = 226\,\mathrm{s}^{-1}$ or $f = 36.0\,\mathrm{Hz}$. At that frequency equal in-phase and quadrature currents do not imply equal dissipated powers; only the resistive component draws nonzero average power in the ideal nondispersive model.

Ask how energy gets into the gap

For $\boldsymbol{E} = E_z \hat{\boldsymbol{z}}$ and $\boldsymbol{H} = H_\varphi \hat{\boldsymbol{\varphi}}$, the Poynting vector is $-E_z H_\varphi \hat{\boldsymbol{s}}$. During charging positive E_z and rising positive D_z give energy flow radially inward. Just inside the cylindrical edge the surface area through which this flow enters is $2\pi a d$, and $H_\varphi(a) = I/(2\pi a)$. Thus incoming electromagnetic power is

$$E_z H_\varphi(a) 2\pi a d = VI,$$

which is $d(CV^2/2)/dt$ at constant capacitance in this approximation. Energy enters the dielectric through field flow, although the mobile charges in the plates arrive through conductors. During discharge the sign

reverses.

At $V = 100$ V and $I = 10\,\mathrm{mA}$ the instantaneous rate is 1.0 W. At half radius the local inward flux density is half its edge value, but the local cylindrical area is also half as large. Their product is one quarter, exactly the fraction of stored electric volume inside half radius. The radial Poynting profile therefore agrees with local energy conservation for every coaxial subvolume. When driving becomes fast enough that substantial magnetic energy also accumulates or radiative effects matter, the simple uniform-field model no longer exhausts this balance.

Move the integration radius and inspect the field and inward power

Check the first spatial correction to a uniform electric field

A strictly time-varying azimuthal magnetic field also enters Faraday's equation in the gap, forcing a small radial dependence of E_z. For an axisymmetric z-directed field with negligible z-variation, $(\nabla \times \boldsymbol{E})_\varphi = -\partial_s E_z = -\partial_t B_\varphi$, hence $\partial_s E_z = \partial_t B_\varphi$. Insert the leading Ampère solution $B_\varphi \simeq \mu\epsilon s\,\dot{E}_z(0,t)/2$ and integrate from zero to s. The first correction is $E_z(s,t) - E_z(0,t) \simeq \mu\epsilon s^2 \ddot{E}_z(0,t)/4$. For a pure harmonic drive it scales as $-(ks)^2 E_z(0,t)/4$, with $k = \omega\sqrt{\mu\epsilon}$. This expansion is valid when $ka \ll 1$; it also predicts no correction from a strictly constant-slope ramp after initial transients have settled.

In an idealised uniform-height cylindrical model the harmonic radial wave equation is

$$s^{-1}\partial_s(s\partial_s E_z) + k^2 E_z = 0.$$

Regularity selects $E_z \propto J_0(ks)$, whose series starts with $1 - (ks)^2/4$. This independent result validates both sign and factor in the iterative argument. For a $5\,\mathrm{cm}$ vacuum radius at $10\,\mathrm{MHz}$, $ka = 0.0105$, giving an edge-to-centre relative difference only 2.75×10^{-5}. At $1\,\mathrm{GHz}$, ka is about 1.05 and the uniform premise itself requires a distributed treatment including leads and boundaries. Small geometrical thickness by itself does not ensure instantaneous behaviour across a broad plate.

Track where the conduction current flows on the plate

The uniform-gap approximation says how much displacement current passes a circle. Current in the metal plate is distributed differently. Put a feed contact at the centre of a thin circular top plate and inject free current $I(t)$ there. Assume that the plate charge density $\sigma_f = Q/(\pi a^2)$ is approximately uniform away from the contact and edge. On the metal surface let the radial sheet current be $K_s(s)$, positive outward. For a disk of radius s, the input current minus the outward sheet flux must equal the rate at which that disk acquires charge. This gives $I - 2\pi s K_s = I s^2/a^2$, hence

$$K_s(s) = \frac{I}{2\pi s}\left(1 - \frac{s^2}{a^2}\right), \qquad 0 < s < a.$$

At $s = a/2$ the outward sheet current across the circle is $0.75I$, whereas the displacement current through its interior is $0.25I$. At the rim the radial sheet current goes to zero because there is no

continuation beyond the ideal metal edge. Near $s = 0$ the point-contact idealisation gives $K_s \propto 1/s$, revealing its own range of validity. A finite feed contact rounds this behaviour; finite conductivity and inductance can change the assumed uniform charging during a rapid ramp. The estimated currents on the metal and in the insulator add consistently without a conduction electron having to traverse the dielectric.

With $a = 50\,\mathrm{mm}$, $s = 25\,\mathrm{mm}$ and $I = 10\,\mathrm{mA}$, the radial sheet current at that point is $0.0477\,\mathrm{A/m}$. A ring segment spanning 30° at that radius carries $K_s s(\pi/6) = 0.625\,\mathrm{mA}$ radially outwards across its arc. The corresponding sector has collected only one twelfth of the enclosed displacement current, $0.208\,\mathrm{mA}$. Summing the two contributions over all twelve sectors restores the full terminal current. Measuring a single thin lead and assigning its whole current to a small patch of plate would give a wrong local Joule-loss density even though a terminal ammeter remained accurate.

For a plate with sheet resistance $R_\square$, the resistive dissipation per area under this low-frequency distribution is $R_\square K_s^2$. Integrate over $b < s < a$ to cut off the finite contact at radius b. The radial spreading contribution is

$$P_{\mathrm{sheet}} = \frac{R_\square I^2}{2\pi}\left[\ln\frac{a}{b} - \frac{3}{4} + \frac{b^2}{a^2} - \frac{b^4}{4a^4}\right].$$

The logarithm comes from the high current concentration near the feed. For $a/b = 10$, the bracket is 1.563 and the effective resistance P/I^2 of this plate region is $0.249 R_\square$. This is a weak-loss estimate: if the voltage drop across the plate becomes comparable to the capacitor voltage, the current and charge distributions must instead be solved jointly. It is nevertheless an instructive correction to an apparently straightforward 'lossless capacitor' measurement, in which dissipation may actually be concentrated in the contacts or electrode sheets.

Recover an unknown dielectric response from a controlled ramp

Use a parallel-plate cell of area A and separation d. Its measured small-signal admittance is $Y = I/V$. In the low-frequency geometrical limit $Y = (A/d)(\sigma + i\omega\epsilon)$ for simple frequency-independent real ϵ, σ and phasor convention $e^{i\omega t}$. Suppose $A = 0.0020\,\mathrm{m}^2$, $d = 0.50\,\mathrm{mm}$, $f = 10\,\mathrm{kHz}$, and after subtracting lead capacitance the measured admittance is $0.80 + i8.90\,\mu\mathrm{S}$. The capacitance is $8.90 \times 10^{-6}/(2\pi 10^4) = 142\,\mathrm{pF}$, giving $\epsilon_r = Cd/(A\epsilon_0) \simeq 4.00$. Treating the real component as DC conduction gives

$$\sigma = Gd/A = 2.0 \times 10^{-7}\,\mathrm{S/m}.$$

A terminal measurement at a single frequency alone cannot confirm that interpretation.

A delayed polarisation can produce the same real terminal component without any mobile charge crossing the entire sample. For example take a single orientational relaxation $\tau\dot{P} + P = \epsilon_0\Delta\chi E$. Then $\widetilde{P} = \epsilon_0\Delta\chi\widetilde{E}/(1 + i\omega\tau)$, and the associated current is $i\omega\widetilde{P}$. Its real coefficient is $\epsilon_0\Delta\chi\omega^2\tau/[1 + (\omega\tau)^2]$, positive in this passive model. The quadrature coefficient is $\epsilon_0\Delta\chi\omega/[1 + (\omega\tau)^2]$. A signal analyser would assign the first part to an equivalent conductance. Its strong frequency dependence distinguishes this process from a constant leakage σ if a sufficient frequency range is measured.

A step up in voltage is another way to separate hypotheses. An instantaneous ideal capacitance responds during the step only, a Debye polarisation adds a decaying current over its characteristic τ, and a true Ohmic path continues to carry current at constant final voltage. In a cell whose polarisation term has $\Delta C = 50\,\mathrm{pF}$, step magnitude $1.0\,\mathrm{V}$ and $\tau = 2.0\,\mathrm{ms}$, the delayed contribution starts at $25\,\mathrm{nA}$ and integrates to $50\,\mathrm{pC}$. The initial current in a measured trace will also contain the finite-risetime vacuum and fast-electronic capacitance currents. Integrating after a separately modelled step response is more robust than inferring charge from a single noisy peak. A steady $0.80\,\mu\mathrm{A}$ residual under 1.0 V would support the constant-conductance interpretation for the above cell; decay toward zero would point to polarisation, instrument offsets or a more complex model.

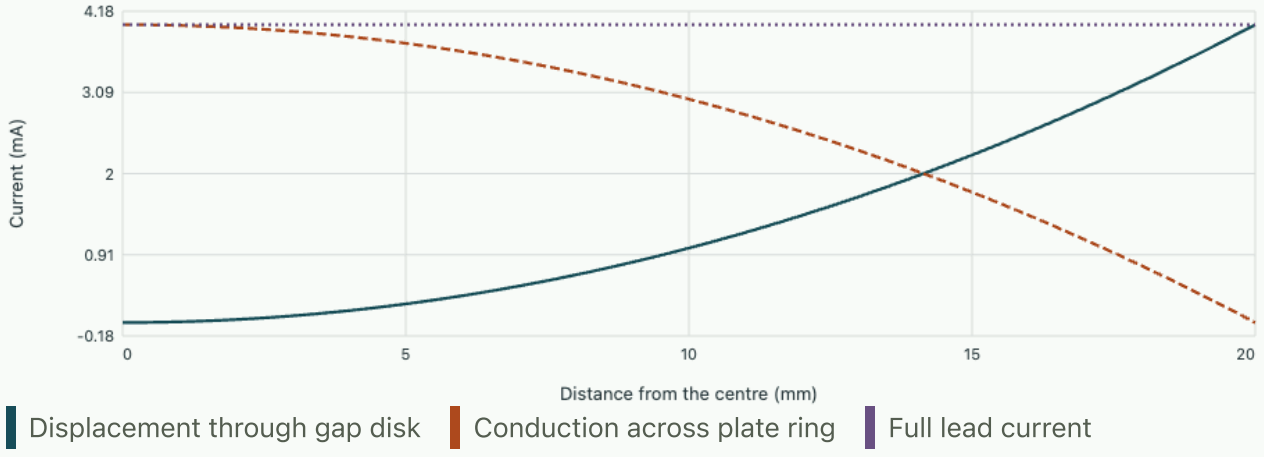
Count displacement current in a charging capacitor

Move the integration surface from wire to dielectric. Check which term in the current balance changes and which loop integral stays the same.

Open this experiment online

Try the calculation

Plate radius (mm)	20
Charging current (mA)	4
Measurement radius / plate radius	0.5



At r = 10 mm the gap encloses 1 mA and the radial plate current is 3 mA. The local azimuthal magnetic field is 20 nT.

Determine the charge that accumulates inside a layered capacitor

Stack two plane dielectric layers in series between extended parallel electrodes. Material 1 has thickness d_1, permittivity ϵ_1 and nonzero conductivity σ_1, with corresponding quantities for material 2. A source suddenly applies a fixed voltage V, and the stack is initially uncharged. The Ohmic current density alone will generally differ between the two layers early in the transient. A conduction-current mismatch deposits

free charge on their internal interface. The total current including displacement remains continuous because equal currents must pass through successive cross sections of this one-dimensional stack.

Choose the field positive from electrode 1 toward electrode 2. Voltage imposes $d_1 E_1 + d_2 E_2 = V$. For positive times after the voltage step, current continuity imposes $\sigma_1 E_1 + \epsilon_1 \dot{E}_1 = \sigma_2 E_2 + \epsilon_2 \dot{E}_2$. Substituting $E_2 = (V - d_1 E_1)/d_2$ and collecting terms gives

$$(\epsilon_1 d_2 + \epsilon_2 d_1)\dot{E}_1 + (\sigma_1 d_2 + \sigma_2 d_1) E_1 = \sigma_2 V.$$

The coefficient of the derivative divided by the coefficient of E_1 is a positive relaxation time τ. The initial split of voltage follows from the absence of pre-existing free interfacial charge, which requires

$$\epsilon_1 E_1(0^+) = \epsilon_2 E_2(0^+).$$

This is the capacitive voltage division of the instantaneous step. At long times, the two steady conduction currents must match, so $\sigma_1 E_1(\infty) = \sigma_2 E_2(\infty)$. This is the resistive voltage division. At intermediate times each field relaxes exponentially from its initial value to its own final value with the same τ, while their thickness-weighted sum remains exactly fixed by the supply.

To see the signs in an example take $d_1 = d_2 = 1.0\,\mathrm{mm}$, $V = 10\,\mathrm{V}$, $\epsilon_1 = 2\epsilon_0$, $\epsilon_2 = 4\epsilon_0$,

$$\sigma_1 = 1.0 \times 10^{-8}\,\mathrm{S/m},$$

$$\sigma_2 = 4.0 \times 10^{-8}\,\mathrm{S/m}.$$

The fields just after charging are 6.67 and $3.33\,\mathrm{kV/m}$, respectively. At long times they are 8.00 and $2.00\,\mathrm{kV/m}$. The relaxation time is

$$6\epsilon_0/(5.0 \times 10^{-8}\,\mathrm{S/m}) = 1.06\,\mathrm{ms}.$$

Increasing a conductivity at fixed permittivity shortens this time because free charge can move faster to redistribute the internal electric field.

The normal directed from material 1 into material 2 gives a free interface charge $q_s = \epsilon_2 E_2 - \epsilon_1 E_1$ per unit area. Its final value is $-7.08 \times 10^{-8}\,\mathrm{C/m^2}$. The initial conduction current into the interface from layer 1 is $6.67 \times 10^{-5}\,\mathrm{A/m^2}$, while $1.33 \times 10^{-4}\,\mathrm{A/m^2}$ leaves through layer 2. Thus our convention predicts decreasing, negative interface charge. Direct differentiation gives $\dot{q}_s = J_{\mathrm{cond},1} - J_{\mathrm{cond},2}$, precisely the same sign. Measuring the total electrode current alone does not tell the experimenter that the current inside each layer is all conductive.

Find the total through-current just after the initial ideal charging impulse, and compare it with the long-time value.

▼ Work through the solution

Immediately after the impulse, layer 1 has $\dot{E}_1 = (8000 - 6667)/(1.06\,\mathrm{ms})$, about $1.25 \times 10^6\,\mathrm{V/(m\,s)}$. Its displacement contribution is therefore about $2.22 \times 10^{-5}\,\mathrm{A/m^2}$. Adding layer 1 conduction gives $8.89 \times 10^{-5}\,\mathrm{A/m^2}$. In layer 2 the derivative has the opposite sign because the layer thicknesses match. The negative displacement contribution subtracts about $4.44 \times 10^{-5}\,\mathrm{A/m^2}$ from its larger conduction

current and gives the same total. At long times only the matched conduction current $8.0 \times 10^{-5}\,\mathrm{A/m^2}$ remains. A real pulse generator gives a finite initial charging burst whose duration should be separated from the material relaxation when fitting data.

Can this sample have nonzero conductivity and still produce no slowly accumulating free interface charge?

▼ **Work through the solution**

It can if $\sigma_1/\epsilon_1 = \sigma_2/\epsilon_2$. In that case the initial capacitive field division also satisfies the steady conduction condition, and the two equilibrium values in the exponential coincide. The voltage step can still deliver an initial charging current and sustain later resistive heating. An absence of the interfacial relaxation feature does not by itself imply perfectly insulating layers. Conversely, a measured slow response from a stack need not indicate slow molecular reorientation in either homogeneous constituent. A change of layer thickness or replacement with a homogeneous control can separate that possibility from charge accumulation at an internal interface.

David Tong, Electrodynamics, on displacement current and field propagation.

Tutorial on two spanning surfaces

A gas-filled capacitor has $C = 35\,\mathrm{pF}$ and rms drive $12\,\mathrm{V}$ at 20 kHz. Determine rms displacement current and the fraction surrounded by a centred circle at radius $0.7a$.

▼ **Work through the solution**

$$I_{\mathrm{rms}} = 2\pi f C V_{\mathrm{rms}} = 52.8\,\mu\mathrm{A}.$$

The fractional area is $0.7^2 = 0.49$, so the smaller circle encloses $25.9\,\mu\mathrm{A}$ in the uniform idealisation. Rms and peak conventions must be used consistently throughout the calculation.

An isolated capacitor initially carries $Q_0 > 0$ on its positive plate. With weak dielectric conductivity and no external leads attached, derive its voltage decay time. What happens to the net current crossing an internal spanning surface?

▼ **Work through the solution**

The parallel conductance $G = \sigma A/d$ drains the free charge, so $C\dot{V} + GV = 0$ and $V = V_0 e^{-t/\tau}$ with $\tau = C/G = \epsilon/\sigma$. Conduction current and displacement current through the gap have opposite signs and equal magnitudes in the uniform model. Their sum vanishes even though Joule heat is produced by the decreasing stored electric energy.

With the second plate connected through a distant return wire, a large probe loop encircles the feed and return together. Should the $1/s$ exterior gap formula be extrapolated to this loop?

▼ **Work through the solution**

The derivation assumed only the charging current was enclosed and the surrounding geometry retained the required symmetry. Including the return changes the enclosed free-current balance and often destroys cylindrical symmetry. Choose a new surface and include its complete total-current flux before predicting any circulation. The local field need not vanish just because the complete circulation is zero.

Magnetism, induction and field dynamics · 10. Energy flow and mechanical force from vacuum fields

An external support holds a charged capacitor plate or a current-carrying coil in place. One can infer the support force from fields just outside the body, without summing forces on every microscopic charge. A second surface integral tracks the energy feeding the same apparatus. Both methods start with Maxwell's equations, but they transport different conserved quantities.

Derive a local energy balance before choosing the surface

In vacuum write $\nabla \times \boldsymbol{E} = -\partial_t \boldsymbol{B}$ and $\nabla \times \boldsymbol{B}/\mu_0 = \boldsymbol{J} + \epsilon_0 \partial_t \boldsymbol{E}$. Dot the second equation with $\boldsymbol{E}$ and the first with $\boldsymbol{B}/\mu_0$. The vector identity $\nabla \cdot (\boldsymbol{E} \times \boldsymbol{B}) = \boldsymbol{B} \cdot \nabla \times \boldsymbol{E} - \boldsymbol{E} \cdot \nabla \times \boldsymbol{B}$ converts their difference into

$$\partial_t u + \nabla \cdot \boldsymbol{S} = -\boldsymbol{J} \cdot \boldsymbol{E},$$

$$u = \frac{\epsilon_0 E^2}{2} + \frac{B^2}{2\mu_0},$$

$$\boldsymbol{S} = \frac{\boldsymbol{E} \times \boldsymbol{B}}{\mu_0}.$$

Integrate over a volume held fixed in the chosen frame. Outward power $\oint \boldsymbol{S} \cdot d\boldsymbol{a} > 0$ represents energy leaving; positive $\int \boldsymbol{J} \cdot \boldsymbol{E}\, dV$ represents electromagnetic work on matter. They do not have to agree instantaneously: their sum plus the rate of increase of contained field energy is zero with these sign conventions. In matter, one can include the total microscopic currents inside the vacuum-form balance, or use an appropriate macroscopic constitutive energy description. A vacuum integration surface enclosing the whole body is particularly useful for a force calculation because it avoids assigning momentum to material polarisation separately.

Consider a long straight resistor with radius a, length L, uniform current along $+z$ and voltage drop V in that direction. Immediately outside its curved face the tangential electric field is V/L under the ideal uniform model, and $H_\varphi = I/(2\pi a)$. Their product makes a radial inward $\boldsymbol{S}$. Integrating over $2\pi a L$ gives VI, equal to the internal $I^2 R$ in steady state. At $V = 4.0\,\mathrm{V}$, $I = 0.20\,\mathrm{A}$, $a = 1.0\,\mathrm{mm}$, $L = 0.10\,\mathrm{m}$, the ideal side flux density is $1273\,\mathrm{W/m^2}$ and the total is $0.80\,\mathrm{W}$. A directed charge drift down the resistor does not imply that its supply of energy travels entirely down the same path.

Convert the Lorentz force into field derivatives

The microscopic force density is

$$\boldsymbol{f} = \rho \boldsymbol{E} + \boldsymbol{J} \times \boldsymbol{B}.$$

Eliminate charge with $\rho = \epsilon_0 \nabla \cdot \boldsymbol{E}$ and current with Ampère–Maxwell. The current term contains $-\epsilon_0 \partial_t \boldsymbol{E} \times \boldsymbol{B}$. Add and subtract the missing piece to make $-\partial_t(\epsilon_0 \boldsymbol{E} \times \boldsymbol{B})$; replace the introduced $\boldsymbol{E} \times \partial_t \boldsymbol{B}$ by $-\boldsymbol{E} \times \nabla \times \boldsymbol{E}$ using Faraday. The spatial terms regroup with $(\boldsymbol{F} \cdot \nabla)\boldsymbol{F} = \nabla(F^2/2) - \boldsymbol{F} \times (\nabla \times \boldsymbol{F})$. The result in Cartesian components is

$$f_i = \sum_j \partial_j T_{ij} - \partial_t g_i,$$

$$T_{ij} = \epsilon_0 (E_i E_j - \tfrac{1}{2}\delta_{ij} E^2) + \mu_0^{-1}(B_i B_j - \tfrac{1}{2}\delta_{ij} B^2),$$

$$\boldsymbol{g} = \epsilon_0 \boldsymbol{E} \times \boldsymbol{B} = \boldsymbol{S}/c^2.$$

For a body enclosed by a fixed vacuum surface S, $F_i = \oint T_{ij} n_j \, dA - d(\int g_i dV)/dt$ if the enclosed Lorentz force is entirely on that body. The traction $\boldsymbol{T} \cdot \boldsymbol{n}$ has units of pressure; the field momentum inside has units of $\mathrm{kg\,m/s}$. The tensor convention used here makes $-T_{ij}$ the outward flux density of i-momentum along j in the conservation equation. Texts that call this outward flux the stress can print the opposite sign. Test the convention on a simple physical force before interpreting a long component sum.

Obtain the attractive pressure on a conducting plate

Take a small vacuum enclosure straddling the surface of an ideal conductor in electrostatic equilibrium. Inside it $\boldsymbol{E} = 0$; just outside $\boldsymbol{E} = E_n \boldsymbol{n}$. Tangential field vanishes. The traction from the exterior face is $\epsilon_0 E_n^2 \boldsymbol{n}/2$; the other thin faces give no limiting contribution. For opposed charged plates this force draws each plate towards the facing gap. Equivalently, a surface charge $\sigma = \epsilon_0 E_n$ experiences the average of the one-sided fields,

$$\boldsymbol{E}_{\text{acting}} = E_n \boldsymbol{n}/2.$$

Using the outside field alone in $\sigma \boldsymbol{E}$ would count the sheet's own field and give twice the measured force.

With area $0.010\,\mathrm{m^2}$, air gap $0.50\,\mathrm{mm}$ and voltage 250 V, $E = 5.0 \times 10^5\,\mathrm{V/m}$, pressure is $1.11\,\mathrm{Pa}$ and force magnitude $0.0111\,\mathrm{N}$. Holding voltage fixed while doubling the gap quarters both pressure and force. If the source is disconnected first, the initial free charge stays fixed and so does the pressure in the ideal large-plate limit. These two protocols agree at the starting configuration and immediately diverge when a plate moves. Neglecting the source when differentiating the fixed-voltage energy $\frac{1}{2}CV^2$ makes a wrong force sign because C changes during the displacement.

Compare force during fixed-charge and fixed-voltage separation

Resolve magnetic tension and side pressure on a surface

For a pure magnetic field of magnitude B along z, the vacuum stress tensor is diagonal with entries $(-B^2/2\mu_0, -B^2/2\mu_0, +B^2/2\mu_0)$. An element whose outward normal is parallel to the field feels positive normal traction in that direction in the tensor convention, while a normal perpendicular to the field gives negative normal traction. For a magnet pole face with a nearly normal gap field and negligible fringe fields elsewhere, integrating the appropriate vacuum surface gives an attraction across the gap of magnitude $AB^2/(2\mu_0)$. For $A = 2.0\,\mathrm{cm^2}$ and $B = 0.20$ T this is $3.18\,\mathrm{N}$. The equal and opposite reaction belongs to the rest of the closed apparatus.

An oblique field produces shear. At a plane with $\boldsymbol{n} = \hat{\boldsymbol{z}}$, $\boldsymbol{B} = B_x \hat{\boldsymbol{x}} + B_z \hat{\boldsymbol{z}}$, one obtains

$$t_x = B_x B_z/\mu_0,$$
$$t_z = (B_z^2 - B_x^2)/(2\mu_0).$$

For $B_x = 0.020$ T, $B_z = 0.050$ T the respective values are 796 Pa and 836 Pa. The sign of the shear reverses when only one component reverses, unlike the pure normal pressure. A single scalar value $B^2/(2\mu_0)$ cannot describe a general loaded face. The surface here is a calculation tool inside vacuum; a proposed net force on a body must still include all the other faces of the enclosure.

Test a radiation-pressure experiment using both integrals

A plane electromagnetic wave in vacuum has transverse electric and magnetic fields with $E = cB$, so its instantaneous energy density is $\epsilon_0 E^2$. The Poynting magnitude is cu, and the momentum density magnitude is u/c. A normally incident pulse of energy U carries momentum U/c towards an absorbing surface. If fully reflected along the reverse direction it changes the field momentum by $2U/c$, giving twice the impulse to the mirror. For a continuous beam of average intensity I on a perfectly absorbing target, pressure is I/c; on a perfect reflector it is $2I/c$. At $I = 600\,\mathrm{W/m^2}$ these are 2.00 and $4.00\,\mu\mathrm{Pa}$. The lower pressure scale relative to the static capacitor example reflects an energy divided by a very large propagation speed.

A pulse striking a surface need not produce its final impulse from the stress term alone at every intermediate time. If the integration box includes part of an incoming or outgoing pulse, $\int \boldsymbol{g}\, dV$ changes. Select a large time interval over which no pulse is present inside at its beginning or end. Then the integrated field-momentum derivative vanishes, and the mechanical impulse equals the time integral of surface traction. Alternatively use a smaller interval and calculate both terms explicitly. Independence of the final impulse from the box location and time window, when all terms are included, is a strong numerical test. It detects spurious forces from artificial absorbing boundaries or inconsistent subtraction of the incoming background.

A standing wave tests energy and momentum at the same surface

A mirror terminating a one-dimensional vacuum beam provides a case in which the cycle-averaged power entering an ideal mirror vanishes but the mechanical force remains finite. Let E_i be the incident peak amplitude polarised along y, and $B_i = E_i/c$ in the appropriate z direction. At an ideal conducting plane the reflected tangential electric field cancels the incident field. The corresponding reflected magnetic field adds there because reversing propagation reverses the sign between E and B. The total magnetic peak is $2E_i/c$, while the total tangential electric field on the surface is zero. Hence the normal instantaneous Poynting vector at the mirror vanishes. There is no energy deposition in a stationary lossless perfect conductor.

The vacuum magnetic stress on the mirror surface still has magnitude $B^2/(2\mu_0)$ at each instant. Averaging its sinusoidal square gives $\langle p \rangle = (2E_i/c)^2/(4\mu_0) = \epsilon_0 E_i^2 = 2I_i/c$, since the incident mean intensity is $\epsilon_0 c E_i^2/2$. A reflector can therefore experience twice the absorbing radiation pressure while dissipating no power. This does not imply a free mechanical power supply: holding a mirror fixed means its

speed and the mechanical power $\boldsymbol{F}\cdot\boldsymbol{u}$ are zero even when its support carries a continuous force. If the mirror moves significantly, the Doppler shift and work on it require an altered reflected field.

Away from the mirror the incident and reflected waves store electric and magnetic energy in spatially alternating regions. At a node of the electric field the magnetic field can be maximal. One quarter wavelength away there is an electric antinode and a magnetic node. The electric field there reaches its temporal maximum a quarter period after the magnetic maximum at the first position. The instantaneous normal Poynting vector generally oscillates and its cycle mean is zero in the lossless standing wave. A detector resolving only the cycle-averaged flux could incorrectly infer that the cavity is empty. A magnetic loop and an electric dipole probe at the same location would give complementary indications of stored energy. They have different coupling and must be calibrated before their raw voltages are compared.

Now consider partial reflection with intensity reflectance $\mathcal{R}$, transmittance $\mathcal{T}$ and absorptance $\mathcal{A}=1-\mathcal{R}-\mathcal{T}$, with vacuum incident and exit regions and no additional scattering channels. At normal incidence the net incident-minus-outgoing momentum flow gives $\langle p\rangle=(1+\mathcal{R}-\mathcal{T})I_i/c=(2\mathcal{R}+\mathcal{A})I_i/c$, while absorbed heating per area is $\mathcal{A}I_i$. If $I_i=300\,\mathrm{W/m^2}, \mathcal{R}=0.60, \mathcal{T}=0.30$, the material absorbs $30\,\mathrm{W/m^2}$ and the pressure is $1.30\,\mu\mathrm{Pa}$. Measuring absorbed power alone would account for only $0.10\,\mu\mathrm{Pa}$; the reversed momentum of the reflected fraction supplies the remainder. If transmitted light emerges at an angle or into a liquid, momentum components and material forces must be included with a surface appropriate to that experiment.

Compare stress and virtual work on a finite electrostatic source

A conductor held at zero potential occupies the lower half-space. Place a point charge q a height a above it. The image method gives the exterior field by adding an opposite charge at $-a$; the conductor's interior field vanishes. Immediately above the conductor, the normal electric field has magnitude $E_n(s)=|q|a/[2\pi\epsilon_0(s^2+a^2)^{3/2}]$ at radial distance s. Its square gives a nonuniform surface attraction $\epsilon_0E_n^2/2$. Integrating it over rings on the full plane checks the force calculation without applying a singular self-field to the point charge.

Specifically, the total attractive force on the conductor has magnitude $q^2a^2/(4\pi\epsilon_0)\int_0^\infty s(s^2+a^2)^{-3}ds$. Substituting $v=s^2+a^2$ gives $1/(4a^4)$ for the integral. The net force is $q^2/(16\pi\epsilon_0a^2)$, exactly the Coulomb force between the real charge and the mathematical image. For $q=1.0\,\mathrm{nC}, a=20\,\mathrm{mm}$, this is $5.62\,\mu\mathrm{N}$. The normal pressure is concentrated near the charge's projection even though the induced surface charge distribution extends indefinitely. Integrating only out to radius b captures a force fraction $1-a^4/(a^2+b^2)^2$. Already at $b=a$ one collects 75 percent of the force.

The electrostatic mechanical potential for the real charge relative to the grounded conductor is $-q^2/(16\pi\epsilon_0a)$, obtained by integrating the attractive force with respect to a. A tempting use of the real-charge–image pair energy $-q^2/(8\pi\epsilon_0a)$ doubles that answer because the image is an induced representation and moves as the real charge moves. The correct assembly work is half $q\phi_{\mathrm{induced}}$ for linear grounded boundary response: as the real charge is built up from zero, its induced potential grows linearly with it. This factor shares its origin with the half in electrostatic field energy, while the force on a test charge already present is the full $q\boldsymbol{E}_{\mathrm{induced}}$. The surface pressure, image force and derivative of

mechanical potential therefore agree when each calculation counts the experiment's degrees of freedom correctly.

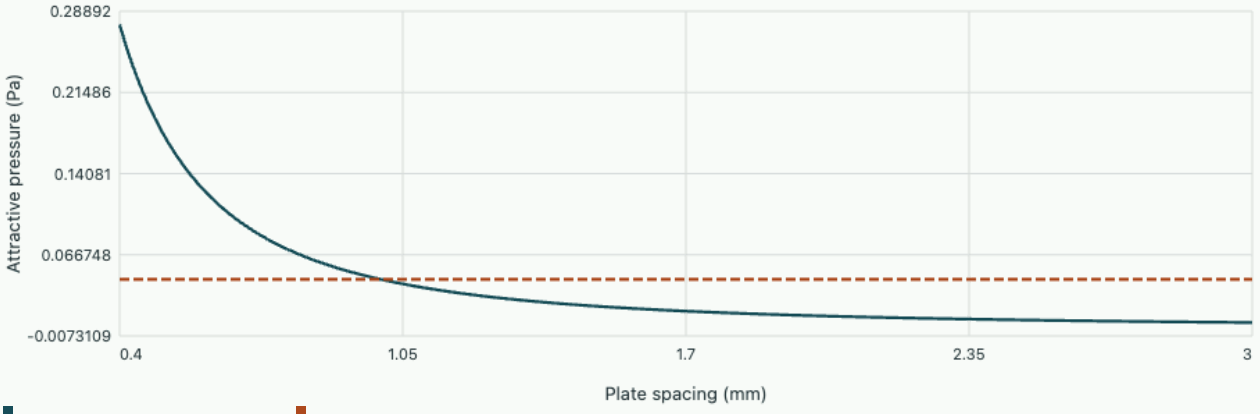
Compare force from stress with energy gradients

Change the capacitor conditions and derive the force independently from field pressure. Work out which external source does work during the displacement.

Open this experiment online

Try the calculation

Reference voltage at 1 mm (V) **100**



At 2 mm the voltage supply gives 0.01107 Pa. With the original charge trapped, the pressure stays at 0.04427 Pa.

Derive the compression of a straight current column

Consider a long nonmagnetic round wire or an ideal cylindrical conducting liquid column of radius a, carrying a uniform axial current I. End regions and any imposed external magnetic field are well removed from the section of interest. For a solid wire the stresses needed to maintain equilibrium depend on elastic properties. A simple fluid model allows us to calculate a scalar pressure difference directly from the magnetic body force. This example distinguishes the total force on a body from substantial internal stresses whose vector sum vanishes by symmetry.

In cylindrical coordinates the current density is

$$J_z = I/(\pi a^2).$$

Applying Ampère's law to an interior circle of radius s encloses the fraction s^2/a^2 of the current and gives

$$B_\phi(s) = \mu_0 I s/(2\pi a^2).$$

The radial Lorentz force density is $f_s = -J_z B_\phi$ because $\hat{z} \times \hat{\phi} = -\hat{s}$. Consequently the magnetic force on a small volume is inward everywhere except on the axis, where the local field vanishes. A static fluid satisfies $0 = -dp/ds + f_s$, so its radial pressure gradient must be negative. Integrating from the axis outwards gives

$$p(s) - p(a) = \frac{\mu_0 I^2}{4\pi^2 a^4}(a^2 - s^2),$$

$$p(0) - p(a) = \frac{B_\phi(a)^2}{\mu_0}.$$

Pressure at the centre must exceed pressure at the rim to balance the inward force. There is no factor one half in this centre-to-rim result. The expression $B^2/(2\mu_0)$ often described as a magnetic pressure is a local Maxwell surface stress in vacuum. Here the field permeates the current-bearing material, and the interior magnetic force is distributed continuously through its volume. The stress on a single outer surface is insufficient to read off the interior fluid pressure profile. Likewise, integrating the radial vector force around a full circular cross section yields zero net translational force, while the scalar pressure difference remains finite.

With $I = 20\,\mathrm{A}$ and $a = 1.0\,\mathrm{mm}$, the edge induction is $4.0\,\mathrm{mT}$, giving a centre-to-rim pressure difference of $12.7\,\mathrm{Pa}$. Doubling the current quadruples the pressure difference; doubling the radius at fixed current reduces it to one quarter. Now consider a thin cylindrical current sheet of the same radius carrying the same total current. Its interior field vanishes, and the exterior field just outside the sheet equals the former edge field. The surface current density $I/(2\pi a)$ experiences the average of the fields on its two sides. The resulting inward surface force has magnitude $B_{\mathrm{out}}^2/(2\mu_0) = 6.37\,\mathrm{Pa}$. The fields outside the two sources agree, yet their internal force distributions differ.

A practical test can exploit the square-law dependence of stress. If $I(t) = I_{\mathrm{pk}} \cos \omega t$, and the magnetic profile follows the current quasistatically with negligible skin effect, the Lorentz force varies as $\cos^2 \omega t = [1 + \cos(2\omega t)]/2$. A pressure or strain sensor can detect a steady signal and a component at twice the current frequency. The actual displacement follows the mechanical response and may resonate or shift in phase. An added external dc field perpendicular to the current gives a force at the current frequency, distinguishable by its angular pattern as well as its frequency.

If the conductor has the same externally measured field at its surface, how could one check whether the uniform-current pressure profile is applicable?

▼ Work through the solution

Estimate the skin depth at the highest significant driving frequency and compare it with the radius. A skin depth comfortably larger than the radius favours a nearly uniform current in a homogeneous conductor far from nearby returns. A much smaller skin depth confines the current to the surface. A measurement of ac resistance or an interior-field measurement in a suitable analogue provides an independent check. A dc current can still be distributed unevenly in a conductor with nonuniform conductivity. Surface field alone measures total enclosed current and cannot distinguish these internal distributions.

A fluid column has an external mechanical pressure p_{ext} and a uniform volume current all the way to the boundary. Must one add $B(a)^2/(2\mu_0)$ as an extra fluid pressure jump at the rim?

▼ **Work through the solution**

In the ideal nonmagnetic volume-current example no singular surface current is present. The tangential induction is continuous just across the rim, so the normal magnetic Maxwell stresses immediately outside and inside are equal there. Ignoring surface tension, other surface forces and fluid motion, the fluid pressure at the rim matches the external mechanical pressure. The inward volume force then builds the pressure toward the centre. In a model with a genuine surface current, the fields on the two sides differ and the magnetic stress jump is real. Checking which current distribution is assumed prevents counting an exterior magnetic traction twice.

David Tong, Electromagnetism and Relativity, on the energy–momentum tensor.

Tutorial on mechanical checks

A lossless test mirror intercepts a pulse of energy $3.0\,\mathrm{mJ}$ at normal incidence. It reflects 75 percent and transmits the remainder in the same direction in vacuum. Calculate its impulse and contrast the result if the remaining quarter is absorbed instead.

▼ **Work through the solution**

Initial momentum is U/c. The outgoing net momentum is $(0.25 - 0.75)U/c = -0.50U/c$, so impulse is

$$1.50U/c = 1.50 \times 10^{-11}\,\mathrm{N\,s}.$$

If the unreflected portion is absorbed, the outgoing momentum is $-0.75U/c$ and impulse rises to

$$1.75U/c = 1.75 \times 10^{-11}\,\mathrm{N\,s}.$$

Absorption changes the energy budget and transfers its own share of incident momentum.

The fields at the opposite sides of an ideal current sheet are $+B_0\hat{\boldsymbol{x}}$ and $-B_0\hat{\boldsymbol{x}}$. Is its own field sufficient to exert a net transverse force on the sheet? What changes in a uniform additional field also along x?

▼ **Work through the solution**

The two self-field magnitudes are equal, so their opposite vacuum-face tractions cancel. An external field changes the total magnitudes differently: the pressure difference is $[(B_{\mathrm{ext}} + B_0)^2 - (B_{\mathrm{ext}} - B_0)^2]/(2\mu_0) = 2B_{\mathrm{ext}}B_0/\mu_0$ in magnitude with orientation set by the actual sheet current. Since $|K| = 2B_0/\mu_0$, this equals $|\boldsymbol{K} \times \boldsymbol{B}_{\mathrm{ext}}|$.

A simulation reports an outward energy flux of 1.2 W and positive work on matter of 0.30 W in a closed fixed region with no other sources there. What must happen to its contained field energy over a 2.0 ms interval if the rates stay constant?

▼ **Work through the solution**

The rate is $-1.2 - 0.30 = -1.50\,\mathrm{W}$, so stored energy must fall by $3.0\,\mathrm{mJ}$. A steady-state claim for the same numbers would violate Poynting's balance unless an additional input or change of sign were identified.

Magnetism, induction and field dynamics · 11. Field momentum, momentum exchange and the quasistatic limits

A system can have static electric and magnetic fields with nonzero local electromagnetic momentum density. A circuit can also be electrically small while either induction or displacement current remains indispensable. Both puzzles become clearer when one follows the relevant conservation law and compares the scales of the discarded terms.

Interpret a local momentum density without invoking radiation

In vacuum $\boldsymbol{g} = \epsilon_0 \boldsymbol{E} \times \boldsymbol{B} = \boldsymbol{S}/c^2$. A plane wave gives an especially simple example because its momentum and energy travel in the same direction at c. The formula itself also applies to near fields, standing fields and stationary crossed fields. A large local $\boldsymbol{g}$ does not imply outward radiation from the entire system: the integral can vanish by symmetry, and neighbouring Poynting vectors can form a circulating pattern. The force law over a fixed region uses the time derivative of the volume integral as well as the Maxwell stress over its boundary.

Consider idealised crossed uniform fields $\boldsymbol{E} = 2.0 \times 10^5 \hat{\boldsymbol{x}}\,\mathrm{V/m}$ and $\boldsymbol{B} = 0.20 \hat{\boldsymbol{z}}\,\mathrm{T}$. The local density is $-3.54 \times 10^{-7} \hat{\boldsymbol{y}}\,\mathrm{N\,s/m^3}$. A $1.0\,\mathrm{litre}$ counted overlap region contributes $-3.54 \times 10^{-10} \hat{\boldsymbol{y}}\,\mathrm{N\,s}$. Treat this number as the integral over that region, not automatically the momentum of the full apparatus: the fringing fields, magnets, sources, support and return currents have not been included. The integrated momentum balance will determine which part changes when a supply is switched off.

Store angular momentum in a coaxial electric field

Take two very long coaxial nonconducting cylindrical shells of radii a, b carrying opposite free charges per length $+\lambda, -\lambda$. Ignore end fields in a central region. For $a < s < b$, the electric field is $\lambda/(2\pi\epsilon_0 s)\hat{\boldsymbol{s}}$; it is negligible elsewhere in the ideal symmetric model. A coaxial long solenoid of radius R with $a < R < b$ creates $B_z = B_0$ for $s < R$ and negligible field outside. In the overlap annulus $\boldsymbol{g} = -\lambda B_0/(2\pi s)\hat{\boldsymbol{\varphi}}$. Integration over angle cancels total linear momentum, yet the z-component of angular momentum density $s g_\varphi$ has the same negative sign everywhere.

$$\frac{L_{\mathrm{field},z}}{\ell} = \int_a^R s g_\varphi \, 2\pi s \, ds = -\frac{\lambda B_0}{2}(R^2 - a^2).$$

The units are angular momentum per length, $\mathrm{kg\,m/s}$. With $\lambda = 2.0\,\mathrm{nC/m}, B_0 = 0.10\,\mathrm{T}, a = 1.0\,\mathrm{cm}, R = 3.0\,\mathrm{cm}$, the field contribution is $-8.0 \times 10^{-14}\,\mathrm{kg\,m/s}$. Its small size makes the thought experiment cleaner than the mechanical measurement, yet the sign can be tested without an ambiguous assignment of a local outward momentum beam.

Now reduce B_0 slowly to zero with the shells free to rotate. Faraday's law gives $E_\varphi = -s\dot{B}_0/2$ for $s < R$, and $E_\varphi = -R^2\dot{B}_0/(2s)$ outside the solenoid. The torque per unit length on a charged shell is $s\lambda_{\mathrm{shell}} E_\varphi$. Time integration gives the inner shell $+\lambda a^2 B_0/2$ and the outer shell $-\lambda R^2 B_0/2$. The combined mechanical angular impulse equals the initial field angular momentum. There was no need to posit a magnetic force on the initially stationary shell charges; the induced azimuthal electric field provides the torque. The calculation assumes negligible friction, enough length for end effects to be small, and shell speeds too low to alter the imposed fields appreciably.

Compare retardation with the strength of two different couplings

Let the apparatus have size L, characteristic variation time T, and typical field magnitudes E_*, B_*. A light-travel time estimate $\eta = L/(cT)$ tests the importance of retardation in vacuum. Faraday's equation compares the induction electric scale LB_*/T against E_*; define $r_F = LB_*/(TE_*)$. Ampère's displacement term compared with a characteristic magnetic curl B_*/L gives

$$r_D = LE_*/(c^2TB_*).$$

Multiplication yields

$$r_F r_D = \eta^2.$$

These are scale estimates, not equalities asserting that every component has its characteristic derivative. Symmetry, very small local fields, and boundary layers can change a local ordering.

When $\eta \ll 1$, at least one of the two ratios is small for the chosen representative scales. In an electroquasistatic description electric fields mainly come from charge distributions, with magnetic induction's correction to their circulation small ($r_F \ll 1$). One can retain displacement current when determining a magnetic field from a charging current. In a magnetoquasistatic description inductive electric fields can dominate, while displacement current contributes negligibly to the magnetic curl ($r_D \ll 1$ in the region of interest). Circuit windings and good conducting bodies often meet this latter condition. Simply setting every time derivative to zero would suppress the very changing flux or charging measurement motivating the approximation.

Carry out two estimates on a physically small apparatus

A parallel-plate vacuum capacitor of radius L with harmonic voltage has $B_\varphi(L) \simeq \mu_0\epsilon_0 L\omega E/2$. Therefore

$$cB/E \simeq \omega L/(2c).$$

With $L = 0.050\,\mathrm{m}$ and $f = 100\,\mathrm{kHz}$, $\eta = \omega L/c = 1.05 \times 10^{-4}$, taking $T = 1/\omega$. Then $r_F \simeq \eta^2/2 = 5.5 \times 10^{-9}$, while $r_D \simeq 2$. The factor two reflects the cylindrical curl geometry; the order-one conclusion is what matters. A scalar-potential calculation of the dominant electric field is excellent here. Dropping displacement current when calculating the small magnetic field in the vacuum gap would give the wrong leading magnetic result.

For a long solenoid with time-dependent uniform interior B_z, Faraday's law gives

$$E_\varphi(L) \simeq L\omega B/2.$$

At the same size and frequency $E/(cB) \simeq \eta/2$, so $r_D \simeq \eta^2/2$ and $r_F \simeq 2$. Induction is now the leading mechanism creating the observed azimuthal electric field. Nevertheless the apparatus is just as electrically small as the capacitor. Comparing the species of source and the requested observable guides the choice

of reduced equations. A single physical circuit can contain a capacitor gap suited to the first analysis and an inductive winding suited to the second.

Preserve causality when an approximation has an instantaneous kernel

The static Coulomb and Biot–Savart integrals use source values at one labelled time. In a slowly changing near zone, evaluating such integrals with instantaneous data may be the leading term of a retarded solution. For a localised source at separation R, source variation enters as $F(t - R/c) = F(t) - (R/c)\dot{F}(t) + \cdots$. A small $L/(cT)$ must apply to the relevant source-to-observer distance and the fastest features actually retained. A step with a nanosecond leading edge can violate the expansion over a metre even if its repetition rate is only one kilohertz. A distant measurement of the radiation of a small antenna cannot inherit the near-zone approximation merely because the antenna itself is small.

For a sinusoidal test $R = 0.30\,\mathrm{m}$ and $f = 10\,\mathrm{MHz}$, the vacuum delay is 1.00 ns and the phase shift is $2\pi f R/c = 0.0629$ rad, or 3.60°. If the measurement tolerates only 1° phase error, this apparently modest delay is already too large. At $R = 0.030\,\mathrm{m}$ the phase error is 0.36°. In a cable the signal speed depends on materials and geometry and can be lower than c. Finally, leading contributions to an observable can cancel by symmetry; then even a small induction or radiation correction can be the first nonzero effect. Estimate the error against the specific signal being measured, not just the largest field anywhere in the apparatus.

Track the angular impulse while a magnetic flux changes

Compute the momentum carried by steady power in a coaxial line

Between perfect concentric conducting cylinders with radii a, b, a TEM line carrying a DC potential difference V and current I has radial electric field $E_s = V/[s\ln(b/a)]$ and azimuthal magnetic field $H_\varphi = I/(2\pi s)$. Use an insulated nondispersive vacuum region and suppose the return carries the opposite current. The Poynting vector is axial and has magnitude $VI/[2\pi s^2 \ln(b/a)]$. Integrating over its annular cross section gives VI because $\int_a^b ds/s = \ln(b/a)$. The magnetic and electric fingerprints are time independent after a DC state is established, yet power flows from the source to a load continuously. Looking only for a time derivative in the local fields would miss this steady energy transport.

Momentum per unit line length in the vacuum field between the conductors is VI/c^2 along the line. At $V = 50\,\mathrm{V}, I = 0.20\,\mathrm{A}$, the power is 10 W and the field momentum per metre is $1.11 \times 10^{-16}\,\mathrm{kg/s}$. In a two-metre section the total field momentum is $2.23 \times 10^{-16}\,\mathrm{kg\,m/s}$. If that section starts or stops transmitting power, changing field momentum belongs in the momentum budget along with stresses at both ends and forces on conductors. A distributed steady system with nonzero power flow can have stationary material when its supports and internal stresses carry the corresponding mechanical momentum balance. One should not infer a centre-of-mass acceleration of the complete apparatus from the field integrand in an arbitrarily selected subvolume.

Compare a matched travelling TEM wave with an imposed DC line state. In vacuum the characteristic impedance is

$$Z_c = (Z_0/2\pi)\ln(b/a);$$

for a pure forward wave $V = Z_c I$. Then magnetic and electric field energies per length are equal and the energy velocity is c. A DC source and independently specified load may have any real ratio V/I compatible with their circuit, so stored electric and magnetic energies need not be equal even though axial power is again VI. With line coefficients L' and C', energy per length is

$$U' = (L'I^2 + C'V^2)/2.$$

Because $L'C' = 1/c^2$, the elementary inequality $L'I^2 + C'V^2 \geq 2|VI|/c$ gives

$$|P|/U' \leq c.$$

Equality holds for $V/I = \pm Z_c$. This is a check on units and a clean example where energy flow and a single pure propagating wave are not interchangeable descriptions.

For a passive matched line with a finite voltage pulse, the electromagnetic momentum of its vacuum portion is U/c in the propagation direction. Two equal counterpropagating pulses have the same local energy sum as two separate pulses but zero net momentum when their integrals agree. During overlap one must combine the electric fields and magnetic fields first. In one transverse dimension the cross terms cancel in the axial Poynting vector and in the sum of the two field energies, yielding the same directional decomposition found from the characteristic fields. Local electric energy alone can show strong interference, so comparing a local electric-field map to the *total* energy sum would manufacture a contradiction where none exists.

Quantify the low-frequency limit of an open line

A lumped open cable looks capacitive at low frequency even though its fields obey the full wave equations. Solve the lossless one-dimensional voltage equation for a line of length ℓ with zero load current at the far end. With $e^{i\omega t}$, one may take $\widetilde{V}(x) = V_L \cos[k(\ell - x)]$, $\widetilde{I}(x) = i(V_L/Z_c)\sin[k(\ell - x)]$, as follows from the lossless telegrapher equations. Therefore $Y_{\text{in}} = i\tan(k\ell)/Z_c$, where $k = \omega/v$. Since $k/Z_c = \omega C'$, the leading term is $i\omega C'\ell$, just the static capacitance per unit length times length. The omitted fractional admittance correction begins at $(k\ell)^2/3$. This is an error estimate for a specified terminal measurement, stricter or weaker than a generic 'small in wavelengths' rule depending on the tolerance.

For $Z_c = 50\,\Omega, \ell = 0.50\,\text{m}, v = 2.0 \times 10^8\,\text{m/s}$, $C' = 1/(vZ_c) = 100\,\text{pF/m}$, so the static input capacitance is 50 pF. At 10 MHz, $k\ell = 0.1571$, giving an apparent capacitance multiplier $\tan(0.1571)/0.1571 = 1.0083$. A one-percent measurement can barely tolerate the lumped approximation. At 50 MHz the same length gives $k\ell = \pi/4$ and multiplier $4/\pi = 1.273$, even though the physical cable geometry and static capacitance have not changed. At a quarter wavelength the ideal open line presents a short input; resolving that resonance is inherently a distributed problem.

The calculation also illustrates why the correct comparison depends on the observable. Travelling from end to end incurs a phase $k\ell$, a first-order effect, while the open input admittance differs from its electroquasistatic approximation only at second order. A line can therefore meet a terminal capacitance tolerance yet fail a tight timing requirement for independently instrumented endpoints. A good test records both reflected waveform and terminal admittance, specifies the same termination for the two models, and checks the leading discrepancy against a frequency sweep. A fit that allows a different effective capacitance at every frequency can reproduce almost any smooth input curve over a narrow interval while concealing the first distributed correction.

Track field momentum during a field change

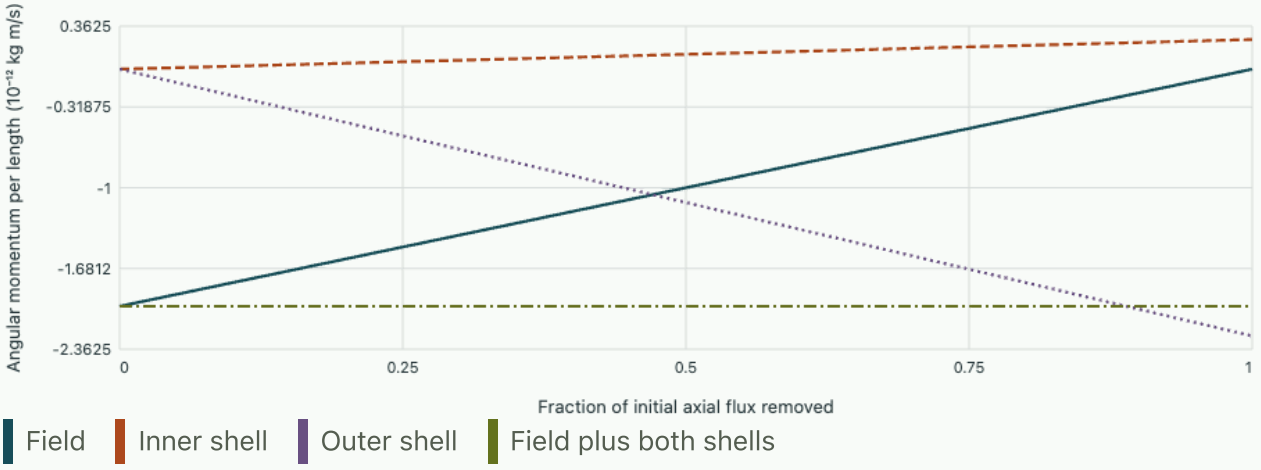
Vary the electric and magnetic fields. Check how a change of field momentum enters a complete mechanical impulse balance.

Open this experiment online

Try the calculation

Outer radius b (mm) **30**

Initial field B (T) **0.25**



After switch-off the inner shell carries 0.25 and the outer shell -2.25 in plotted units. Their sum equals the initial field value -2.

Determine the rest frame of a pair of light pulses

Two short pulses propagate in opposite directions through vacuum and do not overlap at the instant we examine them. A right-moving pulse has energy U_+ and momentum $+U_+/c$. The left-moving pulse has energy U_- and momentum $-U_-/c$. These relations follow by integrating the local densities of a plane travelling wave over its packet. For a broad pulse whose diffraction is negligible one can use them to arbitrary accuracy. Add the energies and the signed momenta before forming the relativistic invariant of the complete two-pulse system. It is

$$U^2 - c^2 P_x^2 = (U_+ + U_-)^2 - (U_+ - U_-)^2 = 4 U_+ U_-.$$

Each ideal one-direction pulse separately has zero invariant mass. The pair has a nonzero invariant mass whenever both energies are positive, because part of its total energy is associated with motion of the packets in opposite directions. A frame exists where the two momenta cancel. Its speed relative to the measuring laboratory is $v = c^2 P_x / U = c(U_+ - U_-)/(U_+ + U_-)$. The centre of energy moves at this speed even though each individual pulse continues to propagate at c. Making one pulse arbitrarily weak sends the centre-of-energy speed toward the speed of the stronger pulse, leaving no finite rest frame in the ideal single-pulse limit.

For $U_+ = 3.0\,\mathrm{mJ}$ and $U_- = 1.0\,\mathrm{mJ}$ the combined energy is $4.0\,\mathrm{mJ}$ and its momentum is

$$2.0\,\mathrm{mJ}/c = 6.67 \times 10^{-12}\,\mathrm{kg\,m/s}.$$

The centre-of-energy velocity is $0.50c$. The rest energy is $\sqrt{12}\,\mathrm{mJ} = 3.46\,\mathrm{mJ}$, corresponding to a mass of about $3.85 \times 10^{-20}\,\mathrm{kg}$. Transform each pulse into that frame as a check. A boost at half the speed of light multiplies the right-moving energy by $\gamma(1 - 0.5) = 1/\sqrt{3}$ and the left-moving energy by $\gamma(1 + 0.5) = \sqrt{3}$. Both then have $1.73\,\mathrm{mJ}$, and their momenta cancel directly.

The same bound has a local form for arbitrary electromagnetic fields. Start with $u = \epsilon_0 (E^2 + c^2 B^2)/2$ and $\mathbf{g} = \epsilon_0 \mathbf{E} \times \mathbf{B}$. The identity $|\mathbf{E} \times \mathbf{B}|^2 = E^2 B^2 - (\mathbf{E} \cdot \mathbf{B})^2$ gives

$$u^2 - c^2 |\mathbf{g}|^2 = \frac{\epsilon_0^2}{4} \left[(E^2 - c^2 B^2)^2 + 4 c^2 (\mathbf{E} \cdot \mathbf{B})^2 \right] \geq 0.$$

Thus the magnitude of the Poynting flux never exceeds cu in vacuum. Equality for a nonzero local field requires perpendicular electric and magnetic fields with $E = cB$, as in a plane travelling wave. Static crossed fields can have a nonzero Poynting vector, yet generally fail to meet that equality. This distinction matters in quasistatic devices. A computation that finds a circulating flux near charges and currents has not, by that observation alone, demonstrated outgoing radiation. To find emitted power measure the net outward energy flux through a surrounding surface at distances where the radiation field dominates.

Put the two ideal pulses on opposite sides of a stationary perfectly absorbing plate, illuminating it simultaneously. Is the impulse on the plate fixed by the total energy alone?

▼ Work through the solution

If the pulses are absorbed and other support impulses are negligible during delivery, the mechanical plus internal momentum of the plate changes by the algebraic sum $P_x = (U_+ - U_-)/c$. Equal incident energies give zero net impulse, even though the plate gains their combined energy. If the two pulse energies differ, reversing the beams reverses the impulse without changing total energy. A subsequent thermal re-emission event introduces additional carried momentum depending on its angular distribution. To compare a complete measured impulse with the initial incoming momentum one must specify the time window and include any significant outgoing flux.

Could a stationary detector sampling both fields at one point determine the invariant mass of an arbitrary radiation packet from the local invariant alone?

▼ **Work through the solution**

The complete energy and momentum are spatial integrals over a common time slice. Squaring their sums produces cross terms between contributions at different locations and directions. In the nonoverlapping example the local invariant expression vanishes wherever a single travelling pulse is present. The invariant of the combined packets is nevertheless positive. A spatially resolving detector or sufficient independent knowledge of the directions and integrated energies is needed to reconstruct the total. The example also explains why one cannot integrate the square root of the local expression and treat that as the invariant mass of an extended electromagnetic system.

David Tong, Electromagnetism and Relativity, on electromagnetic momentum and angular momentum.

Tutorial on choosing what to retain

In the shell calculation, increase the outer shell radius b while keeping $b > R$. Does its final angular impulse change? What happens to the induced tangential impulse per unit charge?

▼ **Work through the solution**

The linked magnetic flux at the outer shell is fixed by R. The time-integrated electric field is proportional to $1/b$, so the tangential impulse per charge decreases. The moment arm b cancels this factor. The outer shell's angular impulse per length is unchanged. The field angular momentum also stays unchanged because its overlap region ends at R.

A measured source has $L = 0.10\,\mathrm{m}$, $\omega = 10^{6}\,\mathrm{s}^{-1}$, and $E_*/(cB_*) = 0.002$. Estimate both induction ratios.

▼ **Work through the solution**

$$\eta = \omega L/c = 3.34 \times 10^{-4}.$$

Hence

$$r_F = \eta/(E_*/cB_*) = 0.167,$$
$$r_D = \eta(E_*/cB_*) = 6.7 \times 10^{-7}.$$

A magnetoquasistatic magnetic-curl approximation is very strong at these scales. A ten-percent electric-circulation accuracy requirement would call for retaining the induction contribution when reconstructing electric fields.

A tiny loop has negligible local retardation, yet its measured signal several kilometres away arrives after an unmistakable propagation delay. Have the near and far measurements contradicted each other?

▼ **Work through the solution**

The small-loop current distribution may be approximated as coherent across the source. This constrains the source size, not the propagation distance to a far detector. A retarded radiation calculation can use a small-source multipole approximation while retaining the full travel delay to the observer.

Magnetism, induction and field dynamics · 12. Posing an electromagnetic evolution problem

A solver cannot evolve 'a pulse in a box' from geometry and a time step alone. It needs material laws, compatible initial fields, a charge-conserving source, and enough boundary data to specify what enters the domain. The same bookkeeping is required when interpreting a real transient measurement.

Separate the evolution equations from the constraints

In vacuum with prescribed sources the evolution system can be written $\partial_t \boldsymbol{B} = -\nabla \times \boldsymbol{E}$ and $\partial_t \boldsymbol{E} = c^2 \nabla \times \boldsymbol{B} - \boldsymbol{J}/\epsilon_0$. The divergence conditions $\nabla \cdot \boldsymbol{B} = 0$ and $\nabla \cdot \boldsymbol{E} = \rho/\epsilon_0$ constrain the initial data. Differentiating the first in time gives zero from the divergence of a curl. Differentiating the second constraint residual gives $-[\nabla \cdot \boldsymbol{J} + \partial_t \rho]/\epsilon_0$. Compatible charge evolution therefore preserves an initially satisfied Gauss constraint. An arbitrarily chosen charge movie and unrelated current movie can violate it before a single boundary reflection occurs.

A numerical solver based only on the two curl equations does not automatically repair bad initial data. In an initial periodic vacuum box a uniform charge density of one sign is inadmissible because the integral of a periodic divergence is zero. One can prescribe a neutralising background as part of a separately justified model, choose net-neutral initial sources, or choose open boundaries that permit net electric flux. Silently discarding the zero Fourier mode of the charge would solve a different problem. Similarly, preserving a constant divergence error exactly during time integration shows consistency of the discrete curls with the chosen derivative, not correctness of the initial physical state.

Construct a right-running pulse from two initial fields

In one-dimensional vacuum choose $\boldsymbol{E} = E_y(x,t)\hat{\boldsymbol{y}},\ \boldsymbol{B} = B_z(x,t)\hat{\boldsymbol{z}}$. Then

$$\partial_t E_y = -c^2 \partial_x B_z,\ \partial_t B_z = -\partial_x E_y.$$

Define $w_+ = E_y + cB_z,\ w_- = E_y - cB_z$. Addition and subtraction give

$$\partial_t w_+ + c\partial_x w_+ = 0,$$
$$\partial_t w_- - c\partial_x w_- = 0.$$

Thus w_+ travels towards $+x$ and w_- towards $-x$. The initial conditions $E_y = E_0 e^{-(x-x_0)^2/(2s^2)}$, $B_z = E_y/c$ specify a right-running packet. Prescribing the same electric shape but $B_z = 0$ launches two oppositely directed electric pulses each of half the original height. Measuring only the initial electric waveform would fail to distinguish the two preparations. At $E_0 = 60\ \mathrm{V/m}$, the peak magnetic field for the first preparation is $0.200\ \mu\mathrm{T}$. An observer 0.90 m farther right sees its peak 3.00 ns later in vacuum.

The component energies provide another independent comparison. Initialising only E_y contains half as much total energy as initialising E_y and its matched B_z. At late times when the two pulses no longer overlap, each half-amplitude pulse has one quarter of the energy of the full matched pulse; the sum is half, as required. In the overlap region it is better to use the characteristic combinations themselves. In these units $u = \epsilon_0(w_+^2 + w_-^2)/4$ and $S_x = \epsilon_0 c(w_+^2 - w_-^2)/4$ for the transverse fields. Equal oppositely running amplitudes can store energy with zero net flux.

Specify only incoming characteristic information at an open end

In $0 < x < L$, w_+ enters at the left end and w_- at the right. Values travelling out are determined by the interior solution. For a pure outgoing plane wave at the right boundary prescribe $w_- = 0$, equivalently $E_y - cB_z = 0$, while allowing w_+ to exit. At a perfectly conducting fixed wall the tangential electric field vanishes, so there $w_- = -w_+$, reflecting the electric amplitude with a sign reversal. A boundary with a real voltage-like reflection coefficient r gives $w_- = r w_+$ in this one-dimensional equal-medium case. The time-integrated returned energy fraction is r^2. Specifying both E_y and B_z independently at all times generally overdetermines a wave boundary and can amount to fixing both incident and reflected waves inconsistently.

Suppose $L = 1.20\,\mathrm{m},\ x_0 = 0.30\,\mathrm{m}$. The centre of a right-running pulse reaches the conducting end at $3.00\,\mathrm{ns}$. A detector at $x = 0.60\,\mathrm{m}$ first sees the incident centre after $1.00\,\mathrm{ns}$, then the reflected centre after $(0.90 + 0.60)\mathrm{m}/c = 5.00\,\mathrm{ns}$. The two electric traces have opposite signs, while their magnetic contributions at the wall add. Swapping the conducting end for a matched boundary removes the reflected peak within this exact one-dimensional model. A local planar absorbing condition on a curved three-dimensional boundary may still reflect oblique waves; the exact statement depended on one normal mode and its wave impedance.

Derive uniqueness from energy

Take two solutions with the same sources, initial data and perfectly conducting closed boundaries in vacuum. Their difference satisfies homogeneous Maxwell equations with zero initial fields and zero tangential electric field at the wall. Poynting's theorem for the difference says $dU_\Delta/dt = -\oint (\boldsymbol{E}_\Delta \times \boldsymbol{H}_\Delta)\cdot \boldsymbol{n}\, dA = 0$ because the flux normal to a conducting wall vanishes. Since U_Δ is a sum of nonnegative squared fields and initially zero, it stays zero. Thus the interior fields agree. A passive outgoing boundary can only remove this difference energy, so the same conclusion holds given no incoming difference signal. This proof shows why specifying incoming boundary data and initial fields correctly is more fundamental than choosing a particular potential gauge.

The argument also supplies a stability target for computation. With no drive and a perfectly reflecting cavity the discrete total energy should remain bounded, and with an absorbing passive end the loss of field energy should match outgoing flux. An unphysical sign in the wall update can inject energy, producing a growing solution even when the interior scheme is stable. Fixing the numerical grid spacing alone cannot cure an active boundary condition. At a material interface Maxwell jump conditions must be applied to the appropriate tangential or normal fields, since setting all Cartesian components continuous would generally violate flux or polarisation physics.

Discretise space and time without destroying wave kinematics

For the one-dimensional system use E_j^n at integer time and integer position, with $B_{j+1/2}^{n+1/2}$ at half steps in both. Integrating Faraday on a small rectangular loop gives

$$B_{j+1/2}^{n+1/2} = B_{j+1/2}^{n-1/2} - (\Delta t/h)(E_{j+1}^n - E_j^n).$$

Ampère then gives $E_j^{n+1} = E_j^n - c^2(\Delta t/h)(B_{j+1/2}^{n+1/2} - B_{j-1/2}^{n+1/2})$ without current. For material sources include the appropriate time-centred current. The staggered layout places discrete curls where the updated component lives. In several dimensions a compatible divergence of the curl vanishes by cancellation of oriented face-edge contributions.

Substitution of waves $e^{i(kjh-\omega n\Delta t)}$ with the matching half-step phases yields $\sin(\omega\Delta t/2) = (c\Delta t/h)\sin(kh/2)$ on the positive-frequency branch. Stability on a one-dimensional uniform grid requires $c\Delta t/h \le 1$. For $h = 5.0\,\mathrm{mm}$, the upper vacuum time step is $16.7\,\mathrm{ps}$. A cubic three-dimensional mesh imposes the more restrictive $c\Delta t/h \le 1/\sqrt{3}$, about $9.63\,\mathrm{ps}$. In one dimension the equality case transports compatible resolved plane waves at the correct speed on this grid. Numerical dispersion generally returns at smaller Courant number or for diagonal propagation in multiple dimensions. Sample a pulse across enough sites that phase and dispersion errors lie below the intended time-delay uncertainty.

A careful initialisation also respects the half time step. Assign $B^{1/2}$ from the analytic field at that shifted time or use a consistent half-step update. Assigning B^0 directly where $B^{1/2}$ is required can create a spurious left-running component in what should have been a one-way pulse. Examine w_- reconstructed at common space and time before attributing a tiny backward peak to the material. For a real calculation, refine both space and time, compare detector arrival times as well as peak amplitudes, and integrate energy flux over complete transit windows. Solver convergence, charge compatibility, and a pulse moving in the intended direction are independent tests.

Compare a matched return and a reflected pulse on a line

Match a travelling pulse at a material boundary

A single homogeneous characteristic impedance was built into the one-dimensional end conditions. Replace the boundary by an interface between lossless homogeneous nondispersive media with positive real impedances Z_1, Z_2. At normal incidence from the left, write electric amplitudes E_i, E_r, E_t and choose the magnetic sign so a right-running wave has $H = E/Z$ and a left-running one $H = -E/Z$. With no free sheet current and no singular electric field at the interface, tangential continuity gives

$$E_i + E_r = E_t,$$
$$(E_i - E_r)/Z_1 = E_t/Z_2.$$

Solving these two equations yields

$$r = \frac{Z_2 - Z_1}{Z_2 + Z_1},$$
$$t = \frac{2Z_2}{Z_2 + Z_1},$$
$$\mathcal{R} = |r|^2,$$
$$\mathcal{T} = \frac{Z_1}{Z_2}|t|^2.$$

The impedance factor in transmitted power is required even when electric continuity makes $t = 1 + r$. With $Z_1 = 300\,\Omega$, $Z_2 = 100\,\Omega$, the electric reflection is $-1/2$ and transmission $+1/2$. The reflected energy is one quarter, while transmitted energy is three quarters. Mistaking both energy fractions for one quarter merely because the two electric amplitudes have equal magnitudes leaves half the input pulse energy unexplained. At $Z_2 \to 0$ the coefficients approach an electric short: the tangential electric field vanishes, even though the magnetic field doubles in the incident medium. Matching the impedances gives zero reflection even if material wave speeds differ, since those speeds depend on the product $\mu\epsilon$ whereas impedance depends on their ratio.

Inject a short pulse with known area and width a sufficient distance to the left that it initially lies entirely in medium one. If the media are nondispersive and impedance independent of frequency across the pulse band, each electric waveform copies the incident temporal shape multiplied by its amplitude coefficient. A transmitted wave travelling at a different speed has a different spatial width for the same temporal width, so comparing stored energy by squaring peak fields and comparing equal geometrical intervals can easily mislead. A fixed detector measuring $\int (EH)\,dt$ across a complete pulse directly measures transmitted energy per area. In a simulation integrate forward and backward flux at separate planes, include any energy still between the planes, and check that the field boundary gives the same answer when the detector positions are changed.

Manufacture a source that obeys charge continuity exactly

A verification case with prescribed charge and current can test different components from a source-free travelling wave. Consider a one-dimensional periodic interval with period $2\pi/k$. Choose a longitudinal electric field $E_x(x,t) = E_a \sin(kx)\cos(\omega t)$ and zero magnetic field. Gauss's law fixes $\rho = \epsilon_0 k E_a \cos(kx)\cos(\omega t)$. Ampère–Maxwell requires

$$J_x = -\epsilon_0 \partial_t E_x = \epsilon_0 \omega E_a \sin(kx)\sin(\omega t).$$

Faraday gives no magnetic-field change because a purely x-directed function of x has zero curl. Differentiating the two source expressions verifies $\partial_t \rho + \partial_x J_x = 0$, and the periodic spatial mean of the charge vanishes. This is a prescribed-source test of the equations, not a claim that an unspecified material spontaneously supplies those currents.

For period $0.20\,\mathrm{m}$, $E_a = 100\,\mathrm{V/m}$, and $f = 2.0\,\mathrm{MHz}$,

$$k = 31.42\,\mathrm{m^{-1}}.$$

The charge-density amplitude is $27.8\,\mathrm{nC/m^3}$ and the free-current-density amplitude $0.0111\,\mathrm{A/m^2}$. At $t = 0$ the longitudinal electric energy is maximal and current vanishes. One quarter period later the electric field and charge are zero while the source current is maximal at the spatial electric-field antinodes. At every point where E, J are nonzero their product is $-\partial_t(\epsilon_0 E^2/2)$; the prescribed external charges exchange energy with the field without a magnetic Poynting flux. A discrete divergence residual and an independently measured field-energy change can therefore check the signs and amplitudes of the source implementation.

Prescribing those continuous formulas sampled separately on a staggered mesh does not necessarily produce exact *discrete* continuity. If the numerical current is sampled at half time and half space, integrate the cell continuity equation to update ρ from numerical boundary currents. Alternatively use the corresponding discrete sinusoidal differentiation factors $2\sin(kh/2)/h$ and $2\sin(\omega\Delta t/2)/\Delta t$ when manufacturing exact discrete data. Comparing the continuous and discrete constructions separates a model truncation error that converges with mesh refinement from a genuinely nonconserving source injection. A code can look excellent on transverse vacuum pulses and still fail this test because those pulses never challenge its charge handling.

Decide what a late echo means

A narrow pulse incident on a material slab can return signals from both interfaces. Separate the two in time when possible. If the slab thickness is d and its propagation speed v_2, the round-trip delay inside it is $2d/v_2$. Later echoes spaced by that delay suggest repeated internal reflections. Changing the outer computational boundary location while keeping the material and detector fixed provides a different test: an artificial wall echo moves in time by twice the boundary displacement divided by exterior speed. A genuine material echo keeps its material-defined delay. A calibrated single detector trace and these two perturbations can thus identify which part of the model causes a late pulse.

For $d = 0.30\,\mathrm{m}$ and $v_2 = 1.5 \times 10^8\,\mathrm{m/s}$ the material echo spacing is $4.0\,\mathrm{ns}$. Moving an external vacuum boundary outward by $0.45\,\mathrm{m}$ would delay a returning boundary echo by $3.0\,\mathrm{ns}$. If the incident pulse lasts 8 ns the echoes overlap strongly and individual peaks are difficult to assign. Broadening the bandwidth with a shorter source or fitting the full complex frequency response is then more informative than selecting the largest local maximum. Any comparison must include the source bandwidth and detector response because a perfectly converged field solution can still predict a different voltage from a finite-bandwidth probe.

Check a discretised field evolution against its phase

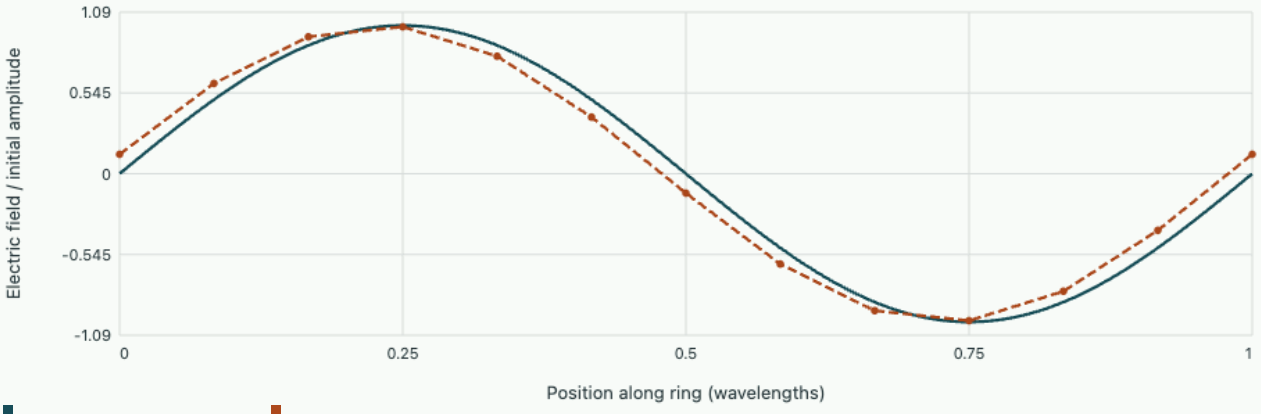
Keep the equations and initial mode fixed while refining the grid. Examine numerical phase delay alongside the physical propagation law.

Open this experiment online

Try the calculation

Points per wavelength 12

Courant number 0.8



Continuous wave Time-stepped nodes

75 steps, 5 physical periods. Phase speed ratio 0.99582. Accumulated phase lag 7.5244°.

Set a cavity ringing from specified initial fields

A one-dimensional lossless cavity consists of perfectly conducting parallel planes at $x = 0$ and $x = L$. Use a polarisation E_y, B_z and ignore lateral edges. With no sources inside, the incoming characteristic at each wall is set by the condition $E_y = 0$. Instead of injecting a pulse during the calculation, prescribe an initial field $E_y(x, 0) = E_0 \sin(\pi x/L)$ and $B_z(x, 0) = 0$. These data obey the tangential electric wall condition. They also obey the divergence constraints because the spatial variation is along x while the fields are transverse. Initial data that satisfy a wave equation but violate either of these checks would model a different physical preparation.

Let $k = \pi/L$ and try $E_y(x, t) = E_0 \sin(kx)F(t)$. The wave equation gives

$$F'' + c^2k^2F = 0.$$

The initial magnetic field has no spatial gradient and the source-free Maxwell–Ampère law therefore sets $F'(0) = 0$. It follows that $F = \cos \omega t$, with $\omega = ck$. Use Faraday's first-order equation to find the sign and phase of B_z, obtaining

$$E_y = E_0 \sin(kx) \cos \omega t,$$

$$B_z = -\frac{E_0}{c} \cos(kx) \sin \omega t.$$

For a first-order time solver this pair gives a stronger check than the electric standing-wave frequency alone. A code with an incorrect sign in one curl can reproduce an electric snapshot but will transfer energy incorrectly on the next advance. The signs of the magnetic fields at the two ends are opposite. Their amplitudes peak one quarter period after the initial electric maximum. Perfectly conducting walls carry the

necessary surface currents. The standing solution has no Poynting flux through them because the tangential electric field at each wall vanishes.

Per unit transverse area the total initial energy is $\epsilon_0 E_0^2 L/4$, since the integral of $\sin^2(kx)$ over the cavity is $L/2$. At a general time the electric share is that total multiplied by $\cos^2 \omega t$, and the magnetic share is the same total multiplied by $\sin^2 \omega t$. The integral of their sum is constant, although energy continually flows within the cavity. For $L = 0.30\,\mathrm{m}$ the fundamental frequency is $c/(2L) = 499.65\,\mathrm{MHz}$ and its quarter period is about $0.500\,\mathrm{ns}$. If $E_0 = 100\,\mathrm{V/m}$, the stored energy is $6.64\,\mathrm{nJ/m^2}$ and the peak induction is $0.334\,\mu\mathrm{T}$.

The experiment can also reveal the discrete wave speed. A staggered one-dimensional mesh with spacing $h = L/10$ and Courant number $C = c\Delta t/h = 0.5$ obeys $\sin(\omega_{\mathrm{num}}\Delta t/2) = C\sin(kh/2)$. The predicted frequency ratio is $2\arcsin[C\sin(kh/2)]/(Ckh)$, approximately 0.9969 here. Measuring too short a time interval to distinguish this phase lag from sampling uncertainty would give little information about spatial accuracy. Track many periods, use consistent time-centred field estimates, and halve h at fixed C. At these small kh, the leading phase error should decrease by roughly a factor of four.

A student prepares a mixture $E_y(x,0) = E_0[\sin(\pi x/L) + 0.30\sin(3\pi x/L)]$, still with initial $B_z = 0$. What should a midplane electric probe observe, and what fraction of the initial total energy is in the third mode?

▼ Work through the solution

At the midplane the fundamental sine equals $+1$, while the third sine equals -1. The probe therefore reads $E_0[\cos(\omega t) - 0.30\cos(3\omega t)]$. Integrated over the full cavity, cross terms vanish because sine modes of different integer index are orthogonal. The third mode's energy fraction is

$$0.30^2/(1 + 0.30^2) = 0.0826.$$

A single spatial probe samples mode amplitudes times their local shape and cannot identify an energy fraction directly. At the midplane all even modes have a node and would escape detection even if they contained substantial energy.

A solver allows an arbitrary spatially uniform constant to be added to B_z while leaving the computed oscillating electric field unchanged. Is this necessarily a bug?

▼ Work through the solution

A constant transverse magnetic induction has zero curl and divergence in the ideal interior. It can be added as a static bias if the physical initial magnetic field and the exterior wall-current preparation permit it. In the test posed above the initial magnetic field was explicitly specified as zero, so a later spontaneously generated offset would be a numerical error. Monitoring the spatial mean of B_z provides a direct test: integrating Faraday's law over the interval gives its rate of change proportional to the difference of the wall electric fields, which is zero. This check tracks an initial-data constraint invisible to a comparison of electric oscillation frequencies alone.

Hermann Haus and James Melcher, Electrodynamic Fields, on wave propagation and standing waves.

Tutorial on initial and boundary compatibility

A vacuum pulse has at its centre $E_y = 100\,\mathrm{V/m}$ and $cB_z = 60\,\mathrm{V/m}$. If both fields have the same spatial shape, infer the two travelling electric amplitudes and the ratio of the energies associated with the two directions.

▼ **Work through the solution**

The characteristics are $w_+ = 160$ and $w_- = 40\,\mathrm{V/m}$. Since $E = (w_+ + w_-)/2$, the component electric amplitudes are 80 and 20 V/m. Their separate directional energy ratio is $160^2/40^2 = 16$, assuming identical envelopes. The total local flux initially points right although neither characteristic is zero.

A numerical field in a periodic box starts with $\nabla \cdot \boldsymbol{B} = 10^{-4}$ in some cells. The code uses a discrete divergence of its discrete curl that is identically zero and carries that pattern unchanged. Does reaching a small time-step tolerance guarantee physically acceptable initial conditions?

▼ **Work through the solution**

No. The curl evolution preserves its initially supplied divergence. A compatible projection or constrained initialisation must remove the error in the correct discrete divergence norm, subject to boundary conditions and any physical mean field. Refining the time integration of inconsistent initial data leaves inconsistency.

A detector sees an amplitude echo 0.15 of an incident Gaussian pulse after reflection at a passive boundary. What is the reflected energy fraction and how much of the input energy remains unaccounted for if measured transmitted energy is 92 percent?

▼ **Work through the solution**

Same-medium, same-shape electric amplitudes give $0.15^2 = 0.0225$ reflected energy. With 0.92 transmitted, the remainder is $1 - 0.0225 - 0.92 = 0.0575$ or 5.75 percent. It may include absorption, energy still in the structure, lateral flux or measurement error. A completed energy audit distinguishes these possibilities.

Magnetism, induction and field dynamics · 13. Why does a Helmholtz pair have a useful uniform region?

A magnetometer averages the field over a finite sample. Derive the coil spacing that suppresses curvature and predict how rapidly the remaining calibration error grows away from the midpoint.

Cancel a spatial derivative with geometry

For an ideal circular loop, Biot–Savart gives the axial field $B(z) = \mu_0 N I R^2/[2(R^2+z^2)^{3/2}]$. It follows by adding equal axial components of the current elements around the circumference. Place two identical loops at $\pm s/2$, carrying currents in the same sense. Superposition and mirror symmetry remove every odd derivative at the midpoint.

Differentiate the single-loop shape $f(u) = (R^2+u^2)^{-3/2}$ twice. The result is $3(4u^2-R^2)(R^2+u^2)^{-7/2}$. The second derivative of the pair changes sign at $s = R$. At this spacing both the linear and quadratic spatial errors vanish. Expanding the even sum one order farther gives

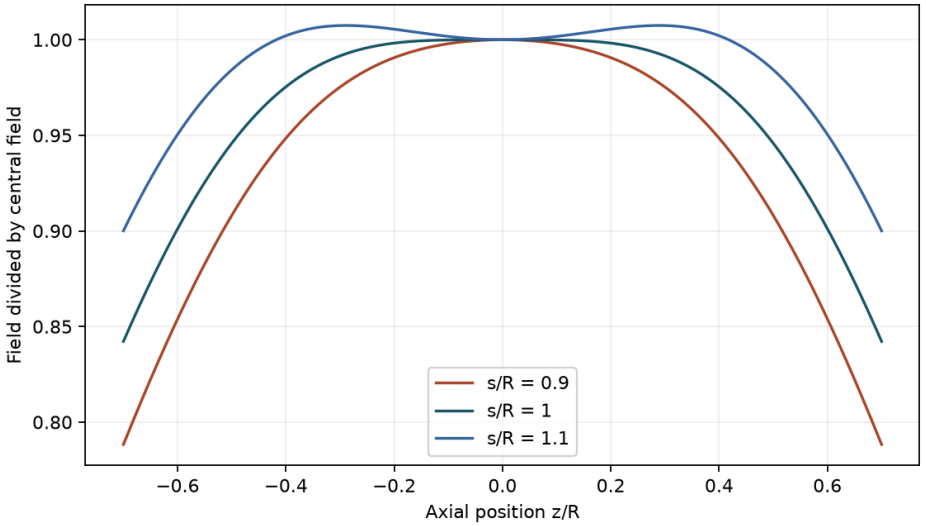
$$\frac{B(z)}{B(0)} = 1 - \frac{144}{125}\left(\frac{z}{R}\right)^4 + O((z/R)^6).$$

Flatness concerns a region of space. A high field value at the centre alone cannot establish that a sample extending on both sides was excited uniformly.

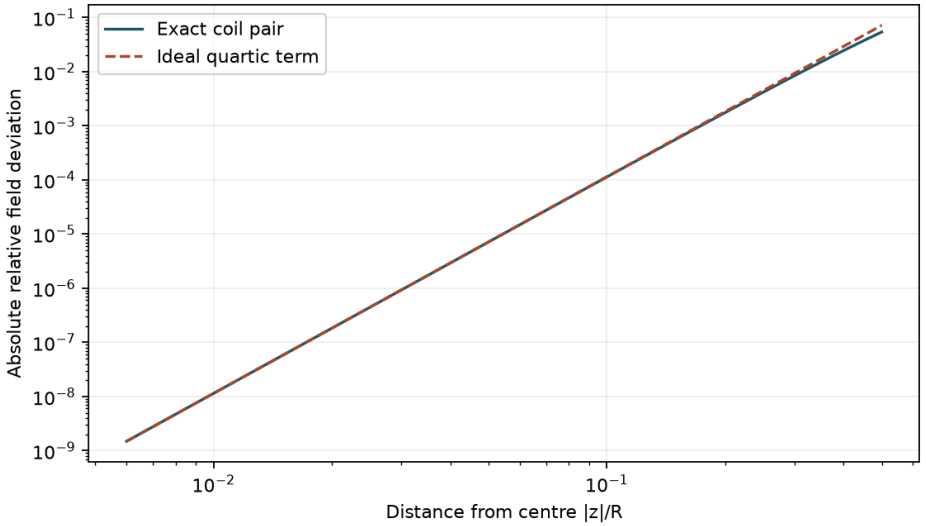
Set the acceptance criterion before choosing a coil

If the allowed pointwise fractional variation is 10^{-3}, the asymptotic half-length of the uniform region is $R(125\times10^{-3}/144)^{1/4} \simeq 0.172R$. With a 100 mm radius this suggests 17 mm on either side. Compare that estimate with the exact field in the notebook. Small spacing errors restore the quadratic term. The accuracy of the coil fixture can therefore matter more than the current-meter resolution when calibrating an extended sample. At 100 ampere-turns per coil and $R = 0.100$ m, the central field is $(4/5)^{3/2}\mu_0 N I/R = 0.899$ mT. A point 10 mm off centre experiences a leading relative deviation of 1.152×10^{-4}, about 0.104 µT in absolute field. A Hall probe with resolution 1 µT would struggle to distinguish a single such deviation. The pointwise specification and the spatial resolution of a verification scan therefore belong in the same calculation.

For a uniformly sensitive sample extending from $-L$ to L, average the quartic term instead of substituting the outermost coordinate. Since the average of z^4 is $L^4/5$, the leading bias of its mean response is a fifth of the edge deviation. A sample that fills the nominal tolerance region can report an apparently excellent mean field while a process sensitive to the edge already operates at the specified limit.



Offset the separation by a tenth of the radius in each direction. The signs of the resulting calibration errors reverse close to the centre.



At ideal spacing the error falls sixteenfold when the sample half-length halves. Move the spacing off its design value to recover quadratic scaling.

Exercise · compare two sample lengths

For an ideal pair, predict the leading error at $z/R = 0.1$ and 0.2.

▼ **Work through the solution**

The quartic terms give 1.152×10^{-4} and 1.8432×10^{-3}. Higher powers change the latter more strongly. Comparing the values on a linear vertical axis can conceal the rapid improvement near the centre.

Exercise · vary the winding current

Will doubling both currents double the length of the region satisfying a relative uniformity tolerance?

▼ **Work through the solution**

Both $B(z)$ and $B(0)$ scale with current, so their ratio and the accepted length stay fixed. An absolute field-error criterion would give a different answer.

Try it · Em coils

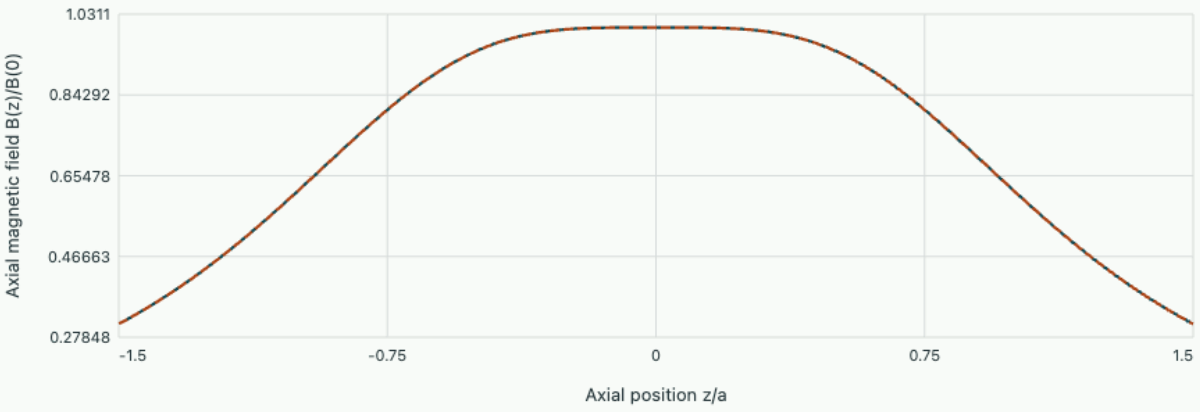
Explore the same physical relation with adjustable inputs. This independent parameter study has its own displayed reference settings; compare its assumptions and units before matching the static example.

Open this live calculation

Try the calculation

Coil separation / radius **1**

Axial test position / radius **0.25**



Selected pair Helmholtz spacing d/a = 1

At z/a = 0.25 the selected field differs from the central field by -0.42%. The sign distinguishes a central peak from a central dip close to the origin.

Download the corresponding editable notebook. The browser recalculates the reduced equations directly; the notebook exposes the full calculation and its checks.

Open editable notebook Download .ipynb

Magnetism, induction and field dynamics · 14. From atomic motion to magnetic susceptibility

Calculate the response of a bound orbit and of an ensemble of existing moments, then follow how interactions permit spontaneous magnetisation.

Perturb an orbit at fixed restoring potential

Use a classical bound electron as a model for the induced orbital response. The field is weak enough that the binding potential remains the dominant influence. A circular orbit of radius r, angular speed ω_0 and electron mass m has an inward restoring force $mr\omega_0^2$. Switch on a uniform axial field B slowly. The magnetic force changes the radial balance, and the induced electric field during switching changes the angular momentum. At first order, with the unperturbed radius used in the magnetic correction, the frequency shift has magnitude eB/(2m). The sign of the induced magnetic moment opposes the applied field for either sense of the original orbit.

$$\Delta\omega = \frac{eB}{2m},$$

$$\Delta m_z = -\frac{e^2 r^2 B}{4m}.$$

The factor one half in the frequency can be recovered by inserting $\omega=\omega_0+\delta\omega$ in the radial balance and keeping $2\omega_0\delta\omega$. The magnetic term is proportional to $\omega_0 B$. The orbit's magnetic moment follows from current times area, with the negative charge determining its direction. For an isotropic shell, the squared radius transverse to the field averages to two thirds of the full squared radius.

$$\chi_{\mathrm{dia}} = -\frac{\mu_0 N e^2}{6m}\sum_j \langle r_j^2\rangle.$$

Here N is the atom number density and the sum runs over the bound electrons in an atom. The susceptibility is negative. This elementary closed-shell result is largely independent of temperature if the atomic size stays fixed. It describes bound orbital response, and its microscopic model should be specified when comparing with conduction-electron magnetism or superconducting screening.

> **Recall · Current direction and the magnetic moment of an orbit**
>
> A charge q circulating with period $2\pi/\omega$ produces current $q\omega/(2\pi)$. Multiplication by the oriented orbit area yields $m=q\omega r^2/2$. Its mechanical angular momentum is $L=mr^2\omega$. Therefore the orbital ratio is $q/(2m)$. Intrinsic spin needs its own gyromagnetic factor and a quantum state description.

Align pre-existing moments statistically

A permanent magnetic moment μ has energy $-\mu\cdot B$. For freely rotating classical moments the orientation integral is identical to the permanent electric dipole calculation, with $x=\mu B/(kBT)$. The density of moments N gives $M=N\mu L(x)$. In weak fields $B\approx\mu_0 H$ for a dilute sample, and expansion yields Curie's law.

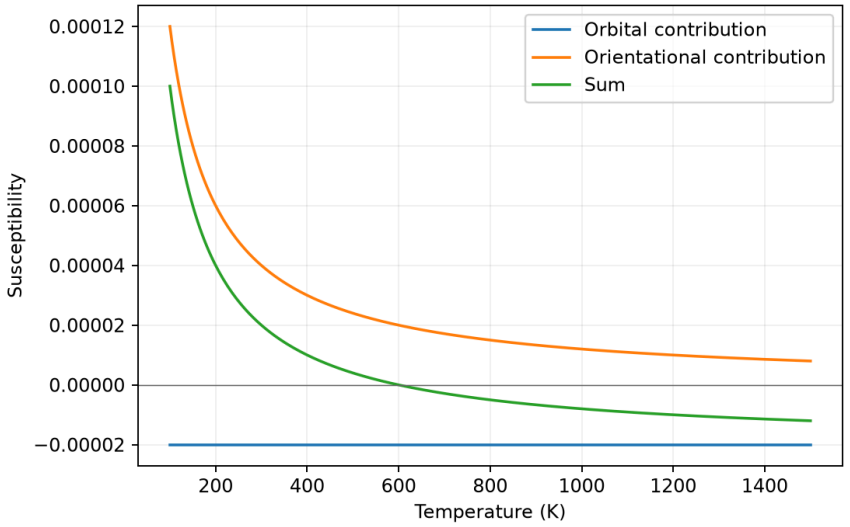
$$\chi_{\text{para}} = \frac{\mu_0 N \mu^2}{3 k_B T}.$$

For quantum angular momentum magnitude J, label the magnetic-moment projections by m=−J,…,J. For an electron with positive g this m is the negative of its angular-momentum projection. The magnetic energy is −gμB mB and the partition function is a finite sum. Expanding the exponential at weak field uses Σm=0 and Σm²=(2J+1)J(J+1)/3.

$$Z = \sum_{m=-J}^{J} e^{g\mu_B m B/(k_B T)},$$

$$\chi = \frac{\mu_0 N (g\mu_B)^2 J(J+1)}{3 k_B T}.$$

For J=1/2 the exact moment is M=N(gμB/2)tanh[gμBB/(2kBT)]. It approaches saturation as temperature falls at fixed field. With μ=μB, T=10 K and B=1 T the argument is 0.06717. A spin-half ensemble is only about 6.71 percent aligned. The classical freely rotating model gives a different weak-field coefficient because its accessible orientations differ.



The two weak-field contributions in exercise 1 have different temperature dependence. Their sum changes sign at 600 K.

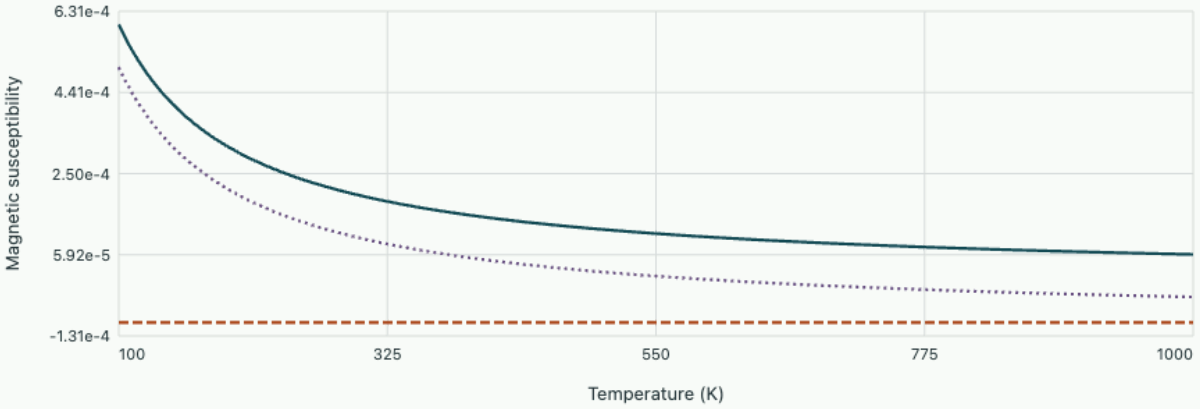
Try it · Curie dia

Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Curie coefficient C (K)	**0.06**
Diamagnetic susceptibility magnitude	**1.00e-4**



Sign change at T = 600 K within the independent-dipole, weak-field model. An ordering transition would invalidate extrapolation of this law.

Allow moments to change each other's effective field

Represent interactions at mean-field level by H_eff=H+λM. Above the transition, insert this in M=(C/T)H_eff and solve the linear equation.

$$\chi = \frac{C}{T - C\lambda} = \frac{C}{T - T_c}.$$

The denominator identifies where the zero-magnetisation solution ceases to be stable. A spin-half model at zero external field gives m=tanh[(Tc/T)m] for reduced magnetisation m=M/Ms. Expand tanh to cubic order near Tc. Besides m=0, the nonzero branch satisfies $m^2 \approx 3(Tc/T-1)/(Tc/T)^3$. This predicts a square-root onset within this approximation.

Remember · A mean-field transition and a measured hysteresis loop answer different questions

The equilibrium equation predicts possible uniform order. Hysteresis requires a history-dependent route through metastable configurations, often involving domain nucleation and motion. A multi-domain specimen can have large local magnetisation and small net moment. A fitted Curie–Weiss temperature alone does not determine coercivity.

A continuum model makes the domain scale explicit. For an easy-axis magnet use energy density $A(\theta')^2+K\sin^2\theta$ in a one-dimensional wall. Translation invariance gives the first integral $A(\theta')^2=K\sin^2\theta$ for a wall joining $\theta=0$ and π. Integration yields $\tan(\theta/2)=\exp[(x-x_0)/\delta]$, with $\delta=\sqrt{(A/K)}$. Substitution in the energy integral gives wall energy per area $4\sqrt{(AK)}$. Exchange penalises rapid variation, whereas anisotropy penalises spending distance away from the easy directions. Magnetostatic energy and sample shape determine whether forming walls lowers the total energy.

Use temperature and shape to discriminate models

1. A sample has $\chi=-2\times10^{-5}+0.012\ \mathrm{K/T}$. Determine the temperature where the net weak-field response changes sign.

▼ **Work through the solution**

Set both terms equal in magnitude. $T=600$ K. At 300 K $\chi=+2\times10^{-5}$, while at 1200 K $\chi=-1\times10^{-5}$. This inference assumes a fixed density and no transition over the interval.

---

2. For $A=10^{-11}$ J/m and $K=10^{4}$ J/m³ find the wall length scale and energy. Predict the effect of quadrupling K.

▼ **Work through the solution**

$\delta=31.6$ nm and wall energy 1.265 mJ/m². Quadrupling K halves δ and doubles the wall energy. The full conventional wall width may include a factor π, so specify the definition when comparing measurements.

---

3. Derive the measured susceptibility of a uniformly magnetised ellipsoid with demagnetising factor N_d and intrinsic χ.

▼ **Work through the solution**

Use $H_{int}=H_{applied}-NdM$ and $M=\chi H_{int}$. Thus $M/H_{applied}=\chi/(1+Nd\chi)$. A strong measured temperature dependence may be suppressed by specimen geometry. Inverting the expression requires the same axis and shape convention used to determine N_d.

Magnetism, induction and field dynamics · 15. The electric field of a rotating magnet

Use charge balance and a surface matching calculation to predict a voltage around a rotating magnetic body, then close the mechanical and electrical power accounts.

Specify the moving material and the electrical constraint

Consider an electrically isolated conducting sphere of radius a, with prescribed uniform axial magnetisation M, rotating slowly at angular speed Ω about that axis. Work to first order in Ωa/c. Neglect deformation and the field produced by the small convective electric current. The magnetostatic interior field is uniform, $B_0=2\mu_0 M/3$. Uniform rotation about the magnetisation axis leaves that magnetic field stationary in the laboratory.

At a material point with cylindrical radius s the velocity is Ωs eϕ. For a highly conducting body in stationary open circuit, mobile charges rearrange until the local motional force vanishes. This gives a laboratory electric field even though the magnetic field is steady.

$$\mathbf{E}+\mathbf{v}\times\mathbf{B}_0=0,$$

$$E_s=-\Omega B_0 s,$$

$$V_{\mathrm{in}}=\frac{\Omega B_0}{2}s^2+C.$$

Gauss's law then gives the total volume charge density $-2\varepsilon_0\Omega B_0$. Its volume integral is $-8\pi\varepsilon_0\Omega B_0 a^3/3$. Since the sphere was neutral, the surface must carry the opposite total charge. In a magnetic material the assignment between free and bound charge needs a constitutive convention. The equations here determine the total charge that produces E.

> **Recall · A moving conductor need not be equipotential in the laboratory**
>
> Static equipotential reasoning assumes charges are at rest in the chosen frame. A charge carried with a rotating conductor feels q(E+v×B). In open circuit it is this total force that vanishes. Stationary voltmeter leads and brushes therefore matter when defining the measured voltage.

Solve the exterior problem using neutrality

At r=a, $\sin^2\theta=2[1-P_2(\cos\theta)]/3$. The surface potential is therefore a constant plus a quadrupole. A neutral isolated sphere has no exterior monopole, so the constant spherical harmonic must vanish when V tends to zero at infinity. This fixes $C=-\Omega B_0 a^2/3$ and determines the exterior field uniquely.

$$V_{\mathrm{out}}=-\frac{\Omega B_0 a^5}{3r^3}P_2(\cos\theta).$$

Differentiate radially on each side. The interior radial field at the surface is $-\Omega B_0 a\sin^2\theta$, while the exterior radial field is $-\Omega B_0 a P_2$. The normal jump gives the total surface charge.

$$\sigma(\theta) = \epsilon_0 \Omega B_0 a \left[\frac{2}{3} - \frac{5}{3} P_2(\cos\theta)\right].$$

Integrating the constant term over the sphere cancels the volume charge exactly, since the angular integral of P_2 vanishes. The tangential derivative of the continuous potential also matches. These are independent checks on the sign and coefficient.

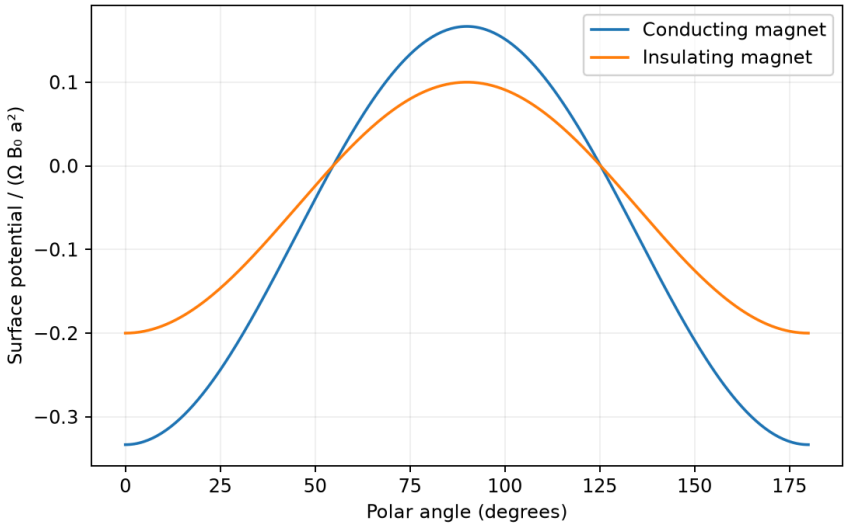
For a=0.02 m, B_0=0.30 T and Ω=100 s^{-1}, the potential difference between equator and pole is $\Omega B_0 a^2/2$=6.0 mV. A laboratory voltmeter with stationary leads and sliding contacts measures that difference in the ideal open-circuit model. Grounding the sphere through a stationary brush at a specified latitude fixes the potential at that contact. This changes the constant harmonic and can supply a net charge.

Track the different constraint in an insulator

An insulating magnet cannot freely redistribute charge until E+v×B vanishes. To leading order, a moving magnetisation contributes electric polarisation P=v×M/c^2. For axial rotation this points radially outward with magnitude ΩMs/c^2. Its bound volume charge is $-2\Omega M/c^2$, and its bound surface charge is $\Omega M a \sin^2\theta/c^2$. Integrating them cancels the net charge. Solve Poisson's equation with these specified sources to determine the field. A particular interior potential is $\Omega M r^2/(3\varepsilon_0 c^2)$. Write the remaining interior quadrupole as D $r^2 P_2$ and the exterior one as $B_2 P_2/r^3$. Continuity of potential gives B_2=D a^5. The quadrupole part of the radial field jump is $5DaP_2$. The prescribed surface charge has quadrupole part $-2\Omega M a P_2/(3c^2)$, so $5Da=-2\Omega Ma/(3\varepsilon_0 c^2)$. Solving for D and substituting B_2 gives the exterior potential below.

$$V_{\mathrm{out,ins}} = -\frac{2\Omega M a^5}{15\epsilon_0 c^2 r^3} P_2(\cos\theta).$$

For the same M and Ω, the exterior quadrupole is three fifths of the conducting result. The distinction is a physical charge-mobility constraint. Rotation of the magnetic pattern alone cannot decide the voltage. The derivation assumes the prescribed magnetisation and ignores additional electric susceptibility of the insulating material.



The exterior quadrupole evaluated at the surface for identical magnetisation and rotation. Electrical mobility changes the amplitude.

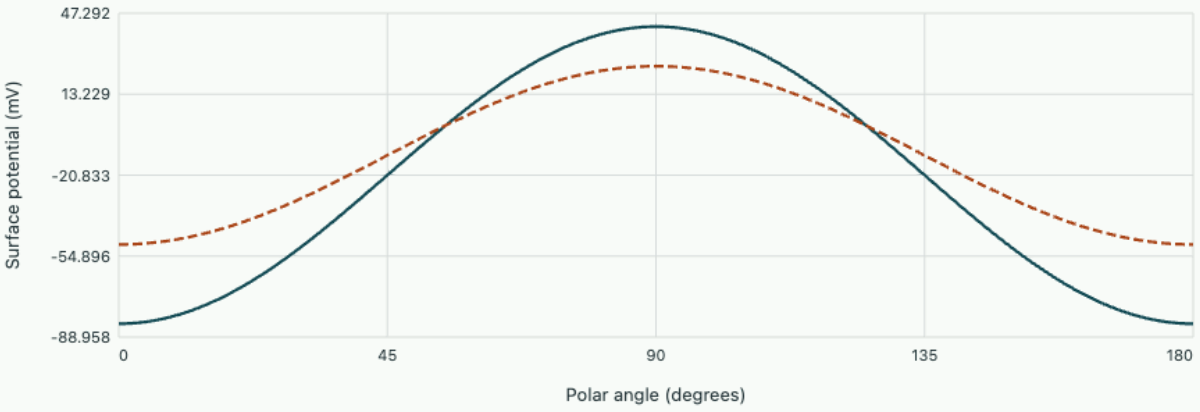
Try it · Rotating magnet

Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Rotation speed (rad/s)	100
Radius (cm)	5
Internal induction (T)	1



Conducting sphere | Insulating sphere, same magnetisation

Exterior quadrupole evaluated at r = a. Potentials are measured relative to infinity; B0 = 2μ0M/3. Insulating/conducting quadrupole ratio is 3/5.

Account for a working generator

Connecting a load permits a current and changes the open-circuit field. In the related conducting-disc model with imposed uniform B, an outward radial current I experiences a braking torque $d\tau_z = -IBs\,ds$ for positive axial B. Integration from axis to rim gives the braking torque $-IBa^2/2$. The driver supplies the opposite torque. Multiplying that driving torque by Ω yields mechanical input power $I\Omega Ba^2/2 = I\mathcal{E}$. The driver supplies the electrical power.

Remember · The magnetic force on each charge is perpendicular to its full velocity

Write a carrier velocity as the material velocity plus drift. The magnetic force transfers momentum between moving lattice and drifting carriers. Its dot product with the full carrier velocity is zero. Mechanical constraints, electric fields and collisions perform the work appearing in the machine-level energy balance. Follow both carrier and body when resolving that balance.

Test a hypothetical magnetic source consistently

In source-free vacuum define F=E+icB. The curl equations combine into i∂tF=c curl F and div F=0. Multiplying F by a constant phase rotates the electric and magnetic fields into one another while preserving the source-free equations. Ordinary electric charges select which combination couples to matter.

A symmetric hypothetical source theory can be written by assigning a magnetic charge the same units as electric charge. Then div(cB)=ρm/ε_0 and curl E=−∂tB−Jm/(ε_0c), while the electric equations retain their usual SI form. Taking a divergence of the modified Faraday equation gives ∂tρm+div Jm=0. The magnetic source convention has to be declared before using a force or energy formula. No magnetic monopole is assumed in the preceding sphere calculation.

Exercises on the observed voltage

1. Reverse Ω while leaving the magnetisation fixed. Which measured quantities change sign?

▼ **Work through the solution**

V, E and the leading induced electric charges change sign. The leading magnetostatic B stays fixed. Reversing both Ω and M leaves the electric signal unchanged.

---

2. A sensor observes the quadrupole at r=2a and at r=4a along the axis. Find the potential and field ratios.

▼ **Work through the solution**

The potential varies as r^{-3}, so the farther reading is one eighth. The radial field varies as r^{-4}, so its magnitude is one sixteenth. A monopole contamination would decay more slowly and spoil these ratios.

---

3. In the disc model let I=0.20 A and use the numerical values above. Calculate torque and power.

▼ **Work through the solution**

τ=12 µN m. Mechanical power is 1.2 mW, equal to I times 6.0 mV. Contact and resistive losses must be included before comparing with a real load.

Waves, radiation and relativity

Follow a disturbance from its source to a measurement. Work through polarisation, dispersion, scattering and changes of observer using measurable quantities.

1. ## From a local field balance to travelling waves and standing modes

   A voltage probe moving along a line finds regularly spaced zeros. A second probe at the same point measures a magnetic field even when the electric reading vanishes. Derive both readings from Maxwell's equations, calculate the energy stored between the zeros, and decide what additional measurement would reveal the direction in which energy travels.

2. ## What an analyser measures about polarisation and coherence

   Two red beams produce the same reading through every rotating linear polarizer. One might be circularly polarized and the other an incoherent mixture. Build the measurements that distinguish them, derive the visibility of their interference fringes, and decide which information is lost on a slow detector.

3. ## Oblique interfaces and films whose reflections cancel

   A film only one quarter of a wavelength thick can suppress reflection from a transparent surface. Follow the boundary conditions all the way from an oblique incident wave to the input admittance of a stack, then calculate how much reflection returns when the wavelength, angle or layer order changes.

4. ## Causal material response and the propagation of a finite pulse

   An optical pulse entering a transparent slab is delayed, broadened and sometimes reshaped. Derive these effects from the motion of bound charge, and learn what a measured group delay actually says about a pulse carrying information.

5. ## A guided mode, a cavity line and the energy they carry

   A microwave pulse sent through a rectangular guide arrives late even though the guide is evacuated. Terminating the guide converts the travelling pattern into a family of resonances. Use the same transverse eigenproblem to predict its cutoff, pulse delay, cavity frequency and measured linewidth.

6. ## Constructing the retarded Green function from an outgoing spherical shell

   A short source pulse reaches two receivers at different times and amplitudes. Derive the signal from a localized spacetime impulse, see why the potential is evaluated at a past time, and identify which observables change when the source occupies an extended region.

7. ## Radiating multipoles and what a detector actually integrates

   A small oscillating source looks point-like from a distant detector, yet its angular pattern reveals its structure. Derive electric and magnetic dipole radiation from a conserved current, identify what

survives when electric dipoles cancel, and turn a detector aperture into a calibrated estimate of total power.

8. Scattering, extinction and the phase of the forward wave

A suspension of transparent particles dims a narrow transmitted beam while absorbing almost none of its energy. Derive the scattered angular power, identify the interference term seen in transmission, and use energy conservation to put a bound on the response of a single small particle.

9. Diffraction from the near field to the focal plane

An illuminated aperture is a boundary condition for a wave. A screen a millimetre away, an intensity camera at a lens focus and a microscope measuring an evanescent field sample different consequences of the same condition. We will derive operators for propagating the field, estimate the approximations each instrument makes and use energy and sampling checks to tell numerical structure from a measurable fringe.

10. What different observers measure

A moving observer carries clocks, rulers and a calibrated test charge. A statement about the electric field must specify which observer held that test charge at rest. In this chapter we derive how different measurements can agree on the force law and yet assign different electric fields, magnetic fields and charge densities to the same physical apparatus. A current-carrying wire and a light pulse will provide independent checks.

11. Relativistic particle dynamics and the field tensor

A circular track measures momentum per charge. A clock following the particle registers proper time, and a detector records how much kinetic energy it has gained. We will put these observations in one dynamical equation, solve that equation in uniform fields and exhibit a numerical pusher that can get the energy right while its orbit phase slowly slips.

12. Radiated power, mechanical work and energy limits

A charge can radiate while its speed is constant. A source can exchange energy with its nearby field before a pulse reaches a receiver. We will use the flux through a distant sphere, the work at the charge and the energy stored in a local field as three operationally distinct measurements. Agreement requires the relevant integration times and the range of validity of the particle model.

13. How can a field that decays in air transmit power to another prism?

Place a thin air gap between identical prisms. Derive the TE transmission from the two interface conditions and test when an exponential attenuation length predicts the actual throughput.

14. Resolve the self-force before taking a point-charge limit

A finite spherical charge distribution introduces a propagation time. Follow its memory kernel to electromagnetic inertia, radiative recoil and the limits of a truncated equation.

15. From light clocks to collision kinematics

Construct spacetime measurements before transforming fields. Use the same invariant interval to compare clocks, add velocities and close a relativistic collision.

Waves, radiation and relativity · 1. From a local field balance to travelling waves and standing modes

A voltage probe moving along a line finds regularly spaced zeros. A second probe at the same point measures a magnetic field even when the electric reading vanishes. Derive both readings from Maxwell's equations, calculate the energy stored between the zeros, and decide what additional measurement would reveal the direction in which energy travels.

Specify a medium, a region and a sign convention

Take a homogeneous, isotropic, transparent medium at rest. In the frequency range of interest its permittivity ϵ and permeability μ are positive, real constants. Neglect free charge and current in the region where we will propagate a field. Sources can still exist outside the region and set its incident wave. The electric and magnetic fields are functions of position and time. Maxwell's equations impose two divergence conditions in addition to the two evolution equations. In particular, $\nabla \cdot \mathbf{E} = 0$ here because ϵ is uniform. At a material interface the normal condition concerns $\mathbf{D}$, so one must reconsider this simplification before using the same wave equation across a boundary.

Consider first a plane field depending only on z and t, with electric polarization along x and magnetic field along y. Using a right-handed coordinate system, the y component of the curl of the electric field is $+\partial_z E_x$. The x component of the curl of the magnetic field is $-\partial_z H_y$. Faraday's law and Ampère–Maxwell's law then give

$$\partial_z E_x = -\mu\, \partial_t H_y,$$
$$\partial_z H_y = -\epsilon\, \partial_t E_x.$$

Imagine a localized positive electric bump travelling to the right. At a point just ahead of its centre the field grows in time. The second equation therefore requires a negative magnetic-field slope there. This checks the relative signs before making a harmonic assumption. With this convention a positive $E_x H_y$ describes energy flow towards increasing z.

Find what travels and what remains constrained

Differentiate the first curl equation with respect to z and substitute the second. Constant ϵ, μ allow derivatives to commute with the coefficients. The electric field obeys $E_{zz} = \epsilon\mu E_{tt}$. The same procedure gives the magnetic wave equation. More generally, taking the curl of Faraday's law produces

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu\epsilon \mathbf{E}_{tt}.$$

The divergence term vanishes under the stated assumptions. A formal solution of three unrelated scalar wave equations can still violate the divergence or original curl conditions. These conditions connect the measurable polarizations and reduce the independent initial data.

$$v = (\epsilon\mu)^{-1/2},$$

$$Z = \sqrt{\mu/\epsilon},$$

$$\partial_t(E_x + ZH_y) + v\partial_z(E_x + ZH_y) = 0.$$

The corresponding combination with a minus sign travels to the left. To see the first equation, expand it and replace $E_t = -H_z/\epsilon$, $H_t = -E_z/\mu$. The coefficients cancel because $vZ = 1/\epsilon$ and $Z/\mu = v$. Hence there exist arbitrary waveform functions F and G such that

$$E_x(z,t) = F(z - vt) + G(z + vt),$$

$$H_y(z,t) = [F(z - vt) - G(z + vt)]/Z.$$

The medium determines two different quantities. Delay measurements determine the product $\epsilon\mu$; measuring electric-to-magnetic amplitude in a single travelling wave determines their ratio. Impedance has units volts per ampere. It relates fields in volts per metre and amperes per metre at one location. A port voltage divided by current additionally depends on how a particular line geometry integrates those fields over a cross-section.

If only the initial electric field $E_x(z,0)$ is known, its decomposition into left- and right-moving components is ambiguous. A simultaneous magnetic measurement supplies the missing combination. For example, suppose a local measurement gives $E_x = 3.0$ V/m and $H_y = 4.0$ mA/m in a medium with $Z = 250\ \Omega$. There $ZH_y = 1.0$ V/m, so the right-moving field equals 2.0 V/m and the left-moving field equals 1.0 V/m. These numbers describe local instantaneous amplitudes. Separating time-averaged directional powers requires squared waveform records or a harmonic representation.

Recover the energy current and average it correctly

Multiply the first curl equation by H, the second by E, and add. The spatial terms form $\partial_z(EH)$, while the time terms form minus the derivative of $\epsilon E^2/2 + \mu H^2/2$. Thus in the absence of work on current the field obeys a local conservation law. The two travelling pieces give an especially useful representation of both sides:

$$u = \frac{\epsilon E^2 + \mu H^2}{2} = \epsilon(F^2 + G^2),$$

$$S_z = EH = \frac{F^2 - G^2}{Z},$$

$$u_t + \partial_z S_z = 0.$$

The energy densities add while the currents subtract. A field carrying equal left- and right-going powers can have substantial stored energy and zero mean transport. A cancellation in E alone says little about stored energy because it often accompanies a maximum in H. For the two local amplitudes of the preceding example the instantaneous net flux is

$$(4 - 1)/250 = 0.012\,\mathrm{W/m^2}.$$

Multiplying the original readings gives $3.0 \times 0.004 = 0.012$ W/m², an independent check.

For harmonic calculations use real physical fields equal to $\mathrm{Re}[\widetilde{\mathbf{E}} e^{-i\omega t}]$, so the phasors denote peak amplitudes. A product of real sinusoids averages to one half the real part of one phasor times the conjugate of the other. Therefore $\langle \mathbf{S} \rangle = \mathrm{Re}(\widetilde{\mathbf{E}} \times \widetilde{\mathbf{H}}^{*})/2$. A single forward sinusoid with electric peak 6.0 V/m in vacuum carries $36/(2Z_0) = 0.0478$ W/m², using $Z_0 = 376.73\ \Omega$. If an instrument already reports RMS electric amplitude, the factor of one half is absorbed in that reading. Mistaking an RMS calibration for a peak calibration changes an inferred power by a factor of two.

Enforce a conductor condition on the two waves

Place a perfect conducting plane at z=0 and illuminate it from negative z. The tangential electric field must vanish at the surface at all times. For peak incident phasor $E_0 e^{ikz}$, with $k = \omega/v$, the reflected field must be $-E_0 e^{-ikz}$; the accompanying magnetic fields add at the surface. Choosing real E_0, the total phasors are

$$\widetilde{E}_x = 2iE_0 \sin(kz),$$

$$\widetilde{H}_y = \frac{2E_0}{Z} \cos(kz).$$

Electric nodes are separated by $\pi/k = \lambda/2$. The magnetic antinodes sit at those same locations. At a fixed interior point the two fields are a quarter cycle out of phase. Their mean product is zero but their instantaneous product alternates sign. Examine a quarter wavelength of the standing wave: energy crossing an internal plane feeds local changes between electric and magnetic storage, while the conductor absorbs no net power in the idealization.

In a nonmagnetic dielectric with $\epsilon_r = 4$, the phase speed is $c/2$ and $Z = Z_0/2$. At 500 MHz the wavelength is 0.2998 m and successive electric zeros lie about 150 mm apart. A 4 V/m incident peak gives an 8 V/m electric antinode and magnetic amplitude $8/(188.4) = 0.0425$ A/m at the conductor. A tiny electric probe positioned at the conductor would read essentially zero. A loop sensing magnetic flux can respond strongly there. The response of a real loop also depends on its orientation, area and the frequency-dependent conversion of flux to voltage.

Build a complete family between two walls

Put conductors at $z = 0, L$. Tangential E now vanishes at both endpoints. Try a separated solution $E = X(z)T(t)$. Dividing the wave equation by XT gives $X''/X = T''/(v^2 T) = -k^2$ for an oscillatory solution. The first condition eliminates the cosine term in X. The second enforces $k_n = n\pi/L$, $n = 1, 2, \ldots$. Frequencies are

$$\omega_n = v k_n.$$

There is no nonzero constant tangential electric mode consistent with two zero boundary values. A static magnetic bias can be independently specified through initial flux; it does not belong to this oscillatory electric-mode series.

$$E(z,t)=\sum_{n=1}^{\infty}\sin(k_n z)[A_n\cos(\omega_n t)+B_n\sin(\omega_n t)].$$

To find coefficients, multiply the initial field by $\sin(k_m z)$ and integrate over the interval. Product-to-sum identities give zero for $n\neq m$, and $L/2$ for $n=m$. Thus $A_m=(2/L)\int_0^L E(z,0)\sin(k_m z)\,dz$. Applying the same projection to the initial time derivative gives

$$B_m=(2/(\omega_m L))\int_0^L E_t(z,0)\sin(k_m z)\,dz.$$

The curl equation supplies E_t from the spatial derivative of the initial H field. These formulas convert measured initial data into mode coordinates. A point reading alone supplies one linear combination of all coordinates.

One can also derive orthogonality without already knowing the sines. Multiply the eigen-equations for X_n and X_m by the opposite functions, subtract, and integrate. The boundary term $[X_n'X_m-X_m'X_n]_0^L$ is zero, so

$$(k_n^2-k_m^2)\int X_n X_m dz=0.$$

Other homogeneous boundary conditions can preserve this cancellation. Spatially varying material parameters generally change the weight in an orthogonality relation. Knowing the inner product matters when estimating an amplitude from measured fields in unlike dielectrics.

Predict which modes a localized preparation actually excites

Imagine an idealized initial tangential electric field with a symmetric top-hat amplitude E_a on $L/3<z<2L/3$, zero elsewhere, and zero initial H. Then $B_n=0$ and integration gives

$$A_n=\frac{2E_a}{n\pi}\left[\cos(n\pi/3)-\cos(2n\pi/3)\right].$$

Symmetry around the middle eliminates even n because their sine eigenfunctions change sign about that point. For $n=1,3,5$ the coefficients are $2E_a/\pi,\ -4E_a/(3\pi),\ 2E_a/(5\pi)$, respectively. Negative A_3 is a relative phase; its contribution to stored energy is positive. Orthogonality gives initial total electric energy per transverse area $\epsilon L\sum A_n^2/4=\epsilon L E_a^2/6$. The first mode carries $6/\pi^2\approx 60.8\%$ of that total, while the third carries $8/(3\pi^2)\approx 27.0\%$. A three-mode truncation reproduces a large share of the energy and still rounds the sharp initial edges. Measured peak error and integrated energy error answer different questions about that approximation.

Take $L=0.60$ m filled with the dielectric above. The first frequency is about 124.9 MHz. A receiver at midspan sees modes 1 and 3 with opposite spatial signs. A second receiver at $L/3$ has a node of mode 3 and will miss it. Detecting only one large peak after changing receiver position is therefore insufficient

evidence that the source was single mode. Predict the source overlap, propagation loss and receiving overlap separately before interpreting a spectrum.

Use an open wave measurement as an independent power check

The linked calculation treats a propagating field and derives an energy-flux observable. Read its amplitude convention and the material properties beside the inputs. Predict which quantity changes quadratically when you vary only field amplitude. If a lab measures a total power, convert the calculated flux using the illuminated area and the angle between the flux vector and the detector normal. For a plane wave in a uniform medium, changing area at fixed field changes total intercepted power without changing any local field. In a narrow cavity beam the local field itself varies across the aperture, so the area product must be replaced by a surface integral.

Compare fields, energy density and transported power in the chapter experiment

An investigation of wave speed on a space–time mesh

A pulse simulation offers a second way to test the continuum mode calculation. Instead of assuming a single-frequency source, initialise a packet in a uniform one-dimensional region and record the times when it crosses two stations. First predict how the numerical method can change this transit measurement. A code that respects the differential equations as its mesh tends to zero can have a noticeably incorrect propagation speed at a finite resolution.

Work with an electric component E_y on integer mesh points and a magnetic component H_z on the intervening half points. Store the magnetic component half a time step later than the electric component. Faraday's law compares the electric field on the two ends of one spatial interval. Ampère's law compares the two magnetic components bounding the electric location. With $d = \Delta x$ and $s = \Delta t$, their integral balances become

$$H_{j+1/2}^{n+1/2} = H_{j+1/2}^{n-1/2} - \frac{s}{\mu d}\left(E_{j+1}^{n} - E_{j}^{n}\right),$$

$$E_{j}^{n+1} = E_{j}^{n} - \frac{s}{\epsilon d}\left(H_{j+1/2}^{n+1/2} - H_{j-1/2}^{n+1/2}\right).$$

For a forward wave the indicated components have the same sign. A change in either sign convention requires changing both updates consistently. Placing the variables on staggered meshes also matters when sampling power. A product of an electric field and the nearest stored magnetic number is not automatically their product at a common event. Interpolate both in time and space for a quantitative flux comparison, or use a discrete energy identity appropriate to the scheme.

Eliminate the magnetic values by taking a second time difference of the electric update. The resulting equation equates the centred second time difference to C^2 times the centred second space difference, where $C = us/d$ and $u = 1/\sqrt{\mu\epsilon}$. For the discrete mode $E_j^n = E_0 e^{i(kjd - \omega n s)}$, the difference of the forward and backward phases becomes a sine. Take the branch connected to the continuum wave at small positive k.

$$\sin(\omega s/2) = C\sin(kd/2),$$

$$\frac{v_{\mathrm{phase}}}{u} = \frac{\arcsin[C\sin(kd/2)]}{Ckd/2},$$

$$\frac{v_{\mathrm{group}}}{u} = \frac{\cos(kd/2)}{\sqrt{1 - C^2\sin^2(kd/2)}}.$$

The group speed follows by differentiating the frequency with respect to wavenumber before choosing a particular packet. The right-hand side of the sine equation must remain within the range of real sines for every mesh mode. This gives $C \le 1$ in this one-dimensional uniform problem. If a short-wavelength mode violates the condition, a growing discrete solution accompanies the decaying one. A seemingly smooth start may seed very little of either, but roundoff or a sharp boundary can excite the growing solution. A stable choice still permits frequency-dependent speed.

For ten samples per wavelength and $C = 0.50$, the phase and group speed ratios are approximately 0.988 and 0.963. Over twenty wavelengths a phase crest is about a quarter of a cycle late, and the envelope peak has an even larger relative lag. With twenty samples per wavelength the errors shrink roughly fourfold. Reducing only the time step at fixed space resolution need not improve the result: the spatial difference retains its own dispersion error, and the temporal error can partially compensate it. At $C = 1$ in this idealised one-dimensional uniform case the phase relation on the resolved branch is exact. A slanted multidimensional wave, a discontinuous medium or a different mesh does not inherit that special cancellation.

A worked time-of-flight measurement with uncertainty

To construct an initial forward pulse, give the electric field a broad Gaussian envelope multiplying its carrier and assign the associated magnetic field $H = E/Z$ at its proper staggered locations and times. Setting the magnetic field to zero instead launches a superposition of left- and right-moving packets of comparable amplitude. A reflecting wall can return one packet through the sensor after the other. Before fitting a speed, identify a time window that excludes every return from the domain boundaries. Check this by increasing the distance to the nearest wall without changing either station.

Suppose the expected continuum transit between stations is 100.0 ns, each threshold crossing is timed to 0.3 ns and the station spacing is known to 0.2 percent. Independent timing errors contribute $\sqrt{2}(0.3)/100 = 0.00424$ relative uncertainty to the difference. Combined in quadrature with the length uncertainty this gives 0.47 percent. The roughly 3.7 percent group-speed bias at ten points per wavelength should be distinguishable, while the smaller phase error calls for careful identification of the same crest at both stations. A common clock offset cancels from the difference, whereas a common timebase scale error does not.

A threshold on the leading edge tracks a changing fraction of peak amplitude if the model disperses the pulse. Use a complex narrow-band signal and fit the slope of the unwrapped phase difference against frequency as a second estimate of group delay. For a single frequency only the phase modulo 2π is observed. An apparent speed that jumps as the station separation changes by a small fraction of a

wavelength can be an unwrapping error, even though the electric time series looks smooth. Begin with stations close enough to choose a branch, then follow the branch as their separation increases.

An additional stringent check uses two packets with different carrier wavelengths but equal initial envelope width in space. If the narrow-band condition holds for both, their group delays should match the two predictions of the mesh dispersion relation. Measured delays that approach the same continuum value with refinement while their difference obeys the predicted wavelength dependence support the diagnosis of numerical dispersion. A nonvanishing delay difference in the refined limit calls for revisiting the assumed physical material law, the timing observable or the boundaries. A single apparent agreement at one wavelength cannot distinguish these possibilities.

Separate flux from standing-wave energy

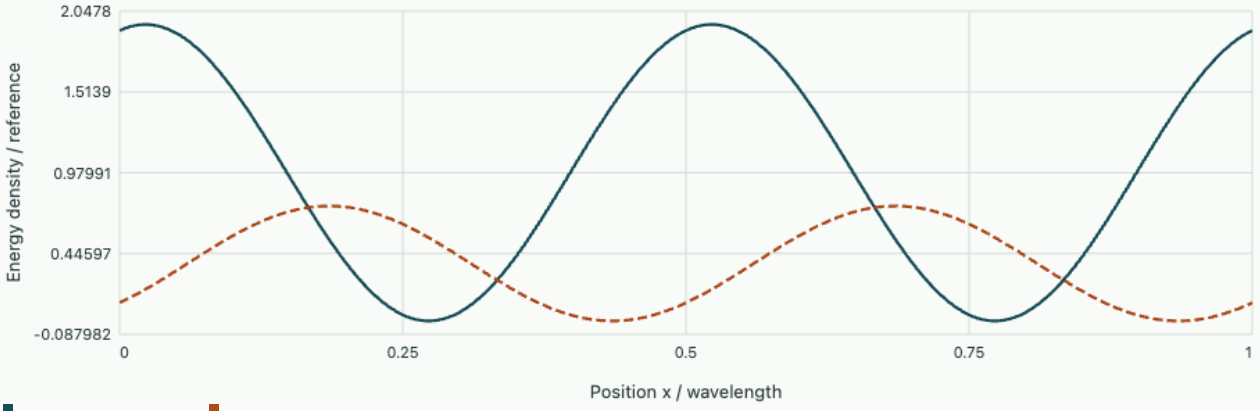
Change the balance of waves travelling in opposite directions. Identify a point where the instantaneous field is large while net power transfer is small.

Open this experiment online

Try the calculation

Reflected electric amplitude / incident **0.6**

Phase of incident wave (degrees) **30**



The cycle-averaged net power is 64.0% of the incident power. Reversing the sign of reflection swaps the pattern at the boundary.

Worked inverse problem · estimate the energy of an invisible mode

A conducting one-dimensional cavity contains its first and third electric standing modes. At a common electric maximum in time let $E(x) = a\sin(\pi x/L) + b\sin(3\pi x/L)$. A small signed probe at $L/2$ reads 2.0 V/m, and a probe at $L/6$ reads 4.0 V/m. Determine the two amplitudes and the fraction of energy in the third mode. Compare this result with what could be concluded from two intensity probes alone.

▼ **Solve the signed and unsigned versions**

The centre gives $a - b = 2.0$. The second station gives $a/2 + b = 4.0$. Summing the equations yields $a = 4.0$ V/m and hence $b = 2.0$ V/m. At this instant the compatible free modes have zero magnetic amplitude. Spatial orthogonality makes integrated electric energy proportional to $a^2 + b^2$, even though there are local interference terms. The third mode therefore carries 20 percent of total energy. This fraction stays fixed during free evolution in an ideal lossless cavity. The instantaneous electric share of each mode alone oscillates in time.

Two intensity probes would give only magnitudes of the two readings. Choose the centre reading positive by a global sign convention. The second could then be -4.0 instead. Solving $a - b = 2.0$, $a/2 + b = -4.0$ gives $a = -4/3$ and $b = -10/3$ V/m, for a third-mode fraction $100/116 = 0.862$. Both field preparations produce the same two intensities. A third carefully chosen position, a phase-sensitive receiver or a frequency-resolved trace distinguishes them. A probe at a location that simply duplicates either existing linear combination adds precision but cannot supply the missing sign.

An eigenmode decomposition is therefore both a prediction tool and an experimental design tool. Before choosing sensors, form the matrix of sampled mode shapes and check whether its columns remain distinguishable within noise. If two stations approach a shared node, the inferred coefficients can grow unstable while the predicted signal at those stations remains an excellent fit. A held-out measurement away from the common node exposes the ambiguity.

Why a sharp preparation resists uniform modal convergence

A rectangular initial displacement has a jump. Any finite sum of the smooth sine modes is continuous and therefore cannot converge uniformly to a function with different one-sided limits. Near the edge a partial sum oscillates. Increasing the cutoff makes the oscillatory region narrower, but its largest overshoot does not tend to zero relative to the jump. Pointwise comparisons a fixed positive distance from the jump and spatially integrated energy tell a different story from the maximum error.

This is physically relevant to preparing an electromagnetic cavity. No source creates a perfectly discontinuous field with unbounded bandwidth. Replace the top hat by its convolution with a smooth narrow preparation profile. The modal coefficients then acquire the Fourier transform of that profile as a high-frequency envelope. Doubling the number of retained modes beyond the preparation bandwidth should change both the integrated energy and the measured field little. A finite-width detector performs its own additional averaging, even when the preparation is sharp.

An informative validation uses two limits in a stated order. Fix the physical source and detector widths while increasing the number of modes until the detector prediction converges. Then vary the physical widths to investigate the sharp-edge limit. Shrinking the model widths each time the numerical resolution increases changes the problem being solved and may keep the apparent peak error from decreasing.

Tutorial · reconstruct a partial reflection from three observations

A coherent line scan in a nonmagnetic medium of known $Z = 200\ \Omega$ finds successive electric maxima 80 mm apart at 1.0 GHz. Maximum and minimum peak electric amplitudes are 5.0 and 1.0 V/m. Determine phase speed, incident and reflected peak fields and mean net flux. Explain exactly what fixes the phase of the reflection coefficient and what remains ambiguous in these observations.

▼ **Derive the two directional amplitudes**

For nonnegative amplitude magnitudes a, b, the maxima and minima of the sum of two opposite travelling waves are $a + b$ and $|a - b|$. If the source is known to lie to the left and net transport is rightward, $a > b$, so $a = 3.0,\ b = 2.0$ V/m. Wavelength is twice the 80 mm maximum spacing and $v = f\lambda = 1.60 \times 10^{8}$ m/s. The reflection magnitude is $2/3$, and the average net flux is

$$(9 - 4)/(2 \cdot 200) = 0.0125\,\mathrm{W/m^2}.$$

A measurement of node spacing constrains the relative phase change over distance. Determining reflection phase at a named load additionally requires the position of a maximum or minimum relative to that reference plane. Without an independent source direction, swapping a and b produces the same electric envelope but reverses the net flux.

Tutorial · test whether a receiver can reconstruct two stored modes

At an instant when both cosine time factors are one, only the first and third electric modes contribute. At midspan the field is 3.0 V/m; at one third span it is 1.732 V/m. Determine A_1, A_3 and the stored electric energy per unit area for $L = 0.60$ m,

$$\epsilon = 4\epsilon_0.$$

Would a second sensor at two thirds span provide independent information if it replaced the midspan sensor?

▼ **Solve the observation system and identify its rank**

At one third span $\sin(\pi/3) = \sqrt{3}/2$ and $\sin\pi = 0$, so $A_1 \simeq 2.00$ V/m. At midspan the combination is $A_1 - A_3 = 3.0$, hence $A_3 = -1.0$ V/m. The energy per area is $\epsilon L(4 + 1)/4 = 2.66 \times 10^{-11}$ J/m². At two thirds both sinusoids have the same values as at one third for these odd modes: the first is $\sqrt{3}/2$ and the third is zero. Those two rows would be identical. More digits in either measurement cannot resolve A_3 from that pair. An offset from a node helps, but near a node a small measured voltage also amplifies calibration noise in the reconstruction.

David Staelin, MIT 6.013 · electromagnetic waves, power and Maxwell boundary conditions

Waves, radiation and relativity · 2. What an analyser measures about polarisation and coherence

Two red beams produce the same reading through every rotating linear polarizer. One might be circularly polarized and the other an incoherent mixture. Build the measurements that distinguish them, derive the visibility of their interference fringes, and decide which information is lost on a slow detector.

Begin with the measured quantity and the averaging time

Fix a beam travelling in positive z and choose perpendicular laboratory axes x and y in the detection plane. Represent a narrowband field by $\mathrm{Re}[\mathbf{e}(t)e^{-i\omega_0 t}]$. The complex envelope $\mathbf{e} = (e_x, e_y)$ varies slowly compared with an optical cycle. A photodiode collecting many cycles measures a time average of energy flux, possibly over many independent changes of the envelope. In a transparent isotropic medium a convenient intensity convention is $I = \langle |e_x|^2 + |e_y|^2 \rangle / (2Z)$. All comparisons in this unit assume equal transmissive losses and detector responsivity unless a calibration factor is stated.

A deterministic monochromatic field has a Jones vector. Its overall complex phase disappears from all single-beam intensities, while a relative phase between components can change the polarization ellipse. For example, use the normalized vectors $(1,0),\ (1,1)/\sqrt{2},\ (1,i)/\sqrt{2}$. The first is x-linear, the second linear at +45°, and the last circular. At a fixed point choose the time origin so the physical electric vector for the last case begins on the positive x axis and then moves towards positive y as t increases. This explicit convention avoids ambiguity in the words left and right circular, which depend on where an observer stands.

A fast field sampler would reconstruct both real components directly. Optical intensity measurements normally come with a response time much longer than a cycle. The data naturally contain squares and correlations of fields. A random global phase changing between shots destroys the inferred absolute phase but need not depolarize a single beam. If $e_y = ie_x$ in every shot, their relative phase stays fixed however rapidly the common phase wanders. A detector can measure that relationship with suitable polarizing optics.

Derive a quadratic form for any analyser setting

An ideal analyser transmitting normalized Jones vector $\mathbf{a}$ first projects the incident field to complex amplitude $\mathbf{a}^\dagger \mathbf{e}$. Squaring and averaging gives a quadratic form in $\mathbf{a}$. Define the coherency matrix in intensity units by $J_{ij} = \langle e_i e_j^* \rangle / (2Z)$. It is Hermitian, its diagonal entries are nonnegative, and its trace is total intensity. For any vector b,

$$\mathbf{b}^\dagger J \mathbf{b} = \langle |\mathbf{b}^\dagger \mathbf{e}|^2 \rangle / (2Z) \geq 0.$$

This positivity will later provide an immediate check on measured data.

$$I_{\mathbf{a}} = \mathbf{a}^{\dagger} J \mathbf{a},$$
$$S_0 = J_{xx} + J_{yy},$$
$$S_1 = J_{xx} - J_{yy},$$
$$S_2 = 2\Re J_{xy},$$
$$S_3 = -2\Im J_{xy}.$$

With this sign convention the circular vector $(1, i)/\sqrt{2}$ has positive S_3. Invert the four definitions to reconstruct $J = \frac{1}{2}\begin{pmatrix} S_0+S_1 & S_2-iS_3 \\ S_2+iS_3 & S_0-S_1 \end{pmatrix}$. A linear polarizer with real $\mathbf{a} = (\cos\alpha, \sin\alpha)$ gives

$$I(\alpha) = \tfrac{1}{2}[S_0 + S_1\cos(2\alpha) + S_2\sin(2\alpha)].$$

This follows by expanding the two squared direction cosines and the cross term. The imaginary part of the correlation drops out for every real analyser. Scanning a linear analyser with arbitrarily fine angular steps cannot supply that missing observable. Measuring the two circular analyser states supplies $I_+ - I_- = S_3$; their sum, like any pair of orthogonal ideal outputs, equals S_0. In a dual-beam measurement it is useful to compare that sum with $I_x + I_y$ as a consistency check on losses.

Build a circular analyser from a retarder

An ideal lossless plate with fast and slow axes x and y introduces a relative phase δ. Factoring out an unobservable common phase, its Jones matrix in those axes is $D = \mathrm{diag}(1, e^{i\delta})$ for the chosen sign of retardance. If the axes are rotated by an angle α, express the incoming vector in the rotated basis, apply D, then transform back. With $R(\alpha)$ the ordinary real two-dimensional rotation, the laboratory matrix is $R(\alpha)DR(-\alpha)$. Reversing the sign of δ interchanges retardation conventions, so calibrate a handedness label using a known input.

For a quarter-wave plate with $\delta = \pi/2$ followed by a +45° linear polarizer, the transmitted amplitude is $(e_x + ie_y)/\sqrt{2}$. The analyzer equivalent to this sequence is the vector $(1, -i)/\sqrt{2}$. Hence it reads I_- in the preceding convention. Interchange the downstream polarizer between +45° and −45° to realize the two orthogonal circular measurements. For a half-wave plate, a linear input at angle ψ relative to its fast axis emerges linearly polarized at angle $-\psi$ relative to that axis, modulo π. Rotating a half-wave plate by 1° rotates the linear output by 2° when the incident direction is held fixed.

Suppose a plate designed at 600 nm has physical birefringence $\Delta n = 0.010$. If the relative phase magnitude is $2\pi\Delta n\, d/\lambda_0$, a first-order quarter wave needs $d = \lambda_0/(4\Delta n) = 15$ µm. At 650 nm, holding Δn fixed, the same plate delivers $0.923(\pi/2)$, about 83.1°. With a +45° linear input, equal component amplitudes are preserved but a residual $S_2/S_0 = \cos\delta \approx 0.121$ remains. A linear analyser scan that varies slightly with angle can therefore reveal retardance error even when the output looks nearly circular.

Reconstruct the degree and axes of polarization

Hermitian J has real nonnegative eigenvalues. Calculating the trace and determinant yields eigenvalues $\lambda_\pm = (S_0 \pm s)/2$, where $s = (S_1^2 + S_2^2 + S_3^2)^{1/2}$. The degree of polarization is $P = s/S_0$ for positive intensity. Positivity requires $P \le 1$. A single deterministic Jones vector gives a rank-one outer product and P=1. A completely unpolarized field gives $J = (S_0/2)I$ and P=0. An isotropic unpolarized contribution adds equally to both eigenvalues. The excess of the larger eigenvalue over the smaller gives the uniquely polarized contribution in this decomposition.

$$2\psi = \operatorname{atan2}(S_2, S_1),$$

$$\sin(2\chi) = S_3/s.$$

Here ψ is the ellipse orientation modulo π; χ is an ellipticity angle between $-\pi/4$ and $+\pi/4$. The ratio of minor to major field axes is $|\tan\chi|$. At circular polarization the orientation is undefined, and near this limit even small noise on S_1, S_2 can rotate a fitted orientation substantially. At P=0 neither angle describes a measurable polarized component.

Consider readings of 7 and 3 mW for horizontal/vertical, 6 and 4 mW for ±45°, and 8 and 2 mW for the positive/negative circular analysers. Every pair sums to 10 mW. The Stokes vector is $(10, 4, 2, 6)$ mW, $s = \sqrt{56} = 7.48$ mW, $P = 0.748$. The polarized-component angle is 13.28°, and $\chi = \frac{1}{2}\arcsin(6/\sqrt{56}) = 26.6^\circ$. A linear analyser alone would have suggested a modulation depth $\sqrt{20}/10 = 0.447$. The recovered eigenvalues, about 8.74 and 1.26 mW, are the largest and smallest transmitted powers over all ideal elliptical analysers. The largest intensity available to a linear analyser is only $(10 + \sqrt{20})/2 = 7.24$ mW.

Derive the two-arm visibility and its ambiguities

Now split a quasimonochromatic scalar field between two interferometer arms with delay τ and relative phase φ. At one output combine complex fields $e_1(t)$ and $e_2(t-\tau)e^{i\varphi}$, absorbing beam-splitter factors into arm intensities I_1, I_2. Expanding the squared sum gives

$$I(\varphi) = I_1 + I_2 + 2\sqrt{I_1 I_2}\,\Re[\gamma_{12}(\tau)e^{i\varphi}],$$

$$\mathcal{V} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2}|\gamma_{12}(\tau)|.$$

Define $\gamma_{12} = \langle e_1^*(t)e_2(t-\tau)\rangle / \sqrt{\langle|e_1|^2\rangle\langle|e_2|^2\rangle}$, provided both intensities are nonzero. Cauchy–Schwarz enforces $|\gamma| \le 1$. For vector beams with constant normalized polarizations $\mathbf{a}_1, \mathbf{a}_2$, an additional overlap $\mathbf{a}_1^\dagger \mathbf{a}_2$ appears. Orthogonal polarizations can erase fringes even when relative phase is perfectly stable. A common output analyser projects both beams into the same polarization and can restore modulation at a reduced transmitted flux.

An interferometer gives $I_1 = 4.0$ mW, $I_2 = 1.0$ mW and observed visibility 0.60 with matched polarization. The intensity imbalance alone limits visibility to $2\sqrt{4}/5 = 0.80$. Thus $|\gamma| = 0.75$. If the polarizations are instead linear at 30° to each other, the polarization overlap is 0.866 and the inferred scalar degree of coherence is

$$0.60/(0.80 \cdot 0.866) = 0.866.$$

Reporting visibility as a direct measure of source coherence requires both arm powers and polarization overlap. Detector backgrounds additionally increase the denominator without increasing fringe modulation.

Translate a spectrum into a coherence envelope

For a stationary scalar field, different frequency components of a long acquisition add in average intensity. A relative delay changes the phase of each by its own frequency times τ. Consequently the mutual coherence of two copies of the same field is proportional to the Fourier transform of the optical power spectrum. If the normalized spectrum is Gaussian in ordinary frequency, with centre ν_0 and standard deviation σ_ν, completing the square in the Gaussian Fourier integral gives

$$|\gamma(\tau)| = \exp(-2\pi^2 \sigma_\nu^2 \tau^2),$$

$$\tau_{1/e} = \frac{1}{\sqrt{2}\pi\sigma_\nu}.$$

The rapid carrier phase varies at ν_0, while finite bandwidth sets the envelope. For $\sigma_\nu = 1.0$ THz the amplitude coherence drops to $1/e$ at 0.225 ps. This is 67.5 µm of vacuum optical path difference. In a Michelson geometry the corresponding mirror travel is half that, since light traverses the additional gap twice. An instrument quoting a coherence length should specify which envelope level, which bandwidth convention and which path geometry it uses. Replacing a standard deviation by a full width at half maximum without conversion changes the answer substantially.

Predict three analyser readings before changing the source

Use the polarimetry calculation with a polarized fraction below one. Set the ellipse orientation to 0°, choose positive ellipticity, and compute the horizontal, diagonal and circular contrasts using the definitions above. A change in polarized fraction scales the three contrast components together. Changing orientation alone rotates the linear contrasts through twice the physical angle while leaving the circular contrast fixed. A separate crystal model follows energy flow when two eigenpolarizations propagate in a birefringent plate. For comparison, separate the output phase retardance of a plate from any spatial displacement that could reduce the overlap of two exiting beams on an interferometer detector.

Change ellipticity and reconstruct an unknown polarization in the Stokes experiment

Compare eigenpolarizations and displacement in a crystal

An investigation with a photon-counting polarimeter

A polarimeter reports a count for each analyser, not an exact Stokes vector. A useful second reconstruction starts with a statistical model for those counts. Let all six exposures have the same

duration and gain after calibration. For a pair of complementary analyser settings labelled $+$ and $-$, the mean counts at one exposure time are $N(1+s)/2$ and $N(1-s)/2$, with $s = S_i/S_0$. Here N is the expected pair total after background removal. Independent Poisson counting models shot noise from a steady source. If the source changes between exposures this independence assumption no longer describes the common intensity drift.

Condition on the observed sum $n = n_+ + n_-$. The number in the positive channel then has a binomial distribution with success probability $(1+s)/2$. The natural estimator and its conditional variance follow without assuming equal noise in bright and dark channels:

$$\hat{s} = \frac{n_+ - n_-}{n_+ + n_-},$$

$$\operatorname{var}(\hat{s} \mid n) = \frac{1 - s^2}{n}.$$

The uncertainty becomes small near a pure eigenpolarisation in the matching analyser, while a transverse component of the same beam has $s = 0$ and variance about $1/n$. Exact intervals are preferable near the boundary, where a symmetric error bar would extend into unphysical values. Conditioning is appropriate for a single pair. If one uses the same intensity measurement to normalise several pairs, their reported components become correlated.

Consider the six counts $(800, 200), (620, 380), (550, 450)$ in complementary linear-horizontal, diagonal and circular bases. Their normalised components are 0.60, 0.24 and 0.10. Raw degree of polarisation is

$$\sqrt{0.6^2 + 0.24^2 + 0.10^2} = 0.654.$$

Individual shot-noise estimates are 0.025, 0.031 and 0.031. To first order, uncertainty in the degree is obtained by projecting the component variance along the measured unit vector. For independent pairs the result is approximately 0.026. Fitting an ellipse to this partially polarised beam recovers the polarised component's ellipse; the unpolarised intensity contributes to every positive and negative analyser reading.

For an unpolarised source, taking a square root of three squared noisy estimates always returns a nonnegative number. In fact the expected squared reported degree is approximately $3/n$ when all pair totals equal n. At $n = 1000$ the root-mean-square value is 0.055, even with a perfect instrument. A positive fitted degree around 0.04 in a single exposure therefore cannot establish weak polarisation. Repeat the acquisition at four times the counts. A statistical excess falls in scale by about two, whereas stable true components persist with their signs. A forced projection onto the physical Stokes sphere cannot remove calibration drift or substitute for this uncertainty calculation.

Which-path information encoded in a polarisation state

A balanced interferometer supplies a different quantitative probe. Suppose its path lengths are equal, the beam in arm one has unit horizontal polarisation $\mathbf{e}_1$, and a rotator in arm two produces $\mathbf{e}_2 = \cos\alpha\,\mathbf{e}_1 + \sin\alpha\,\mathbf{e}_y$. The detector receives equal-amplitude contributions with relative phase δ. Squaring the sum in the two output ports of an ideal lossless beamsplitter gives complementary fringes.

$$I_{\pm} = \frac{I_{\mathrm{tot}}}{2}(1 \pm \cos\alpha \cos\delta).$$

For $\alpha = 60^{\circ}$, sweeping the phase from constructive to destructive moves one port between $0.75 I_{\mathrm{tot}}$ and $0.25 I_{\mathrm{tot}}$; the visibility is 0.50. Their sum should remain constant throughout the sweep. If both measured ports fade together, look for path loss, a changing source or detector gain. Polarisation distinguishability moves fringe energy between the two outputs in this ideal result. It does not remove energy from both.

At $\alpha = 90^{\circ}$, complete temporal coherence does not suffice for an unpolarised intensity detector to see fringes. The two field vectors are orthogonal and their scalar product vanishes. Insert an analyser at 45° in front of one detector. It projects both contributions onto the same unit vector with equal positive amplitude $1/\sqrt{2}$, so the maximum visibility returns while the phase-averaged brightness of that port halves. An analyser at -45° gives a fringe displaced by half a period. Summing those two orthogonal analyser intensities restores the original featureless signal. The recovered fringes do not require a change in either source clock or path length.

For unequal arm intensities and a time delay, the interference coefficient additionally contains $2\sqrt{I_1 I_2}/(I_1 + I_2)$ and the magnitude of the normalised temporal coherence evaluated at that delay. This factorisation assumes the temporal waveform and polarisation can be separated in each arm. A broadband retarder with appreciable frequency-dependent axes couples the two. Then compute the spectral integral of the polarisation-resolved field products, preserving the complex phase at each frequency before integration. Multiplying three averaged magnitudes can give the wrong result even if every factor lies between zero and one.

A practical discrimination uses three sweeps at fixed geometry. First measure both single-arm intensities by blocking one arm at a time. Second sweep delay near zero with the analyser removed, and third repeat with the two orthogonal diagonal analysers. If the two diagonal traces regain equal but opposite fringes and sum to the original low-visibility result, distinguishability is a plausible explanation. If none shows variation after path intensities have been balanced, check whether the coherence envelope is narrower than the available delay step. Finite detector bandwidth during a fast phase sweep also suppresses every trace. Repeating the sweep more slowly and reversing its direction distinguishes that response delay from a static polarisation transformation.

Determine what a rotating analyser can detect

Compare projections of a polarised field. Identify which polarisation information is missing from a scan with only a linear analyser.

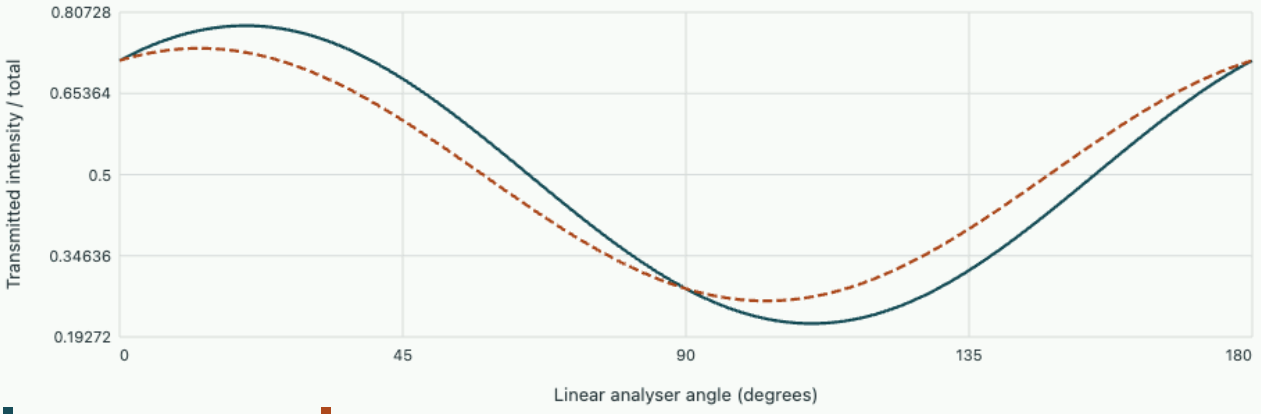
Open this experiment online

Try the calculation

Polarised fraction P **0.6**

Azimuth ψ (degrees) **20**

Ellipticity χ (degrees) **10**



Linear analyser alone Quarter-wave plate before analyser

Measured complementary pairs (horizontal, diagonal, circular): 0.716 / 0.284 · 0.6812 / 0.3188 · 0.6026 / 0.3974. Reconstructed $(S_1,S_2,S_3)/S_0$ = (0.4319, 0.3624, 0.2052); degree 0.6.

Worked inverse problem · distinguish a retarder error from source ellipticity

A linear input is sent through a retarder whose fast axis defines the horizontal direction. The input makes 45° to that axis. A diagonal analyser records a normalised Stokes component $s_2 = 0.100$; a circular analyser gives $|s_3| = 0.990$. Determine possible retardance values, then design a measurement separating imperfect retardance from a slightly elliptical input. Assume no diattenuation for the first inference.

▼ **Track the phase and design an independent preparation**

The equal input components acquire a relative phase δ while keeping equal intensities. The outgoing Stokes components are $s_1 = 0$, $s_2 = \cos\delta$, with $|s_3| = |\sin\delta|$. The first measurement gives $\delta = 84.26^\circ$ or 275.74° modulo one full turn. Both predict $|s_3| = 0.995$, close to the reported 0.990. A signed, calibrated circular analyser selects the sense of the quarter-wave action, provided the basis and time convention are stated. Measuring two intensities more precisely cannot tell a one-quarter from a five-quarter retardance at a single frequency.

If the input already possessed relative phase δ_i, the observation measures $\delta + \delta_i$. Bypass the retarder and reconstruct the input. Then rotate the retarder by 90° to exchange its two eigenaxes while holding the input and detector fixed. The retarder-induced phase changes sign; the source phase does not. Two signed phase measurements can recover the sum and difference, modulo ambiguities eliminated by a nearby wavelength or a known approximate plate order.

Record the calibrated horizontal and vertical transmissions as well. Differential loss makes the two output amplitudes unequal and reduces the accessible circular Stokes component even at perfect quarter-wave phase. A measured degree less than one can result from averaging several slightly different retardances across a finite spectrum. Changing the spectral filter tests that hypothesis. These checks illustrate how independent preparations constrain a device and source separately, where one very precise polarisation measurement of their combination cannot.

A rotating source exposes the detector's averaging time

Imagine an otherwise perfectly linearly polarised beam whose axis rotates uniformly during an exposure. At every instant its degree of polarisation is one. A horizontal analyser transmits an intensity proportional to the squared cosine of the axis angle, and a diagonal one samples the same function displaced by forty-five degrees. Over a complete rotation their average intensities are equal. A circular analyser also receives half the total power. Time-integrated tomography would therefore assign zero net polarisation to the exposure.

For a shorter exposure covering an angular sweep Δ about a central angle, average the complex combination $s_1 + i s_2 = e^{2i\theta}$ uniformly over the sweep. Its magnitude becomes $|\sin\Delta/\Delta|$. A ninety-degree sweep retains a degree of approximately 0.637, whereas a full half-turn yields zero because an unoriented linear polarisation axis repeats every half-turn. The reported degree can thus identify neither instantaneous purity nor temporal stability without an exposure model.

Repeat the measurement with shorter synchronised gates and accumulate many exposures starting at the same rotation phase. Stable rotation should recover the predicted phase-resolved Stokes trajectory. A genuinely random uncorrelated orientation cannot be recovered by simply synchronising to an unrelated clock. If the analyser states are acquired sequentially, synchronise their gates too; otherwise different components may sample different portions of the trajectory and combine into a physically impossible vector.

Tutorial · recover a missing analyser setting and audit positivity

A pair of equal-throughput linear analysers reports $I_x = 4,\ I_y = 2$ and $I_{45} = 4.5$ in the same power units. Find the predicted −45° reading and bound a possible circular contrast. A circular analyser then reports $I_+ = 6.0$. Can these measurements describe a single stationary ensemble?

▼ Reconstruct only what the measured subspace determines

The orthogonal sum fixes $S_0 = 6$, hence $I_{-45} = 1.5$, $S_1 = 2$, and $S_2 = 3$. Remaining positivity requires $S_3^2 \le 6^2 - 2^2 - 3^2 = 23$, so $|S_3| \le 4.796$. A positive circular reading of 6 would imply $I_- = 0$ and $S_3 = 6$. That would make the reconstructed coherency matrix have a negative smaller eigenvalue. Merely normalizing all six readings by their common sum would retain the inconsistency. Check gain, background, plate retardance and whether the source changed during sequential measurements. If data have known

statistical error, fitting J subject to positive semidefiniteness can enforce physical consistency while exposing the residuals.

Tutorial · distinguish delay decorrelation from an unresolved doublet

Two equal narrow spectral lines are separated by 80 GHz, the arms have equal powers, and polarizations match. Obtain the fringe-visibility envelope. Find the first null and first revival in optical path difference. A single visibility reading at the null is compatible with several physical explanations. Propose a scan that tests the doublet hypothesis.

▼ **Add the complex correlations before taking the magnitude**

Each narrow line contributes a phase at its optical frequency. Averaging the two phasors gives $e^{i2\pi\nu_0\tau}\cos(\pi\Delta\nu\,\tau)$, so visibility is $|\cos(\pi\Delta\nu\,\tau)|$. The first null occurs at $1/(2\Delta\nu) = 6.25$ ps, or 1.87 mm path difference. The first revival is at 12.5 ps or 3.75 mm. In a Michelson instrument those mirror travels are 0.937 and 1.87 mm. For real spectral lines of finite width the revival sits under their own coherence envelope. Scan beyond the first null, independently match arm powers and verify polarization. A clean revival distinguishes a narrow doublet from a single broad Gaussian spectrum reproducing the same initial contrast loss.

MIT 2.71 Optics · polarisation, interference and spatial coherence

Waves, radiation and relativity · 3. Oblique interfaces and films whose reflections cancel

A film only one quarter of a wavelength thick can suppress reflection from a transparent surface. Follow the boundary conditions all the way from an oblique incident wave to the input admittance of a stack, then calculate how much reflection returns when the wavelength, angle or layer order changes.

Tangential phase is an observable condition at an extended interface

Let the plane z=0 separate two homogeneous, lossless, isotropic dielectrics. Incident light lies in the xz plane and travels towards positive z. Use phasors $\exp[i(k_x x + k_z z) - i\omega t]$; choose square roots for transmitted waves so energy flows away from the interface, or evanescent fields decay into the second medium. Linear time-independent boundary conditions hold at every point x and time t. If their Fourier factors had unequal k_x or ω, cancellation of the tangential boundary fields at every x,t would be impossible. Thus incident, reflected and transmitted waves share both quantities. The reflected z component reverses sign. The transmitted normal component follows from its own medium's dispersion relation.

$$k_x = k_0 n_1 \sin\theta_1 = k_0 n_2 \sin\theta_2,$$

$$k_{zj} = \sqrt{k_0^2 n_j^2 - k_x^2},$$

$$k_0 = \omega/c.$$

The angles are measured from the normal. Equality of tangential phase gives Snell's law independently of a ray argument, and remains useful when the normal wavenumber becomes imaginary and a real ray angle no longer exists. For example, vacuum light at 600 nm falling at 45° onto a dielectric of index 1.50 has $k_x = 7.40 \times 10^6$ m⁻¹ and $\theta_2 = 28.1^\circ$. Relative phase across a pair of boundary probes at fixed separation along x must be the same whether they detect the incident or transmitted wave. Applying a different frequency-dependent normal propagation constant in each material does not violate that observation.

Reduce each polarization to one tangential electric amplitude

Define s polarization with E along y and p polarization with H along y. For s let $e = E_y,\ h = -H_x$; for p let $e = E_x,\ h = H_y$. Both choices make normal average flux $\mathrm{Re}(eh^*)/2$. Maxwell's curl relation $\mathbf{H} = (\mathbf{k}\times\mathbf{E})/(\mu\omega)$, together with $\mathbf{k}\cdot\mathbf{E} = 0$, gives for a forward wave $h = Ye$ and for a reverse wave $h = -Ye$, where

$$Y_s = \frac{k_z}{\mu\omega} = \frac{\cos\theta}{Z},$$

$$Y_p = \frac{\epsilon\omega}{k_z} = \frac{1}{Z\cos\theta}.$$

For p polarization e is the tangential projection of the total E vector. Many optics tables use a reflected p basis whose tangential component flips sign on reflection. The resulting p reflection coefficient differs in sign from this tangential-e coefficient; predicted powers and a consistent full electric field agree. Here all

layers use the same laboratory tangential axes, which removes unnecessary sign conversions between interfaces.

In the absence of free surface currents the two tangential boundary conditions give

$$e_i + e_r = e_t,\; Y_1(e_i - e_r) = Y_2 e_t.$$

Solving for reflected and transmitted tangential amplitudes yields

$$r = \frac{Y_1 - Y_2}{Y_1 + Y_2},$$

$$t = 1 + r = \frac{2Y_1}{Y_1 + Y_2},$$

$$R = |r|^2,$$

$$T = \frac{\Re Y_2}{\Re Y_1}|t|^2.$$

For lossless propagating waves Y is real and $R + T = 1$ follows by expanding the two squared binomials. The amplitude transmission can exceed one when transmitted admittance is smaller, because energy is proportional to admittance times squared tangential field. Omitting the admittance factor leads to failed power balance even if the two field equations were solved perfectly. At normal incidence into index 1.5 from air, r=−0.20, t=0.80, R=0.04, and T=$1.5(0.80)^2 = 0.96$.

Distinguish Brewster cancellation from total internal reflection

For positive nonmagnetic indices, p reflection vanishes when $n_1/\cos\theta_1 = n_2/\cos\theta_2$. Combined with Snell's law this yields $\tan\theta_B = n_2/n_1$ at a nontrivial interface. At that angle p admittances match even while phase speed and ray direction differ across the boundary. At air-to-glass incidence with index 1.5 the Brewster angle is 56.31°. The s reflection at that same angle is substantial because s admittance varies with cosine in the opposite direction. A measurement of an intensity minimum only approximately zero can reflect polarization impurity, absorption or a surface film; a minimum angle alone cannot establish all three properties.

If $n_1 > n_2$, transmission at incidence above $\theta_c = \arcsin(n_2/n_1)$ gives

$$k_{z2} = i\alpha.$$

The decaying exponential satisfies the same tangential phase matching. For a single semi-infinite lossless second medium its admittance is imaginary, so T vanishes and $|r| = 1$. A finite evanescent field remains on the low-index side and carries energy parallel to the surface. For glass index 1.50 facing air at 50° incidence and 600 nm vacuum wavelength, $\alpha = k_0\sqrt{(1.5\sin 50^\circ)^2 - 1} \approx 5.93 \times 10^6$ m⁻¹. The electric amplitude decays to $1/e$ in about 169 nm; intensity in about 84.5 nm. A second high-index surface within that region changes the boundary condition and can admit transmitted power.

Locate the Brewster zero and total-reflection threshold for two dielectric indices

Propagate a forward and a backward field through one film

For a film of thickness d the entrance phasors are $e_0 = A + B,\ h_0 = Y(A - B)$. At the exit they are $e_d = Ae^{i\delta} + Be^{-i\delta},\ h_d = Y(Ae^{i\delta} - Be^{-i\delta})$, with $\delta = k_z d$. Eliminate A and B to propagate the exit fields back to the entrance:

$$\begin{bmatrix} e_0 \\ h_0 \end{bmatrix} = \begin{bmatrix} \cos\delta & -i\sin\delta/Y \\ -iY\sin\delta & \cos\delta \end{bmatrix} \begin{bmatrix} e_d \\ h_d \end{bmatrix}.$$

A one-line formula conceals useful checks. At zero thickness the matrix is the identity. Its determinant is one because $\cos^2\delta + \sin^2\delta = 1$. At a half-wave thickness the matrix is minus the identity; both amplitudes change sign and their ratio is unchanged. For a stack, form the product in physical order from entrance to exit because each matrix here maps its own exit backward. Reversing two layers can change the response. For very thick evanescent layers transfer matrices contain exponentially large and small entries; a scattering matrix or scaled recursion can avoid catastrophic loss of precision.

If the substrate is semi-infinite and admits only an outgoing wave, its boundary admittance is $Y_b = h_d/e_d$. Divide the two matrix rows to obtain

$$Y_{\text{in}} = \frac{Y_b\cos\delta - iY\sin\delta}{\cos\delta - i(Y_b/Y)\sin\delta}.$$

Reflection from the whole film at its entrance now follows from the same single-interface equation with $Y_2 = Y_{\text{in}}$. At a quarter wave, $\delta = \pi/2$, this simplifies to $Y_{\text{in}} = Y^2/Y_b$. From the incident medium the film and substrate look perfectly matched if $Y = \sqrt{Y_0 Y_b}$. The field in the film still contains both travelling directions; zero reflected field at the entrance fixes a coherent sum of all returns. Adding reflected intensities interface by interface would destroy the interference responsible for the matching.

Design a coating and predict an out-of-band reading

At normal incidence choose $n_0 = 1,\ n_b = 2.25$, with no magnetic response. The required coating index is 1.50. At $\lambda_0 = 600$ nm choose $d = \lambda_0/(4n) = 100$ nm. The uncoated substrate reflects $R = [(1 - 2.25)/(1 + 2.25)]^2 = 0.148$. The ideal coated substrate has R=0 at design wavelength. A half-wave optical thickness instead returns the original admittance and restores the 14.8% reflection. These two limits can detect a factor of two in a code that confuses a one-way phase with a round-trip phase.

For a single film the same result follows by summing multiple reflections. Let $r_{ij} = (Y_i - Y_j)/(Y_i + Y_j)$. Each additional round trip contributes $r_{10}r_{12}e^{2i\delta}$. Summing the geometric series and using $r_{10} = -r_{01}$ yields

$$r = (r_{01} + r_{12}e^{2i\delta})/(1 + r_{01}r_{12}e^{2i\delta}).$$

In the quarter-wave example both interface coefficients are −0.20. If wavelength shifts to 500 nm at fixed indices, $\delta = 0.6\pi$. The reflected power is $0.04[2 + 2\cos(1.2\pi)]/[1 + 0.04^2 + 0.08\cos(1.2\pi)] \approx 0.0163$

. The coating still reduces reflection relative to the bare substrate, but a strictly zero field requires the design phase and amplitudes to match.

At oblique incidence the phase uses the film's normal wavenumber, giving $\delta = 2\pi n d \cos\theta_{\text{film}}/\lambda_0$. Increasing incidence angle shortens the design wavelength that achieves the same phase. At the same time the matching admittances change differently for s and p, so keeping only the phase shift is incomplete. For illumination by a focused beam a range of tangential wavenumbers reaches the substrate. A single normal-incidence calculation cannot predict its whole cone's reflection without an angular/polarization weighting.

Build a finite dielectric mirror from admittance transformations

Suppose the incident side sees an alternating high-index H and low-index L layer, each quarter-wave thick at normal incidence. Starting at the substrate admittance Y_b, the last low layer gives Y_L^2/Y_b; the preceding high layer transforms that to $Y_b(Y_H/Y_L)^2$. Repeating N pairs yields

$$Y_{\text{in}} = Y_b(Y_H/Y_L)^{2N}.$$

The growth is geometric in pair count despite each individual interface reflecting only a fraction of the wave. This is a coherent reflection measurement and depends on all pair phases staying near their design values.

Take $n_H = 2,\ n_L = 1.5,\ n_0 = n_b = 1.5$. With three HL pairs the normalized input admittance is $(4/3)^6 = 5.62$ times its external value, giving R about 0.487. With six pairs the ratio is $(4/3)^{12} = 31.6$, giving R about 0.881. The same total number of random-thickness layers need not achieve those values. Swapping HL for LH in the symmetric exterior configuration inverts the input admittance ratio. It reverses the reflected field phase while leaving the centre-wavelength reflected intensity unchanged. Phase-sensitive reflection can therefore distinguish two stacks with identical intensity at one wavelength.

A stack on a substrate different from the input medium removes that simple intensity degeneracy. The layer order can then alter centre-wavelength reflected intensity as well. Fabrication tolerances introduce additional phase shifts and can split spectral extrema. To identify a real film from measured data, compare phase or several wavelengths and polarization states, and account for instrument spectral width. One scalar normal-incidence reflectance generally admits many index-thickness pairs.

From a stack of interfaces to a spectral ruler

A pair of parallel partially reflecting surfaces makes more than a mirror. A short incident pulse emerges as a train of diminishing replicas, and a continuous wave samples their coherent sum. Work first at normal incidence in a lossless gap of refractive index n and thickness L. Identical ideal mirrors have intensity reflectance R. An extra round trip therefore multiplies a transmitted field by $Re^{i\delta}$, where $\delta = 2n\omega L/c$ plus any frequency-independent mirror phase. At a resonance this phase is an integer multiple of 2π. The directly transmitted field magnitude contains the product of the two mirror transmissions, $1 - R$ in the symmetric lossless case. Sum the geometric series and take the magnitude squared.

$$\frac{I_{\mathrm{out}}}{I_{\mathrm{in}}} = \frac{(1-R)^2}{|1-Re^{i\delta}|^2} = \frac{1}{1+F\sin^2(\delta/2)},$$

$$F = \frac{4R}{(1-R)^2}.$$

The unit transmission at resonance follows from interference and symmetry. It does not imply that the mirrors transmit the whole incident wave on the first encounter. At the front surface a leaked internal wave cancels the directly reflected wave. During switch-on energy accumulates in the gap, so imposing the steady resonance formula on the leading edge of a finite pulse would violate the assumed steady state. If the two mirrors have different transmissions or the gap absorbs, the two reflected amplitudes at the entrance no longer necessarily cancel at resonance.

Adjacent maxima differ in ordinary frequency by $c/(2nL)$ if dispersion and mirror phase slopes are negligible. If the material phase $n\omega L/c$ is dispersive, differentiate the round-trip phase. The local mode spacing becomes the inverse of the round-trip group delay, including the frequency derivative of reflective phase. Replacing n by a phase index measured far from that frequency can then bias an inferred length. At an oblique angle, decide whether the scan holds external angle or tangential wavevector fixed before differentiating the phase.

For $L = 10.0$ mm, $n = 1.50$ and $R = 0.81$, the free spectral range is 9.99 GHz. The half-power points satisfy

$$\sin^2(\delta/2) = 1/F.$$

Their full frequency separation divided by one free spectral range is $2\arcsin(1/\sqrt{F})/\pi$. Here $F = 89.75$ and the linewidth is about 0.673 GHz. The finesse, spacing divided by linewidth, is 14.9. The familiar high-reflectance approximation $\pi\sqrt{R}/(1-R)$ also gives 14.9. For a very weak mirror there need not be half-maximum points at all because the minimum remains above half the peak. The exact formula tells you when a linewidth extracted by that convention is undefined.

Verify the same length in time and find a small hidden loss

The energy of a freely decaying intracavity pulse is multiplied by R^2 each round trip for identical mirrors. With round-trip time $t_r = 2nL/c$, the continuous exponential through the discrete samples has energy lifetime $\tau_E = -t_r/[2\ln R]$. For the above gap $t_r = 100.1$ ps and $\tau_E = 238$ ps. Its reciprocal $1/(2\pi\tau_E)$ is 0.670 GHz, close to the direct linewidth. This equality is an approximation to the isolated narrow resonance. A short enough pulse can resolve the individual replicas and reveal directly that the idealised trace is a train, not a continuously measured exponential at every instant.

Now introduce a weak internal loss. Let a be the intensity fraction remaining after one pass through the material. An additional round trip multiplies the circulating field by Ra, while the direct transmitted intensity is multiplied by a. Thus the resonant transmission and freely decaying energy factor are

$$T_{\mathrm{peak}} = \frac{(1-R)^2 a}{(1-Ra)^2},$$

$$g_E = R^2 a^2.$$

If the measured ratio of successive isolated replicas is 0.630, assuming the mirrors still have $R = 0.81$ gives

$$a = \sqrt{0.630}/0.81 = 0.980.$$

The predicted peak transmission is about 0.83, substantially below one despite the small two-percent single-pass loss. Multiple passes amplify the effect of absorption. Energy reflected in the mode, absorbed in the gap and transmitted are separate channels. A detector collecting only the central transmitted lobe can also show a low peak when the incident beam couples poorly to the cavity, even with very low material absorption.

An inference procedure uses an intensity-calibrated bypass beam to determine the incident power, a direct time spacing to estimate $n_g L$ and ringdown ratios to estimate Ra. If R is unknown, ringdown alone constrains only the product. A resonant transmission measurement supplies a second equation in the ideal symmetric setup, but beam mismatch and mirror imbalance add unknowns. Block one mirror or measure it separately to obtain an independent reflectance. Small thermal changes of gap length shift the resonance between repeated readings. Sweep the frequency and fit the full peak instead of observing its height at a single nominally resonant wavelength.

Finally examine a common interpretation trap with broadband illumination. A spectrometer pixel integrating many evenly spaced narrow resonances reports their average transmission, while a narrow-line source scanned sufficiently slowly recovers their maxima and minima. The difference is predicted by averaging the intensity spectrum with the instrument response, not by summing the fields emitted from mutually incoherent source frequencies. Conversely the repeated passes of a single resolved frequency are mutually coherent and must be summed as fields. The same device contains both averaging operations, in a definite order. This order follows from the detector integration time and the source correlations.

Apply the interface conditions to a measurable power ratio

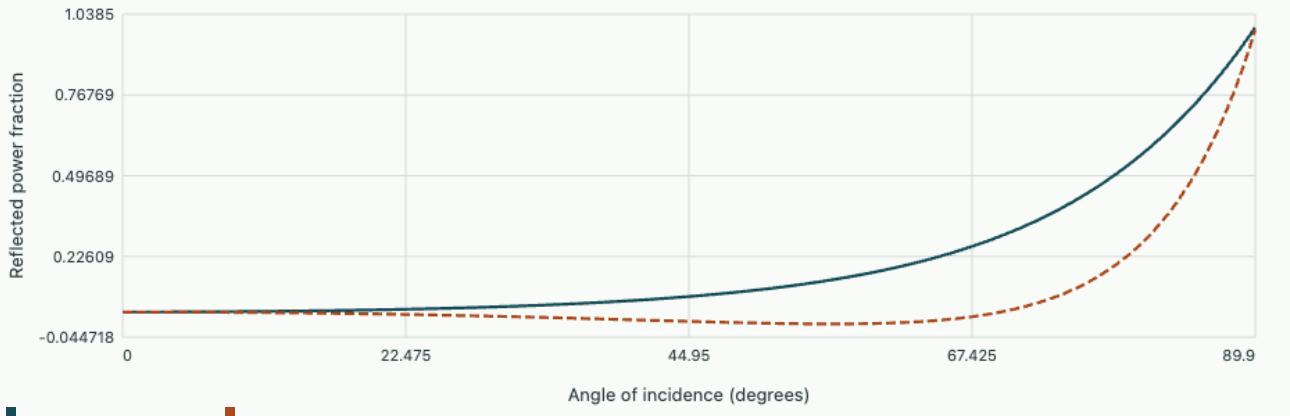
Change incidence angle and the receiving medium. Compare amplitude reflection with the power that actually crosses the interface.

Open this experiment online

Try the calculation

Incident refractive index 1

Transmitted refractive index 1.5



The Brewster minimum occurs at 56.3°. At 45°, reflected power is 9.2% (s) and 0.847% (p).

Worked boundary problem · where does the energy of a totally reflected wave go?

A lossless glass prism with refractive index 1.50 is illuminated internally at 50°, adjacent to air. Vacuum wavelength is 600 nm. Calculate the phase propagation along the interface, the normal field decay length in air and the associated one-sided intensity decay length. Can a detector in the second medium remove power without changing the reflection problem?

▼ **Evaluate the complex wavevector and its physical flux**

Continuity of tangential phase gives

$$k_x/k_0 = 1.50 \sin 50^\circ = 1.149.$$

The normal air wavenumber has imaginary part $\kappa = k_0\sqrt{1.149^2 - 1} = 0.566\,k_0$, giving amplitude decay length $1/\kappa = 169$ nm. Squared field decays over about 84.5 nm. The tangential phase period is

$$\lambda_0/1.149 = 522\,\text{nm}.$$

Treating this as an air plane wave with a real propagation direction would imply an impossible sine and lose the normal decay that satisfies the boundary condition.

For s polarisation take $E_y \propto e^{ik_x x - \kappa z}$. The wave equation gives

$$H_x = -i\kappa E_y/(\mu\omega),$$
$$H_z = k_x E_y/(\mu\omega).$$

In $\mathrm{Re}(\mathbf{E}\times\mathbf{H}^*)/2$, the normal component vanishes because the needed product is purely imaginary, but the tangential component is positive. A finite evanescent field thus accompanies unit reflected power at this ideal isolated interface. Zero time-averaged normal flux does not require a zero field in the adjacent material.

Introducing a second prism or a genuinely absorbing near-field probe changes the boundary condition. A field returning from the second boundary permits interference that can carry normal energy. A perfectly calibrated detector nevertheless perturbs the solution if it extracts appreciable power. Increase the air gap and compare the recovery of unit reflectance as well as the detected signal. A decay in detected intensity alone could also reflect geometric collection changes, while agreement with the complementary reflected-power change constrains the power path.

A wedge turns phase dispersion into a spatial measurement

Two plates forming a small air wedge provide a complementary way to measure a layer without scanning the source frequency. Let the gap rise linearly with transverse position at small angle α. In reflection at nearly normal incidence successive equivalent fringes correspond to a full additional round-trip phase, hence to an added gap of half a wavelength in the intervening medium. For a 633 nm source and a measured fringe spacing of 0.80 mm in air, the inferred wedge slope is about 0.396 milliradians. The phase reversals on the plates choose whether the contact end is bright or dark but do not change this spacing.

If the refractive index is unknown, a single spacing constrains only the product of gap slope and index. An independent surface-profile measurement provides the geometric slope. Alternatively fill the same wedge with a transparent fluid and compare its spacing against air, keeping geometry and temperature fixed. The ratio of air spacing to fluid spacing measures the ratio of indices within the local parallel-film approximation.

Several identifiable departures can occur. A finite camera pixel averages neighbouring fringes, a broad source loses contrast with increasing path difference and a curved plate produces a changing fringe spacing. The first two principally reduce visibility; the third changes local phase gradient. Fit fringe positions and their visibility as separate observables before attributing either effect to absorption inside the film.

Tutorial · measure a Brewster index and bound an unwanted polarization

A transparent nonmagnetic sample in air has an observed p minimum at 58.0°. Assume for this estimation that it is a single dielectric interface. Infer its index, calculate the s reflectance there, and estimate what

fraction of incident power could be s-polarized if the measured minimum reflectance is 0.0040 and all other backgrounds are negligible.

▼ **Derive and use the second polarization as a calibration**

The inferred index is

$$\tan 58^\circ = 1.600.$$

At the Brewster angle the transmitted angle is 32°. Relative s admittance is $n\cos 32^\circ/\cos 58^\circ = n^2$ for air incidence. Thus $R_s = [(1-n^2)/(1+n^2)]^2 \approx 0.192$. An incident s fraction f contributes fR_s to reflected power while the ideal p component contributes zero. Matching 0.0040 gives $f = 0.0208$, about 2.1%. This is an upper bound if another nonnegative reflected background is present. Absorption or a coating can also displace the minimum, invalidating the simple index fit itself. Measuring the complementary s curve and checking transmitted power provide independent tests before interpreting the impurity.

Tutorial · separate a thickness error from an index mismatch

A nominal 600 nm quarter-wave coating with design index 1.5 on an index 2.25 substrate is measured to be 102 nm thick. First hold all indices fixed. Predict the wavelength of its exact null, then estimate R at 600 nm to leading order. If its index is actually 1.55, can changing only the thickness recover an exact null?

▼ **Expand the coherent reflection and apply the admittance condition**

The quarter-wave wavelength is $4nd = 612$ nm. At 600 nm the phase deviates from $\pi/2$ by $\Delta = \pi(0.02)/2 = 0.0314$. With $r_{01} = r_{12} = r_s = -0.20$, the small-amplitude reflection is $r \simeq -2ir_s\Delta/(1-r_s^2)$. Thus $R \simeq 4r_s^2\Delta^2/(1-r_s^2)^2 = 1.71\times10^{-4}$. An index of 1.55 gives unequal magnitudes at the two interfaces. Varying the phase can put their return fields in opposition but cannot equalize their magnitudes. At a quarter wave the best input admittance is proportional to $1.55^2/2.25 = 1.0678$, giving R about 0.00107. A wavelength scan can locate the minimizing optical thickness while the depth of that minimum constrains the impedance mismatch.

MIT 6.013 · boundary conditions, oblique waves, impedance transformations and layered systems

Waves, radiation and relativity · 4. Causal material response and the propagation of a finite pulse

An optical pulse entering a transparent slab is delayed, broadened and sometimes reshaped. Derive these effects from the motion of bound charge, and learn what a measured group delay actually says about a pulse carrying information.

Turn the idea of delayed response into an equation

At a point in a stationary isotropic dielectric, a short electric impulse displaces bound charge. The displacement can continue after the applied field disappears. In a linear time-translation-invariant material write the polarization as a weighted sum over earlier values of the local electric field:

$$P(t) = \epsilon_0 \int_0^\infty h(\tau) E(t-\tau)\, d\tau,$$

$$\chi(\omega) = \int_0^\infty h(\tau) e^{i\omega\tau}\, d\tau.$$

The lower limit encodes causality. Switching to steady harmonic fields yields $\widetilde{P} = \epsilon_0 \chi(\omega) \widetilde{E}$ under the $e^{-i\omega t}$ convention. An instantaneous contribution could also be represented by a delta function at zero delay. The response kernel h has dimensions of inverse time when susceptibility is dimensionless. A short memory compared with the period can sometimes be approximated by a constant susceptibility over a limited band. Extending a constant complex response to all frequencies would generally violate either causality or the requirement of a real time-domain polarization.

For real h, $\chi(-\omega) = \chi(\omega)^*$, so its real part is even and its imaginary part odd. For a decaying kernel, extending ω into the upper half of the complex plane introduces $e^{-\operatorname{Im}\omega\,\tau}$, ensuring convergence. Analyticity there links dispersion and absorption. Conversely, a fitted pole in the upper half-plane in this convention would imply growth, an incompatible time convention, or an unstable model. Report the phasor convention beside an empirical complex permittivity to assign the sign of loss correctly.

Obtain a susceptibility with a recognizable resonance

Represent a bound charge of magnitude q and mass m by a displacement x, restored with frequency ω_0 and damped at rate γ. Choose x along the direction of the force on a positive effective charge; the resulting susceptibility depends on q^2. For a density N of independent oscillators the polarization is Nqx. The equation of motion and its phasor solution give

$$m(\ddot{x} + \gamma\dot{x} + \omega_0^2 x) = qE,$$

$$\chi = \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega},$$

$$\omega_p^2 = \frac{Nq^2}{m\epsilon_0}.$$

Writing $D = (\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2$ gives

$$\chi' = \omega_p^2(\omega_0^2 - \omega^2)/D,$$
$$\chi'' = \omega_p^2\gamma\omega/D.$$

A positive damping rate makes $\chi'' > 0$ at positive frequency. At low frequency the displacement follows the drive almost in phase, while far above resonance inertia produces the opposite displacement. At resonance the displacement lags the drive by a quarter cycle, making the velocity align with force and maximizing the ability of damping to absorb average work for a given displacement amplitude.

The energy balance follows directly by multiplying the equation by $N\dot{x}$. The field supplies $E\dot{P}$, equal to the rate of change of oscillator kinetic plus restoring energy and the dissipated power $Nm\gamma\dot{x}^2$. Averaging a harmonic cycle eliminates the change of stored mechanical energy. With peak phasors,

$$\langle E\dot{P}\rangle = \tfrac{1}{2}\omega\epsilon_0\chi''|\widetilde{E}|^2 = \tfrac{1}{2}Nm\gamma\omega^2|\widetilde{x}|^2.$$

This supplies a practical sign check on both the material fit and the assumed time dependence. The imaginary part that attenuates a passive forward wave must also give nonnegative cycle-averaged heating.

Reconstruct the dispersive shift from its loss spectrum

For a response that vanishes sufficiently fast at high frequency, analyticity in the upper half-plane permits a contour integral of $\chi(z)/(z-\omega)$. Close the contour above and indent around the pole on the real axis. The large arc vanishes under the decay assumption. Taking real and imaginary parts of the resulting principal-value relation, then using the symmetry between positive and negative frequencies, gives

$$\chi'(\omega) = \frac{2}{\pi}\,\mathcal{P}\int_0^\infty \frac{\Omega\chi''(\Omega)}{\Omega^2 - \omega^2}\,d\Omega.$$

The principal value means combine contributions on both sides of the singular frequency before taking the limit. At zero frequency there is no interior singularity and $\chi'(0) = (2/\pi)\int_0^\infty \chi''(\Omega)/\Omega\,d\Omega$. A passive absorption band produces a finite static polarization contribution even when it lies far above the optical frequencies being measured. If instantaneous high-frequency polarization is included in a model, subtract that constant background before using this vanishing-at-infinity form.

As a simple check, let a measured narrow absorption band have area $A = \int \chi''\,d\Omega = 3.0 \times 10^{14}$ s⁻¹ centered at $\Omega_0 = 2.0 \times 10^{15}$ s⁻¹. Across the band $1/\Omega$ varies little, giving an estimated contribution $2A/(\pi\Omega_0) = 0.0955$ to $\chi'(0)$. The same numerical area translated to twice the frequency contributes only half as much. A spectrometer limited to one window cannot recover an absolute static susceptibility without assumptions about absorption outside its window. Tail models and independent low-frequency measurements constrain this extrapolation.

Distinguish the carrier phase from the envelope delay

In a nonmagnetic bulk medium define $k(\omega) = \omega n(\omega)/c$, choosing the branch of n that attenuates a wave travelling in positive z. A component accumulates $\exp[ik(\omega)L]$, so its field amplitude decays as

$\exp[-\operatorname{Im} k\,L]$ and its intensity twice as rapidly in the exponent. In a transparent band take k effectively real and expand around carrier ω_c:

$$k(\omega_c+\Omega) = k_c + k_1\Omega + \tfrac{1}{2}k_2\Omega^2 + \cdots,$$

$$k_1 = \frac{1}{v_g} = \frac{n + \omega_c\, dn/d\omega}{c}.$$

The constant $k_c L$ rotates the carrier phase. The term $k_1 L\,\Omega$ multiplies the envelope spectrum by a linear phase, corresponding to translation in time by $k_1 L$. The second derivative produces different delays for different spectral components and generally changes the pulse shape. These three operations can be measured separately with an interferometer, a time-of-flight experiment and a pulse-duration measurement. A continuous-wave phase measurement at one frequency does not independently determine k_1 or k_2.

At 800 nm suppose an index measurement gives $n = 1.50$ and $dn/d\lambda_0 = -0.050$ per µm, where λ_0 is vacuum wavelength. By differentiating $\omega = 2\pi c/\lambda_0$,

$$n_g = n - \lambda_0\, dn/d\lambda_0 = 1.54.$$

A 100 mm slab has carrier phase travel time $nL/c = 0.5003$ ns and narrowband envelope time $n_g L/c = 0.5137$ ns. Compared with a vacuum path of equal length the envelope delay is 0.1801 ns. The 13.3 ps difference between phase and group times is small relative to a nanosecond but enormous compared with a 50 fs pulse width.

Derive the width of a dispersed Gaussian by completing a square

Use an input envelope $A(0,t) = A_0\exp[-t^2/(2T_0^2)]$, so its intensity is proportional to $\exp[-t^2/T_0^2]$. The amplitude spectrum is proportional to $\exp[-T_0^2\Omega^2/2]$. In a frame at retarded time $u = t - k_1 L$, multiplication by the quadratic propagation phase gives a Gaussian integral with complex coefficient $T_0^2 - ik_2 L$. Evaluate that integral with its physical square-root branch:

$$\frac{A(L,u)}{A_0} = \frac{1}{\sqrt{1 - iD}}\exp\left[-\frac{u^2}{2T_0^2(1-iD)}\right],$$

$$D = \frac{k_2 L}{T_0^2}.$$

Hence $I(L,u) = I_0(1+D^2)^{-1/2}\exp[-u^2/(T_0^2(1+D^2))]$. The intensity width broadens by $\sqrt{1+D^2}$ and the peak falls by the reciprocal factor. Integrating over u recovers the original pulse energy per area, as expected for lossless propagation. The phase inside the exponential is quadratic in u, so the instantaneous frequency varies across the envelope. Measuring only output intensity cannot tell the sign of k_2, since the width depends on its square. A phase-sensitive measurement or a known initial chirp can reveal that sign.

Take $T_0 = 20$ fs and integrated dispersion $k_2 L = 400$ fs². Then D=1. The incident intensity FWHM is $2\sqrt{\ln 2}\,T_0 = 33.3$ fs, and the emerging FWHM is 47.1 fs. The peak intensity falls to 70.7% while the energy

remains fixed. In a second slab of the same length with $k_2 L = -400$ fs², the quadratic phases cancel, provided loss and higher-order dispersion remain negligible and the beams couple without appreciable filtering. Propagating with group velocity alone would predict the arrival of the centre but miss these measurable width changes.

When a measured peak time no longer tracks an unaltered envelope

Near an absorption line both real and imaginary parts of k can vary strongly over a pulse spectrum. A simple propagation experiment then applies a complex filter $H(\omega) = t_{\text{interfaces}}(\omega) e^{ik(\omega)L}$, where interface reflections may also matter. A severely attenuated output maximum may move forward because the trailing part of the incoming waveform is preferentially removed. A negative slope of the transmission phase can therefore coexist with a causal impulse response. The arrival of a newly imposed nonanalytic wavefront requires the high-frequency behaviour of the full causal transfer function. A band-limited Gaussian already has tails at every finite time and has no sharp front.

There is also a distinction between phase delay and energy deposition in an opaque medium. If $n = n' + in'' = 1.5 + 0.02i$ at 600 nm, $\operatorname{Im} k = 2\pi(0.02)/\lambda_0 = 2.09 \times 10^5$ m⁻¹. Ignoring reflections, intensity falls to $1/e$ in $1/[2 \operatorname{Im} k] = 2.39$ µm. After 5.0 µm the intensity fraction is

$$e^{-2.094} = 0.123.$$

A pulse delay inferred from the remaining photons is a conditional timing measurement on transmitted energy; it does not describe energy that was deposited at shallower depths.

Separate a resonance shift, a loss peak and a spectral weight

In the oscillator calculation, hold oscillator strength fixed and compare two damping rates. Predict the imaginary susceptibility at resonance from $\omega_p^2/(\gamma\omega_0)$; halving damping should double this height. The static susceptibility ω_p^2/ω_0^2 stays fixed. A narrower taller peak can preserve its weighted spectral contribution. Record real susceptibility on both sides of the resonance, where its sign changes. Then change the oscillator strength and explain why this change scales absorption and dispersion together. Use the separate half-space calculation to check that the absorbed power integrated over depth accounts for the part of incoming flux that passes the surface.

Change damping in the bound-charge susceptibility experiment

Check how integrated heating closes an energy balance

A causal relaxation whose spectrum can be integrated exactly

As a second material model, consider permanent molecular dipoles that reorient toward a weak applied field. When rotational inertia relaxes much faster than the times of interest, a single exponential approach to equilibrium is useful. Let the additional static susceptibility of these dipoles be $\Delta\chi$, and let τ be the

relaxation time. For the extra polarisation write $\tau\dot{P} + P = \epsilon_0 \Delta\chi E$. This is a different material approximation from a weakly damped oscillator. It still has a retarded response and provides an independent test of the dispersion–loss relation.

Solve the inhomogeneous differential equation by multiplying it by $e^{t/\tau}$, integrating from an unpolarised remote past and dividing by the same factor. A unit step turns on exponentially. A harmonic drive gives a complex susceptibility with one pole on the lower imaginary frequency axis for the chosen $e^{-i\omega t}$ sign convention.

$$P(t) = \frac{\epsilon_0 \Delta\chi}{\tau} \int_{-\infty}^{t} e^{-(t-t')/\tau} E(t')\, dt',$$

$$\chi' = \frac{\Delta\chi}{1 + \omega^2\tau^2},$$

$$\chi'' = \frac{\Delta\chi\, \omega\tau}{1 + \omega^2\tau^2}.$$

At $\omega\tau = 1$ the in-phase and lagging components are equal to $\Delta\chi/2$, and the loss component reaches its maximum. A material with $\tau = 8.0$ ns therefore peaks near 19.9 MHz. A step reaches 63.2 percent of its final additional polarisation after 8.0 ns and 95.0 percent after 24.0 ns. If the chosen observable is current rather than polarisation, the relaxation current is proportional to $\dot{P}$ after the step and decays, while P itself rises. An instantaneous background polarisation can deliver a narrow current spike in addition to this resolved tail.

Substitute χ'' into the zero-frequency dispersion relation. With $x = \omega\tau$, the integral becomes

$$(2\Delta\chi/\pi) \int_0^\infty dx/(1 + x^2) = \Delta\chi.$$

The high-frequency background, if present, must be added separately to obtain the total static response. This exact integral is also a test for loss-spectrum simulations on a logarithmic grid. The integration measure is $d\omega/\omega = d\log\omega$. Equal weights per ordinary frequency or a missing conversion between natural and decimal logarithms gives an incorrect static increment even if the plotted peak is perfect.

A spectrometer covering $0.1 \le \omega\tau \le 10$ captures $2[\arctan 10 - \arctan(0.1)]/\pi = 0.873$ of the static increment. Two full decades centred on the peak miss almost thirteen percent. The missing low and high frequency contributions are equal in this particular symmetric logarithmic loss model. A superposition of slower relaxations or a second unrelated absorption feature destroys this equality. Report a static measurement or bounds on those extra contributions before interpreting a small discrepancy as failure of causality.

Drive the relaxation twice and distinguish memory from propagation delay

Apply a rectangular field E_0 from time zero to T, then switch it off. Continuity of polarisation at both switching instants gives

$$P(T) = \epsilon_0 \Delta\chi E_0 (1 - e^{-T/\tau}).$$

After switch-off the value decays as $P(T)e^{-(t-T)/\tau}$. For $T = \tau$ the maximum is 0.632 of the static value and at 2τ the remainder is 0.233. A pure delay with the same time of maximum would instead translate the rectangular waveform with unchanged amplitude and abrupt edges. One can test this distinction from the complete waveform even if a slow instrument makes identifying the first nonzero response difficult.

Follow the first pulse by an opposite pulse of equal height starting after a zero-field gap g. At the second start the memory is $P_s = \epsilon_0 \Delta\chi E_0 (1 - e^{-T/\tau}) e^{-g/\tau}$. The second pulse drives the polarisation toward the opposite equilibrium. It crosses zero during the pulse after a further time $\tau \ln[1 + P_s/(\epsilon_0 \Delta\chi E_0)]$. For $T = \tau$ and $g = \tau$ the result is 0.209τ, or 1.67 ns with the above time constant. This cancellation depends on the gap. A nonlinear saturation or a detector recovery effect can also produce interactions between pulses, so repeat with smaller field and with a reference cell.

Because the material equation is linear, scaling both pulses by the same factor leaves the zero-crossing time unchanged. The delay-based measurement rejects a constant gain error, but an additive baseline shifts the crossing. An experimentally estimated timing uncertainty is the noise level divided by the local slope magnitude there. At zero polarisation within the negative pulse the equation gives $\dot{P} = -\epsilon_0 \Delta\chi E_0/\tau$, making the conversion from amplitude noise especially simple. This also explains why decreasing the field too far increases the timing noise even though it improves the validity of a linear material response.

To separate propagation from memory, take data through two cell thicknesses using the same cables and electronics. A phase accumulating in the material scales with thickness, while a common cable or detector lag cancels in their phase difference. Reflections give additional echoes whose times scale with thickness too, but follow round-trip distances and can interfere with the direct signal. Gate a sufficiently isolated first arrival or include interface scattering in the transfer function. Taking a ratio of Fourier spectra near a zero of the reference pulse magnifies noise; discarding or regularising those frequencies changes the reconstructed temporal resolution and should be stated with the result.

Compare absorption and phase near a resonance

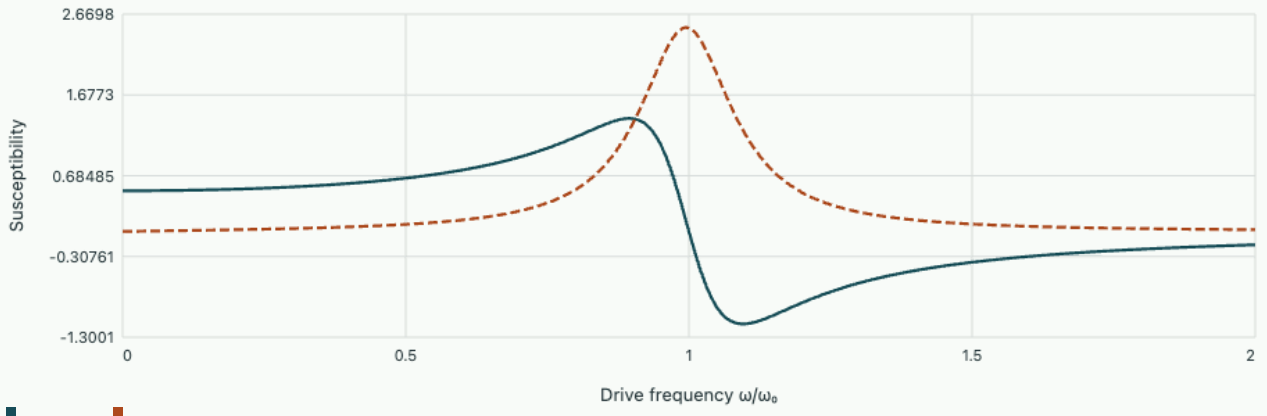
Sweep the drive through a material resonance. Follow the phase response and loss separately before interpreting a pulse arrival time.

Open this experiment online

Try the calculation

Damping γ/ω_0 **0.2**

Oscillator strength ω_p^2/ω_0^2 **0.5**



At $\omega = \omega_0$ the susceptibility is purely imaginary with Im χ = 2.5. At zero frequency the real susceptibility is 0.5.

Worked pulse problem · bound a group-delay measurement

A pulse traverses two samples cut from one transparent material, of lengths 10.0 and 30.0 mm. The measured difference between the carrier-envelope peak times is 103 ps. The extra material replaces air in a fixed total optical path. Estimate the group index. A student obtains a substantially different value from the phase velocity at the carrier. Specify two physical measurements and one numerical test that can resolve the disagreement.

▼ **Use the replaced air and examine the pulse spectrum**

The incremental delay is $(n_g - n_{g,\mathrm{air}})\Delta L/c$, with $n_{g,\mathrm{air}}$ near one in this idealised setting. Thus $n_g \simeq 1 + c(103\,\mathrm{ps})/(20.0\,\mathrm{mm}) = 2.54$. Treating the extra material path as an added path in an otherwise unchanged optical train would give 1.54 instead. A genuinely moved detector would require that different time reference and geometry. A 2 ps independent uncertainty in this differential delay alone contributes about 0.030 to n_g.

First measure phase at several closely spaced frequencies and unwrap its slope. The slope gives group delay, whereas phase at one frequency determines a phase index only after fixing an integer number of waves. Second, shorten the spectrum while keeping its central frequency fixed and compare peak delay with a cross-correlation delay and a centre-of-energy delay. A strongly distorted pulse can yield three different answers even with perfect electronics. Record a transmitted spectrum for each sample to establish whether selective attenuation is significant.

As a numerical test, propagate the measured incident spectrum using the fitted complex wavenumber and reconstruct the physical detector observable, including its impulse response. Compare with the analytic

linear-phase translation after suppressing higher dispersion derivatives and frequency-dependent attenuation. Keep the computational time window large when narrowing the spectrum, since the pulse then extends further in time. A wrapped pulse in a periodic FFT window can acquire an apparently early peak. Looking at the first and last portions of the unwrapped time trace is an inexpensive way to rule that out.

A deliberately chirped input reveals the sign of dispersion

The temporal width of an initially transform-limited Gaussian after propagation depends on squared group-delay dispersion. Measuring one broadened width therefore loses the sign of the quadratic phase. Add a known adjustable quadratic spectral phase before the sample. For a Gaussian with fixed spectral intensity, the broadening now depends on the square of the sum of the imposed and sample phase curvatures. Plot measured pulse width versus the imposed curvature. The minimum identifies the setting that compensates the sample and hence gives the opposite of its signed dispersion.

For example, suppose independent calibration of the phase device assigns its minimum output width to an imposed minus four hundred femtoseconds squared after inserting a plate. If the optical train without the plate minimises at minus one hundred femtoseconds squared, the plate adds plus three hundred femtoseconds squared. The displacement of the minimum removes a common phase offset from the optical train. The widths at their respective minima need not be exactly equal if the plate changes the transmitted spectrum.

Measure that spectrum before interpreting a displaced or broadened minimum. A filter narrowing the spectrum increases even the best possible transform-limited pulse width. Cubic spectral phase and detector response may make the measured width curve asymmetric or leave a long tail near the nominal optimum. Comparing a reconstructed complete temporal profile with its spectrum gives information that a single width at one propagation distance cannot contain.

Tutorial · infer a material slope with a pair of delay measurements

Two samples of the same transparent glass have lengths 20.0 and 70.0 mm. With the receiver and external air path arranged so that the added glass replaces an equal vacuum length, the narrowband arrival time differs by 90.0 ps. At 1.00 µm the independently measured phase index is 1.500. Determine group index and the local index slope. Estimate how an unaccounted fixed electronic time offset affects a two-length measurement.

▼ **Use the difference to eliminate common delays**

With $\Delta L = 0.0500$ m and $\Delta t = (n_g - 1)\Delta L/c$, the inferred index is

$$n_g = 1 + c\Delta t/\Delta L = 1.540.$$

Therefore $dn/d\lambda_0 = (n - n_g)/\lambda_0 = -0.040$ per µm. A fixed additive trigger latency appears in both raw arrival times and cancels in the difference. A sample-dependent connector or surface coating that

changes the pulse transmission phase need not cancel. Comparing a third length tests the assumed affine dependence of delay on material length, while measuring transmitted spectra can reveal frequency-dependent filtering.

Tutorial · distinguish dispersion from simple attenuation

An initially transform-limited Gaussian pulse has intensity FWHM 50 fs. After a transparent segment its width is 80 fs and its peak intensity is 0.50 times the original. Estimate $|k_2 L|$ under the Gaussian model, then infer the additional frequency-independent intensity transmission needed to account for the peak. Could the measured peak alone determine dispersion?

▼ **Recover a width and then an energy ratio**

The initial amplitude parameter is $T_0 = 50/(2\sqrt{\ln 2}) = 30.0$ fs. The width ratio 1.60 gives

$$|D| = \sqrt{1.60^2 - 1} = 1.249.$$

Thus $|k_2 L| \simeq 1.13 \times 10^3$ fs². At conserved energy the peak would be $1/1.60 = 0.625$, so an additional intensity transmission of $0.50/0.625 = 0.80$ reconciles both observations. Equivalently the energy ratio is peak ratio times width ratio, 0.80. A peak value alone admits many combinations of filtering, dispersion and attenuation. The inference additionally assumes that higher-order spectral phase and spectral amplitude filtering do not change the Gaussian shape.

David Tong, University of Cambridge · electromagnetism in matter, dispersion and causality

Waves, radiation and relativity · 5. A guided mode, a cavity line and the energy they carry

A microwave pulse sent through a rectangular guide arrives late even though the guide is evacuated. Terminating the guide converts the travelling pattern into a family of resonances. Use the same transverse eigenproblem to predict its cutoff, pulse delay, cavity frequency and measured linewidth.

State the longitudinal dependence before choosing a transverse basis

Let a hollow rectangular perfect-conductor guide occupy $0 < x < a,\ 0 < y < b$ and extend in z. Fill it uniformly with a lossless dielectric. A single-frequency field varies as $e^{i\beta z - i\omega t}$. In the interior each Cartesian component satisfies the Helmholtz equation. Separating the z dependence leaves the transverse equation $\nabla_t^2 F + k_c^2 F = 0$, with $k_c^2 = k^2 - \beta^2,\ k = \omega\sqrt{\epsilon\mu}$. Conducting walls impose zero tangential E at the perimeter. They do not impose zero H there, since surface current can flow and supply the required magnetic discontinuity.

A transverse-magnetic (TM) mode has $H_z = 0$ and uses E_z as its longitudinal scalar. Because z lies tangential to every side wall, $E_z = 0$ there. The allowed nonzero separable solutions are sine-sine products with positive integers m,n. A transverse-electric (TE) mode has $E_z = 0$ and uses H_z with vanishing outward normal derivative on the side walls. Its scalar modes are cosine-cosine products. One index may be zero but both cannot be zero in this hollow single-conductor guide. The apparent constant H_z TE00 solution fails to reconstruct a nonzero propagating transverse field from Maxwell's equations. A TEM mode can exist in a coaxial line with two conductors maintained at different potentials. A connected perfect conductor surrounding empty simply connected space admits only constant electrostatic boundary potential, precluding a nontrivial transverse potential difference for TEM.

$$k_{c,mn}^2 = (m\pi/a)^2 + (n\pi/b)^2,$$

$$\beta^2 = \omega^2\epsilon\mu - k_{c,mn}^2,$$

$$f_{c,mn} = \frac{1}{2\sqrt{\epsilon\mu}}\sqrt{(m/a)^2 + (n/b)^2}.$$

The cutoff emerges from the room a transverse standing field needs to fit between walls. It is a condition on frequency, geometry and filling material together. Below cutoff a decaying longitudinal solution can still store near fields and mediate coupling over finite lengths. A reading on the far side of a short below-cutoff section therefore need not be exactly zero.

Reconstruct one measurable field from the vector equations

Take $a > b$ and the lowest TE10 mode. A useful choice is $\widetilde{E}_y = E_0 \sin(\pi x/a) e^{i\beta z}$, with $E_x = E_z = 0$. It vanishes at x walls as required. At the y walls E_y is normal and can end on surface charge. Use $\nabla \times \widetilde{\mathbf{E}} = i\omega\mu\widetilde{\mathbf{H}}$ to reconstruct both magnetic components:

$$\widetilde{H}_x = -\frac{\beta}{\omega\mu}\widetilde{E}_y,$$

$$\widetilde{H}_z = -\frac{i\pi E_0}{\omega\mu a}\cos(\pi x/a)e^{i\beta z}.$$

Ampère's law then reproduces the dispersion relation. The negative x direction of H for positive E_y and positive β gives positive S_z. The TE wave impedance is $Z_{\mathrm{TE}} = \omega\mu/\beta$ using transverse magnitudes. Unlike a plane-wave impedance in a homogeneous bulk medium, this one changes with frequency even when ϵ, μ are constant. At cutoff transverse magnetic H_x vanishes for a fixed finite E_y, while H_z remains finite. This is why large field amplitude near cutoff need not imply substantial transmitted axial power.

Integrate power through the cross-section using $\langle S_z\rangle = \beta|\widetilde{E}_y|^2/(2\omega\mu)$. The integral of $\sin^2(\pi x/a)$ is $a/2$. The mode therefore carries $P = ab\,|E_0|^2/(4Z_{\mathrm{TE}})$. Using simply the area times the central peak flux would overestimate power by a factor of two. A small receiving probe samples a spatially weighted field. Its calibration can be checked by moving it from the centre to x=a/6, where field amplitude halves and local axial flux falls by a factor of four.

Compute two velocities and check them by energy transport

With bulk phase speed $v = 1/\sqrt{\epsilon\mu}$, the dispersion relation is

$$\omega^2 = v^2(\beta^2 + k_c^2).$$

Taking the frequency-to-wavenumber ratio and its derivative gives

$$v_{\mathrm{ph}} = \frac{\omega}{\beta} = \frac{v}{\sqrt{1-(\omega_c/\omega)^2}},$$

$$v_g = \frac{d\omega}{d\beta} = v\sqrt{1-(\omega_c/\omega)^2},$$

$$v_{\mathrm{ph}}v_g = v^2.$$

Phase fronts can move faster than the bulk wave speed along the guide. They are surfaces of equal phase and carry no independent message. A narrowband unattenuated envelope travels at the smaller group speed. An energy calculation provides a second route to the same quantity. Insert all three TE10 field components into $\overline{u} = (\epsilon|\widetilde{E}|^2 + \mu|\widetilde{H}|^2)/4$, integrate over the cross-section, and use $\beta^2 + (\pi/a)^2 = \omega^2\epsilon\mu$. Electric and magnetic energies per length are each $\epsilon ab|E_0|^2/8$. Dividing P by total energy per length yields

$$\beta/(\omega\epsilon\mu) = v_g.$$

Keeping only the transverse magnetic component would spoil the equality.

For an evacuated guide $a = 30.0$ mm, $b = 12.0$ mm, $f_{c,10} = 4.997$ GHz. At 8.00 GHz, $\sqrt{1-(f_c/f)^2} = 0.781$, $v_g = 2.34\times10^8$ m/s and guide wavelength $2\pi/\beta = 48.0$ mm. A 2 m section gives 8.54 ns group delay, compared with 6.67 ns in bulk vacuum. TE20 turns on at 9.99 GHz and TE01 at

12.49 GHz, so the chosen frequency is inside the single-propagating-mode band. With central electric peak 100 V/m, $Z_{\mathrm{TE}} \simeq 482\ \Omega$ and $P \simeq 1.87$ mW. Knowing frequency is essential to infer power from the same central probe in another band.

Close the guide and quantize its longitudinal phase

Add perfect end walls at z=0 and z=L. Tangential E_y must vanish at each end, so superpose equal counterpropagating modes into $\sin(p\pi z/L)$ with $p = 1, 2, \ldots$ for this TE10 family. The resonance frequencies become

$$f_{10p} = \frac{v}{2}\sqrt{a^{-2} + (p/L)^2}.$$

The common estimate $c/(2L)$ is the free spectral range for a nondispersive one-dimensional TEM cavity. It omits the transverse wavenumber in this example. For the guide above closed at $L = 60$ mm, $f_{101} = 5.59$ GHz and $f_{102} = 7.07$ GHz. Their separation is about 1.48 GHz, considerably smaller than the nominal 2.50 GHz TEM spacing. At sufficiently high longitudinal order the separation approaches $v/(2L)$, though other transverse families can then populate the spectrum.

The mode's field has nodes and antinodes in two coordinates. A tiny probe on an electric node may leave a resonance undetected, and moving a lossy probe can also shift its frequency and damping. Two nearby spectral peaks could be adjacent p indices, two different transverse modes or a degeneracy split by manufacturing errors. Compare measured spatial patterns and polarization selection rules before indexing solely by a frequency spacing.

Derive amplitude decay, stored-energy decay and linewidth

In practice wall conduction, dielectric loss and deliberate port coupling all remove stored energy. If a lightly damped mode retains the same shape as it decays, energy obeys $dU/dt = -\kappa U$. Then $U = U_0 e^{-\kappa t}$, field amplitude falls as $e^{-\kappa t/2}$, and loaded quality factor is

$$Q_L = \omega_0/\kappa.$$

Energy lost per cycle is approximately $2\pi U/Q_L$ for small fractional loss. The exact integrated energy lost over a cycle is $U[1 - e^{-2\pi/Q_L}]$. These expressions distinguish a rate in radians per second from linewidth in ordinary hertz.

Write a complex amplitude in a frame rotating at ω_0. Under a weak frequency-swept harmonic drive its steady amplitude is proportional to $[\,\kappa/2 - i(\omega - \omega_0)\,]^{-1}$. The power response is a Lorentzian. Its half-maximum points have $|\omega - \omega_0| = \kappa/2$, so the full power width is κ in angular frequency, or f_0/Q_L in hertz. A field-amplitude reading falls to $1/\sqrt{2}$ at those same frequencies. Half the field magnitude corresponds to a different bandwidth.

Consider $f_0 = 6.00$ GHz and $Q_L = 6000$. The power-line FWHM is 1.00 MHz, the energy e-folding time $1/\kappa = 159$ ns, and the field e-folding time 318 ns. Observe a field e-folding time of 159 ns instead, and the inferred Q would be 3000, even though the number printed on the oscilloscope is unchanged. Always

determine whether an instrument's output voltage is proportional to field, instantaneous power or a filtered logarithm of either.

Separate a feed port from dissipative losses

Normalize a mode amplitude a so $|a|^2 = U$, and an incident port amplitude s_+ so $|s_+|^2 = P_{\mathrm{in}}$. In a simple one-port cavity the internal loss rate κ_i and external escape rate κ_e add. A consistent choice of port phase gives

$$\dot{a} = [i\Delta - (\kappa_i + \kappa_e)/2]a + \sqrt{\kappa_e}s_+,$$

$$s_- = s_+ - \sqrt{\kappa_e}a.$$

At steady on-resonance drive the reflected amplitude fraction is $(\kappa_i - \kappa_e)/(\kappa_i + \kappa_e)$. Exact cancellation at critical coupling is compatible with continued reflection when driving off resonance, and with finite resonant energy stored in the cavity. Once the drive is removed that energy decays through both channels.

With rates $\kappa_i/(2\pi) = 0.30$ MHz and $\kappa_e/(2\pi) = 0.70$ MHz, the resonant reflected power fraction is 0.16. The remaining 84% of the steady incoming power is dissipated internally in this one-port model. Interchanging the two rates preserves both the linewidth and reflected power at resonance but reverses the reflected phase. A calibrated coherent reflection or a second measurement after deliberately changing coupling can resolve this ambiguity.

Compare field and energy ring-down and linewidth at several loaded quality factors

Change guide width and follow cutoff and axial propagation

Use a small bead to map a cavity mode

A resonator offers a way to measure a field without placing a conventional field sensor at its maximum. Insert a small dielectric perturbation, record the change in resonance frequency, then scan its position. The derivation requires a closed, lossless cavity or a sufficiently high-Q nearly closed cavity for which one isolated mode is meaningful. Take fixed conducting walls, nondispersive permittivity and permeability, and a perturbation small enough to change the frequency and the shape smoothly. A nearby degeneracy requires a coupled-mode treatment.

In the source-free eigenproblem $\nabla \times \mu^{-1}\nabla \times \mathbf{E} = \omega^2\epsilon\mathbf{E}$, multiply by the complex conjugate of the unperturbed mode and integrate over the volume. Subtract the conjugate unperturbed equation multiplied by the perturbed mode. Integration by parts transfers the double curl with no surface contribution at a fixed ideal conducting wall. Expand ω^2 and $\epsilon + \delta\epsilon$ through first order. The unknown first-order change of mode shape cancels from the difference, leaving

$$\frac{\delta\omega}{\omega} = -\frac{1}{2}\frac{\int \delta\epsilon(\mathbf{r})|\mathbf{E}(\mathbf{r})|^2\,dV}{\int \epsilon(\mathbf{r})|\mathbf{E}(\mathbf{r})|^2\,dV}.$$

A positive permittivity increment lowers the resonance frequency. The denominator is proportional to total electric energy for a nondispersive mode, and the fractional shift is the electric-energy overlap with the

perturbation. The result describes a small perturbation of the permittivity distribution, not automatically a high-contrast conducting bead. A high-index sphere changes its own internal field even when it occupies a tiny volume; its electrostatic polarizability can then replace the simple weak-contrast volume approximation. The distinction matters if the shift is used for absolute calibration.

Suppose a weak dielectric insert raises local permittivity by 10.0 percent and covers a region containing 0.40 percent of the unperturbed electric energy. At 3.00 GHz the predicted fractional shift is -0.00020, or -600 kHz. Moving it to a position where the electric amplitude is half as large gives a 150 kHz shift for the same orientation, volume and local background. A map of resonance displacement measures squared field magnitude, so by itself it cannot determine the sign of an oscillating electric component. Two neighbouring antinodes separated by a nodal plane can produce nearly equal shifts with opposite electric phases.

A mode with loaded $Q = 3000$ at 3.00 GHz has a full linewidth of about 1.00 MHz. Its centre can still be fitted more accurately than the linewidth if signal-to-noise ratio and frequency stability are adequate. Record an unloaded baseline between bead positions and keep the probe coupling fixed. A systematic drift of 50 kHz is a small part of a linewidth but already one third of the expected shift at the lower-field position. If two modes draw close when the bead is moved, fitting one Lorentzian over their sum can imitate a position-dependent loss as well as a frequency shift.

A finite below-cutoff length is a boundary problem

A guide section whose infinite-guide propagation constant is imaginary does not simply have zero field. It supports decaying and growing spatial solutions. In a semi-infinite region an outgoing bounded solution picks the decaying branch. A finite narrowed section joined to a second propagating guide needs both branches to satisfy the two interfaces. This is the electromagnetic version of transmitting through a classically forbidden interval in a one-dimensional wave equation. No energy flows out of an isolated purely decaying transverse mode of the uniform lossless section, yet a properly driven combination can carry power through the interval.

Write a scalar model of one isolated mode as $\psi'' + k^2\psi = 0$ in identical leads and $\psi'' - \kappa^2\psi = 0$ in the central interval of length d. Assume continuity of ψ and ψ' at both seams. This mimics a single matched constitutive parameter and deliberately neglects higher transverse modes excited at a real step. Propagation of the pair (ψ, ψ') across the barrier uses hyperbolic functions. Matching it to an incident plus reflected wave on the left and one transmitted wave on the right gives

$$T = \left[1 + \frac{(k^2 + \kappa^2)^2}{4k^2\kappa^2}\sinh^2(\kappa d)\right]^{-1}.$$

At $k = \kappa$ this simplifies to $T = \operatorname{sech}^2(\kappa d)$. For $\kappa = 40$ per metre and lengths 25 and 50 mm, the transmitted powers are 0.420 and 0.0707 times incident. Doubling length does not square the transmission, because the reflected wave and the coupling at both interfaces matter. At large κd an exponential remains, multiplied by a matching-dependent prefactor. Extrapolating a long-section straight line on a log-power plot to zero length should therefore not be interpreted as an additional physical loss in a zero-length guide.

For a real device use modal matching, or a numerical Maxwell solution retaining evanescent transverse modes at a discontinuity. Check convergence against the number of included modes at each step, not just the resolution along the uniform section. Measurements with several lengths can separate the uniform attenuation constant from entrance and exit effects. Measuring reflection as well as transmission discriminates a lossless cutoff barrier, where rejected energy returns to the source, from an absorber that dissipates it. Finally use a time-domain envelope narrow enough in frequency that the fraction above the cutoff is negligible. A small above-cutoff spectral tail can dominate a pulse recorded after a long barrier and give a misleading apparent decay length.

Infer a resonator timescale from its line

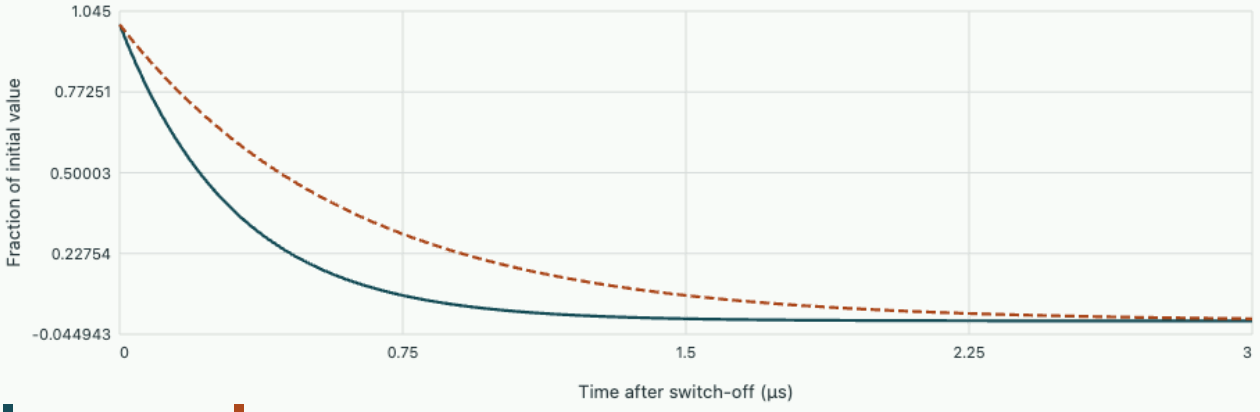
Change loss at a selected cavity mode and compare linewidth with energy decay. State whether each reported rate describes amplitude or energy.

Open this experiment online

Try the calculation

Loaded quality factor Q **6000**

Cavity length d (mm) **80**



The mode frequency is 3.123 GHz with full power linewidth 520 kHz. The energy e-folding time is 0.306 µs.

Worked resonance problem · a square guide with a slight width error

A nominally square air-filled guide of side 20.0 mm supports two orthogonal fundamental patterns with identical cutoff in the ideal square section. The fabricated horizontal dimension is 20.10 mm; the other remains 20.00 mm. Estimate the two cutoffs and the polarisation beating length at 10.0 GHz. How could a single scalar power measurement miss the effect?

▼ Follow the phases of the two modes

For the rectangular TE modes with one half-cycle across either side, the cutoffs are $c/(2a)$ and $c/(2b)$. These are 7.458 and 7.495 GHz, separated by about 37 MHz. At 10.0 GHz the longitudinal propagation constants are $k_{zj} = (2\pi/c)\sqrt{f^2 - f_{cj}^2}$. Their difference is about 0.88 per metre, corresponding to a full relative phase cycle over approximately 7.1 m. It takes only one quarter of this length to build a quarter-cycle retardance between equally excited components.

Near cutoff the propagation-constant sensitivity to a small width change is enhanced because the longitudinal wavenumber appears in the denominator of its derivative. Approaching the higher cutoff too closely also makes the interpretation vulnerable to wall loss, poorly matched launchers and higher-order evanescent contributions at the ends. Above both cutoffs in a sufficiently long uniform section, project the exit field on a pair of known orthogonal receiving patterns and measure their coherent phase difference. Two power readings along the eigenaxes remain unchanged in the lossless model even while a diagonal analyser shows oscillating power.

Rotating the input launcher onto one eigenpattern should suppress the beating, since no second component then accumulates a relative phase. Repeating the experiment at several frequencies tests the predicted dispersion of $k_{z1} - k_{z2}$, distinguishing a distributed width mismatch from a nearly constant phase offset in two receiver cables. A slow twist of the guide can additionally couple the patterns instead of merely advancing their phases. In that situation first measure the full two-by-two complex transfer matrix; treating every disappearing component as attenuation would misidentify coherent mode conversion.

A vanishing resonance can be a coupling zero

A cavity spectrum measures modes through a particular feed and pickup. For a short aligned electric probe, coupling to a mode is proportional at leading order to the projection of the modal electric field onto the probe's current distribution. A probe at a nodal plane or oriented perpendicular to that field can therefore fail to display a well-defined cavity mode. Moving the receiver away from the node or rotating it can restore the spectral line without changing the cavity eigenfrequency.

As a quantitative illustration consider a mode whose electric field varies as a sine across the width. Move a small probe from one quarter of that width to one twelfth. The ratio of its squared coupling magnitudes is approximately 0.134. With an unchanged weak drive and an effectively nonloading pickup, the transmitted power in that mode can shrink by almost an order of magnitude. If the same port also substantially loads the cavity, its movement changes the linewidth and the resonant intracavity field as well.

A mapping experiment should record peak frequency, linewidth and integrated spectral response alongside peak power. A frequency shift with little change in linewidth supports a predominantly reactive perturbation, whereas linewidth changes signal modified damping or port coupling. A changing peak alone cannot distinguish a weak field region, a rotated polarisation and an altered loaded resonance. Combining two feed orientations with a weak scanning pickup supplies a more informative field map.

Tutorial · infer guide width from a delay without assuming a vacuum transit

A pulse at 9.00 GHz traverses 1.50 m of evacuated rectangular guide in 6.00 ns after cable delays are removed. Assume a narrowband TE10 packet with negligible loss and $b = 10.0$ mm. Infer its cutoff and guide width, check the next two candidate cutoffs and estimate its modal impedance.

▼ **Invert the dispersion and verify the mode assignment**

The group speed is 2.50×10^8 m/s, or 0.834 of c. Hence $f_c = f\sqrt{1-(v_g/c)^2} \simeq 4.97$ GHz and $a = c/(2f_c) \simeq 30.2$ mm. TE20 cuts on at about 9.94 GHz and TE01 at 15.0 GHz, both above the test. The TE impedance is $Z_0/(v_g/c) \simeq 452\ \Omega$. A broad pulse overlapping 9.94 GHz would require a mode-launch and dispersion analysis before interpreting one transit time with this inversion. Measuring delay at a second centre frequency tests the inferred cutoff without another width parameter.

Tutorial · separate an internal Q from a port Q

A one-port cavity has $f_0 = 5.00$ GHz and measured power linewidth 0.80 MHz. Halving the port coupling rate with negligible resonance shift reduces linewidth to 0.60 MHz. Infer both initial rates in hertz, both quality factors, the initial resonant reflected fraction and the energy lifetime after the coupling change.

▼ **Combine the rate sums from both tests**

Let $g_i = \kappa_i/(2\pi)$ and $g_e = \kappa_e/(2\pi)$. Solve $g_i + g_e = 0.80,\ g_i + g_e/2 = 0.60$ MHz to get $g_e = g_i = 0.40$ MHz. Both initial individual quality factors are $5.00\,\mathrm{GHz}/0.40\,\mathrm{MHz} = 12500$, and initial loaded Q is 6250. Initial reflection vanishes by critical coupling. After the change the resonant reflected amplitude is 1/3 and reflected power 1/9. The energy lifetime is

$$1/(2\pi\,0.60\,\mathrm{MHz}) = 265\,\mathrm{ns}.$$

This result relies on a fixed internal loss. A coupler adjustment that introduces dielectric loss would invalidate the two-equation inference.

David Staelin, MIT 6.013 · chapter 9 on waves guided by planar and rectangular structures

Waves, radiation and relativity · 6. Constructing the retarded Green function from an outgoing spherical shell

A short source pulse reaches two receivers at different times and amplitudes. Derive the signal from a localized spacetime impulse, see why the potential is evaluated at a past time, and identify which observables change when the source occupies an extended region.

Turn the curl constraints into wave equations with sources

Use vacuum SI units, a localized charge density ρ and current density $\mathbf{J}$ satisfying $\partial_t \rho + \nabla \cdot \mathbf{J} = 0$. Introducing $\mathbf{B} = \nabla \times \mathbf{A}$ enforces the absence of magnetic charge. Setting $\mathbf{E} = -\nabla\Phi - \partial_t \mathbf{A}$ then enforces Faraday's law. These definitions leave freedom to add $\nabla\Lambda$ to $\mathbf{A}$ while subtracting $\partial_t \Lambda$ from Φ. An individual value of scalar or vector potential therefore depends on convention, whereas electric and magnetic fields constructed from them do not.

Adopt the Lorenz condition $\nabla \cdot \mathbf{A} + c^{-2}\partial_t \Phi = 0$. Insert the potentials into Gauss's and Ampère's laws and use $\nabla \times \nabla \times \mathbf{A} = \nabla\nabla \cdot \mathbf{A} - \nabla^2 \mathbf{A}$. The gradient and time-mixed terms cancel in this gauge:

$$(\nabla^2 - c^{-2}\partial_t^2)\Phi = -\rho/\epsilon_0, \qquad \left(\nabla^2 - c^{-2}\partial_t^2\right)\mathbf{A} = -\mu_0 \mathbf{J}.$$

The same scalar propagation operator acts on each component. Taking a divergence and time derivative of these equations shows that the Lorenz-condition combination itself obeys a source-free wave equation when current is conserved. If it vanishes initially with compatible time derivative and boundary conditions it remains zero. Independently prescribing an oscillating point charge and zero current would violate continuity and can undermine both this consistency and any computed radiation field.

Determine the radial form and its source strength

Define a causal distribution $G(\mathbf{R}, \tau)$ solving $(\nabla^2 - c^{-2}\partial_\tau^2)G = -\delta^3(\mathbf{R})\delta(\tau)$ and vanishing for $\tau < 0$. In free space rotational symmetry makes the impulse response depend on $R = |\mathbf{R}|$. For $R > 0$,

$$\nabla^2 G = R^{-1}\partial_R^2 (RG).$$

Thus RG satisfies the ordinary one-dimensional wave equation. The solution for an outward-only signal from the origin has the form $g(\tau - R/c)$. An incoming term $g_{\text{in}}(\tau + R/c)$ would specify radiation arriving from infinity. It can occur with an incident field, but is excluded for the initially quiet source experiment.

To fix g, integrate the equation over a small sphere of radius a and pair its time dependence with a smooth test function. The flux of ∇G through that sphere has leading term $-4\pi g(\tau - a/c)$. Additional terms proportional to $a\, g'$ tend to zero as distributions; the integral of the time derivative likewise vanishes as the sphere shrinks. The prescribed delta requires $-4\pi g(\tau) \to -\delta(\tau)$. Consequently

$$G(\mathbf{R}, \tau) = \frac{\delta(\tau - R/c)}{4\pi R}.$$

This normalization depends on which operator was given unit source. If instead one uses $\partial_t^2 - c^2\nabla^2$, the corresponding unit impulse response carries an additional $1/c^2$. The energy spreading as R^{-2} does not

replace the field amplitude R^{-1}; the two concern different observables. In three spatial dimensions the free-space impulse is concentrated on the light cone. This sharp shell property need not hold in a material with memory, a waveguide, or a two-dimensional scalar wave problem.

Perform the time integral before interpreting retardation

A distributed source is a superposition of shifted impulses. Convolve each charge or current component with G over $\mathbf{r}'$ and t', keeping the observation point $\mathbf{r}, t$ fixed. The delta sets $t' = t - R/c$, with $R = |\mathbf{r} - \mathbf{r}'|$. Because t' occurs linearly for a fixed spatial source coordinate, this integration has unit Jacobian. The potentials are

$$\Phi(\mathbf{r}, t) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}', t - R/c)}{R}\, d^3 r',$$

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r}', t - R/c)}{R}\, d^3 r'.$$

Every source point is sampled at its own retarded time. Replacing the whole extended source by its history at $t - r/c$ discards phase variation across it. This may be acceptable only when that variation is small or has been explicitly integrated. A charge distribution fixed in space and time recovers the ordinary Coulomb potential after the initial transient passes. The specified radiation condition selects a particular solution; a homogeneous incident pulse may be added if the actual experiment includes one.

For the $e^{-i\omega t}$ convention the frequency-domain kernel is $e^{ikR}/(4\pi R)$ with $k = \omega/c$. Substituting a harmonic source into the time-domain potential gives the positive outward radial phase directly. The incoming Helmholtz kernel $e^{-ikR}/(4\pi R)$ solves the same differential equation away from the origin but has the opposite boundary condition at infinity. One physical check is that outgoing radiation transports energy through a large sphere away from the source.

Background · Why does an ideal line source leave a tail after the front?

The three-dimensional scalar retarded kernel for the wave operator $c^{-2}\partial_t^2 - \nabla^2$ has support on the light cone. Now switch on an impulse simultaneously along an ideal infinite straight line. At perpendicular distance ρ, the nearest segment contributes first, at $t = \rho/c$. More distant segments continue to arrive even though each point source emitted only once. Integrating the point kernel along the line gives, for $t > \rho/c$,

$$G_{\text{line}}(\rho, t) = \int_{-\infty}^{\infty} \frac{\delta(t - \sqrt{\rho^2 + z^2}/c)}{4\pi\sqrt{\rho^2 + z^2}}\, dz = \frac{1}{2\pi\sqrt{t^2 - \rho^2/c^2}}.$$

Before the first arrival the integral vanishes. At later times its two contributing segments lie at $z = \pm\sqrt{c^2 t^2 - \rho^2}$. This geometrical tail can be present even with a nondispersive propagation speed. A finite line length cuts it off when the contributions from the ends have arrived.

Jefimenko equations from differentiating the source clock

Let brackets denote source quantities at $t_R = t - R/c$ and $\widehat{\mathbf{R}} = (\mathbf{r} - \mathbf{r}')/R$. At fixed observation time,

$$\nabla R = \widehat{\mathbf{R}},$$

$$\nabla t_R = -\widehat{\mathbf{R}}/c.$$

Therefore differentiating $[\rho]/R$ gives both a derivative of inverse distance and a derivative of charge history. Combining $-\nabla\Phi$ with $-\partial_t\mathbf{A}$ leads to

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\int\left(\frac{[\rho]\widehat{\mathbf{R}}}{R^2} + \frac{[\dot\rho]\widehat{\mathbf{R}}}{cR} - \frac{[\dot{\mathbf{J}}]}{c^2R}\right)d^3r'.$$

Similarly take the curl of $[\mathbf{J}]/R$. The gradient of $1/R$ crossed into J gives $[\mathbf{J}]\times\widehat{\mathbf{R}}/R^2$, while the retarded-time derivative gives $[\dot{\mathbf{J}}]\times\widehat{\mathbf{R}}/(cR)$. The terms individually contain charge and current derivatives, but exact continuity produces important cancellations. For example, a spherically symmetric source with conserved total charge cannot radiate a time-varying electric monopole in vacuum. Outside a symmetric redistribution Gauss's law fixes the radial field by the constant total charge. Predicting radiation from just one term of the retarded electric integral would obscure that cancellation.

Use delay and geometrical spreading as separate diagnostics

Take an ideal compact scalar emitter as a model of a single outgoing far-field polarization trace, with waveform $s(t) = s_0 e^{-(t/T)^2}$. A receiver measuring that field with negligible bandwidth limitation reads $Cs(t - r/c)/r$. At 0.60 m the peak is delayed 2.00 ns and at 1.80 m it is delayed 6.00 ns after emission. The field peak falls by three, so power flux falls by nine. Multiplying each waveform by its distance and shifting its time origin back by r/c should collapse the two records. A mismatch in width after doing so suggests dispersion, source extent, detector bandwidth, multipath or near-field contributions. A single peak-height comparison does not distinguish these effects.

Suppose the emission has compact duration 4 ns and receivers are separated radially by 1.20 m. Their onset times differ by 4.00 ns. If identical receivers have a 1 ns electrical delay, their differential time of flight remains unchanged; if their amplifiers have different group delays, the difference directly biases the inferred propagation speed. A pair of distances can determine one offset and one speed if their absolute radii are known. Three or more distances also permit checking whether a straight-line timing model fits the data.

Translate and rescale two records of a compact retarded pulse

Keep the directional phase when the source is far away

Write $\mathbf{r} = r\mathbf{n}$ with source extent d and $r \gg d$. Expand the distance as $R = r - \mathbf{n}\cdot\mathbf{r}' + [r'^2 - (\mathbf{n}\cdot\mathbf{r}')^2]/(2r) + \cdots$. Relative amplitude variation in $1/R$ is order d/r. The first directional phase $k\mathbf{n}\cdot\mathbf{r}'$ can remain large even when d/r is tiny. For a harmonic current,

$$\widetilde{\mathbf{A}}(\mathbf{r}) \simeq \frac{\mu_0 e^{ikr}}{4\pi r} \int \widetilde{\mathbf{J}}(\mathbf{r}') e^{-ik\mathbf{n}\cdot\mathbf{r}'}\, d^3r'.$$

Neglecting the next, quadratic, phase requires kd^2/r small up to a geometry-dependent coefficient. This is a separate far-zone condition from $r \gg d$. The integral is a spatial Fourier component of the current, evaluated on the sphere of radius k. Measurements from one propagation direction reveal one projected spatial frequency and generally cannot reconstruct an arbitrary extended source.

For two equal emitters separated by d along z, observed at angle θ to z, the two contributions differ by delay $d\cos\theta/c$. In harmonic operation with equal drive phase the amplitude factor is $2\cos(kd\cos\theta/2)$. At 1.00 GHz, $d = 0.150$ m is almost half a free-space wavelength. Along the source axis the two arrivals differ by 0.500 ns, nearly half a period, so the equal contributions cancel. Broadside they arrive together and add. For pulsed excitation with a duration much shorter than 0.500 ns, the end-on receiver sees two time-separated pulses instead of a permanent null across the whole spectrum. The time-domain and frequency-domain statements agree after applying the appropriate bandwidth.

Recover source time from a moving point instead of a fixed grid

The earlier time integral held the source coordinate fixed. A point following a trajectory contributes $\rho(\mathbf{r}', t') = q\,\delta^3[\mathbf{r}' - \mathbf{r}_q(t')]$. If one instead integrates the spatial delta first, the remaining condition is

$$g(t') = t - t' - |\mathbf{r} - \mathbf{r}_q(t')|/c = 0.$$

The source-to-observer distance now depends on the integration time. Its derivative is $-1 + \widehat{\mathbf{R}} \cdot \boldsymbol{\beta}$, giving

$$\int f(t')\delta(g(t'))dt' = \frac{f(t_{\mathrm{ret}})}{1 - \widehat{\mathbf{R}} \cdot \boldsymbol{\beta}_{\mathrm{ret}}}.$$

For subluminal motion this denominator is positive. It records how emission intervals map into observation intervals and must appear when evaluating retarded point-source potentials. Simply replacing position by retarded position in a static Coulomb formula misses this Jacobian. A charge moving towards a distant receiver compresses successive arrival intervals; transverse motion has no first-order time compression along that line of sight. Later units relate this observable to Doppler shifts and relativistic angular radiation patterns.

As a numerical example, two features emitted 10 ns apart by a compact laboratory source whose radial velocity component is $0.20c$ arrive about 8.0 ns apart at a distant fixed receiver, provided the direction and velocity vary negligibly during the interval. The same proper-time interval of a fast moving clock would also require time dilation in converting to the laboratory emission interval. Keeping those two clocks separate avoids applying the Doppler factor twice.

A causal Green function in front of a grounded plane

The free-space Green function encodes an outgoing wave from a point impulse. A boundary changes the response and can be built into the Green function before integrating over sources. As an exactly soluble example choose the scalar wave equation in $z > 0$ with a homogeneous Dirichlet condition on $z = 0$. This is a scalar field problem, and can describe a selected tangential electric component in a geometry whose conducting boundary sets that component to zero. It is not a universal image rule for every vector potential component. The boundary conditions on electromagnetic scalar and vector potentials depend on the chosen gauge and physical surface.

Place a source at $\mathbf{s} = (x_s, y_s, h)$ and its mirror at $\mathbf{s}^* = (x_s, y_s, -h)$. Every point on the plane is equidistant from these points. Hence two opposite-sign delayed spherical responses cancel there at every time. The mirror source lies outside the physical domain and introduces no delta function inside it. Both contributions vanish before their respective light cones. Uniqueness of the retarded boundary-value problem identifies their difference as the required Green function.

$$G_D(\mathbf{r}, \mathbf{s}, t) = \frac{\delta(t - R/c)}{4\pi R} - \frac{\delta(t - R^*/c)}{4\pi R^*},$$
$$R = |\mathbf{r} - \mathbf{s}|,$$
$$R^* = |\mathbf{r} - \mathbf{s}^*|.$$

A receiver at height 0.40 m is displaced horizontally 1.20 m from an impulse source at height 0.30 m. The direct path is $\sqrt{1.20^2 + 0.10^2} = 1.204$ m and the reflected path is

$$\sqrt{1.20^2 + 0.70^2} = 1.389\,\mathrm{m}.$$

The reflected pulse follows the direct one by 0.617 ns, with opposite sign and amplitude magnitude about 0.867 of the direct amplitude for this scalar geometrical model. A receiver that only rectifies the signal will lose this sign information. For two overlapping pulses their sign then still affects intensity, so rectification before adding the predicted pulses is incorrect.

For a scalar Neumann condition one adds the mirror instead. Although the two path lengths are identical at the boundary, their normal derivatives there are opposite, including the derivative of the retarded-time argument. Thus the sum has zero normal derivative. A rigid acoustic wall and a pressure-release surface furnish experimentally familiar examples of different scalar boundary rules, but the field variable must be named before assigning a reflection sign. Two surfaces can yield the same return time despite opposite returned amplitudes.

If a source is monochromatic, the two impulses become terms e^{ikR}/R and $-e^{ikR^*}/R^*$. Where the amplitudes are nearly equal, a path difference equal to one wavelength gives nearly complete cancellation, while a half wavelength can give reinforcement because the boundary reversal supplies the other half cycle. The image solution therefore connects the time-domain echo test to a spatial fringe prediction. Moving the receiver a few centimetres may change interference dramatically without changing the material reflectance.

Can two receiver traces determine a short source waveform?

Let $g(t)$ denote the waveform emitted by a small scalar source. After correcting a receiver's direct-path amplitude and arrival time, the ideal plane response above has the form $y(t) = g(t) - b\,g(t-\Delta)$. Both $b = R/R^*$ and $\Delta = (R^* - R)/c$ are known from geometry. The transfer function for this normalised signal has modulus squared $1 + b^2 - 2b\cos\omega\Delta$. Frequencies with $\omega\Delta = 2\pi n$ have poor signal. Merely subtracting an independently smooth fitted baseline does not recover their phase.

For the preceding geometry $b = 0.867$, so response amplitude ranges from 0.133 to 1.867 times the corrected direct-path amplitude. Unregularised division at a minimum magnifies amplitude noise by about 7.5. The spacing between minima is roughly 1.62 GHz. A second receiver at a different height changes b and Δ. Choose a geometry for which its minima avoid the first receiver's vulnerable frequencies across the band of interest. Two receivers at symmetric points with equal path lengths give more counts but no new spectral constraint.

Given complex measurements $Y_j(\omega) = H_j(\omega)G(\omega) + \eta_j$ with independent mean-square noise levels s_j^2, minimise $\sum_j |Y_j - H_jG|^2/s_j^2$ at each frequency. Differentiating in the real and imaginary parts of G gives

$$\widehat{G} = \frac{\sum_j H_j^* Y_j / s_j^2}{\sum_j |H_j|^2 / s_j^2}.$$

The reciprocal denominator sets the estimated error variance under the model. If both instruments have a common phase calibration error it does not average down like independent noise. If both transfer functions are small, probabilistic regularisation or a finite source-support constraint may stabilise a fit, but the inferred high-frequency waveform then depends on the extra assumptions. Present the bandwidth over which the data genuinely constrain the result.

A clean time-domain cross-check is available when the pulse is shorter than the known path separation. Window the first arrival before any reflected contribution enters and predict both later echoes using the measured first pulse. If the first pulse is not fully separated, the retarded equation can still be solved forward in time by $g(t) = y(t) + b\,g(t-\Delta)$, starting from a known zero prehistory. Its noise after many recurrences accumulates, particularly for b near one. A frequency fit and this delayed recursion use the same physics and have related conditioning. Agreement over the accurately measured band with disagreement confined to a spectral minimum identifies an inverse problem limitation before it is mistaken for a new source feature.

Recover the source phase at the correct emission time

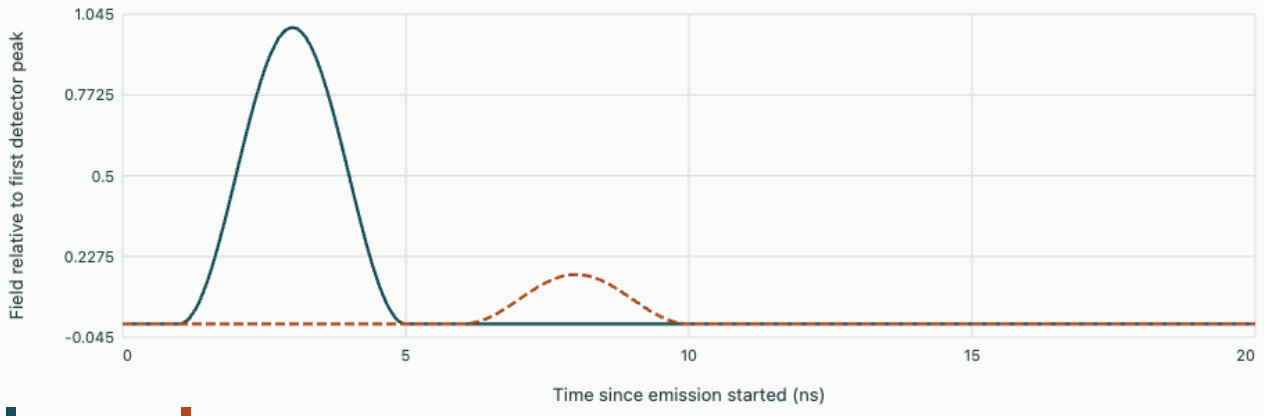
Move the observation point and compare its signal with the source evaluated at the observation time. Quantify the phase error of that shortcut.

Open this experiment online

Try the calculation

Second detector radius (m) 1.8

Pulse duration (ns) 4



At 0.30 m Second detector

The second trace starts 5 ns after the first. Its peak is 16.7% of the first peak.

Worked propagation problem · a cell-centred retarded-time approximation

A compact source is represented by cubic cells of edge $d = 2.0$ mm at frequency 5.0 GHz. Each cell is replaced by its centre value in the far-field sum. Estimate the error for a cell with uniform complex current when observed along a grid axis. Distinguish this amplitude error from uncertainty in arrival time.

▼ **Integrate the phase across a cell before summing**

Across one cell the leading far-zone phase is linear, e^{-ikx}, for observation along the local x axis. Averaging exactly across its width multiplies the centre approximation by $\sin(kd/2)/(kd/2)$. Here wavelength is approximately 60.0 mm and $kd/2 = 0.1048$, giving a factor 0.99817 or about 0.18 percent amplitude reduction. The two transverse integrals have no phase variation in this particular direction. An oblique look multiplies three sinc factors evaluated at the separate direction-cosine components. Spatially varying current or strongly curved wavefronts require additional corrections.

The travel time changes by $d/c = 6.67$ ps across the cube along the axis. The centre convention therefore leaves a maximum individual delay discrepancy of 3.34 ps. Its effect on the amplitude of a symmetric uniform cell begins at second order because the first-order phase terms cancel. That cancellation can disappear for a strongly varying current or a clipped boundary cell. Adding the same delay bound over every cell in the source would grossly overestimate a coherent error; conversely, relative error in a near-null total field can be very large even when each absolute contribution is accurate.

Refine spatial resolution while separately refining interpolation of the source history at non-grid retarded times. The former controls geometric phase and boundary representation, the latter controls the recovered temporal spectrum. Compare a signed field at a generic angle, integrated radiated power and

absolute field near an interference minimum. No single relative-error statistic is satisfactory for all three. Reporting a tiny percentage error for the total power can hide a phase error that materially changes the inverse source estimate at selected receivers.

How a finite measurement surface can test retardation

A closed surface of time-resolved field probes can test a proposed source history against local energy conservation. Integrate the measured outward Poynting vector over the surface at each receiver time. Between two concentric surfaces in free space with no charges, the difference between outward powers equals the rate at which field energy in the intervening shell changes, with the sign determined by whether the shell is filling or emptying. During a leading pulse edge the outward flux at the smaller sphere can exceed that at the larger one without any violation of conservation.

With a compact central source and far-zone probes, rescale radiation amplitudes by radius and shift their signals by the radial travel-time difference before imposing pointwise agreement. For a radial separation of thirty centimetres the shift is approximately one nanosecond. A modest timing mismatch can produce an apparent amplitude discrepancy on a steep edge even when both probes are correctly calibrated. Integrating through the complete pulse removes that sensitivity if the integration windows include its entire support.

Residual disagreement can identify omitted directional dependence, environmental reflections or nonradiative near-field terms. Repeat at increasing radii and compare their scaling. A coherent field proportional to inverse squared radius becomes relatively weaker compared with the outgoing radiation, while a reflected contribution may vary nonmonotonically as phase changes. A delayed free-space Green function makes distinct predictions for these different tests.

Tutorial · infer an unknown delay and test the amplitude law

A detector moved successively to 0.45, 0.90 and 1.35 m records field-peak times of 2.20, 3.70 and 5.20 ns and field amplitudes 6.0, 3.0 and 2.2 mV/m. Estimate speed and common timing offset. Which amplitude is least consistent with a single outgoing far field, and what additional observation could distinguish an additive background from a receiver gain change?

▼ **Fit the delay and audit the independent product**

Each 0.45 m step adds 1.50 ns, giving $v = 3.00 \times 10^{8}$ m/s. Extrapolation to r=0 gives 0.70 ns offset. The products rE are 2.70, 2.70 and 2.97 mV, with the last 10% higher than the common value. A stationary additive background field could interfere coherently with the local signal; a gain error would scale the waveform. A source-off record tests an additive background under the same trigger. Moving the detector back to the first radius or swapping two calibrated receivers tests whether the excess follows the position or the instrument. Arrival-time agreement alone cannot verify absolute amplitude calibration.

Tutorial · decide when one common retarded time suffices

A 60 mm source stands 3.0 m from a receiver near broadside. It radiates near 20 GHz in vacuum. Bound the relative amplitude variation from replacing $1/R$ by $1/r$. For a source lying in a transverse disk of radius 30 mm, estimate the largest omitted quadratic phase when evaluating a far-field integral. Would neglecting the entire directional phase be justified for observation 20° away from broadside?

▼ **Compare three different expansion parameters**

A conservative bound using the full size is order $0.060/3 = 0.020$; the actual broadside amplitude variation across a transverse disk is much smaller. At the rim the omitted optical path is approximately $0.030^2/(2 \cdot 3) = 0.00015$ m. Wavelength is about 15.0 mm, so the omitted phase is 0.063 rad. Off broadside along a diameter the retained linear path difference over 60 mm is $0.060 \sin 20^\circ = 0.0205$ m, or 8.60 rad. Constant amplitude and linearized phase can therefore be excellent approximations while treating all parts of the source as simultaneous at the receiver is seriously wrong.

David Tong, University of Cambridge · electromagnetic radiation, Green functions and retarded potentials

Waves, radiation and relativity · 7. Radiating multipoles and what a detector actually integrates

A small oscillating source looks point-like from a distant detector, yet its angular pattern reveals its structure. Derive electric and magnetic dipole radiation from a conserved current, identify what survives when electric dipoles cancel, and turn a detector aperture into a calibrated estimate of total power.

Name the small parameters before discarding a part of the field

Let d be the characteristic source size, r the source-to-detector distance, and $k = \omega/c$. The condition $d/r \ll 1$ permits geometric expansions of the radial amplitude. The condition $kr \gg 1$ selects the radiation field over induction and quasistatic fields. The condition $kd \ll 1$ permits a multipole expansion of the source's directional phase. These inequalities concern different pairs of lengths. A source can be electrically small while a probe sits in its near field. A large optical aperture can be many wavelengths wide while its far field is measurable at a sufficiently large r, though it then requires many multipoles or a full Fourier integral.

For harmonic localized sources in vacuum the outgoing far-zone vector potential is proportional to $\mathbf{I}(\mathbf{n}) = \int \widetilde{\mathbf{J}}(\mathbf{r}')e^{-ik\mathbf{n}\cdot\mathbf{r}'}d^3r'$. Differentiating the rapidly varying e^{ikr}/r factor gives the magnetic field; using the radiation plane-wave relationship gives the electric field. This removes the component of $\mathbf{I}$ parallel to the observation direction:

$$\widetilde{\mathbf{E}}_{\mathrm{rad}} = \frac{i\omega\mu_0 e^{ikr}}{4\pi r}\,[\mathbf{I} - \mathbf{n}(\mathbf{n}\cdot\mathbf{I})],$$

$$\widetilde{\mathbf{B}}_{\mathrm{rad}} = \frac{1}{c}\mathbf{n}\times\widetilde{\mathbf{E}}_{\mathrm{rad}}.$$

The transverse projection is essential. It forces zero electric-dipole radiation along the dipole axis. Treating the current integral as an already transverse scalar misses this geometrical zero and produces the wrong total power.

Eliminate an integrated current with charge conservation

The electric dipole moment is

$$\mathbf{p} = \int \mathbf{r}'\rho\, d^3r'.$$

Differentiate this definition and use $\dot\rho = -\nabla'\cdot\mathbf{J}$. For component j, integration by parts changes $-\int x'_j\partial'_l J_l$ into $\int J_j$, since the surface term vanishes outside a localized current. Thus $\int\widetilde{\mathbf{J}} = -i\omega\widetilde{\mathbf{p}}$. For $kd \ll 1$ the leading term of the phase expansion is one. Inserting it gives

$$\widetilde{\mathbf{E}}_p = \frac{k^2 e^{ikr}}{4\pi\epsilon_0 r}[\mathbf{p} - \mathbf{n}(\mathbf{n}\cdot\mathbf{p})],$$

$$\frac{d\overline{P}}{d\Omega} = \frac{\omega^4|\mathbf{n}\times\mathbf{p}|^2}{32\pi^2\epsilon_0 c^3}.$$

The phasor p is a peak moment. For a linearly oscillating real-axis dipole along z the intensity follows $\sin^2\theta$, independent of azimuth. Integrating $\sin^2\theta\sin\theta\,d\theta d\phi$ gives $8\pi/3$. Hence the cycle-averaged total is $\overline{P} = \omega^4|\mathbf{p}|^2/(12\pi\epsilon_0 c^3)$. A complex vector p describing an elliptical source is allowed in the vector formula. With equal orthogonal components in temporal quadrature, radiation along the normal to their plane is circularly polarized and there is no direction parallel to both components at once.

Consider an effective moment $p_0 = 1.0\times10^{-12}$ C m at 100 MHz. A vacuum far-field receiver at 10 m on the equator sees peak $E_0 = k^2 p_0/(4\pi\epsilon_0 r) = 3.95$ mV/m. Mean flux is $E_0^2/(2Z_0) = 20.7$ nW/m². Total power is $(8\pi/3)r^2 S_{\mathrm{eq}} = 17.4$ µW. At the same distance but a 30° polar angle the field peak halves and flux falls to a quarter. Increasing frequency while holding moment fixed raises power as the fourth power. Maintaining that same moment with a specified driven current introduces its own frequency dependence, so a frequency sweep must state the held source variable.

Recover the magnetic dipole from the antisymmetric current moment

Expand the phase as $1 - ikn_j x'_j + \cdots$. The first directional correction involves $M_{ij} = \int x'_j J_i\, d^3r'$. Its symmetric and antisymmetric parts have distinct origins. Define $\mathbf{m} = \frac{1}{2}\int \mathbf{r}'\times\mathbf{J}\,d^3r'$. Then the antisymmetric contraction with n gives a current contribution $ik\mathbf{n}\times\mathbf{m}$ for the index convention above. Substituting into the far-field electric expression yields

$$\widetilde{\mathbf{E}}_m = -\frac{Z_0 k^2 e^{ikr}}{4\pi r}\mathbf{n}\times\mathbf{m},$$

$$\overline{P}_m = \frac{\mu_0\omega^4|\mathbf{m}|^2}{12\pi c^3}.$$

A small planar current loop has $\mathbf{m} = I\,\mathbf{A}$ with oriented area vector. For a loop in the xy plane its magnetic-dipole intensity has the same $\sin^2\theta$ dependence as a z electric dipole, while its electric polarization is azimuthal instead of meridional. An intensity detector without polarization analysis can therefore confuse two fundamentally different sources. A rotating linear analyser placed at the equator can distinguish them even if each delivers the same incident power.

For a loop of radius 10 mm carrying a 10 mA peak sinusoidal current at 100 MHz, $m_0 = 3.14\times10^{-6}$ A m² and the equivalent effective electric-moment magnitude for equal power is $m_0/c = 1.05\times10^{-14}$ C m. Its power is about $17.4\,\mu\mathrm{W}(0.0105)^2 = 1.91$ nW. The current and geometry are assumed uniform around the loop. Larger loops require retaining retardation along the current path and a feed model before scaling this result.

Keep the symmetric part when the electric dipole vanishes

Define $D_{ij} = \int \rho x'_i x'_j\, d^3r'$. One integration by parts gives

$$\dot{D}_{ij} = \int (x'_i J_j + x'_j J_i) d^3r'.$$

Therefore $M_{(ij)} = -i\omega D_{ij}/2$ for harmonic sources. Split D into trace and traceless parts, defining $Q_{ij} = 3D_{ij} - \delta_{ij}\,\mathrm{tr}\,D$. The trace contribution after contraction is parallel to n and vanishes under transverse projection. Relative to the electric dipole field the effective transverse source is consequently

$$\widetilde{\mathbf{E}}_{p,Q} = \frac{k^2 e^{ikr}}{4\pi\epsilon_0 r}\,\mathsf{P}_\perp\left[\mathbf{p} - \frac{ik}{6}Q\mathbf{n}\right],$$

$$\mathsf{P}_\perp\mathbf{v} = \mathbf{v} - \mathbf{n}(\mathbf{n}\cdot\mathbf{v}).$$

This coefficient belongs to the stated quadrupole convention. Changing the factor three in Q changes the prefactor in the radiation field. Comparing numerical tensors from different sources therefore requires checking their definitions. Electric quadrupole and magnetic dipole fields both arise at the same first correction to a general localized current. One cannot always describe a source whose electric dipole vanishes as purely quadrupolar without examining the current's antisymmetric moment.

A transparent demonstration uses two opposite, equal z-directed electric dipoles separated along z by distance d. Their individual small-source patterns are proportional to $\sin\theta$. The pair's far-field phase sum is $-2i\sin(kd\cos\theta/2)$ up to an overall phase. When $kd \ll 1$ the field pattern is proportional to $\sin\theta\cos\theta$, with zeros on the axis and equator and maxima at 45° and 135°. Expansion of the exact array factor gives a controlled test of the approximation. At $kd = 0.4$, $\theta = 45^\circ$, the argument of the sine is 0.1414. Replacing the sine by its argument overestimates amplitude there by about $0.1414^2/6 = 0.33\%$. A small amplitude imbalance in the opposite dipoles can reintroduce an electric-dipole contribution even where the ideal quadrupole vanishes.

Change the symmetry of a dipole pair and test which angular nulls survive

Integrate over the detector aperture

A detector collecting directions within a cone normally measures an integral of $dP/d\Omega$ times throughput and polarization acceptance. For a z electric dipole and a complete azimuthal cone of half-angle θ_a centred on +z, the fraction is

$$f_{\mathrm{pole}} = \frac{\int_0^{\theta_a} \sin^3\theta\, d\theta}{\int_0^{\pi} \sin^3\theta\, d\theta} = \frac{2 - 3\cos\theta_a + \cos^3\theta_a}{4}.$$

For small θ_a expand the integrand rather than subtracting nearly equal cosines in floating point: $f_{\mathrm{pole}} \simeq 3\theta_a^4/16$. Although a true axial ray carries no power, a finite cone collects a small positive flux varying as the fourth power of its opening angle. An isotropic-emission approximation would instead scale as $\theta_a^2/4$, badly overestimating the near-null signal. An equatorial belt between $\theta = \pi/2 - \delta$ and $\pi/2 + \delta$, spanning all azimuths, captures fraction $(3\sin\delta - \sin^3\delta)/2$. For $\delta = 10^\circ$ this is 0.258, even though the belt covers only 0.174 of the full solid angle.

To convert a small planar detector into solid angle, include its projected area. For a detector of area $A_d = 1.0$ cm² normal to the line of sight at 2.00 m, $\Delta\Omega \simeq A_d/r^2 = 2.5 \times 10^{-5}$ sr. At an equatorial maximum the normalized dipole density is $3/(8\pi)$ per steradian, giving a collected fraction 2.98×10^{-6}. Rotating its face 60° to the line of sight halves this fraction before considering polarization-sensitive

throughput. An extrapolation from a small aperture to total emission is therefore a model-based estimate, whose validity can be assessed with readings at additional angles.

Separate geometrical zeros from coherence and background

Suppose equal geometry and detectors at polar angles 90°, 45° and 0° return powers 100, 50 and 5 in arbitrary units. An ideal point dipole with no offset would suggest 100, 50 and zero. Fitting a constant isotropic additive background using the axial reading leaves 95, 45 and zero. A pure dipole would predict 47.5 at 45° after subtraction. The discrepancy is small but informative if calibration uncertainty is smaller. Check finite aperture before interpreting the axial excess as a second emitter; make the aperture twice as narrow and see whether its captured fraction near the null falls by roughly sixteen, as for an ideal axial dipole cone.

If two excited source moments are coherent, add their complex fields before squaring. Their cross term can create a fore-aft asymmetry that disappears upon incoherent averaging. If their relative phase changes randomly over the integration window, add the angle-dependent intensities instead. A single angular sweep measures some combinations of magnitudes and relative phases and leaves others degenerate. Repeating with an analyser or a known phase reference helps distinguish a weak quadrupole contribution from an independent weak background. A total-power reading alone discards even more directional information.

Compare radiated power and angular dependence against a dipole calculation

Directionality from coherent electric and magnetic moments

A compact source can have two comparable leading moments of different type. This can happen if a resonant dielectric particle supports electric and magnetic responses at the same optical frequency. The long-wavelength multipole expansion still applies to the exterior observation, but symmetry or resonance can change the relative magnitudes of the moments. Add the electric fields before calculating power whenever the moments are coherently driven. Adding two separately calculated angular powers would erase a measurable directional effect.

Take an electric moment $p\mathbf{e}_x$ and a magnetic moment $m\mathbf{e}_y$, both complex peak amplitudes with the same time convention. Their leading radiation fields share the common factor $k^2 e^{ikr}/(4\pi\epsilon_0 r)$. The electric vector is $\mathbf{p} - \mathbf{n}(\mathbf{n}\cdot\mathbf{p})$. The magnetic contribution in these units is $-\mathbf{n}\times\mathbf{m}/c$. Thus in the $+z$ and $-z$ directions their electric polarisations coincide with x, but one magnetic contribution changes sign. After removal of the common amplitude their intensities are proportional to

$$I_+ = |p + m/c|^2,$$

$$I_- = |p - m/c|^2,$$

$$\frac{I_+ - I_-}{I_+ + I_-} = \frac{2\,\mathrm{Re}(p^* m/c)}{|p|^2 + |m/c|^2}.$$

If $m/c = p$, the intensity vanishes exactly in the backward direction. A quadrature phase between equal magnitudes removes the forward–backward difference, even though both moments radiate. A pair of

intensity measurements therefore constrains a real cross term but does not determine a unique pair of complex moments. Rotating analysers and adding another angular direction can supply additional independent constraints.

To understand the full angular pattern, sum the squared magnitudes of the two vector contributions for an arbitrary direction. The separate terms are $|p|^2(1-n_x^2)$ and $|m/c|^2(1-n_y^2)$; their cross term is $2n_z\,\mathrm{Re}(p^* m/c)$. For the equal in-phase case the sum reduces to $|p|^2(1+\cos\theta)^2$. Its integral over the sphere is $16\pi|p|^2/3$, twice the integral for one electric dipole alone at this amplitude. The cross term integrates to zero because it reverses sign between opposite hemispheres. Directional suppression redistributes existing radiation while leaving the sphere-integrated powers of these two orthogonal multipole types additive.

The mean angular intensity is proportional to $4|p|^2/3$, and the forward value to $4|p|^2$. The corresponding directivity is three. The backward hemisphere still contains radiation even when its axial intensity is exactly zero. Integrating $(1+u)^2$ from $u=-1$ to 0 and dividing by the integral from -1 to 1 gives a fraction $1/8$. A detector measuring integrated backward hemispherical power would report a clear finite signal. Two teams reporting an axial null and one eighth backward power are describing consistent observations if their angular acceptances differ.

The extinction of a null is exceptionally sensitive to field errors

Set $m/(cp) = (1+\epsilon)e^{i\varphi}$, with real small amplitude imbalance ϵ and phase error φ in radians. Expanding the backward intensity to the first nonzero order gives

$$I_-/|p|^2 \simeq \epsilon^2 + \varphi^2;$$

the forward intensity is about $4|p|^2$. Misalignment, background scattering and angular acceptance can also fill the nominal null. The ratio alone cannot identify an electric–magnetic amplitude mismatch unless phase and background are independently known.

A 5.0 percent amplitude error together with a five-degree phase error gives a leading estimate $I_-/I_+ \simeq 0.00253$. Using the exact formula gives about 0.00250. With 100,000 forward counts in matched exposures a perfectly calibrated backward detector would register about 250 signal counts. A background of 200 counts is comparable to this signal, and subtracting it adds uncertainty from the background exposure. Doubling the integration time does not remove an incorrectly modelled constant fraction of forward stray light reaching the backward detector.

Finite acceptance near an exact null has a different scaling. In the ideal balanced case write $\theta = \pi - \delta$. The intensity $(1-\cos\delta)^2$ grows as $\delta^4/4$. Integrated over a cone of half-angle α it scales as α^6, because $d\Omega$ contributes a further factor of δ. A small constant background integrated over the same cone instead scales as α^2. Measuring with several aperture stops can distinguish these trends only if the instrument's point-spread function and angular centring are known. For a simple electric dipole aligned with the viewing direction, the zero is quadratic in angle, giving yet another scaling.

A set of checks can distinguish a physical interference null from an insensitive receiver. Rotate the source by half a turn about the y axis so that its electric moment changes sign but its magnetic moment does not. The predicted dark direction should reverse. Reverse one measurement cable's reference phase in a

coherent receiver and verify that the reported phase changes while intensity does not. Use an independent isotropic calibration source or a known dipole orientation to calibrate gain in both directions. Finally scan frequency through the resonances that create the two moments. The null should follow the relative amplitude and phase inferred from independently measured spectral responses, not just one accidentally quiet detector channel.

Cancel the leading dipole contribution

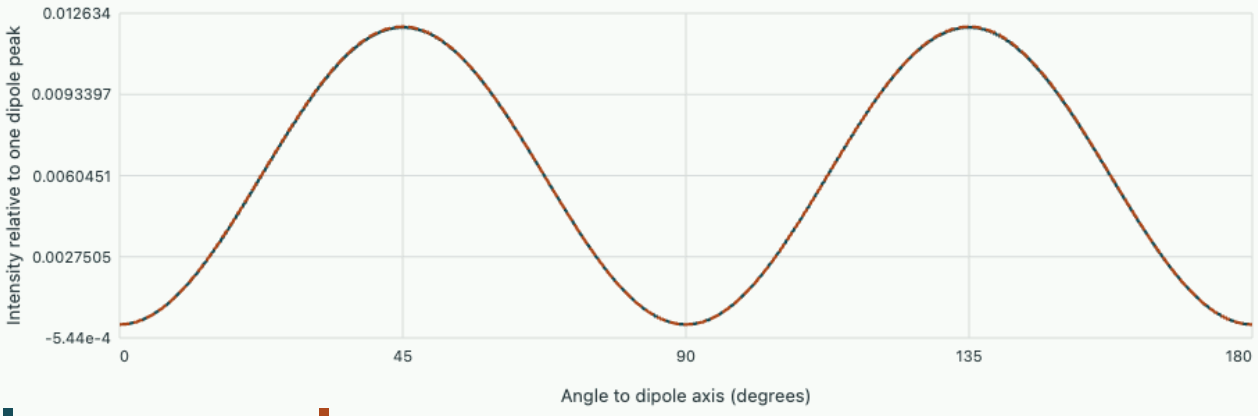
Change the separation or relative phase of the source elements and locate directions in which the leading radiation vanishes.

Open this experiment online

Try the calculation

Separation d/λ **0.035**

Opposing amplitude deficit **0**



Integrating over solid angle gives 0.9656% of the total power of one dipole. At 90° the intensity relative to its equatorial peak is 0.

Worked angular problem · fit an axis tilt from two detectors

A linearly oscillating electric dipole intended to lie along z is tilted by a small unknown angle ϵ toward $+x$. Matched detectors in the xz plane are placed at $+45^\circ$ and -45° relative to z. Their background-subtracted counts are 4200 and 5800 in equal exposures. Infer the tilt and its approximate counting uncertainty. Would reversing the electric dipole's oscillation phase change the answer?

▼ **Keep the two detector directions separate**

With positive angles toward $+x$, the expected intensities vary as $\sin^2(\pi/4 - \epsilon)$ and $\sin^2(\pi/4 + \epsilon)$. Their sum is independent of tilt, and the signed difference divided by their sum is $\sin(2\epsilon)$ when the difference is

defined as $I_- - I_+$. Its measured value is 0.160. The small solution is

$$\epsilon = \arcsin(0.160)/2 = 4.60^\circ.$$

Additional mathematical solutions are excluded by knowing the source is close to the desired initial axis.

Treat the counts as independent Poisson measurements with background uncertainties negligible for this calculation. Condition on their total of ten thousand. The normalised difference has estimated standard error $\sqrt{(1-0.160^2)/10000} = 0.00987$. Propagating through the inverse sine gives a tilt error of about 0.0050 radians or 0.29 degrees. A relative detector-gain error can imitate part of this difference. Swap the detectors physically while preserving source orientation or interchange the source tilt to separate a fixed gain bias.

Reversing the dipole phase multiplies the entire electric field by minus one, leaving both intensities unchanged. A phase-sensitive receiver with a common reference could detect the reversal. A finite aperture averages a range of angles, reducing the sensitivity if the apertures are wide. Integrate the angular predictions across the measured fields of view before applying the point-detector formula to a precision tilt.

A translated expansion can mix multipole labels

The exact far field is independent of where the expansion origin is drawn. Individual moments need not be. A translation of the source origin by a vector changes the electric dipole by the total source charge times that vector, with sign depending on the translation convention. For an electrically neutral compact source the dipole is therefore origin independent. A corresponding expansion of the quadrupole contains products of the translation with the dipole as well as the monopole. If both lower moments vanish, the first surviving electric quadrupole is likewise robust under a shift.

In a directional Fourier integral, changing the origin also changes the external phase that multiplies the expanded moments. At a fixed observation direction this common phase has no effect on intensity, but at a given truncation order its series must be kept consistently when comparing complex amplitudes. An origin-dependent angular intensity at the nominal retained order is a warning that a translated lower-order contribution or a same-order magnetic term may have been dropped.

A numerical check translates the integration grid by a known small vector while preserving the physical source and detector positions. Recompute both the unexpanded far-field integral and its truncated multipole prediction. The exact physical comparison should agree after phase referencing. The discrepancy between a leading-order expansion and the exact result should scale with the neglected power of source size over wavelength. This test probes moments, phase conventions and coordinate bookkeeping with one calculation.

Tutorial · infer a dipole moment and estimate detector uncertainty

A calibrated equatorial probe at 5.0 m observes a 300 MHz linearly oscillating dipole with RMS electric amplitude 2.00 mV/m. Determine peak moment and total emitted power. The receiver angle is uncertain by

±5° around the equator. Estimate the peak possible relative bias in a power estimate that assumes exact equatorial placement.

▼ **Keep the RMS calibration through the flux calculation**

Peak E is $\sqrt{2}(0.00200) = 0.002828$ V/m. With $k = 2\pi f/c \simeq 6.29$ m⁻¹, $p_0 = 4\pi\epsilon_0 r E_{\text{peak}}/k^2 \simeq 3.98 \times 10^{-14}$ C m. Alternatively use RMS E directly to find

$$S = E_{\text{rms}}^2/Z_0 = 1.06 \times 10^{-8}\,\text{W/m}^2.$$

The inferred total is $(8\pi/3)(25)S = 2.22$ μW. At a 5° displacement from the maximum the measured flux is multiplied by $\cos^2 5^\circ = 0.9924$, so the inferred total can be low by 0.76%. A comparable angular error near an axial null produces a much larger relative uncertainty.

Tutorial · use a symmetry-breaking parameter as a diagnostic

Two opposing parallel dipoles have moment magnitudes $p(1+\delta)$ and $p(1-\delta)$, with $kd = 0.20$. Derive the exact direction-dependent pair factor, estimate the equatorial-to-45° intensity ratio for $\delta = 0.010$, and explain why this experiment cannot determine the sign of a small real imbalance from total intensities alone in this geometry.

▼ **Retain the phase of each current element**

With $q = kd\cos\theta/2$, choose pair factor $(1+\delta)e^{-iq} - (1-\delta)e^{iq} = 2\delta\cos q - 2i\sin q$. The intensity pattern is proportional to $4\sin^2\theta(\delta^2\cos^2 q + \sin^2 q)$. At the equator q=0, so the intensity factor is

$$4\delta^2 = 0.0004.$$

At 45° q=0.07071 and the factor is about 0.01018. The ratio is 0.0393. The residual electric-dipole and leading quadrupole terms here occur in quadrature, eliminating their cross term for a real imbalance. Reversing its sign leaves each intensity unchanged. A phase-sensitive reference at the equator could recover that sign.

David Tong, University of Cambridge · radiated fields of localized currents and multipole radiation

Waves, radiation and relativity · 8. Scattering, extinction and the phase of the forward wave

A suspension of transparent particles dims a narrow transmitted beam while absorbing almost none of its energy. Derive the scattered angular power, identify the interference term seen in transmission, and use energy conservation to put a bound on the response of a single small particle.

Build the cross sections from a specified incident beam

Illuminate an isolated target in a uniform transparent host by a monochromatic plane wave of incident intensity I_0. A far-away receiver collecting solid angle $d\Omega$ measures scattered power dP_{sca}. Define $d\sigma_{\mathrm{sca}}/d\Omega = (dP_{\mathrm{sca}}/d\Omega)/I_0$ and integrate over direction to obtain σ_{sca}. Dividing internally dissipated power by the same incident intensity defines σ_{abs}. Extinction is $\sigma_{\mathrm{ext}} = \sigma_{\mathrm{sca}} + \sigma_{\mathrm{abs}}$ for a passive isolated object in a lossless surrounding medium. All three have dimensions of area, even for a particle that has no sharply defined physical footprint.

The geometrical area of an object and its cross section answer different questions. A transparent object much smaller than its wavelength may scatter only a small fraction of the power geometrically incident on its silhouette. A resonant dipole can draw energy from an effective area comparable to a wavelength squared, much larger than a nanoparticle's silhouette. The incident field and the disturbance it generates coexist in the forward direction. Removing a fixed area from a ray diagram cannot represent the interference between those fields.

Represent the incident polarization by a normalized complex vector $\mathbf{e}_0 \perp \widehat{\mathbf{z}}$. Define a vector scattering amplitude of dimensions length by $\widetilde{\mathbf{E}}_{\mathrm{sca}} = E_0 e^{ikr}\mathbf{f}(\mathbf{n})/r + O(r^{-2})$ for the same $e^{-i\omega t}$ convention. The scattered wave is transverse to n; for a polarization-insensitive receiver $d\sigma/d\Omega = |\mathbf{f}|^2$. With a receiving analyser $\mathbf{a}$ only $|\mathbf{a}^{\dagger}\mathbf{f}|^2$ is detected. Thus an intensity pattern alone loses information about phase and possibly polarization. The forward extinction will depend on one particular quadrature of that phase.

Derive the small sphere's polarizability from two boundary equations

In vacuum take a dielectric sphere of radius a with scalar relative permittivity ϵ_r at the driving frequency. For $ka \ll 1$ and negligible retardation inside the sphere, solve the instantaneous-looking boundary problem for a phasor incident field $E_0\hat{z}$. Axial symmetry and the degree-one angular dependence admit $\Phi_{\mathrm{in}} = -Cr\cos\theta$ and $\Phi_{\mathrm{out}} = -E_0 r\cos\theta + D\cos\theta/r^2$. At r=a continuity of tangential E is equivalently continuity of these potentials; continuity of normal D supplies the other equation. They give

$$C = E_0 - D/a^3,$$

$$\epsilon_r C = E_0 + 2D/a^3,$$

$$D = a^3\frac{\epsilon_r - 1}{\epsilon_r + 2}E_0.$$

Comparing the decaying term with a dipole potential gives

$$p = 4\pi\epsilon_0 D.$$

Define the volume polarizability by $p = \epsilon_0 \alpha E_0$, so $\alpha_0 = 4\pi a^3(\epsilon_r - 1)/(\epsilon_r + 2)$ in m³. Some conventions absorb ϵ_0 in α. Naming the convention avoids spurious permittivity factors when scattering formulas are converted between sources. At $\epsilon_r \to 1$ the particle becomes indistinguishable from its environment and the dipole vanishes. In the formal electrostatic conducting limit the coefficient tends toward $4\pi a^3$.

Radiation from this induced electric dipole gives

$$\mathbf{f}(\mathbf{n}) = \frac{k^2 \alpha}{4\pi} [\mathbf{e}_0 - \mathbf{n}(\mathbf{n} \cdot \mathbf{e}_0)],$$

$$\sigma_{\mathrm{sca}} = \frac{k^4}{6\pi} |\alpha|^2.$$

The projection shows why scattered light from unpolarized illumination becomes partially polarized at an angle: average the power of two orthogonal transverse incident polarizations. For observation angle θ to the incident wavevector, the normalized angular factor is $(1 + \cos^2\theta)/2$. At 90° an analyser can suppress the surviving polarization. The total cross section is independent of the orientation of a single linear polarization because the isotropic sphere can be rotated with that polarization.

Predict an absolute scale as well as a wavelength slope

For a transparent sphere $a = 20$ nm, index 1.50 so $\epsilon_r = 2.25$, and $\lambda_0 = 600$ nm, the contrast factor is 0.2941 and $ka = 0.2094$. The quasistatic volume polarizability is 2.96×10^{-23} m³. Substitution gives $\sigma_{\mathrm{sca}} \simeq 5.6 \times 10^{-19}$ m² compared with geometric area 1.26×10^{-15} m². At otherwise fixed properties, halving radius divides $|p|$ by eight and σ_{sca} by 64. Changing wavelength from 600 to 500 nm instead increases the leading cross section by $(600/500)^4 = 2.074$. This scaling assumes that material dispersion and finite-size corrections have stayed sufficiently small.

For a dilute suspension with $N = 10^{16}$ particles/m³ path loss from independent scattering gives

$$dI/dz = -N\sigma_{\mathrm{ext}} I.$$

Over 0.10 m at the 600 nm parameters, optical depth is about 5.6×10^{-4}, or 0.056% collimated intensity loss. Collecting a wider transmitted cone can admit some of the scattered light and reduce that apparent extinction. At high density fields scattered by one particle drive others and positional correlations can alter angular scattering. The isolated-particle prediction and exponential independent-particle extrapolation should be tested separately.

Change particle radius and compare a calculated scattering cross section

Find an interference signal that survives at arbitrarily large distance

Write the total field outside the scatterer as incident plane wave plus outgoing spherical wave. The Poynting flux through a large sphere contains incident-only, scattered-only and cross terms. The net incident-only flux through a closed sphere vanishes. The scattered-only term tends to $I_0\sigma_{\mathrm{sca}}$. The interference term accounts for the reduction of incident energy in the forward beam. To see its scaling, examine a narrow forward cap; in a scalar polarization-preserving calculation the relevant angular

oscillation is $e^{ikr(1-\cos\theta)}$. The interference field is order $1/r$, but the integration area is order r^2; oscillatory phases cancel over most directions and leave a forward contribution of finite order.

Near the cap $\theta \ll 1$, $1-\cos\theta \simeq \theta^2/2$, $d\Omega \simeq 2\pi\theta d\theta$. Introduce a smooth angular acceptance equal to one near zero. The leading forward factor in the integral is

$$2\pi \int_0^\infty \theta e^{ikr\theta^2/2} d\theta = 2\pi i/(kr),$$

with arbitrarily weak positive damping understood before taking the large-radius limit. Magnetic and electric interference terms have the same forward polarization and supply twice the real cross product in total flux. A backward stationary contribution is cancelled by the opposing propagation directions in the net flux. With the outgoing-wave convention this gives

$$\sigma_{\mathrm{ext}} = \frac{4\pi}{k} \Im[\mathbf{e}_0^* \cdot \mathbf{f}(\hat{\mathbf{z}})].$$

This is the optical theorem. A forward detector sensitive only to $|\mathbf{f}(0)|^2$ is insufficient; the unscattered incident field acts as a coherent reference that picks out the appropriate quadrature. Changing the scattering-amplitude convention changes corresponding signs. Inserting a known passive dipole is a quick way to verify the convention used in a computational implementation.

There is an equivalent small-particle derivation. The cycle-averaged work extracted from a prescribed incident electric field is

$$\langle \mathbf{E}_{\mathrm{inc}} \cdot \dot{\mathbf{p}} \rangle = (\omega/2)\,\mathrm{Im}(\widetilde{\mathbf{E}}_{\mathrm{inc}}^* \cdot \widetilde{\mathbf{p}}).$$

Dividing by $I_0 = |E_0|^2/(2Z_0)$ gives

$$\sigma_{\mathrm{ext}} = k\,\mathrm{Im}\,\alpha.$$

In the forward direction the dipole formula gives $\mathbf{e}_0^* \cdot \mathbf{f}(0) = k^2\alpha/(4\pi)$, reproducing that result exactly. Comparing integrated scattered power with interference-derived extinction tests both amplitude normalization and global energy accounting.

Repair the apparent contradiction of a real static polarizability

A strictly real static α_0 gives positive dipole radiation in σ_{sca}, yet inserting that same real number into $k\Im\alpha$ would give zero extinction. The two formulas then mix orders of approximation. Replacing the driving field by its electrostatic value determined the leading in-phase moment. The small imaginary part needed to supply radiated energy arises at the next electrodynamic order. For a lossless purely dipolar particle energy conservation requires $\sigma_{\mathrm{ext}} = \sigma_{\mathrm{sca}}$, hence

$$\Im\alpha = \frac{k^3}{6\pi}|\alpha|^2,$$

$$\Im(1/\alpha) = -\frac{k^3}{6\pi}.$$

Retaining the static inverse real part yields a simple radiation correction $\alpha = [\alpha_0^{-1} - ik^3/(6\pi)]^{-1}$. For a real α_0 this expression obeys the lossless optical theorem. It captures the radiative imaginary correction while ignoring possible real dynamic depolarization and higher multipoles; it is not an exact all-size sphere solution. If the material itself absorbs, a passive α_0 contributes additional loss, and $\sigma_{\mathrm{abs}} = k\Im\alpha - k^4|\alpha|^2/(6\pi)$ must remain nonnegative.

Define $b = k^3\alpha/(6\pi)$. Passivity is $\Im b \geq |b|^2$, which is the disk $(\Re b)^2 + (\Im b - 1/2)^2 \leq 1/4$ in a complex-amplitude diagram. Points on the boundary are purely radiative dipoles; interior points have additional internal absorption. The largest possible $|b|$ is one at $b = i$, giving $\sigma_{\mathrm{sca}} = 6\pi/k^2 = 3\lambda^2/(2\pi)$. At that point the particle is lossless and extinction is all scattering. The largest absorption occurs instead at $b = i/2$. There $k\Im\alpha = 3\pi/k^2$, scattering is $1.5\pi/k^2$, and the difference $1.5\pi/k^2$ is absorbed. The result applies to one ideal electric dipole channel with the stated normalization. Additional channels can contribute separately for a larger or more complex object.

Choose the detector geometry that answers the desired question

A narrow-aperture transmission experiment measures an extinction-like signal if it isolates the original mode. A large integrating sphere estimates total outgoing scattered flux over its acceptance and provides an independent σ_{sca}; a temperature-rise measurement estimates heat deposition. For the same incident beam, the combination should satisfy

$$P_{\mathrm{ext}} = P_{\mathrm{sca}} + P_{\mathrm{abs}}.$$

Measuring extinction and assuming it is all absorption can exaggerate heating by orders of magnitude for transparent dielectric particles. Conversely reflection plus a small-angle detector may miss side-scattered light and turn that deficit into an apparent unexplained source term.

Suppose a calibrated intensity is 2.0×10^4 W/m², transmission implies 60 pW extracted per particle, and integrated scattering gives 45 pW. The inferred cross sections are 3.0×10^{-15} and 2.25×10^{-15} m², leaving 0.75×10^{-15} m² for absorption and 15 pW thermal load. If each of the two power readings has independent 5 pW standard uncertainty, the difference has about 7.1 pW uncertainty. The inferred absorption is then much less certain in relative terms than either directly measured quantity. A negative fitted difference within that uncertainty should prompt an examination of collection angles and background subtraction rather than an immediate interpretation as optical gain.

How energy conservation fixes a scattering phase

The forward optical theorem can appear surprising because it relates a single angular amplitude to a sphere-integrated power. For a spherical scatterer this relationship can be seen one angular momentum

channel at a time. Begin with a scalar plane wave scattering elastically from a spherically symmetric object in a lossless exterior. This scalar example exposes the mathematics; electromagnetic spherical waves have separate electric and magnetic polarisation families and start at angular momentum one. A scalar zeroth channel should never be inserted directly into a vector light-scattering calculation.

Outside a sphere containing the object, the radial equation for each Legendre component has independent incoming and outgoing spherical solutions. Define S_ℓ as the ratio of outgoing to incoming wave amplitudes relative to its value for free propagation. The radial energy flux in a channel is proportional to the squared magnitude of the outgoing amplitude minus that of the incoming amplitude. For an elastic scatterer flux balance demands $|S_\ell| = 1$. Set $S_\ell = e^{2i\delta_\ell}$ with real phase shift δ_ℓ. The scattered outgoing wave is the change from the unscattered outgoing part, proportional to $S_\ell - 1$.

$$f(\theta) = \frac{1}{k}\sum_{\ell=0}^{\infty} (2\ell+1) e^{i\delta_\ell}\sin\delta_\ell\, P_\ell(\cos\theta),$$

$$\sigma_{\mathrm{sca}} = \frac{4\pi}{k^2}\sum_{\ell} (2\ell+1)\sin^2\delta_\ell.$$

The second result follows by integrating $|f|^2$ and using $\int_{-1}^{1} P_\ell P_m\, du = 2\delta_{\ell m}/(2\ell+1)$. At forward angle all Legendre functions equal one. The imaginary part of $e^{i\delta}\sin\delta$ is $\sin^2\delta$, exactly the quantity appearing in the integrated cross section. This equality in each uncoupled channel proves the elastic optical theorem in the scalar example. It does not require the forward intensity $|f(0)|^2$ to have any particular value. The calibrated phase of the forward amplitude is essential.

For a hypothetical acoustic sphere at wavenumber $k = 20$ per metre, suppose only the scalar zeroth phase shift is appreciable and has $\delta_0 = \pi/6$. Its scattering amplitude has real part 0.0217 m and imaginary part 0.0125 m, independent of viewing angle. Integrating its differential cross section gives $4\pi(0.025\,\mathrm{m})^2 = 0.00785$ square metres. The forward identity gives $4\pi(0.0125)/20 = 0.00785$ square metres as well. Shifting the phase through $\pi/2$ at this wavenumber increases the cross section by four, to the single-channel elastic maximum.

What an incomplete outgoing channel tells the experimenter

For an absorbing or internally converting target write $S_\ell = \eta_\ell e^{2i\delta_\ell}$, $0 \le \eta_\ell \le 1$. The missing outgoing flux in that exterior channel contributes to absorption or to unobserved channels, depending on the physical system. Define $a_\ell = (S_\ell - 1)/(2i)$, so the scalar wave amplitude above uses $a_\ell = e^{i\delta_\ell}\sin\delta_\ell$ in the purely elastic limit. Elementary algebra gives

$$\operatorname{Im} a_\ell - |a_\ell|^2 = \frac{1 - |S_\ell|^2}{4},$$

$$\sigma_{\mathrm{abs}} = \frac{\pi}{k^2}\sum_{\ell} (2\ell+1)(1-\eta_\ell^2).$$

For a single channel with $\eta = 0$, the scattered amplitude is $i/2$. The residual coherent outgoing component relative to its incident plane-wave background corresponds to a finite scattered power even though this channel has zero total outgoing radial amplitude. Its absorbed and scattered cross sections

are equal, while extinction is their sum. This distinction is the angular-channel counterpart of interference between incident and scattered fields in the forward direction.

An experiment may infer $\eta < 1$ because the detector fails to collect another outgoing polarisation, a converted elastic mode or a large-angle part of the pattern. To assign true heat absorption one needs the full exterior channel accounting or a calorimetric measurement. For a homogeneous liquid suspension distinguish energy dissipated in the nominal particles from energy dissipated in the surrounding medium. The preceding cross-section definitions assume a nonabsorbing reference host, so a lossy host requires an explicit incident-reference subtraction and an agreed control surface.

Suppose a measurement models the target only as an isotropic scalar scatterer, yet detects a pronounced forward lobe. A one-channel fit can reproduce a forward complex amplitude by choosing an appropriate complex a_0, but it also predicts the same intensity at every other angle. Add a possible $\ell = 1$ component and measure at $\theta = 0, \pi/2, \pi$. At right angles the term proportional to P_1 vanishes, which helps isolate the $\ell = 0$ component. Sum and difference of complex forward and backward amplitudes separate even and odd parity. Intensity-only data lose signs and relative phase and may require additional incident conditions or interferometry.

To check a numerical partial-wave calculation, increase the highest retained order at fixed geometry and compare the integrated angular quadrature with the coefficient sum. These diagnostics test the angular grid and truncation separately. A narrow forward peak may contribute little at any grid point if no nodes resolve its centre, yet contribute substantially to extinction. Gauss–Legendre quadrature in $u = \cos\theta$ automatically includes the solid-angle measure. Finally compare the coefficient-sum extinction against the imaginary forward value with the chosen time convention. A sign error in outgoing radial waves can produce a beautifully converged negative extinction for a passive sample.

Track the frequency dependence of a scattering signal

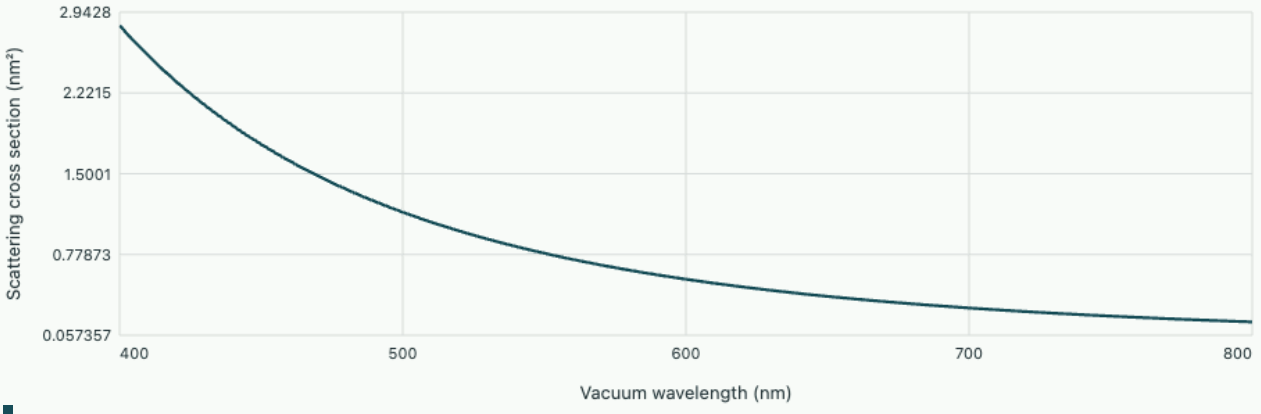
Change the ratio of forcing to resonance frequency. Decide which part of the response determines a scattering measurement and compare it with the power balance in the text.

Open this experiment online

Try the calculation

Sphere radius (nm) **20**

Relative permittivity **2.25**



Rayleigh sphere

The leading cross section at 600 nm is 0.558 nm². The shortest plotted wavelength gives ka = 0.314; finite-size corrections grow as this ratio grows.

Worked suspension problem · infer and check an optical depth

A dilute suspension has independently measured single-particle extinction cross section 5.0×10^{-15} square metres and particle density 3.0×10^{14} per cubic metre. A narrow stop rejects the scattered halo. Predict the direct transmission through 20 mm and estimate how many particles occupy a cylindrical illuminated volume of radius 0.10 mm. Which approximation is most vulnerable if the radius is increased?

▼ **Derive the transport law and identify the role of the aperture**

In an infinitesimal length dz, the direct-beam intensity loses a fraction $n\sigma_{\mathrm{ext}}dz$ in the independent, uniform, dilute limit. Integrating gives

$$I/I_0 = e^{-n\sigma_{\mathrm{ext}}L}.$$

The extinction coefficient is 1.50 per metre and optical depth 0.0300, implying direct transmission 0.9704. A first-order calculation gives 0.9700. The small difference is consistent with neglect of higher powers in optical depth, but the exponential by itself does not describe multiple scattering into the receiver.

The illuminated volume is $\pi(10^{-4})^2(0.020) = 6.28 \times 10^{-10}$ cubic metres and contains about 188,000 particles on average. Increasing beam radius at fixed density and path length raises this count but leaves ideal optical depth unchanged. Counting more independent scatterers reduces relative sample fluctuations. It can, however, change beam divergence, spatial illumination and what scattered angles reach a finite detector. There is no general reason for local diluteness to fail merely because the beam area is larger.

Interchange cells of several lengths and measure the slope of the logarithm of transmitted intensity. An intercept different from zero suggests interface loss or a calibration offset if both incident and transmitted powers are referenced outside the cell. Repeat the length scan with a wider angular stop. A reduced slope with the wide stop signals that the detector collects part of the scattering halo; a genuine absorption contribution persists. At higher densities, evaluate interactions between particles or coherent structure before extrapolating the dilute-solution cross section from the same attenuation law.

A scattering measurement weights the large end of a distribution

The small-particle scattering law grows as the sixth power of radius for fixed material contrast and wavelength. A suspension with a modest spread of radii can therefore scatter substantially more than a calculation using the mean radius for every particle. If radii are independently distributed with mean a_0 and small fractional standard deviation s, expand $a^6 = a_0^6(1+\delta)^6$. The first correction to the ensemble mean is $15s^2$ times the monodisperse value when the centred deviations average to zero. At ten percent relative spread, the leading increase is fifteen percent.

Separately measuring total particle material changes the comparison. Material volume scales as the third power, whose corresponding correction is $3s^2$. At fixed material volume, the scattering per unit volume therefore increases by approximately $12s^2$ relative to spheres of mean radius, within the same narrow-distribution approximation. Reporting the quantity held fixed changes the numerical prediction meaningfully.

A few larger outliers can invalidate this expansion and even take the sample beyond the Rayleigh regime. Measure a size distribution independently and average the single-particle cross section over that distribution, including detector acceptance if it varies with size. A fit to a single apparent radius may match total intensity at one wavelength while failing its spectral slope or angular pattern. Comparing two wavelengths and a separate volume estimate gives a practical check on that hidden weighting.

Tutorial · examine two proposed dipole fits for passivity

At a fixed wavelength two fits to an angular intensity pattern return $b_1 = 0.30 + 0.10i$ and $b_2 = 0.30 - 0.10i$, with $b = k^3\alpha/(6\pi)$. Can angular intensity distinguish them? Find normalized scattering, extinction and absorption by dividing each cross section by $6\pi/k^2$. Which fit can describe a passive particle?

▼ **Compare interference with scattered intensity**

Both have $|b|^2 = 0.10$, giving the same angular scattering and normalized scattered cross section 0.10. Extinction in these units is $\Im b$, hence +0.10 or −0.10. Absorption is the difference, zero or −0.20. The first sits on the lossless passivity circle. The second requires an active energy supply if described by these conventions. Even an angular pattern measured over the whole sphere would retain that ambiguity because it contains squared magnitudes. The forward interference phase or an independent extinction reading discriminates.

Tutorial · recover a particle radius with a stated systematic uncertainty

At fixed wavelength and index a dilute sample scatters 1.50 times the reference signal per unit volume. Assume initially the same number density and a Rayleigh regime. Infer the radius ratio. Then allow number density to be 20% higher than assumed. Estimate how much of the signal difference could instead be a concentration effect, and propose a phase-sensitive check of the dipole interpretation.

▼ **Track the powers of radius and concentration separately**

Scattered power is proportional to Na^6 under these assumptions, so the radius ratio is

$$1.50^{1/6} = 1.070.$$

If number density is actually 1.20 times the reference, the ratio becomes

$$(1.50/1.20)^{1/6} = 1.038.$$

An apparently precise intensity can therefore constrain a combined quantity quite well without determining either factor separately. Measure an independent particle count or concentration. A forward extinction measurement should satisfy the optical theorem, while the real part of the dilute coherent refractive response scales mainly with $N\Re\alpha \sim Na^3$. Combining these observables is better conditioned than repeating a single intensity measurement.

Barton Zwiebach, MIT · optical theorem and the forward scattering amplitude

David Tong, University of Cambridge · scattering, induced dipoles and electromagnetic radiation

Waves, radiation and relativity · 9. Diffraction from the near field to the focal plane

An illuminated aperture is a boundary condition for a wave. A screen a millimetre away, an intensity camera at a lens focus and a microscope measuring an evanescent field sample different consequences of the same condition. We will derive operators for propagating the field, estimate the approximations each instrument makes and use energy and sampling checks to tell numerical structure from a measurable fringe.

How each spatial frequency propagates

Take a monochromatic scalar field with physical value given by the real part of $U(x,y,z)e^{-i\omega t}$. This approximates one transverse component when the polarisation remains nearly uniform and angles are small. In a homogeneous transparent half-space it obeys the Helmholtz equation. Specify the complex amplitude in the plane $z=0$ and require outgoing or decaying waves for positive z. Fourier transform the two transverse coordinates, leaving a differential equation in the third.

$$\partial_z^2 \widehat{U} + (k^2 - k_x^2 - k_y^2)\widehat{U} = 0,$$

$$\widehat{U}(k_x,k_y,z) = \widehat{U}(k_x,k_y,0)e^{ik_z z},$$

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2}.$$

Choose the root with positive real part for propagation and positive imaginary part for decay. Each propagating component advances in phase. Hence power integrated over a plane is unchanged for a paraxial beam with no absorbing aperture downstream. For angles approaching grazing incidence the weight for normal flux contains the direction cosine. Components with transverse wavenumber exceeding k decay exponentially. An aperture map thus contains structures invisible to a distant measurement even if the map itself is known exactly.

In vacuum a sinusoidal field structure of period 400 nm illuminated at 600 nm has $k_x/k = 1.50$. Its amplitude has fallen to $1/e$ after $1/(k\sqrt{1.5^2-1}) = 85.4$ nm. At 500 nm the remaining amplitude is 0.00287 and the squared amplitude is roughly eight parts in a million of its plane value. A numerical back-propagation that reinstates the original component also multiplies measurement errors by about 350. Improving detector gain does not restore the unknown phase.

A propagation kernel with a controlled phase error

For small transverse wavenumber, expand $k_z = k - k_\perp^2/(2k) - k_\perp^4/(8k^3) + \cdots$. The first correction becomes a chirp in Fourier space. Completing the square in its inverse transform gives the Fresnel kernel. Write $\mathbf{r}$ for a transverse observation point and $\boldsymbol{\xi}$ for a point in the aperture.

$$U(\mathbf{r},z) \simeq \frac{e^{ikz}}{i\lambda z}\int U(\boldsymbol{\xi},0)\exp\left[\frac{ik}{2z}|\mathbf{r}-\boldsymbol{\xi}|^2\right]d^2\xi.$$

This expression retains phase curvature across the aperture. The coefficient ensures that a spatially uniform boundary amplitude reproduces a uniform plane wave. An uncovered infinite plane is a useful test

for a numerical convolution, although a finite computation needs padding or a truly periodic uniform test. A measured slit has sharp edges with broad spatial spectra, so peak error near an edge can exceed the error suggested by a small angle evaluated for the central lobe.

At an observation angle θ a propagation distance z accumulates neglected phase of order $kz\sin^4\theta/8$. For $z = 0.10$ m, $\lambda = 500$ nm and $\theta = 0.01$ rad this is 0.00157 rad. At 0.05 rad it is close to one radian. A faint sidelobe at a moderately large angle can require exact angular-spectrum propagation even while a paraxial estimate describes the central spot beautifully.

Which quadratic term can the far field lose?

Expand the kernel's square into an observation phase, a source phase and a bilinear term. Only the source phase can be dropped without changing the predicted intensity directly through that factor. For source radius a its maximum value is $\pi a^2/(\lambda z)$. The remaining integral is the Fourier transform of the aperture at spatial frequencies $x/(\lambda z), y/(\lambda z)$. Define the Fresnel number using the half-width a. A convention based on full slit width gives a value four times larger.

$$U(\mathbf{r}, z) \simeq \frac{e^{ikz + ikr^2/(2z)}}{i\lambda z}\, \widehat{U}\!\left(\frac{kx}{z}, \frac{ky}{z}, 0\right),$$

$$N_F = \frac{a^2}{\lambda z}.$$

For a uniformly illuminated slit of full width $w = 0.20$ mm at 532 nm and $z = 1.0$ m, $N_F = 0.0188$. Direct integration of the linear aperture phase gives

$$I/I_0 = \operatorname{sinc}^2(\pi w x/\lambda z),$$

where $\operatorname{sinc}(u) = \sin(u)/u$. The first zeros should be 2.66 mm either side of centre. The edge phase is 0.059 rad. Moving the screen to 30 mm gives $N_F = 0.627$ and edge phase 1.97 rad. Substituting the smaller distance in the far-field formula yields a number but no longer a controlled approximation.

Compare near- and far-field predictions in the slit integral

In that experiment reduce the mesh interval before moving the screen. If the exposed variable counts panels, doubling panels at fixed aperture halves the interval. A converged difference between the full and far-field integrals is a physical failure of the approximation. A feature that moves each time the number of source samples doubles is a numerical failure. Keep the same physical horizontal axis and the same normalisation when comparing adjacent screens.

From a circular pupil to a resolution measurement

When the pupil is a disk, polar coordinates considerably simplify the Fourier integral. Align the transverse transform vector along the horizontal axis. The angular average of $e^{-iqr\cos\phi}$ defines $J_0(qr)$. One radial integration, using $d[uJ_1(u)]/du = uJ_0(u)$, then gives the result. Divide by the on-axis amplitude, proportional to aperture area.

$$\frac{U(q)}{U(0)} = \frac{2}{a^2}\int_0^a r J_0(qr)\,dr = \frac{2J_1(qa)}{qa},$$

$$q = k\sin\theta.$$

At $qa = 0$ use $J_1(s) \sim s/2$. The first zero is at $qa = 3.8317$. Hence a disk of diameter D has its first dark ring at $\sin\theta = 1.22\lambda/D$. Two mutually incoherent identical points contribute the sum of their intensities. Rayleigh's conventional separation places each peak on the first zero of its neighbour. A coherent pair instead has an interference term that depends on the unknown relative phase. The same separation criterion cannot predict its measured trough without that phase.

A 5.0 mm pupil on a 100 mm focal-length lens at 550 nm gives a first-zero radius of 13.4 micrometres on the focal plane. A camera with 6.5 micrometre pixels spans the full bright core with about four pixels. Fitting the central pixel alone to find the image position would depend strongly on whether the spot lay on a pixel centre or a corner. A joint fit to integrated pixel intensities can estimate a position more precisely than the image width when enough photons are detected. That precision describes one isolated spot and does not directly give a two-source resolution.

Why a lens puts the spectrum close at hand

A thin positive lens adds the quadratic phase $-kr^2/(2f)$ to a field immediately in front of it. At the plane $z = f$, this cancels the source term of the Fresnel kernel exactly within paraxial theory. The focal intensity therefore samples a Fourier spectrum even for a tabletop lens standing much nearer than the free-space far-field distance. The physical exit pupil remains part of the amplitude. Closing it removes high transverse spatial frequencies in an imaging arrangement and broadens the spot of a point emitter.

For illumination with Gaussian amplitude $U \propto e^{-r^2/w^2}$, completing the transverse Fourier integrals gives another Gaussian, with focal-plane amplitude radius $w_0 = \lambda f/(\pi w)$ for a broad lens and an incident flat wavefront. If $w = 1.0$ mm, $f = 100$ mm and $\lambda = 633$ nm, $w_0 = 20.1$ micrometres. This is the $1/e$ field radius, also the $1/e^2$ intensity radius. The Rayleigh length $\pi w_0^2/\lambda$ is 2.02 mm. Moving a detector one Rayleigh length off the waist increases the radius by $\sqrt{2}$ and halves its on-axis intensity for fixed total beam power.

Use the ray-bundle experiment to locate the geometric focus before estimating its diffractive width

From a point image to transmitted object contrast

A microscope may resolve individual test objects yet deliver very little contrast for a periodic specimen at a similar length scale. The reason is that incoherent imaging acts linearly on intensity, while coherent imaging acts linearly on field amplitude. Start with mutually incoherent object points. Their expected image intensities add, so the image is the convolution of object intensity with a nonnegative intensity point-spread function $h = |h_A|^2$, where h_A is the coherent amplitude impulse response. The Fourier transform of h, normalised to one at zero spatial frequency, is the optical transfer function. Its magnitude describes attenuation of a sinusoidal intensity modulation.

For a focused unaberrated system the coherent amplitude response in Fourier space is the pupil function. The convolution theorem applied to $|h_A|^2$ shows that the incoherent transfer function is the pupil autocorrelation. For a clear circular pupil this amounts to asking how much two equal disks overlap as their centres separate. The overlap consists of two circular sectors minus their triangular parts. With normalised frequency $s = f/f_c$ and incoherent cutoff $f_c = 2\,\mathrm{NA}/\lambda_0$ in object space, the result is

$$M(s) = \frac{2}{\pi}[\arccos s - s\sqrt{1-s^2}], \quad 0 \le s \le 1,$$

$$M(s) = 0 \quad (s > 1).$$

Here λ_0 is vacuum wavelength and $\mathrm{NA} = n \sin\theta_{\max}$. At $s = 0$ the disks coincide and the full mean intensity passes. At the cutoff their edges just touch and sinusoidal contrast vanishes. The coherent pupil transmits field spatial frequencies only through NA/λ_0. These two cutoffs answer different questions about different object and illumination statistics. A coherent phase grating viewed in exact focus can yield an almost uniform intensity image despite transmitting several field components.

For $\mathrm{NA} = 0.25$ at 550 nm, the incoherent cutoff is 909 cycles/mm. A sinusoidal target of 450 cycles/mm has $s = 0.495$ and $M \simeq 0.396$. An object with 0.60 sinusoidal intensity visibility is predicted to show about 0.238 before detector averaging. If the object-referred pixel width is 0.75 micrometres, rectangular pixel integration multiplies this contrast by $\mathrm{sinc}(\pi f p) = 0.823$; the recorded visibility is about 0.196. Nyquist frequency for this pixel spacing is 667 cycles/mm. Hence even some frequencies that the optics can transmit lie above the detector Nyquist limit and can alias back to lower apparent frequencies.

This prediction can be tested with an adjustable magnification while keeping entrance pupil and object fixed. Object-referred pixel width falls as magnification rises; a pixel-limited measurement should approach the optical transfer curve. Raising magnification also spreads light over more pixels, changing per-pixel count noise, so compare fitted contrasts with uncertainty. At frequencies where the ideal transfer is small, a high-contrast target and low read noise matter. Dividing by an arbitrarily small transfer function to sharpen the final image can turn noise into a convincing periodic pattern.

A pupil-phase error redistributes energy without absorbing it

Introduce a known defocus into a clear circular pupil. Let the phase relative to its centre vary as $\phi(r) = \Phi r^2/a^2$. The on-axis focal amplitude is proportional to the pupil average of $e^{i\phi}$, since all ideal contributions arrive in phase there. Integrate in polar coordinates and substitute $u = r^2/a^2$. Relative to the perfect on-axis intensity the result is

$$S = \left| \int_0^1 e^{i\Phi u} du \right|^2 = \left[\frac{\sin(\Phi/2)}{\Phi/2} \right]^2.$$

This special-case Strehl ratio assumes a uniform pupil and pure quadratic defocus. Removing a phase piston makes no change to intensity. Expanding for small Φ yields

$$S = 1 - \Phi^2/12 + \cdots.$$

The pupil-phase variance after removing the mean piston is $\Phi^2/12$, which explains the leading connection between reduced peak intensity and wavefront variance without assuming that every type of aberration has the same pattern.

A quarter of a wave of centre-to-edge optical path error gives $\Phi = \pi/2$ and $S = 0.811$. One half wave gives 0.405. In an ideal phase-only pupil the missing on-axis intensity appears elsewhere in the image; total transmitted power over a sufficiently wide observation plane is unchanged. Contrast this with an absorbing pupil that removes a uniform fraction of intensity everywhere. Measuring only the brightest pixel confuses the two. Measure integrated power over increasing radii and fit a spatial profile as well as a peak.

To turn this into an axial calibration, derive the edge phase from the distance mismatch. For a paraxial marginal ray in air making angle approximately NA, an axial shift Δz introduces a relative phase magnitude near $k\,|\Delta z|\,\mathrm{NA}^2/2$. At $\mathrm{NA} = 0.10$, $\lambda = 500$ nm and $|\Delta z| = 25$ micrometres this is 1.57 rad, the quarter-wave case. With a larger numerical aperture use actual longitudinal wavevectors instead of the quadratic approximation and model polarisation where required. The formulas predict the same on-axis drop for positive and negative defocus. A single peak reading cannot give the sign of an axial misplacement; scanning through focus supplies that missing sign and can reveal an asymmetric aberration.

Distinguish aperture physics from quadrature error

Refine the aperture samples and vary their physical extent independently. A moving numerical null and a true diffraction minimum have different refinement signatures.

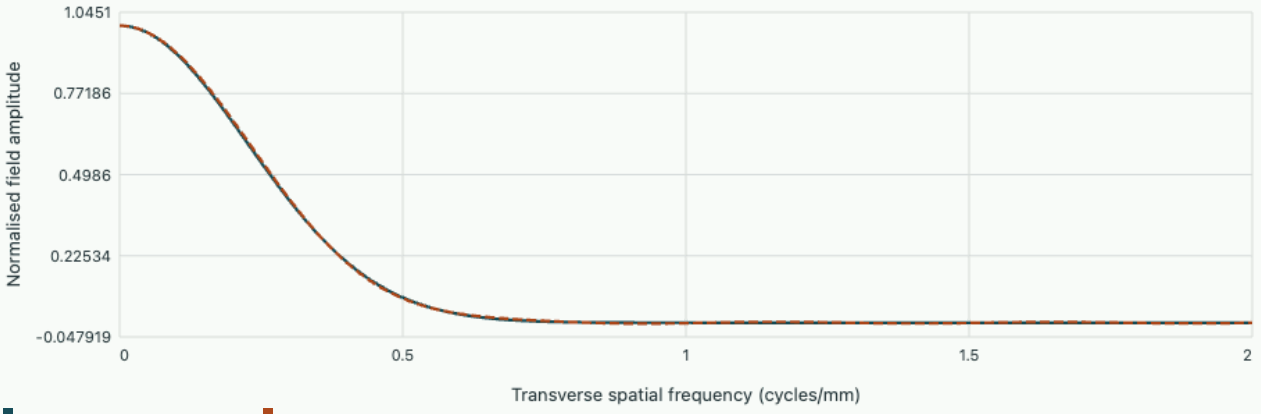
Open this experiment online

Try the calculation

Gaussian aperture radius w (mm) 1

Window half-width / w 2

Samples across window 30



Infinite Gaussian Sampled finite window

Sample spacing 0.13333 mm, Nyquist frequency 3.75 cycles/mm. Measured amplitude at 1 cycle/mm: -0.0014811.

Worked optical problem · determine how much light fits into one Airy disk

Two groups compare image quality using respectively the peak intensity and the energy contained inside the first dark ring. Derive the encircled-energy fraction for a uniform circular pupil and obtain its value at the first ring. Would closing the aperture increase the useful energy fraction on a detector of fixed physical radius?

▼ **Integrate the radial intensity with its area measure**

Write $v = ka\sin\theta$, and use the small-angle focal-plane variable proportional to image radius. Focal intensity varies as $[2J_1(v)/v]^2$. The annular area element supplies a factor $v\,dv$. Use $J_0' = -J_1$ and $J_1' = J_0 - J_1/v$ to differentiate $J_0^2 + J_1^2$. The derivative is $-2J_1^2/v$, giving the fraction inside radius v as $1 - J_0(v)^2 - J_1(v)^2$. It tends to zero at the origin and to one at very large radius, which fixes the integration constants and normalisation without needing an absolute incident amplitude.

At $v = 3.8317$, $J_1 = 0$ and $J_0 \simeq -0.4028$. The encircled fraction is therefore 0.838. The remaining 16.2 percent is outside the central disk even in a perfect ideal image. If the physical detector radius corresponded to this value before closing the pupil, reducing the pupil radius decreases v at the detector edge and therefore decreases the collected fraction. It also ordinarily reduces the total incident power transmitted by the pupil.

Peak-normalised plots rescale each pupil's image to one at its centre, which hides both effects. For a fair power comparison record the total incident flux, the pupil transmission and the energy on the fixed

detector. A phase aberration or obscuration can leave the nominal first-zero position a poor boundary for comparing different pupils. The same physical encircled radius offers a reproducible metric in those cases, provided the beam centre and background are estimated independently.

A phase ramp moves a diffraction pattern without changing its shape

Illuminate a thin aperture with a tilted plane wave in the paraxial regime. At the aperture plane this multiplies its complex amplitude by a linear phase ramp. The Fourier shift theorem then moves the entire focal-plane amplitude by a transverse amount $f \sin \theta_i$ along the tilt direction. Its intensity pattern retains its shape as long as the illumination envelope fills the pupil uniformly and the lens has enough acceptance. Tilting a narrow 532 nm beam by one milliradian in front of a 100 mm lens moves the pattern by about one tenth of a millimetre.

This provides a practical separation of phase tilt from aperture narrowing. A tilt moves both first zeros in the same direction, whereas changing the width of a centred slit moves the zeros apart or together. A position fit using one zero alone cannot separate these perturbations. Jointly fitting both zeros yields a mean position sensitive to tilt and a separation sensitive to width. A scanned detector with unequal position calibration on the two sides can mix the two observables.

For a spatially incoherent angular source, integrate intensities of the shifted patterns over the angular source distribution. An extended source then blurs zeros that would be sharp under one coherent angle. A coherent superposition of the same angular components instead requires adding the shifted amplitudes and can produce additional fringes. The illumination statistics determine which calculation describes the recorded image.

Tutorial · Distinguish a pupil defect from pixel integration

In a slit focal pattern the central lobe has one-sided first-zero distance 40 micrometres. An analyst samples precisely at that first zero using a centred pixel 8 micrometres wide and finds positive intensity. Estimate the finite-pixel fraction and propose a measurement that would expose an additional scattered background.

▼ Show the pixel calculation and a discriminating test

Let $x = x_0 + s$, where $|s|$ is much smaller than x_0. The sinc amplitude near its zero is approximately $-s/x_0$, giving intensity $I/I_{\mathrm{peak}} \simeq s^2/x_0^2$. Averaging across a pixel of width p gives $p^2/(12x_0^2)$, about 3.33×10^{-3}. This estimate uses a flat response within each pixel and a sample centred on the true minimum. A two-micrometre centring error contributes about another 0.0025 at leading order.

Use magnification that doubles the first-zero distance in pixel units while keeping pupil and wavelength fixed. Pixel leakage at a centred zero should then fall by four relative to the peak. A background proportional to local scattered light may respond differently and a dark offset can be estimated from

opaque-frame counts. Calibrate differing exposure and throughput, subtract dark counts prior to normalisation and fit both flanks to locate the true minimum.

Tutorial · Design a propagation-grid check

A 512-point transverse grid with 2 micrometre spacing represents an aperture at 600 nm. What angle can the grid represent? Could its periodically repeated boundary contaminate a nominal 0.08 rad bundle after 10 mm propagation, if the beam spreads from the centre? Explain two refinements that test distinct errors.

▼ **Show the angular and spatial budgets**

Nyquist gives $|k_x| \leq \pi/\Delta x$, hence

$$|\sin\theta| \leq \lambda/(2\Delta x) = 0.15.$$

The 0.08 rad direction is resolved in principle, although modes near Nyquist would need a separate resolution study. Its lateral displacement over 10 mm is about 0.802 mm, larger than the 0.512 mm half-window even for a beam initially confined to its centre. A periodic FFT can therefore wrap light around into the opposite side.

Double the number of points with spacing fixed. This doubles the window while keeping the represented angular limit fixed. Then halve the spacing while doubling the number again to leave this enlarged window fixed and test transverse discretisation independently. Compare conserved flux, central intensity and the location of the first minimum, retaining absolute normalisation. Flux conservation in a periodic window can hold perfectly even when light has wrapped across its boundary. A plot of that single diagnostic would miss the contamination.

MIT 2.71 Optics · Fourier optics, apertures and propagation

Waves, radiation and relativity · 10. What different observers measure

A moving observer carries clocks, rulers and a calibrated test charge. A statement about the electric field must specify which observer held that test charge at rest. In this chapter we derive how different measurements can agree on the force law and yet assign different electric fields, magnetic fields and charge densities to the same physical apparatus. A current-carrying wire and a light pulse will provide independent checks.

Events, synchronisation and directed boosts

Use two inertial coordinate systems with aligned axes. The primed origin moves at speed v in the positive laboratory x direction. Choose coincident origins when both clocks read zero and synchronise distant clocks in each frame by light signals. Homogeneity makes the transformation linear. The worldline of the primed origin requires

$$x' = A(x - vt).$$

A light ray moving along either axial direction must still obey $x' = \pm ct'$. Substituting its two possible worldlines into the linear map fixes the relative coefficients of x and t in t'. Reversing the transformation by changing the sign of v fixes the remaining scale.

$$x' = \gamma(x - vt),$$

$$t' = \gamma(t - vx/c^2),$$

$$y' = y,$$

$$z' = z,$$

$$\gamma = (1 - v^2/c^2)^{-1/2}.$$

The squared interval $c^2\Delta t^2 - |\Delta \mathbf{r}|^2$ is unchanged, a direct algebra check that detects a surprising number of sign choices. Two sparks separated by 3.0 m along a laboratory bench and simultaneous on laboratory clocks have $\Delta t' = -7.51$ ns for $v = 0.60c$, if the second event is at larger x. Their ordering in this frame is consistent with their spacelike separation. Events joined by a light signal are null-separated and retain their order under this boost.

Length measurements need two simultaneous endpoint readings in the chosen frame. A bar at rest in the lab and parallel to the axis has length L/γ in the primed frame. Substituting the simultaneous laboratory endpoint events into the boost instead gives γL, but their primed times differ. This distinction reappears when we count charge per unit length in a moving wire. The result of that count depends on the spacetime slice used to perform it.

A current is a rate of charge crossing a surface

Begin with a single population of particles, each carrying charge q, whose proper number density is n_0. Their counting volume is Lorentz-contracted only along their motion. The lab assigns density $n = \gamma_u n_0$ and current density $qn\mathbf{u}$. Both are components of $j^\mu = qn_0 u^\mu$, where $u^\mu = \gamma_u(c, \mathbf{u})$ is the four-velocity.

Independent species can be superposed. A total four-current can consequently be spacelike even though every material particle moves more slowly than light.

$$\rho' = \gamma(\rho - vJ_x/c^2),$$
$$J'_x = \gamma(J_x - v\rho),$$
$$J'_y = J_y,$$
$$J'_z = J_z.$$

The usual continuity law becomes $\partial_\mu j^\mu = 0$ with $\partial_0 = c^{-1}\partial_t$. Transformation of the derivative and of the current cancels exactly in this contraction. An isolated finite charge distribution has an observer-independent total charge if each observer integrates through a complete spacelike slice. One way to see this is to connect the two slices by a large timelike enclosing wall. Four-dimensional flux conservation sets the difference between the two charge integrals equal to the current crossing the wall. For a wall outside all the charge trajectories that last contribution vanishes.

For a charge-neutral straight wire carrying conventional current $I = 2.0$ A along $+x$, a frame moving parallel to the wire at $0.20c$ assigns line density $\lambda' = -\gamma v I/c^2 = -1.36$ nC/m. Transverse cross-section is unchanged. There is no requirement that electrons themselves travel at $0.20c$. Both the mobile and immobile species contribute to the transformed density. The ideal wire is infinitely long, so one cannot integrate its transformed charge density out to infinity and apply the finite isolated system argument.

Obtaining field components from measured forces

Place a test charge at rest at an event in the primed frame. In the lab it has velocity $v\mathbf{e}_x$ and force $q(\mathbf{E} + \mathbf{v}\times\mathbf{B})$. To transform the force, transform the corresponding momentum change. For a small transverse kick, $dp'_y = dp_y$ while $dt' = dt/\gamma$ on the test charge's path. Thus $F'_y = \gamma F_y$ and similarly for z. For the parallel component use $dp'_x = \gamma(dp_x - v\,d\mathcal{E}/c^2)$ and $d\mathcal{E}/dt = vF_x$. The result is

$$F'_x = F_x.$$

Dividing by q gives the electric field. Requiring the same force law for charges moving transversely, or transforming the magnetic flux measured by a moving loop, gives the magnetic field.

$$\mathbf{E}'_\parallel = \mathbf{E}_\parallel,$$
$$\mathbf{B}'_\parallel = \mathbf{B}_\parallel,$$
$$\mathbf{E}'_\perp = \gamma(\mathbf{E} + \mathbf{v}\times\mathbf{B})_\perp,$$
$$\mathbf{B}'_\perp = \gamma(\mathbf{B} - \mathbf{v}\times\mathbf{E}/c^2)_\perp.$$

Imagine a uniform $+y$ electric field of 90.0 kV/m and a $+z$ magnetic field with $cB = 150$ kV/m. For a boost along $+x$, the y component of $\mathbf{v}\times\mathbf{B}$ is negative. At $\beta = 0.60$ the electric field vanishes. The remaining magnetic field has $cB' = 1.25(150 - 54) = 120$ kV/m or 0.400 mT. A particle comoving with that frame feels no force at that event. A different particle moving across the magnetic field can still curve.

Vary the observer speed and compare the electric and magnetic field invariants

Compare two force measurements near an ideal wire

Return to the 2.0 A wire and put the observation point 1.0 cm along $+y$. Ampère's law gives $B_z = \mu_0 I/(2\pi r) = 40.0$ microtesla. The lab radial electric field is zero in this idealised symmetric model. A positive test charge moving along $+x$ at $0.20c$ experiences an inward magnetic force per unit charge of 2.40 kV/m. In its momentary rest frame the wire's negative line charge produces an inward radial electric field. The field from Gauss's law is $|\lambda'|/(2\pi\epsilon_0 r) = 2.45$ kV/m. The difference in the reported transverse forces is precisely the factor $\gamma = 1.0206$ derived from the momentum kick.

There is no contradiction between a field with zero laboratory divergence outside the wire and a different electric field in the primed frame. The charge density at the wire itself also transforms. If the laboratory wire carries a real driving surface charge and a return conductor, add their fields and charges before transforming. A claimed measurement of this ideal result without specifying the return path and electrostatic shielding needs a separate model for those additional fields.

When can a field be made purely electric or magnetic?

Direct substitution in the transverse transformation leaves the scalars $B^2 - E^2/c^2$ and $\mathbf{E}\cdot\mathbf{B}$ unchanged. The first is a difference, so retain its sign. If the second is nonzero, no observer sees either field vanish. If the fields are orthogonal and $E < cB$, choose $\mathbf{v} = \mathbf{E}\times\mathbf{B}/B^2$. The vector identity for $(\mathbf{E}\times\mathbf{B})\times\mathbf{B}$ cancels $\mathbf{E}$, and the required speed is below c. With $E > cB$, choosing $c^2\mathbf{E}\times\mathbf{B}/E^2$ instead cancels $\mathbf{B}$.

When both invariants vanish one must still distinguish a field identically zero from a nonzero null field. A vacuum travelling plane wave satisfies $E = cB$ and orthogonality everywhere. The advertised boost would then require speed c. An observer can lower its measured frequency and amplitude continuously by following the wave at increasing speed but can never transform away its nonzero field. A superposition of waves need not be null. A standing wave has positions and times where only one of the fields is present.

Phase matching across inertial frames

Every observer agrees on whether a particular spacetime event is a crest. The phase $\mathbf{k}\cdot\mathbf{r} - \omega t$ therefore labels the same event in each frame. Substituting the inverse coordinate transformation gives the transformation of frequency and wavevector. For vacuum propagation at laboratory angle θ to the boost, divide the transformed longitudinal wavenumber by the transformed magnitude.

$$\omega' = \gamma\omega(1 - \beta\cos\theta),$$

$$\cos\theta' = \frac{\cos\theta - \beta}{1 - \beta\cos\theta}.$$

An observer chasing a collinear 600 nm light wave at $0.60c$ sees half its frequency and a wavelength of 1200 nm. A beam entering at 90° to the boost appears at $\cos\theta' = -0.60$ and frequency 1.25ω. Asking for a transverse Doppler shift without saying in whose frame the angle is a right angle leaves the problem

under-specified. A detector counts phase fronts crossing its own trajectory and its own clock sets the elapsed time.

A boosted light beam also transports momentum

Frequency alone does not specify the energy measured by a moving detector. Its collecting surface and clock must be assigned to an observer. For a monochromatic collimated beam in vacuum, one route is to transform both fields and calculate their new Poynting vector. Another route is to transform the stress-energy tensor, whose energy density, momentum density and momentum flux describe the same field together. This second route shows explicitly which measured quantity is being transformed and extends to superposed beams with zero mean momentum in one frame.

For a right-moving plane wave along x, the energy density is $u = S_x/c$, momentum density is u/c and normal momentum flux is u. In coordinates with $x^0 = ct$, the relevant block has $T^{00} = T^{01} = T^{10} = T^{11} = u$. Boost both indices using the same Lorentz matrix as for coordinates. The transformed energy density follows directly:

$$T'^{00} = \gamma^2(T^{00} - 2\beta T^{01} + \beta^2 T^{11}) = \gamma^2(1-\beta)^2 u.$$

The frequency Doppler factor enters twice. One factor corresponds to lower energy per photon and one to the changed number per unit volume. For a finite pulse, its length at a common primed time grows for an observer following it, so the total energy changes by only one Doppler factor. Volumes measured between equal-time spatial endpoints must be computed on the appropriate observer's simultaneity slice, just as for a moving rod.

Now superpose two equal-intensity mutually incoherent plane waves in opposite axial directions. In the laboratory $T^{01} = 0$, while $T^{11} = T^{00} = u$ for the combined radiation. An observer moving at $0.60c$ obtains $u' = \gamma^2(1+\beta^2)u = 2.125u$, with net flux $S'_x = -2\gamma^2\beta\, cu = -1.875\, cu$. The counterpropagating beam that the observer meets head-on is blueshifted, while the other is redshifted. Energy density remains greater than the magnitude of momentum density times c, as required for a superposition of distinct propagation directions. This check fails if one transforms a zero Poynting vector as though it were the spatial component of a four-vector with no stress information.

The direction-independent radiation field in an ideal cavity has $T^{ii} = u/3$ for each spatial axis in its rest frame rather than $T^{11} = u$. The one-third comes from averaging n_i^2 over a sphere. The moving observer then sees $u' = \gamma^2 u(1 + \beta^2/3)$ and a net longitudinal flux. At $\beta = 0.60$ the energy-density factor is 1.75. An isotropic thermal field becomes directionally distorted; assigning one Doppler factor to the whole sky would give neither the correct integrated energy density nor its directional spectrum.

A moving parallel-plate capacitor tests surface sources

Choose large parallel conducting plates at $z = \pm d/2$ carrying opposite uniform surface densities $\pm\sigma$. Ignore edge effects in the central region. In the lab the field between them has magnitude $E_0 = \sigma/\epsilon_0$ and there is no magnetic field. A frame boosted along $+x$ sees both plates move at $-v$. Their areas contract along x, so their surface charge magnitudes become $\gamma\sigma$; their equal and opposite surface currents point

in opposite senses. Use Gauss's law and the magnetic jump condition on each sheet to reconstruct the transformed fields. The current sheets reinforce magnetically between the plates and cancel outside, just as electric sheets of opposite charge do electrostatically.

For laboratory electric field along $+z$, the moving observer has $E'_z = \gamma E_0$ and $B'_y = +\gamma v E_0 / c^2$. Cross the primed plate velocity $-v\mathbf{e}_x$ with this magnetic field to find a force along $-z$ on a positive surface charge, partly cancelling the enhanced electric attraction towards the opposite plate. When finding the force on a sheet, use the fields of the other sheet or the arithmetic mean of the two limiting fields. Substituting the total interior field without removing the self contribution would double the result.

At $E_0 = 2.00$ MV/m and $\beta = 0.60$, the primed interior field is 2.50 MV/m and its magnetic field 5.00 mT. Per positive surface charge the transverse attraction from the opposite sheet has magnitude $\gamma E_0 (1 - \beta^2)/2 = E_0/(2\gamma)$. This is 0.800 MV/m, compared with 1.00 MV/m in the laboratory. The transformation factor is inverse γ because the charge in this example is at rest in the original laboratory; the earlier force derivation placed the test charge at rest in the primed frame. Interchanging those initial conditions without changing the time conversion reverses the factor.

Multiplying the primed force per charge by the larger surface density $\gamma\sigma$ restores the same force per transverse plate area in this tangential boost. The cancellation supplies a local cross-check among length contraction, field transformation and mechanical force. A complete calculation of the total energy or momentum of an accelerated finite capacitor must also account for stresses in its supports and for the fields near its edges. Taking the uniform central field times a transformed volume is adequate for a local comparison, but does not by itself describe that whole closed mechanical system.

Test a field transformation with an invariant

Change observer speed, compare the separately transformed fields and inspect the invariant combination. Agree on the event before comparing measurements.

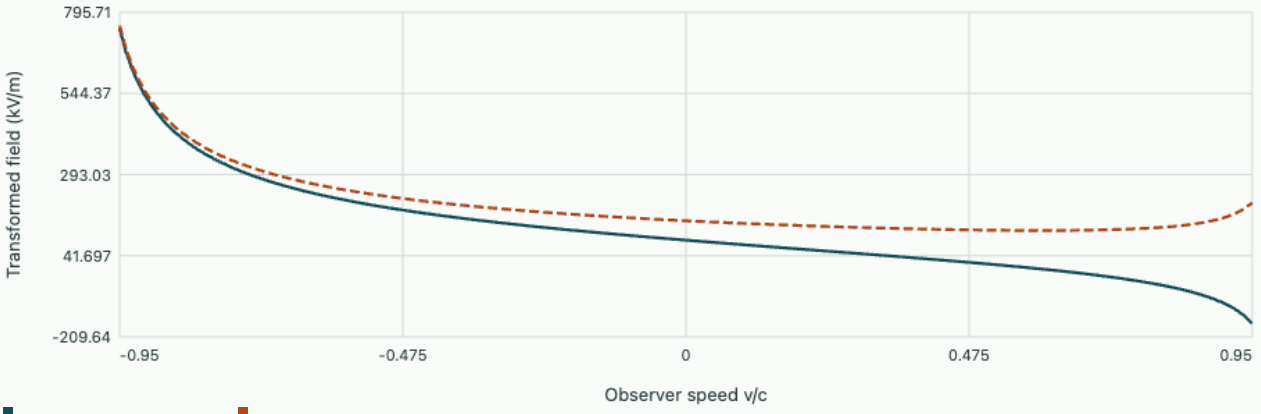
Open this experiment online

Try the calculation

Electric field E (kV/m) **90**

Magnetic field cB (kV/m) **150**

Observer speed β **0.6**



At the selected speed E′ = 0 kV/m and cB′ = 120 kV/m. E′² − c²B′² = -14400 (kV/m)².

Worked observer problem · a null of one component does not imply a vanishing field

At one event a laboratory finds $\mathbf{E} = (30, 40, 0)$ kV/m and $c\mathbf{B} = (0, 0, 50)$ kV/m. Another laboratory reports that it can remove the magnetic field by boosting along the direction of $\mathbf{E} \times \mathbf{B}$. Test the claim and predict the fields for a boost of speed $0.80c$ along this direction. Does the local measurement establish that this is an infinite plane wave?

▼ **Evaluate the invariants and transform the transverse field**

The electric and scaled magnetic fields both have magnitude 50 kV/m and are perpendicular. Both quadratic Lorentz invariants vanish, while the fields themselves are nonzero. With $\mathbf{v}$ along $\mathbf{E} \times \mathbf{B}$, both field vectors are perpendicular to the boost. The cross terms reduce their magnitudes by a common factor $\gamma(1 - \beta) = 1/3$ for this speed, so $E' = cB' = 16.7$ kV/m. Their directions in this transverse configuration are unchanged. No subluminal boost removes either nonzero field, as follows also by applying the inverse transformation to an alleged zero field.

A pulse at one point can have these local field values, as can the field near a suitable superposition at a particular event. A local tensor and its invariants contain no information about all the field derivatives elsewhere. To establish a plane-wave model measure spatial phase across transverse probes, verify the propagation relation at separated stations and bound amplitude variation over the required observation region. Infinity is an idealisation established only to a scale specified by the experiment.

As an error check, compare the measured difference $E^2 - c^2B^2$ with the propagated uncertainties of both meters. Small nonzero estimated differences near a null may be statistically insignificant, while a

systematic alignment error in $\mathbf{E}\cdot\mathbf{B}$ can obscure orthogonality. Classifying a field as electric-dominated or magnetic-dominated from a difference of nearly equal large squares requires an uncertainty budget. The exact mathematical classification and the experimental confidence in assigning it answer different questions.

Recover a source direction from two clocks and a known speed

Two perfectly synchronised receivers travel with equal speed in opposite directions along the laboratory axis. They sample the same monochromatic plane wave. Their measured frequencies are proportional to $1-\beta\cos\theta$ and $1+\beta\cos\theta$, with a common Lorentz factor. Taking their difference divided by their sum determines $\beta\cos\theta$ without prior knowledge of emitted frequency. Their sum then fixes that frequency once the speed is calibrated. The construction combines two observer assignments rather than treating a transverse Doppler correction as an unknown empirical shift.

For speeds of half the speed of light and a measured frequency ratio of three to two, with the receiver travelling toward negative x recording the higher value, the inferred cosine is 0.40. If their average measured frequency is 600 THz, the emitted laboratory frequency is about 520 THz. A second direction on the opposite side of the boost axis has the same cosine, so the two receivers alone cannot determine the full incoming unit vector.

A third receiver with a nonparallel velocity supplies another directional projection. Detector clocks must be calibrated in their respective rest frames and each must see the same stable source frequency. If the source itself chirps between noncoincident measurements, that chirp can imitate an angular shift. Using overlapping signals and comparing phase histories addresses this limitation more directly than quoting only two peak frequencies.

Tutorial · Infer charge and current from two observers

A uniform channel carries an unknown lab line charge λ and current I along its axis. Observers moving at $+0.30c$ and $-0.30c$ measure line densities 1.0 and 3.0 nC/m respectively. Recover the laboratory quantities. Can a third observer travel alongside this current so that the measured current vanishes?

▼ **Recover the two independent quantities and classify the current**

Adding the readings gives $2\gamma\lambda = 4.0$ nC/m, so $\lambda = 1.908$ nC/m. Subtraction gives $2\gamma\beta I/c = 2.0$ nC/m and hence $I = 0.953$ A. Reinsert both signs in the original equations as a check. The speed needed to cancel the current would be $I/\lambda = 5.00\times10^{8}$ m/s. It exceeds the speed of light. The current four-vector is spacelike since $I^2 > (c\lambda)^2$.

A frame with zero line charge can instead move at $v = c^2\lambda/I = 0.600c$. A spacelike total current is possible for superposed charge carriers of opposite sign. The calculation would rule out a model consisting solely of one sign of carriers all propagating with a common subluminal drift velocity. It does not rule out the measured pair of densities.

Tutorial · What does a moving pulse detector count?

A 10.0 ns rectangular-envelope light pulse moves in $+x$ in vacuum. A detector moves along the same direction at $0.60c$. Compute the interval between its encounters with the leading and trailing edges using laboratory kinematics, then its own clock. Predict its measured peak intensity and fluence relative to a stationary detector for a broad collimated beam.

▼ Track the two crossing events and transform the flux

The pulse is $c(10.0\,\mathrm{ns})$ long in the lab. Its back approaches the moving detector at coordinate closing speed $c - v = 0.40c$, so the two events occur 25.0 ns apart in laboratory time. Both events lie on the detector worldline, so its proper interval is 20.0 ns. This agrees with the inverse Doppler factor of two. A pulse of enough cycles to define a carrier now has half the frequency and twice its envelope duration.

For a transverse plane wave the field amplitude transforms as $\gamma(1 - \beta) = 0.50$ times the laboratory value. The primed Poynting flux is therefore one quarter as large. Integrating over twice the proper duration gives one half the original energy per unit transverse area. The cross section perpendicular to the boost is unchanged. Photons are registered with half their laboratory energy and the same total number if the beam is entirely collected, an independent check on the fluence result.

Cambridge Electromagnetism · electromagnetism and relativity

Waves, radiation and relativity · 11. Relativistic particle dynamics and the field tensor

A circular track measures momentum per charge. A clock following the particle registers proper time, and a detector records how much kinetic energy it has gained. We will put these observations in one dynamical equation, solve that equation in uniform fields and exhibit a numerical pusher that can get the energy right while its orbit phase slowly slips.

What the track, clock and calorimeter measure

Consider a particle of invariant mass m, signed charge q and ordinary velocity $\mathbf{v}$. Inertial observers agree on the proper time between neighbouring events on its timelike worldline,

$$d\tau = dt\sqrt{1 - v^2/c^2}.$$

Adopt coordinates $x^\mu = (ct, x, y, z)$ and metric $g = \mathrm{diag}(1, -1, -1, -1)$. Dividing the displacement by proper time produces a vector of fixed Minkowski norm. Multiplying it by invariant mass gives four-momentum.

$$u^\mu = \gamma(c, \mathbf{v}),$$
$$p^\mu = m u^\mu = (\mathcal{E}/c, \mathbf{p}),$$
$$\mathcal{E}^2 - c^2|\mathbf{p}|^2 = m^2 c^4.$$

Energy includes the rest term. Momentum is $\gamma m \mathbf{v}$. Differentiating the last equation on a trajectory gives

$$d\mathcal{E}/dt = \mathbf{v} \cdot d\mathbf{p}/dt.$$

This identity replaces the nonrelativistic step of dotting $m\mathbf{a}$ with velocity. A magnetic field can rotate momentum with fixed magnitude and energy, whereas an electric field can change energy by doing work.

Suppose an electron begins at rest and gains 10.0 keV of kinetic energy across an electrostatic potential difference in vacuum. Its rest energy is 511 keV, giving $\gamma = 1.01957$ and $v/c = \sqrt{1 - 1/\gamma^2} = 0.1950$. The nonrelativistic estimate is 0.1978. In a modest energy measurement the discrepancy could lie below the experimental spread. For a precise orbit calculation the 1.96 percent change in relativistic cyclotron period is measurable independently of that discrepancy in speed.

A linear map that produces the Lorentz force

Seek a local equation linear in field and four-velocity and covariant under boosts. For the energy component to give $q\mathbf{E} \cdot \mathbf{v}$, and for the spatial components to give the electric force and the oriented cross product, construct the antisymmetric field tensor. With the stated metric and $\epsilon^{123} = 1$, set

$$F^{0i} = -E_i/c,$$

$$F^{i0} = E_i/c,$$

$$F^{ij} = -\epsilon^{ijk} B_k,$$

$$\frac{dp^\mu}{d\tau} = qF^{\mu\nu} u_\nu.$$

Lowering the index on u reverses the signs of its spatial components. The spatial product becomes $q\gamma(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, and division by γ converts the proper-time derivative to $d\mathbf{p}/dt$. The time component similarly yields

$$d\mathcal{E}/dt = q\mathbf{E} \cdot \mathbf{v}.$$

Contract the force with u_μ. The symmetric product $u_\mu u_\nu$ annihilates the antisymmetric tensor, which proves that the evolution preserves the mass shell.

To recover Maxwell's kinematic constraints from potentials, put $A^\mu = (\phi/c, \mathbf{A})$ and $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. Then $F^{0i} = (\partial_t A_i + \partial_i \phi)/c = -E_i/c$, while the spatial-spatial components are the magnetic curl. Adding a four-gradient to A^μ leaves the tensor unchanged because ordinary derivatives commute. The particle action $S = -mc^2 \int d\tau - q \int A_\mu dx^\mu$ gives the same force after integrating variations by parts with fixed endpoints. Gauge changes add only an endpoint term to this action. A trajectory measured with local forces cannot distinguish such potential representatives.

Compare the potentials and field-based description in the field-tensor chapter

Constant force, variable coordinate acceleration

Put a particle initially at rest in a uniform static electric field along x. Let $a_0 = qE/m$ be signed. Integration of $dp_x/dt = qE$ gives $p_x = ma_0 t$, so the speed tends to c even while momentum increases linearly. Proper-time evolution is simplest in terms of rapidity η, where $u^0 = c\cosh\eta$, $u^x = c\sinh\eta$. The two tensor equations imply $d\eta/d\tau = a_0/c$.

$$t = \frac{c}{a_0}\sinh(a_0\tau/c),$$

$$x = \frac{c^2}{a_0}\{\cosh(a_0\tau/c) - 1\},$$

$$\frac{dv_x}{dt} = \frac{a_0}{\gamma^3}.$$

The initial time and position are zero. The work $qEx = mc^2(\cosh\eta - 1)$ equals kinetic energy, which checks both the integration constant and the dimensional factor in the displacement. For an electron use the direction of the electric force as positive to simplify signs. At 20.0 kV/m it gains 1.0 keV in 5.0 cm. The energy gain follows from the static potential even though the speed does not grow exactly as the square root of distance. The field region must last long enough for the particle to cross it and fringe fields can alter the entry trajectory.

Compare momentum, coordinate time and proper time in the charged-worldline experiment

Helical motion, radius and period

A uniform magnetic field along z keeps p_z, energy and $|\mathbf{p}_\perp|$ fixed. Write $p_x + ip_y$ as a complex number. The Lorentz force reduces to a first-order rotation equation and a single time integration gives the corresponding position. With positive B, positive charges starting along $+x$ initially turn towards $-y$.

$$\frac{d}{dt}(p_x + ip_y) = -i\frac{qB}{\gamma m}(p_x + ip_y),$$

$$R = \frac{p_\perp}{|q|B},$$

$$T = \frac{2\pi\gamma m}{|q|B}.$$

The helical pitch is $v_z T$. Radius gives transverse momentum and a calibrated period gives total energy through γ. A track seen only in projection along the field gives no direct evidence that longitudinal momentum vanishes. A second view or a timing measurement can resolve this ambiguity.

At 10.0 keV the electron has $pc = \sqrt{K(K + 2mc^2)} = 101.6$ keV. With entry perpendicular to a 2.0 mT field its radius is 0.169 m and its period 18.2 ns. Doubling the field halves both radius and period at fixed injection energy. Doubling momentum at fixed field doubles radius but changes period by the ratio of Lorentz factors. These are different tests of a track calibration. A beam with a spread of energies will dephase relative to a narrow timing gate even when every single particle remains in a perfectly uniform field.

A drift frame gives an exact crossed-field solution

For constant crossed fields with $E < cB$, transform to the frame moving at $\mathbf{v}_d = \mathbf{E} \times \mathbf{B}/B^2$. There $E' = 0$ and $B' = \sqrt{B^2 - E^2/c^2}$; calculate the transformed initial momentum and integrate a helix. Transform two successive equal-phase events back to the lab. Their perpendicular displacement in the electric-free frame is zero. If the motion has no component parallel to $\mathbf{B}$, the entire displacement there is zero, and the average lab cycle velocity is exactly $\mathbf{v}_d$. The detailed lab trace need not be circular and the instantaneous lab energy may oscillate.

Take $E = 30.0$ kV/m along y and $B = 0.50$ mT along z. The drift is along $+x$ at 6.00×10^7 m/s. A particle injected at that velocity crosses the uniform fields on a straight line because the forces cancel. An initially stationary positive particle instead begins by accelerating along $+y$. These observations can use the same field with different injected particles. If E exceeds cB, this electric-free frame is unavailable. Transform to a magnetic-free frame and integrate electric acceleration instead.

Continuous rotation and a discrete experiment

A naive forward step $\mathbf{p}_{n+1} = \mathbf{p}_n + q\Delta t\, \mathbf{v}_n \times \mathbf{B}$ adds a vector perpendicular to the previous momentum. Its squared norm increases by the squared magnitude of that kick. The apparent heating will persist in perfectly static uniform magnetic fields. For a purely magnetic update, evaluate the direction using the mean of old and new momenta. At fixed γ, the resulting implicit two-dimensional linear map can be solved algebraically and is an exact rotation by a modified angle.

$$\delta = 2\arctan\left(\frac{\omega_c \Delta t}{2}\right),$$

$$\omega_c = \frac{qB}{\gamma m}.$$

The sign of rotation follows the earlier complex equation. Dotting the update with the mean momentum gives zero change in squared norm, which verifies the energy property even for relatively large time steps. A common electromagnetic pusher surrounds the magnetic rotation by two electric half-kicks. In spatially varying fields, specify at what time and position each field is sampled, and check spatial interpolation and time convergence. For ultrarelativistic nearly cancelling electric and magnetic fields, preserving rotation in a pure field alone does not ensure cancellation is represented accurately.

Solve the constant tensor without transforming every time step

When uniform electric and magnetic fields are constant in spacetime, the Lorentz equation is a linear constant-coefficient equation for four-velocity as a function of proper time. Nonlinearity in ordinary time hides that simple structure, because the rate at which proper time accumulates changes along the orbit. Define the mixed tensor $M^{\mu}{}_{\nu} = (q/m)F^{\mu\alpha}g_{\alpha\nu}$. In a uniform field it does not vary along the trajectory. A formal solution is then

$$u(\tau) = e^{M\tau}u(0),$$

$$x(\tau) - x(0) = \int_0^{\tau} e^{Ms}u(0)\, ds.$$

The coordinate formula is written as an integral since M can be singular, as it is in a purely magnetic field. Blindly replacing this integral by $M^{-1}(e^{M\tau} - I)$ fails in exactly such simple cases. A power series for the integral or a block matrix exponential treats zero eigenvalues without an artificial inverse. Antisymmetry of the tensor implies $M^{\mathsf{T}}g + gM = 0$. Differentiating $e^{M^{\mathsf{T}}\tau}ge^{M\tau}$ shows that it stays equal to g, making preservation of $u^{\mu}u_{\mu} = c^2$ a structural property of the exact mapping.

For parallel constant fields along z, the matrix decomposes into a boost in the ct, z subspace and a rotation in x, y. Write $\alpha = qE/(mc)$, $\Omega = qB/m$, and denote the initial four-velocity components by u_0^{μ}. The proper-time solution is

$$u^0 = u_0^0\cosh\alpha\tau + u_0^z\sinh\alpha\tau,$$

$$u^z = u_0^z\cosh\alpha\tau + u_0^0\sinh\alpha\tau,$$

$$u^x + iu^y = (u_0^x + iu_0^y)e^{-i\Omega\tau}.$$

The longitudinal and transverse operations commute because they act in different subspaces. In lab time the transverse rotation slows as the electric field increases energy, although the proper-time rotation frequency stays constant. A particle therefore traces a helix whose transverse radius $p_\perp/(|q|B)$ is fixed while the pitch evolves. A single measured turn time without a tagged position in the accelerating region is insufficient for an energy reconstruction.

As a numerical test use an electron, quoting magnitudes, with $B = 1.0$ mT, $E = 10.0$ kV/m and initial purely transverse speed $0.10c$. One magnetic turn takes 35.72 ns in proper time and accumulates electric rapidity magnitude $|\alpha|\tau \simeq 0.2096$. The final Lorentz factor is $\gamma_i \cosh(0.2096) = 1.0272$, giving final kinetic energy about 13.9 keV, compared with 2.57 keV initially. The lab-time turn duration is $\gamma_i \sinh(|\alpha|\tau)/|\alpha| \simeq 36.2$ ns. Reversing the electric field changes the direction of the axial displacement but leaves final energy and turn duration unchanged for this initially transverse particle. Reversing the initial longitudinal velocity would generally break that energy symmetry.

Infer uncertainty and test units in a momentum reconstruction

A track fitter commonly reports field, radius, pitch, time and a fitted covariance matrix. The dynamical relation turns these observed quantities into momentum. Let L_z denote a signed pitch per turn. With known signed direction of circulation and calibrated absolute field magnitude,

$$|\mathbf{p}| = |q|B\sqrt{R^2 + (L_z/2\pi)^2}.$$

For small errors, logarithmic differentiation gives weights for fractional radius and pitch errors:

$$d\log p = d\log B + w_R\, d\log R + w_L\, d\log |L_z|,$$

$$w_R = \frac{R^2}{R^2 + (L_z/2\pi)^2},$$

$$w_L = 1 - w_R.$$

For $R = 0.20$ m and $L_z = 0.30$ m the radius accounts for 94.6 percent of the squared transverse-plus-longitudinal geometric factor. Independent relative uncertainties of 0.20 percent in field, 0.50 percent in radius and 2.0 percent in pitch yield about 0.525 percent uncertainty in total momentum. A common detector scale controlling both radius and pitch would enter with correlated signs and should not be counted as two independent errors. When the pitch tends to zero, use the original derivatives with respect to L_z; a reported fractional error on a quantity consistent with zero is ill defined.

Differentiating $K = \sqrt{m^2c^4 + p^2c^2} - mc^2$ gives

$$d\log K = (\gamma + 1) d\log p/\gamma.$$

The coefficient approaches two at low energy and one at very high energy. At the approximately 3.70 keV energy of the tutorial the above momentum error would imply about a 1.05 percent energy error. A kinetic-energy estimate from a small difference of nearly equal numbers can also lose numerical precision at much lower energies. Rationalise it as $K = p^2c^2/(\sqrt{m^2c^4 + p^2c^2} + mc^2)$ when working near rest.

Several independent checks use this reconstruction without assuming the track fitter is correct. The turn time inferred from the fitted γ can be compared with a separate timing channel. Reverse the magnetic field to reverse the direction of curvature while keeping the speed and energy approximately fixed. Translate the beam within a mapped uniform field and look for changing fitted radius that correlates with position, an indication of geometric or field-map bias. Finally plot the residual of the magnetic force law separately for transverse magnitude and angle. A time integrator may conserve momentum magnitude for symmetry reasons even when its field interpolation biases the phase or shifts the guiding centre systematically.

Advance a charge using relativistic momentum

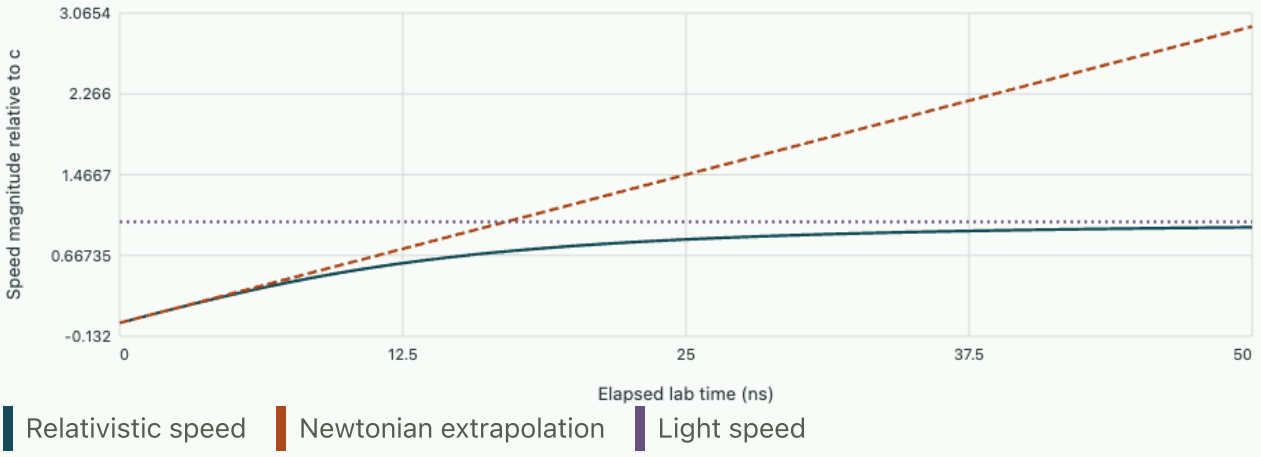
Increase the impulse while following speed, energy and proper time. Identify the range where a Newtonian momentum update becomes inaccurate.

Open this experiment online

Try the calculation

Electric field magnitude (kV/m) **100**

Lab time for a reading (ns) **17**



At 17 ns laboratory time the electron speed is 0.706 c. Its clock has advanced 15 ns and it has gained 211 keV.

Worked timing problem · infer an accelerating voltage without a radius

A tagged electron bunch completes ten nearly circular turns in a calibrated 5.0 mT magnetic field. Two time stamps indicate an elapsed 73.0 ns. The field is perpendicular to the orbit. Assuming complete turns and negligible electrical work during the timed interval, estimate the injection energy. What timing precision would distinguish a one-percent change in that energy?

▼ Use the relativistic period sensitivity rather than the energy fraction directly

The nonrelativistic turn period is

$$2\pi m/(eB) = 7.145\,\mathrm{ns}.$$

The measured period is 7.30 ns, giving $\gamma \simeq 1.0217$ and $K = (\gamma - 1)mc^2 \simeq 11.1$ keV. The relative energy sensitivity is much larger than the relative period change when γ is close to one. A one-percent change in K corresponds to $\Delta\gamma \simeq 2.17 \times 10^{-4}$, or relative period change 2.12×10^{-4}. Over the ten turns this gives approximately 15.5 ps.

To resolve that change spectroscopically, uncertainties in the elapsed time and field scale must together be significantly smaller than the equivalent change. A constant pickup cable delay cancels if the same electronics register both ends of the interval, while a drift of its timebase scale does not. Timing over more turns improves the total accumulated signal but also increases exposure to field inhomogeneity, energy loss and dephasing among particles with different injection energies. A narrower timing histogram does not automatically eliminate a systematic field calibration.

Finally verify the integer turn count independently. Missing a pickup crossing would cause a much larger change in inferred energy than the desired resolution. The measured wave packet may broaden so strongly that later crossings overlap. Recording the full pickup waveform and tracking ordered crossings, along with an independent rough radius estimate, can constrain the turn assignment before fitting the small relativistic shift. A perfect circular image with no time labels cannot supply that missing count.

Why momentum and kinetic energy do not gain fixed fractions together

In an electrostatic accelerating gap, the kinetic energy increment for a particle passing between two equipotential surfaces is its charge times the potential drop, regardless of injection energy. This statement follows by integrating the Lorentz power law over the actual path. A magnetic guiding field can alter the path and transit time while doing no direct work. A time-dependent electric field, a missed equipotential endpoint or appreciable radiation changes the simple inference and must be addressed separately.

For an electron entering with 10 keV and acquiring an additional 10 keV, kinetic energy doubles. Its initial momentum times c is about 101.6 keV and its final value about 144.4 keV. The radius in the same subsequent perpendicular magnetic field increases by a factor approximately 1.42. Doubling the radius would instead require doubling momentum, which demands a substantially greater energy increment at this low initial energy. Comparing radii before and after a known voltage supplies an independent check on the sign and magnitude of a particle's charge-to-mass model.

A measured energy gain below the nominal voltage does not by itself diagnose radiation. The potential actually traversed, finite angular acceptance, entrance or exit fringe fields, collision losses and biases in reconstructing longitudinal momentum can all matter. Repeat with reversed field polarity or several injection energies and compare the kinetic-energy difference rather than the two radii directly.

Tutorial · Reconstruct a track with two measurements

A particle known to be an electron or positron crosses a uniform $+z$ field of 1.0 mT. The projected radius is 0.200 m and the pitch magnitude is 0.300 m. Recover the transverse and longitudinal momentum, energy and period. What else must the detector record to determine the sign of charge?

▼ **Reconstruct both momentum components**

The transverse momentum is $|q|BR = 3.20 \times 10^{-23}$ kg m/s. During a cycle the parallel displacement is $p_z T/(\gamma m) = 2\pi p_z/(|q|B)$, so $|p_z| = 7.65 \times 10^{-24}$ kg m/s. The quadrature total is 3.29×10^{-23} kg m/s, yielding $p/(mc) = 0.1206$, $\gamma = 1.00724$, and approximately 3.70 keV kinetic energy. The period is about 36.0 ns.

Neither radius nor pitch magnitude carries the signed circulation. Record the direction of traversal relative to the calibrated direction of magnetic field. Reversing both charge and the direction of traversal yields the same unlabelled transverse ring. If timing or two ordered detectors supply a direction, the sign follows from the first Lorentz kick.

Tutorial · Can a constant-energy circle be trusted?

A pusher uses 20 uniform steps per exact cyclotron period in a uniform magnetic field. It reproduces initial kinetic energy to machine precision for 50 turns. Estimate the phase drift from the midpoint rotation formula. Design a refinement diagnostic that distinguishes a time-discretisation bias from a field calibration error.

▼ **Compare the exact and discrete angles**

The exact unsigned angle per step is

$$h = 2\pi/20 = 0.31416.$$

The discrete one is

$$2\arctan(h/2) = 0.31161.$$

The accumulated deficit over 1,000 steps is about 2.55 rad, roughly 0.41 turn. The method can keep every track point on a closed momentum circle yet miss the pickup electrode by a large fraction of a revolution. Coordinate radius may depend on how positions were staggered, so perform the angular comparison on momentum first.

Double the number of steps at fixed physical end time and take the difference in unwrapped phases. The leading angular error $h^3/12$ per step becomes a quadratic global error, so halving the step should reduce it by about four. A calibration error in $qB/(\gamma m)$ persists under time refinement. Track both the expected frequency and the convergence of the time series; comparing the coordinates at a single late time can confuse complete turns with near agreement.

Cambridge Electromagnetism · four-vectors, fields and charged matter

Waves, radiation and relativity · 12. Radiated power, mechanical work and energy limits

A charge can radiate while its speed is constant. A source can exchange energy with its nearby field before a pulse reaches a receiver. We will use the flux through a distant sphere, the work at the charge and the energy stored in a local field as three operationally distinct measurements. Agreement requires the relevant integration times and the range of validity of the particle model.

Integrate the far field of one slowly moving charge

For a source moving slowly compared with c, the radiation field at a large distance is transverse to the viewing direction $\mathbf{n}$. To leading order in v/c, its electric amplitude is $q\mathbf{n}\times(\mathbf{n}\times\mathbf{a})/(4\pi\epsilon_0 c^2 r)$ evaluated at retarded time. The magnetic radiation field is $\mathbf{n}\times\mathbf{E}/c$. Form the radial Poynting flux and integrate over the sphere. Choose the polar axis along the instantaneous acceleration.

$$\frac{dP}{d\Omega}=\frac{q^2a^2}{16\pi^2\epsilon_0 c^3}\sin^2\theta,$$

$$\int \sin^2\theta\, d\Omega=\frac{8\pi}{3},$$

$$P=\frac{q^2a^2}{6\pi\epsilon_0 c^3}.$$

The angular integral contains a factor of $\sin\theta$ from solid angle in addition to the square from transverse projection. The result is instantaneous emission power, not the average over a sinusoid. For acceleration $a_0\cos\omega t$, the average is one half the power computed with peak acceleration a_0. No acceleration component along the velocity is required. A charge in circular motion radiates even when the externally applied magnetic field does no work.

A trapped electron executing small prescribed oscillations at 100 MHz with amplitude 1.0 micrometre has peak speed 628 m/s and peak acceleration 3.95×10^{11} m/s squared. Its mean radiated power is 4.45×10^{-31} W. The energy scale of this illustrative classical orbit is only 1.80×10^{-25} J, much smaller than one oscillator quantum at 100 MHz. The field calculation is meaningful as a classical scaling or for an appropriately prepared mean current. Interpreting it as a literal well-defined single-electron mechanical orbit would require addressing the state preparation. Raising the amplitude to 1.0 mm raises energy and mean emitted power by a factor of a million, and keeps v/c below 0.003.

Integrate the measured dipole pattern and compare with the total radiated power

Transform instantaneous power using the rest frame

In the instantaneously comoving frame the nonrelativistic expression applies at that event, with the proper acceleration. The dipolar pattern there carries zero net three-momentum when integrated over all directions, by front–back symmetry. The radiation four-momentum emitted during $d\tau$ is therefore $(P_0 d\tau/c,\mathbf{0})$. A boost multiplies its energy by γ, but also multiplies the emission-event time interval by γ. Thus the total laboratory power allocated per emitter coordinate time equals P_0. The directional detector signal need not share this invariance.

At the instant of rest the proper acceleration is the Euclidean norm of the spatial part of $du^\mu/d\tau$. Its square in any frame follows from the invariant $-a^\mu a_\mu$. Differentiating $u^\mu = \gamma(c, \mathbf{v})$, with $d\gamma/dt = \gamma^3 \mathbf{v}\cdot\mathbf{a}/c^2$, and separating acceleration parallel and perpendicular to velocity gives

$$P = \frac{q^2}{6\pi\epsilon_0 c^3}\left(\gamma^6 a_\parallel^2 + \gamma^4 a_\perp^2\right) = \frac{q^2\gamma^6}{6\pi\epsilon_0 c^3}\left[a^2 - |\boldsymbol{\beta}\times\mathbf{a}|^2\right].$$

For a fixed applied parallel electric force,

$$a_\parallel = qE/(\gamma^3 m).$$

The γ^6 cancels, so the ideal point-charge radiated power in a constant longitudinal field is independent of instantaneous speed. This conclusion keeps the field fixed. Holding coordinate acceleration fixed while increasing speed would require an increasing input force, producing a different power law. Always identify what is actually kept fixed before applying the exponents.

At fixed perpendicular magnetic field, $a_\perp = |q|vB/(\gamma m)$, so power varies as $\gamma^2 v^2 B^2$. A gyrating electron gradually loses energy if there is no replenishing electric field. Over a short tracked interval one can treat its radius and frequency as constant only after comparing radiated energy with kinetic energy. The large-mass dependence makes the corresponding correction much smaller for heavy particles at comparable speeds and charges.

Arrival time and a detector's angular window

Let radiation from emission position $\mathbf{r}_s(t_e)$ travel to a fixed distant receiver. Its arrival time is

$$t_o = t_e + |\mathbf{R} - \mathbf{r}_s(t_e)|/c.$$

Differentiating this relation gives

$$dt_o/dt_e = 1 - \mathbf{n}\cdot\boldsymbol{\beta}.$$

A forward-moving charge can deliver the energy emitted over a long source time in a much shorter arrival interval. Angular radiation patterns specified per emission time and a directional Poynting flux specified per arrival time then differ by this Jacobian.

Explore how observation direction and motion change the received field

At $\beta = 0.90$ directed toward a distant receiver, radiation from two source events separated by 5.0 ns arrives about 0.50 ns apart if the direction remains nearly unchanged. The same events viewed transversely arrive 5.0 ns apart. A slow photodiode could register comparable integrated pulse energies after geometric corrections while reporting very different peak shapes from the true fields. If the acceleration also changes direction during the interval, account for the angular antenna pattern before attributing the whole difference to arrival-time compression.

A finite-time mechanical energy balance

Poynting's theorem balances electromagnetic energy, its outward flux and work on charges. In a small region surrounding an idealised particle one can separate a far-field flux associated with radiation from energy temporarily stored with the near field. In the nonrelativistic, weak-reaction expansion the leading self-force is $m\tau_0\dot{\mathbf{a}}$, where $\tau_0 = q^2/(6\pi\epsilon_0 mc^3)$. Dotting with velocity and applying the product rule gives

$$\mathbf{F}_{\mathrm{self}}\cdot\mathbf{v} = \frac{d}{dt}(m\tau_0\mathbf{a}\cdot\mathbf{v}) - m\tau_0 a^2,$$

$$P_{\mathrm{rad}} = m\tau_0 a^2.$$

Adding the external mechanical work shows that $W_{\mathrm{ext}} = \Delta K + E_{\mathrm{rad}} - [m\tau_0\mathbf{a}\cdot\mathbf{v}]_{t_1}^{t_2}$. This last boundary term accounts for reversible energy exchange with the bound field in this separation. It vanishes if acceleration is zero at both endpoints, or over complete periods when the mechanical state repeats. During a limited interval positive radiated energy and instantaneous mechanical damping need not coincide.

A prescribed interval of constant acceleration illustrates the distinction. Within that interval $\dot{\mathbf{a}} = 0$ and the local leading self-force vanishes, yet the far-zone power is positive. The endpoint term changes because $\mathbf{a}\cdot\mathbf{v}$ grows with time. A physical pulse has switching intervals and any comparison of initial and final energies must include them. Treating an endlessly accelerated point charge as a mechanical system with ordinary finite reservoir energies introduces boundary questions absent in a compact laboratory pulse.

Compare emitted energy, mechanical work and the endpoint contribution in a pulse

When a weak-reaction or classical model stops answering the question

The electron time τ_0 is approximately 6.27×10^{-24} s. For a driving angular frequency ω the ratio $\tau_0\omega$ estimates the size of the leading reaction force relative to inertial force in a harmonic nonrelativistic problem. At 100 MHz it is 3.94×10^{-15}, tiny per cycle. Many cycles can sum to an appreciable energy loss when all other damping mechanisms are very weak. For a nonrelativistic harmonic oscillator the cycle-averaged Larmor loss equals $\tau_0\omega^2$ times its stored mechanical energy, giving an energy lifetime $1/(\tau_0\omega^2)$.

A perturbative approximation replaces acceleration inside the small self-force by the acceleration from the external force to leading order. For a prescribed, smooth, nonrelativistic $F_{\mathrm{ext}}(t)$, this yields

$$ma \simeq F_{\mathrm{ext}} + \tau_0\dot{F}_{\mathrm{ext}}.$$

Errors scale with further time derivatives. Treating the unexpanded third-order point-particle equation as an arbitrary initial-value solver admits a homogeneous exponentially growing acceleration. That mode varies on a timescale the weak-reaction expansion did not resolve in the first place.

For a collective source also check dimensions against wavelength. A cloud much smaller than the wavelength can add fields coherently when all charges have the same driven phase. Then total power can scale as the squared number of radiators, while random phases usually leave only a contribution proportional to their number after averaging. At larger dimensions include the spatial Fourier transform of

the source. Finally, photon recoil and quantisation can matter even when a classical average flux describes an ensemble. Compare $\hbar\omega$ with the mechanical energy change that the measurement actually resolves.

Recover a scattering cross section from an energy balance

A free nonrelativistic electron in a weak nearly monochromatic incident wave is accelerated by the electric field. Let its peak incident field be E_0, assume a quiver speed small compared with c, and ignore the leading magnetic force in calculating the quiver. Then $a_0 = |q|E_0/m$. The time-averaged Larmor power divided by incident time-averaged intensity $\epsilon_0 c E_0^2/2$ gives a quantity with the dimensions of area. The amplitude cancels:

$$\sigma_T = \frac{q^4}{6\pi\epsilon_0^2 m^2 c^4} = \frac{8\pi}{3} r_e^2,$$

$$r_e = \frac{q^2}{4\pi\epsilon_0 m c^2}.$$

For an electron $r_e = 2.82$ femtometres and $\sigma_T = 6.65 \times 10^{-29}$ square metres. The radius here is a combination of charge, mass and speed of light, not an independently measured mechanical size of the electron. For an incident intensity of 100 watts per square metre the weak-drive prediction is 6.65×10^{-27} W per free electron. Calculating an expected photon count requires wavelength, collection solid angle and detector efficiency as well as that power.

There is a subtle question about work done by the incident wave. At lowest order the acceleration is exactly proportional to the applied field and the velocity lags it by a quarter cycle, giving zero average $q\mathbf{E}\cdot\mathbf{v}$. Yet the radiation formula predicts positive time-averaged outgoing energy. The resolution is an order count. Radiation power is already of the order of the small reaction correction, and an input-work calculation to that same order must include the correction to velocity that is in phase with the drive. The perturbatively reduced self-force introduces that component. Equivalently, far from the particle the incident–scattered interference contributes extinction equal to scattering in the absence of absorption.

For a charge bound by a spring, the quiver changes near and below its resonance. Well below resonance the displacement is nearly independent of driving frequency, so radiation power falls with the fourth power of frequency at fixed field. Far above resonance, but well within the classical low-recoil regime, the inertial free-electron response is recovered. This explains why applying the Thomson cross section indiscriminately to all the electrons in a transparent dielectric can give a misleading attenuation. Bound response, correlations, interfaces and coherent forward propagation can dominate a bulk optical measurement.

Count coherent radiators through a spatial form factor

Consider N nominally identical dipoles driven at positions $\mathbf{r}_j$ by a plane wave with wavevector $\mathbf{k}_i$. Their induced moments acquire incident phases $e^{i\mathbf{k}_i\cdot\mathbf{r}_j}$. Observation in direction $\mathbf{n}$ supplies the outgoing source factor $e^{-ik\mathbf{n}\cdot\mathbf{r}_j}$, so define the transfer wavevector $\mathbf{Q} = \mathbf{k}_i - k\mathbf{n}$. With equal amplitudes and a fixed observed polarisation, the field relative to one dipole is $\sum_j e^{i\mathbf{Q}\cdot\mathbf{r}_j}$. Squaring gives N self terms and ordered cross terms. If positions are independently drawn from a normalised spatial distribution $w(\mathbf{r})$,

$$\left\langle \left| \sum_j e^{i\mathbf{Q}\cdot\mathbf{r}_j} \right|^2 \right\rangle = N + N(N-1)|\widehat{w}(\mathbf{Q})|^2.$$

At exact forward scattering $\mathbf{Q} = 0$ and all phases add regardless of source extent. Away from the forward direction a broad cloud can lose mutual coherence even though its drive remains perfectly monochromatic. The total power integrated over a sphere need not follow the on-axis scaling, because the enhanced forward cone narrows as the cloud grows. At sufficiently large density the incident field on one dipole includes scattered fields from others, invalidating the independent-dipole derivation. The formula alone cannot predict arbitrarily large emitted power from an optically thick weakly fed sample.

For a Gaussian cloud with one-coordinate standard deviation s,

$$|\widehat{w}|^2 = e^{-s^2 Q^2},$$

$$Q = 2k\sin(\theta/2).$$

Take wavelength 600 nm, $s = 5.0$ micrometres and $N = 100$. At one degree, $sQ = 0.914$ and the averaged intensity factor is about 4.39×10^3 rather than either 100 or 10,000. At five degrees $sQ = 4.57$; the coherent excess becomes negligible and the factor approaches 100. An aperture spanning both these angles requires an angular integral, not the central value multiplied by its nominal solid angle.

One way to measure this distinction uses two nominally identical clouds with different sizes but equal independently counted particles. Record angular intensity through a series of apertures, keeping polarisation and illumination profile fixed. A growing forward peak with a narrowing angular width points toward a coherent form factor. If broad-angle intensity changes too, check whether the local density modifies each dipole's response or the sample heats. Compare a phase-sensitive measurement of forward extinction with a calibrated total scattered power rather than treating the forward peak as a direct measure of lost beam energy.

An energy audit of the complete apparatus must include the drive. If radiation from N synchronised elements is substantially larger than the sum for isolated elements, maintaining specified oscillations can demand additional work from the source through mutual fields. If the source instead supplies fixed power, the element amplitudes generally change. The external energy supply, field distribution and measurement solid angle are all part of the physical question. This accounting links a many-source optics measurement back to the finite-time mechanical work balance for an individual charge.

Relate an accelerating source to the energy it radiates

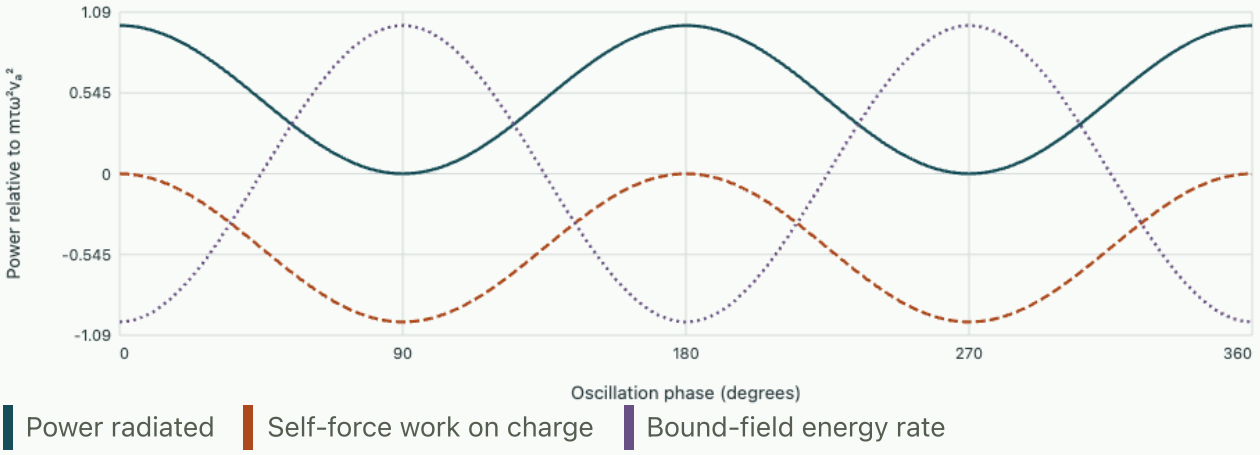
Change the time scale of a smooth motion. Compare instantaneous power with the integrated balance over the complete event.

Open this experiment online

Try the calculation

Mean velocity / oscillating speed 0

Emission phase (degrees) 0



At 0°, radiation, self-force work and bound-field rate are 1, 0, -1 in the plotted units. Their sum is zero.

Worked detector problem · energy in a cone around a dipole axis

A nonrelativistic charge acceleration pulse has known total radiated energy 1.00 pJ. An ideal detector collects the entire cone of half-angle 15° about the direction of acceleration, at a distance large compared with the source and the radiation wavelength. Estimate the recorded energy. Compare it with a detector of the same small solid angle centred perpendicular to the acceleration.

▼ **Integrate the projected electric field before applying the photon count**

For acceleration with fixed direction during the pulse, the radiated energy per unit solid angle is proportional to $\sin^2\theta$. The axial fraction is

$$(3/4)\int_0^{\alpha}\sin^3\theta\,d\theta = (3/4)(2/3-\cos\alpha+\cos^3\alpha/3).$$

At fifteen degrees this gives about 8.61×10^{-4} and a collected energy of 0.861 fJ. The small-angle approximation $3\alpha^4/16$ gives 0.881 fJ, consistent at this angle. A detector that assigns its on-axis zero intensity to the whole cone would incorrectly predict no signal.

At right angles, intensity varies slowly over a narrow aperture. Its leading fraction is $3\Omega/(8\pi)$, where $\Omega = 2\pi(1-\cos 15^\circ)$. This gives about 0.0256 or 25.6 fJ, with a small finite-angle correction. For an exact transverse cone average n_z^2 in coordinates about the detector axis. The ring average is $\sin^2\delta/2$ at polar offset δ, so the exact fraction becomes $(3/8)[(1-\cos\alpha)+(1-\cos^3\alpha)/3]$, or approximately 0.0251. Equal solid angle alone does not imply comparable collected energy.

These estimates assume unit efficiency across the entire pulse spectrum and a stable acceleration direction. A real detector with frequency-dependent response measures an integral over frequency and collection angle. Calibrate both before scaling its reading to total pulse energy. If the acceleration changes direction during the event, integrate the angular pattern appropriate to each emission time instead of assigning all the energy to a single measured peak direction.

A radiofrequency oscillator's finite lifetime creates a spectral width

An energy loss mechanism can be inferred from a time trace and independently tested in the frequency domain. Suppose an isolated weakly damped oscillator loses mechanical energy exponentially with time constant τ_E. Its position amplitude then falls as $e^{-t/(2\tau_E)}$. Near a well-resolved oscillation frequency its Fourier amplitude has a simple pole shifted below the real axis by $1/(2\tau_E)$. The squared Fourier amplitude is Lorentzian, with full width in ordinary frequency $1/(2\pi\tau_E)$.

If the measured energy lifetime in the earlier trapped-electron example were limited by its predicted radiation alone, the resonance width would be roughly four tenths of a microhertz. A measurement lasting one second cannot directly resolve such a line through the uncomplicated Fourier width of that record. Truncating a long-lived signal after that interval multiplies it by the acquisition window and convolves its spectrum with a much broader instrumental response. A broad fitted peak from the short record is consequently insufficient evidence of a short physical lifetime.

Compare records of several durations and perform a phase-coherent fit where appropriate. Technical drive-frequency jitter broadens an ensemble spectrum while a single undriven decay can remain long. Additional true energy dissipation shortens the ringdown itself. Pairing a spectral-width measurement with a time-domain energy measurement thus distinguishes loss of phase coherence from loss of oscillator energy, two effects that simple line shapes alone often confound.

Tutorial · Does the field energy explain a measured ringdown?

A nonrelativistic electron is driven at 100 MHz in a harmonic potential. After the drive ceases, the measured energy decays with time constant 0.10 s. Calculate the radiative energy lifetime, the energy fraction radiated in 0.10 s under radiation alone and whether a measured exponential proves radiative damping.

▼ **Construct and compare the decay rates**

Using $\omega = 2\pi 10^8$ rad/s and the electron τ_0, the radiation-only energy decay rate is 2.47×10^{-6} per second and its time constant is 4.04×10^5 s, about 4.68 days. The fraction radiated in 0.10 s is $1 - \exp(-2.47 \times 10^{-7}) \simeq 2.47 \times 10^{-7}$. The observed decay rate is about four million times larger. Under the assumptions of this single-particle free-space model it calls for further coupling or dissipation.

The amplitude decay time of an oscillator is twice its energy decay time. Confirm whether the measuring circuit reports amplitude, squared amplitude or extracted energy before fitting. A conducting environment can modify the density of radiation modes, and a measurement circuit can add a load. Compare separate

runs with different coupling to that circuit and compare a calculated angular power with the collected signal. A single exponential identifies a rate, not a physical mechanism.

Tutorial · Extract a finite-pulse balance from an acceleration trace

During a short nonrelativistic interval the acceleration is $a(t) = a_0 t/T$, the velocity initially vanishes and motion is one-dimensional. Calculate the radiated energy and integrated self-force work over that interval. Then explain what additional information an experiment needs to compare final mechanical energy with radiation after the entire pulse.

▼ **Retain the endpoint before interpreting the sign**

Integrating the acceleration gives

$$v(t) = a_0 t^2/(2T).$$

The radiated energy is $m\tau_0 a_0^2 T/3$. At the final endpoint $m\tau_0 a v = m\tau_0 a_0^2 T/2$. Thus integrated self-force work is the boundary term minus radiated energy, $+m\tau_0 a_0^2 T/6$. A direct integration of $(m\tau_0 a_0/T)v(t)$ gives the same positive answer, despite an outward flux throughout the ramp.

At T the acceleration is still nonzero. A smooth continuation back to zero changes both the radiated energy and boundary contribution. To compare initial and final states without a bound-field endpoint term, measure through that continuation and wait for the pulse to reach the collecting surface. Specify angular coverage and collection efficiency if only a fraction of the radiation reaches the apparatus. An on-axis detector along the acceleration could see no leading dipole signal even though the sphere integral is positive.

Cambridge Electromagnetism · accelerating charges and radiation

Waves, radiation and relativity · 13. How can a field that decays in air transmit power to another prism?

Place a thin air gap between identical prisms. Derive the TE transmission from the two interface conditions and test when an exponential attenuation length predicts the actual throughput.

Propagate both boundary data through the gap

A monochromatic TE wave in a prism of index n has tangential wavenumber $k_x = nk_0 \sin\theta$, normal wavenumber $k = nk_0 \cos\theta$, and $k_0 = 2\pi/\lambda$. Above the critical angle the air equation becomes

$$E'' = \kappa^2 E,$$

where $\kappa = (k_x^2 - k_0^2)^{1/2}$. Matching E and its derivative across each nonmagnetic interface retains both independent exponential solutions. Discarding the growing solution would preclude matching the outgoing wave in the second prism.

From the exit to the entrance the transfer matrix has diagonal entries $\cosh\kappa d$ and off-diagonal entries $-\sinh(\kappa d)/\kappa$ and $-\kappa\sinh(\kappa d)$. Insert (t, ikt) at the exit and $(1+r, ik(1-r))$ at the entrance. Eliminating r and taking the squared modulus gives the power transmission:

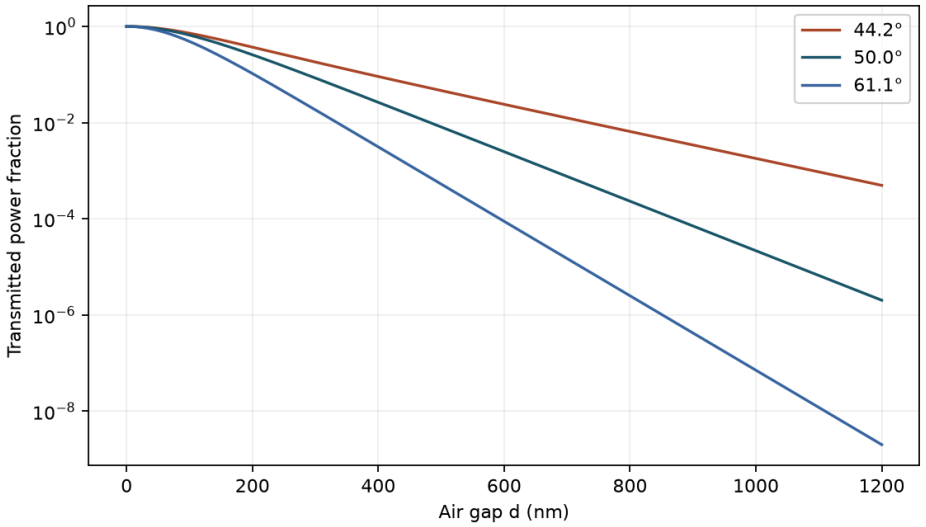
$$T = \left[1 + \frac{(k^2+\kappa^2)^2}{4k^2\kappa^2}\sinh^2(\kappa d)\right]^{-1}.$$

Both exterior media are identical, so transmitted power fraction equals squared transmitted amplitude. The finite flux through the gap comes from interference between its two evanescent components. The single exponential at an isolated interface carries no average normal power.

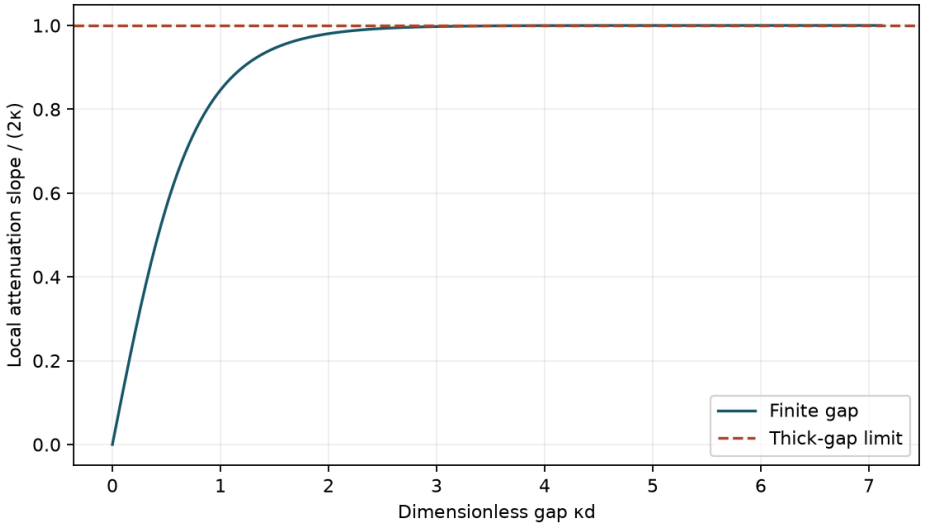
The penetration length sets a slope

At a large gap the logarithmic slope $-d\ln T/dd$ tends to 2κ. At contact $T = 1$ with zero slope, because the leading correction in gap thickness is quadratic. Measuring a straight-line slope over a very small gap would therefore return a biased estimate of the penetration depth. Use the first graph to choose the window for a fit and the second graph to assess the bias. Take a vacuum wavelength of 600 nm, index 1.50 and incidence angle 50°. The critical angle is 41.81°. The amplitude penetration length $1/\kappa$ is 168.72 nm and the coefficient of $\sinh^2(\kappa d)$ is 1.312. For gaps of 50, 100 and 200 nm the transmitted fractions are 0.894, 0.659 and 0.257. Doubling the gap therefore does not simply square the measured transmission. Only the sufficiently thick-gap logarithmic slope approaches a constant.

One can invert a single transmission reading when the angle and indices are independently known. Write $A = (k^2+\kappa^2)^2/(4k^2\kappa^2)$. Then $d = \kappa^{-1}\,\mathrm{asinh}\,\sqrt{(T^{-1}-1)/A}$. The reading 0.659 returns a gap of about 100 nm. Near $T = 1$ a small absolute power error makes a large relative error in the inferred gap. This follows directly from the quadratic expansion $1 - T \simeq A(\kappa d)^2$.



A higher transverse wavenumber causes a steeper thick-gap decay. At contact every angle has unit transmission.



Normalising by twice the decay constant turns the thick-gap limit into one. The early points reveal why a thin-gap slope fit fails.

Exercise · halve a thick-gap signal

How much extra separation is required to halve the transmitted power deep in the exponential regime?

▼ **Work through the solution**

The ratio tends to $\exp(-2\kappa\Delta d)$. Setting it to one half gives

$$\Delta d = \ln 2/(2\kappa).$$

This uses a power attenuation length. Confusing amplitude and intensity introduces a factor of two.

Exercise · choose a boundary at the far side

What changes if the second prism is removed while the incident prism and angle remain the same?

▼ **Work through the solution**

The gap becomes a half-space and boundedness removes the growing component. The transmitted evanescent amplitude can still be finite at the interface, while the average normal power flux becomes zero. The limit of the two-prism transmission agrees as d tends to infinity.

Try it · Evanescent gap

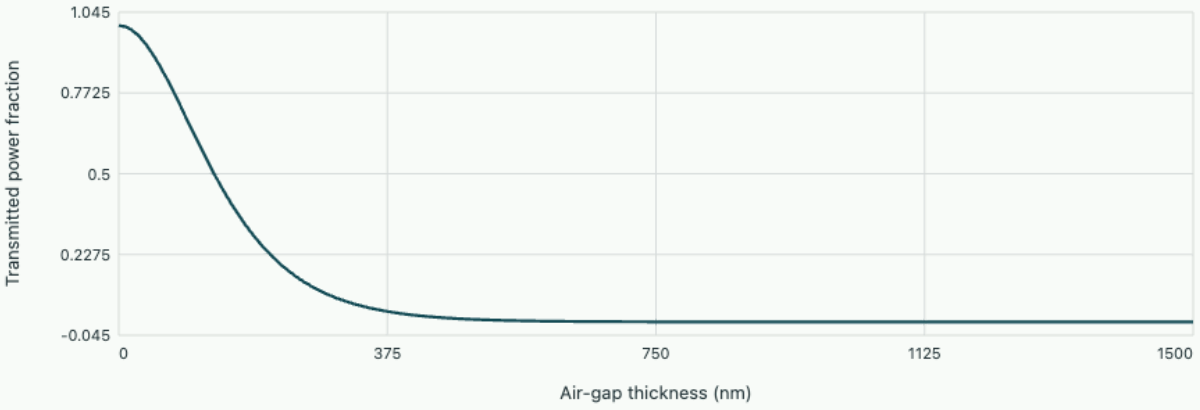
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Incidence angle (degrees) **50**

Wavelength (nm) **600**



Frustrated total internal reflection

Symmetric TE gap with surrounding index 1.5. Amplitude decay length 168.72 nm; critical angle 41.810°. At zero gap transmission is one.

Download the corresponding editable notebook. The browser recalculates the reduced equations directly; the notebook exposes the full calculation and its checks.

Open editable notebook Download .ipynb

Waves, radiation and relativity · 14. Resolve the self-force before taking a point-charge limit

A finite spherical charge distribution introduces a propagation time. Follow its memory kernel to electromagnetic inertia, radiative recoil and the limits of a truncated equation.

Give the charge a specified size and mechanical support

Take a thin spherical shell of radius a carrying fixed uniform total charge q. Prescribe rigid translational motion at speed much smaller than c, and linearise the self-field in that velocity. Mechanical stresses maintain the spherical distribution. Rotation, relativistic contraction and deformable internal charge modes lie outside this model. A finite radius is essential because two different surface elements communicate after a finite delay.

In the linearised retarded-field integral, the instantaneous longitudinal Coulomb pair forces cancel in the net force when using Coulomb gauge. The transverse inertial kernel contains the angular projector I–nn. Spherical averaging gives 2I/3. The resulting vector force is expressible as a pair-distance integral of the retarded acceleration.

$$\mathbf{F}_{\mathrm{self}}(t) = -\frac{2}{3c^2}\iint \frac{dq\,dq'}{4\pi\epsilon_0 R}\,\mathbf{a}(t-R/c).$$

This is the linear, rigid, spherically symmetric kernel. An independently incident electromagnetic field belongs to the applied force. A deformable shell would require an additional internal response model. The delayed time appears inside the acceleration. Pulling it out before averaging would retain electromagnetic inertia but remove the radiative correction. Here R is the distance between two surface points, evaluated in the unperturbed shell geometry.

Evaluate the pair distribution geometrically

Fix one point on the shell. A second point at polar angle θ from it is separated by R=2a sin(θ/2). The fraction of surface area in the annulus is sinθ dθ/2. Since R dR=a² sinθ dθ, the normalised distribution of pair distances is R dR/(2a²) on 0≤R≤2a. This cancels the kernel's inverse distance. Integrating the acceleration over the remaining retarded-time interval gives the exact linear delay formula for this model.

$$\mathbf{F}_{\mathrm{self}}(t) = \frac{q^2}{12\pi\epsilon_0 a^2 c}\,[\mathbf{v}(t-2a/c)-\mathbf{v}(t)].$$

Recall · A delay equation needs a history

A second-order particle equation is normally supplied with position and velocity at one instant. The shell equation samples velocity over the preceding interval of length 2a/c. A compatible history and the incoming electromagnetic field are additional information. Replacing that history by extra arbitrary derivatives changes the initial-value problem.

Separate inertia from the first dissipative correction

If the acceleration varies on a time scale T much longer than 2a/c, Taylor-expand the delayed velocity. The first term is proportional to acceleration and the second to its derivative.

$$\mathbf{F}_{\text{self}} = -m_{\text{em}}\mathbf{a} + \frac{q^2}{6\pi\epsilon_0 c^3}\dot{\mathbf{a}} + O(q^2 a\ddot{\mathbf{a}}/(\epsilon_0 c^4)),$$

$$m_{\text{em}} = \frac{q^2}{6\pi\epsilon_0 a c^2}.$$

The electrostatic shell energy is U=q²/(8πε₀a), so this electromagnetic inertial coefficient is 4U/(3c²). The coefficient refers to the electromagnetic part of a constrained shell. Covariant energy-momentum accounting also includes the stresses that maintain its shape. Identifying this isolated coefficient with the complete rest mass would omit those stresses.

Move the inertial term to the left of m_bare a=F_ext+F_self and define the measured inertial mass m=m_bare+m_em. The remaining leading local term is the Abraham–Lorentz force mτ dot a, with τ=q²/(6πε₀mc³). In the limit a→0 at fixed q the electromagnetic inertia diverges. Holding the measured mass fixed then requires a singular adjustment of the mechanical parameter. The finite-shell model therefore does not justify an unrestricted point limit.

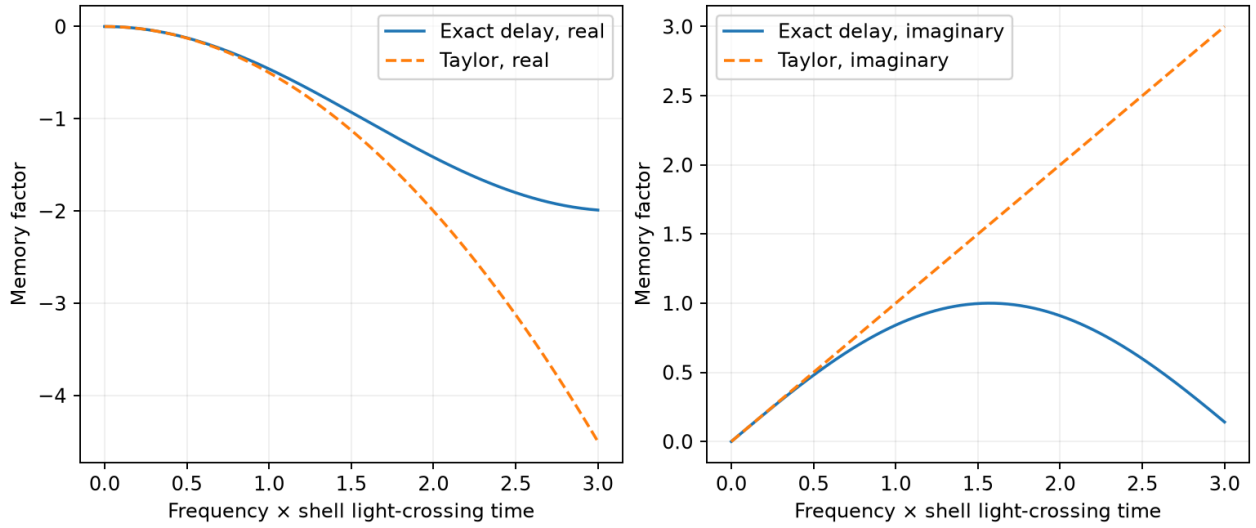
Check the mechanical power and the stored near field

Dot the recoil term with velocity and use the product rule. The first term is a total derivative. The second is the Larmor power with a minus sign.

$$m\tau\dot{\mathbf{a}}\cdot\mathbf{v} = \frac{d}{dt}(m\tau\mathbf{a}\cdot\mathbf{v}) - m\tau a^2.$$

Defining the near-field energy contribution E_S=−mτa·v gives d(K+E_S)/dt=F_ext·v−P_rad. Over a pulse whose endpoint acceleration vanishes, the integrated near-field change is zero. At an intermediate instant it can be nonzero, so mechanical power loss and radiated power need not agree instant by instant.

> **Remember · A derivative expansion has a frequency range**
>
> For harmonic velocity v∝e^{−i}ωt, the exact delay factor is e^{i}ω(2a/c)−1. Its Taylor expansion is controlled by |ω|2a/c. Solving the truncated differential equation can create modes outside that range. Check the frequency of a proposed solution against the approximation that produced the equation.



The low-frequency approximation agrees near the origin. At order-one light-crossing phase the exact delay and its truncated differential approximation differ substantially.

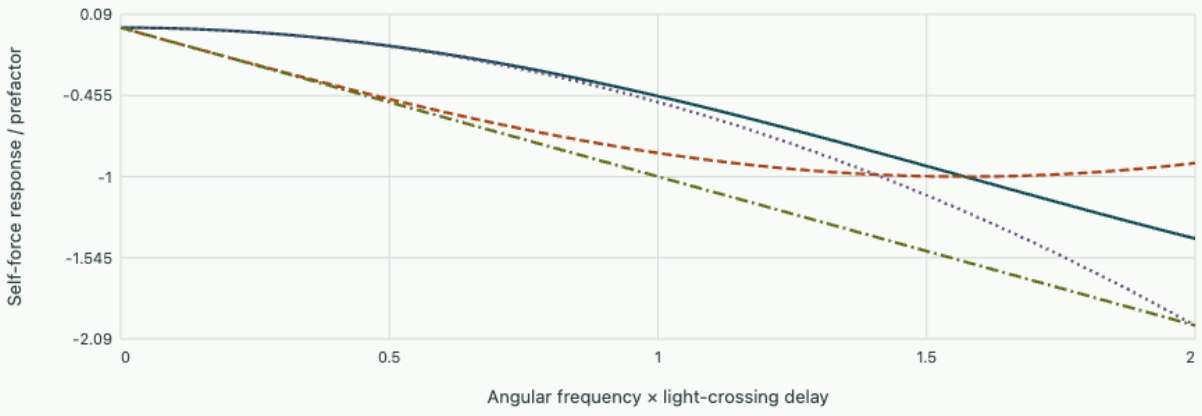
Try it · Delay force

Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Maximum light-crossing phase 2



For velocity proportional to exp(iωt), the delayed shell factor is exp(–iωτ)–1 with τ=2a/c. The differential approximation requires |ωτ| ≪ 1.

Reduce the order on the physical slow branch

The truncated force-free equation $a=\tau$ dot a admits an exponentially growing acceleration. Its time scale is τ, while the local expansion was intended to describe slowly varying applied forcing. On the perturbative branch replace a on the small correction's right-hand side by F_ext/m. To first order in τ this gives m $a=F_{ext}+\tau$ dF_ext/dt, with a total derivative along the leading trajectory.

For a harmonic restoring force $F_{ext}=-m\omega_0^2x$, order reduction gives $\ddot{x}+\tau\omega_0^2\dot{x}+\omega_0^2x=0$. The amplitude decay rate is $\tau\omega_0^2/2$. The average Larmor power divided by oscillator energy gives the same energy decay rate $\tau\omega_0^2$. This comparison tests both the sign and the factor two between amplitude and energy.

Exercises on limits and energy

1. For an electron, $\tau\approx6.27\times10^{-24}$ s. Estimate the amplitude-decay time for $\omega_0=10^{15}$ s^{-1} within the classical oscillator model.

▼ **Work through the solution**

The amplitude rate is 3.135×10^{6} s^{-1} and the time is 3.19×10^{-7} s. The small parameter $\tau\omega_0$ is 6.27×10^{-9}. A specific atomic transition additionally requires quantum state and coupling information.

---

2. Evaluate the ratio of the exact shell memory factor to its first-order inertial approximation when $\omega 2a/c=0.1$. Which next term captures loss?

▼ **Work through the solution**

The exact factor is $\cos 0.1-1+i\sin 0.1$. Relative to $i0.1$ it is approximately $0.998334+i0.049958$. The real negative term $-(\omega 2a/c)^2/2$ in the force factor gives the first dissipative correction. Keeping only the linear imaginary term gives reactive inertia.

---

3. A force has compact support and the acceleration vanishes at both endpoints. Show that net radiated energy equals external work minus the change in kinetic energy.

▼ **Work through the solution**

Integrate the balance for $K+E_S$. The endpoint values of E_S vanish, leaving $W_{ext}-\Delta K=\int P_{rad}\,dt$. Truncating the interval before acceleration vanishes would require its near-field endpoint contribution.

For the finite-history model and its asymptotic limits, see the original analysis by Herbert Spohn. Radiative damping, a case study.

Waves, radiation and relativity · 15. From light clocks to collision kinematics

Construct spacetime measurements before transforming fields. Use the same invariant interval to compare clocks, add velocities and close a relativistic collision.

Make the two postulates operational

Assume inertial experiments obey the same physical laws and that vacuum light signals have the same speed c in each inertial frame. A light clock consists of two parallel mirrors separated transversely by distance d in its rest frame. One round trip takes Δτ=2d/c. An observer seeing the clock move at speed v measures a longer diagonal light path. For half of the round trip, the right triangle gives $(c\Delta t/2)^2=d^2+(v\Delta t/2)^2$. Solving yields Δt=γΔτ.

Here proper time means the reading of a clock carried along the two events. For arbitrary smooth timelike motion, split the trajectory into short inertial pieces and integrate.

$$d\tau = dt\sqrt{1 - v^2/c^2},$$

$$\Delta\tau = \int_{t_1}^{t_2} \sqrt{1 - v(t)^2/c^2}\, dt.$$

> **Recall · A reading at one place differs from comparing distant clocks**
>
> Proper time uses successive events at the same moving clock. Measuring the length of a moving rod instead requires its two endpoints at the same time in the observer's frame. Changing that simultaneity condition changes the pair of events used in the measurement.

Construct the Lorentz map and its diagram

Homogeneity makes the map between aligned inertial coordinates linear. The moving origin has x=vt, fixing x′=A(x−vt). Requiring each ray x=±ct also to satisfy x′=±ct′ fixes t′=A(t−vx/c²). Symmetry under reversing v determines A=γ. The y and z coordinates stay unchanged for this boost.

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}, \qquad (c\Delta t)^2 - \Delta x^2 = (c\Delta t')^2 - \Delta x'^2.$$

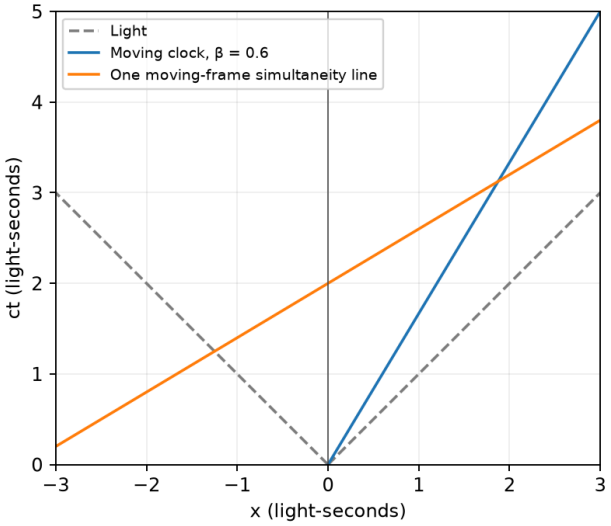
On a diagram with x horizontal and ct vertical, the moving time axis is x=βct. The moving simultaneity axis is ct=βx. A light ray always has slope ±1. For events separated by Δx=3 m and Δt=0, a frame with β=0.6 assigns Δt′=−7.505 ns. Their interval is spacelike, so their order can change without reversing the order of a causal signal. For timelike events, a subluminal boost preserves the temporal order.

A rod of rest length L_0 has endpoint separation $\Delta x'=L_0$. Its laboratory length uses Δt=0, giving $\Delta x'=\gamma\Delta x$ and $L=L_0/\gamma$. Requiring Δt′=0 instead would describe another pair of events. Differentiating the transformation gives the velocity composition rule.

$$u'_x = \frac{u_x - v}{1 - u_x v/c^2},$$

$$u'_y = \frac{u_y}{\gamma(1 - u_x v/c^2)}.$$

Two speeds of 0.8c in the same direction compose to 1.6c/1.64=0.97561c. Light still gives c. Introducing rapidity η through β=tanhη turns successive collinear boosts into addition of η, which is convenient when several observer changes are chained.



A moving time axis lies inside the light cone. Its simultaneity lines tilt relative to the laboratory horizontal.

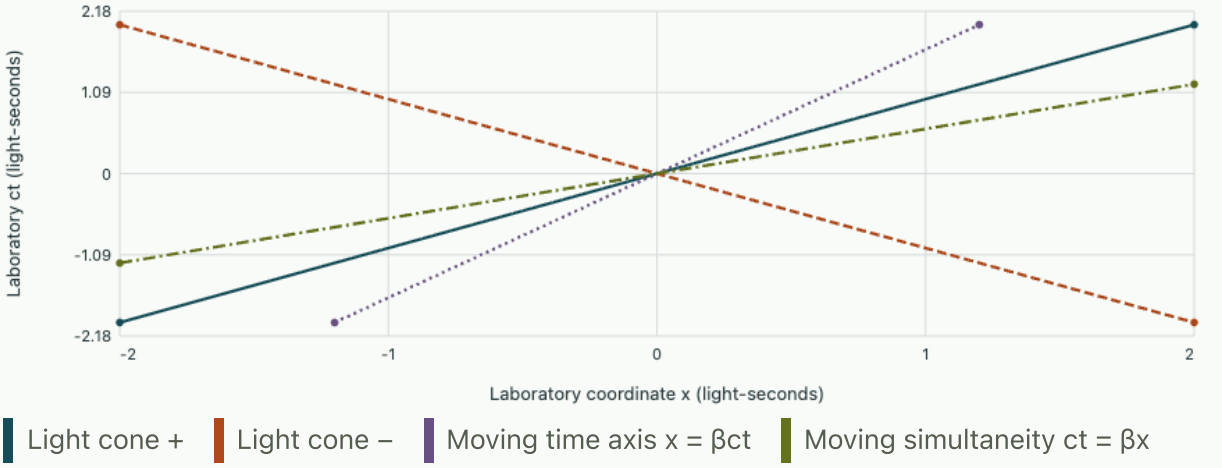
Try it · Spacetime

Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Observer speed v/c **0.6**



Lorentz factor γ = 1.25. Axes tilt with opposite roles; light directions remain invariant.

Compare clocks that meet again

In a chosen laboratory frame, let one clock remain at the origin while another travels out and back at speed 0.6c for a total laboratory duration of 10 years, with a short idealised turnaround. The stay-at-home clock records 10 years, and the travelling one records 8 years. Both readings can be compared at the same final event. The travelling clock uses two inertial frames separated by the turnaround. Its change of simultaneity prevents combining the outward and inward time-dilation arguments as if there were one travelling inertial frame.

Acceleration can be spread over a finite time and inserted into the proper-time integral. The measured age difference follows the whole path. Merely knowing a peak acceleration or a final relative speed does not determine it.

Derive the quantities that survive a collision

For a free particle use the invariant action $S=-mc^2\int d\tau$, giving $L=-mc^2\sqrt{(1-v^2/c^2)}$. Differentiate with respect to v to obtain $p=\gamma mv$. The Legendre transform $p\cdot v-L$ gives $E=\gamma mc^2$. Eliminate v to obtain the invariant mass

relation.

$$E^2 - c^2p^2 = m^2c^4,$$

$$K = (\gamma - 1)mc^2.$$

For an isolated collision add the incoming four-momenta and equate them to the outgoing sum. The invariant of the total four-momentum defines the centre-of-momentum energy, including energy that becomes internal excitation. The velocity of that frame along the total momentum is V=c²P/E_total.

As an example, two identical particles of mass m approach with speeds ±0.6c and stick. Each has γ=1.25. The total momentum vanishes, and the final invariant mass is 2.5m. The difference from 2m is stored internal energy 0.5mc². In another inertial frame the composite moves, but its invariant mass is the same.

Find a production threshold with one stationary target

Let an incident particle with mass m and total energy E hit an identical stationary target. Squaring the total four-momentum gives s=2m²c⁴+2mc²E. Consider endothermic production with minimum final total rest mass Mf≥2m. At threshold it has no relative kinetic energy in its centre-of-momentum frame, so s=Mf²c⁴.

$$E_{\text{threshold}} = \frac{(M_f^2 - 2m^2)c^2}{2m}.$$

If Mf=3m, the required incident total energy is 3.5mc² and its kinetic energy is 2.5mc². The newly created rest energy is only mc², but momentum conservation forces part of the available laboratory energy into overall recoil. A collider with oppositely directed equal momenta uses the incoming kinetic energy differently.

> **Remember · Use total energy when constructing a four-vector**
>
> The time component is E/c with E=K+mc². Substituting K/c would fail the invariant relation even for a particle at rest. Kinetic energy is convenient for a calorimeter, while total energy is required in the Lorentz transformation and threshold invariant.

Exercises with events and four-momenta

1. A moving clock at 0.8c records 3 μs between two flashes. Find their laboratory time and displacement.

▼ **Work through the solution**

γ=5/3 gives Δt=5 μs. The displacement is 0.8c times 5 μs≈1199.17 m. Direct substitution gives c²Δt²−Δx²=c²(3 μs)².

2. Two photons of energies 2 MeV and 1 MeV travel in opposite directions. Determine the invariant mass of their combined system and its centre-of-momentum velocity.

▼ **Work through the solution**

Total energy is 3 MeV and $cP = 1$ MeV. Hence $Mc^2 = \sqrt{8}$ MeV≈ 2.828 MeV and $V = c/3$ toward the more energetic photon. Each individual photon has zero rest mass.

---

3. At $\beta = 0.6$ a rod has laboratory length 0.80 m. Determine its rest length, then find the laboratory time difference for endpoint events simultaneous in the rod frame.

▼ **Work through the solution**

$L_0 = \gamma L = 1.00$ m. For $\Delta t' = 0$ and $\Delta x' = L_0$, the inverse transformation gives $\Delta t = \gamma v L_0 / c^2 = 2.502$ ns. Those events cannot be used directly for the laboratory length measurement.

Fields and the response of matter

Connect electric and magnetic measurements to relaxation times, carrier motion, thermal fluctuations and collective order. Use frequency, temperature and geometry to distinguish competing explanations.

1. ## Measure how quickly a dielectric follows an electric field

   Derive the time and frequency responses of a relaxing polarisation, locate the loss maximum and decide whether a measured sample contains more than one relaxation process.

2. ## Identify mobile charge with a transverse field

   Derive a Hall-voltage measurement from transverse force balance, reverse current and field to reject offsets and explore what fails when two carrier types conduct.

3. ## Measure equilibrium field fluctuations with a capacitor

   Derive an electrical noise measurement in two ways, determine what fraction reaches a finite-bandwidth detector and use agreement to calibrate an experiment.

4. ## When does a polar crystal change its preferred state?

   Derive stable and unstable dielectric responses from a free-energy landscape. Separate a reversible small-signal capacitance from a switching experiment.

5. ## Distinguish magnetic screening from electrical diffusion

   Compare the frequency dependence, phase and thickness dependence of field penetration in a normal metal and a superconducting slab.

6. ## When a periodic dielectric has no propagating Bloch wave

   Derive a band gap using two lossless transparent materials and distinguish evanescent Bloch propagation from absorption of optical power.

7. ## Why rotating a layered dielectric changes a capacitance measurement

   Derive the two principal permittivities of a fine laminate, rotate the constitutive tensor and distinguish the applied electrode condition from an assumed field direction.

8. ## How a magnetic field rotates the polarisation of transmitted light

   Resolve a linearly polarised wave into circular eigenmodes, infer a refractive-index difference from an analyser and distinguish nonreciprocal rotation from a simple change of reference axes.

9. ## Can a flat interface guide light without mirrors?

   Send light along a plane boundary between a dielectric and a metal. Seek a wave that propagates parallel to the interface while decaying away on both sides. The signs in the magnetic and electric boundary conditions determine whether any such wave can exist.

10. Infer charge discreteness from fluctuations in a steady photocurrent

A detector can deliver a perfectly steady mean current over a long interval while its charge per short exposure fluctuates. Derive both a counting prediction and its spectral version, then use a bandwidth-limited measurement to separate partition fluctuations from thermal and readout noise.

11. Why a nearby gate changes the speed law of a two-dimensional plasma

A current pulse excites a travelling density modulation in a thin conducting sheet. Move a metal gate towards the sheet and the shape of the dispersion curve changes. Rebuild the restoring electric field from the surface charge before using any oscillator analogy.

12. Can a capacitance measurement reveal the fraction of embedded particles?

A sample contains well separated dielectric spheres in an insulating matrix. Its capacitance changes when the particle fraction changes. Derive a dilute estimate from one sphere and a self-consistent approximation from a local-field model. Use the inverse calculation to discover when the fraction becomes poorly determined.

13. Can two ohmic layers imitate a relaxing dielectric?

Model a film and its contact layer in series. Derive the frequency scale of interfacial charge accumulation and decide which features of a measured spectrum can arise from geometry alone.

14. From molecular dipoles to a dielectric susceptibility

Derive induced and orientational polarisation separately, then determine when temperature or local fields invalidate a constant dielectric coefficient.

Fields and the response of matter · 1. Measure how quickly a dielectric follows an electric field

Derive the time and frequency responses of a relaxing polarisation, locate the loss maximum and decide whether a measured sample contains more than one relaxation process.

From a delayed polarisation to a constitutive equation

A slab of insulating material is placed between parallel electrodes. At first the measured charge increases very quickly when a small voltage is switched on. It then continues to rise slowly even though the applied voltage is fixed. A constant real relative permittivity can describe the final charge but cannot predict this record. The experiment reveals at least two relevant time scales. For a measurement window that resolves only the slower process, treat the fast electronic response as instantaneous and describe the slower polarisation P by its approach to an equilibrium value.

At fixed weak field the equilibrium slow polarisation is $\epsilon_0 \Delta\epsilon E$, with $\Delta\epsilon > 0$. Assume the rate of change is proportional to the distance from that equilibrium, with one positive relaxation time τ. Then

$$\tau \dot{P} + P = \epsilon_0 \Delta\epsilon E(t).$$

This is a phenomenological material law. Its assumptions include linear response about the prepared state, a time-independent τ, and a sample whose constituents can be described by one dominant relaxation. Testing those assumptions is part of the experiment.

For $E(t) = E_s$ switched on at zero from an unpolarised state, solve using the integrating factor $e^{t/\tau}$. The result is

$$P = \epsilon_0 \Delta\epsilon E_s (1 - e^{-t/\tau}).$$

After one τ, 63.21% of the slow response has developed; after 3τ, 95.02%. Measuring an absolute fraction requires subtracting the fast initial jump and knowing the eventual plateau. A long baseline drift can look like a second relaxation if those references are not controlled.

Derive the complex dielectric response

Now use peak phasors with time dependence $e^{-i\omega t}$. Differentiation contributes $-i\omega$, so

$$P/E = \epsilon_0 \Delta\epsilon / (1 - i\omega\tau).$$

Include the high-frequency relative permittivity ϵ_∞, where “high” means above the slow process but below additional electronic resonances. The real and imaginary parts of relative permittivity are

$$\epsilon' = \epsilon_\infty + \frac{\Delta\epsilon}{1 + (\omega\tau)^2},$$

$$\epsilon'' = \frac{\Delta\epsilon\, \omega\tau}{1 + (\omega\tau)^2}.$$

At fixed field amplitude the polarisation current is

$$\dot{P} = -i\omega P.$$

The mean electrical work per unit time and volume is

$$\tfrac{1}{2}\mathrm{Re}(E^{*}\dot{P}) = \tfrac{1}{2}\omega\epsilon_0\epsilon''|E|^2.$$

Loss therefore has positive ϵ'' in the chosen time convention. Using $e^{+i\omega t}$ reverses the sign of the imaginary part of permittivity but leaves mean absorbed energy positive when the current convention is changed consistently.

Set $u = \omega\tau$. Differentiating $u/(1+u^2)$ gives $(1-u^2)/(1+u^2)^2$. Hence ϵ'' peaks at $u = 1$, with height $\Delta\epsilon/2$, and the real part lies halfway between its low- and high-frequency plateaux. The peak identified here is in the imaginary *permittivity*. A plot of dissipated power at fixed field includes an additional factor of frequency and need not have a peak at the same location.

Tie a frequency sweep to a step test

Suppose a material has $\epsilon_\infty = 2$, $\Delta\epsilon = 6$, and a measured loss peak at 1 kHz. Then $\tau = 1/(2\pi\,1000) \simeq 159\ \mu$s. At the peak the relative permittivity is $5 + 3i$ in our convention. A time-domain voltage step should show a slow charge transient reaching 63.2% of its added charge in roughly 159 µs. This comparison connects two experiments that weight electrode offsets and parasitic impedances differently.

An independent shape check follows by eliminating u. Let $X = \epsilon' - \epsilon_\infty$ and $Y = \epsilon''$. From the formulas $X^2 + Y^2 = \Delta\epsilon X$. Thus $(X - \Delta\epsilon/2)^2 + Y^2 = (\Delta\epsilon/2)^2$: a single Debye relaxation traces a semicircle in the complex plane. Departures from a semicircle can reflect several relaxation times, spatially inhomogeneous samples, leakage or imperfect fixture correction. They do not uniquely identify the microscopic mechanism.

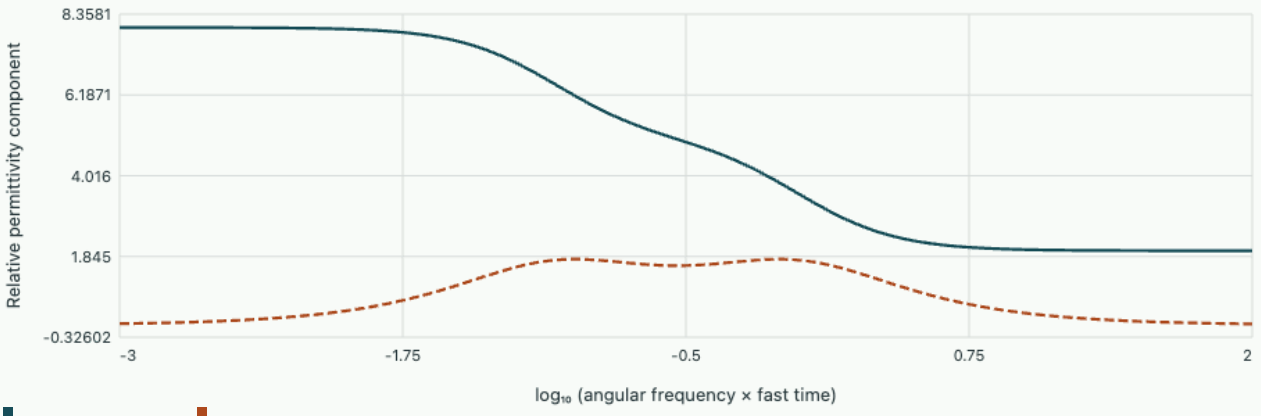
Resolve a second process only when the data require it

Open the live experiment

Try the calculation

Slow fraction of total dielectric strength **0.5**

Slow / fast relaxation time **12**



Largest sampled loss 1.7724 near ωτ_fast = 0.1. Relative permittivity tends to 8 at low frequency and 2 at high frequency.

The synthetic sample has two groups contributing a fixed total $\Delta\epsilon = 6$. Change the fraction in the slower group and its relaxation time relative to the faster group. At equal relaxation times the groups give exactly the single-process response whatever their fractions. When the times differ, the imaginary curves broaden and may separate. Observe the real dispersion steps as well. A broad lossy peak alone can be fitted by more than one pair of times and strengths.

In a practical measurement the low-frequency tail can contain DC conductivity. In the $e^{-i\omega t}$ convention,

$$\sigma E - i\omega\epsilon_0\epsilon E = -i\omega\epsilon_0[\epsilon + i\sigma/(\epsilon_0\omega)]E.$$

This added imaginary part is proportional to $1/\omega$, whereas the low-frequency side of an isolated Debye relaxation rises as ω. Change electrode separation and use a DC conduction measurement to challenge the interpretation. An electrode interface can add a further time scale, so an observed fit parameter deserves an uncertainty and a physical validation step.

Tutorial

The loss maximum is at 2 kHz and its height is 4 in relative-permittivity units. The high-frequency plateau is 3. Predict the low-frequency plateau and the slow response at $t = 100\ \mu s$ after a step.

▼ **Work through the solution**

The single-time model gives $\Delta\epsilon = 8$ and the low-frequency plateau 11. The relaxation time is $1/[2\pi(2000)] \simeq 79.6\,\mu$s. The slow fraction after 100 µs is $1 - e^{-100/79.6} \simeq 0.715$. Repeat at another step amplitude to check linearity.

At the loss peak, is the phase difference between the *total* displacement and the electric field always 45°?

▼ **Work through the solution**

The slow polarisation alone has a 45° phasor angle at $\omega\tau = 1$. The total displacement includes $\epsilon_\infty E$. Its angle is $\arctan[(\Delta\epsilon/2)/(\epsilon_\infty + \Delta\epsilon/2)]$, generally less than 45°.

Why does observing only a high-frequency tail constrain a ratio of parameters more tightly than the individual parameters?

▼ **Work through the solution**

For large $\omega\tau$,

$$\epsilon'' \simeq \Delta\epsilon/(\omega\tau).$$

Many values with the same ratio $\Delta\epsilon/\tau$ have very similar tails. Measure the peak or both plateaux to distinguish the two parameters.

Compare with material measurements

Inspect the orientational response and dielectric-loss treatment in MIT's material-properties course. Use the phasor convention of each source to predict the sign assigned to loss.

Fields and the response of matter · 2. Identify mobile charge with a transverse field

Derive a Hall-voltage measurement from transverse force balance, reverse current and field to reject offsets and explore what fails when two carrier types conduct.

Choose signs from a drawing

A rectangular conducting sample has its long axis along x, width w along y, and thickness t along z. A uniform magnetic field $B\hat{\mathbf{z}}$ is applied. An external circuit maintains conventional current $I > 0$ in $+x$. High-impedance transverse probes draw negligible current. Define the measured signal as $V_H = V(y = 0) - V(y = w)$. For uniform transverse field this convention gives

$$V_H = E_y w.$$

Reversing the voltmeter leads reverses the reported sign; writing a carrier sign without a lead convention is incomplete.

With one charge carrier species of number density n and signed charge q, the drift velocity satisfies

$$j_x = nqv_x = I/(wt).$$

If the transverse velocity vanishes, its magnetic force component is $-qv_xB$, opposed by qE_y. Thus $E_y = v_xB = j_xB/(nq)$. The resulting port measurement is

$$V_H = \frac{IB}{nqt}.$$

The width cancels for a uniform bar; changing the contact geometry or using a current source with a nonuniform injection profile can spoil that simplification near the terminals. The cancellation also exposes the sign. For electrons $q < 0$, positive conventional current means that the mobile charge drifts in $-x$. The inferred transverse field is negative for positive B in the chosen convention.

Reconstruct density and mobility from separate observations

Take $t = 0.50$ mm, $I = 5.0$ mA, $B = 0.20$ T and a transverse signal of -1.25 mV. The inferred signed Hall coefficient is

$$R_H = tV_H/(IB) = -6.25 \times 10^{-4}\ \mathrm{m^3/C}.$$

Using singly charged electrons gives

$$n = 1/(e|R_H|) \simeq 1.00 \times 10^{22}\ \mathrm{m^{-3}}.$$

A longitudinal voltage measurement, after correcting for probe separation, can yield the material conductivity. If $\sigma = 80$ S/m, the single-carrier mobility is

$$\mu = \sigma|R_H| = 0.050\ \mathrm{m^2/(V\,s)}.$$

These operations use different pieces of information. The longitudinal measurement alone gives only the product $n|q|\mu$. The transverse measurement supplies a sign and, in the one-carrier model, the density independently of mobility. This is an inverse experiment whose two observables have different sensitivities, not a more precise measurement of the same scalar.

One can derive the same result from a stationary Drude equation $0 = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) - m\mathbf{v}/\tau$. The longitudinal drag gives

$$E_x = m v_x/(q\tau);$$

force balance in y gives the transverse field just obtained. In a homogeneous single-carrier sample, $E_y/E_x = q\tau B/m$, the signed Hall angle tangent. At fixed sample and field, doubling I doubles both voltages in this linear regime and leaves their ratio unchanged. Significant current-dependent apparent density can indicate heating or nonlinear contacts.

Isolate the transverse voltage from a misaligned contact

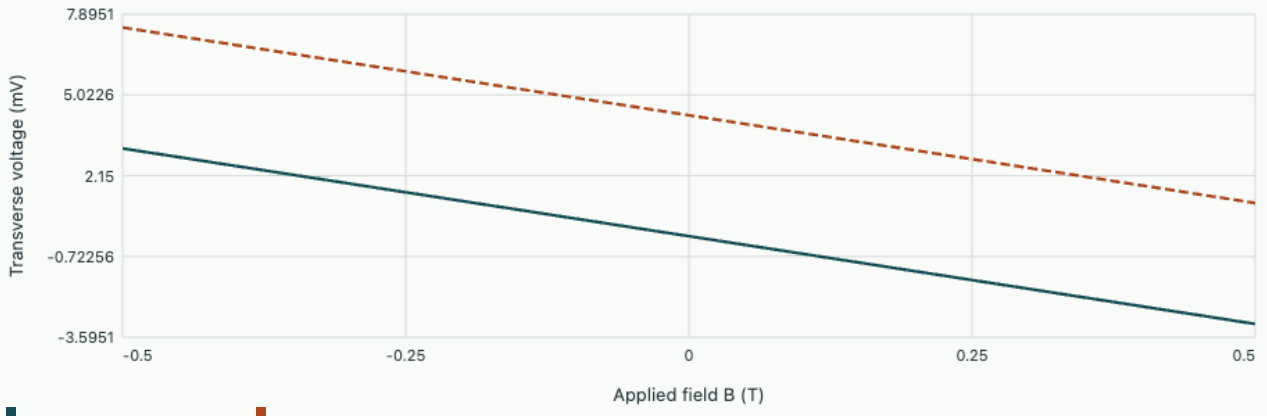
Practical side contacts seldom sit at exactly the same x. A fraction of the longitudinal voltage can leak into the transverse pair. Write the readout at a fixed current as $V_\mathrm{m}(B) = V_0 + \alpha V_{xx} + V_H(B)$. A background that is even in field cancels in $[V_\mathrm{m}(B) - V_\mathrm{m}(-B)]/2$. For example, readings of 3.05 mV and 5.55 mV at $+B, -B$ return -1.25 mV Hall signal and 4.30 mV of even background. Taking a single measured sign would have diagnosed the wrong carriers.

Open the live experiment

Try the calculation

Electron density (10^{22} m^{-3}) 1

Current (mA) 5



At +0.2 T the instrument reads 3.0517 mV; at −0.2 T it reads 5.5483 mV. Half-difference -1.2483 mV.

The plot deliberately superposes a constant 4.3 mV pickup on a signed Hall signal. Vary density and current independently before reading the pair at $B = \pm 0.2$ T. Large density weakens the signal; current increases it. Verify that field reversal recovers the correct transverse voltage even when both instrument readings have the same sign. If the unwanted background changes between readings, reverse the field several times and check for repeatable drift.

With both current and field reversal, the double-odd combination $[V(+I,+B) - V(+I,-B) - V(-I,+B) + V(-I,-B)]/4$ isolates terms odd in both. A thermal emf that is independent of current cancels along with a stable contact offset. This symmetry decomposition is a physical assumption about the parasitics, so preserve all four readings and inspect the rejected channels too.

When a Hall density stops being a particle count

With electrons and holes, longitudinal conductivities add. In a weak magnetic field their transverse contributions also depend on squared mobilities. For positive mobilities μ_e, μ_h and electron and hole densities n, p, the two-carrier weak-field coefficient in the sign convention above is $R_H = (p\mu_h^2 - n\mu_e^2)/[e(p\mu_h + n\mu_e)^2]$. To see where the square appears, note that one factor of mobility converts the longitudinal field into drift; a second converts the resulting transverse magnetic force into a drift contribution. Summing species currents and imposing zero total transverse current gives the denominator management through the summed longitudinal conductivity.

Equal electron and hole densities with equal mobilities give zero Hall signal while current still flows readily. A single-species inversion would suggest unbounded density. At equal densities but unequal mobilities, the faster species can set the sign even though neither species is more numerous. Additional field

dependence, temperature variation and independent structural information are needed before treating a Hall number as a carrier census.

Tutorial

A measurement doubles the width w at fixed thickness, total current and field in a uniform sample. Predict both E_y and V_H.

▼ **Work through the solution**

The current density halves, so v_x and E_y halve. The voltage integrates that halved field over twice the width and stays the same. If fixed current density were held instead, the transverse voltage would double.

A nominal thickness is overestimated by 5%. What bias does this cause in the single-carrier density inferred from a fixed transverse signal?

▼ **Work through the solution**

$n = |IB|/(et|V_H|)$, so using $1.05t$ gives $n_{\mathrm{fit}} = n/1.05$, approximately 4.8% low. This systematic error is common to all repeated field sweeps.

Why might a metal yield a smaller Hall voltage than a semiconductor with the same dimensions at the same current?

▼ **Work through the solution**

In the one-carrier regime the signal is inversely proportional to carrier density. A large density allows the imposed current at a very small drift speed, requiring a correspondingly small transverse balancing field. Conductivity alone cannot determine this comparison without mobility information.

Follow the transport model

Compare the one-carrier assumptions with the treatment of mobile carriers in David Tong's solid-state physics course. Connect an independently measured conductivity with the signed transverse response before interpreting an effective mass or scattering time.

Fields and the response of matter · 3. Measure equilibrium field fluctuations with a capacitor

Derive an electrical noise measurement in two ways, determine what fraction reaches a finite-bandwidth detector and use agreement to calibrate an experiment.

A zero-average voltage can store measurable energy

Connect a capacitor C across a resistor R at a uniform absolute temperature T, with no external drive. Ideal circuit equations predict that any deliberately deposited charge decays on time scale RC. At thermal equilibrium, microscopic fluctuations continually exchange charge and energy with the capacitor. Symmetry gives zero mean voltage, but the stored energy involves V^2. In the classical limit a quadratic degree of freedom has mean energy $k_B T/2$, giving $\langle V^2 \rangle = k_B T/C$. The resistor sets the time scale of equilibration; equilibrium mean-square voltage does not depend on resistance in this ideal model.

Check the dimensions: $k_B T$ is an energy in joules and capacitance has units coulombs per volt. Their ratio is volts squared. At $T = 300$ K and $C = 100$ pF the RMS voltage is $\sqrt{k_B T/C} \simeq 6.44\ \mu\mathrm{V}$. A voltmeter that averages the signed voltage for a long time reports a value near zero. Measuring the noise requires an RMS, variance or appropriately calibrated spectral measurement.

Derive the same variance from a spectral measurement

Use a one-sided voltage power spectral density per hertz at positive frequencies. The classical Johnson–Nyquist open-circuit result for a resistor is $S_R(f) = 4k_B T R$ V²/Hz. Its square root is an amplitude spectral density in V/$\sqrt{\mathrm{Hz}}$, not a voltage variance. Use an equivalent series voltage noise source and the RC transfer function $H(f) = 1/[1 + 2\pi i f RC]$. A stationary linear system multiplies the voltage power spectrum by $|H|^2$, so

$$S_C(f) = \frac{4k_B T R}{1 + (2\pi f RC)^2},$$

\int_0^\infty S_C(f)\,df=\frac{k_BT}{C}.

The integral follows by $u = 2\pi RCf$ and $\int_0^\infty du/(1+u^2) = \pi/2$. It furnishes a detailed check on the factor four, the use of one-sided density, the conversion from hertz to angular frequency and the amplitude-versus-power distinction. If your result disagrees by a constant factor of two or 2π, audit these conventions before assigning a new noise temperature.

For $R = 10$ kΩ, $T = 300$ K, the low-frequency amplitude spectral density is approximately 12.9 nV/$\sqrt{\mathrm{Hz}}$. The RC corner at $C = 100$ pF is $f_c = 1/(2\pi RC) \simeq 159$ kHz. The open-circuit low-frequency density rises with R while the bandwidth shrinks in inverse proportion. Their integrated product gives the earlier resistance-independent variance.

Include a real observation bandwidth

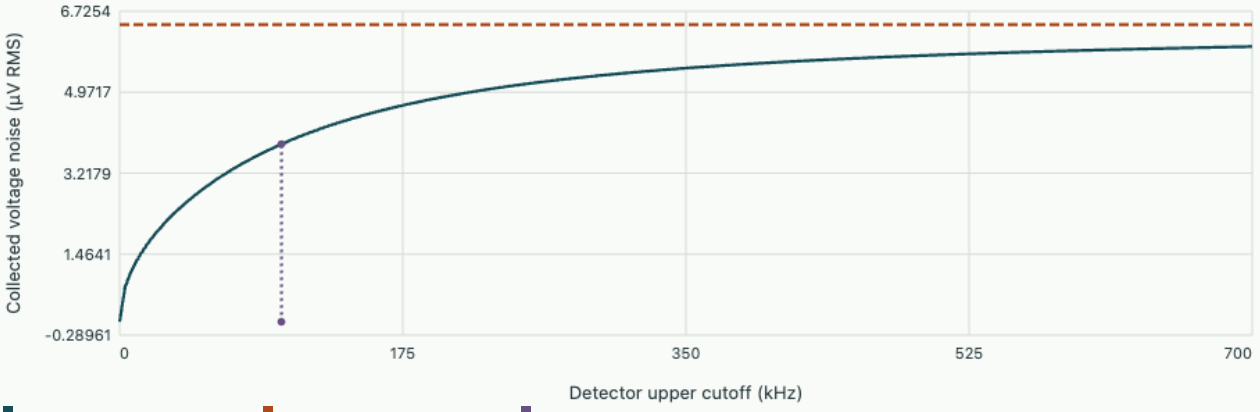
Suppose a first measurement adds an ideal low-pass cutoff B. The recorded fraction of equilibrium variance is $F = (2/\pi)\arctan(B/f_c)$. Half of the noise variance lies below the RC corner. That may be surprising if the corner was introduced as the frequency at which the *spectral density* reaches half its zero-frequency value. These two properties coincide for this single-pole spectrum for different mathematical reasons.

At $B = 0.1 f_c$, $F \simeq 0.06345$, so the recorded RMS is about 25.2% of the infinite-band RMS. At $B = 10 f_c$, $F \simeq 0.93655$, giving about 96.8% of the RMS. In an experiment, taking a square root one step too early changes this comparison. A real detector transfer function replaces the sharp cutoff by an integral of $S_C(f)\lvert H_{\mathrm{det}}(f)\rvert^2$.

Open the live experiment

Try the calculation

Resistance (kΩ)	**10**
Temperature (K)	**300**
Detector bandwidth (kHz)	**100**



RC corner 159.15 kHz. Detector reads 3.8461 µV RMS, collecting 35.713% of equilibrium variance.

Sweep the resistance by a factor ten. Observe how the corner moves along the frequency axis and how the fraction measured in a fixed bandwidth changes. Then vary temperature. In the classical regime spectral density and integrated variance are proportional to T, while RMS voltage grows only as $\sqrt{T}$. A plot labelled “noise” without units cannot distinguish these two dependencies.

Background · How does a longer voltage exposure alter the measured noise?

An equilibrium resistor–capacitor circuit has voltage covariance $C_v(s) = (k_B T/C)\exp(-|s|/\tau)$, where $\tau = RC$. Suppose a digitiser returns the mean voltage during an exposure T_m. Squaring a sample of this mean measures a different variance from averaging the instantaneous voltage squared. In the classical stationary model the two-time integral becomes

$$\operatorname{var}(\bar{v}) = \frac{2}{T_m^2}\int_0^{T_m}(T_m - s)C_v(s)\,ds = \frac{2k_BT}{C}\,\frac{u - 1 + e^{-u}}{u^2},$$

$$u = \frac{T_m}{\tau}.$$

Short exposures recover k_BT/C. At ten correlation times only about 18% of that variance remains. For very long exposures the variance decreases as $2\tau/T_m$. Taking more closely spaced samples of one exposure cannot undo that averaging.

A physical check and a calibration check

An RC time-domain experiment provides a second observable beyond integrated power. Under the same single-pole equilibrium model the voltage autocovariance is

$$\langle V(t)V(t+\tau)\rangle = (k_BT/C)e^{-|\tau|/(RC)}.$$

Its variance at zero lag agrees with equipartition, and its decay time is obtained from the homogeneous charge-relaxation equation. Repeating the measurement with a known extra parallel capacitance lowers the RMS and lengthens the correlation time. Seeing only one of those changes suggests that the observed noise may include substantial amplifier input voltage or an independent background.

The classical approximation assumes $hf \ll k_BT$ over the relevant band and a resistor that can be treated as frequency independent. A MHz room-temperature experiment is far inside this regime. At sufficiently high frequencies or cryogenic temperatures, a quantum noise model, an explicit definition of the measured ordering of the spectrum and the detector's energy exchange are needed. Extending an ideal classical white source to infinite frequency is only permissible as an approximation inside the convergent low-frequency RC calculation; it is not a microscopic model at arbitrarily high frequency.

Tutorial

At fixed C, T, doubling R in a detector with $B \ll f_c$ changes the measured RMS by what factor? Will the same factor apply when $B \gg f_c$?

▼ **Work through the solution**

At very small bandwidth, $S_C \simeq 4k_BTR$ is nearly constant, so variance doubles and RMS increases by $\sqrt{2}$. At large bandwidth the measurement approaches k_BT/C and becomes independent of R.

A two-sided spectrum in hertz is used instead. What density and integration limits preserve the same variance?

▼ **Work through the solution**

For this real voltage, the density for each sign of frequency is half the one-sided positive-frequency value. Integrate it over $(-\infty, \infty)$. Integrating the one-sided density over both signs would count the power twice.

A measured RMS is 10 µV, including independent amplifier noise of 6 µV RMS. Can 6 µV simply be subtracted to infer the sample contribution?

▼ **Work through the solution**

Independent zero-mean voltages add in variance. The inferred sample RMS is

$$\sqrt{10^2 - 6^2} = 8\ \mu\mathrm{V}.$$

A deterministic offset or correlated noise would require a different treatment.

Compare with an actual laboratory

The MIT Junior Lab noise experiment treats the detector as part of the measurement. Identify its calibration observable before comparing the theoretical density with an instrument reading.

Fields and the response of matter · 4. When does a polar crystal change its preferred state?

Derive stable and unstable dielectric responses from a free-energy landscape. Separate a reversible small-signal capacitance from a switching experiment.

Why one voltage can support several measured polarisations

Some insulating crystals develop a spontaneous electric polarisation as temperature is lowered. The high-temperature phase may be symmetric under inversion so that states with polarisation $+P$ and $-P$ are equivalent in zero field. Near a continuous transition, treat one component of polarisation as an order parameter. A local free-energy density compatible with the symmetry is $f(P) = a(T - T_c)P^2/2 + bP^4/4$ with $a, b > 0$. The omitted terms limit how far in temperature or field this truncated expansion is reliable. The positive quartic term keeps the energy bounded when the quadratic coefficient is negative.

For a device whose applied electric field E is controlled, the appropriate potential in this simple model is $g(P; E) = f(P) - EP$. The linear term records external work; changing what is held fixed changes the energy to minimise. Differentiation gives the constitutive equation $E = a(T - T_c)P + bP^3$. An algebraic root is an equilibrium candidate. A stable local response also requires positive second derivative $a(T - T_c) + 3bP^2$. A very accurate rootfinder can return an unstable stationary point if this second test is omitted.

Find the equilibrium branches and their response

Introduce scaled polarisation p, field e and temperature parameter s such that $g = p^4/4 + sp^2/2 - ep$. Then $e = p^3 + sp$. For $s > 0$, the slope $3p^2 + s$ stays positive. Each field has one stable solution. At $e = 0$, $p = 0$ and the small-signal susceptibility $dp/de = 1/s$ increases when approaching the ideal critical point from above.

For $s = -1,\ e = 0$, stationary values are $-1, 0, +1$. The middle root has curvature –1 and is a local maximum. Each outer root has curvature 2 and is a minimum. The susceptibility about a selected outer state is $1/2$ in the scaled units. The fact that the two domains can carry opposite zero-field polarisation does not make the harmonic response negative. The mere existence of a negative-slope segment in the cubic constitutive curve is insufficient to infer a stable isolated small-signal device with that differential response.

Derive the limit of a metastable homogeneous state

Imagine preparing the negative-polarisation state at $s = -1$, then increasing the applied field slowly. A positive field lowers the energy of the positive state. The negative state can persist locally behind a barrier even after it ceases to be the global minimum. A model of perfectly homogeneous evolution without early nucleation follows this metastable branch until the curvature vanishes. Solve $3p^2 - 1 = 0$, choose $p = -1/\sqrt{3}$, and insert it into $e = p^3 - p$. The positive limiting field is

$$e = 2/(3\sqrt{3}) \simeq 0.3849.$$

At the opposite spinodal the signs reverse.

This is a limit of the particular energy model. Real switching can occur earlier through domain walls, nucleation, defects or thermally activated barrier crossing. Observing a smaller coercive field in a specimen is not evidence that the cubic equation was solved inaccurately. To compare an actual hysteresis loop with a calculation, state the sweep rate, boundary circuit, sample preparation and a dynamical or nucleation model in addition to equilibrium free energy.

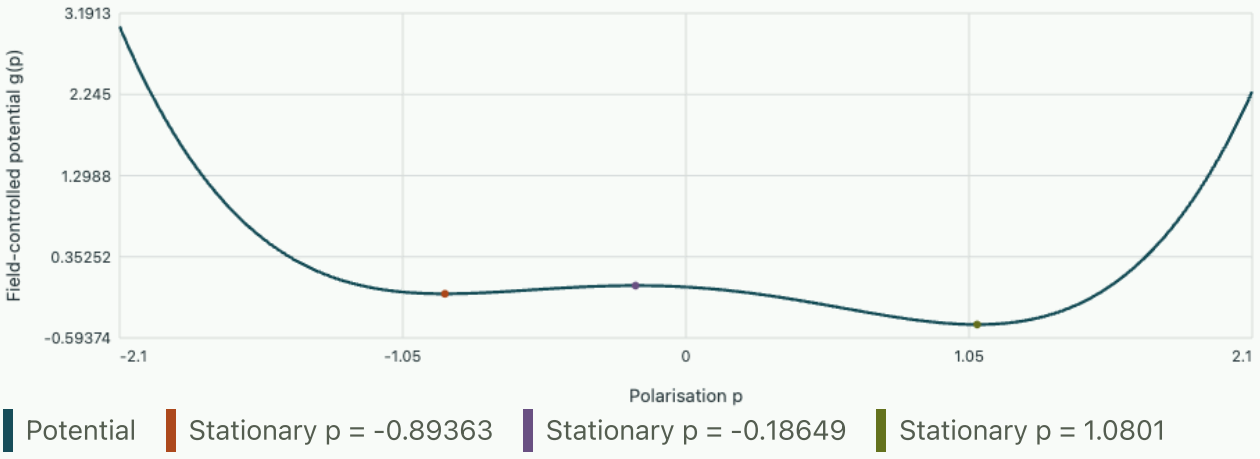
Read a landscape rather than tracing an unlabelled loop

Open the live experiment

Try the calculation

Temperature coefficient s **-1**

Applied field e **0.18**



Stationary polarisations: -0.89363 (minimum), -0.18649 (maximum), 1.0801 (minimum).

Set $s = -1$. At zero applied field, locate the two equal wells. Increase positive e gradually and watch the negative well become shallower. The browser locates real stationary points of the cubic, classifies them by curvature and reports their positions. Its plotted potential shows the energy barrier separating basins. Choose $s > 0$ and repeat to see why fitting the same S-shaped hysteresis loop above a continuous transition would be inappropriate.

A below-transition sample can also form spatially varying domains. Adding a positive gradient term $K(\nabla P)^2/2$ penalises interfaces, while depolarising fields can favour a mixture with reduced net surface charge. This competition sets new length scales absent from the homogeneous model. An electrode measuring total charge can return near-zero mean polarisation from oppositely oriented domains even though each domain is strongly polarised. A local probe and a bulk capacitance measurement therefore constrain different parts of the model.

Turn a capacitance measurement into a consistency test

For a small harmonic voltage imposed about a prepared local minimum, the differential response is $\delta P/\delta E = 1/[a(T - T_c) + 3bP^2]$, provided the driving frequency is slow enough for this local variable to respond and the amplitude is small enough to stay in the same well. Add the vacuum and fast polarisation terms when converting to total displacement D. In a parallel-plate geometry capacitance per area then follows from $\delta D/\delta E$ divided by thickness.

A large-signal charge–voltage measurement crosses different branches and may dissipate work around a loop. The energy per volume over a closed polarisation cycle is $\oint E\,dP$ with the sign convention for electrical work. A single equilibrium graph of $f(P)$ returns to its initial state with no prescribed loop area. Measured hysteresis thus carries dynamical information that a polynomial fit to one reversible branch cannot uniquely reconstruct.

Tutorial

For $s = -4$ in the scaled model, find zero-field minima and the positive homogeneous spinodal field.

▼ **Work through the solution**

Zero-field minima lie at $p = \pm 2$. The negative spinodal is

$$p = -\sqrt{4/3} = -2/\sqrt{3}.$$

Substitution gives positive field $16/(3\sqrt{3}) \simeq 3.08$. The $|s|^{3/2}$ dependence follows from the cubic.

A Newton solver initialised at $p = 0$ reports exact convergence for $s = -1, e = 0$. What additional observation rejects that result as the prepared stable state?

▼ **Work through the solution**

Its residual is exactly zero but the curvature $3p^2 + s = -1$ is negative. A small perturbation lowers the potential. Check a measured weak-field slope or the relaxation after a perturbation, and compare the energy of the two minima.

Why can the homogeneous coercive field be much larger than the measured coercive field?

▼ **Work through the solution**

The calculation removes a local barrier only when a spatially uniform state ceases to exist. A real material can switch by creating a small favourable domain and moving a wall; electrodes and defects can lower the needed activation energy. The relevant material and time scales differ.

Compare material descriptions

Use MIT's ferroelectrics and polarisation material to distinguish a microscopic source of displacement from the macroscopic free-energy model used here.

Fields and the response of matter · 5. Distinguish magnetic screening from electrical diffusion

Compare the frequency dependence, phase and thickness dependence of field penetration in a normal metal and a superconducting slab.

Perfect conductivity and magnetic expulsion are different statements

A normal conductor has current density $\mathbf{J} = \sigma \mathbf{E}$. Combining Faraday's and low-frequency Ampère laws in a homogeneous medium gives

$$\partial_t \mathbf{B} = (\mu\sigma)^{-1} \nabla^2 \mathbf{B}.$$

A step in field diffuses inwards over a length of order $\sqrt{t/(\mu\sigma)}$. Under sustained sinusoidal drive the skin depth is $\delta = \sqrt{2/(\mu\sigma\omega)}$. As ω tends to zero, this skin depth tends to infinity. An ordinary nonmagnetic conductor in its final static equilibrium generally admits a uniform applied magnetic field.

The London description of an isotropic superconductor in a Meissner state makes a distinct, equilibrium constitutive assumption:

$$\nabla \times \mathbf{J}_s = -\mathbf{B}/(\mu_0 \lambda_L^2),$$

with penetration length λ_L. In a homogeneous specimen carrying supercurrent and at frequencies where displacement current and normal losses may be neglected,

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}_s.$$

Curl once more and use $\nabla \cdot \mathbf{B} = 0$. The resulting field satisfies

$$\nabla^2 \mathbf{B} = \mathbf{B}/\lambda_L^2.$$

Unlike the diffusive skin depth, the equilibrium length remains finite for a static applied field.

Taking the formal limit $\sigma \to \infty$ in a normal conductor suppresses the electric field for a finite current and can freeze a pre-existing flux during changes. That argument does not determine which static field a specimen adopts when cooled through its transition. Magnetic expulsion is extra thermodynamic physics. The comparison below assumes a weak enough field and a suitable state, omitting flux vortices, trapped flux and nonlinear or nonlocal effects.

Solve a slab with equal applied fields on both faces

Let a broad slab fill $-d/2 < x < d/2$, and impose identical tangential magnetic fields $B_a \hat{\mathbf{z}}$ on its two faces. Symmetry requires $B'(0) = 0$. Applying both boundaries to the London equation gives

$$B(x) = B_a \frac{\cosh(x/\lambda_L)}{\cosh[d/(2\lambda_L)]},$$

\frac{\overline B}{B_a}=\frac{2\lambda_L}{d}\tanh\frac{d}{2\lambda_L}.

Differentiating gives the screening current $J_y = -B_z'/\mu_0$. Currents flow in opposite directions at opposite faces; their net transport current across the thickness is zero. For a thick slab each face approaches a half-space with exponential decay. For $d \ll \lambda_L$, the two surface layers overlap. Expanding the average gives $1 - d^2/(12\lambda_L^2) + \cdots$, so even a superconductor can have a substantial mean internal field in a sufficiently thin specimen under this parallel-field geometry.

Take $d = 100$ nm and $\lambda_L = 50$ nm. The centre field is $B_a/\cosh 1 \simeq 0.6481 B_a$; the mean is $\tanh 1 \simeq 0.7616 B_a$. The distinction between centre value and volume average is already larger than many instrumental error bars. A measurement of magnetic moment in a finite non-slab geometry additionally requires its demagnetising field. Do not directly substitute the infinite-slab mean into an arbitrary sample-shape calibration.

Obtain the equilibrium length as an energy minimum

The same static profile follows from minimising $\int_{-d/2}^{d/2} [B^2 + \lambda_L^2 (B')^2]\, dx/(2\mu_0)$ with boundary values imposed. The first term is magnetic field energy; the second is the energy associated with the screening current in this effective description. Vary B by a function η vanishing at the faces. Integrating $B'\eta'$ by parts leaves an interior factor $B - \lambda_L^2 B''$. Its vanishing returns the London equation. This route makes the physical tradeoff clear: complete expulsion would require extremely sharp gradients and costly currents, while uniform penetration pays magnetic energy throughout the bulk.

The reduction assumes a fixed superconducting condensate amplitude. Near a critical temperature or critical field, allowing that amplitude to change introduces an additional coherence length and may alter both the field state and the validity of London theory. A numerical fit giving a stable value of λ_L at one temperature does not prove this approximation over an entire phase diagram.

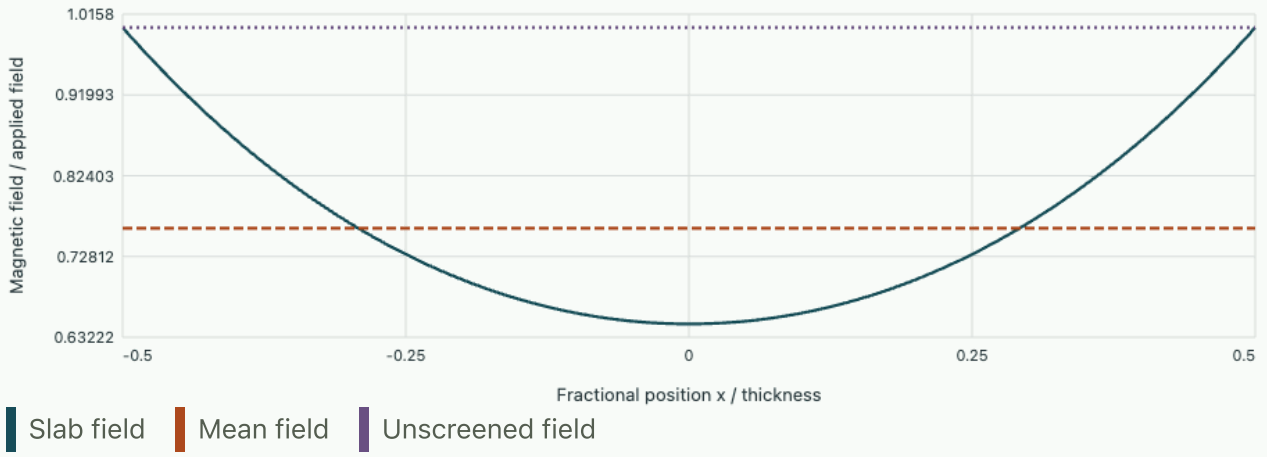
Compare two experimental controls

Open the live experiment

Try the calculation

Slab thickness d (nm) **100**

London length λ (nm) **50**



At d/λ = 2 the centre retains 64.805% of the applied field and the thickness average 76.159%.

Change the thickness and the penetration length independently. Find two different specimens with the same d/λ_L; their normalised profiles against fractional thickness match, even though their local physical gradients and current densities can differ. If all you can measure is a single normalised profile, geometry calibration is part of the inference of λ_L.

A frequency sweep supplies another distinction. In a normal sheet the complex field has both attenuation and phase lag with depth, and δ changes as $f^{-1/2}$. The static London profile above is real and determined by a material length. At higher frequencies superconductors can also exhibit complex response and dissipation, so measuring a single attenuation length in one frequency band cannot by itself establish static flux expulsion. Use the normal-film impedance experiment for a controlled comparison with its different rear boundary condition.

Tutorial

In a thick specimen, how far from one isolated face must a probe be to see 1% of the applied static field? How does the answer depend on field amplitude in this model?

▼ **Work through the solution**

The half-space profile is e^{-x/λ_L}. Set it to 0.01 to obtain

$$x = \lambda_L \ln 100 \simeq 4.605\lambda_L.$$

Linearity makes this scaled distance independent of B_a, until the assumed Meissner and weak-field conditions fail.

For a symmetric slab with $d = 4\lambda_L$, compute the centre and mean fractions.

▼ **Work through the solution**

The centre gives $1/\cosh 2 \simeq 0.266$. The mean is $\tanh 2/2 \simeq 0.482$. Both surface layers contribute substantially to the average even when the centre field is much smaller.

A phase-sensitive measurement detects a substantial depth-dependent phase lag in a slowly driven slab. Does the static London solution predict this?

▼ **Work through the solution**

The static real solution has no phase lag. Assess normal-fluid losses, vortex motion, measurement phase reference and the finite drive frequency. Comparing only magnitudes can conceal these distinct mechanisms.

Read the equilibrium model

The MIT applied-superconductivity course places London's laws alongside microscopic and order-parameter treatments. Use that ordering to identify the assumptions needed when taking a static limit.

Fields and the response of matter · 6. When a periodic dielectric has no propagating Bloch wave

Derive a band gap using two lossless transparent materials and distinguish evanescent Bloch propagation from absorption of optical power.

Construct a unit cell whose waves interfere

Stack two transparent, lossless, nonmagnetic dielectrics periodically in one dimension. Their refractive indices are n_1, n_2, their physical thicknesses d_1, d_2, and the period $a = d_1 + d_2$. Work at normal incidence, so polarisation directions are equivalent for isotropic layers. Each separate material admits a propagating wave at the chosen optical frequency, yet the periodic stack may not. The effect comes from coherent interface reflections, not from a negative refractive index or an absorptive constituent.

Describe the local field by its tangential electric and magnetic components. Propagation through one homogeneous layer connects them by a 2×2 transfer matrix with diagonal $\cos\phi_j$, off-diagonal terms proportional to $iZ_j \sin\phi_j$ and $i\sin\phi_j / Z_j$, where $\phi_j = \omega n_j d_j / c$ and $Z_j = Z_0 / n_j$. Changing the chosen propagation direction or time convention may reverse the two imaginary signs together. The real trace and unit determinant of the cell matrix below are unaffected.

Read the dispersion relation from the trace

At an interface both tangential components are continuous in the absence of imposed sheet currents. Multiply the two propagation matrices to form one cell. Its determinant is one because each layer's determinant is

$$\cos^2\phi + \sin^2\phi = 1.$$

Thus the two eigenvalues are reciprocal. For a Bloch mode they are $e^{\pm i k_B a}$ on a pass band. Taking half the trace gives

$$\cos(k_B a) = \cos\phi_1 \cos\phi_2 - \frac{1}{2}\left(\frac{n_1}{n_2} + \frac{n_2}{n_1}\right)\sin\phi_1 \sin\phi_2.$$

When the right side lies between –1 and 1 a real Bloch phase exists. If its magnitude exceeds one, a decaying and a growing Bloch factor replace a pair of unit-modulus factors. For a bounded excitation of a semi-infinite stack choose the decaying direction. The attenuation per cell is $\operatorname{arcosh}\left|\operatorname{trace}(M)/2\right|$. A lossless semi-infinite crystal in a stop band reflects incident power. Spatial attenuation of a Bloch amplitude in this region is compatible with zero volumetric heat production.

Obtain a quantitative stop-band width

Choose $n_1 d_1 = n_2 d_2 = \lambda_0/4$. Writing $s = \omega/\omega_0$, both layer phases are $\phi = \pi s/2$. At $s = 1$ the half trace is $-\frac{1}{2}(n_2/n_1 + n_1/n_2)$. Recognising $\cosh(\log r) = (r + r^{-1})/2$ gives centre attenuation $|\log(n_2/n_1)|$ per cell. For $n_1 = 1.5, n_2 = 2.0$ this is 0.2877, so a decaying eigencomponent has

amplitude factor 0.75 per cell. This statement about an eigencomponent should not be confused with the exact transmittance of a specific finite stack, which also depends on entrance and exit impedances.

The first gap edges obey half trace –1. With $\cos^2\phi = 1 - \sin^2\phi$, solve for the edge sine or cosine. The two edge phases are symmetric about $\pi/2$ and satisfy

$$|\cos\phi| = |n_2 - n_1|/(n_1 + n_2).$$

Hence the fractional angular-frequency width is $(4/\pi)\arcsin(|n_2 - n_1|/(n_1 + n_2))$, about 0.183 for the example. The lowest gap width tends continuously to zero as index contrast vanishes. A gap in a plotted numerical spectrum at $n_1 = n_2$ would indicate an implementation or plotting problem.

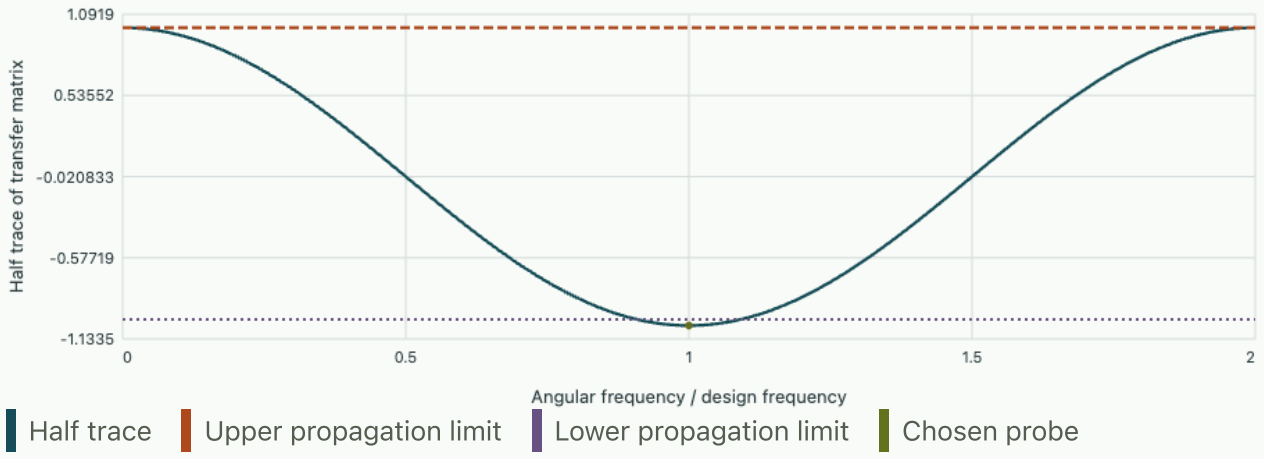
Compare a frequency scan with a finite mirror

Open the live experiment

Try the calculation

High index n_2 (n_1 = 1.5) 2

Probe frequency / design frequency 1



First gap width / design frequency 0.18252. Probe half trace -1.0417. Attenuation per cell 0.28768.

The graph plots half the transfer-matrix trace against normalised frequency and marks the admissible interval for a real Bloch phase. Change the index contrast first, keeping the layers quarter wave at the same reference frequency. The physical layer thickness adjusts in this family; it is a different experiment from changing a material while manufacturing all thicknesses fixed. Use the frequency control to read an attenuation exponent at the centre and just outside the gap.

At low frequency a wavelength spans many cells and a homogenised description can predict the long-wave speed. At the first gap the field changes phase by order one across a cell, exactly where the homogenisation scale separation fails. Conversely, an electromagnetic calculation showing multiple bands

at very low frequency should be checked against a uniform-medium or long-wavelength effective-index limit. Compare a finite stack's reflection with the Bragg-mirror experiment. The bulk trace determines whether a propagating state can exist. A finite sample adds interference between boundaries and possible surface or defect effects.

Tutorial

If the two refractive indices become equal, simplify the trace formula for arbitrary thicknesses.

▼ **Work through the solution**

The impedance factor becomes one, so the half trace is $\cos\phi_1\cos\phi_2 - \sin\phi_1\sin\phi_2 = \cos(\phi_1+\phi_2)$. It never exceeds unit magnitude. Labelling a uniform medium by an artificial period folds the dispersion into zones but does not create a physical gap.

At the central quarter-wave frequency for index ratio 2, how many periods reduce a decaying Bloch component's amplitude by a factor of 100?

▼ **Work through the solution**

The amplitude factor per cell is

$$e^{-\log 2} = 1/2.$$

Require $2^{-N} \le 0.01$, giving $N \ge 7$. A prediction for transmitted optical power also needs termination and impedance matching.

A measured stack exhibits both a transmission minimum and measurable heating. Can the minimum be attributed entirely to a band gap?

▼ **Work through the solution**

A lossless band gap can reflect power without heating. Measure reflected and transmitted flux under consistent normalisation, then infer absorption from $1 - R - T$ after accounting for diffuse scattering and other exit channels. The bulk stop-band model and material absorption can coexist.

Compare with band calculations

MIT's mathematical nanophotonics course computes full periodic Maxwell modes. The one-dimensional trace above is a small benchmark with known pass and stop boundaries that a numerical eigenproblem should reproduce.

Fields and the response of matter · 7. Why rotating a layered dielectric changes a capacitance measurement

Derive the two principal permittivities of a fine laminate, rotate the constitutive tensor and distinguish the applied electrode condition from an assumed field direction.

Compare the same material in two orientations

Build a specimen by repeating layers with scalar permittivities $\epsilon_1, \epsilon_2 > 0$, occupying fractions f and $1 - f$. The period is much smaller than both the specimen scale and the spatial variation of the applied field. Suppose both materials are linear, free of mobile charge and effectively lossless in the measurement band. An observer using a large capacitor wants one law relating averaged displacement and electric field. A scalar average cannot describe both arrangements of the same layers relative to the plates.

Place the field tangential to all internal interfaces first. A narrow rectangular electrostatic loop straddling an interface gives continuity of tangential E. Both layers see the same tangential field, but their displacements differ in proportion to permittivity. Averaging gives

$$\overline{D}_t = [f\epsilon_1 + (1 - f)\epsilon_2]\overline{E}_t.$$

The tangent-direction effective permittivity is an arithmetic mean. This is a parallel connection of capacitances at shared voltage in the special capacitor geometry.

Apply a purely normal displacement instead. No free charge lies at the internal interfaces, so their normal displacement is continuous. Normal fields differ as $E_i = D_n/\epsilon_i$. Averaging the voltage drops gives

$$\overline{E}_n = [f/\epsilon_1 + (1 - f)/\epsilon_2]\overline{D}_n,$$

with a harmonic-mean effective permittivity. The local continuity condition changes which field is shared between constituents and thereby changes the relevant averaging operation.

Check the amount of anisotropy from field energy

For $\epsilon_1 = 2\epsilon_0, \epsilon_2 = 10\epsilon_0$ at equal fractions,

$$\epsilon_t = 6\epsilon_0,$$
$$\epsilon_n = (10/3)\epsilon_0.$$

The arithmetic-to-harmonic ratio is 1.8. A constituent does not need to be crystalline or anisotropic for the composite to show distinct principal responses. The geometry of the internal boundaries is enough.

Average the energy density for the tangential experiment,

$$u = f\epsilon_1 E_t^2/2 + (1 - f)\epsilon_2 E_t^2/2 = \epsilon_t E_t^2/2.$$

For the normal experiment,

$$u = D_n^2[f/\epsilon_1 + (1 - f)/\epsilon_2]/2 = D_n^2/(2\epsilon_n).$$

These expressions supply a check independent of the capacitor-circuit analogy and identify which energy should be used when field or free charge is controlled. Energy positivity follows directly from the positive constituent permittivities. It does not require guessing an effective law from the largest measured component.

Rotate the law, then specify what the electrodes control

Use principal axes $\hat{n}, \hat{t}$ and suppose a macro-field makes angle θ to the layer normal. If $\overline{\mathbf{E}} = E(\cos\theta\,\hat{n} + \sin\theta\,\hat{t})$, then $\overline{\mathbf{D}} = E(\epsilon_n \cos\theta\,\hat{n} + \epsilon_t \sin\theta\,\hat{t})$. Project along the imposed field and divide by E:

$$\epsilon_E(\theta) = \epsilon_n \cos^2\theta + \epsilon_t \sin^2\theta.$$

The perpendicular displacement is $E(\epsilon_t - \epsilon_n)\sin\theta\cos\theta$. It is generally nonzero. A device that instead fixes the direction of averaged displacement while allowing a transverse field measures $\epsilon_D(\theta) = [\cos^2\theta/\epsilon_n + \sin^2\theta/\epsilon_t]^{-1}$. These two directional responses correspond to different constraints. They agree at principal axes and for isotropic material. A pair of unspecified “angled capacitor” measurements cannot be interpreted uniquely without describing the contacts and edge conditions.

At 45° in the example, the field-controlled projection gives about $4.667\epsilon_0$, while the displacement-controlled projection gives $4.286\epsilon_0$. These are both legitimate scalar ratios from one symmetric positive-definite tensor. Neither supplies a universal scalar permittivity for the rotated material. The complete vector response also predicts displacement tilt and field energy for any orientation.

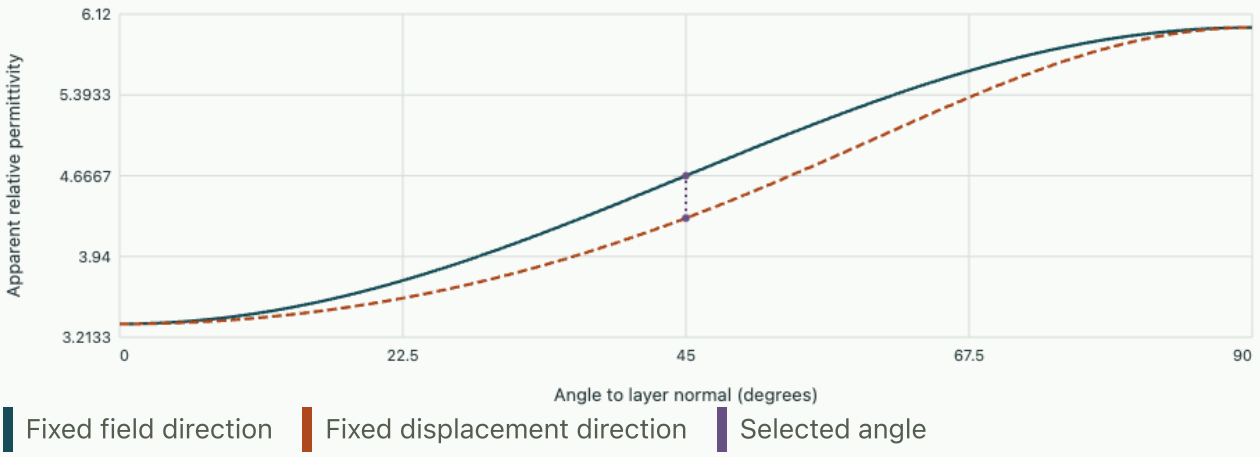
Test the measurement against the unit-cell model

Open the live experiment

Try the calculation

Fraction of low-permittivity phase **0.5**

Measurement angle (degrees) **45**



Fixed field direction | Fixed displacement direction | Selected angle

Principal relative permittivities: normal 3.3333, tangential 6. At 45° the two responses are 4.6667 and 4.2857.

Change the volume fraction while keeping the constituent values fixed. At a pure phase all orientations and both control conditions must collapse to a single value. At intermediate fractions, find the angle where the transverse component would be largest. It occurs where $|\sin\theta\cos\theta|$ is largest, but its magnitude also reflects the material contrast. Now reduce the contrast mentally to one and predict all curves without recomputing a single cell.

A numerical field model offers an independent homogenisation experiment. Use periodic tiling in the material directions and impose a macroscopic voltage difference plus periodic fluctuations on opposite faces. For a unit imposed field along each coordinate direction, average the resulting displacement over the cell. The resulting vectors supply the tensor columns. A grounded box on every side is a different boundary problem and can suppress the tangential response. Increase the number of material periods at fixed specimen scale and refine the individual period separately. These two limits test homogenisation and numerical representation, respectively.

How far can the effective law travel?

Electrostatic homogenisation assumes the cell fields respond quasistatically. In a wave experiment near the Bragg gap, phase changes substantially across one period and a local constant tensor is inadequate. A finely layered capacitor and a dielectric mirror can be made from the same constituents while requiring different models at different frequencies. Test the Bloch spectrum to identify the regime where layering begins selecting propagation bands. A low-frequency effective index alone cannot predict that gap.

Large spatial contrast also places a practical demand on mesh and interface handling. A normal gradient can jump strongly even when potential and displacement are continuous. A smooth-looking interpolated

voltage surface is weak evidence for an accurate energy density. Compare constituent fields, total flux and energy consistency before accepting the extracted effective tensor.

Tutorial

Show that $\epsilon_t - \epsilon_n$ is nonnegative for every positive pair of constituents without appealing to a named mean inequality.

▼ **Work through the solution**

Put both terms over the positive denominator $f\epsilon_2 + (1-f)\epsilon_1$. The numerator after cancellation is $f(1-f)(\epsilon_1 - \epsilon_2)^2$, so the difference is nonnegative and vanishes for identical materials or a pure phase.

Equal fractions give measured relative values 6 tangentially and $10/3$ normally. Can you infer both constituents if the laminate assumptions hold?

▼ **Work through the solution**

Their sum is 12. The harmonic mean is $2ab/(a+b) = 10/3$, giving product 20. The positive roots of $x^2 - 12x + 20 = 0$ are 2 and 10. The data do not identify which label belongs to which phase without another observation.

Why can a constituent occupy very little volume yet strongly affect the normal response?

▼ **Work through the solution**

Every normal field line in this lamination crosses each layer. The inverse response weights f/ϵ; even a thin, low-permittivity layer can contribute a large voltage drop. A field along the layers distributes displacement differently and may be much less sensitive to that same phase.

Fields and the response of matter · 8. How a magnetic field rotates the polarisation of transmitted light

Resolve a linearly polarised wave into circular eigenmodes, infer a refractive-index difference from an analyser and distinguish nonreciprocal rotation from a simple change of reference axes.

Derive rotation from propagation rather than drawing an arrow

Let monochromatic light propagate along $+z$ through a transparent magneto-optic sample, with an externally prepared axial magnetic field. Define fixed laboratory vectors $\mathbf{e}_\pm = (\hat{x} \pm i\hat{y})/\sqrt{2}$ and use phasors $e^{i(kz-\omega t)}$. These definitions specify the handedness algebraically. Calling a vector simply “right circular” without stating the viewing and time conventions can reverse a reported rotation sign.

An input linearly polarised in x is $(\mathbf{e}_+ + \mathbf{e}_-)/\sqrt{2}$. Suppose the medium supports wave numbers $k_\pm = 2\pi n_\pm/\lambda_0$ for the two bases, with negligible absorption. After a distance L, factor out the common phase $e^{i(k_+ + k_-)L/2}$. The remaining Cartesian phasor is $\hat{x}\cos\Delta - \hat{y}\sin\Delta$, where $\Delta = (k_+ - k_-)L/2$. Thus the polarisation angle in the fixed x, y plane is

$$\theta = \frac{\pi L}{\lambda_0}(n_- - n_+).$$

Equal transmissions of the two components preserve a linearly polarised output. Unequal absorption makes the circular amplitudes unequal and generally produces an ellipse. Rotating a linear analyser then finds a finite minimum, even for a perfectly coherent field. Residual intensity at the nominal extinction angle can also arise from stray light or imperfect polariser extinction, which must be measured independently.

Derive a magneto-optic eigenbasis

For axial magnetisation, a simple lossless transverse relative-permittivity block is $\begin{pmatrix} \epsilon & ig \\ -ig & \epsilon \end{pmatrix}$ with real g at the frequency of observation. Acting on $(1, i)$ gives $(\epsilon - g)(1, i)$; acting on $(1, -i)$ gives $(\epsilon + g)(1, -i)$. For relative permeability near one, $n_+^2 = \epsilon - g$, $n_-^2 = \epsilon + g$. At small g their difference $n_- - n_+ \simeq g/\sqrt{\epsilon}$. Reversing magnetisation reverses the sign of the off-diagonal coefficient and interchanges the two eigenvalues. Both the existence of circular modes and the sign-reversal test now follow from a constitutive model.

If $n_- - n_+$ is proportional to magnetic field over a specified range, write $\theta = VBL$, with Verdet constant V in rad/(T m). This is a measurable material and wavelength dependent coefficient. It is not determined by the mean refractive index alone. A strong nearby absorption line can give substantial frequency dependence and dichroism, invalidating a lossless narrow-band extrapolation.

Work through an analyser calibration

For $\lambda_0 = 600$ nm, $L = 0.20$ mm and measured index difference 5.0×10^{-4}, the rotation is $\pi(2.0 \times 10^{-4})(5.0 \times 10^{-4})/(6.0 \times 10^{-7}) = \pi/6$, or 30°. An ideal analyser at angle α transmits $I/I_{\text{in}} = \cos^2(\alpha - \theta)$. Its maxima occur at 30° and 210°; an intensity measurement determines a linear

polarisation angle modulo 180°. For a sample thick enough to rotate several half turns, an intensity maximum alone cannot give an unwrapped θ. Vary field or path length continuously from a known reference.

If the applied field is independently 0.40 T in that example, $V = \theta/(BL) \simeq 6.54 \times 10^3$ rad/(T m). The arithmetic is dimensionally correct, but the large value warrants a physical check of thickness, field calibration, units and material spectrum before reporting it as an ordinary glass coefficient. Degrees substituted for radians would make the inferred material coefficient too large by $180/\pi$.

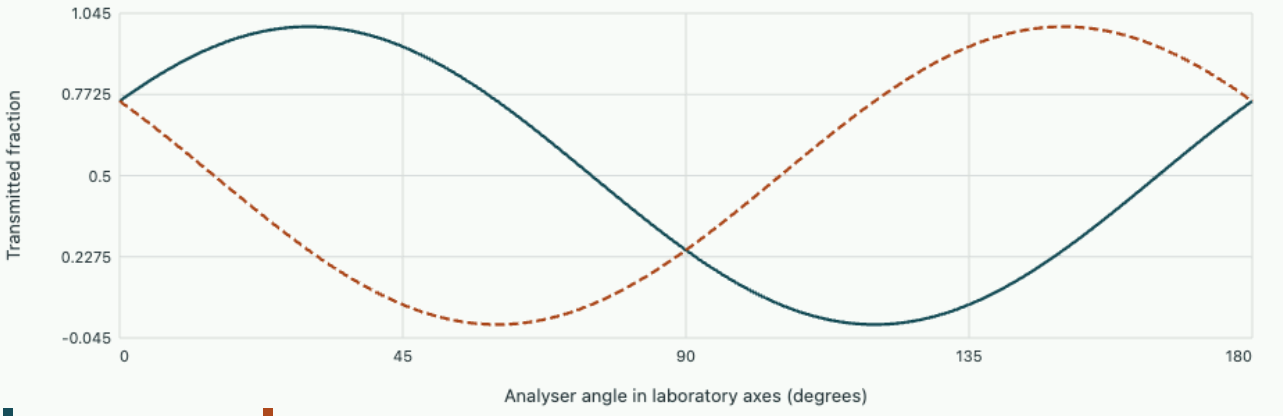
Use field reversal to reject alignment errors

Open the live experiment

Try the calculation

Index difference (10^{-4}) **5**

Sample length (mm) **0.2**



Unwrapped single-pass rotation 30°. At a 45° analyser the two transmitted fractions are 0.93301 and 0.066987.

The experiment shows the predicted analyser traces for opposite index splitting, with the same input polariser. Increase sample thickness to see why tracking the maximum by one measurement can select the wrong branch. Reversing the splitting reverses the tilt in fixed laboratory axes. A fixed analyser offset shifts both inferred angles equally and cancels in half their difference, provided the branch has been unwrapped consistently.

For small rotation, an analyser exactly at a crossed position transmits intensity proportional to θ^2. Its response has no sign sensitivity close to zero. Set the analyser at 45° instead:

$$\cos^2(\pi/4 - \theta) = [1 + \sin(2\theta)]/2 \simeq 1/2 + \theta.$$

The difference between fields $+B$ and $-B$ then measures a signed signal linear in small rotation. This trades extinction of the background for sensitivity to sign. Balance the two orthogonal analyser outputs if source intensity fluctuates appreciably during a sweep.

Ask what happens on a return pass

For the ideal Faraday medium at fixed magnetisation, the circular eigenmodes relative to a fixed laboratory basis are determined by its transverse material tensor. On propagation back through the sample, the accumulated positive propagation path again distinguishes the same laboratory circular basis after accounting for the reflection boundary. With a suitable simple mirror at normal incidence, the resulting material rotation doubles on a round trip in fixed transverse coordinates. In a reciprocal optically active material, the eigenvalue assignment with respect to the laboratory basis reverses on reversing propagation, and the single-pass rotation cancels on the ideal return. Write both transformations in one fixed basis before comparing signs.

A real double-pass setup includes mirror phase shifts, stress birefringence of windows and possible ellipticity. Reversing magnetic field, removing the sample and recording both analyser quadratures provide separate checks. A nonzero rotation in one transmitted beam proves only that its polarisation changed; establishing nonreciprocity requires comparing a reversed propagation experiment with a stated polarisation convention.

Tutorial

Two analyser maxima occur at 47° and 39° for opposite weak fields. If the response is linear and the same branch is tracked, what field-induced rotation and static offset do they imply?

▼ **Work through the solution**

Half the difference gives 4 degrees of signed rotation for the first field relative to zero. The mean is 43 degrees, the combined input polarisation and analyser zero. Conversion to a Verdet constant additionally needs the applied field and path length.

In the 600 nm example, what thickness uncertainty corresponds to 1 degree of rotation uncertainty at fixed index difference?

▼ **Work through the solution**

Rotation is proportional to thickness. Relative uncertainty $1/30$ therefore gives about $0.20/30 = 0.0067$ mm, or 6.7 µm. Uncertainty in wavelength or field would add additional contributions.

An ideal analyser gives a minimum 4% of its maximum. Can this be explained solely by a rotation of an ideal linearly polarised monochromatic input in a lossless circularly birefringent sample?

▼ **Work through the solution**

Pure rotation preserves a linear state, so an ideal analyser should extinguish it completely. Differential loss, ellipticity, depolarisation, finite bandwidth or a detector background are possible alternatives. Use circular-analysis or full Stokes measurements and an empty-path baseline to distinguish them.

Compare with a full polarisation measurement

Use the Stokes reconstruction experiment to determine which intensities reveal a circular component. A fitted linear rotation alone cannot measure that missing component.

Fields and the response of matter · 9. Can a flat interface guide light without mirrors?

Send light along a plane boundary between a dielectric and a metal. Seek a wave that propagates parallel to the interface while decaying away on both sides. The signs in the magnetic and electric boundary conditions determine whether any such wave can exist.

Decide what confinement requires

Put a homogeneous dielectric with positive relative permittivity ϵ_d above $z = 0$, a lossless dispersive metal of relative permittivity $\epsilon_m(\omega)$ below it and take both permeabilities to be μ_0. Use phasors with $e^{i\beta x - i\omega t}$. For the transverse magnetic polarisation write $H_y = H_0 e^{-\kappa_d z}$ in the dielectric and $H_y = H_0 e^{+\kappa_m z}$ in the metal. Both κ must be positive real numbers for this bound, lossless mode. Substitution into the homogeneous wave equation gives

$$\kappa_j^2 = \beta^2 - \epsilon_j k_0^2,$$

$$k_0 = \omega / c.$$

Continuity of tangential magnetic field without an externally imposed surface current identifies the two amplitudes H_0. The metal's mobile charges are already included in its bulk permittivity. Introducing the same current a second time as an arbitrary sheet would describe another boundary. Continuity of normal electric displacement follows here from the same magnetic amplitude and common tangential wavenumber.

The sign that selects a material pair

Use $\nabla \times \mathbf{H} = -i\omega\epsilon_0\epsilon_j \mathbf{E}$. Its x component is $-\partial_z H_y$. Hence $E_x = -i\partial_z H_y / (\omega\epsilon_0\epsilon_j)$. Continuity of this tangential electric field gives

$$\kappa_d / \epsilon_d = -\kappa_m / \epsilon_m.$$

No two positive real permittivities can satisfy that equation with two positive decay constants. Squaring and eliminating κ_j gives a candidate propagation constant.

$$\beta^2 = k_0^2 \frac{\epsilon_m \epsilon_d}{\epsilon_m + \epsilon_d}.$$

Check it in the unsquared equation. With positive dielectric permittivity the physical branch requires $\epsilon_m < -\epsilon_d$. On this branch $\beta > \sqrt{\epsilon_d}\, k_0$, so the mode lies outside the propagation cone of a plane wave in the dielectric. For transverse electric polarisation the corresponding condition replaces permittivities by permeabilities. With two ordinary positive permeabilities the sign requirement fails. The wave exists because of a coupled electric field and charge response at a surface.

Work an example before using the dispersion diagram

Let $\epsilon_d = 1$, $\epsilon_m = -4$ at the chosen frequency. Then $\beta / k_0 = \sqrt{4/3} = 1.155$,

$$\kappa_d/k_0 = 1/\sqrt{3},$$

$$\kappa_m/k_0 = 4/\sqrt{3}.$$

At vacuum wavelength 600 nm the electric and magnetic field amplitudes in the dielectric fall by a factor e after 165 nm. The corresponding distance in the metal is 41 nm. An intensity measurement falls by a factor e at half these distances. A wave sent directly through air has a parallel wavevector no larger than k_0. An index-1.50 prism could match the parallel component at $\theta = \arcsin(1.155/1.50)$, about 50.4 degrees. A practical evanescent coupler also needs finite overlap with the surface field.

Identify the surface resonance without assuming an infinite measured signal

Take the lossless Drude law $\epsilon_m = 1 - \omega_p^2/\omega^2$. The guided branch ends at $\omega_{\mathrm{sp}} = \omega_p/\sqrt{1+\epsilon_d}$, where $\epsilon_m + \epsilon_d$ tends to zero. Approaching this frequency from below increases β at finite ω. In the electrostatic limit the restoring electric field and inertia of mobile charge select the limiting frequency. At small ω, by contrast, β approaches the dielectric light line and the wave spreads far into the dielectric. A measured resonance has a finite width. Collisions, surface roughness, microscopic nonlocal response and the optical coupling geometry matter before any divergence could be observed.

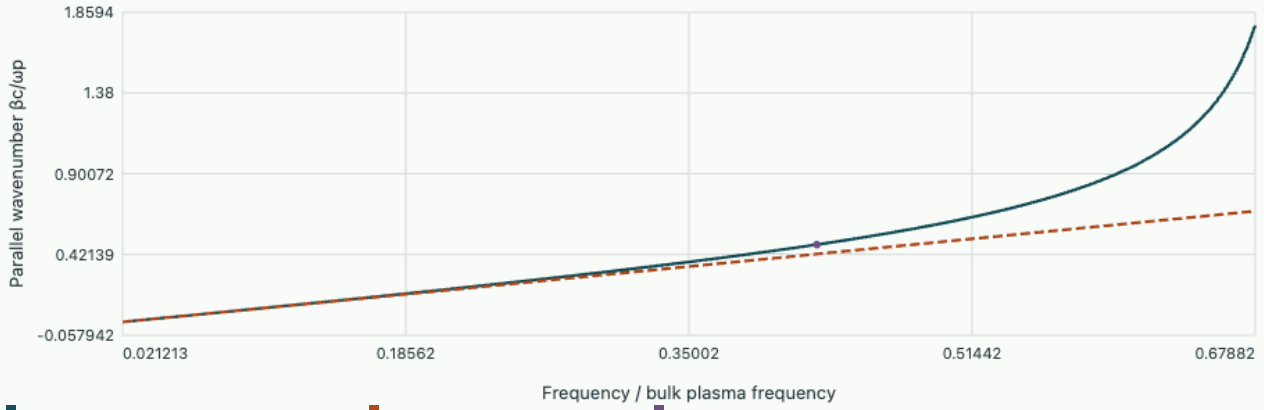
Change the dielectric permittivity. Predict separately the light-line slope and the frequency approached by the surface mode. The plot shows the propagation and dielectric light lines against frequency in units of the bulk plasma frequency, with a marker at the chosen fraction of the surface resonance.

Open the live experiment

Try the calculation

Dielectric relative permittivity **1**

Selected frequency / surface resonance **0.6**



Bound TM surface mode Dielectric light line Selected frequency

Surface resonance ωsp/ωp = 0.70711. At the selected frequency β/k₀ = 1.1319.

Tutorial

If the dielectric permittivity becomes 2.25 with the same ideal metal, what bulk-normalised frequency does the surface mode approach? Does an air beam now have enough tangential momentum near the resonance?

▼ **Work through the solution**

The asymptote moves from $1/\sqrt{2} = 0.7071$ to $1/\sqrt{3.25} = 0.5547$. The parallel wavenumber exceeds the dielectric light line $1.5k_0$, which already exceeds the air limit. Direct incidence from air cannot provide the necessary parallel phase gradient.

Two near-field probes lie 120 nm apart normal to an interface at the 600 nm example. Their measured intensities have ratio 0.48. Is that consistent with sampling only the bound mode in the dielectric?

▼ **Work through the solution**

Its predicted intensity ratio is

$$\exp[-2(120/165)] = 0.234.$$

A ratio near 0.48 is close to the amplitude ratio. Check whether the detector is actually a linear heterodyne field measurement. Background light would also flatten an intensity decay.

Suppose a calculation squares the boundary equation and reports a propagating surface wave for two positive dielectrics. How can it be rejected without repeating the whole derivation?

▼ **Work through the solution**

Compute both positive normal decay constants and substitute them into $\kappa_d/\epsilon_d + \kappa_m/\epsilon_m = 0$. The sum cannot vanish. Even a real positive β^2 alone does not establish decay plus the correct boundary sign.

Compare a single-interface solution with the two-interface standing-wave condition in the dielectric slab chapter.

Fields and the response of matter · 10. Infer charge discreteness from fluctuations in a steady photocurrent

A detector can deliver a perfectly steady mean current over a long interval while its charge per short exposure fluctuates. Derive both a counting prediction and its spectral version, then use a bandwidth-limited measurement to separate partition fluctuations from thermal and readout noise.

Start with what the counter actually reports

Let statistically independent detection events bring charge q at mean rate ν. Count events during a rectangular exposure of duration T. Subdivide it into small bins of length δt in which a single event has probability $\nu\delta t$ and two events have negligible probability compared with δt. Independent Bernoulli bins have additive means and variances. Taking δt towards zero while T is fixed gives equal count mean and variance, both νT. The integrated charge is $Q = qN$, and the indicated average current is Q/T.

$$I = q\nu,$$

$$\operatorname{var}(Q) = qIT,$$

$$\operatorname{var}(\bar{I}_T) = qI/T.$$

The relative counting uncertainty is $1/\sqrt{\nu T}$, which falls with exposure. The variance of the integrated charge itself grows with exposure. Confusing the sum and its average reverses the predicted trend with integration time.

Derive the spectral density from a known measurement window

Define the one-sided current spectrum $S_I(f)$ for $f > 0$, so that current variance at an instrument output equals its integral weighted by the squared gain. Averaging over the rectangular exposure uses impulse response $h(t) = 1/T$ between 0 and T, zero elsewhere. Its Fourier amplitude is a sinc times a phase factor. Parseval's identity, applied to the real impulse response, gives the one-sided equivalent noise bandwidth $\int_0^\infty |H(f)|^2 df = 1/(2T)$. Equate the counting result to $S_I/(2T)$.

$$S_I = 2qI,$$

$$[S_I] = \mathrm{A}^2/\mathrm{Hz}.$$

The factor of two belongs to the one-sided convention. The square root is an amplitude spectral density, measured in $\mathrm{A}/\sqrt{\mathrm{Hz}}$. A one-pole filter with −3 dB frequency f_c instead has gain squared $1/[1 + (f/f_c)^2]$. Its equivalent noise bandwidth is $\pi f_c/2$. Replacing that integral by f_c silently changes a noise prediction by a measurable factor.

Separate uncorrelated sources by changing the mean

A photodiode feeds an ideal current-measuring amplifier with a $1.0\ \mathrm{M\Omega}$ feedback resistor at 300 K. Assume photon arrivals have Poisson statistics, unity internal charge gain, negligible dark-current fluctuations and an otherwise ideal amplifier. Refer the resistor noise to the input as a parallel current source, with density $4k_B T_\mathrm{R}/R$. Independent variances add. At $I = 0.80$ microampere and equivalent

bandwidth $B = 10.0$ kHz the photocurrent gives 50.63 pA rms and the resistor contributes 12.87 pA. The total current noise is their quadrature sum, about 52.24 pA. Adding their rms amplitudes would predict 63.50 pA despite the assumed independence.

Keep the resistor temperature and filter fixed while sweeping the photocurrent. Plot measured variance divided by equivalent bandwidth against the actual mean current. The slope is $2q$ and the intercept measures the current-independent floor. In the more general model $\mathrm{var}(N) = F\langle N\rangle$ for long exposures, the low-frequency slope is $2qF$. A measured slope alone cannot separate charge from a Fano factor. Establishing independent electrons and unity gain is part of inferring the elementary charge.

Run the measurement with controlled non-Poisson counting

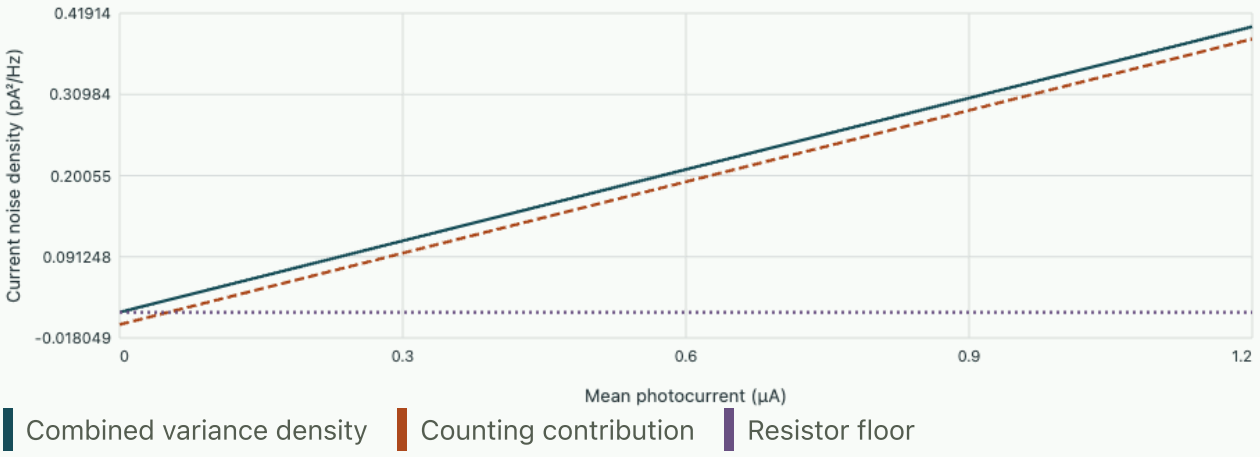
The graph plots variance across a current sweep, quoted as spectral density in $\mathrm{pA}^2/\mathrm{Hz}$. Vary the Fano factor and the readout resistance independently. Predict which operation affects the intercept. Read the current at which the signal fluctuations equal the readout fluctuations and check whether increasing the detection bandwidth changes this crossover for two white sources.

Open the live experiment

Try the calculation

Fano factor 1

Feedback resistance (MΩ) 1



The two noise contributions are equal at 51.704 nA. The zero-current density is 0.016568 pA²/Hz.

A detector impulse $i_p(t)$ with finite transit time has integrated area q. Poisson arrivals give a low-frequency plateau because the current pulses look like point charges when averaged slowly. At frequencies comparable with the inverse transit time the spectrum gains a factor $|\tilde{i}_p(f)/q|^2$. A curved spectrum can arise from that single-event shape even with independent arrival times. Compare the curve with a calibrated impulse response before attributing it to correlations of separate events.

Tutorial

An exposure of 20 microseconds collects a mean of 200 electrons. Predict the standard deviation in electrons and the indicated mean current. What happens after averaging four independent exposures?

▼ **Work through the solution**

The standard deviation is $\sqrt{200} = 14.14$ electrons and the mean current is

$$200e/(20\,\mu\mathrm{s}) = 1.60\,\mathrm{pA}.$$

Averaging four independent readings halves the rms fluctuation in indicated current. Adding their raw counts gives a different observable with mean 800 and standard deviation 28.28.

For $R = 1\ \mathrm{M\Omega}$, $T = 300$ K and $F = 1$, at what current do resistor and shot noise have equal variance?

▼ **Work through the solution**

Equate $2eI$ and $4k_BT/R$ to obtain

$$I = 2k_BT/(eR) = 51.7\,\mathrm{nA}.$$

Both variances scale with a common equivalent noise bandwidth, so it cancels in this white-noise model.

A colleague multiplies a measured noise density by the optical carrier frequency to obtain the receiver noise. Which frequency interval belongs in this calculation?

▼ **Work through the solution**

Integrate the electrical noise spectrum over the measurement filter's squared gain. The optical field creates charge carriers, but the transimpedance circuit and estimator determine which fluctuations in their arrival rate reach the output. A carrier frequency and an electrical equivalent noise bandwidth have the same dimension and different operational meanings.

Use the RC noise experiment to derive the equilibrium resistor contribution independently.

Fields and the response of matter · 11. Why a nearby gate changes the speed law of a two-dimensional plasma

A current pulse excites a travelling density modulation in a thin conducting sheet. Move a metal gate towards the sheet and the shape of the dispersion curve changes. Rebuild the restoring electric field from the surface charge before using any oscillator analogy.

Solve the electrostatic problem for one Fourier mode

Place a two-dimensional electron sheet at $z = 0$ in a homogeneous dielectric of absolute permittivity ϵ. The equilibrium positive background is fixed. Let the electron number perturbation be $\delta n\, e^{iqx - i\omega t}$ with $q > 0$, so the perturbation of charge per area is $\sigma = -e\,\delta n$. Away from the sheet the potential solves $(\partial_z^2 - q^2)\phi = 0$. Decay on both sides selects $A e^{-q|z|}$. Across zero, the normal field discontinuity gives

$$2\epsilon q A = \sigma.$$

Consequently the bare sheet has $\phi(0) = \sigma/(2\epsilon q)$. The inverse power of q reflects the field extending a distance of order $1/q$ away from the sheet.

Ground a parallel, perfectly conducting plane at $z = d$. Place an image of opposite surface charge at $2d$ to enforce zero potential at the conductor. The transverse exponential suppresses the contribution that travels from this image to the physical sheet. Linear superposition therefore gives

$$\phi(0) = \frac{-e\,\delta n}{2\epsilon q}\left(1 - e^{-2qd}\right).$$

This result assumes that charge is free to flow into the gate from a ground reservoir and that the dielectric is homogeneous throughout the gap. A floating finite plate would require an additional global charge condition. The zero mode also needs its own electrostatic reference and is excluded from this derivation.

Combine continuity with the acceleration equation

Writing u for the sheet electron velocity along x, number conservation gives

$$-i\omega\delta n + iqn_0 u = 0.$$

The electric force on a single electron is $-eE_x = +ieq\phi(0)$, and the acceleration equation is

$$-i\omega m^* u = ieq\phi(0).$$

Solving for u, substituting the potential and cancelling the common density perturbation gives

$$\omega^2(q) = \frac{n_0 e^2 q}{2\epsilon m^*}\left(1 - e^{-2qd}\right).$$

Check dimensions using n_0 as a number per square metre. A three-dimensional density in this formula would leave a spurious inverse length. In an ungated plane the frequency scales as $\sqrt{q}$. Its group velocity

$d\omega/dq$ is half the phase velocity. When qd is small, the exponential has first-order difference $2qd$ and the field is largely confined in the gate gap. Then $\omega \simeq sq$ with $s = \sqrt{n_0 e^2 d/(\epsilon m^*)}$. The field and inertia now behave like a distributed capacitance and kinetic inductance. Both waves are collective charge oscillations even though one has the same long-wavelength dispersion exponent as sound.

Put numbers on the length-scale assumption

Use sheet density $2.0 \times 10^{15}\,\mathrm{m}^{-2}$, effective mass $0.067 m_e$ and relative permittivity 12.9. For a 5.0 micrometre density wavelength,

$$q = 2\pi/\lambda = 1.257 \times 10^{6}\,\mathrm{m}^{-1}.$$

A 50 nm gate separation gives $qd = 0.0628$. Taking a square root after evaluating the electrostatics gives 342.4 GHz without the gate and 117.6 GHz with it. The small-qd linear speed $6.07 \times 10^{5}\,\mathrm{m\,s}^{-1}$ would predict roughly 121.4 GHz. The few-percent difference can be seen well before the field has ceased to be substantially gated.

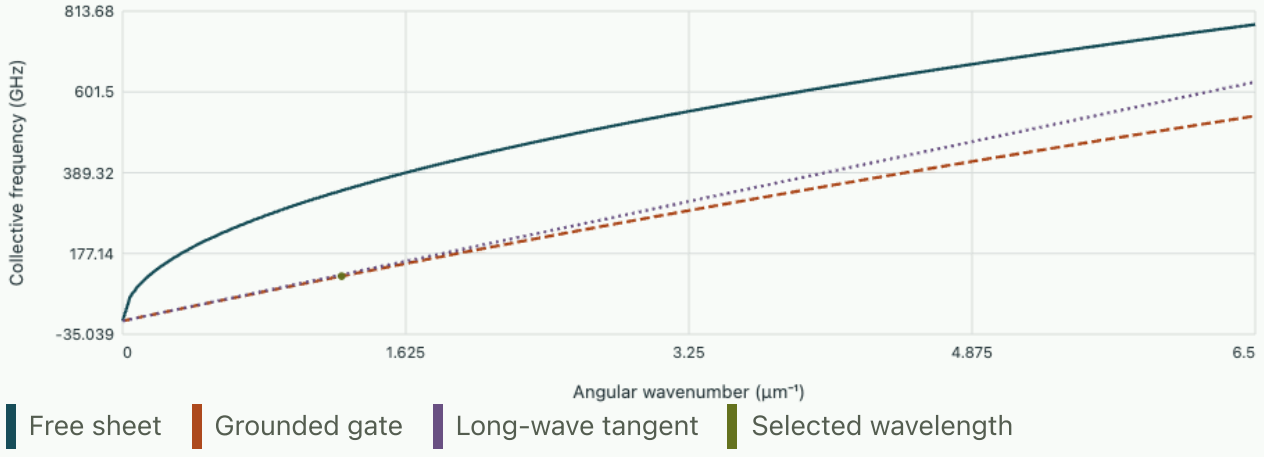
The experiment lets you vary the distance and the selected wavelength. Compare the full dispersion with its gate-screened tangent. The horizontal coordinate is inverse micrometres for the angular wavenumber. Predict where $qd = 1$ lies before interpreting a bend. Check the reported selected frequency by converting q to an ordinary wavelength with the 2π included.

Open the live experiment

Try the calculation

Gate separation (nm) **50**

Selected wavelength (µm) **5**



At the chosen wavelength qd = 0.062832. Gated mode 117.65 GHz. Free sheet 342.36 GHz.

Infer a parameter and test whether it is identifiable

A low-wavenumber velocity measurement determines $n_0 d/(\epsilon m^*)$. Doubling n_0 and halving d gives the same leading speed, so one such measurement cannot separate density and spacing. At sufficiently large qd, ω^2/q becomes independent of d and determines $n_0/(\epsilon m^*)$. Combining low and high wavenumber patterns can therefore identify d separately. In a real sheet, collisions, scattering, electron pressure and a nonlocal conductivity can invalidate the simple acceleration law. An inferred gate dimension is credible only if a model fitted in one regime predicts data in the other.

Tutorial

Derive the ratio of exact group to phase velocity from the dispersion and state both limiting values.

▼ **Work through the solution**

Differentiate $\ln \omega$ with respect to $\ln q$. The answer is

$$v_g/v_p = \tfrac{1}{2}[1 + 2qd/(\exp(2qd) - 1)].$$

It tends to one for a closely screened long wave and to one half when the gate is far from the evanescent field.

A measured long-wave speed increases by ten percent with unchanged density and permittivity. Under the ideal gate model, how much must the spacing have changed?

▼ **Work through the solution**

Since $s^2 \propto d$, the inferred spacing increases by $1.10^2 - 1 = 0.21$, or 21 percent. A simultaneous change of effective mass would invalidate this inference. Check the large-q coefficient independently.

Explain why an electromagnetic calculation of waves propagating only at $c/\sqrt{\epsilon_r}$ misses this mode, and why the present electrostatic derivation cannot be extrapolated to arbitrary phase speeds.

▼ **Work through the solution**

The moving surface charge supplies a dynamic constitutive response with its own inertia. A constant bulk permittivity alone omits that degree of freedom. Conversely our use of Laplace's equation neglects propagation delay. If the mode approaches the dielectric light line one must keep the retarded Maxwell fields and solve for both charge motion and electromagnetic propagation.

Compare the bulk charge-density experiment and the electrostatic image construction.

Fields and the response of matter · 12. Can a capacitance measurement reveal the fraction of embedded particles?

A sample contains well separated dielectric spheres in an insulating matrix. Its capacitance changes when the particle fraction changes. Derive a dilute estimate from one sphere and a self-consistent approximation from a local-field model. Use the inverse calculation to discover when the fraction becomes poorly determined.

What does a single particle actually contribute?

Let the host and sphere have positive, frequency-independent relative permittivities h and s and let each sphere have radius a. The imposed field near an isolated sphere is approximately uniform. The exterior potential is the sum of a uniform-field term and a dipole term proportional to $\cos\theta/r^2$; the regular interior solution is linear in $r\cos\theta$. Impose continuity of potential and of normal electric displacement at $r = a$. The internal field then equals $3hE_{\mathrm{loc}}/(s+2h)$. The induced dipole is its integrated excess polarisation relative to the host:

$$p = \frac{4\pi a^3}{3}\epsilon_0(s-h)E_{\mathrm{in}} = 4\pi\epsilon_0 h a^3 \frac{s-h}{s+2h}E_{\mathrm{loc}}.$$

A low-permittivity inclusion can have a dipole opposite to its local field. It has removed more polarisable host material. This sign need not imply any negative permittivity. Define the contrast $b = (s-h)/(s+2h)$. The denominator includes two components of the host contribution from the exterior field. Simply averaging permittivities as if every particle experienced the applied field would ignore this boundary calculation.

Leading density and a local-field closure

For number density n, the inclusion fraction is $f = 4\pi a^3 n/3$. To first order in f, replace E_{loc} by the macroscopic field and add the dipoles. The predicted effective relative permittivity is $h(1+3fb)$, up to terms of second and higher order. To obtain the Maxwell–Garnett approximation, imagine a spherical exclusion region containing a representative particle. Approximate the far distribution of particles by its smooth excess polarisation $P = np$. Removing the polarised sphere produces a local field increment $P/(3\epsilon_0 h)$ along P. Substituting $E_{\mathrm{loc}} = E + P/(3\epsilon_0 h)$ into $P = n\alpha E_{\mathrm{loc}}$ gives

$$\frac{\epsilon_{\mathrm{eff}} - h}{\epsilon_{\mathrm{eff}} + 2h} = fb,$$

$$\epsilon_{\mathrm{eff}} = h\,\frac{1+2fb}{1-fb}.$$

The closure resums some higher-order terms. It does not solve the exact many-particle boundary-value problem at arbitrary density. Short-range correlations, connected clusters, nonspherical particles and anisotropy can matter. Identifying a distinguished host and separated inclusions is an experimental assumption. Interchanging host and particle labels while preserving their volume fractions generally gives another microstructure and another prediction.

Infer a fraction from a measured capacitor

Take $h = 2$, $s = 8$ so $b = 0.5$. At a fraction of 0.15, Maxwell–Garnett gives an effective relative permittivity $2(1.15)/(0.925) = 2.4865$. The first-order estimate gives 2.45. For electrodes of area $100\,\mathrm{mm}^2$ separated by $0.5\,\mathrm{mm}$, and negligible edge fields, these numbers give 4.40 pF and 4.34 pF. A stray capacitance of 0.1 pF is larger than their difference. A regression reporting many fractional digits cannot decide between these approximations until the empty fixture and lead capacitances are known.

Invert the ratio before inserting a measured value. If $\epsilon_{\mathrm{eff}} = 2.50$ the inferred fraction is

$$(2.50 - 2)/[(2.50 + 4)0.5] = 0.154.$$

The derivative of inferred fraction with respect to the measured effective value is $3h/[b(\epsilon_{\mathrm{eff}} + 2h)^2]$. An absolute uncertainty of 0.03 on the effective permittivity corresponds to approximately 0.0085 on the fraction for this example. If s approaches h, b approaches zero and the same measurement loses sensitivity. Increasing the number of data points only helps random scatter. It cannot create constituent contrast.

Compare geometry-dependent predictions at the same volume fraction

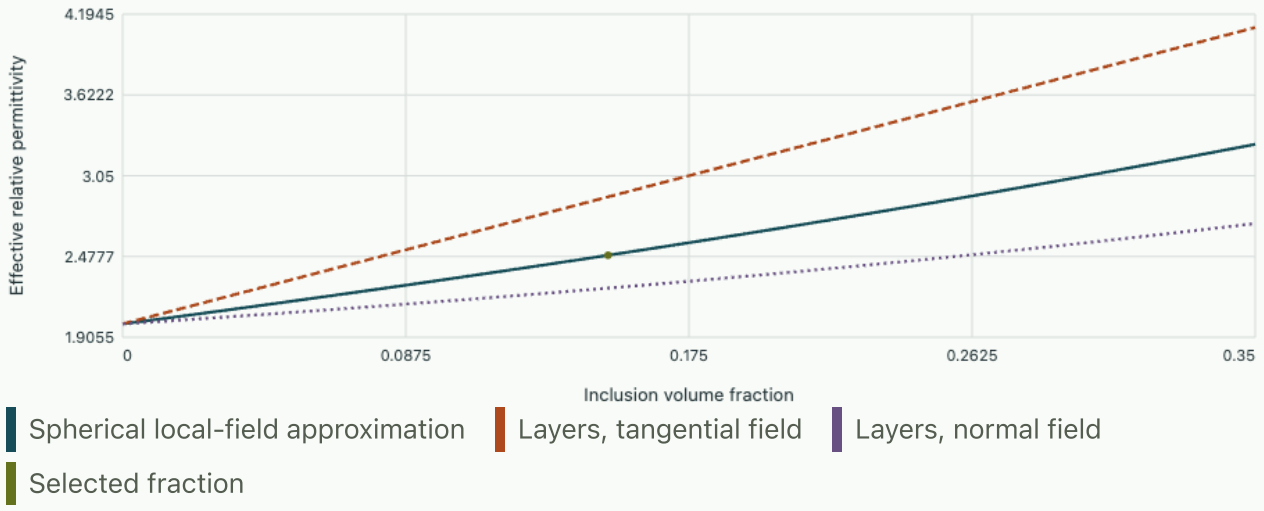
The graph uses a host permittivity of two and lets the particle permittivity and selected volume fraction vary independently. Compare the sphere estimate to the exact arithmetic and harmonic values for a laminate with fields parallel and perpendicular to its layers. All curves coincide for equal constituents and pure host. Separated spheres have no preferred direction under isotropic sampling, whereas a laminate does. Agreement of one measured scalar with a mixing formula cannot establish which geometry was present.

Open the live experiment

Try the calculation

Particle relative permittivity **8**

Selected volume fraction **0.15**



At the selected fraction εeff = 2.4865. The first-order dilute estimate gives 2.45.

Tutorial

If the particle radius is halved at fixed dilute volume fraction, what happens to the predicted effective permittivity? What unmodelled effect could reveal the smaller radius?

▼ **Work through the solution**

There are eight times more particles per volume and each ideal dipole has one eighth of its original polarizability. Their product is unchanged in quasistatic volume mixing. A finite-thickness interfacial layer occupies an increasingly large fraction as radius shrinks. Dynamic scattering at non-negligible size can also distinguish the samples.

What does the local-field formula predict for cavities of relative permittivity one embedded in a host of four at $f = 0.10$?

▼ **Work through the solution**

Here $b = (1 - 4)/(1 + 8) = -1/3$. The effective result is

$$4(1 - 2/30)/(1 + 1/30) = 3.613.$$

The sign and magnitude lie between the pure constituent values. A value larger than four from this lossless model would require rechecking the sign of the excess dipole.

A series of samples has the same measured capacitance in one direction. Propose a measurement that can separate isotropically dispersed spheres from a stack of layers.

▼ Work through the solution

Reorient the electrodes or the sample and measure normal and tangential components at fixed geometry and frequency. A laminate generally exhibits different arithmetic and harmonic responses. An isotropic distribution of dilute spheres has a scalar effective response. Repeat with a second fixture orientation to exclude direction-dependent stray capacitance.

Work through the single-particle calculation in the dielectric sphere chapter and compare the planar geometry in the layered material experiment. The University of Utah electromagnetic materials course gives the local-field construction.

Fields and the response of matter · 13. Can two ohmic layers imitate a relaxing dielectric?

Model a film and its contact layer in series. Derive the frequency scale of interfacial charge accumulation and decide which features of a measured spectrum can arise from geometry alone.

Match total current at a leaky interface

Take harmonic fields proportional to $e^{i\omega t}$. In layer j, Ohm's law and Maxwell's displacement current give $J = Y_j E_j$ with $Y_j = \sigma_j + i\omega\epsilon_j$. Use absolute permittivity in SI units. Total current is continuous across the interface. If the first layer occupies thickness fraction f, the imposed mean field is

$$\bar{E} = fE_1 + (1-f)E_2.$$

Eliminating the layer fields yields

$$Y_{\mathrm{eff}} = \frac{Y_1 Y_2}{fY_2 + (1-f)Y_1},$$

$$\tau = \frac{f\epsilon_2 + (1-f)\epsilon_1}{f\sigma_2 + (1-f)\sigma_1}.$$

The denominator is proportional to $1 + i\omega\tau$. A pole has appeared even though each material has frequency-independent conductivity and permittivity. The free interfacial charge per unit area is $\epsilon_2 E_2 - \epsilon_1 E_1$ with the normal from layer 1 to layer 2. At low frequency the less conducting layer takes a larger field. At high frequency the displacement currents determine the field sharing.

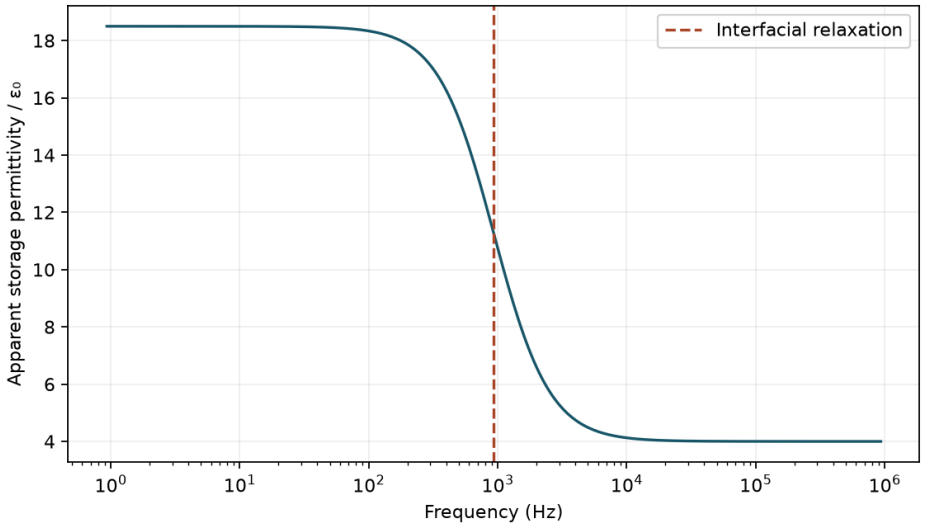
Separate dielectric storage from direct conduction

Infer the relative storage permittivity as $\mathrm{Im}(Y_{\mathrm{eff}})/(\omega\epsilon_0)$. A DC current is real and drops out of this quantity. For layers with the same relative permittivity, their high-frequency effective value must equal that common value. In the DC limit the effective conductivity is the thickness-weighted harmonic mean. These two checks are independent. The interface can nevertheless produce a large rise in apparent storage on the low-frequency side. Use a first layer that fills 20% of the gap and has conductivity 10^{-8} S/m. Give the other layer conductivity 10^{-6} S/m and both layers relative permittivity 4. The DC conductivity is 4.81×10^{-8} S/m. An applied mean field of 1000 V/m therefore drives 4.81×10^{-5} A/m². The weakly conducting layer needs about 4808 V/m to carry it and the second layer only 48.1 V/m. Their thickness-weighted mean recovers the imposed field.

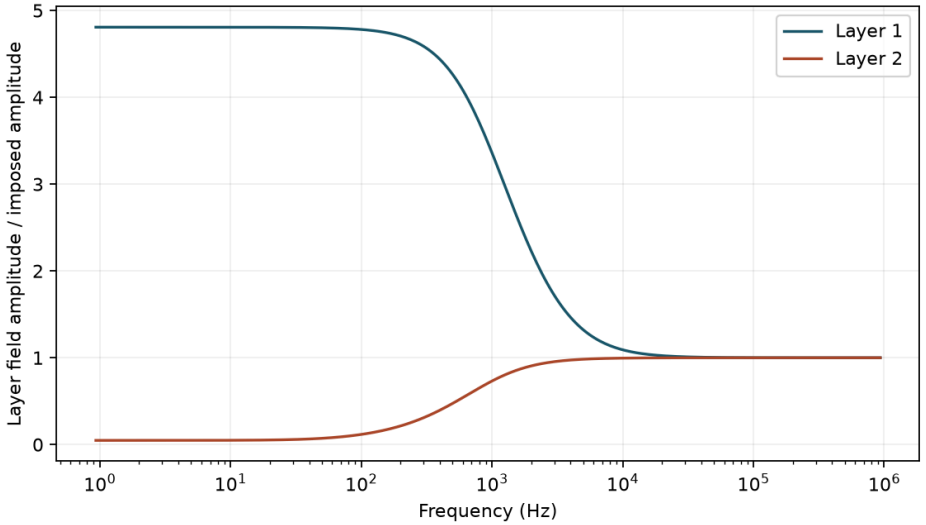
The time constant is 0.170 ms, corresponding to 935 Hz. Expanding the two complex conductivities to first order in frequency gives the low-frequency relative storage permittivity

$$4[f\sigma_2^2 + (1-f)\sigma_1^2]/[f\sigma_2 + (1-f)\sigma_1]^2 = 18.50.$$

At high frequency it falls to 4. Make the conductivities equal in the notebook and predict the interfacial charge before running it. The denominator still defines a time scale algebraically but its pole cancels out of the measured admittance.



Use the measured transition frequency and a thickness series to distinguish interfacial storage from an intrinsic molecular relaxation.



Each curve is an amplitude. Add the complex fields with their phases before checking that the thickness-weighted average equals the applied field.

Exercise · remove the relaxation

Which choice of layer properties makes the interfacial charge vanish for every frequency?

▼ Work through the solution

The charge vanishes if $\epsilon_2/Y_2 = \epsilon_1/Y_1$. The imaginary terms cancel identically after cross multiplication, leaving $\epsilon_2\sigma_1 = \epsilon_1\sigma_2$. The two dielectric relaxation times are then equal. Check the limit by making both layers identical.

Exercise · exchange the materials

Relabel the layers and their thickness fractions. Which measurements can distinguish the order of this stack?

▼ **Work through the solution**

The two-terminal admittance is unchanged under $(1, 2, f) \mapsto (2, 1, 1 - f)$. It cannot identify which material is nearer the upper electrode. A local field measurement at a known depth or a contact-specific measurement can distinguish the configurations.

Try it · Leaky layers

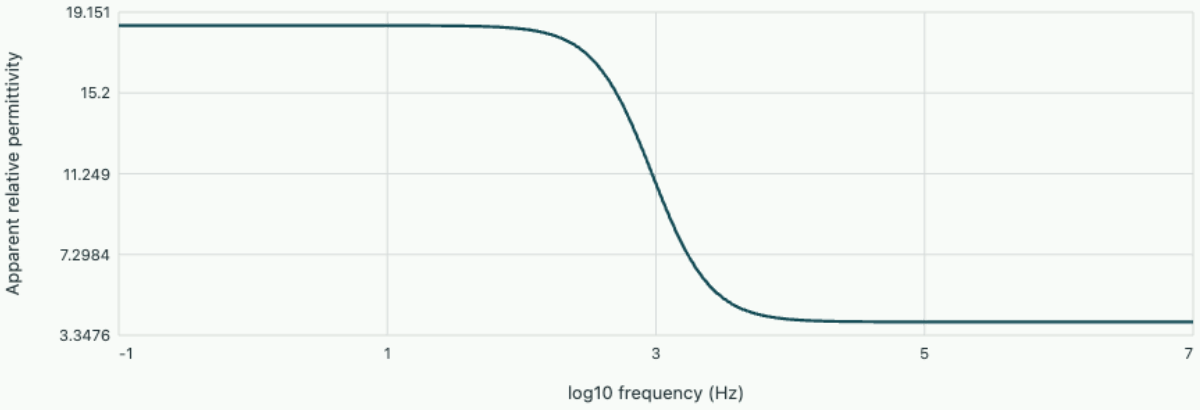
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Fraction occupied by layer 1 **0.2**

Conductivity ratio σ2/σ1 **100**



Effective storage permittivity

Both intrinsic relative permittivities equal 4. Interfacial relaxation frequency 934.71 Hz. A large low-frequency apparent permittivity can arise from conductivity contrast.

Download the corresponding editable notebook. The browser recalculates the reduced equations directly; the notebook exposes the full calculation and its checks.

Open editable notebook Download .ipynb

Fields and the response of matter · 14. From molecular dipoles to a dielectric susceptibility

Derive induced and orientational polarisation separately, then determine when temperature or local fields invalidate a constant dielectric coefficient.

Displace a bound charge before defining a material coefficient

Consider a neutral molecule whose negative charge cloud can move a small distance relative to its positive centre. Represent the restoring force along the applied field by a spring constant k and the displaced charge magnitude by q. Static balance gives kx=qE. The dipole moment is p=qx, so its proportionality coefficient is fixed by the restoring physics.

$$p = \alpha E,$$
$$\alpha = q^2/k.$$

The energy includes both deformation and the coupling to the external field. Minimising kx²/2−qxE gives the equilibrium displacement and the equilibrium energy −αE²/2. The factor one half follows from creating the moment while raising the field. A permanent, already established dipole instead has orientation energy −pE cosθ. The difference matters in a force gradient and in a thermodynamic partition sum.

> **Recall · Dipole moment, torque and field energy**
>
> Two charges ±q separated by vector d have p=qd directed toward positive charge. Expanding their potential energy in a slowly varying external potential gives U=p·∇V=−p·E. Vary the angle to obtain torque p×E. Vary position to obtain the gradient force. This expansion requires the charge separation to be small compared with the scale on which the external field changes.

For N independently responding molecules per unit volume, P=NαE if the local and macroscopic fields agree. Combining D=ε₀E+P with D=εE gives $\chi_e = N\alpha/\varepsilon_0$. Each equality has an assumption. A dense solid can change the local field, and a large displacement can change the spring constant.

Count orientations with the correct angular measure

Now take a freely rotating permanent dipole of fixed magnitude p in thermal equilibrium at temperature T. Put u=cosθ and x=pE/(kBT). A thin ring on the orientation sphere has area 2π sinθ dθ. Consequently equal intervals in u have equal orientation measure. Boltzmann weighting gives the normalised density below.

$$Z_u(x) = \int_{-1}^{1} e^{xu}\, du = \frac{2\sinh x}{x},$$
$$f(u) = \frac{e^{xu}}{Z_u(x)}.$$

Differentiate the integral with respect to x. The derivative inserts u in the numerator, so the alignment is the logarithmic derivative of Z. Differentiating a second time gives a variance and therefore a nonnegative

differential response.

$$\langle u\rangle = \frac{d\log Z_u}{dx} = \coth x - \frac{1}{x} = L(x),$$

$$\frac{d\langle u\rangle}{dx} = \langle u^2\rangle - \langle u\rangle^2.$$

Series expansion of sinh and cosh gives L(x)=x/3−x³/45+… . Thus the orientational susceptibility of a dilute gas is Np²/(3ε₀kBT). At large x the alignment approaches 1−1/x, saturating at P=Np. A linear law cannot remain valid in that limit.

$$P = N\left(\alpha + \frac{p^2}{3k_BT}\right)E + O(E^3).$$

For p=3.0×10⁻³⁰ C m, T=300 K and E=10⁶ V/m, x=7.24×10⁻⁴. The mean orientation is only 2.41×10⁻⁴. A visibly nonzero macroscopic polarisation can therefore coexist with almost random individual molecular directions. Doubling E doubles the weak-field response, while doubling T halves its orientational part. The induced contribution in this elementary model stays fixed.

OneSimulation · When does the linear orientation law fail?

Try alignment and the angular measure

Change pE/(kBT) to concentrate the probability near the field direction, then lower the integration resolution. Equal intervals in u = cosθ represent equal solid-angle bands. The zero-field density is therefore 1/2 on this axis. The reported mean compares a trapezoidal integral with the exact Langevin law and its weak-field expansion.

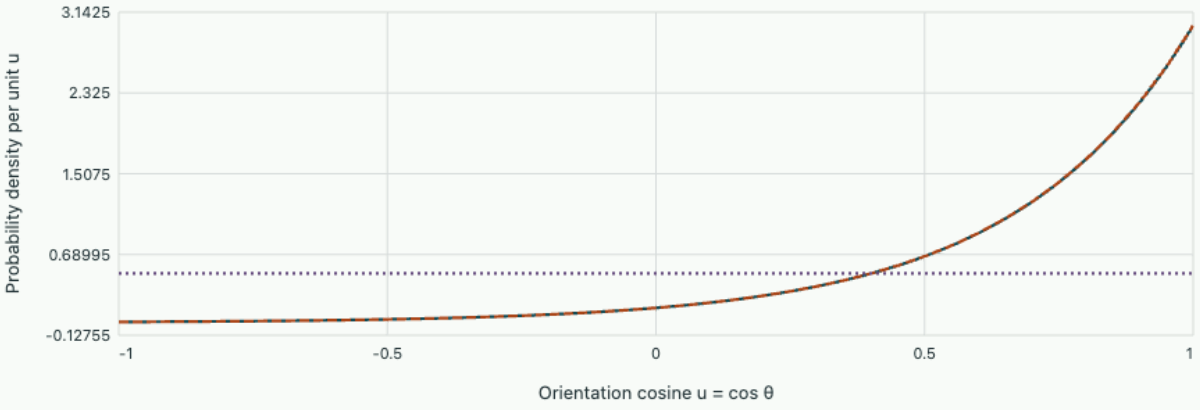
The browser uses a simple composite trapezoidal rule so that resolution error is visible. The OneSimulation notebook below obtains Gaussian quadrature nodes from an eigenproblem and provides an independent, higher-order calculation of the same integral. Increase `order` there and compare convergence at fixed field energy. The model assumes freely rotating, noninteracting permanent dipoles in thermal equilibrium.

Open this experiment online

Try the calculation

Dipole energy / thermal energy pE/(kBT) **3**

Trapezoidal integration intervals **40**



Boltzmann orientation density | Integration samples | Zero-field isotropic density

Exact mean cos θ = 0.671636; trapezoidal mean = 0.672883; weak-field estimate = 1. Absolute quadrature error = 1.25e-3. Density normalisation from this grid = 1.00187. The downloadable notebook independently uses Gaussian quadrature through OneSimulation.

Reproduce with OneSimulation

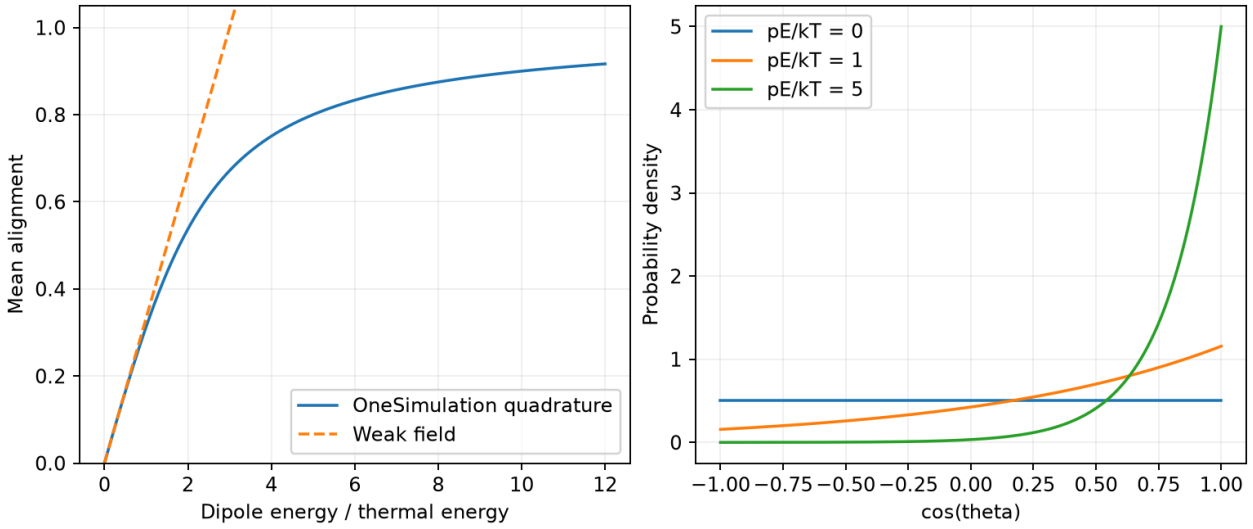
The code obtains Gaussian integration nodes and weights from a symmetric eigenproblem using OneSimulation. It then integrates the probability density derived above. The analytic Langevin law supplies an independent check. Run this locally with OneSimulation, NumPy and Matplotlib installed.

```
"""Orient freely rotating permanent dipoles and compare quadrature with the exact law."""
import numpy as np
import matplotlib.pyplot as plt
from onesim.numerics.linear.matrices.dense_linear_algebra import symmetric_eigensystem

# The eigenvectors give Gaussian weights for the integral over u = cos(theta).
order = 32
indices = np.arange(1, order, dtype=float)
off_diagonal = indices / np.sqrt(4 * indices**2 - 1)
jacobi = np.diag(off_diagonal, 1) + np.diag(off_diagonal, -1)
nodes, vectors = symmetric_eigensystem(jacobi)
weights = 2 * vectors[0, :]**2
field_energy = np.linspace(0, 12, 121)  # x = p E / (k_B T)
mean_alignment = []
for x in field_energy:
    probability = weights * np.exp(x * (nodes - 1))
    mean_alignment.append(np.dot(nodes, probability) / probability.sum())
mean_alignment = np.asarray(mean_alignment)
exact = np.zeros_like(field_energy)
exact[1:] = 1 / np.tanh(field_energy[1:]) - 1 / field_energy[1:]
print(f"Maximum quadrature error: {np.max(abs(mean_alignment - exact)):.2e}")
for x in (0.2, 1.0, 3.0, 10.0):
    index = np.argmin(abs(field_energy - x))
    print(f"pE/kT = {x:4.1f}, <cos(theta)> = {mean_alignment[index]:.6f}, linear = {x/3:.6f}")
assert np.max(abs(mean_alignment - exact)) < 1e-12
```

```
fig, axes = plt.subplots(1, 2, figsize=(9, 3.8), layout="constrained")
axes[0].plot(field_energy, mean_alignment, label="OneSimulation quadrature")
axes[0].plot(field_energy, field_energy / 3, "--", label="Weak field")
axes[0].set(xlabel="Dipole energy / thermal energy", ylabel="Mean alignment", ylim=(0, 1.05))
axes[0].legend()
for x in (0, 1, 5):
    u = np.linspace(-1, 1, 301)
    normalizer = 2 if x == 0 else 2*np.sinh(x)/x
    axes[1].plot(u, np.exp(x*u)/normalizer, label=f"pE/kT = {x}")
axes[1].set(xlabel="cos(theta)", ylabel="Probability density")
axes[1].legend()
for axis in axes:
    axis.grid(alpha=.2)
fig.savefig("electrodynamics-orientation.png", dpi=170)
```

Output from the supplied parameters

```
Maximum quadrature error: 8.88e-16
pE/kT =  0.2, <cos(theta)> = 0.066490, linear = 0.066667
pE/kT =  1.0, <cos(theta)> = 0.313035, linear = 0.333333
pE/kT =  3.0, <cos(theta)> = 0.671636, linear = 1.000000
pE/kT = 10.0, <cos(theta)> = 0.900000, linear = 3.333333
```



The probability concentrates near the field direction. At x=3 the weak-field law already predicts complete alignment, while the integral gives 0.672. Change the quadrature order and find when numerical error becomes visible separately from the linear-model error.

Download Python · Download Jupyter notebook

Close the local-field problem

In an isotropic, approximately cubic environment imagine a spherical cavity around the chosen molecule. Let n point from the centre into the material. The material normal at the cavity wall is −n, giving bound charge σb=−P·n. A surface patch contributes dE(0)=−σb n dΩ/(4πε₀). Integrating with $\int n_i n_j \, d\Omega = 4\pi\delta_{ij}/3$ gives the cavity contribution P/(3ε₀) at its centre. Under the Lorentz-cavity approximation, which neglects the remaining discrete near-neighbour correction, E_loc=E+P/(3ε₀). Write α_eff=α+p²/(3kBT) in the weak-field regime and solve P=Nα_eff E_loc for P.

$$\frac{\epsilon_r - 1}{\epsilon_r + 2} = \frac{N\alpha_{\text{eff}}}{3\epsilon_0}.$$

For Nα_eff/(3ε₀)=0.10 this gives εr=1.3333. The independent-molecule estimate gives 1.30. The discrepancy increases with density. The divergence predicted as that parameter approaches unity indicates the breakdown or instability of this simple constitutive closure. Molecular correlations, anisotropy, or nonlinear ordering must then be considered before assigning a physical permittivity.

Remember · A cavity field is a model of the surroundings

The factor one third belongs to a spherical cavity in the isotropic construction. The macroscopic depolarising factor of the entire specimen is a different geometric quantity. The local-field calculation addresses what one molecule feels. The boundary-value calculation addresses how the specimen and electrodes determine the macroscopic E.

Follow the ambiguity of the unit cell in a polar crystal

Take an infinite alternating chain of charges +q at na and −q at na+a/2, with one chain per cross-sectional area A. A neutral cell containing these two sites has moment −qa/2 and polarisation −q/(2A). Choose the equivalent negative ion at na−a/2 instead, keeping the same infinite crystal. That cell has moment +qa/2 and polarisation +q/(2A). The two assignments differ by q/A although every actual charge position is unchanged.

A finite sample supplies the missing information. Its end charges depend on how it is cut and on whether charges are supplied at the surfaces. Reassigning the cell changes the book-keeping of bound and free end charge together. The measurable total surface charge and external field remain determined by the physical termination. A bulk dipole-per-cell convention alone therefore does not determine the surface charge of an ionic crystal. In three dimensions the corresponding polarisation increment is qR/Ω for lattice vector R and cell volume Ω.

Exercises from response to inference

1. At the same number density, the measured dilute susceptibility is 0.010 at 300 K and 0.007 at 600 K. Separate a temperature-independent contribution from a contribution proportional to 1/T.

▼ **Work through the solution**

Write χ=A+B/T. Subtraction gives 0.003=B/600, hence B=1.8 K and A=0.004. At 300 K the orientational contribution is 0.006. To infer p from B also requires N. If density changes with temperature, correct that effect before fitting this law.

2. A linear extrapolation predicts P=1.5Np. Determine what assumption has failed and estimate the large-field response for x=10.

▼ **Work through the solution**

A mean cosine cannot exceed unity. The small-x expansion has been used outside its domain. L(10)=0.900000004…, giving P≈0.900Np for freely rotating permanent dipoles. The induced contribution is a separate term and can change the total moment magnitude.

---

3. Repeat the unit-cell construction after replacing the cross-sectional area by 2A. Explain which quantity changes and what measurement must specify the crystal termination.

▼ **Work through the solution**

The polarisation branch spacing becomes q/(2A). Surface charge per area and the external field are intensive quantities, so the chain density matters. An end-charge measurement with specified electrodes and termination determines the physical surface preparation.

Electromechanics, optics and measurement

Follow the field into an instrument. Predict which mechanical displacement, temperature change or optical signal can be observed and determine what can be learned from it.

1. ## When a suspended capacitor loses its stable equilibrium

   An electrode suspended on a spring can remain motionless at a bias that makes the calculated electrostatic force grow as it moves. Static force balance alone misses the voltage at which that equilibrium disappears. Deriving the small-signal frequency gives an independent measurement of the same instability.

2. ## Which part of a sample changes a four-electrode impedance measurement?

   Driving one electrode pair and measuring another yields a transfer impedance whose spatial sensitivity can be positive, negative or nearly zero. A second field solution supplies the sensitivity without perturbing every element of a numerical mesh.

3. ## Measuring the force on a sliding dielectric

   A suspended strip slides between two electrodes without touching them. Predict the pull, infer the air clearance from a capacitance trace and find where an electrical measurement can detect a change in mechanical stiffness.

4. ## Designing a displacement bridge at its noise limit

   A moving electrode changes two capacitances in opposite directions. Derive the displacement signal, reconstruct its noise in metres and select a carrier voltage consistent with measured mechanical stiffness and electrical losses.

5. ## A moving-coil galvanometer as an electrical and mechanical system

   A scale division records an angle, yet the quantity marked on the dial is electric current. A spring, a magnetic field and the resistance of the external measuring circuit connect the two. Derive the calibration, predict how a pointer approaches its reading, and find which experiment measures the charge of a short pulse.

6. ## Eddy-current braking, magnetic penetration and the heat left behind

   A conducting loop moves above a row of magnetic poles. The drag first increases with speed, but a rapid crossing need not generate proportionally greater braking. Trace the work done against the Lorentz force into circuit energy, heat and a measurable temperature change. Then decide which features of a solid plate the single-loop model can capture.

7. ## Measuring magnetic core loss by a hysteresis loop and by calorimetry

A closed loop on a graph of magnetic induction against field encloses an area with units of energy per unit volume. A thermometer measures kelvins. Bring these two observations onto the same scale, account for the winding heat, and determine when the two experiments actually test the same material loss.

8. A diamagnetic specimen in a gravitational and magnetic trap

A suspended specimen can have exactly the right vertical magnetic force and still slide sideways out of its apparent equilibrium. Derive the force on induced magnetisation, impose gravity, and use the curvature of a measured field to decide whether small disturbances return the specimen to its starting point.

9. Optical trapping and Brownian calibration of a dielectric bead

An image of a bead remaining near a focus does not establish the strength of its trap. Derive the force pulling a small dielectric sphere into a bright region, compare it with the forward force from scattering and learn what a motion record can tell you about stiffness, drag and temperature. The model here uses an inert dielectric sphere in a homogeneous Newtonian liquid.

10. Coupled optical and mechanical resonances

A vibrating mirror can imprint its position on an optical field, while the stored light pushes back on the mirror. An apparent pair of peaks can indicate two hybrid modes, interference at a readout port or two unrelated resonances. Starting from energy in a moving cavity, derive the coupled response and learn which measurements establish a coupling strength and its uncertainty.

11. Thermal radiance, étendue and detector calibration

A square millimetre of a heated object emits thermal radiation into an entire hemisphere. A detector sees only the rays accepted by its optics and only wavelengths to which it responds. Derive the equilibrium spectrum, follow it through an optical collection system and infer what temperature or emissivity a calibrated reading can actually determine.

12. Moving-mirror Doppler metrology

A moving mirror changes optical phase, shifts the reflected frequency and exchanges energy and momentum with the light. In many displacement measurements the classical phase rule is extraordinarily accurate. Derive its exact inertial-frame counterpart, identify which time a detector samples and find the practical uncertainties that usually matter well before a relativistic correction does.

13. Weighing a small deposit with a piezoelectric resonator

Two nearby frequencies appear in the electrical response of one mechanical mode. Derive the series and parallel resonances, then use their different sensitivities to recognise added mass and stray capacitance.

Electromechanics, optics and measurement · 1. When a suspended capacitor loses its stable equilibrium

An electrode suspended on a spring can remain motionless at a bias that makes the calculated electrostatic force grow as it moves. Static force balance alone misses the voltage at which that equilibrium disappears. Deriving the small-signal frequency gives an independent measurement of the same instability.

Start with the electrical and mechanical control variables

A movable conducting plate of area A faces a fixed electrode. Its unstressed gap is g. Define positive displacement x toward the fixed plate, with spring force $-kx$, mass m and viscous damping coefficient b. For $0 \le x < g$ assume ideal parallel plates, negligible tilt and fringing, constant ϵ, and a linear elastic spring. The capacitance is $C(x) = \epsilon A/(g-x)$. The model ends when the plates approach contact or when the suspending structure departs from its specified elastic law. It cannot by itself calculate adhesion, contact resistance or post-contact motion.

Before differentiating electrical energy specify whether the source holds voltage or charge. At fixed charge Q, electrical energy is $Q^2/[2C(x)]$, and its negative displacement derivative gives the force toward increasing x. At fixed voltage V, a source supplies $V\,dQ$ as the plates move. The appropriate potential for mechanical equilibrium is the spring energy plus the field energy minus QV, which here is $\mathcal{U}_V = kx^2/2 - C(x)V^2/2$. Differentiating gives the outward spring reaction and the attractive electrical force $F_e = V^2 C'(x)/2 = \epsilon A V^2/[2(g-x)^2]$. An independent derivation multiplies the conductor pressure $\epsilon E^2/2$ by area with $E = V/(g-x)$. Both give a positive force. Replacing the source potential with the stored energy at fixed voltage would reverse its sign.

Locate all equilibria before choosing a branch

Define $s = x/g$ and $\Lambda = \epsilon A V^2/(2kg^3)$. The static force equation reduces to

$$\Lambda = s(1-s)^2.$$

At a specified voltage one can find s by scanning the curve $s \mapsto s(1-s)^2$ over the physically admissible interval $0 \le s < 1$. Its derivative is $(1-s)(1-3s)$, giving an interior maximum $4/27$ at $s = 1/3$. For $0 < \Lambda < 4/27$ there are two physical roots. The smaller is a local minimum of $\mathcal{U}_V$ while the larger is a local maximum. Their merger is a fold of the static equilibrium branch:

$$x_{\mathrm{pi}} = g/3,$$

$$V_{\mathrm{pi}} = \sqrt{\frac{8kg^3}{27\epsilon A}}.$$

The often-quoted one-third displacement is the result for a linear spring, ideal voltage drive and this capacitance geometry. It is not a universal travel limit determined by the gap alone. Additional mechanical restoring forces or a different electrical boundary condition change the fold. Neither positive net electrical force nor the mere existence of a real solution is evidence of stable operation. A root finder started near

the almost closed gap can converge to the unstable large-displacement root with an extremely small static residual.

The two branches can be distinguished experimentally without deliberately reaching contact. At static equilibrium $kx = F_e$. Linearise the total force about x_0 while maintaining voltage at its prescribed value. The incremental restoring stiffness and corresponding weak-damping natural frequency are

$$k_{\text{eff}} = k - \frac{\epsilon A V^2}{(g - x_0)^3} = k\,\frac{1 - 3s_0}{1 - s_0},$$

$$f_{\text{small}} = \frac{1}{2\pi}\sqrt{\frac{k_{\text{eff}}}{m}}.$$

A small push toward the fixed electrode is restored only when $k_{\text{eff}} > 0$. The familiar electrostatic softening is the reason the smaller equilibrium is stable. The small-signal frequency tends to zero as the fold is approached. With appreciable damping the observed amplitude resonance and ringdown oscillation have slightly different frequencies; fit the full mechanical response rather than identifying an arbitrary peak with the undamped f_{small}.

Background · Can a sudden step pull in a statically stable actuator?

An ideal voltage step applied to a plate initially at rest can push it past the unstable equilibrium. During the motion at fixed voltage the effective potential per kg^2 is $u(s) = s^2/2 - b/[2(1 - s)]$, where $s = x/g$ and $b = \epsilon A V^2/(kg^3)$. Neglect damping and begin at $s = 0$. At the dynamical threshold the initial energy equals the saddle energy and the slope vanishes at that saddle.

$$u(s) = u(0) \implies b = s(1 - s),$$

$$u'(s) = 0 \implies b = 2s(1 - s)^2.$$

The two equations meet at $s = 1/2$, $b = 1/4$. The quasistatic fold gives $b = 8/27$. Consequently the undamped step threshold is only $\sqrt{27/32} \simeq 0.919$ times the slow-ramp pull-in voltage. A device operated between the two voltages can survive a ramp but collapse after an abrupt step. Dissipation and the actual rise time change this prediction.

Reconstruct the prepared state from a frequency measurement

Take a 100 µm square electrode in vacuum, $g = 2.00$ µm, $k = 2.00$ N/m and $m = 5.00 \times 10^{-10}$ kg. The predicted pull-in voltage is 7.32 V, the zero-bias capacitance 44.3 fF and the zero-bias natural frequency 10.07 kHz. On the low branch $s_0 = 0.20$ gives $x_0 = 0.400$ µm and voltage 6.80 V. Then $k_{\text{eff}} = 1.00$ N/m and the small-signal frequency is 7.12 kHz. A measured frequency ratio f/f_0 can be inverted via $s = (1 - r)/(3 - r)$, where $r = (f/f_0)^2$. The result is independent of absolute optical displacement calibration if the mass and underlying spring stiffness have not changed during the bias sweep.

At the same voltage the other physical root is $s = 0.48769$. A numerical force balance at this root would give the same supply reading but a negative tangent stiffness. The ideal voltage-controlled potential in units kg^2 is $u(s) = s^2/2 - \Lambda/(1 - s)$. The barrier height $kg^2[u(0.48769) - u(0.20)]$ for this example is

7.26×10^{-14} J. A sufficiently energetic disturbance can carry the plate out of the stable basin even below static pull-in. This is different from loss of the static equilibrium itself. A quasistatic model predicts the existence and local stiffness of the equilibrium; a transient simulation must additionally specify initial displacement, velocity, drive ramp and damping.

Change the preparation to fixed charge

Disconnect a perfect isolated capacitor after placing a known charge Q on it, neglecting leakage and parasitic capacitance. Then $F_Q = Q^2/(2\epsilon A)$ is independent of the remaining gap, because $D = Q/A$ is fixed. Static equilibrium is $x_0 = Q^2/(2\epsilon A k)$, with incremental stiffness exactly k. At the same initial voltage and gap the instantaneous force agrees with the voltage-controlled value, since $Q = C(x_0)V$. The derivatives governing subsequent motion differ. At $x_0 = 0.400$ µm the required charge in the numerical example is 0.376 pC. After ideal electrical isolation the small-signal frequency returns to the underlying elastic value even though the loaded static position stays the same.

Ideal charge isolation is delicate in a small device. A fixed parasitic capacitance C_p from the isolated node to a reference means the conserved charge belongs to $C(x) + C_p$. Motion changes the fraction of total charge residing on the actuator. The electrical energy becomes $Q_{\text{total}}^2/[2(C + C_p)]$, so the force is $Q_{\text{total}}^2 C'/[2(C + C_p)^2]$, generally displacement dependent. If $C_p \gg C$, the node voltage changes little with a small displacement and behavior is closer to voltage control. Interpreting an observed softening after disconnecting the supply as evidence against electromagnetic force balance would overlook this boundary change.

A different practical preparation puts an intentional series capacitor C_s between a fixed voltage supply and the actuator. Its series combination is $C_{\text{ser}} = \epsilon A/[g - x + \epsilon A/C_s]$. For static mechanical work at fixed supply voltage use $F = V_s^2 C'_{\text{ser}}/2$. Defining an effective electrical spacing $g_e = g + \epsilon A/C_s$, the static fold occurs at $x = g_e/3$ if this lies within the open physical gap. If $\epsilon A/C_s \geq 2g$, contact would occur before any such static fold in the ideal model. This calculation shows a direct connection between passive circuit design and allowed mechanical travel, while making clear that the maximum actuator voltage can differ from the voltage printed on the supply.

Why a sudden voltage change can reach contact below static pull-in

For a rapid electrical step the plate has a position and velocity inherited from the previous preparation. Suppose it starts at rest at its zero-bias position, the supply steps to a constant V much faster than a mechanical period, and mechanical damping is initially negligible. Immediately after the step, energy in the mechanical potential $u(s) = s^2/2 - \Lambda/(1 - s)$ is $u(0) = -\Lambda$ in units of kg^2. Motion under this constant voltage can reach the force barrier if the initial value exceeds the potential at the unstable stationary point. This question can be solved by comparing two energies without integrating an entire time history.

Let s_b be the unstable root of $\Lambda = s_b(1 - s_b)^2$. At threshold the initial state is on the separatrix, the trajectory separating oscillations in the well from excursions to the closing gap. Set $u(0) = u(s_b)$, subtract the two expressions and cancel s_b from the nontrivial solution. The two conditions become

$$\Lambda = \frac{s_b(1-s_b)}{2},$$

$$\Lambda = s_b(1-s_b)^2.$$

Thus $s_b = 1/2$, $\Lambda = 1/8$, and the ideal undamped step threshold is

$$V_{\mathrm{step}}/V_{\mathrm{pi}} = \sqrt{27/32} = 0.919.$$

In our example this is 6.72 V. A slow increase to 6.80 V leads to the observed stable 0.40 µm deflection, whereas a sudden ideal voltage step to the same value from rest at the origin contains enough energy to escape the well. Starting on the biased equilibrium before applying a small perturbation is a different experiment. The direction of the threshold shift makes physical sense: damping removes mechanical energy during the first excursion, and a slow ramp allows the equilibrium to move with the plate. The undamped value specifies a particular ideal initial-value experiment, rather than a voltage rating for an unspecified drive waveform.

One can check a simulated step by monitoring $m\dot{x}^2/2 + \mathcal{U}_V$ after the supply settles. With viscous damping its derivative is $-b\dot{x}^2$, so any sustained increase during a constant-voltage segment points to a numerical or modelling error. Before the step ends there is an explicit drive term $-CV\dot{V}$ in this energy balance. The mechanical potential already includes the source work. Adding $V\dot{Q}$ to the same balance a second time gives an artificial energy imbalance. A transient solver should stop at a physically specified minimum gap instead of marching across the singularity; a statement about contact time depends on what counts as contact.

A finite source resistance changes what a ringdown measures

A dc supply with series resistance R can maintain the static voltage accurately while failing to hold the actuator voltage fixed over a mechanical cycle. Let its steady node voltage be V, with equilibrium capacitance $C = C(x_0)$. A small sinusoidal displacement δx causes current $i\omega(C\,\delta v + VC'\delta x)$ into the actuator. The small-signal supply voltage is zero, so $\delta v = -Ri$. Eliminating current gives

$$\delta v = -\frac{VC'}{C}\,\frac{i\omega\tau}{1+i\omega\tau}\delta x,$$

$$\tau = RC.$$

Linearise $F_e = V^2C'/2$ allowing both voltage and position to vary. The direct derivative is $K_{e0} = V^2C''/2$, while the change from node voltage is $VC'\,\delta v$. For an ideal parallel plate $C''/2 = (C')^2/C$, so the two terms combine to the simple complex relation $\delta F_e = K_{e0}\delta x/(1+i\omega\tau)$. At low frequency the electrical force follows displacement with its fixed-voltage slope. At high frequency the slope vanishes, as it must for an isolated ideal plate carrying fixed charge over one oscillation. At intermediate frequency a motion-induced current dissipates heat in the resistor.

Bring δF_e to the left side of the mechanical equation. Its real part subtracts $K_{e0}/[1+(\omega\tau)^2]$ from the mechanical stiffness; its imaginary part adds $K_{e0}\tau/[1+(\omega\tau)^2]$ to the viscous damping coefficient. This

damping is positive for a passive resistor. At $s = 0.20$ the static electrical stiffness is 1.00 N/m, the operating capacitance is 55.3 fF and $R = 400$ MΩ gives

$$\tau = 22.1\,\mu\mathrm{s}.$$

A 7.2 kHz interrogation has $\omega\tau \simeq 1$. Only about half the static electrical slope survives and the added damping is roughly $11\,\mu\mathrm{N\,s/m}$. Treat these as frequency-dependent coefficients in the driven susceptibility. Substituting the low-frequency stiffness into a ringdown fit at this frequency would give the wrong displacement when the fitted frequency is inverted. Fitting the complex response at several frequencies avoids that ambiguity.

An experiment can separate the circuit effect from mechanical nonlinearity. Keep the same static bias and displacement, insert two known series resistances, then measure both in-phase and quadrature response to a small sinusoidal force. The displaced equilibrium should remain essentially unchanged at dc if leakage stays small. A stiffness dispersion and dissipative peak that shift in proportion to $1/R$ support an electrical cause. Then repeat at a smaller oscillation amplitude. Amplitude-dependent frequency shift that survives the change in resistance can arise from nonlinear spring or force curvature, which the first-order susceptibility leaves out. A simple change to the external circuit has turned the slope of an energy curve into a prediction for a measured phase lag.

Compare force intersections for voltage and charge control

The initial gap is 2.00 µm, the area is a 100 µm square and the mass is 0.50 µg. Use the voltage ratio to move the force curve through the fold. At 93% of pull-in the low-displacement fixed-voltage intersection has a positive restoring slope. The high-displacement intersection has a negative slope. The charge-controlled line is prepared at the selected voltage at the original open gap. It gives the same initial force but a different subsequent curve. The readout reports both open-gap roots even if the upper root lies beyond the plotted interval. At the baseline stiffness, static pull-in is 7.32 V.

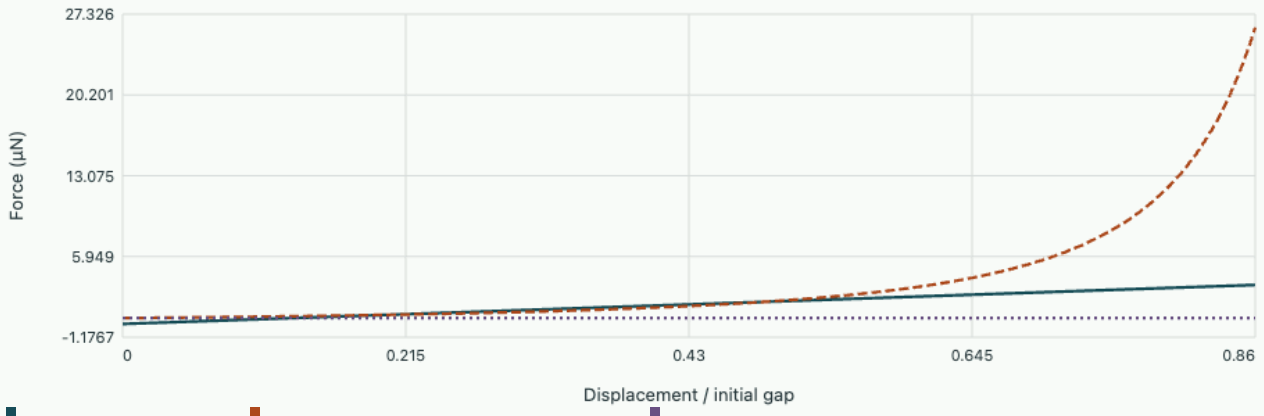
Change the spring stiffness while holding the voltage ratio fixed. The dimensionless locations of the roots stay in place, but their dimensional force and small-signal frequency change. This checks the scaling that eliminated the physical dimensions in the fold equation. Then set the ratio between the dynamic step threshold and unity. Use the force intersections to locate the potential maximum and sketch the area between attraction and spring reaction from the initial position. Its signed integral supplies the work which distinguishes a bounded excursion from crossing the maximum. The displayed curves predict initial-value behaviour only when combined with a choice of initial kinetic energy and damping.

Open this experiment online

Try the calculation

Bias as fraction of static pull-in **0.93**

Spring stiffness (N/m) **2**



Bias 6.805 V; static pull-in 7.317 V. Stable x/g = 0.200, unstable x/g = 0.487, small-signal frequency 7.11 kHz.

Tutorial · distinguish stability from force balance

Infer a gap fraction and a bias from a resonance

An otherwise identical actuator shows $f/f_0 = 0.80$ at a stable fixed voltage. Find the deflection as a fraction of gap, the voltage relative to pull-in, and the independent capacitance change. What could spoil this comparison?

▼ **Work through the solution**

Here $r = 0.64$ and $s = (1 - 0.64)/(3 - 0.64) = 0.1525$. The bias ratio is $\sqrt{(27/4)s(1-s)^2} = 0.860$ approximately. The parallel-plate capacitance should rise by $1/(1-s) - 1 = 18.0\%$. Measure the change relative to the active capacitance after removing any constant wiring capacitance. Disagreement can come from changing mode shape, a nonlinear spring, a tilted plate, a measurement frequency above the electrical equilibration bandwidth, or an incorrect extraction of mechanical stiffness from a strongly damped resonance. With a measured gap and mode shape these explanations yield different quantitative corrections.

Choose a series capacitor to move the static fold

The original 44.3 fF zero-bias actuator should remain on its stable slope as far as $x = 0.60g$, ignoring contact and other limits. Find the series capacitance whose fold falls exactly at this displacement. How much supply voltage appears across the actuator at that point?

▼ **Work through the solution**

Set $g_e/3 = 0.60g$, giving $\epsilon A/C_s = 0.80g$ and $C_s = C(0)/0.80 = 55.4$ fF. At $0.60g$ the actuator capacitance is $C(0)/0.40 = 110.7$ fF. Series charge is identical in both capacitors, so

$$V_{\text{act}}/V_{\text{supply}} = C_s/(C_s + C_{\text{act}}) = 1/3.$$

Choosing exactly this value places operation at vanishing ideal stiffness, so a strictly positive measured restoring slope at that displacement requires a smaller series capacitance or a different margin. The result is sensitive to additional capacitance connected between the series node and ground.

Further reading

- Derive capacitor pressure from field stress and work

Electromechanics, optics and measurement · 2. Which part of a sample changes a four-electrode impedance measurement?

Driving one electrode pair and measuring another yields a transfer impedance whose spatial sensitivity can be positive, negative or nearly zero. A second field solution supplies the sensitivity without perturbing every element of a numerical mesh.

Distinguish the current terminals from the voltmeter terminals

Consider a bounded two-phase electrical sample in a benign laboratory characterisation cell. Drive harmonic current $+I$ through electrode a and collect $-I$ through b. Two independently selected electrodes c, d read the difference $U_c - U_d$ with negligible voltmeter current. The measured transfer impedance is

$$Z_{cd,ab} = (U_c - U_d)/I.$$

A four-terminal measurement reduces some lead-voltage errors, but finite-area electrodes still alter the field and contact processes can affect it. Reversing either pair reverses the signed transfer reading. Interchanging both pairs leaves it unchanged in a reciprocal linear sample.

Adopt phasors $e^{i\omega t}$ and a local complex admittivity $\gamma = \sigma + i\omega\epsilon$, with real positive scalar σ, ϵ for this first model. The total current density is $\mathbf{J} = -\gamma\nabla u$. In the electroquasistatic regime $\nabla \cdot (\gamma\nabla u) = 0$ away from injected current at the boundary. Magnetic induction and propagation are negligible at the chosen cell dimensions and frequencies; checking electrode and cable dynamics remains necessary even when the sample itself is quasistatic. Perfectly conducting measurement electrodes are equipotentials. Inactive electrodes have zero *net* current, which allows charge redistribution over their finite surfaces; a strictly insulating uncovered wall instead requires zero normal current locally.

Use an outward normal $\mathbf{n}$ from the sample. Define positive terminal current into the sample by $I_i = \int_{S_i} \gamma \partial_n u \, dA$; it equals minus the outward conduction-plus-displacement flux. The pair currents sum to zero, and only voltage differences matter, so fix one reference or a mean voltage. Integrate by parts against any test voltage v constant on each ideal electrode, giving

$$\int_\Omega \gamma \, \nabla u \cdot \nabla v \, dV = \sum_i I_i \, v_i.$$

This is a complex bilinear identity for reciprocity and linear sensitivity. Real average heating is obtained separately by conjugating the physical fields when computing power. Confusing those two products can give an incorrect phase or an incorrect sign for an imaginary admittivity perturbation.

Perturb the sample while keeping injected currents fixed

Let ψ_{ab} be the potential per unit current driven from a to b, and ψ_{cd} the solution when current is instead driven between the measurement electrodes. Using each as test function for the other shows directly that

$Z_{cd,ab} = Z_{ab,cd}$. Now change material admittivity by a small distribution $\delta\gamma$ and let the drive field change by $\delta\psi_{ab}$. Holding all currents and electrode geometry fixed, differentiate the weak equation:

$$\int_\Omega \gamma \nabla \delta\psi_{ab} \cdot \nabla v \, dV = -\int_\Omega \delta\gamma \, \nabla \psi_{ab} \cdot \nabla v \, dV.$$

Choose $v = \psi_{cd}$ and apply the unperturbed measurement-pair equation to $\delta\psi_{ab}$. Its left side becomes the change in voltage difference per drive current, hence

$$\delta Z_{cd,ab} = -\int_\Omega \delta\gamma(\mathbf{x}) \, \nabla \psi_{ab}(\mathbf{x}) \cdot \nabla \psi_{cd}(\mathbf{x}) \, dV.$$

The minus sign has an immediate check in a conducting bar: increasing conductivity lowers its resistance at fixed current. The sensitivity density is the overlap of the drive and reciprocal fields, not simply the magnitude of the drive field. If the fields are locally perpendicular a small isotropic admittivity perturbation is invisible to first order in this channel. If the fields are opposed the local real sensitivity can have the opposite sign to nearby regions. For tensor perturbations replace the product by $-\nabla\psi_{cd} \cdot \delta\boldsymbol{\gamma} \nabla\psi_{ab}$. Reciprocity in the above form requires symmetric material tensors.

A one-dimensional experiment tests the sign, normalisation and scale

Take a uniform real-conductivity bar of length 20 mm, cross-section $A = 100$ mm² and conductivity 0.20 S/m. Assume end electrodes cover its faces and ideal voltage taps select the full length in this limiting four-wire experiment. The impedance is

$$L/(\sigma A) = 1000\ \Omega.$$

The drive and reciprocal gradients per ampere have magnitude $1/(\sigma A)$. Increase conductivity in a full cross-section segment 2.0 mm long by 0.020 S/m. The first-order formula predicts

$$\delta R = -\ell \delta\sigma / (\sigma^2 A) = -10.0\ \Omega.$$

Direct series integration of $1/(\sigma A)$ gives a change from 100.0 to 90.91 Ω in the segment, or $-9.09\ \Omega$.

The 0.91 Ω difference is consistent with truncating at first order for a 10% conductivity change. Recompute at half the material perturbation: the exact change becomes $-4.762\ \Omega$, only 0.238 Ω away from the first-order $-5.00\ \Omega$. The residual decreases by about four when the step halves. Repeating at several positive and negative perturbations and taking a centred difference provides a sharper test of the adjoint code. Varying only the mesh resolution with a fixed large material contrast cannot remove a genuine nonlinear remainder.

The same one-dimensional geometry also demonstrates an absolute identifiability limit. Interchanging two equal-length full-cross-section inserts with different conductivities leaves the end-to-end resistance unchanged, since the integral of reciprocal conductivity is independent of ordering. Perfect measurement

precision cannot recover the order. A voltage tap between the two inserts or an alternative contact geometry is needed to break that symmetry.

Interpret real and imaginary voltage readings together

Even if each local σ, ϵ is real, the fields ψ in an inhomogeneous sample usually carry spatially varying phase. A change $\delta\sigma$ can therefore change both the real and imaginary transfer impedance, as can $i\omega\delta\epsilon$. Treat a complex reading as two real measurements with their calibrated covariance. To first order one finite-element cell with piecewise-constant parameters contributes coefficients $-\int_{\text{cell}} \nabla\psi_{ab} \cdot \nabla\psi_{cd}\, dV$ for $\delta\sigma$ and $i\omega$ times that coefficient for $\delta\epsilon$. The units of the two unknowns differ, so rescale with plausible physical parameter changes before comparing singular values.

As a simple frequency check choose a uniform engineered dielectric-conductor composite with $\sigma = 0.010$ S/m and $\epsilon_r = 1000$. At $f = \sigma/(2\pi\epsilon) \simeq 180$ kHz, $\gamma = \sigma(1 + i)$. The same bar geometry then gives $Z = (L/A)/\gamma = 10.0(1 - i)$ kΩ, a phase of -45°. At low frequency the phase tends to zero and the real impedance to 20.0 kΩ. A constant cable delay produces phase proportional to frequency independent of the sample parameters, so record a calibrated reference at each frequency before attributing an observed phase change solely to the specimen.

Derive a spatial sensitivity that can be checked without a mesh

Consider a homogeneous, thick conducting calibration block with a flat, electrically insulating top except at four small contacts on the same line. Coordinates are x, y along the surface and $z > 0$ into the block. If a unit current enters through a point at the surface, it spreads over hemispheres of radius r. Uniform normal current through a hemisphere requires

$$J_r = 1/(2\pi r^2).$$

The radial field is J_r/σ; integrating from r to infinity gives potential $1/(2\pi\sigma r)$. The coefficient differs from a point source in the interior where an entire sphere carries the current. Subtract a corresponding solution for the terminal extracting the current to satisfy conservation.

Let the drive pair be at $x = \mp a$ and the sense pair at $x = \mp c$, with positive terminal at negative x in each case. At surface points separated from the point contacts the baseline transfer impedance becomes

$$Z_0 = \frac{1}{\pi\sigma}\left(\frac{1}{a - c} - \frac{1}{a + c}\right) \qquad (0 < c < a).$$

For $a = 15$ mm, $c = 5$ mm and $\sigma = 0.20$ S/m this is 79.58 Ω. A 1 mA real current gives 79.58 mV. The same number follows on interchanging the drive and measurement pairs. The large individual contact potentials of an ideal point source are singular at the driving contact itself, yet the transfer potential evaluated at a distinct contact is finite. A two-terminal impedance obtained by setting $c = a$ is divergent in this point-contact approximation. To make such a measurement calculable one must assign physical electrode radii, contact laws and current distributions.

Place a small spherical inclusion of radius 1.0 mm at $x = y = 0$ and depth z, with conductivity just 0.020 S/m above background. At the centreline the vertical field components cancel in each paired solution. Differentiating the two point-source potentials horizontally gives a gradient magnitude $a/[\pi\sigma(a^2 + z^2)^{3/2}]$ for a drive half-spacing a; replace a by c for the reciprocal measurement field. The first-order change in impedance is the product of the two signed gradients times minus the conductivity perturbation and the sphere volume. The centre sample approximates the volume integral when both background gradients vary slowly over the sphere.

At $z = 8$ mm the result is -3.86 mΩ, a $-3.86\,\mu\mathrm{V}$ perturbation at 1 mA, about 48 parts per million of baseline. Increasing depth to 16 mm leaves only -0.320 mΩ. A point midway between one source and one sensing contact can experience a very different gradient alignment from a point on the common centreline. This is why a general inclusion shift should be calculated by differentiating the vector field before taking the dot product. Replacing each gradient by its magnitude would forbid the experimentally possible positive changes. A point inclusion moved twice as deep does not, in general, change the answer by one universal inverse power. Each field samples a different combination of depth and contact separation.

Retain unknown contact impedance as a separate physical parameter

A metal current terminal is nearly equipotential internally but can have an interface voltage drop at the sample. Let U_i be its metal potential, u the potential on the adjacent sample, $j_i = \gamma \partial_n u$ positive into the body and z_i an area-specific contact impedance in Ω m². A simple local interface law is $U_i - u = z_i j_i$, together with $\int j_i \, dA = I_i$. At an inactive metal terminal only this integral is zero. Current can enter on one side and return on the other if the sample potential varies appreciably across its face. A high-impedance voltmeter does not by itself make that metal patch equivalent to an absent electrode.

For sample and electrode test fields (v, W), integration by parts and the contact law give

$$B((u,U),(v,W)) = \int_\Omega \gamma \nabla u \cdot \nabla v \, dV + \sum_i \int_{S_i} \frac{(U_i - u)(W_i - v)}{z_i} \, dA = \sum_i I_i W_i.$$

Here the current and potential gauges are chosen consistently on both drive and reciprocity solves. The contact product is as symmetric as the bulk product, so finite reciprocal contact impedances retain equality upon drive/sense exchange. If the contact changes, differentiating $1/z_i$ introduces a minus, and the minus from differentiating the full field equations cancels it. The additional first-order sensitivity is $+\sum_i \int \delta z_i j_{i,d} j_{i,m} \, dA$, with both current densities normalised to unit terminal current. In a uniform two-terminal contact of area A, the two normalised currents are $1/A$ and one contact contributes $+\delta z/A$, as an ordinary series contact should. This independent limit fixes both the positive sign and the square-metre convention.

For a controlled reconstruction, form one column per candidate sample perturbation and another per uncertain contact coefficient. Weight measured channels by their noise covariance. Two nearly proportional columns then warn that bulk variation and degrading contact cannot be separated by those channels over that bandwidth. Reciprocity catches lead swaps and some drift, but a reciprocal change in a contact can pass this check. Repeat with a different electrode pattern, or calibrate against a

homogeneous reference block at the same frequency. To verify a numerical sensitivity, perturb the contact alone by equal positive and negative amounts and compare the centred impedance difference with the surface integral. Then perturb the sample alone and compare with the volume integral. A successful combined test could otherwise conceal equal and opposite implementation errors.

Move a weak inclusion beneath four surface contacts

The block is treated as a homogeneous half-space except for a 1 mm radius sphere with a positive conductivity perturbation. Each plotted point uses the two unperturbed analytical vector gradients at the sphere centre. The second curve repeats the calculation at twice the selected depth. At 8 mm and 5 mm sensing half-spacing the predicted centre perturbation is −3.86 mΩ against a baseline of 79.58 Ω. Look for lateral positions where the sign changes before interpreting increased conductivity as a uniformly smaller transfer impedance.

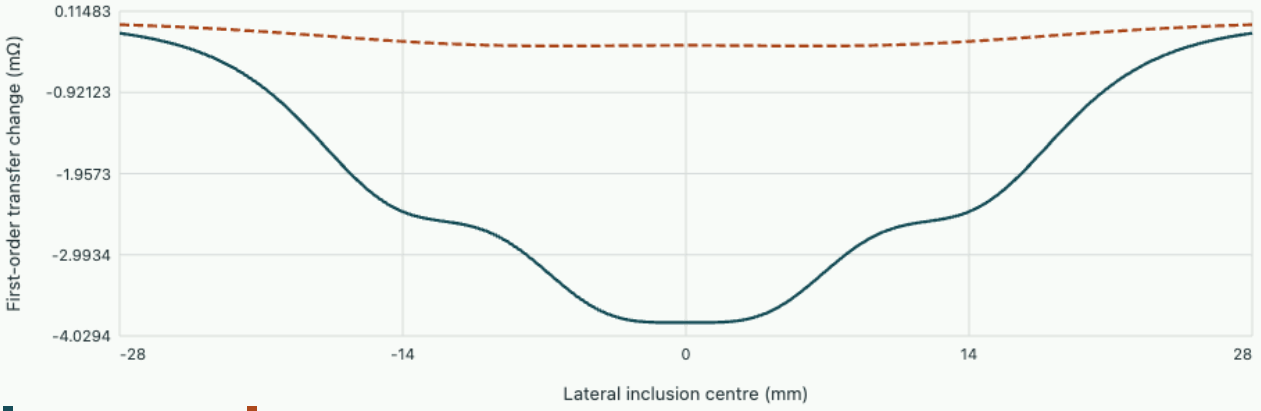
Widen the sensing pair and predict how both the transfer baseline and the local sensitivity will respond. A larger transfer voltage does not guarantee a proportionately larger fractional change for a chosen deep inclusion. The ideal point-contact fields vary steeply close to electrodes. When the inclusion radius approaches the distance over which the gradients change, volume quadrature supplies a better approximation than evaluating a single point. A finite laboratory specimen also needs boundaries sufficiently distant compared with the patterned contact spacing.

Open this experiment online

Try the calculation

Inclusion depth (mm) 8

Sensing half-spacing (mm) 5



Centre perturbation -3.858 mΩ; unperturbed transfer 79.58 Ω. Sphere radius 1 mm, conductivity increase 0.020 S/m.

Tutorial · bound a perturbation and design another observation

Predict a transfer change without solving again

A field computation for unit drive and unit reciprocal current in a small inclusion gives approximately parallel real field gradients of magnitudes 1500 and 400 V/(m A). The inclusion volume is 2 mm³. Its conductivity rises by 0.010 S/m. Predict the resistance change and the observed voltage change at a 5 mA drive. What changes if only the measurement electrodes are reversed?

▼ **Work through the solution**

The overlap is $6.0 \times 10^{5}\ \mathrm{V}^2/(\mathrm{m}^2\mathrm{A}^2)$. Multiplying by $2 \times 10^{-9}\ \mathrm{m}^3$ and the material perturbation gives $\delta Z = -1.2 \times 10^{-5}\ \Omega$, or -12 μΩ. At 5 mA the transfer voltage changes by -60 nV. Reversing only c, d reverses the reciprocal field and the transfer change. The result assumes nearly uniform overlap inside the inclusion and small perturbation. Subdividing the inclusion or independently perturbing the forward model tests these assumptions differently. A voltmeter's microvolt accuracy would be inadequate for the proposed single measurement.

Do more electrodes resolve two unknown defects?

Two calibrated real channels have sensitivity rows proportional to $(1, 0.8)$ and $(0.8, 1)$ for two dimensionless defect amplitudes. Which combination is measured poorly? Would two additional measurements that simply reverse the existing electrode signs supply a new material constraint?

▼ **Work through the solution**

The common pattern has gain 1.8 and the difference pattern gain 0.2; their singular-value ratio is nine in these units. Distinguishing which defect changed is harder than detecting a common material change. Sign reversals duplicate the same rows up to a minus sign. They can cancel offsets or improve precision by averaging independent noise, but do not increase the rank of the noise-free sensitivity matrix. A drive/measurement pair with opposing sensitivities to the two spatial regions would directly strengthen the weak difference direction.

Further reading

- Derive electrostatic reciprocity from a boundary-value problem
- Check complex field conventions with an independent power balance
- Somersalo, Cheney and Isaacson · existence and uniqueness for the complete electrode model

Electromechanics, optics and measurement · 3. Measuring the force on a sliding dielectric

A suspended strip slides between two electrodes without touching them. Predict the pull, infer the air clearance from a capacitance trace and find where an electrical measurement can detect a change in mechanical stiffness.

Reduce a three-dimensional field to a capacitance gradient

A light insulating strip is attached to a translation stage by a spring. Two stationary metal plates extend a length L in the direction of motion and have width w and separation d. The leading face of the strip is at coordinate x, measured from the entrance to the plates. The strip is longer than the active length, has thickness $t < d$, and fills the width in the interior to good approximation. The combined air clearance above and below it is $d - t$. Use V for the terminal voltage and increasing x for insertion. Electrode lengths and widths appreciably exceed the gap. Which information about the detailed electrostatic field is needed to predict a force in the horizontal direction?

First consider one square metre of plate well inside the overlap region. A pillbox straddling an air–solid interface contains no free charge, so the normal displacement D has the same value in air and solid. The fields are D/ϵ_0 and $D/(\epsilon_0\epsilon_r)$. Integrating the field through the layers gives the series law per unit area. Another square metre outside the strip has the gap filled entirely with air.

$$c_f = \frac{D}{V} = \frac{\epsilon_0}{d - t + t/\epsilon_r},$$

$$c_a = \frac{\epsilon_0}{d}.$$

Adjacent columns share the same voltage. For a small interior displacement dx, area $w\,dx$ changes from the air value to the filled value. Thus $dC = w(c_f - c_a)\,dx$. Define $K = w(c_f - c_a)$, measured in farads per metre. For a rigid straight edge well inside the plates it is nearly constant. Any fixed capacitance from leads adds to the measured C and drops out of K. Guard electrodes, if fitted, must be included in the electrical boundary conditions. Moving the guard or measuring a different pair of terminals can change the gradient itself.

For $L = 30$ mm, $w = 20$ mm, $d = 0.50$ mm, $t = 0.40$ mm and $\epsilon_r = 4.0$, the series thickness is 0.20 mm. The effective relative permittivity of a filled column is 2.50. The ideal empty capacitance is 10.63 pF and the filled capacitance is 26.56 pF. The slope is 0.5313 pF/mm. A naive average of permittivity by volume gives 3.4 and overestimates the force. It would apply to columns side by side in a uniform tangential field. The orientation of the interface supplies the missing piece of information.

Account for the power supply during a displacement

Imagine changing x very slowly at fixed voltage. The source transfers $dQ = V\,dC$ through voltage V, supplying $V^2 dC$. The energy in the field increases by $\frac{1}{2}V^2 dC$. Their difference is the mechanical work $F\,dx$ delivered to the moving strip. For an electrically isolated pair, Q is constant and energy $Q^2/(2C)$ decreases with increasing insertion. Differentiating either description gives

$$F_V = \frac{1}{2} V^2 C'(x),$$

$$F_Q = \frac{Q^2}{2C(x)^2} C'(x).$$

Their instantaneous values agree when $Q = C(x)V$. Their evolution after a displacement generally differs. At 80 V the plateau force is $1.700\,\mu\mathrm{N}$. A spring of $0.20\,\mathrm{N/m}$ allows an extra $8.50\,\mu\mathrm{m}$ of insertion if the slope stays constant. The source supplies 102.0 nJ over a complete ideal 30 mm sweep. Mechanical work is 51.0 nJ and the field gains 51.0 nJ. Make this independent check by evaluating each endpoint energy separately. With the charge prepared on empty plates then disconnected, the final voltage is

$$80/2.5 = 32.0\,\mathrm{V}.$$

The final plateau force approached from the interior falls to $0.272\,\mu\mathrm{N}$. Once the slab is fully inside, the edge term makes the force decay toward zero.

How can a vertical field pull a slab horizontally? Deep in either uniform column the local translational symmetry prevents a longitudinal push on a uniform piece of material. There is a transition near boundaries where the field curves and the polarisation changes. The work calculation moves material through this transition while keeping the remote uniform regions trackable. It predicts the integral of the force without locating each contribution. A force calculation based only on a vertical component times a horizontal normal lacks the tangential and normal stress components on a closed surface. Evaluating a vacuum Maxwell-stress integral on a contour around a finite strip gives a physically independent check of the net result when the supporting mechanical contact is accounted for separately.

Infer an edge profile from a measured trace

An abrupt switch in C' at the end of a plate would imply an unlimited second derivative. There is a finite interaction length around any real edge. A convenient interpolation for fitting data starts from the ideal rectangular overlap derivative and smooths each step over a length ℓ. Take the measured empty capacitance, including any fixed wiring, to be C_b. A symmetric fit for the measured gradient and its integral is

$$C'(x) = \frac{K}{2}\left[\tanh\!\left(\frac{x}{\ell}\right) - \tanh\!\left(\frac{x-L}{\ell}\right)\right],$$

$$C(x) = C_b + \frac{K}{2}\left[L + \ell \ln\cosh(x/\ell) - \ell \ln\cosh((x-L)/\ell)\right].$$

At large negative x the bracket tends to zero, at large positive x it tends to $2L$. The area under the gradient therefore produces the correct total increment KL for every ℓ. This normalisation is important. A fit that reproduces the centre of each transition but changes the area also changes the predicted work in a complete insertion. The two edges can have different widths near asymmetric supports. A fit should then keep the capacitance endpoints and the physical separation of the stations as independent experimental checks.

For $L \gg \ell$ the first isolated transition reaches 10% and 90% at $x/\ell = \mp\,\mathrm{atanh}(0.8)$. Their separation is 2.1972ℓ. An observed 1.10 mm separation implies $\ell = 0.501$ mm. At the entrance the voltage-controlled gradient of force is

$$F_V' \simeq V^2 K/(4\ell);$$

the spring sees $k_{\mathrm{eff}} = k - F_V'$. At the exit the slope reverses and the oscillation frequency increases. For the numerical stage the entrance subtracts 0.00170 N/m from its 0.20 N/m spring. The ratio of loaded to zero-bias resonance frequencies is about 0.9957, if the moving mass is unchanged. The middle of the strip gives the largest static displacement but essentially no frequency shift under ideal voltage control. Measuring displacement and frequency at the same point distinguishes a force from a force gradient.

Estimate a derivative with enough resolution to trust its sign

A bridge returns C at positions separated by a calibrated step h. For the central derivative $D_h = [C(x+h) - C(x-h)]/(2h)$, expansion of both observations about their midpoint gives

$$D_h = C'(x) + h^2 C'''(x)/6 + \cdots.$$

With independent capacitance noise of standard deviation s_C per reading, the standard deviation of the slope is $s_C/(\sqrt{2}h)$. Shrinking h reduces truncation bias but magnifies random fluctuations. This competition reverses the familiar intuition that the smallest encoder step necessarily gives the most accurate force. A slowly drifting bridge offset can cancel between close readings, whereas a drift proportional to time can imitate a slope if all positions are visited in only one direction.

Suppose each measurement has 0.005 pF rms noise and $h = 0.25$ mm. The slope noise is 0.0141 pF/mm, 2.7% of the plateau. At fixed 80 V this corresponds to 45 nN rms in a single force estimate. Twenty-five independent pairs bring this contribution down to 9 nN. An uncertainty of 0.5% in voltage causes a 1% scale uncertainty in force shared by all the pairs. Repeated measurement cannot cancel a common scale uncertainty. At the exact centre of an isolated symmetric edge the symmetric difference of the integrated profile returns exactly $K/2$: opposite gradient values sum to K. The surrounding curve can still suffer marked smoothing. A correct midpoint alone provides a weak validation of a fitted transition length.

A complementary measurement clamps the stage at positions separated by several millimetres and obtains the force from a calibrated spring deflection. Fit the measured forces against V^2 at each position with an intercept free to capture a mechanical preload. A force caused by a contact potential V_c would instead have its parabola centred at $V = V_c$. Measurements of both voltage polarities reveal this offset; a single magnitude and a zero point cannot distinguish it from a biased slope. Perform both insertion and withdrawal sweeps. Hysteresis in the measured spring position with almost identical $C(x)$ indicates mechanical friction or creep rather than a change of the electrostatic energy law.

From a fitted edge to a field calculation

A two-dimensional cross-section can test which part of the interpolation is determined by geometry. In the air and in the strip solve $\nabla \cdot (\epsilon \nabla \phi) = 0$. Prescribe metal potentials, extend the exterior region far enough

to include fringing field and match continuous potential and normal displacement at each material boundary. At each x, calculate $C(x)$ in two independent ways. Integrate the normal displacement on the driven electrode to get Q/V. Integrate $\epsilon|\nabla\phi|^2$ over every dielectric region and divide by V^2 to get the energy definition of capacitance. Both quantities per unit width must be multiplied by the same model width. Integrating only over the material strip removes the energy of the air gap, precisely the storage mechanism that reduced the effective permittivity.

To find force by differentiating numerical energies, use the same domain, mesh quality and quadrature on displacements either side of the evaluation point. A remeshing error uncorrelated from point to point enters like the measurement noise in the preceding calculation. Vary the displacement step and repeat with a finer mesh. A range of steps where the force moves little under both changes is stronger evidence than obtaining a smooth potential plot. Resolve thin clearance layers explicitly or represent them using a validated thin-layer boundary condition. Test the sharp-edge idealisation by changing the radius of the metal edge. The integrated force can converge even while a pointwise maximum field remains sensitive to this radius.

Finally, ask which geometric parameters can be determined from the plateau. Its single slope measures $d - t + t/\epsilon_r$ if w and d are known. It cannot separately determine thickness and relative permittivity. The electrical endpoint change provides KL, which is the same measurement multiplied by L; including both in a fit as independent material constraints can produce an artificially small confidence interval. Measuring t mechanically, or repeating with material samples of independently known thickness, resolves the ambiguity. The observed edge widths provide other geometric information but require a field model of the support and nearby shields before they can be interpreted as a direct permittivity measurement.

Measure curvature with a small carrier modulation

Capacitance-gradient force can be checked while a stage is held near a prescribed operating position. Bias the plates at V_0 and add a small sinusoidal voltage $v\cos\omega t$, where v is a peak amplitude in this experiment. With the stage almost stationary, expand the square of this voltage. The constant force is $C'(V_0^2/2 + v^2/4)$, the force at ω has peak amplitude $C'V_0 v$, and the force at 2ω has amplitude $C'v^2/4$. The latter persists at zero dc bias. Reversing the bias reverses the first harmonic while leaving the second harmonic unchanged. These are experimentally separable predictions of the local energy law even if the absolute mechanical preload is poorly known.

On the plateau of the example, choose $V_0 = 40.0$ V and $v = 2.0$ V. The first-harmonic force is 42.5 nN and the second-harmonic force is 0.531 nN. Well below mechanical resonance a 0.20 N/m spring turns these into peak oscillations of 0.213 µm and 2.66 nm. Close to resonance the spring response instead multiplies each force by its own complex susceptibility at ω or 2ω. A measured ratio of displacement harmonics should therefore be divided by the ratio of those susceptibilities before attributing any difference to dielectric nonlinearity. Sweep the modulation frequency slightly or use a mechanical shaker calibration to measure the transfer function independently.

One subtle source of false harmonics is the motion itself. Near the entrance C' changes over a distance of order ℓ. A vibration that appreciably samples this length modulates the force even at constant bias. To use the harmonic test above, verify that the measured oscillation is small compared with ℓ and that any mean position shift due to bias has been accounted for. On withdrawing the dielectric entirely, both force

harmonics should tend to the background associated with moving cables or the support. This empty-stage control tests the often implicit assumption that the wire capacitance is independent of stage position. A fixed added cable capacitance changes a charge-controlled experiment, but only a position-dependent contribution can imitate a capacitance gradient at fixed voltage.

Compare a connected and an isolated moving strip

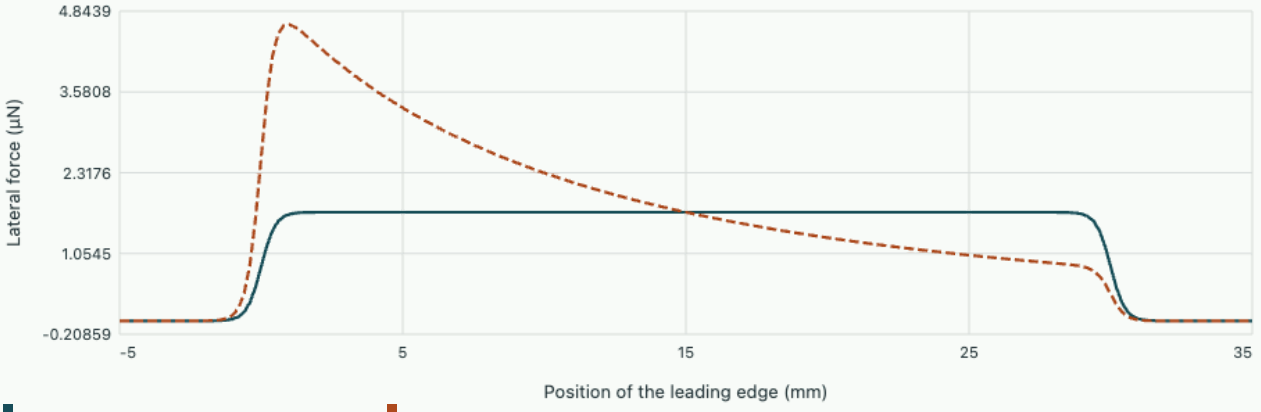
Follow the leading edge from outside the plates to beyond the exit. In the connected experiment the interior pull stays near 1.70 µN at 80 V. In the isolated experiment the charge is prepared at half insertion and kept fixed; no extra cable capacitance is included here. Find the common point of the two curves, then increase the fitted transition length. The peak displacement derivative of force changes even though the integrated voltage-controlled work over a complete traversal remains the same. Predict the sign of the resonance shift on each side of the device.

Open this experiment online

Try the calculation

Bias at preparation (V) **80**

Edge transition length (mm) **0.5**



Interior fixed-voltage pull 1.700 µN; entrance force gradient 1.700 mN/m. The two preparations coincide at 15 mm.

Tutorial · infer a material property and test the energy balance

Recover a clearance and assign an uncertainty

The plateau is 0.500 ± 0.015 pF/mm. Width and plate separation retain their nominal values; a mechanical gauge gives $t = 0.400$ mm. What permittivity follows? If instead $\epsilon_r = 4$ is accepted, what total clearance follows?

▼ **Work through the solution**

Calculate the equivalent thickness from $d_e^{-1} = d^{-1} + K/(\epsilon_0 w)$. The observed K gives $d_e = 0.2073$ mm. With known thickness,

$$\epsilon_r = t/(d_e - d + t) = 3.73.$$

Differentiation gives $|\delta d_e| = d_e^2 |\delta K|/(\epsilon_0 w) = 0.00364$ mm, corresponding to approximately 0.13 in relative permittivity with other quantities held fixed. A positive fluctuation of slope gives smaller equivalent thickness and larger inferred permittivity.

With known permittivity and unknown clearance $a = d - t$, write $d_e = a + (d - a)/4$. Then $a = (4d_e - d)/3 = 0.1097$ mm, with uncertainty 0.0049 mm. Substitution into the original series expression reproduces the observed slope. These two inferences answer different experimental questions. The same pair of endpoints cannot establish both estimates simultaneously without the extra information specified in each part.

Compare two preparations over one short motion

At half insertion connect 80 V, then isolate the terminals including 5.00 pF of fixed wiring. Use the ideal interior geometry. Estimate force after a further 5.00 mm motion and calculate the released mechanical work directly from endpoint energies.

▼ **Work through the solution**

The initial total capacitance is $10.63 + 7.97 + 5.00 = 23.60$ pF, so $Q \simeq 1.888$ nC. The final value is 26.25 pF and the terminal voltage falls to 71.9 V. The final force is $1.37\,\mu\mathrm{N}$. The energy released is $Q^2(1/C_i - 1/C_f)/2 \simeq 7.65$ nJ. Independently integrate $Q^2 K/[2(C_i + Kx)^2]$ over the 5 mm path. The antiderivative is $-Q^2/[2(C_i + Kx)]$ and gives the same result. The mean force of $1.53\,\mu\mathrm{N}$ lies between the initial 1.70 and the final 1.37. Leaving the wiring out of this calculation leaves the voltage-controlled plateau unchanged but predicts an incorrect final voltage and energy release in isolation.

Further reading

- Field force from a closed stress surface
- MIT 6.013, chapter 6 · capacitor work and dielectric forces

Electromechanics, optics and measurement · 4. Designing a displacement bridge at its noise limit

A moving electrode changes two capacitances in opposite directions. Derive the displacement signal, reconstruct its noise in metres and select a carrier voltage consistent with measured mechanical stiffness and electrical losses.

Find the voltage of the moving electrode before calculating sensitivity

Consider three plane electrodes with common active area. At the centred position the middle conductor is separated by g from each outer electrode and each gap has capacitance C_0. Displacement x toward the positive drive electrode reduces one gap and increases the other. Use $s = x/g$, where $|s| < 1$. Drive the outer electrodes with coherent sinusoidal voltages $+V_d$ and $-V_d$, with V_d the rms amplitude of either electrode relative to ground. The middle conductor leads to a high-impedance amplifier. Its cable and input provide a fixed total capacitance C_p to ground. Leakage is initially negligible over one carrier period; the input draws negligible signal current. All formulas for carrier voltages in this project use rms amplitudes.

$$C_+ = \frac{C_0}{1 - s},$$

$$C_- = \frac{C_0}{1 + s}.$$

The middle node has zero net charge at the carrier frequency. Its instantaneous ac charge is $C_+(V_n - V_d) + C_-(V_n + V_d) + C_p V_n$. Set the sum to zero and solve the resulting linear equation for V_n. After combining the gap capacitances the result simplifies considerably:

$$\frac{V_n}{V_d} = \frac{C_+ - C_-}{C_+ + C_- + C_p} = \frac{2C_0 s}{2C_0 + C_p(1 - s^2)}.$$

For zero parasitic capacitance the response is exactly $V_d x/g$ throughout the ideal geometry. Each individual capacitance is strongly curved near an electrode. Charge conservation divides one nonlinear combination by another and removes that curvature from their voltage ratio. Additional capacitance reduces the small-signal gain and introduces an odd cubic term. Let $\eta = 2C_0/(2C_0 + C_p)$. Expanding about the centre gives

$$V_n/V_d = \eta s[1 + (1 - \eta)s^2 + \cdots].$$

The derivative at zero is $K = \eta V_d/g$ in volts per metre. A small-signal calibration taken only at the origin implicitly accepts this cubic distortion at larger displacements.

For $C_0 = 2.00$ pF, $C_p = 4.00$ pF, $g = 100\,\mu\text{m}$ and $V_d = 1.00$ V, $K = 5000$ V/m. A displacement of 100 nm produces 0.5000 mV. At 20 µm, the exact signal is 0.10204 V against the linear prediction 0.10000 V. The 2.04% curvature would be apparent in a calibration over the wider range. A single proportional calibration cannot remove it. A fit to the ratio itself has a physical parameter, $C_p/(2C_0)$, which can also be checked by measuring total input capacitance with the mechanical element held at centre.

Translate each readout fluctuation into equivalent displacement

Represent the amplifier near the carrier by an input series voltage noise density e_n and a shunt current noise density i_n. With the drives treated as small-signal grounds for a noise perturbation, the central node presents $C_T = 2C_0 + C_p$ at the centre. The impedance magnitude for current fluctuations is $1/(\omega C_T)$. For mutually independent noise sources and a narrow demodulation band the corresponding input-referred displacement amplitude spectral density is

$$S_{x,e}^{1/2} = \frac{1}{K}\sqrt{e_n^2 + \left(\frac{i_n}{\omega C_T}\right)^2}.$$

Use noise densities and gain referred to the same detection channel. An rms sine amplitude, a peak-amplitude display and a lock-in reading with an unreported normalisation can differ by factors involving $\sqrt{2}$. Here noise figures denote the effective one-sided densities referred to the recovered rms carrier channel. For a white spectrum and an equivalent noise bandwidth B_e the displacement standard deviation equals the density times $\sqrt{B_e}$. The displayed cut-off frequency of a multi-pole lock-in filter need not equal its equivalent noise bandwidth. A calibration using a known injected sideband measures both the gain convention and the actual band response.

At 100 kHz, $e_n = 10\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$, $i_n = 10\,\mathrm{fA}/\sqrt{\mathrm{Hz}}$ and $C_T = 8.00$ pF, current noise produces $1.99\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$. Adding variances gives $10.20\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$, or $2.04\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$ at the stated gain. A 100 Hz equivalent noise bandwidth gives 20.4 pm rms of electrical readout noise. Cable capacitance reduces the voltage made by current noise but also reduces signal gain. In fact $C_T K = 2C_0 V_d/g$, so their ratio for current noise is independent of C_p in this model. The voltage-noise contribution grows with cable capacitance. Higher carrier frequency initially reduces the current contribution but reaches a plateau set by voltage noise.

Derive backaction for a floating sensing plate

Electrical drive can change the mechanical stiffness of the sensor. For each gap take half the squared rms gap voltage times its capacitance derivative. This follows either by averaging the instantaneous plate pressure over a carrier cycle or by differentiating electrical coenergy with the middle node at its charge-balanced voltage. Mechanical motion must be slow compared with the carrier. Since C_-' is negative, the two signed forces cancel at the centre. For a small positive s, write $V_n \simeq \eta V_d s$,

$$C_+' \simeq (C_0/g)(1+2s),$$
$$C_-' \simeq -(C_0/g)(1-2s).$$

The leading imbalance is

$$F = \tfrac{1}{2}C_+'(V_d - V_n)^2 + \tfrac{1}{2}C_-'(-V_d - V_n)^2 \simeq \frac{2C_0V_d^2}{g^2}(1-\eta)x.$$

The electromagnetic stiffness $K_e = 2C_0V_d^2(1-\eta)/g^2$ subtracts from the mechanical restoring stiffness k. If $C_p = 0$, the equal field magnitudes on both sides cancel exactly for every displacement in ideal plane

geometry. To check directly, the positive gap field is

$$[V_d(1-s)]/[g(1-s)] = V_d/g.$$

The negative gap field has the same magnitude. A charge amplifier holding the central electrode at virtual ground would change this condition and have a different backaction; its calibration must use its own circuit. For a finite C_p, the measured change of resonance with carrier voltage can identify electrical coupling independently of the signal voltage.

With $k = 0.020$ N/m the nominal configuration has $K_e = 0.00020\,\mathrm{N/m}$ at 1 V and $0.00080\,\mathrm{N/m}$ at 2 V. At 2 V the resonance ratio for the same effective mass is

$$\sqrt{1 - 0.0008/0.020} = 0.980.$$

A 2% frequency decrease measures roughly a 4% reduction of stiffness. If the displacement is imposed by a very stiff external positioner this effect changes the force needed to hold position. If the sample is itself a force sensor, the same effect changes the mechanical conversion between applied force and displacement. Quoting an electronic noise floor without specifying which mechanical situation is intended can hide the larger uncertainty.

Choose a feasible drive and include Brownian motion

Raising the carrier at fixed frequency increases displacement signal. It may also heat the dielectrics, exceed amplifier headroom or significantly soften the spring. Suppose each gap at centre has an effective loss tangent $\tan\delta$, with a real conductance $\omega C_0 \tan\delta$ at the carrier. Its dissipated power is conductance times the squared rms voltage. Add both gaps to find

$$P = 2\omega C_0 \tan\delta V_d^2.$$

Fixed parasitic capacitance at the stationary middle node has negligible carrier voltage at exact balance in this model. An actual shield or cable dielectric connected directly to one of the drive nodes will have its own loss. A calorimetric or current phase measurement should include that extra load.

With effective $\tan\delta = 0.010$, the two active gaps dissipate 25.13 nW per squared volt at 100 kHz. A 100 nW heat budget gives $V_d \le 1.995$ V. A driver that remains linear to 2.50 V and a softening allowance $K_e \le 0.1k$ permit 2.50 V and 3.16 V respectively. The thermal bound sets the largest practical drive under these assumptions. For the electrical readout alone, the best feasible setting is thus 1.995 V and the electrical floor is about $1.02\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$. A separate limit on the electric field or measurement amplitude could be introduced as another experimentally justified constraint.

Now let the mechanical element have viscous damping $b = 1.00 \times 10^{-6}\,\mathrm{N\,s/m}$ at $T = 300$ K. In thermal equilibrium the one-sided Brownian force density is

$$\sqrt{4k_B T b} = 1.29 \times 10^{-13}\,\mathrm{N}/\sqrt{\mathrm{Hz}}.$$

Far below mechanical resonance its displacement density is this force divided by $k - K_e$. At 1 V it is $6.50\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$, already greater than the electrical term. At 1.995 V it is about $6.70\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$. The combined density adds in quadrature, giving about 6.81 at 1 V and 6.78 at the heat limit. A minimum of the total can occur between them. For an imposed position of a rigid stage the Brownian motion of this compliant probe may be outside the measurement model; for a freely suspended sensing mass it is part of the signal uncertainty. The optimisation criterion follows the physical quantity the instrument measures.

Use the phase and polarity of a calibration to track false signals

Middle-node leakage to ground with conductance G changes the circuit equation to $[G + i\omega C_T]V_n = i\omega(C_+ - C_-)V_d$ at small displacement. The ideal gain acquires a factor $i\omega C_T/(G + i\omega C_T)$. For a 10 MΩ leakage path and 8 pF, the corner is 1.99 kHz. At 100 kHz the magnitude reduction is about 0.02% and the phase lead about 1.14°. At 2 kHz both quadratures matter. A drift in humidity that lowers the resistance can look like a gain change if only the in-phase channel is recorded. Measure a known displacement at two carrier frequencies and fit both quadratures before interpreting it as a change in geometry.

Suppose the two actual drives are $V_d(1 + a)$ and $-V_d(1 - a)$. At centre the output becomes $\eta V_d a$, equivalent in the small-signal calibration to displacement ag. An amplitude imbalance of one part in ten thousand gives a 10 nm offset at $g = 100\,\mu\mathrm{m}$, many orders above a short-band noise floor expressed in picometres. Reverse the mechanical calibration displacement to isolate odd signal, measure the voltages at the electrode terminals and compare the nulled bridge under exchange of the drive channels. Reversing both carrier phases simply reverses both desired signal and this coherently driven offset; by itself it cannot distinguish them. An offset stable over a short scan can be estimated from two symmetric calibration displacements, whereas a drifting imbalance requires a reference or simultaneous drive measurement.

Find the minimum for a compliant probe instead of guessing at it

It is possible to optimise the free probe analytically up to a scalar root. Write $y = V_d^2$, $K_e = hy$, and a for the readout displacement density at unit drive multiplied by one volt. Let $d_F = \sqrt{4k_B T b}$ be the Brownian force density. At displacement frequencies well below the resonance, the total variance density has an electrical term falling as $1/y$ and a mechanical term that grows as the spring softens:

$$S_x(y) = \frac{a^2}{y} + \frac{d_F^2}{(k - hy)^2},$$

$$\frac{dS_x}{dy} = -\frac{a^2}{y^2} + \frac{2h d_F^2}{(k - hy)^3}.$$

On the stable interval the first derivative strictly increases when $h > 0$. The first term starts at negative infinity and the second diverges at the stiffness boundary. Hence there is one unconstrained minimum. Solve a bracketed scalar equation for y, then compare with each experimental upper bound in volts squared. At zero parasitic capacitance the ideal floating plate has $h = 0$; total noise then falls monotonically toward an asymptotic Brownian floor. A numerical routine that divides by h would fail precisely at this informative physical limit. In either case, a force-noise specification referred to the

external sample can produce a different optimum because force conversion multiplies displacement noise by stiffness.

With the baseline values $a = 2.039 \times 10^{-12}\ \mathrm{m\,V/\sqrt{Hz}}$, $h = 2.00 \times 10^{-4}\ \mathrm{N/(m\,V^2)}$, $d_F \simeq 1.287 \times 10^{-13}\ \mathrm{N/\sqrt{Hz}}$. Substitution brackets the stationary point between 1.45 and 1.50 V. At roughly 1.47 V the density is about $6.72\,\mathrm{pm/\sqrt{Hz}}$, a shallow minimum compared with the heat-limited result. A tiny uncertainty in amplifier noise or in the support damping could shift the exact optimum appreciably while having very little effect on achievable resolution. Reporting a numerically optimised voltage to six decimals would imply a knowledge of the physical system absent from these inputs.

A direct bench experiment can test whether the excess really comes from compliance. Compare the demodulated spectrum with the centre plate tightly held and with it free. The electrical floor should be similar if the clamping keeps cable capacitance and carrier coupling fixed. The free probe should acquire a mechanical resonance which moves with V_d^2, and the low-frequency component should track the inferred static stiffness. A large peak locked to carrier-drive amplitude modulation in both experiments suggests coherent feedthrough. Also check that the noise estimate comes from a frequency band above drift yet below resonance, and that the inferred amplitude density times the square root of the bandwidth predicts repeated stationary readings. These practical checks distinguish a readout-limited displacement estimate from a calibrated measurement of actual thermal motion.

Explore the usable voltage range of the bridge

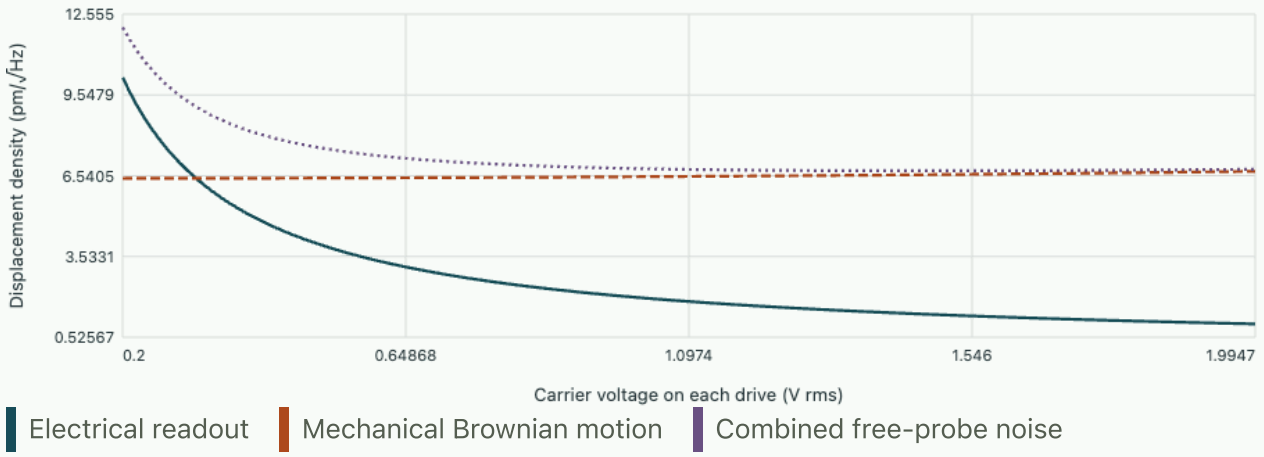
The horizontal axis ends at the first of the heat, driver and mechanical limits. With 4 pF fixed parasitic capacitance at 100 kHz the heat ceiling is 1.995 V rms, while a freely suspended probe reaches its minimum combined noise near 1.47 V rms. Reduce the added capacitance to zero and explain why a finite optimum arising from mechanical softening disappears. Then lower carrier frequency. The current-noise term grows and dielectric heating at a given voltage falls. Observe which effect changes the allowed voltage range and which changes the height of the electrical noise curve.

Open this experiment online

Try the calculation

Input and stray capacitance (pF) 4

Carrier frequency (kHz) 100



Maximum allowed drive 1.995 V rms. Minimum combined noise 6.72 pm/√Hz at 1.472 V rms.

Tutorial · separate electronic resolution from mechanical response

Find the frequency at which current and voltage noise contribute equally

Retain the baseline input capacitances and noise sources. Calculate the crossover frequency and the effect of doubling cable capacitance from 4 pF to 8 pF. What happens to input-referred current-noise displacement at fixed frequency?

▼ **Work through the solution**

Set $i_n/(2\pi f C_T) = e_n$. With $C_T = 8$ pF the crossover is 19.9 kHz. The longer cable gives $C_T = 12$ pF and a crossover at 13.3 kHz. At 100 kHz, the current noise at the input falls to $1.33\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ but gain falls from 5000 to 3333 V/m. The current contribution remains $0.398\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$. The voltage contribution increases from 2.00 to 3.00 pm/√Hz. Both results follow independently by substituting $K = 2C_0V_d/(gC_T)$ into the noise formula and cancelling C_T in only one of the two terms.

Diagnose a frequency shift and a noise claim

At 2.00 V the resonance changes from 500.0 Hz to 490.0 Hz. Determine the implied stiffness reduction. A second experiment quotes 10 pm rms in a 100 Hz noise bandwidth for the freely suspended probe at 300 K. Can thermal equilibrium and the measured spring and damping account for this number?

▼ **Work through the solution**

At unchanged moving mass $k_{\mathrm{eff}}/k = (490/500)^2 = 0.9604$. The measured electrical stiffness is 0.000792 N/m, consistent with the 0.00080 N/m prediction. At frequencies well below resonance the Brownian density is approximately 6.70 pm/√Hz, and a flat 100 Hz equivalent band would contain approximately 67 pm rms of thermal motion alone. Ten picometres rms would correspond to about 2.2 Hz equivalent noise bandwidth at this density, or require a changed mechanical condition, frequency band or calibration. To discriminate experimentally, measure the mechanical resonance spectrum with a known input-force calibration and integrate the noise with the actual filter transfer. Include mechanical susceptibility over the band if the filter extends close to resonance.

Further reading

- Maxwell stress and electrical coenergy in mechanical force balances
- MIT · Woodson and Melcher, lumped-parameter electromechanics, chapter 3

Electromechanics, optics and measurement · 5. A moving-coil galvanometer as an electrical and mechanical system

A scale division records an angle, yet the quantity marked on the dial is electric current. A spring, a magnetic field and the resistance of the external measuring circuit connect the two. Derive the calibration, predict how a pointer approaches its reading, and find which experiment measures the charge of a short pulse.

Convert the force on two straight sides into a torque

A light rectangular coil has N turns, active length l, width w, and area per turn $A = lw$. It pivots about an axis parallel to its long sides. Shaped pole pieces confine an approximately radial magnetic field of magnitude B to an air gap. Over the working angular range the magnetic field is perpendicular to the active wire and lies approximately in the plane of the coil. An elementary length carrying conventional current i experiences $d\mathbf{F} = i\,d\mathbf{l} \times \mathbf{B}$. The two long sides carry current in opposite directions and experience forces of equal magnitude $NilB$. Acting with lever arms $w/2$, the forces make a couple whose torque is $NiAB$. Their net force is approximately zero. The short ends supply negligible torque about the pivot under the ideal geometry.

Write $K = NAB$, and choose the positive sense of current to give positive torque. Its units are $\mathrm{N\,m/A}$. In a uniform Cartesian field an ordinary rigid loop instead gives torque $NiAB\sin\alpha$, where α is the angle between its normal and the field. The radial gap maintains the factor equivalent to $\sin\alpha$ close to unity as the pointer turns. This is a design choice that produces a nearly linear scale. If a measured calibration curve flattens at a large angle, field nonuniformity and a nonlinear torsion spring are candidate causes that can be tested separately.

For a soft torsion spring with stiffness κ, the return torque is $-\kappa\theta$ at small deflections from the zero-current setting. At static equilibrium $\theta = Ki/\kappa$. With $N = 200$, $A = 1.00\,\mathrm{cm^2}$, $B = 0.250\,\mathrm{T}$ and $\kappa = 8.0 \times 10^{-4}\,\mathrm{N\,m/rad}$, the torque constant is $0.0050\,\mathrm{N\,m/A}$ and the current sensitivity is $6.25\,\mathrm{rad/A}$. A $5.0\,\mathrm{mA}$ current turns the pointer through $0.03125\,\mathrm{rad}$, about 1.79°. A mirror mounted on the same axle deflects a reflected laser through twice this angle. At a screen distance of $0.50\,\mathrm{m}$ the spot shifts $0.50\tan(2\theta) = 31.3\,\mathrm{mm}$, making the small mechanical rotation easy to resolve.

Recover the reciprocal voltage from work

Now rotate the coil through a small positive angle while it carries positive current. The electromagnetic torque does mechanical work $Ki\,d\theta$. The voltage associated with the motion must therefore draw electrical work $ie_{\mathrm{back}}\,dt$ of equal magnitude when the motor torque assists that motion. This gives the back-emf relation $e_{\mathrm{back}} = K\dot\theta$, measured as a positive drop in the motor sign convention. Its coefficient has units $\mathrm{V\,s/rad}$, numerically the same SI coefficient as the torque constant. One can also derive it from the motional emf on the two active sides. Each moves at speed $(w/2)\dot\theta$, crosses field B, and generates a voltage $Bl(w/2)\dot\theta$ along its wire. The two contributions add over every turn.

Let J denote the rotational inertia, b the mechanical viscous damping coefficient, L the coil inductance and R the sum of winding plus external series resistance seen by a voltage source $v(t)$. The electromechanical equations are

$$J\ddot{\theta} + b\dot{\theta} + \kappa\theta = Ki,$$

$$L\dot{i} + Ri + K\dot{\theta} = v(t).$$

Multiplying the first equation by $\dot{\theta}$ and the second by i eliminates the coupling power $Ki\dot{\theta}$. The result says that source power vi equals the rate of increase of $J\dot{\theta}^2/2 + \kappa\theta^2/2 + Li^2/2$, plus $b\dot{\theta}^2 + Ri^2$. This identity fixes the relative signs of the torque and back emf. If a sign mistake turns a passive closed coil into a self-accelerating pointer, the same mistake produces a residual coupling term in this power balance. Mechanical friction and winding heat represent distinct sinks even though both can damp the observed displacement.

Predict overshoot and find a critical external resistance

If the electrical relaxation time L/R is short compared with the oscillation time $1/\sqrt{\kappa/J}$, set $i \simeq (v - K\dot{\theta})/R$. Substituting into the mechanical equation increases its effective damping to $b_{\mathrm{eff}} = b + K^2/R$. Set $v = 0$ when discussing free relaxation with the electrical terminals connected through a resistor. A moving shorted coil generates a current that opposes its motion. Opening the circuit instead suppresses that current and leaves the mechanical damping. An ideal current source holding specified current through all pointer motions also changes the back-emf response because its output voltage adjusts to enforce that current.

The undamped angular frequency is $\omega_0 = \sqrt{\kappa/J}$ and the damping ratio is

$$\zeta = b_{\mathrm{eff}}/(2\sqrt{\kappa J}).$$

Take $J = 2.0 \times 10^{-6}\,\mathrm{kg\,m^2}$, $b = 1.6 \times 10^{-5}\,\mathrm{N\,m\,s/rad}$, and the preceding K, κ. Then $\omega_0 = 20.0\,\mathrm{rad/s}$, $2\sqrt{\kappa J} = 8.0 \times 10^{-5}\,\mathrm{N\,m\,s/rad}$, and an open circuit gives $\zeta = 0.20$. With $R = 1.00\,\Omega$ the electric contribution to damping is $2.50 \times 10^{-5}\,\mathrm{N\,m\,s/rad}$, giving $\zeta = 0.5125$. For a voltage step the ideal fractional overshoot in the reduced model is

$$\exp[-\pi\zeta/\sqrt{1-\zeta^2}] = 0.153.$$

The first peak occurs after $\pi/(\omega_0\sqrt{1-\zeta^2}) = 0.183\,\mathrm{s}$.

Critical damping occurs when $b + K^2/R = 2\sqrt{\kappa J}$, which requires a total resistance of about $0.391\,\Omega$. A winding resistance of $0.300\,\Omega$ therefore leaves $0.091\,\Omega$ for external source and leads. If the winding alone exceeds the target resistance, a simple passive external series resistor cannot reach this damping in the reduced model. A finite L must also be checked: with $L = 5.0\,\mathrm{mH}$ and $R = 1.0\,\Omega$, $\omega_0 L/R = 0.10$, so the approximation provides a useful starting estimate. Reducing R indefinitely eventually invalidates the elimination of current. The full pair of equations then predicts the damping and frequency dependence.

Distinguish a charge impulse from a slow current reading

Suppose a brief imposed current pulse ends long before the pointer has turned appreciably. Integrating the torque equation over that interval gives

$$J\Delta\dot{\theta} \simeq K\int i\,dt = Kq.$$

The initial signal is now angular velocity proportional to total transmitted charge. For weak subsequent damping, the first excursion is approximately $Kq/(J\omega_0)$. A dose of $20\,\mu\mathrm{C}$ gives $2.50\,\mathrm{mrad}$ with the instrument above, corresponding to about $2.50\,\mathrm{mm}$ on the mirror scale. Two current pulses of identical charge but different shapes give nearly the same reading if both are short on the mechanical time scale. Their peak currents can be entirely different.

For appreciable damping, use the impulse solution $\theta(t) = Kq\,e^{-\zeta\omega_0 t}\sin(\omega_d t)/(J\omega_d)$, with $\omega_d = \omega_0\sqrt{1-\zeta^2}$ for an underdamped pointer. Its maximum obeys $\tan(\omega_d t_p) = \omega_d/(\zeta\omega_0)$ in the first quarter cycle. Substitution corrects the ballistic calibration instead of treating the damped peak as an undamped amplitude. The state of the external circuit after the pulse matters because it determines ζ during the throw. A calibration pulse should leave the measuring circuit connected in the same way as the unknown pulse.

Measure the torque constant without assuming the magnetic field

The preceding design calculation multiplies turn count, active area and a nominal air-gap induction. An independent calibration can measure the same constant without knowing the field distribution along every wire. Move the pointer through a small controlled angle while measuring its open-circuit voltage with a high-impedance instrument. In the same working interval the voltage is $K(\theta)\dot{\theta}$ up to a polarity choice. Integrating the voltage over the sweep gives $\int_{\theta_1}^{\theta_2} K(\theta)\,d\theta$. If K is constant this is simply $K\Delta\theta$, independent of sweep speed. It is often easier to integrate a recorded voltage pulse accurately than to differentiate the angle trace to obtain instantaneous speed.

For a sweep through $0.0800\,\mathrm{rad}$, suppose the measured voltage integral is $0.401\,\mathrm{mV\,s}$. The generator constant is $5.01\times10^{-3}\,\mathrm{V\,s/rad}$. Independently add a known small torque to the spring and measure its static angle. A torque $32.0\,\mu\mathrm{N\,m}$ rotates the pointer by $0.0400\,\mathrm{rad}$, giving $\kappa = 8.00\times10^{-4}\,\mathrm{N\,m/rad}$. Combining the two measurements predicts a current sensitivity of $6.27\,\mathrm{rad/A}$. Splitting the sweep into smaller equal angular intervals tests whether the measured K depends on angle. Sweeping backwards should reverse the integrated voltage. A voltage offset instead contributes its offset times duration, so repeating at two sweep speeds helps separate it.

A dynamic cross-check uses free decay under two load conditions. Measure the exponential decay rate α_{open} with a high-impedance voltage probe and α_R with a known total resistance in the closed circuit. In the short-electrical-time limit, subtraction of the two damping laws yields

$$K^2 = 2JR(\alpha_R - \alpha_{\mathrm{open}}).$$

With $J = 2.0\times10^{-6}\,\mathrm{kg\,m^2}$, $R = 1.00\,\Omega$, and decay rates 4.0 and $10.25\,\mathrm{s^{-1}}$, this gives

$$K = 0.0050\,\mathrm{N\,m/A}.$$

The method needs two identifiable decay rates and independent inertia. If the loaded response is overdamped, the envelope of two oscillation peaks is no longer available; fit the full two-real-pole response to the measured displacement.

Keep the third dynamical variable in a frequency-response experiment

For a sinusoidal voltage at angular frequency ω the exact linear transduction law follows by eliminating the complex current after transforming both differential equations. The mechanical impedance in angular-displacement form is $\kappa - J\omega^2 + i\omega b$; the winding impedance is $R + i\omega L$. Using phasors with time dependence $e^{i\omega t}$ gives

$$\frac{\Theta}{V} = \frac{K}{(R + i\omega L)(\kappa - J\omega^2 + i\omega b) + i\omega K^2}.$$

At zero frequency the gain is $K/(R\kappa)$ and the phase tends to zero, as the static calibration requires. At a sufficiently high frequency both the electrical and inertial impedances grow. For $L > 0$ the magnitude approaches $K/(LJ\omega^3)$. The reduced model with the current eliminated algebraically falls only as $1/\omega^2$. The change of the asymptotic slope provides an experimental way to distinguish finite inductance from an incorrectly adjusted viscous damping constant. Beyond the working band of a real instrument, elastic modes of the pointer or parasitic capacitance can enter before that asymptote is resolved.

At the nominal undamped frequency $20\,\mathrm{rad/s}$, with $R = 1.0\,\Omega$ and $L = 5.0\,\mathrm{mH}$, the stiffness and inertia terms cancel. The denominator becomes $-\omega^2 L b + i\omega(Rb + K^2)$, or approximately $-0.000032 + i\,0.000820$ in consistent SI units. Its magnitude gives about $6.09\,\mathrm{rad/V}$. A sinusoid of $1.0\,\mathrm{mV}$ peak produces about $6.09\,\mathrm{mrad}$ peak, whereas the static response to $1.0\,\mathrm{mV}$ is $6.25\,\mathrm{mrad}$. The phases differ radically despite their similar amplitudes: near this frequency the pointer lags by about 92.2°. A gain-only calibration at a single frequency would miss the phase lag.

Change the electrical load and compare two step models

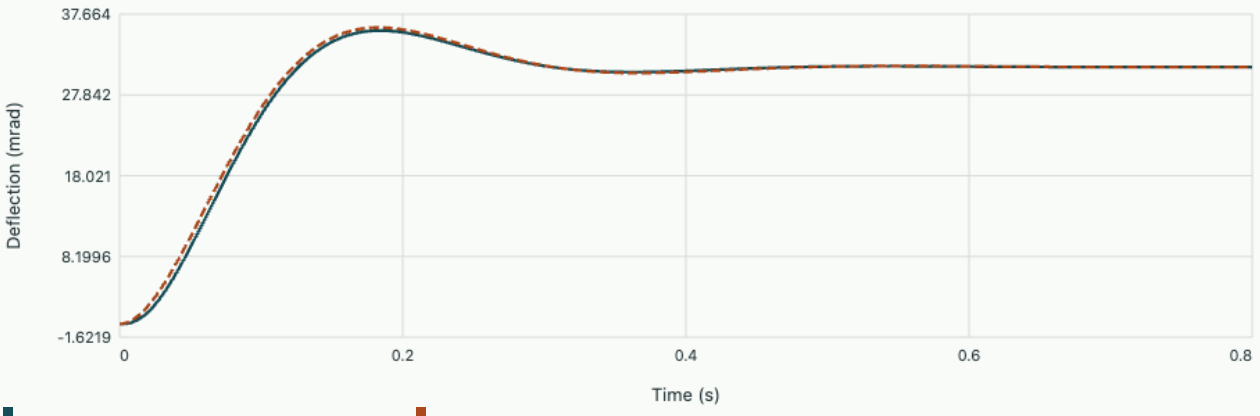
In the live graph the source voltage is chosen anew at each resistance to give a final current of $5.0\,\mathrm{mA}$ after the pointer settles. That convention keeps the final angle common to all curves and isolates the transient response of the load. The full model retains $i, \theta, \dot\theta$; the second curve sets $L\dot i = 0$. Predict first which trace should start flatter. With finite inductance the current initially vanishes and increases at a finite rate. The initial torque and angular acceleration therefore vanish, while the reduced model applies a finite torque immediately. Increasing the inductance makes that distinction easier to see, especially when the electrical response time becomes comparable to the period of mechanical oscillation.

Open the live experiment

Try the calculation

Total resistance (Ω) 1

Winding inductance (mH) 5



The reduced damping ratio is 0.512. The largest deflection in the first 0.80 s is 35.67 mrad. The common static limit is 31.25 mrad.

How could a student change the static range of the instrument without changing its suspension or its torsional stiffness?

▼ **Work through the solution**

For a dc ammeter, connect a calibrated shunt resistance in parallel with the coil, allowing only a known fraction of total current into it. With a coil winding resistance $0.300\,\Omega$ and a $0.0333\,\Omega$ shunt, the coil carries about ten percent of a steady terminal current and the indicated full-scale terminal current increases about tenfold. The total circuit seen by a motional voltage during a pointer oscillation depends on the shunt and source impedance, so the static range multiplier does not by itself give the new damping. For a voltmeter, a series resistance instead sets the current produced by an applied terminal voltage. In both cases record the resistance network whenever reporting a dynamic response.

Estimate the back action on the signal being measured

A galvanometer becomes part of the circuit it observes. Suppose a source with open-circuit voltage $10.0\,\mathrm{mV}$ and internal resistance $2.00\,\Omega$ drives a coil of resistance $0.300\,\Omega$. After settling, the current is $4.35\,\mathrm{mA}$ and the pointer turns by $27.2\,\mathrm{mrad}$. The undisturbed short-circuit current of the source would have been $5.00\,\mathrm{mA}$. Interpreting the indicated value as the current that would flow with a perfect zero-resistance meter underestimates that quantity by about thirteen percent. Measuring the open-circuit voltage instead requires a sufficiently large input resistance, another electrical boundary condition.

During pointer motion the disturbance includes the back emf as well as winding resistance. If the same coil is briefly turning at $0.40\,\mathrm{rad/s}$ it develops a $2.0\,\mathrm{mV}$ back emf. Neglecting inductance for that estimate,

the resulting current is

$$(10.0 - 2.0)\,\mathrm{mV}/2.30\,\Omega = 3.48\,\mathrm{mA}.$$

The deflection therefore contains information about both the source and the pointer's earlier motion. A static calibration alone cannot recover a rapid source waveform. A known source resistance and the two coupled dynamic equations provide a way to estimate the waveform from synchronised angle and current measurements.

Tutorial on calibration and transient identification

A current source produces a static mirror shift of $12.5\,\mathrm{mm}$ on a screen $0.50\,\mathrm{m}$ away. Estimate the current and identify a reversible test for an angular zero offset.

▼ **Work through the solution**

Use $\theta = \tfrac{1}{2}\arctan(0.0125/0.50) = 0.01250\,\mathrm{rad}$ and $i = \kappa\theta/K = 2.00\,\mathrm{mA}$. Repeat with reversed current of equal magnitude. Half the difference of the two angles cancels a fixed angular zero offset and isolates the response odd in current. A changing temperature during reversal can alter the spring or winding resistance, so alternate the order and record the zero afterwards if the required uncertainty is small.

For an open-circuit free decay the ratio of two successive same-side positive peak angles is 3.60. Their time separation is $0.321\,\mathrm{s}$. Recover the mechanical parameters that can be inferred if J is known.

▼ **Work through the solution**

The logarithmic decrement is $\ln 3.60 = 1.281$. The decay rate is $\alpha = 1.281/0.321 = 3.99\,\mathrm{s}^{-1}$, and the damped angular frequency is

$$2\pi/0.321 = 19.57\,\mathrm{rad/s}.$$

Then $\omega_0 = \sqrt{\omega_d^2 + \alpha^2} = 19.97\,\mathrm{rad/s}$,

$$b = 2J\alpha = 1.60 \times 10^{-5}\,\mathrm{N\,m\,s/rad},$$

$$\kappa = J\omega_0^2 = 7.98 \times 10^{-4}\,\mathrm{N\,m/rad}.$$

This identification assumes viscous damping and a sufficiently high impedance to make the electrical damping negligible.

Explore how motion generates voltage in the induction chapter.

Electromechanics, optics and measurement · 6. Eddy-current braking, magnetic penetration and the heat left behind

A conducting loop moves above a row of magnetic poles. The drag first increases with speed, but a rapid crossing need not generate proportionally greater braking. Trace the work done against the Lorentz force into circuit energy, heat and a measurable temperature change. Then decide which features of a solid plate the single-loop model can capture.

Generate a voltage from a spatially periodic field

Let a rectangular loop of height h along y and width w along x translate to the right above a stationary magnetic pattern. At the loop let the normal induction be $B_z(x) = B_0 \cos(kx)$, with $k = 2\pi/\lambda$. Its two active sides are centred symmetrically about the coordinate x_c. Assume the field varies negligibly along y over this height. Integrating through the loop gives

$$\Phi = h \int_{x_c - w/2}^{x_c + w/2} B_0 \cos(kx)\, dx = [2hB_0 \sin(kw/2)/k] \cos(kx_c).$$

A uniform translation $x_c = ut$ therefore produces an external emf with peak amplitude $E_{\mathrm{pk}} = Ku$, where $K = 2hB_0 |\sin(kw/2)|$.

This spatial averaging explains an initially surprising design feature. A loop spanning precisely one full spatial period always links zero net flux from the sinusoidal field and therefore generates no emf in this ideal geometry. A loop of half a period has maximum coupling at a specified height. If w is much smaller than the period,

$$\sin(kw/2) \simeq kw/2,$$

$$K \simeq hwB_0 k.$$

This is exactly the small-loop result from using its area times the local gradient of induction. Reversing velocity reverses the emf at a fixed position. At a fixed positive speed the temporal frequency is $f = u/\lambda$, so the same conductor experiences different electrical frequencies at different mechanical speeds.

Include finite inductance and derive the drag curve

Model the current path by a constant resistance R and self-inductance L. Its steady periodic current solves $L\dot{i} + Ri = \mathcal{E}(t)$. In phasor notation its magnitude is $E_{\mathrm{pk}}/\sqrt{R^2 + (\omega L)^2}$, lagging the emf by $\arctan(\omega L/R)$, where $\omega = ku$. The current develops field energy $Li^2/2$ and dissipates Ri^2. Multiplying the circuit equation by i and averaging over a period eliminates the derivative of stored energy. The average mechanical power delivered against the drag is therefore the mean Joule heating $P = E_{\mathrm{pk}}^2 R/[2(R^2 + \omega^2 L^2)]$. Dividing by the speed gives the drag magnitude

$$F(u) = \frac{K^2 u}{2R[1 + (u/u_c)^2]},$$

$$u_c = \frac{R}{kL}.$$

At low speed the drag is proportional to velocity. At the crossover $u = u_c$ it reaches $K^2 u_c/(4R)$, and at much larger speeds it falls approximately as $1/u$ within the assumed lumped model. The incoming mechanical power then tends to a constant rather than growing without bound. Why is there a maximum? Increasing speed raises the emf, but it also raises the reversal frequency. An inductive current increasingly lags the driving emf, and its component in phase with the emf does the net dissipative work. A measured turnover therefore tests both an electrical time scale and the spatial scale of the magnet pattern.

Use $h = 50\,\mathrm{mm}$, $w = 25\,\mathrm{mm}$, $\lambda = 0.100\,\mathrm{m}$, $B_0 = 0.20\,\mathrm{T}$, $R = 0.050\,\Omega$ and $L = 0.150\,\mathrm{mH}$. Then $K = 0.01414\,\mathrm{T\,m}$, $u_c = 5.31\,\mathrm{m/s}$, and maximum mean drag is $5.31\,\mathrm{mN}$. At $u = 2.0\,\mathrm{m/s}$ the field alternates at $20\,\mathrm{Hz}$, giving an emf of $28.3\,\mathrm{mV}$ peak. The current is $0.529\,\mathrm{A}$ peak and its phase lag 20.7°. The resistor receives approximately $7.00\,\mathrm{mW}$, and the motor maintaining translation supplies a mean force of $3.50\,\mathrm{mN}$. Using peak current in I^2R without its factor one half would overstate the mechanical force by a factor two.

Determine whether the current penetrates a real plate

A continuous metal plate contains many possible current paths, unlike a preformed loop. Their resistances, inductances and mutual couplings depend on the plate boundary, magnet geometry and frequency. For a good homogeneous conductor the magnetic diffusion equation gives a skin depth $\delta = \sqrt{2/(\mu\sigma\omega)}$ under sinusoidal excitation. At a depth δ a field entering a locally planar half-space has lost a factor e in amplitude and acquired a phase delay of one radian. The frequency to use is the frequency experienced by the moving material; the fact that the magnets are stationary in the laboratory does not make this frequency zero.

For copper with $\sigma = 5.8 \times 10^7\,\mathrm{S/m}$ and permeability close to μ_0, the skin depth is about $14.8\,\mathrm{mm}$ at $20\,\mathrm{Hz}$ and $1.48\,\mathrm{mm}$ at $2.0\,\mathrm{kHz}$. A $2.0\,\mathrm{mm}$ plate therefore sees little through-thickness attenuation in the first experiment but may need a diffusion calculation in the second. The plate's lateral current pattern can still matter when the through-thickness current density is nearly uniform. Slots interrupt wide current loops and usually reduce drag even when the mass and the bulk conductivity change little. Comparing a slotted and an unslotted plate directly tests the importance of current closure.

If a plate experiences sharply separated poles, its applied field in the moving frame contains a fundamental and higher harmonics. Harmonic n has a skin depth smaller by $\sqrt{n}$, so checking only the fundamental can miss surface concentration associated with narrow features. A Hall scan along the track supplies the actual spatial field spectrum. Multiplying each spatial wavenumber by the speed converts it into a temporal angular frequency for a constant-speed pass. Once the speed changes significantly over one pole spacing, a time-domain calculation using the measured position trace is preferable to this fixed-frequency conversion.

Close a mechanical and thermal energy budget

Take a rotating brake as a second experiment. A flywheel of known inertia $J_{\mathrm{rot}} = 0.020\,\mathrm{kg\,m^2}$ slows from 120 to $80\,\mathrm{rad/s}$ while no drive supplies power. The released rotational kinetic energy is

$$J_{\mathrm{rot}}(120^2 - 80^2)/2 = 80.0\,\mathrm{J}.$$

A separate run with the magnet removed measures bearing and air losses. Suppose the corresponding speed-dependent losses integrated over the magnetic braking trajectory total $8.0\,\mathrm{J}$. Then the electromagnetic channel receives $72.0\,\mathrm{J}$. Using the reference run's total energy over an arbitrary equal time could give a different answer, since the two speed histories affect the parasitic losses.

A calorimeter enclosing the conducting rotor has effective heat capacity $C = 60\,\mathrm{J/K}$ and a thermal link of conductance $G = 0.12\,\mathrm{W/K}$ to an environment at constant temperature T_a. Its balance is

$$C\,dT/dt = P_{\mathrm{deposited}} - G(T - T_a).$$

During a $20\,\mathrm{s}$ run, suppose the time-averaged rise above ambient is $0.50\,\mathrm{K}$. Then $1.20\,\mathrm{J}$ flowed through the thermal link. If all electromagnetic heat entered the calorimeter, its final rise from an ambient start is

$$(72.0 - 1.20)/60 = 1.18\,\mathrm{K}.$$

A measured rise near $0.80\,\mathrm{K}$ would leave an energy discrepancy of approximately $23\,\mathrm{J}$. Check conduction into a shaft, heat stored outside the instrumented volume, and the estimate of parasitic torque before adjusting the electromagnetic theory.

Find the spatial features that dominate a measured braking signal

A single Fourier component of field already produces an informative turnover. In a tracked experiment the field often has a measurable third harmonic. Let $B(x) = B_1 \cos kx + B_3 \cos(3kx + \varphi)$, with constant translation speed. Flux integration across the finite loop multiplies each harmonic by its own width factor. The corresponding emf peak is $2h|B_n|u|\sin(nkw/2)|$, and its temporal angular frequency is nku. The linear RL equation treats the two drives independently. Orthogonality over a complete fundamental period makes the mean of their product vanish, so total mean heat can be found by adding the individual resistive powers. Within one period the interference terms do affect instantaneous torque and power.

Suppose $B_3/B_1 = 0.20$ and keep the preceding $w = \lambda/4$. Here $\sin(kw/2)$ and $\sin(3kw/2)$ both have magnitude $1/\sqrt{2}$. At low speed the third harmonic therefore adds about four percent to mean drag. At $u = u_c$, however, its angular frequency sees inductive impedance $3R$ while the fundamental sees R. The third-harmonic share relative to the fundamental becomes $0.20^2(1+1)/(1+9) = 0.008$, only 0.8%. A peak-to-peak field reading can change visibly without a comparable change in mean drag. Conversely, if the loop width happens to null one spatial harmonic, another smaller harmonic can supply most of the signal. A field map and the loop geometry together determine what source drives the circuit.

In a solid plate, predicting force from field and speed needs a field solution that enforces current closure. In the slowly moving conductor, $\mathbf{J} = \sigma[-\nabla\phi + \mathbf{u}\times\mathbf{B}]$ when the imposed magnetic pattern is effectively time independent in the laboratory and the induced magnetic field can be neglected. Charge conservation then gives

$$\nabla\cdot[\sigma(-\nabla\phi + \mathbf{u}\times\mathbf{B})] = 0.$$

On an electrically insulated plate edge, enforce $\mathbf{n}\cdot\mathbf{J} = 0$. Solving this potential problem accounts for charge accumulation that redirects currents at a boundary. Simply setting $\mathbf{J} = \sigma\mathbf{u}\times\mathbf{B}$ can predict

current flowing straight out of an insulating edge. The force follows from integrating $\mathbf{J} \times \mathbf{B}$ through the conducting volume, and Joule heat follows from $\int J^2/\sigma \, dV$.

A useful check on such a field calculation is a finite plate translating wholly inside an exactly uniform static magnetic field. Charges can rearrange until a uniform electrostatic field cancels $\mathbf{u} \times \mathbf{B}$. An isolated plate in this steady limit then has zero transport current and zero drag. Motion alone does not guarantee braking. At an entry edge, the moving conductor spans a spatial change of induction, and one scalar potential generally cannot cancel the motional drive around every closed path. Slits, thickness variations, a conducting contact at a boundary, and the induced field at higher speeds each change this test problem in an identifiable way.

Interpret a thermal transient without assuming it is instantaneous

If the brake runs at a stable speed and equilibrium electrical conditions, a local single-temperature model sees a nearly constant power. Its step response from ambient is $\Delta T = (P/G)[1 - \exp(-Gt/C)]$. Removing the drive gives an exponential decay with the same time constant C/G. Measuring that cooling tail fixes the ratio of heat capacity to conductance but does not set either one separately. A known heater power or an independent heat-capacity measurement supplies the missing absolute scale. A slope evaluated over a long interval without correcting leakage will underestimate the input power as the sample warms.

Take a small bench calorimeter enclosing the single test loop with $C = 1.20\,\mathrm{J/K}$ and $G = 0.020\,\mathrm{W/K}$. The $7.00\,\mathrm{mW}$ low-speed run gives a $0.350\,\mathrm{K}$ eventual rise and a $60\,\mathrm{s}$ time constant. After $30\,\mathrm{s}$ the rise is about $0.138\,\mathrm{K}$, whereas a thermally isolated calculation would give $0.175\,\mathrm{K}$. Its stored heat $0.166\,\mathrm{J}$ is less than the $0.210\,\mathrm{J}$ mechanical work over that interval; the thermal link accounts for the remainder. Loading the conductor thermally with an additional copper holder increases the response time and changes the initial slope without changing the final rise if conductance to the bath remains the same.

Compare predictions at a chosen travel speed

The live graph displays the force predicted by the finite-inductance loop alongside the resistive low-frequency extrapolation. Choose a speed, calculate the apparent excitation frequency, and compare the returned skin depth to the actual conductor thickness before interpreting the prediction as a plate measurement. Try choosing an inductance twice as large. The location of maximum force halves in speed, but the low-speed slope at fixed geometry and resistance is unchanged. A real plate whose high-speed drag does not follow the single-loop curve can still agree with the model on its low-speed slope. Plotting the measured residuals against speed can reveal the point where extra current paths or through-thickness diffusion become important.

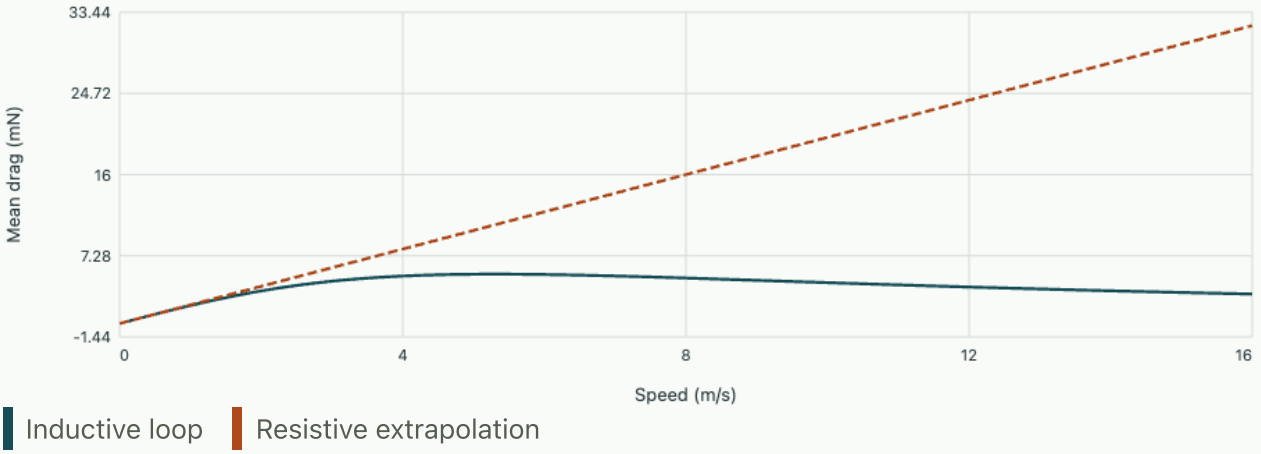
Open the live experiment

Try the calculation

Travel speed (m/s) **2**

Loop inductance (mH) **0.15**

Peak field (T) **0.2**



At 2.00 m/s the drag is 3.50 mN and the mean heat is 7.00 mW. The copper skin depth at 20.0 Hz is 14.78 mm. Peak drag occurs at 5.31 m/s.

A plate and holder cool from $0.240\,\mathrm{K}$ above ambient to $0.120\,\mathrm{K}$ in $42\,\mathrm{s}$ after a run. Can these two readings establish the braking power?

▼ **Work through the solution**

The half-life gives

$$C/G = 42/\ln 2 = 60.6\,\mathrm{s},$$

which agrees approximately with the small calorimeter example. It determines only a ratio. If an independent steady $10.0\,\mathrm{mW}$ heater then produces a $0.500\,\mathrm{K}$ temperature rise, $G = 0.0200\,\mathrm{W/K}$ and $C = 1.21\,\mathrm{J/K}$. A magnetic run that settles at $0.350\,\mathrm{K}$ now corresponds to $7.00\,\mathrm{mW}$, provided the thermal boundary and deposition pattern are sufficiently similar to the calibration.

Check the sign of instantaneous electromagnetic work

At a phase where the applied emf has just crossed zero, a finite-inductance loop may still carry considerable current. During part of a cycle the emf and current can have opposite signs. The product $\mathcal{E}i$ is then negative even though Ri^2 remains positive. The difference is a negative derivative of the magnetic store $Li^2/2$. The loop briefly returns energy mechanically while its stored field decreases, and some of that release also becomes heat. A force transducer resolving separate pole passages can therefore see instantaneous variations that a mean braking-force estimate suppresses.

For the numerical $2.0\,\mathrm{m/s}$ case the peak field energy is $LI_{\mathrm{pk}}^{2}/2 \approx 21.0\,\mu\mathrm{J}$. The dissipated energy per full pole period is

$$7.00\,\mathrm{mW}/20\,\mathrm{Hz} = 350\,\mu\mathrm{J}.$$

These quantities serve different purposes. The first is a reversible instantaneous store; the second is net irreversible energy delivered each time the conductor moves one spatial period. Dividing the second by the spatial period of $0.100\,\mathrm{m}$ recovers the mean drag of $3.50\,\mathrm{mN}$ without using the translational speed explicitly.

Tutorial on braking measurements

At $u = 2.0\,\mathrm{m/s}$ the loop above experiences a mean drag of $3.50\,\mathrm{mN}$. Predict what halving the induction and doubling the speed does, keeping its geometry and electrical properties fixed.

▼ **Work through the solution**

Halving induction quarters K^{2}. The speed factor contributes $2[1+(2/5.31)^{2}]/[1+(4/5.31)^{2}] \approx 1.46$. The combined change gives a drag about 0.364 times the original, or $1.27\,\mathrm{mN}$. The mechanical power decreases less because speed has doubled. A linear-speed extrapolation would miss the inductive phase lag, already significant at $4\,\mathrm{m/s}$.

In a repeated flywheel test the brake becomes warmer with every run. Which electrical parameter change can explain a change of drag even with the magnets held fixed?

▼ **Work through the solution**

The resistance of an ordinary metal generally increases with temperature. At a fixed speed F is proportional to $R/[R^{2}+(\omega L)^{2}]$. Raising resistance reduces drag on the low-frequency resistive side and can increase it on the high-frequency inductive side. The maximum occurs at $R = \omega L$ if electrical geometry and inductance stay fixed. Measuring the temperature and resistance before each run helps distinguish this effect from a weakening field or a changed gap.

Compare a step-driven diffusion profile with sinusoidal skin penetration.

Electromechanics, optics and measurement · 7. Measuring magnetic core loss by a hysteresis loop and by calorimetry

A closed loop on a graph of magnetic induction against field encloses an area with units of energy per unit volume. A thermometer measures kelvins. Bring these two observations onto the same scale, account for the winding heat, and determine when the two experiments actually test the same material loss.

Infer the field and the induction from independent windings

Wind a toroidal specimen with a primary winding of N_p turns and a lightly loaded pickup of N_s turns. Let A be the representative iron cross-sectional area, ℓ the mean magnetic path length and $V_c = A\ell$ the magnetic volume. Neglect a gap and leakage initially. Ampère's law around a path through the core gives

$$H \simeq N_p i_p / \ell.$$

This relation is accurate when the toroid is slender enough that the spread of path radii is small and the pickup current is negligible. In a thick toroid $H(r) = N_p i_p/(2\pi r)$, so one point on a nominal loop can actually represent a range of local field values.

Faraday's law supplies the second observable. Choose winding polarity so the reported voltage is

$$v_s = N_s A\, dB/dt.$$

Integrating the voltage gives the change in cross-section-averaged magnetic induction. The pickup cannot determine a constant induction offset, because static flux induces no voltage. A closed loop area is insensitive to adding a constant to B or H, but a dc offset in the voltage integrates into a time-dependent ramp and prevents the reconstructed trajectory from closing. In steady periodic operation one can subtract the voltage mean over a complete period before integration. Applying that correction to a genuinely evolving magnetisation process would remove part of a real flux change.

Multiply the two measured field quantities in the appropriate order. The field work supplied per unit material volume on a small induction change is $H\,dB$. The source term associated with the magnetising flux is $i_p v_{p,\mathrm{induced}}\,dt$. Using $v_{p,\mathrm{induced}} \simeq (N_p/N_s)v_s$, its value is

$$i_p N_p A\, dB = V_c H\, dB.$$

Integrating around a settled cycle gives the magnetic loss

$$W_c = V_c \oint H\, dB,$$

$$P_c = f V_c \oint H\, dB.$$

The field can absorb positive work and return some of it later in the same cycle. The area measures the algebraic remainder. Reversing the polarity of a single signal flips the loop orientation. A passively heated specimen must receive positive net energy when signs are defined for power entering the core. Simply taking the absolute value of all instantaneous products would count reversible storage twice and can grossly overestimate the heat.

Work an elliptical loop from time-domain signals

Suppose a nearly sinusoidal field is $H = H_p \cos \omega t$ and the induction lags it by a positive angle δ, $B = B_p \cos(\omega t - \delta)$. Differentiate the latter and multiply by H. Over one cycle $\cos \omega t \sin \omega t$ averages to zero, and $\cos^2 \omega t$ averages to one half. The resulting mean power density is $\omega H_p B_p \sin \delta / 2$, giving $\oint H \, dB = \pi H_p B_p \sin \delta$. With zero lag the two signals trace a line and the area vanishes, even if the peak values are large. A saturation-distorted loop need not be elliptical; the original time integral remains valid with synchronised data.

Take $H_p = 120\,\mathrm{A/m}$, $B_p = 0.200\,\mathrm{T}$ and $\delta = 12^\circ$. The energy per cycle and volume is $15.68\,\mathrm{J/m^3}$. A toroid with $\ell = 0.20\,\mathrm{m}$, $A = 2.0 \times 10^{-4}\,\mathrm{m^2}$ and $V_c = 40\,\mathrm{cm^3}$ dissipates $0.627\,\mathrm{mJ}$ per cycle. At $f = 200\,\mathrm{Hz}$ its power is $0.125\,\mathrm{W}$. With $N_p = 100$, the necessary current peak is $0.240\,\mathrm{A}$. With $N_s = 50$, the pickup voltage has peak $N_s A (2\pi f) B_p = 2.51\,\mathrm{V}$. These last two predictions allow the raw oscilloscope traces to be checked before a calculated loss in watts is trusted.

For sampled periodic data use $w = \sum_n (H_n + H_{n+1})(B_{n+1} - B_n)/2$, including the segment from the last point back to the first. This trapezoidal line integral computes an oriented polygon area. Plot the starting and ending points before adding the closing segment. Connecting widely separated endpoints by hand hides drift and can generate a fictitious wedge whose area depends on acquisition timing. Repeat at higher sampling rates to see whether corners or switching features carry appreciable extra area. When a periodic excitation has harmonics the time increment must resolve the highest significant harmonic in both measured signals.

Compare a thermal measurement with the magnetic area

Consider first a calorimeter that receives heat only from the core. Let the heat capacity of specimen, holder and sensor be C_{eff}. Let the leakage conductance to an ambient bath be G, and let $\Delta T = T - T_a$. The thermal balance $C_{\mathrm{eff}} \dot{\Delta T} = P_c - G \Delta T$ predicts a steady offset P_c / G and an initial heating slope P_c / C_{eff}. A small thermometer lag or poor thermal contact can make the earliest measured slope misleading even though environmental losses are then small. A separate electrical heater placed where the core deposits its heat can calibrate the effective thermal response.

For $C_{\mathrm{eff}} = 20\,\mathrm{J/K}$, $G = 0.015\,\mathrm{W/K}$, and a $2.0\,\mathrm{K}$ rise above ambient, the magnetic loss $0.125\,\mathrm{W}$ would give a slope of about $0.00477\,\mathrm{K/s}$. The heat already leaking to the bath is $0.030\,\mathrm{W}$, too large to discard in a ten-percent comparison. If the winding is inside the same calorimeter, its Ohmic heat must enter the balance as well. With primary resistance $1.50\,\Omega$ and the computed sinusoidal current, winding heating is

$$(0.240/\sqrt{2})^2 (1.50) = 0.0432\,\mathrm{W}.$$

The expected slope of the combined package at the same temperature is now $0.00693\,\mathrm{K/s}$. Subtraction of winding heat recovers the core result only when the estimated resistance represents the winding at its operating temperature and frequency.

A thermal estimate drawn from a whole interval is often more robust than a derivative from two adjacent temperature samples. Integrating gives

$$Q_{\mathrm{deposited}} = C_{\mathrm{eff}}[\Delta T(t_2) - \Delta T(t_1)] + G\int_{t_1}^{t_2} \Delta T\, dt.$$

Divide by the duration for mean power, then remove other measured heat sources. A drifting ambient must be recorded explicitly because the calorimeter stores energy relative to its own absolute temperature. If the core and coil equilibrate slowly with each other, a single fitted heat capacity can depend on the fitting time window. A two-temperature model or a longer dwell then supplies an independent diagnostic.

Separate reversible exchange from hysteretic and induced currents

A single measured loop at finite frequency gives the total energy dissipated by the core at that prescribed waveform and temperature. The loop can include irreversible domain processes, eddy currents and other rate-dependent effects. It does not by itself split these contributions. To probe their different scaling hold the peak induction and waveform shape fixed, then change frequency. Under a restricted low-frequency model the quasi-static hysteresis energy per cycle is approximately constant, while classical thin-lamination eddy-current power scales as $f^2 B_p^2$. Its energy per cycle therefore scales roughly as f. Plotting the measured loss per cycle and volume against f tests whether such an intercept-plus-slope description fits the chosen range.

Suppose the area is $12\,\mathrm{J/m^3}$ at $100\,\mathrm{Hz}$ and $20\,\mathrm{J/m^3}$ at $300\,\mathrm{Hz}$, at the same peak induction. A straight line gives intercept $8\,\mathrm{J/m^3}$ and slope $0.040\,\mathrm{J\,s/m^3}$. It predicts $16\,\mathrm{J/m^3}$ at $200\,\mathrm{Hz}$, near the elliptical-loop calculation. Two points always determine a line. A third and fourth frequency, together with uncertainties, are required to test the law. Different losses can share similar frequency scaling over a short interval, and temperature-dependent conductivity can bend a nominally simple curve.

Use the pickup to reject the reversible part of terminal power

Measuring primary terminal voltage and current with a power analyser provides a separate electrical route to loss, but the wanted number can be a small difference between terms. The applied voltage contains winding resistance, induced magnetising voltage, leakage-inductance voltage and possible parasitic contributions. In steady periodic operation the energy stored in a fixed linear leakage inductance returns after each cycle, while winding resistance dissipates energy permanently. The secondary pickup preferentially measures the core-linked flux. Forming $p_c(t) = (N_p/N_s) i_p(t) v_s(t)$ and averaging over a full cycle estimates core power without directly including the primary's copper voltage drop. It relies on both windings linking the same relevant flux and on negligible power drawn by the pickup.

For the elliptical example the primary magnetising voltage has peak $5.03\,\mathrm{V}$ and leads the induction by ninety degrees. Its phase relative to the current is

$$90^\circ - \delta = 78^\circ.$$

The core branch's apparent power is $V_{\mathrm{rms}} I_{\mathrm{rms}} \approx 0.603\,\mathrm{VA}$; its mean dissipative part is about $0.125\,\mathrm{W}$, and the reactive part is about $0.590\,\mathrm{var}$. Only the mean part contributes net heating of the magnetic volume. Including the winding voltage adds about $0.043\,\mathrm{W}$ of real power to the terminal reading. A simple

product of peak current and peak secondary voltage, without the turns ratio, factor one half and phase factor, has no reason to agree with either measured loss.

Consider a nearly ideal core with the same peak signals but a lag of only one degree. Its loss is about $10.5\,\mathrm{mW}$, while the much larger reversible power exchange is almost unchanged. A time offset of $10\,\mu\mathrm{s}$ at $200\,\mathrm{Hz}$ adds 0.72°, causing a relative loss error of approximately 72% for that small lag. Calibrate the relative channel delays using a reference signal through the actual current transducer and voltage chain, including their filters. Adjusting traces on screen until a core loop appears narrow assumes the very loss one is trying to measure. A synchronous measurement over many periods suppresses random noise but does not remove a persistent delay bias.

Test reconstruction on a waveform with a third harmonic

Let $H(t) = H_1 \cos \omega t + H_3 \cos 3\omega t$ while $B(t) = B_1 \cos(\omega t - \delta_1) + B_3 \cos(3\omega t - \delta_3)$. With these definitions δ_n is the phase lag of the nth harmonic in its own sinusoid. Multiplying H by $\dot{B}$ and integrating over an integer number of fundamental periods eliminates products of different harmonics. Each surviving harmonic contributes the number of its own cycles in one fundamental period. The area per fundamental period is therefore $\pi[H_1 B_1 \sin \delta_1 + 3 H_3 B_3 \sin \delta_3]$. This identity tests whether a numerical area calculation tracks the full waveform. It does not assume that a real material is linear; here the two signals are supplied as measured periodic functions.

Use the earlier fundamental values and add $H_3 = 20\,\mathrm{A/m}$, $B_3 = 0.030\,\mathrm{T}$ and $\delta_3 = 30^\circ$. The third harmonic contributes $2.83\,\mathrm{J/m^3}$ in each fundamental period. This represents an additional $0.0226\,\mathrm{W}$ for the chosen specimen and frequency. A fundamental-only sine fit would miss nearly fifteen percent of the resulting total area. An apparent jagged loop can also arise from asynchronous sampling or an offset. Distinguish genuine harmonics by checking that they are phase locked to the drive and persist under a change in sampling rate and analogue bandwidth.

Sweep phase and read the oriented loop

The live experiment reconstructs the H, B trajectory and compares it with a zero-lag reference. It uses the same peak field and induction for the two curves. At fixed peaks changing frequency leaves an assigned phase lag and therefore the prescribed energy per cycle unchanged, while watts rise with the number of cycles per second. A real core can change its lag and waveform with frequency; the graph separates this independent variable so that the meaning of the two scaled quantities is clear. Add a timing offset and predict the direction of the area error before checking its reported size. Near zero lag even the direction of an apparent loop can be controlled by a small electronic offset.

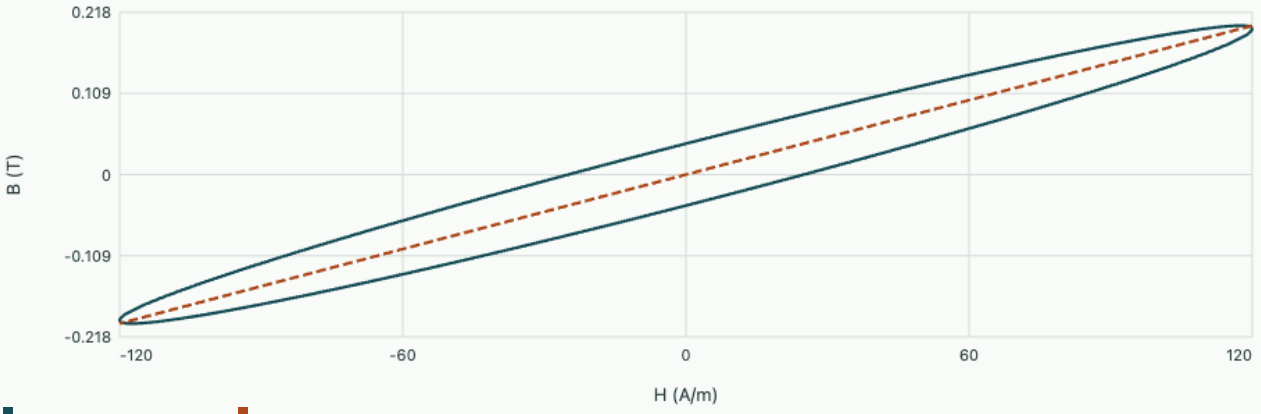
Open the live experiment

Try the calculation

Physical induction lag (°) **12**

Frequency (Hz) **200**

Induction channel delay (µs) **0**



True loop area 15.68 J/m³ per cycle. The recorder indicates 15.68 J/m³. Physical heating of the 40 cm³ specimen is 0.125 W.

Two toroids are made of the same nominal material with equal cross sections but different mean path lengths. How should a result reported only as watts be compared?

▼ **Work through the solution**

Determine actual magnetic volume for each toroid and calculate the energy per unit volume per cycle, $P/(fV_c)$. Also check that both experienced the same induction waveform, frequency and temperature. Equal primary currents need not give equal field when the path lengths differ; equal primary applied voltages need not give equal core induction if turns, winding voltage drops or area differ. Reproducing the actual $B(t)$ with feedback on the pickup is a more direct way to match excitation. Report dimensions and peak induction with the result so a reader can reconstruct what material condition was tested.

A saturated sample shows a small loop during a biased excitation that never crosses zero induction. Does the area formula still apply?

▼ **Work through the solution**

A settled periodic minor loop encloses an area representing the net cyclic energy under exactly that bias and excursion. Adding a static induction offset does not change $\oint H\,dB$, but the bias can alter the trajectory, differential permeability, and width of the minor loop itself. Do not infer the full major-loop energy from its area. The pickup returns only changes of induction unless an independent dc measurement or a known magnetic preparation supplies the offset. A bias winding may additionally heat the calorimeter, so include its measured power when comparing with thermal loss.

Tutorial on comparing two loss measurements

A data logger reports H_p and B_p correctly but adds an uncorrected $10\,\mu\mathrm{s}$ delay to the induction channel. Estimate the fractional area error of the elliptical example.

▼ **Work through the solution**

At $200\,\mathrm{Hz}$ the extra apparent lag is

$$2\pi f \Delta t = 0.01257\,\mathrm{rad} = 0.72^\circ.$$

The reported area is multiplied by $\sin(12.72^\circ)/\sin(12^\circ)$, about 1.059. An uncorrected delay of only one hundred-thousandth of a second overstates the core loss by about six percent. For a very low-loss specimen with much smaller physical lag, the same delay produces a far larger relative error.

A thermal run on the combined winding and core reports a slope $0.0070\,\mathrm{K/s}$ at $\Delta T = 2.0\,\mathrm{K}$. Recover the core power and the volume-specific loss area at $200\,\mathrm{Hz}$.

▼ **Work through the solution**

Heat deposition in the package is

$$20(0.0070) + 0.015(2.0) = 0.170\,\mathrm{W}.$$

Subtract the estimated $0.0432\,\mathrm{W}$ of winding heating to obtain $0.1268\,\mathrm{W}$ in the core. Dividing by $200 \times 4.0 \times 10^{-5}$ gives $15.85\,\mathrm{J/m^3}$ per cycle. Its closeness to $15.68\,\mathrm{J/m^3}$ should be evaluated against the uncertainties of heat capacity, thermal conductance, resistance, dimensions and electrical phase alignment.

Vary the magnetic circuit and identify the induction scale used by the specimen.

Electromechanics, optics and measurement · 8. A diamagnetic specimen in a gravitational and magnetic trap

A suspended specimen can have exactly the right vertical magnetic force and still slide sideways out of its apparent equilibrium. Derive the force on induced magnetisation, impose gravity, and use the curvature of a measured field to decide whether small disturbances return the specimen to its starting point.

Derive the force on an induced magnetic dipole

Take a small, isotropic, weakly magnetic particle of volume V, mass density ρ and SI volume susceptibility $\chi < 0$. In vacuum or air the imposed field is $\mathbf{B}$, and to first order in χ its magnetic moment is

$$\mathbf{m} = \chi V \mathbf{B}/\mu_0.$$

The field must vary little over the particle for a centre-of-mass dipole description. A permanent magnetic moment aligned with a field would have interaction energy $-mB$. Applying that formula directly with an induced m varying with B doubles the predicted force. Build the induction from zero at fixed source-current control: the energy change at a small increment of imposed field is $-\mathbf{m}\cdot d\mathbf{B}$. Since the induced moment itself increases linearly with field, integration gives

$$U_{\mathrm{mag}} = -\frac{\chi V}{2\mu_0} B^2,$$

$$\mathbf{F}_{\mathrm{mag}} = \frac{\chi V}{2\mu_0} \nabla B^2.$$

Because χ is negative, energy rises with squared induction and the particle moves toward weaker field. The factor one half can also be checked using the instantaneous force $m_j \partial_i B_j$ on a small dipole in a current-free region, inserting the actual induced moment only after the field derivative. For an anisotropic solid the scalar susceptibility must be replaced by its orientation-dependent tensor and the material can rotate. For a strongly magnetisable body the internal field is appreciably changed by its own magnetisation. These effects alter both equilibrium and stiffness. The weak isotropic model provides a controlled starting point for a specimen with calibrated response.

Include gravity by choosing z upward and taking $U = \rho V g z - \chi V B^2/(2\mu_0)$. Equilibrium requires zero horizontal gradients of B^2, and $\partial_z B^2 = 2\mu_0 \rho g/\chi$. The right-hand side is negative: the imposed squared field must decrease upwards at the equilibrium point. Volume cancels because gravity and the weak induced moment are both proportional to volume. A larger sample of identical density and susceptibility ideally levitates at the same height. Real finite-size changes in field averaging supply a way to test whether the point-particle approximation has broken down.

Solve a three-dimensional quadrupole trap

A local source-free field satisfying both divergence and curl constraints is $\mathbf{B} = b(x\,\hat{x}/2 + y\,\hat{y}/2 - z\,\hat{z})$, where b has units of $\mathrm{T/m}$. Its trace of derivatives is

$$b/2 + b/2 - b = 0.$$

Squaring gives

$$B^2 = b^2[(x^2 + y^2)/4 + z^2].$$

In zero gravity the induced magnetic energy therefore has an isolated minimum at the origin. Although each Cartesian component of a static current-free field is harmonic, the sum of their squares need not be. This mathematical fact permits a stable low-field-seeking diamagnet in three dimensions.

Completing the square in the vertical potential puts the equilibrium at $z_e = -\mu_0 \rho g/(|\chi| b^2)$. The two horizontal spring constants are $k_x = k_y = |\chi| V b^2/(4\mu_0)$ and the vertical one is $k_z = |\chi| V b^2/\mu_0$. Divide by mass and take square roots to obtain

$$\omega_z = b\sqrt{|\chi|/(\mu_0 \rho)},$$

with $\omega_x = \omega_y = \omega_z/2$. Gravity shifts the equilibrium in this exactly quadratic potential without changing the frequencies. More general magnetic geometries have curvature that changes with position; in those a gravity-induced shift can move the particle into a region with different stiffness.

Take an illustrative isotropic specimen with $\rho = 2200\,\mathrm{kg/m^3}$, $\chi = -4.0 \times 10^{-4}$, and $b = 200\,\mathrm{T/m}$. The predicted equilibrium is $1.70\,\mathrm{mm}$ below the field zero. The vertical oscillation frequency is $12.1\,\mathrm{Hz}$, the horizontal frequencies are about $6.05\,\mathrm{Hz}$, and the local induction magnitude at the equilibrium is $0.339\,\mathrm{T}$. A check independent of susceptibility is $g = \omega_z^2 |z_e|$ in this ideal quadrupole, obtained by eliminating $b^2|\chi|/(\mu_0\rho)$. A measured $12.1\,\mathrm{Hz}$ and $1.70\,\mathrm{mm}$ give approximately $9.83\,\mathrm{m/s^2}$. Agreement tests the consistency of position and dynamical measurements with the quadratic trap assumption.

Check radial stability from an axial field map

Practical magnets rarely generate a pure quadrupole over the entire useful volume. Suppose a smooth axisymmetric current-free field has on-axis induction $B_0(z)\hat{z}$. Expand about the axis using reflection symmetry: $B_r = a(z) r + O(r^3)$ and $B_z = B_0(z) + c(z) r^2 + O(r^4)$. The zero divergence condition gives $2a + B_0' = 0$. The azimuthal curl condition $\partial_z B_r - \partial_r B_z = 0$ then gives $a' - 2c = 0$, hence $a = -B_0'/2$ and $c = -B_0''/4$. Squaring both components yields the near-axis expansion

$$B^2 = B_0^2 + r^2\left[\frac{(B_0')^2}{4} - \frac{B_0 B_0''}{2}\right] + O(r^4).$$

Thus the axial curvature of squared induction is $2[(B_0')^2 + B_0 B_0'']$, and each transverse curvature is $(B_0')^2/2 - B_0 B_0''$. A weak isotropic diamagnet needs both to be positive for stable harmonic motion. An axial field profile can satisfy vertical balance and produce positive axial curvature while the transverse curvature is negative. With no symmetry one needs the full Hessian matrix of B^2. Positive diagonal elements alone are insufficient if off-diagonal couplings tilt the unstable direction. Its eigenvectors give normal-mode directions; eigenvalues multiplied by $|\chi| V/(2\mu_0)$ give the corresponding linear restoring constants.

At one test position take $B_0 = 0.800\,\mathrm{T}$,

$$B_0' = -84.75\,\mathrm{T/m}.$$

The product is approximately $-67.8\,\mathrm{T^2/m}$, close to the value required for the numerical specimen above. Let $B_0'' = 5000\,\mathrm{T/m^2}$. The transverse curvature is about $3591 - 4000 = -409\,\mathrm{T^2/m^2}$, while axial curvature is about $22365\,\mathrm{T^2/m^2}$. The particle can satisfy vertical balance yet move away under a small horizontal displacement. Reducing B_0'' to $1000\,\mathrm{T/m^2}$ at the same local induction and slope makes the transverse curvature about $2791\,\mathrm{T^2/m^2}$, while the axial curvature remains positive. A field value and a gradient alone cannot distinguish these two cases.

Design measurements that can resolve second derivatives

An axial Hall scan at equally spaced positions $z - h, z, z + h$ estimates B_0' from the centred first difference and B_0'' from the centred second difference. For independent induction noise with standard deviation σ_B, the noise of the gradient estimate is $\sigma_B/(\sqrt{2}h)$, and the noise of the second derivative is $\sqrt{6}\,\sigma_B/h^2$. Halving the spacing while leaving the noise fixed improves the leading truncation error from smooth spatial variations but multiplies second-derivative noise by four. Repeated measurements, a justified fit across several points and a probe smaller than the desired spatial scale can improve the estimate. A fitted polynomial with small residuals at the recorded points does not guarantee a sharply determined curvature.

For $h = 0.50\,\mathrm{mm}$ and $\sigma_B = 0.10\,\mathrm{mT}$, the three-point second-derivative uncertainty is approximately $980\,\mathrm{T/m^2}$. In the first test geometry above the radial curvature lies only $409\,\mathrm{T^2/m^2}$ below zero, while its contribution from B_0'' uncertainty alone is $0.8 \times 980 \approx 784\,\mathrm{T^2/m^2}$. Such a scan cannot establish the sign convincingly from those three readings. Independently displacing a specimen sideways and tracking whether it returns is a direct stability check. Air drag may make a growing instability look slow, so fit a damped or unstable equation over sufficient time and monitor nearby surfaces that might mechanically support the specimen.

Measure susceptibility from dynamics and compare with force balance

An oscillation measurement can determine susceptibility if the field curvature and density are known. For a weak isotropic sample the matrix of squared angular frequencies is $|\chi|\,\mathcal{H}(B^2)/(2\mu_0\rho)$, with $\mathcal{H}$ evaluated at the equilibrium position. An observed oscillation along an eigenvector with curvature $h_j > 0$ gives

$$|\chi| = 2\mu_0\rho\,\omega_j^2/h_j.$$

Simultaneously, the static force-balance equation gives

$$\chi = 2\mu_0\rho g/(\partial_z B^2).$$

The two estimates use different derivatives of the same field map. Their agreement is more informative than fitting both readings by adjusting unmeasured field parameters independently.

For a measured pure quadrupole gradient $b = (200 \pm 2)\,\mathrm{T/m}$, density $2200 \pm 22\,\mathrm{kg/m^3}$, and vertical oscillation frequency $12.10 \pm 0.10\,\mathrm{Hz}$, the dynamical estimate is

$$|\chi| = \mu_0 \rho (2\pi f_z / b)^2 \approx 4.00 \times 10^{-4}.$$

If the quoted errors are independent standard uncertainties, the relative uncertainty is $\sqrt{(0.01)^2 + (2 \times 0.10/12.10)^2 + (2 \times 2/200)^2}$, about 2.8%. The gradient calibration enters squared. Recording more frames of the oscillation improves the frequency measurement only until the field calibration and density dominate. A traceability statement for b, the camera time base and the measurement position is therefore part of the susceptibility result.

If the motion is viscously damped, the observed oscillation frequency is

$$\omega_d = \sqrt{\omega_0^2 - \alpha^2},$$

where the amplitude envelope decays as $e^{-\alpha t}$. A fit to oscillation frequency alone therefore underestimates the undamped stiffness. If $\omega_d = 70\,\mathrm{s^{-1}}$ and $\alpha = 20\,\mathrm{s^{-1}}$, the undamped angular frequency is $72.8\,\mathrm{s^{-1}}$. Since the inferred susceptibility is proportional to frequency squared, using the uncorrected frequency would give only $4900/5300 = 92.5\%$ of the correct value. Visually observing a stable position gives no direct measurement of this dynamic correction. A repeated small displacement at several amplitudes also tests whether cubic restoring terms contaminate the fit.

Account for the finite volume and orientation of a specimen

For a particle that is weakly perturbed by the field but too large for point sampling, the magnetic energy is the spatial integral of the local energy density over its translated volume. For a uniform spherical specimen of radius a centred at $\mathbf{R}$, Taylor expand $B^2(\mathbf{R} + \mathbf{s})$. Inversion symmetry eliminates the linear term and $\langle s_i s_j \rangle = a^2 \delta_{ij}/5$, so the average is $B^2(\mathbf{R}) + a^2 \nabla^2 B^2(\mathbf{R})/10 + O(a^4)$. Differentiating this average yields the leading force correction. In a perfectly linear quadrupole B^2 is quadratic and its Laplacian is spatially constant, so a spherical specimen adds an energy offset but leaves the translational force and stiffness unchanged. A large particle in a nonquadratic trap need not share that property.

Consider an additional on-axis quartic contribution to the field energy density, proportional locally to z^4, as one illustrative component of a measured three-dimensional expansion. Averaging it over a sphere gives $Z^4 + 6Z^2 a^2/5 + \langle s_z^4 \rangle$. The last term is independent of position and drops out of the force. The term proportional to $Z^2 a^2$ modifies even the small-amplitude stiffness. Such a size dependence could be mistaken for different susceptibilities if two samples have materially different diameters. A realistic numerical calculation should average the measured three-dimensional squared field over the actual shape. Extending a one-dimensional polynomial to three dimensions arbitrarily can violate the source-free Maxwell constraints used above.

For an anisotropic small specimen the energy is $-V\,\mathbf{B} \cdot \boldsymbol{\chi} \mathbf{B}/(2\mu_0)$ in the material's orientation. A thin crystalline platelet can experience a torque that changes which susceptibility component contributes to the force. For a weak isotropic ellipsoid with demagnetising factor N_d along an axis, uniform magnetisation satisfies $M = \chi(H_{\mathrm{applied}} - N_d M)$, giving $\chi_{\mathrm{apparent}} = \chi/(1 + N_d \chi)$. This correction is

small for the numerical weak susceptibility used here, but the shape and tensor response of a real sample still need to be recorded for a quantitative comparison across different orientations.

Test a local field map before declaring a trap stable

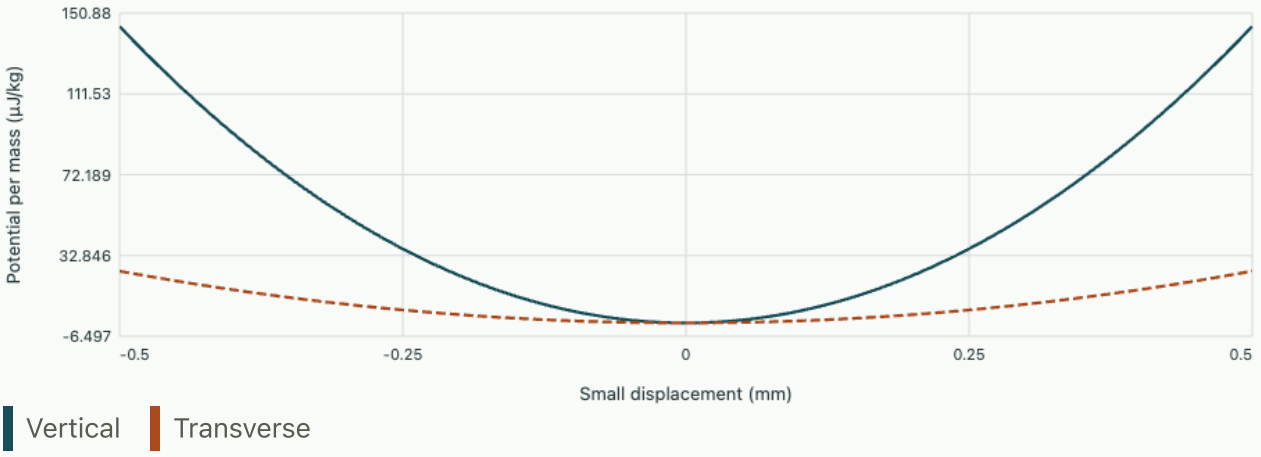
The live calculation varies axial curvature at fixed specimen response, equilibrium field and gradient. It plots the change of total potential per unit mass for a small vertical displacement and a transverse displacement. For the vertical curve the linear magnetic term cancels gravity by construction. Predict which curve will first turn down as positive axial field curvature is increased. At a transverse curvature of zero the quadratic plot is locally flat and higher powers or experimental imperfections decide the outcome. The calculation cannot declare stability at that marginal point. Move the control clearly onto each side of it and compare the resulting real oscillation frequency or exponential growth rate.

Open the live experiment

Try the calculation

Axial B curvature (T/m²) **1000**

Equilibrium field (T) **0.8**



Force balance fixes B′ at -84.75 T/m. Vertical motion: 5.41 Hz oscillation. Transverse motion: 2.26 Hz oscillation.

Is a vertically balanced particle exactly at a magnetic field zero possible in this static weak-diamagnet model with gravity?

▼ **Work through the solution**

Where a smooth imposed field vanishes,

$$\nabla B^2 = 2\sum_j B_j \nabla B_j = 0.$$

The induced magnetic force therefore vanishes at that point, while gravity remains finite. A pure quadrupole displaces the equilibrium away from its zero and creates the required nonzero gradient of squared induction. If a photograph seems to show a specimen at the field zero, test the location of the zero independently, check the imaging perspective and consider other forces such as buoyancy, electrostatic charging or a support. None follows merely from the existence of a zero of magnetic induction.

Test the effect of a nearby fluid

A specimen immersed in a weakly magnetic carrier fluid displaces that fluid as it moves. To first order in the susceptibilities, the magnetic energy change on exchanging equal volumes is proportional to the susceptibility contrast $\chi_{\mathrm{particle}} - \chi_{\mathrm{fluid}}$. The gravitational energy similarly depends on the density contrast. A neutrally buoyant particle has almost no net gravitational weight in the fluid even though it retains inertia and can experience a substantial magnetic force. Oscillatory experiments in liquid also require drag and possibly added inertia of surrounding fluid, so the simple vacuum frequency cannot be applied merely by substituting a density difference everywhere.

For example, take a particle of density $2200\,\mathrm{kg/m^3}$ in a liquid of density $1000\,\mathrm{kg/m^3}$ with magnetic susceptibility small relative to the particle. Its leading gravitational force is only $1200/2200 = 0.545$ of that in air. Under the same ideal linear quadrupole and negligible magnetic-fluid contrast correction, its equilibrium sag shrinks from $1.70\,\mathrm{mm}$ to about $0.93\,\mathrm{mm}$. This change in a static measurement could otherwise be misidentified as an increase in particle susceptibility. Comparing in-air and immersed equilibria therefore requires conditions for both buoyancy and magnetisation of the surrounding material.

Tutorial on force balance and stability

A particle oscillates vertically at $10\,\mathrm{Hz}$ in a pure quadrupole trap. What gravitational displacement is predicted, and how would a measured displacement twice that large challenge the model?

▼ Work through the solution

The vertical angular frequency is $2\pi(10) = 62.83\,\mathrm{s^{-1}}$, so the predicted shift is $g/\omega_z^2 = 2.48\,\mathrm{mm}$ below the field zero. A measured shift near $5.0\,\mathrm{mm}$ would require testing the location of the field zero, the measured frequency, and the curvature variation over the gravitational displacement. Adjusting susceptibility alone cannot reconcile both observations in the quadratic model because it changes stiffness and inverse shift together.

Two specimens have equal volume and identical density. The magnitude of susceptibility of the second is one quarter that of the first. What changes in the ideal quadrupole?

▼ Work through the solution

The second specimen sags four times as far because the restoring force at a given displacement is one quarter as large. Each frequency falls to one half. Its larger sag can take it outside the region where the

quadrupole expansion is accurate even if the first specimen is well described. A volume change alone would leave both predicted quantities intact for specimens of equal density and susceptibility.

Michael Berry and Andre Geim derive local magnetic conditions for diamagnetic levitation.

Connect induced magnetisation to its current and external field.

Electromechanics, optics and measurement · 9. Optical trapping and Brownian calibration of a dielectric bead

An image of a bead remaining near a focus does not establish the strength of its trap. Derive the force pulling a small dielectric sphere into a bright region, compare it with the forward force from scattering and learn what a motion record can tell you about stiffness, drag and temperature. The model here uses an inert dielectric sphere in a homogeneous Newtonian liquid.

From two dielectric boundaries to the energy of an induced dipole

A sphere of radius a and relative permittivity ϵ_b sits in a liquid of relative permittivity ϵ_m. Assuming the sphere is small compared with a wavelength in the liquid, solve its local electric problem in an almost uniform applied field E. Rotational symmetry about the field admits an interior potential proportional to $r\cos\theta$. The exterior consists of the imposed potential $-Er\cos\theta$ and an induced dipolar term proportional to $\cos\theta/r^2$. Equality of potential and of normal displacement at the surface gives the local-field and dipole factors

$$E_{\mathrm{in}} = \frac{3\epsilon_m}{\epsilon_b + 2\epsilon_m} E,$$

$$p = \alpha E,$$

$$\alpha = 4\pi\epsilon_0\epsilon_m a^3 \frac{\epsilon_b - \epsilon_m}{\epsilon_b + 2\epsilon_m}.$$

The field doing the polarising work is the external field in the host at the particle position before inserting it. To turn that work into a potential, build up the dipole from zero rather than assuming a permanent dipole moment. Quasistatic assembly gives

$$-\int_0^E \alpha E'\, dE' = -\alpha E^2/2.$$

For a harmonic peak field average the square over an optical period. In a nearly lossless liquid of refractive index n_m, the peak field is related to travelling-wave intensity by $I = n_m\epsilon_0 c|E|^2/2$. The resulting conservative force is

$$U = -\frac{\mathrm{Re}\,\alpha}{4}|E|^2 = -\frac{\mathrm{Re}\,\alpha}{2n_m\epsilon_0 c} I,$$

$$\mathbf{F}_{\mathrm{grad}} = -\nabla U.$$

The two factors of two have different origins: induced-dipole assembly and time averaging of a sinusoid. Using a permanent-dipole potential with a peak field would miss both. A bead of higher refractive index than the liquid has positive real polarizability and is attracted to increasing intensity. A lower-index inclusion reverses this local tendency. For large particles, strong absorption or rapidly structured fields, a dipole with real static polarizability no longer specifies the force; a field or multipole calculation of the complete momentum balance is required.

Investigate the small-sphere dipole response and its dependence on particle size

Compute a transverse stiffness and test how local it is

Assume the beam near its focal plane has intensity $I(x) = I_0 e^{-2x^2/w^2}$, with power P, waist w defined at $1/e^2$ intensity and $I_0 = 2P/(\pi w^2)$. Put $U_0 = \mathrm{Re}\,\alpha I_0/(2n_m\epsilon_0 c) > 0$. The potential measured above its minimum is $U_0[1 - e^{-2x^2/w^2}]$. Expanding to second order and equating its leading term to $\kappa_x x^2/2$ gives

$$\kappa_x = 4U_0/w^2.$$

Stiffness therefore scales linearly with power, as the third power of sphere radius and as the inverse fourth power of waist at fixed indices. Each scaling applies within the stated approximations.

For a 40 nm radius bead with $n_b = 1.59$, take $n_m = 1.33$, $w = 0.65$ micrometres, vacuum wavelength 1064 nm, $P = 90$ mW and $T = 298$ K. The index contrast factor is 0.1252 and the waist intensity is 1.36×10^{11} watts per square metre. The predicted depth 3.03×10^{-20} J is 7.36 times the thermal energy. The lateral stiffness is 2.87×10^{-7} N/m. A displacement of 50 nm then requires approximately 14.3 fN of transverse balancing force. At a displacement of half a waist, the exact restoring force is the harmonic estimate multiplied by $e^{-1/2} = 0.607$; a Hookean calibration cannot be extrapolated that far.

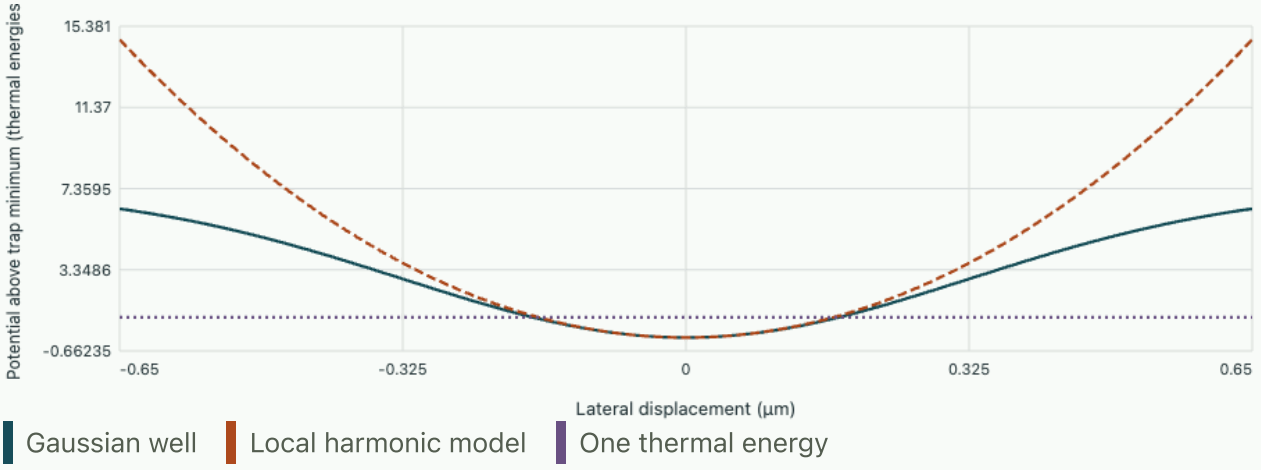
Open the live experiment

Try the calculation

Bead radius (nm) **40**

Optical power (mW) **90**

Beam waist (µm) **0.65**



The well depth is 7.36 kBT. The local model gives an rms displacement of 119.8 nm and a corner frequency of 60.5 Hz at 298 K.

Use the two curves to identify the interval in which the local quadratic energy is within ten percent of the Gaussian prediction. First lower power at fixed waist. Both curves shrink together in energy units while the harmonic thermal displacement grows. Next increase bead radius at fixed power and watch the well deepen, but also evaluate the size parameter $2\pi n_m a/\lambda_0$. It is 0.314 at the default and grows linearly with

radius. Stronger trapping predicted by this simplified model does not establish the accuracy of the dipole approximation at arbitrary sizes.

Compare axial restoring force with radiation pressure

An induced dipole reradiates a portion of the incident field. In a uniform host use the Rayleigh scattering cross section $\sigma_s = (8\pi/3)k_m^4 a^6 |(\epsilon_b - \epsilon_m)/(\epsilon_b + 2\epsilon_m)|^2$. In a common travelling-wave momentum-flux model for a bead in a stationary nondispersive liquid the axial scattering force is approximately $F_s = n_m I \sigma_s / c$; absorption would add a further contribution. This estimate uses the total scattering cross section because the leading dipole pattern has vanishing mean forward cosine. For a larger forward-scattering sphere a transport cross section weights the deflected momentum instead.

In the default case the cross section is roughly 2.05×10^{-18} square metres. The forward force at the waist is about 1.23 fN. A Gaussian axial profile has $I(0,z) = I_0/[1 + (z/z_R)^2]$ with $z_R = \pi n_m w^2/\lambda_0 = 1.66$ micrometres. Its local axial stiffness $2U_0/z_R^2 = 2.20 \times 10^{-8}$ N/m is much smaller than its transverse stiffness. Balancing the constant forward force with that local stiffness predicts an equilibrium about 56 nm downstream of the waist. The small displacement relative to the Rayleigh length supports the linearisation.

Both the axial stiffness and forward scattering force scale with power in this simple model. A power sweep can thus increase confinement while barely shifting the equilibrium offset, a useful discriminating prediction. At fixed focus the offset grows rapidly with bead size because the scattering force scales with the sixth power and the gradient force with the third power. A stable lateral position alone does not establish stable three-dimensional capture. Check the axial restoring slope at a force zero and verify that thermal escape and fluid flow are negligible over the acquisition interval.

Extract stiffness from fluctuations or from a relaxation time

Near a stable point in one Cartesian direction, let x denote displacement from the mean and ζ the viscous drag. At sufficiently low frequencies the inertial term is negligible and the overdamped Langevin equation is

$$\zeta \dot{x} = -\kappa x + \xi.$$

At equilibrium the fluctuating force has zero mean and white covariance $2k_B T \zeta \delta(t - t')$ over the frequencies considered. The deterministic relaxation time is

$$\tau = \zeta/\kappa.$$

Multiplying the equation at positive delay by $x(0)$ and averaging removes the uncorrelated later noise and gives the decaying covariance. Equipartition sets its initial value.

$$C(t) = \frac{k_B T}{\kappa} e^{-|t|/\tau},$$

$$f_c = \frac{\kappa}{2\pi\zeta},$$

$$S_x^{(1)}(f) = \frac{4 k_B T \zeta}{\kappa^2 + (2\pi f \zeta)^2}, \quad f \geq 0.$$

The stated spectrum is one-sided in ordinary frequency. Its integral from zero to infinity equals $C(0)$; doubling another time would double the variance. In an unbounded fluid with viscosity 1.00 mPa s, Stokes drag is $6\pi\eta a = 7.54 \times 10^{-10}$ N s/m. The default predicts $f_c = 60.5$ Hz and rms displacement 120 nm. The latter is roughly 0.18 waist, so the outer part of a long thermal trajectory may sample appreciable anharmonicity. A finite-depth well in an unbounded bath does not have a globally normalisable trapped equilibrium. These local equilibrium predictions apply to a sufficiently long-lived, locally sampled trapped state.

If a voltage detector gives $V = gx$, the spectral corner and normalised autocorrelation time do not depend on the displacement gain g. They determine only κ/ζ. Equipartition using a calibrated spatial variance gives κ independently of drag, provided mechanical fluctuations are in equilibrium and sufficiently resolved. If the viscosity or nearby-wall drag is uncertain, combining a known stage oscillation with the Brownian spectrum can determine further calibration quantities. Keep a stage-driven coherent peak out of the noise integral used for equipartition.

Account for exposure, sample size and unobserved motion

A camera typically reports the position averaged during an exposure of duration t_e. Its variance is the double integral of the covariance over the exposure interval divided by t_e^2. Evaluating the exponential local model and introducing $u = t_e/\tau$ gives a multiplicative factor $2(u - 1 + e^{-u})/u^2$. At $u = 1$ the reported physical-motion variance is only 0.736 of the instantaneous variance. Interpreting it as the full variance would overestimate stiffness by approximately 36 percent. Independent pixel-localisation noise adds additional variance with the opposite sign of bias.

Estimate that localisation noise with an effectively immobilised reference bead at a comparable brightness. Test the averaging correction by changing exposure duration at constant sampling frequency where possible. A camera that reports overlapping exposures has correlated measurement errors and needs a different covariance model. A long series of closely spaced frames also contains fewer independent thermal configurations than its frame count suggests. Compare estimates across several blocks substantially longer than the relaxation time; residual low-frequency drift can dominate even when high-frequency position precision looks excellent.

A separate check uses a known uniform liquid flow. At a steady small displacement x_s, force balance gives

$$\kappa x_s = \zeta u_{\mathrm{flow}}.$$

For the default drag and stiffness a flow of 10 micrometres per second gives an offset of 26.3 nm. Repeat with reversed flow and use half the displacement difference to cancel an unknown zero-force position. Near a wall, the local flow at the bead and wall-corrected drag may differ from a stage velocity and the

unbounded Stokes value. Those effects alter the mechanical calibration even if the optical field is unchanged.

A known stage oscillation calibrates position in a fluid

Brownian fluctuations alone determine the product of several unknown calibration parameters. A weak sinusoidal translation of the sample cell adds a known mechanical signal at a chosen frequency. Work in laboratory coordinates with a fixed optical trap and with fluid moving uniformly at velocity $u(t)$ near the bead, well away from walls in the idealised model. The drag is now $-\zeta(\dot{x}-u)$. Superposing a stage displacement $A\cos\omega t$ on thermal fluctuations gives a deterministic response whose complex amplitude satisfies $(\kappa - i\omega\zeta)\widetilde{x} = -i\omega\zeta A$ in the negative-exponential convention. Therefore its amplitude ratio is

$$\frac{|\widetilde{x}|}{A} = \frac{\omega\tau}{\sqrt{1+\omega^2\tau^2}}.$$

At the trap corner the bead displacement amplitude is the known stage amplitude divided by the square root of two. A calibrated 80.0 nm stage oscillation there produces a predicted bead amplitude of 56.6 nm. If a position-sensitive receiver records a coherent peak amplitude of 0.170 V, its displacement gain is 3.01 mV/nm. At very low drive frequencies the fluid velocity is small for fixed stage amplitude and the trap holds the bead nearly stationary. At high enough frequency the bead follows the fluid displacement within the overdamped model, but finite hydrodynamic, optical and electronic bandwidths eventually matter.

After determining the gain, convert the surrounding voltage spectrum into spatial units and use the area of the stochastic fluctuation spectrum to find stiffness. Exclude the calibrated line and account for the finite spectral averaging window; a narrow drive that leaks into neighbouring bins can inflate the inferred thermal variance. Dividing measured stiffness by the independently determined corner rate then gives drag, which can be compared with the known viscosity and sphere radius. This avoids assuming free-space Stokes drag in the first step of a near-wall measurement.

There are two limiting systematic effects that a good fit to a single drive frequency cannot resolve. A frequency-dependent stage calibration changes the actual fluid displacement, and motion of the optics relative to the cell generates a direct detector background even with the bead immobilised. Repeat the drive above and below the corner, compare coherent phase as well as amplitude and acquire a trapped and immobilised reference. Correct conversion should explain the entire complex response with one spatial gain and one corner time. A single calibration number obtained at one frequency can hide a phase lag in the stage controller or in the optical position channel.

One last cross-check reverses the stage drive phase by half a period. The coherent bead response must reverse, while thermal background power remains unchanged. A component that fails this reversal could come from ambient vibration unrelated to the stage. Averaging the two traces with and without a sign change separates these contributions before either is used in an optical-force calibration.

Tutorial · Separate a unit-conversion error from excess noise

In a single-axis trap at 298 K a calibrated detector reports a variance of 2500 square nanometres after removal of localisation noise and exposure correction. An independent spectral fit gives a corner of 175 Hz. Infer stiffness, drag and the equivalent unbounded hydrodynamic radius for a liquid of viscosity 1.00 mPa s. Compare against a nominal 80 nm radius bead.

▼ **Infer the three quantities and test their consistency**

A nanometre squared is 10^{-18} square metres, so the variance is 2.50×10^{-15} square metres. Equipartition gives 1.65×10^{-6} N/m. Dividing by $2\pi f_c$ gives drag 1.50×10^{-9} N s/m and radius $\zeta/(6\pi\eta) = 79.5$ nm. The agreement supports the local overdamped description under the assumed temperature and viscosity. Forgetting the factor of 2π would predict a radius roughly six times larger and could be mistaken for a huge near-wall correction.

Now imagine that additional uncorrelated read noise had been left in the variance. This raises apparent variance, lowers the equipartition stiffness and would lower the inferred drag at the same fitted corner. Nearby-wall drag for a given radius instead normally raises hydrodynamic resistance for parallel translation. Their effects on the inferred radius thus differ in sign under the stated assumptions. Varying height and independently calibrating the measurement noise are appropriate diagnostics before attributing a discrepancy to bead dimensions.

Tutorial · Test a force-law inference with power and concentration controls

At fixed bead radius a trap is operated at two powers differing by a factor of two. Measured corner frequencies increase from 42 to 84 Hz, but rms displacement falls only from 110 to 87 nm. What does the harmonic model with fixed drag and temperature predict? Estimate a power-independent additive position-noise variance that could reconcile the two readings.

▼ **Use two variances to separate scale and offset**

The doubled corner gives doubled stiffness if drag is unchanged. Physical thermal variance should halve. The predicted rms at higher power would be 77.8 nm if the lower-power value were entirely physical. Let the low-power measured variance be $v + b$ and the high-power one $v/2 + b$, where b is a constant additive noise variance. Solving from 12,100 and 7,569 square nanometres gives $v = 9062$ and $b = 3038$ square nanometres, or approximately 55.1 nm rms noise.

This is a large inferred detector noise and demands an independent immobilised reference check. Power-dependent heating of the liquid changes temperature and viscosity, so an unchanged product inferred from only two quantities cannot prove that heating is absent. Test several powers, reverse their order to reveal slow drift and vary bead concentration to exclude overlapping particles in the tracking window. The inferred additive term should remain constant across those controls if it comes from the specified detector model.

Tolić-Nørrelykke and colleagues · position and force calibration using thermal and driven measurements

Electromechanics, optics and measurement · 10. Coupled optical and mechanical resonances

A vibrating mirror can imprint its position on an optical field, while the stored light pushes back on the mirror. An apparent pair of peaks can indicate two hybrid modes, interference at a readout port or two unrelated resonances. Starting from energy in a moving cavity, derive the coupled response and learn which measurements establish a coupling strength and its uncertainty.

From a moving boundary to a force per photon

Take a nearly empty Fabry–Pérot cavity of optical length L with one moving end. A longitudinal resonance has $\omega_c = m\pi c/L$ for integer m, neglecting frequency-dependent mirror phase. For a small extension x, $\delta\omega_c = (\partial\omega_c/\partial x)x$ with signed slope $G_x = -\omega_c/L$. At fixed photon number n the optical energy is $n\hbar\omega_c(x)$. Its derivative gives the radiation-pressure force on the coordinate. For this mirror $F = -n\hbar G_x > 0$ pushes outward. The derivative must be at fixed photon number. Differentiating a driven steady-state photon number as though it were a closed-system energy would incorrectly include work exchanged with the pump.

For a 10 mm cavity resonant near 1064 nm the frequency slope magnitude in ordinary frequency is about 28.2 MHz per nanometre. A static shift of 10 pm therefore changes the optical resonance by approximately 282 kHz. One photon exerts a mean force $\hbar\omega_c/L \simeq 1.87 \times 10^{-17}$ N. An equivalent classical derivation counts a momentum change twice per normal reflection for each relevant round trip. The two descriptions yield the same mean force. The mirror displacement is the properly normalised mode coordinate; in a flexural device the optically sampled position and the total physical mass are generally different from centre-of-mass displacement and mass.

Introduce a single mechanical eigenmode with effective mass m_{eff}, angular frequency Ω_m and displacement scale $x_{\mathrm{zpf}} = \sqrt{\hbar/(2m_{\mathrm{eff}}\Omega_m)}$. The single-photon coupling magnitude is $g_0 = |G_x|x_{\mathrm{zpf}}$. This normalisation is convenient even when the observed vibrations are entirely classical. For $m_{\mathrm{eff}} = 10$ ng and $\Omega_m/(2\pi) = 1.0$ MHz, $x_{\mathrm{zpf}} = 9.16 \times 10^{-16}$ m and $g_0/(2\pi) = 25.8$ Hz for the stated optical geometry. The small displacement scale explains why a coherent pump is normally used to make the interaction measurable.

How a laser couples frequencies that are otherwise far apart

An optical carrier oscillates at hundreds of terahertz, whereas the mirror may vibrate at only megahertz. Direct comparison of these two bare frequencies would never suggest a resonant exchange. A strong coherent pump supplies a phase reference and converts a mechanical displacement into optical sidebands. Move into a frame rotating at the pump frequency and write the cavity amplitude as its large steady part plus a small perturbation. The leading frequency-shift interaction is proportional to photon number times displacement. Expanding the intensity around the large pump keeps terms linear in both the probe field and displacement. The resulting coupling is $G = g_0\sqrt{n_c}$, where n_c is mean intracavity pump photon number.

Tune the pump about one mechanical frequency below the optical cavity resonance. In the resolved-sideband and weak-fluctuation limit a near-resonant process moves one excitation between the mechanical mode and an upper optical sideband. The rapidly rotating opposite-sideband terms average

out. Now the two coupled frequencies are an optical sideband offset and the mechanical frequency, measured relative to a common rotating reference. The large carrier frequency has already been accounted for by the pump. The pump phase can be absorbed into the mode definitions, so G may be chosen positive for this isolated pair.

If instead the pump is placed near the upper sideband, the surviving interaction creates or annihilates excitations in pairs. The passive exchange equations below cease to apply and strong pumping can produce amplification or instability. A bare two-oscillator machine driven mechanically near a shared natural frequency often has analogous exchange equations without a carrier pump. Identical matrices can describe different energy supplies; identify the drive and rotating reference before assigning a literal photon-to-phonon energy conversion.

Derive complex poles and the effect of unequal damping

Let a, b be small energy-normalised optical and mechanical complex amplitudes. In a frame where the uncoupled optical sideband has zero frequency, let the mechanical detuning be δ. Let κ, Γ denote energy linewidths in angular frequency, so amplitude decay rates are half these values. With an optical probe s at offset ω, the linear system is

$$\begin{pmatrix} \kappa/2 - i\omega & iG \\ iG & \Gamma/2 - i(\omega - \delta) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} s \\ 0 \end{pmatrix}.$$

Zeros of the determinant give autonomous complex mode frequencies. In the lossless case the eigenvalues are $\delta/2 \pm \sqrt{G^2 + (\delta/2)^2}$. Far off resonance each branch approaches a bare mode. At zero detuning the two eigenvectors have equal optical and mechanical weights and the splitting is $2G$. The eigenvectors rotate continuously through the avoided crossing, an important discriminator from two uncoupled lines simply passing through one another.

At zero detuning with losses, completing the square in the determinant gives

$$\omega_{\pm} = -i(\kappa + \Gamma)/4 \; \pm \sqrt{G^2 - (\kappa - \Gamma)^2/16}.$$

The real parts separate once $G > |\kappa - \Gamma|/4$. The shared imaginary part above that threshold corresponds to a hybrid energy linewidth $(\kappa + \Gamma)/2$. Resolving two intensity features on a given instrument is a stronger, readout-dependent condition. The dimensionless cooperativity $C = 4G^2/(\kappa\Gamma)$ is yet another measure, comparing interaction against the product of dissipation scales. A value greater than one does not automatically prove two resolved normal-mode poles in a device with very unequal linewidths.

Use frequency numbers divided by 2π consistently from here for examples. With optical linewidth 240 kHz, mechanical linewidth 40 kHz and coupling 160 kHz, the real pole separation is

$$2\sqrt{160^2 - 50^2} = 304\,\mathrm{kHz}.$$

Both hybrid energy linewidths are 140 kHz, and cooperativity is 10.7. If coupling falls to 45 kHz, cooperativity is still about 0.844 and no real zero-detuning pole separation appears. At 50 kHz the square

root reaches zero. Near that threshold even a small linewidth or detuning calibration error has a large effect on a coupling estimate inferred from a single splitting.

Compute what an optical or mechanical receiver actually sees

Invert the matrix. With $D = (\kappa/2 - i\omega)[\Gamma/2 - i(\omega - \delta)] + G^2$, the small amplitudes are $a = s[\Gamma/2 - i(\omega - \delta)]/D$ and $b = -iGs/D$. The poles are shared. Their numerators are different, so the two channels can show different apparent maxima. A port in reflection also has a directly reflected field that interferes with the leaking intracavity amplitude. Its intensity can display a narrow dip or asymmetric feature without a matching resonance-height pattern inside the cavity. A legitimate forward model must describe the measured port.

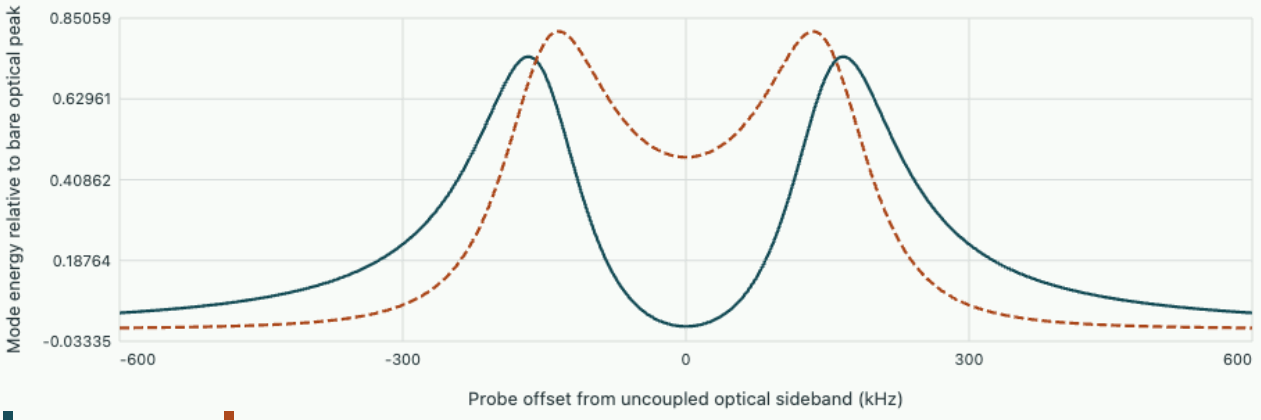
Open the live experiment

Try the calculation

Coupling G/(2π) (kHz) **160**

Optical linewidth κ/(2π) (kHz) **240**

Mechanical detuning (kHz) **0**



At zero bare detuning the real pole separation is 304.0 kHz. For the selected coupling and linewidths the cooperativity is 10.67. The mechanical linewidth is fixed at 40 kHz.

The graph plots the internal squared amplitudes relative to the bare on-resonance optical probe energy. Set the mechanical detuning to zero and reduce the coupling through 50 kHz. Compare the optical dip with the predicted real pole separation. Then set a substantial positive detuning and follow the transfer of mechanical response onto one branch. As a control set the coupling exactly to zero. The optical curve recovers the bare cavity Lorentzian and the mechanical probe-induced response vanishes. A thermally excited mechanical background, absent from this driven calculation, could remain in a real detector.

At zero detuning and zero probe offset,

$$a/a_{\mathrm{bare}} = 1/(1 + C).$$

For the default parameters the internal optical probe intensity at the centre is only $1/(1+10.67)^2 = 0.00735$ of the bare peak. This deep feature alone still does not establish the pole spacing. An incoherent background, a frequency-dependent detector transfer or coherent feedthrough can alter peak and dip positions. Fit complex quadratures across a detuning series when available, and verify bare linewidths by independent ringdown or far-detuned spectra.

Check how energy linewidth and amplitude ringdown are related for one isolated resonance

Watch energy exchange in time and identify the linewidth convention

In an ideal equal-linewidth system starting with optical perturbation $a(0) = a_0$, $b(0) = 0$ at zero detuning, change to symmetric and antisymmetric modes. Recombination gives

$$a(t) = a_0 e^{-\kappa t/2} \cos Gt,$$

$$b(t) = -i a_0 e^{-\kappa t/2} \sin Gt.$$

Their normalised total energy decays as $e^{-\kappa t}$, while energy moves back and forth between optical sideband and vibration. The first full transfer occurs at $t = \pi/(2G)$ if the rotating-wave approximation holds. For $G/(2\pi) = 100$ kHz, this is 2.50 microseconds. With a shared energy linewidth of 20 kHz in ordinary frequency the surviving energy fraction then is

$$e^{-2\pi(20\ \mathrm{kHz})(2.50\ \mu\mathrm{s})} = 0.730.$$

A gated excitation shorter than the exchange period but long enough to avoid unwanted modes can prepare an approximately optical initial perturbation. Independent detection of an optical revival and a delayed mechanical response tests coherent exchange more strongly than the presence of two peaks alone. Comparing times of maxima directly with one half the reciprocal measured optical linewidth misidentifies the exchange physics. Check the decay of the sum of appropriately calibrated energies and the relative phase of the exchange. Electrical or optical feedthrough can produce a prompt apparent response with no mechanical build-up.

Infer a coupling with propagated measurement errors

When the two hybrid pole frequencies have been obtained from a justified fit and zero bare detuning is established, the coupling follows from the measured real separation S:

$$G = \sqrt{S^2/4 + (\kappa - \Gamma)^2/16}.$$

All quantities here can be expressed in the same ordinary-frequency units. If $S = 300 \pm 8$ kHz, $\kappa = 240 \pm 10$ kHz and $\Gamma = 40 \pm 4$ kHz, the estimate is 158.1 kHz. Independent small errors give

$$\partial G/\partial S = S/(4G) = 0.474,$$
$$\partial G/\partial \kappa = -\partial G/\partial \Gamma = (\kappa - \Gamma)/(16G) = 0.0791.$$

Combining in quadrature gives about 3.9 kHz uncertainty. A common frequency calibration can correlate all three inputs, and a fitted detuning and splitting from one scan may also be correlated. Use their covariance rather than multiplying independent error bars blindly.

A second inference varies intracavity pump power. Since $G^2 = g_0^2 n_c$, plot the squared fitted coupling against independently calibrated n_c. A nonzero intercept could indicate an additional direct coupling, baseline fitting bias or a mode incorrectly assigned at low power. The linewidths themselves can vary with pump power because of heating or backaction, so fit or measure them across the scan. Incident optical power is not identical to stored photon number; cavity coupling, detuning and loss determine their relation.

A measured frequency shift and linewidth change reveal optical feedback

The same instrument can be operated below the resolved splitting regime and read out a dispersive shift of the mechanical response. Algebraically eliminate the optical amplitude in the two-mode equations while driving the mechanical mode weakly. Its inverse susceptibility acquires a self-energy contribution proportional to the optical susceptibility. Let x be probe frequency measured from the bare optical sideband and use the same linewidth convention as above. In the mechanical denominator the additional term is

$$\Sigma(x) = \frac{G^2}{\kappa/2 - i x} = \frac{G^2(\kappa/2 + i x)}{(\kappa/2)^2 + x^2}.$$

Its real part adds to the amplitude damping, while its imaginary part shifts the resonant condition. At very weak coupling one can evaluate it at the bare mechanical frequency $x = \delta$, because variation across the narrow mechanical peak is small. Near an avoided crossing or when the mechanical peak samples a large change in optical susceptibility, solve the coupled poles directly; simply adding a linewidth and shift then neglects the very hybridisation being tested.

In ordinary-frequency units choose $G = 10$ kHz, optical linewidth 200 kHz, mechanical linewidth 2 kHz and detuning $+100$ kHz. At that offset the optical amplitude half-width is 100 kHz, giving $\mathrm{Re}\,\Sigma = 0.50$ kHz and $\mathrm{Im}\,\Sigma = 0.50$ kHz. The full mechanical linewidth increases from 2.0 to approximately 3.0 kHz. Its centre moves upwards by about 0.50 kHz with the adopted matrix signs. Moving to minus 100 kHz reverses the shift and preserves the added damping in this symmetric model. At zero detuning the predicted shift vanishes but the added linewidth is largest.

A scan on both sides is more discriminating than measuring broadening alone. A background rise in mechanical temperature changes integrated thermal noise and may change intrinsic loss, but need not reproduce this odd shift and even broadening pattern. Measure intrinsic linewidth far from the optical response and subtract that baseline. An unknown constant frequency offset can be removed with the two symmetric detunings, but an optical power dependence of cavity frequency can move the actual symmetry centre. Monitor the optical resonance during the scan.

The effective susceptibility describes classical feedback through the delayed optical field. It does not by itself specify the final mechanical temperature. A bath occupation or noise spectrum is also required to convert a damping change into an energy distribution. A driven response peak that broadens and lowers in amplitude is evidence for a changed susceptibility; thermal cooling additionally requires a calibrated measurement of the fluctuating mechanical displacement and its optical transduction.

Tutorial · A high cooperativity without resolved modes

A device has ordinary optical linewidth 1.00 MHz, mechanical linewidth 1.0 kHz and pump-enhanced coupling 30 kHz. Calculate cooperativity, compare the coupling with the real-splitting threshold and estimate the relative intracavity probe intensity at the central frequency. What observation would distinguish a narrow interference window from two separated poles?

▼ **Evaluate the three criteria independently**

In kHz units the cooperativity is

$$4(30)^2/(1000 \cdot 1) = 3.6.$$

The real-splitting threshold is $999/4 = 249.75$ kHz, far above the actual coupling. Yet the zero-detuning centre intensity is $1/(1+3.6)^2 = 0.0473$ relative to the bare optical peak. A narrow, deep structured feature is possible with no separated real pole frequencies in this model. The linewidth contrast permits mechanical response to interfere strongly within a broad optical susceptibility.

Acquire both amplitude and phase while scanning the absolute mechanical detuning and compare a coupled complex-response model with independently measured bare parameters. A pair of well-resolved poles should trace two consistent trajectories in both observables and can generate underdamped exchange in the time domain. A single optical intensity scan containing two flanks around a dip is insufficient to make that inference.

Tutorial · Diagnose an erroneous estimate of the single-photon coupling

An independent linewidth analysis gives pump-enhanced couplings 80 and 160 kHz at photon numbers 4.0×10^6 and 1.6×10^7. Estimate $g_0/(2\pi)$ and predict the coupling at 9.0×10^6 photons. A calibration revision doubles all inferred photon numbers while leaving the measured spectra unchanged. Which conclusions move?

▼ **Use coherent amplitude scaling and separate calibration from spectroscopy**

Divide the respective couplings in ordinary frequency by the square roots of their photon numbers. Both give 40 Hz and the intermediate prediction is 120 kHz. Doubling all inferred photon numbers lowers the inferred single-photon coupling by $1/\sqrt{2}$, to 28.3 Hz. The observed enhanced couplings, pole locations and cooperativities computed from measured linewidths stay the same. A prediction labelled by fixed

physical photon number changes; a prediction at a repeatable laser setting can still be calibrated empirically.

To test the absolute calibration, compare the inferred single-photon value with the independent cavity frequency slope times the mechanical zero-point displacement. This requires an effective mass appropriate to the measured displacement and the actual cavity optical overlap. An incorrect choice of mechanical mode normalisation can mimic an error in stored photon number. Agreement across a power ratio tests the square-root scaling but cannot by itself determine the absolute optical or mechanical calibration.

Aspelmeyer, Kippenberg and Marquardt · cavity optomechanics, linearisation and strong coupling

Electromechanics, optics and measurement · 11. Thermal radiance, étendue and detector calibration

A square millimetre of a heated object emits thermal radiation into an entire hemisphere. A detector sees only the rays accepted by its optics and only wavelengths to which it responds. Derive the equilibrium spectrum, follow it through an optical collection system and infer what temperature or emissivity a calibrated reading can actually determine.

Count optical states in a large equilibrium cavity

Take a cubic airless cavity of volume V in thermal equilibrium at temperature T. For its bulk mode density we can use periodic boundary conditions. Allowed wavevectors are separated by $2\pi/L$ in each coordinate. Count a spherical shell in wavevector space, divide by the volume per state and include two independent transverse polarisations. Substituting $k = 2\pi\nu/c$ gives

$$g(\nu)d\nu/V = 8\pi\nu^2 d\nu/c^3.$$

Using standing waves in an ideally reflecting large box gives the same leading volume density when directions and polarisations are counted consistently.

One oscillator mode of frequency ν can hold $m = 0, 1, 2, \ldots$ photons. Its thermal occupation has probabilities proportional to $e^{-mh\nu/k_BT}$. Summing the geometric series and taking its logarithmic derivative gives

$$\langle m\rangle = 1/(e^{h\nu/k_BT} - 1).$$

Multiplying by the density of modes and energy per photon gives spectral energy density u_ν. The zero-point contribution of a harmonic oscillator is not part of the thermal power exchanged between these black surfaces. At low frequency the result approaches an energy k_BT per mode; extending that classical limit to arbitrarily high frequency would make the integrated power diverge.

An isotropic distribution of energy travelling at speed c produces a normal flux through a small aperture weighted by the direction cosine. Per unit solid angle its radiance is $B_\nu = cu_\nu/(4\pi)$. Transform to equal wavelength intervals using $B_\lambda d\lambda = B_\nu|d\nu|$. The two useful spectral forms are

$$B_\nu = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/k_BT} - 1},$$

$$B_\lambda = \frac{2hc^2}{\lambda^5}\frac{1}{e^{hc/(\lambda k_BT)} - 1}.$$

The wavelength expression has units of watts per square metre per steradian per metre of wavelength. To express it per micrometre multiply by one millionth. At 400 K and 10 micrometres the result is 33.56 W per square metre per steradian per micrometre. At 295 K it is 9.14 in the same units. Equal wavelength and equal frequency bins carry different widths, so their distributions reach their maxima at different converted wavelengths. A spectrum's peak should always name its horizontal variable and spectral-density convention.

Distinguish emitter power from brightness in a chosen direction

For a black surface with angle-independent radiance into its outward hemisphere, integrate $\cos\theta\, d\Omega$. The result is π, not 2π. The emitted spectral power per unit area is πB_λ. Integrate over frequency using $x = h\nu/(k_B T)$. The remaining integral $\int_0^\infty x^3/(e^x - 1)\, dx = \pi^4/15$ follows by expanding $1/(e^x - 1)$ as a sum of exponentials and integrating term by term. This yields Stefan–Boltzmann exitance $M = \sigma T^4$ and identifies the constant $\sigma = 2\pi^5 k_B^4/(15 h^3 c^2)$.

A black patch of 1.0 square millimetre at 400 K emits about 1.45 mW into its entire outward hemisphere over all wavelengths. The same patch at 295 K emits about 0.429 mW. Replacing one by the other changes all-wavelength hemispherical power by about 1.02 mW. This is not the power a small infrared receiver sees through a narrow filter. The integrated emitter number is valuable for energy conservation, but by itself contains no receiver area, direction, passband or optical transmission.

The surface spectrum of a body in thermal equilibrium has directional spectral emissivity equal to its absorptivity in the corresponding reciprocal optical modes under the usual reciprocal-material conditions. An opaque diffuse grey surface with emissivity ϵ in an isotropically illuminated background can be modelled by $\epsilon B_\lambda(T) + (1 - \epsilon) B_\lambda(T_{\mathrm{bg}})$. It emits less than a black surface but also reflects background. A shiny object viewed in a warm room may consequently display a radiance temperature nearer the reflected scene than its own physical temperature. Specular surfaces require explicit directional knowledge of that scene.

Derive the geometric acceptance without counting the angle twice

An area element dA normal to the collection axis emits into $d\Omega$ at angle θ a spectral power $B_\lambda \cos\theta\, dA\, d\Omega$. For a uniformly filled cone of half-angle α and a uniform patch of area A, integrating gives $A\pi \sin^2\alpha$. This is its geometrical throughput in air when there is no vignetting. Optical étendue more generally uses $n^2 dA \cos\theta\, d\Omega$. A lossless passive optical relay in geometrical optics preserves this phase-space measure; correspondingly L/n^2 is conserved along a ray in transparent media under the usual stationary conditions.

A direct lens example makes the restriction tangible. An image relay enlarging a uniformly bright patch by a factor of two in each transverse direction forms an image four times the area. In a lossless paraxial design filling the original pupil, each image point then occupies one quarter the solid angle. A detector covering the whole enlarged image and accepting that cone can receive the original power. Placing a fixed small sensor at the image plane intercepts less area instead. Higher magnification has not created additional radiance, even if it makes a small detail easier to resolve.

For the same 1.0 square millimetre patch seen within ten degrees, the projected acceptance is 9.47×10^{-8} square metre steradian. A wavelength-independent combined transmission and detector absorption of 0.65 leaves effective acceptance 6.16×10^{-8} square metre steradian. Integrating the Planck curve from 8 to 12 micrometres gives source and background band radiances 133.74 and 35.43 W per square metre per steradian. Their difference yields only 6.05 microwatts in the receiver. Keeping a single midpoint radiance times four micrometres happens to be fairly close here, but there is no fixed universal midpoint correction across temperatures or filters.

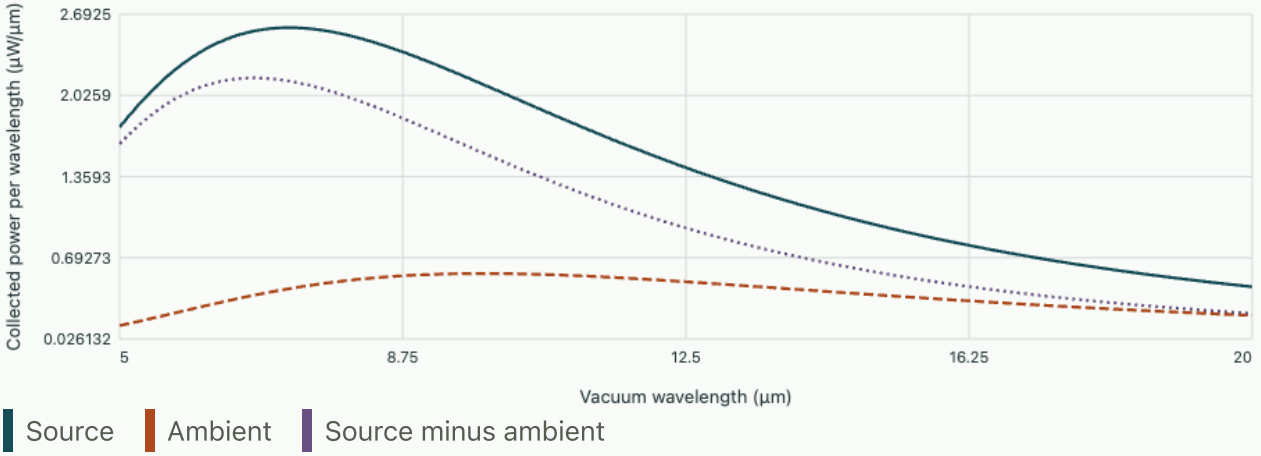
Open the live experiment

Try the calculation

Source temperature (K) **400**

Ambient temperature (K) **295**

Collection half-angle (°) **10**



The net 8–12 µm signal is 6.053 µW for a 1 mm² area and 65% throughput. Ideal source peak wavelength is 7.24 µm.

Set source and ambient temperatures equal. The emitted and ambient curves remain finite while the differenced instrument reading vanishes. Increase the cone half-angle from ten to twenty degrees and predict the scale factor $\sin^2 20^\circ / \sin^2 10^\circ = 3.88$. The spectral shape should stay fixed if transmission is separable from angle. If real data shift spectral shape with aperture, inspect filter angle dependence, source nonuniformity and stray light. The integral in the readout is computed on the specified 8–12 micrometre band even though the graph shows wavelengths beyond it.

Convert received power to calibrated voltage or counts

A thermal detector may respond to absorbed power through a temperature rise and an electrical responsivity. A photon-counting detector measures absorbed photons, so its count rate contains a factor $\lambda/(hc)$ multiplying spectral power before integration. Two receivers with identical geometric acceptance and nominal quantum efficiency can respond differently to a changing source temperature if one output is proportional to energy and the other to photon number. Calibrate the actual electrical transfer and make the time interval explicit for accumulated counts.

For an ideal linear power detector let $V = V_0 + gP$, with offset V_0 and gain g. Two calibrated sources spanning the use range determine both. A black reference port at the same ambient temperature can also reject a common background in a chopped measurement. A filter, window or warm lens that absorbs thermal radiation also emits it. If these optics remain unchanged and stable, part of their own emission cancels in a reference difference. A reference traversing a different path generally leaves a residual input that is not captured by an electronic dark measurement.

Suppose effective source-minus-ambient powers of 2.00 and 8.00 microwatts yield 0.120 and 0.420 V. The fitted slope is 0.0500 V per microwatt with an offset of 0.020 V. A reading of 0.320 V implies 6.00 microwatts after offset correction. Substituting the same reading directly into a single-point gain with zero assumed offset would bias the inferred temperature. Rechecking a third reference near six microwatts tests linearity, while repeated zero-input readings at realistic instrument temperature test offset drift.

Invert the radiance and propagate uncertainties

A narrowband blackbody radiance measurement can be inverted analytically by solving Planck's expression for T. In a broad passband use numerical quadrature across the calibrated spectral response and solve a scalar monotone equation when emissivity and background are known. The derivative at a single wavelength is particularly informative for estimating precision. With $x = hc/(\lambda k_B T)$,

$$\frac{\partial \log B_\lambda}{\partial \log T} = \frac{x}{1 - e^{-x}},$$

$$T = \frac{hc/(\lambda k_B)}{\log[1 + 2hc^2/(\lambda^5 B_\lambda)]}.$$

At 10 micrometres and 400 K the logarithmic sensitivity is 3.70. A one-percent relative radiance uncertainty therefore gives about 1.08 K temperature uncertainty if all other quantities are exact. An unknown five-percent emissivity error can dominate that estimate. A two-colour temperature ratio may reduce uncertainty from an emissivity that is constant across the two narrow bands, but cannot cancel an arbitrary spectral emissivity or a large reflected ambient component.

As an example consider a grey object at 400 K and ambient 295 K with emissivity 0.70. At ten micrometres its outgoing radiance is $0.70(33.56) + 0.30(9.14) = 26.24$ in the earlier spectral units. Inverting as a perfect blackbody gives a brightness temperature of about 375 K. The difference is a physical radiance–temperature degeneracy. Repeat at known ambient temperatures while keeping the object thermally stabilised to constrain the reflected fraction. If heating the background changes the actual object temperature, a second independent temperature sensor or a transient thermal model is necessary.

When detector heat capacity sets the measured signal

Consider a thermal receiver of heat capacity C_d coupled to its surroundings with thermal conductance G_d. For small excursions its temperature obeys $C_d \dot{\theta} + G_d \theta = P_{\mathrm{abs}}(t)$, where θ is the excess above its operating point. A voltage thermometer with responsivity r_T converts this into output. Its frequency-dependent power responsivity is $r_T / (G_d - i\omega C_d)$. The geometric and spectral calculations determine absorbed optical power, while this thermal equation determines how much of its modulation appears at the electrical output.

Take $C_d = 2.0$ microjoules per kelvin and $G_d = 0.10$ milliwatt per kelvin. The thermal time constant is 20 ms and ordinary corner frequency is 7.96 Hz. A sinusoidal absorbed-power amplitude of 6.0 microwatts would give a static temperature rise of 0.060 K. At 20 Hz the amplitude is reduced by $1/\sqrt{1 + (2\pi f \tau)^2} = 0.370$, leaving about 22.2 mK with a phase lag magnitude near 68 degrees. A receiver calibrated with a steady optical power is not automatically calibrated for this chopped experiment.

A square-wave chopper also changes the relation between time-domain peak differences and a lock-in's fundamental reading. For equal hot and reference dwell times, a square wave switching between powers P_1, P_0 has fundamental sinusoidal peak amplitude $2(P_1 - P_0)/\pi$, with different factors for rms and peak reports. The thermal transfer attenuates this component at the chopper frequency and the odd harmonics at their respective frequencies. Substituting the full hot-minus-reference power directly into the first-harmonic response would overestimate the expected signal.

Measure a frequency sweep on a stable reference and test whether amplitude and phase share a single time constant. A geometrical error rescales the amplitude but ordinarily leaves this thermal phase unchanged. A second thermal mass or an electronic filter can introduce another pole. Lowering chop frequency increases thermal signal until the static limit but can expose low-frequency environmental drift. Choose the frequency by comparing the calibrated response with the measured noise spectrum.

Finally, a reference and an unknown source must illuminate the same active area if responsivity varies across the receiver. Imaging one source as a small spot and the other as a filled patch can create a calibration bias even when total spectral powers agree. Translating a reference spot across the detector or imaging the same aperture successively at both temperatures tests that spatial uniformity assumption.

If the sensor is used on a source much smaller than the calibrated reference aperture, record an image or a field-stop scan to determine its filling fraction. Otherwise even exact knowledge of filter transmission and thermodynamic temperature leaves an unknown geometrical scale factor. Aperture filling is part of the measurement equation, not an electronic gain setting.

Tutorial · Infer emissivity using two controlled backgrounds

An opaque diffuse sample at fixed temperature is observed at 10 micrometres against two isotropic backgrounds with calibrated spectral radiances 8.00 and 16.00 W per square metre per steradian per micrometre. The resulting outgoing readings are 25.00 and 27.00 in the same units. Infer emissivity and the sample's blackbody-equivalent radiance. Estimate the significance of independent reading errors of 0.10 units.

▼ **Separate the reflected and emitted contributions**

The increased outgoing radiance is 2.00 for an 8.00 increase of background. Thus reflectivity is 0.250 and emissivity 0.750. The first measurement gives intrinsic blackbody radiance $[25.00 - 0.25(8.00)]/0.75 = 30.67$. The second returns the same value. With 0.10 independent uncertainty on each outgoing reading and exact backgrounds, the slope uncertainty is $\sqrt{2}(0.10)/8 = 0.0177$, as is the absolute emissivity uncertainty. Larger background separation can improve this slope estimate if sample temperature stays fixed.

If the sample transmits radiation, the residual fraction does not all correspond to reflection. A controlled change behind the sample helps detect that additional path. A directional specular reflector exposed to a warm enclosure also does not necessarily see the same angular average used in the diffuse model. Without these checks, a formally precise fitted emissivity can describe the experimental surroundings instead of the material.

Tutorial · Design a numerical spectral-integration check

An engineer tabulates spectral radiance per micrometre on a wavelength mesh. Another rewrites the detector model on an equally spaced frequency mesh, obtaining a signal that differs by several percent after independent mesh refinement. Explain what must transform, what stays dimensionless and how to compare a photon-counting channel with a power channel.

▼ Transform the complete measurement integral

The differential power $B_\lambda \, d\lambda$ must equal $B_\nu \, |d\nu|$, with $|d\lambda/d\nu| = c/\nu^2$. A dimensionless transmission curve becomes the same physical transmission at $\lambda = c/\nu$; it acquires no additional Jacobian. The integration bounds reverse when increasing wavelength is replaced by increasing frequency. A response specified per unit wavelength, or a calibrated gain integrating a particular finite bin, must instead be interpreted from how it was defined.

Evaluate a unit-transmission test using a very narrow band, then the full 8–12 micrometre band whose radiance values are given above. Refine each grid independently, treating sharp filter edges consistently and converting per-micrometre densities before using metres. For a photon counter divide each spectral power contribution by photon energy before summation. Replacing that energy by a single central-wavelength value is an approximation that should be checked against the full band and source temperature. Close agreement between two different meshes using the same incorrect density is not a valid units check.

NIST radiance-temperature calibrations · spectral radiance, sources and uncertainty

Electromechanics, optics and measurement · 12. Moving-mirror Doppler metrology

A moving mirror changes optical phase, shifts the reflected frequency and exchanges energy and momentum with the light. In many displacement measurements the classical phase rule is extraordinarily accurate. Derive its exact inertial-frame counterpart, identify which time a detector samples and find the practical uncertainties that usually matter well before a relativistic correction does.

Match phases on a boundary that moves

In a vacuum laboratory let an incident plane wave travel along $+x$ with phase $kx - \omega t$. A perfectly reflecting plane mirror has position $X(t)$, receding for positive velocity. The reflected wave travels along $-x$ with phase $-k_r x - \omega_r t + \phi_0$. At the moving conducting surface the tangential electric fields must cancel in the mirror rest frame. For a constant mirror speed this requires equality of the two phase rates evaluated on its trajectory. The phase reversal at reflection supplies a constant offset and disappears on differentiation.

$$kv - \omega = -k_r v - \omega_r,$$

$$\frac{\omega_r}{\omega} = \frac{1-\beta}{1+\beta},$$

$$\beta = v/c.$$

Here $k = \omega/c$ and $k_r = \omega_r/c$ are positive magnitudes. A receding mirror redshifts the reflected light and an approaching mirror blueshifts it. Another derivation boosts into the mirror frame, where incident and reflected frequencies are equal, reverses the wavevector and boosts back. This yields a product of two Doppler factors, one for each transformation. Multiplying the first factor by its inverse without reversing the wavevector would predict no shift.

Define the signed beat $f_b = f_i - f_r$. The exact and first-order relations with vacuum wavelength λ_0 are

$$f_b = \frac{2v}{\lambda_0(1+\beta)},$$

$$f_b \simeq \frac{2v}{\lambda_0},$$

$$v = \frac{\lambda_0 f_b}{2 - \lambda_0 f_b/c}.$$

At 633 nm and $v = 2.00$ mm/s the beat is 6319.12 Hz. The relative correction to the first-order expression is only 6.67×10^{-12}. Frequency scale, geometry and air refractive index are ordinarily much larger uncertainties in an uncorrected bench experiment. The same exact expression makes a genuinely different prediction at $\beta = 0.10$: $f_r/f_i = 0.81818$, whereas the first-order frequency ratio is 0.80. This latter value is a theoretical benchmark for the formula rather than a required moving-bench experiment.

Connect reflection time with the detector's interference phase

Put an ideal stationary beamsplitter and detector at the origin, and let the source continuously illuminate the mirror along the measurement arm. A wave arriving back at laboratory time t was reflected at t_r satisfying $t = t_r + X(t_r)/c$. It left the beamsplitter at $t_0 = t_r - X(t_r)/c$. Combining these gives

$$t - t_0 = 2X(t_r)/c.$$

With a stable source and a fixed reference arm, the accumulated return phase relative to the reference is proportional to $2kX(t_r)$ up to a constant and an electronic quadrature sign.

$$\phi(t) = \frac{4\pi X(t_r)}{\lambda_0} + \phi_{\mathrm{off}},$$

$$\frac{1}{2\pi}\frac{d\phi}{dt} = \frac{2v(t_r)}{\lambda_0[1 + v(t_r)/c]}.$$

The denominator comes from differentiating the arrival-time condition. It is not obtained from differentiating a geometrical arm length at detector time while keeping the reflection time fixed. For very slow motion and a modest arm one can replace the retarded trajectory by the present one to high accuracy. At higher bandwidth a fixed path delay can still matter even if the velocity correction is tiny. A one-metre mirror distance gives roughly 3.34 ns from reflection event to detection.

A sinusoidally driven mirror with amplitude 100 nm and ordinary mechanical frequency 1.0 kHz has phase amplitude $4\pi A/\lambda_0 = 1.99$ rad at 633 nm. The peak speed is 0.628 mm/s and peak signed beat about 1.99 kHz. The instantaneous beat changes sign at each mechanical turning point. A frequency counter returning a single unsigned average misses that directional information. Demodulating the phase or fitting the sinusoidal motion to both quadratures can recover position, speed and reversal times.

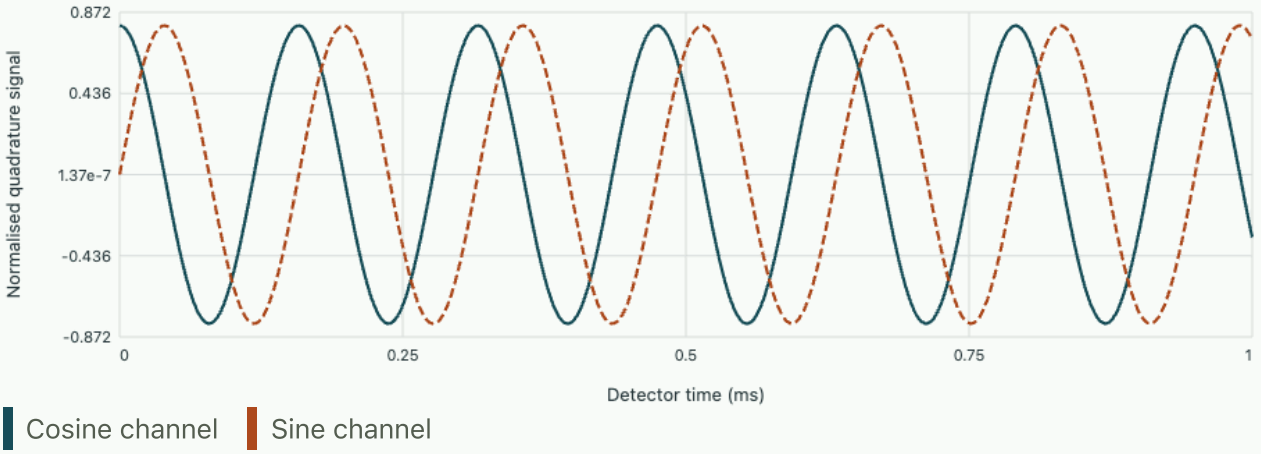
Infer displacement or signed speed without losing complete fringes

If stationary reference and returned fields overlap in polarisation, spectrum and space, their combined intensity contains a cross term proportional to $\cos\phi$. A single intensity channel cannot distinguish increasing from decreasing phase at an intensity extremum. Two calibrated quadrature outputs ideally read $X_s = V\cos\phi$, $Y_s = V\sin\phi$ after removal of offsets and gain normalisation. Recover wrapped phase with the two-argument arctangent. Unwrap successive phase differences only when the sampling interval and noise allow the branch to be identified.

Open the live experiment

Try the calculation

Mirror speed, positive receding (mm/s)	2
Vacuum wavelength (nm)	633
Fringe visibility	0.8



The signed beat is 6319.12 Hz. In 1 ms the mirror travels 2.00 µm and the unwrapped phase changes by 39.70 rad.

The signed velocity control reverses the direction traced by the quadratures. Hold its absolute value fixed and change sign. The cosine channel alone remains almost identical at laboratory speeds; the sine channel reverses the progression from its initial value. Reduce the visibility and keep in mind that the graph shows noise-free curves. With fixed electronic amplitude noise a smaller circle makes angular uncertainty larger, even though the exact underlying phase still advances at the same rate.

One full phase turn represents a physical mirror displacement of half the vacuum wavelength at normal incidence. A 20-fringe movement at 633 nm is 6.33 micrometres. If both endpoints retain phase residuals, add their signed difference to the complete turn count before converting to distance. A missed fringe biases displacement by 316.5 nm regardless of how accurately sub-fringe phases are fitted. For the default speed and wavelength the magnitude of frequency is 6.32 kHz. To unwrap a plain uniformly sampled phase with less than half a turn between adjacent samples one needs a sampling frequency above 12.64 kHz, and usually a margin for noise and acceleration.

Quadrature imbalance distorts the ideal circle into an ellipse shifted away from the voltage origin. Taking an arctangent of raw voltages then introduces a periodic position error. Record a slow multi-fringe calibration, fit channel offsets, scales and nonorthogonality, and test the recovered phase on held-out data with reversed motion. Calibration using only a tiny arc of the ellipse makes offset and scale poorly identifiable even if the residuals on that arc are small.

Close the energy and momentum balance at the mirror

One incident photon has energy hf_i and momentum $+hf_i/c$. On reflection it has energy hf_r and momentum $-hf_r/c$. The mirror acquires momentum $h(f_i + f_r)/c$ and energy $h(f_i - f_r)$. Their ratio is

exactly v when the prescribed constant-speed reflection formula applies. A recoil-free moving boundary can exchange mechanical work $v\,\Delta p$ while a drive maintains its velocity. For a freely recoiling low-mass mirror illuminated by a small number of photons the fixed-velocity approximation must be replaced by energy and momentum conservation for the joint system.

For an incident beam with power P_i crossing a stationary plane in the lab, the number of photons encountering the receding surface per laboratory time is $P_i(1-\beta)/(hf_i)$. Multiply by transferred momentum per encounter to get the force on that moving surface. The result simplifies to $F = (2P_i/c)(1-\beta)/(1+\beta)$. At rest, a 10 mW perfectly reflecting beam gives 66.7 pN. For a 2.0 mm/s receding mirror at that power, the corresponding mechanical work rate is only 1.33×10^{-13} W to leading order.

A detector fixed upstream sees reflected-wave intensity reduced by the square of the reflected frequency ratio for an ideal steady plane wave. Its received power change compared with a fixed-plane incident flux is not in general equal to the instantaneous work on the mirror. The region of field between the fixed reference plane and the moving surface grows or shrinks and stores energy. Instead compare powers crossing the moving boundary with their proper encounter rates, or account for the change of field energy in a specified control volume. For a finite pulse the integrated photon count is unchanged under perfect reflection and the total energy loss per photon is the Doppler shift already derived.

Estimate realistic errors before pursuing small corrections

At normal incidence in vacuum the first-order speed estimate is $\lambda_0 f_b/2$. For a beam at angle θ to the translation direction, with an appropriately arranged reflection/return geometry interrogating mirror-normal motion, the usual small-speed sensitivity along the normal includes $\cos\theta$. A tilt of two degrees gives a relative cosine factor 0.99939, a 610 parts-per-million reduction. A plane mirror tilted relative to a fixed coaxial receiver can additionally send the reflected beam outside the collection path, so specify how oblique reflected and reference fields are recombined. Blindly interpreting reduced fringe visibility as a velocity correction conflates these two effects.

In a homogeneous stationary transparent medium a quasistatic phase measurement measures optical path nL. If n changes during a displacement measurement, the phase derivative includes both $n\dot{L}$ and $L\dot{n}$ when the simple uniform-medium approximation holds. A refractive-index change of one part in a million over a 0.5 m arm imitates about 0.50 micrometres of path-length change. A short differential arm, a refractive-index reference or a controlled environment can separate it from actual motion. The vacuum relativistic formula cannot simply absorb an arbitrary time-varying medium by replacing c everywhere with c/n.

For a phase-difference velocity estimate over a record of duration T, the displacement noise from two independent endpoint phase errors of standard deviation s_ϕ is $\lambda_0\sqrt{2}s_\phi/(4\pi)$. At 633 nm, $s_\phi = 0.010$ rad and $T = 2$ ms, the resulting velocity uncertainty is about 0.356 micrometres per second. Fitting a slope over many samples can improve white-noise precision, but laser drift, mechanical acceleration and missed branches require their own models. Report the time interval over which a constant velocity is being claimed.

Recover a small vibration amplitude from optical sidebands

A phase-sensitive receiver is not the only way to recover periodic motion. Let a mirror vibrate at angular frequency Ω about a fixed mean, with displacement $A\cos\Omega t$ and speed sufficiently small for a first-order optical path description. The reflected analytic signal relative to its carrier contains $e^{im\cos\Omega t}$ with modulation index $m = 4\pi A/\lambda_0$. Expand the exponential in its Fourier series. The power of sideband order j relative to the incident unmodulated carrier is $J_j(m)^2$, where J_j is a Bessel function. Energy conservation of the phase factor requires that all the sideband powers sum to one in this phase-only approximation.

For a 10.0 nm displacement amplitude at 633 nm the index is 0.1985. To leading small-index order, each first sideband carries $(m/2)^2 = 0.00985$ of the original carrier power. The central component loses approximately twice this amount. A receiver collecting only one sideband would infer the wrong vibration amplitude if it treated that signal as the sum of both. At larger amplitudes higher Bessel orders become important and a single sideband amplitude is no longer a monotone function of displacement. A multi-sideband fit or an independent coarse amplitude resolves the ambiguity.

With a balanced optical heterodyne reference offset far enough from zero frequency, both sidebands can be separated electronically from the carrier beat. Measuring their power ratio to the central beat cancels a common optical gain, but any electronic response slope between their frequencies needs calibration. Pure phase modulation predicts symmetric sideband powers. Asymmetry can arise from simultaneous amplitude modulation or an unequal frequency response rather than from a change in mirror speed. Measuring their complex phases distinguishes amplitude and phase modulation more directly.

A useful consistency check uses three independent measured quantities. Infer displacement amplitude from the first two sidebands, obtain the mechanical frequency from their separation and predict peak Doppler beat from the product of those two values. Then compare with a separate unwrapped quadrature phase trace. For a rigid sinusoidal mirror these should agree within timing and wavelength uncertainties. An interferometric signal containing two different moving surfaces can fit a few sideband powers yet fail the predicted time-dependent phase; changing the illuminated spot or one surface position tests that possibility. A spectrum and a time trace jointly constrain motion in a way that a single fringe visibility cannot.

Tutorial · Resolve a branch ambiguity in a vibrating mirror

A mirror returns a signed mean beat of 8.0 kHz while moving uniformly over a short time window. The vacuum wavelength is 632.8 nm. The detector samples quadrature at 10.0 kHz and a naive nearest-branch unwrap reports minus 2.0 kHz. Explain the discrepancy, infer both candidate speeds and design a measurement to select the correct one.

▼ Track the sampled phase and use an independent constraint

A true phase advance of 0.80 cycle per sample is indistinguishable from minus 0.20 cycle per sample when only consecutive wrapped angles are available. The corresponding physical speeds are plus 2.531 mm/s and minus 0.6328 mm/s. Both give exactly the same ideal complex samples at that sampling rate

under the first-order laboratory formula. A smaller least-squares residual on one constant-speed branch cannot resolve the ambiguity without another observation.

Increase the sampling frequency above 16 kHz with margin or use an independently known direction and bounded speed range that excludes alternative aliases. A second wavelength also changes the beat while preserving physical velocity and can resolve the branch if the two ambiguous frequency ladders do not coincide over the admissible range. For accelerated motion verify the condition at the maximum instantaneous beat rather than at its signed average over a complete vibration period.

Tutorial · Can this beat measurement verify a relativistic correction?

A translation standard operates at 1.00 m/s and an interferometer at 500 nm records a nominal 4.00 MHz beat. Estimate the difference between exact and first-order vacuum predictions. A timebase has a fractional uncertainty 10^{-8}, and wavelength is uncertain by 2×10^{-8}. Assess the test and state one way to reject an offset without confusing it with the desired correction.

▼ Calculate the small difference without subtracting nearly equal optical frequencies

Use the exact beat formula or expand it directly. The correction is approximately $-2v^2/(\lambda_0 c) = -0.0133$ Hz for the receding mirror. The stated timebase contributes 0.040 Hz and the wavelength scale contributes 0.080 Hz to a 4 MHz prediction, even before including the velocity standard or alignment. This setup does not resolve the correction with those absolute uncertainties. Subtracting two floating-point optical frequencies near hundreds of terahertz is also a poor numerical route to such a small difference.

Reverse the velocity with equal magnitude. The leading signed beat is odd in velocity, while its leading relativistic correction is even. Adding the two signed beats cancels the ideal leading term and leaves about minus 0.0267 Hz. This reversal also preserves a fixed electronic frequency offset, which would add twice to the same sum. Take stationary-mirror reference readings and vary the speed magnitude; the wanted even term scales as speed squared. Different arm alignments or asymmetric velocity calibration during reversal can still imitate part of that effect.

Cambridge Electromagnetism · wave propagation, radiation momentum and Lorentz transformations

Electromechanics, optics and measurement · 13. Weighing a small deposit with a piezoelectric resonator

Two nearby frequencies appear in the electrical response of one mechanical mode. Derive the series and parallel resonances, then use their different sensitivities to recognise added mass and stray capacitance.

Convert mechanical motion into a terminal admittance

A piezoelectric plate converts terminal voltage into force and velocity into current. With a linear reciprocal coupling coefficient n, $F = nV$ and $I_m = nv$. The mechanical velocity impedance is $b + i\omega m + k/(i\omega)$. Dividing it by n^2 maps it to a series branch with $R_m = b/n^2$, $L_m = m/n^2$ and $C_m = n^2/k$. The electrodes also have a static capacitance C_0 in parallel. Thus

$$Y(\omega) = i\omega C_0 + \left[R_m + i\omega L_m + \frac{1}{i\omega C_m} \right]^{-1}.$$

For negligible loss the series branch impedance vanishes at $\omega_s = 1/\sqrt{L_m C_m}$. Just above this frequency its susceptance is inductive and can cancel the parallel capacitance. Setting $Y = 0$ gives

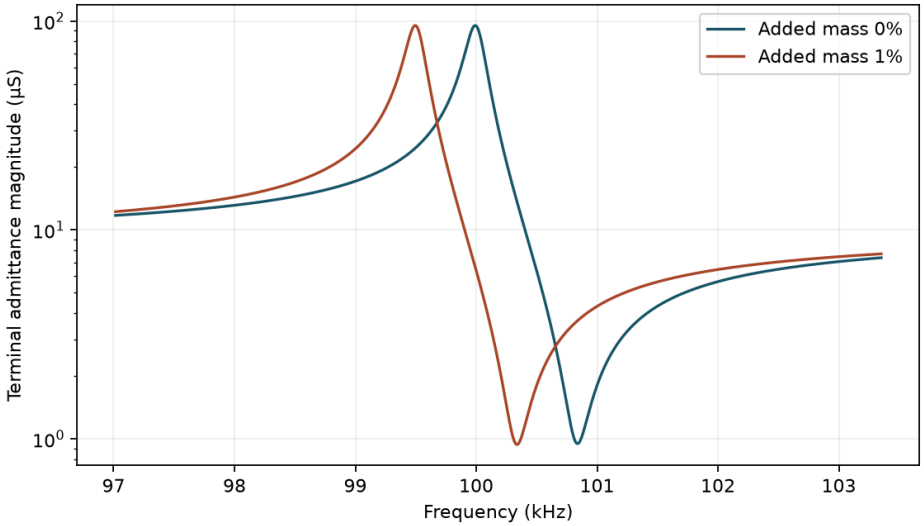
$$\omega_p = \omega_s \sqrt{1 + C_m/C_0}.$$

With finite damping a peak and a dip replace the singular limits. Their extrema need not coincide exactly with the lossless frequencies.

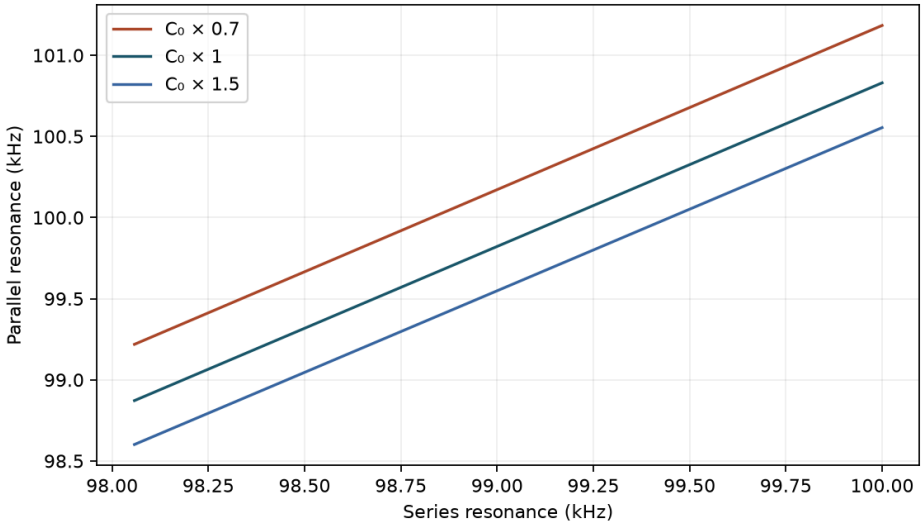
Isolate the perturbation

A thin rigid deposit that changes only the effective mass multiplies L_m by $1 + \eta$. Both ideal frequencies scale as $(1 + \eta)^{-1/2}$; their ratio is unchanged. Additional cable capacitance increases C_0, changing the ratio while leaving the mechanical series frequency fixed. At a load fraction of 0.01 the expected series shift is about minus 0.496 percent. A viscoelastic deposit can also dissipate energy and alter stiffness, so the two-frequency test includes a physical assumption about the coating. For $f_s = 100.000$ kHz, $C_m = 0.25$ pF and $C_0 = 15$ pF, the unloaded parallel frequency is 100.830 kHz. A 1% increase in effective mass puts the series resonance at 99.504 kHz. If the same shift came from extra cable capacitance alone, the series frequency would remain at 100.000 kHz. At quality factor 600 the approximate unloaded resonance linewidth is $f_s/Q = 167$ Hz. The predicted mass shift spans about three linewidths. The height of the peak by itself would confound damping and transduction.

Conversely, fit the two frequencies and infer the capacitance ratio as $(f_p/f_s)^2 - 1$. A small random error in f_p/f_s has a large relative effect on this estimate when the resonances nearly coincide. Zoom the frequency grid and compare the uncertainty from bin spacing with the shift under investigation.



A terminal sweep contains the mechanical resonance and the electrical cancellation. Follow both features before inferring added mass.



Changing cable capacitance displaces the parallel resonance at fixed series frequency. Mass loading moves both frequencies along a diagonal path.

Exercise · recover a capacitance ratio

A nearly lossless device gives $f_p/f_s = 1.01$. What ratio of motional to static capacitance is consistent with it?

▼ **Work through the solution**

Squaring and subtracting one gives $C_m/C_0 = 0.0201$. The difference between the squares matters when the shift is used for accurate capacitance inference.

Exercise · challenge a mass reading

After changing a fixture the dip moves but the series frequency remains within the experimental uncertainty. Does that support a pure mass increase?

▼ **Work through the solution**

A purely inertial change shifts both resonance frequencies. The observed pattern is consistent with an increased shunt capacitance if the dip moved toward the series peak. A sweep of the linewidth and an independent capacitance check would test the model.

Try it · Piezo load

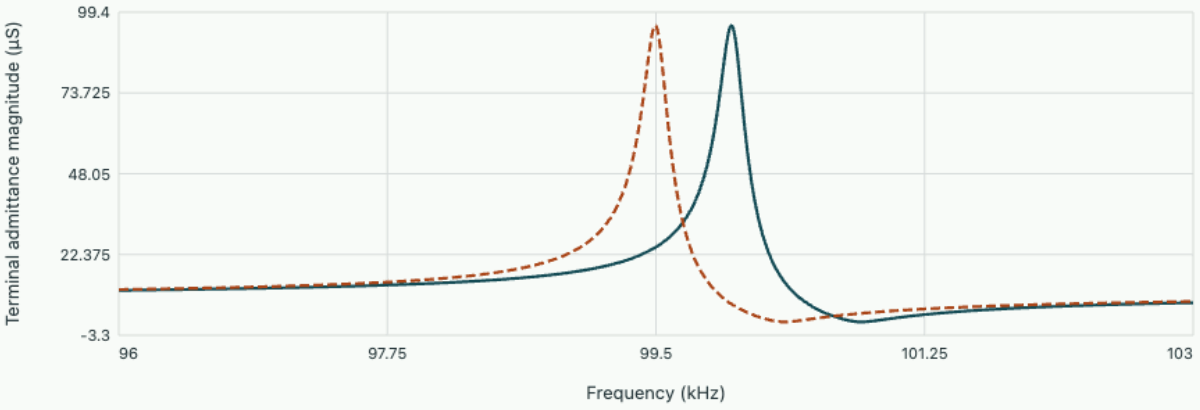
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Added mass fraction **0.01**

Parallel capacitance (pF) **15**



Loaded lossless series/parallel frequencies 99504 / 100330 Hz. Quality factor 600; motional capacitance 0.25 pF.

Download the corresponding editable notebook. The browser recalculates the reduced equations directly; the notebook exposes the full calculation and its checks.

Open editable notebook Download .ipynb

Numerical field laboratory

Turn a prediction into a computational experiment. Derive a reference, solve the field problem, challenge the model with a diagnostic and interpret the residual.

1. ## A converged field solver can still give the wrong voltage

   Verify a two-dimensional electrostatic calculation with two independent charge patterns. Use symmetry to distinguish a transposed output, a coarse mesh and a poorly converged linear solve.

2. ## Measure the capacitance of a continuously graded dielectric

   Find the exact solution of a variable-coefficient field equation, then estimate the same capacitance from charge, voltage drop and electrostatic energy.

3. ## Is a computed screening layer resolved?

   Use a measured decay length to select a mesh, derive the exact discrete penetration constant and decide whether a discrepancy belongs to the physical screening law or the numerical grid.

4. ## What can a mesh measure about a point charge?

   Distinguish conserved charge and measurable exterior fields from a numerical peak that diverges as the source becomes more localised.

5. ## Measure the phase speed of a numerical electromagnetic wave

   Start with a lossless Maxwell wave, build a second-order time update and compare its actual travel time with what the continuum wave equation predicts.

6. ## Move the reference plane of a wave measurement

   Recover a load reflection coefficient from a noncontact voltage measurement and explain why identical amplitude data can represent physically different terminations.

7. ## When a film is thinner than a skin depth

   Solve a finite conducting layer in complex notation and independently account for its losses using a surface impedance and a volume integral.

8. ## Separate diffraction from a finite observation window

   Recover the far-field angular spectrum of a smooth aperture, then isolate the errors from truncating and sampling the aperture field.

9. ## Estimate a cavity linewidth from a finite, biased record

   Connect the time and frequency descriptions of one damped mode, then diagnose why a weak detector background systematically biases the inferred quality factor.

10. ## A magnetic orbit that spirals outward in a perfect vacuum

Track a charged particle in a prescribed uniform field, derive a rotation update and separate numerical energy growth from a real magnetic force.

11. When differentiating a measured dipole invents radiated power

Infer dipole radiation from a sampled displacement and show quantitatively why position noise can dominate a seemingly well-resolved power estimate.

12. Choose a frequency that identifies a coil's stored and dissipated energy

Design a small-signal experiment to estimate resistance and inductance, propagate amplitude and phase uncertainty and decide when a distributed field model is needed.

13. Where should a numerical electrostatic domain end?

A finite mesh needs an outer boundary even when the physical field extends to infinity. Solve the artificial-boundary error exactly for a dipole and test whether enlarging the numerical domain or matching the exterior field is more effective.

Numerical field laboratory · 1. A converged field solver can still give the wrong voltage

Verify a two-dimensional electrostatic calculation with two independent charge patterns. Use symmetry to distinguish a transposed output, a coarse mesh and a poorly converged linear solve.

Choose a voltage before choosing a solver

A very long grounded tube has a square cross section $0 \le x, y \le L$, with $L = 1$ m. The contents are uniform along the axis. A model of its transverse charge density should predict what a voltage probe placed inside would measure. Three numerical statements that can sound similar are actually quite different: the discrete equations have a small residual, the grid solution approximates the continuous solution, and the physical source describes the tube. We can test the first two in a laboratory problem where the last is specified exactly.

Start by prescribing a voltage that satisfies all four walls. Introduce $s_m(x) = \sin(m\pi x/L)$ and choose $V_*(x, y) = [s_1(x) + g s_2(x)] s_1(y)$ volts. The parameter g is the signed amplitude of the second spatial pattern relative to the first. A probe at the centre sees exactly 1 V for every g, since $s_2(L/2) = 0$. A calculation checked at that point alone could entirely lose the second pattern and still pass. At $x = L/4$ on the centreline, $V_* = 1/\sqrt{2} + g$; at $3L/4$ it is $1/\sqrt{2} - g$. Their difference is $2g$ volts. First write down all three expected readings for $g = 0.5$.

Manufacture the source by differentiating the field

The physical problem uses Gauss's law and the electrostatic relation $\mathbf{E} = -\nabla V$. For uniform vacuum permittivity they yield

$$-\nabla^2 V = \rho/\epsilon_0.$$

A product $s_m(x) s_1(y)$ returns the factor $(m^2 + 1)\pi^2/L^2$ when acted on by the negative Laplacian. Applying this separately to the two patterns gives

$$\rho/\epsilon_0 = (\pi/L)^2 [2 s_1(x) + 5 g s_2(x)] s_1(y) \ \mathrm{V}.$$

The right side has dimensions V/m² after including the factor L^{-2}; multiplication by permittivity gives C/m³. When g becomes large enough, part of the imposed density changes sign. That is permissible for a manufactured source. The calculation represents a solution of the stated Poisson problem. Judging whether a particular apparatus could maintain that charge requires its own charge-transport model.

Uniqueness gives the reference solution special status. If there were a different voltage for the same walls and charge, their difference w would satisfy Laplace's equation with zero boundary value. Integrate $w(-\nabla^2 w)$ over the square. Integration by parts leaves $\int |\nabla w|^2$, because w is zero on the boundary. The integral must vanish; w is constant and its boundary makes that constant zero. This is why agreement with the manufactured voltage is a test of the solver and its input, rather than a choice between two physical solutions.

Background · How can the same source test a second boundary condition?

A smooth manufactured solution that vanishes on every edge tests a homogeneous Dirichlet solve. A code that incorrectly sets every requested boundary voltage to zero may still pass. Add a harmonic lift to the known field on a square of width L. For a constant permittivity the charge depends on the Laplacian, so the exact source is unchanged.

$$\widetilde{V}(x,y) = V(x,y) + V_0\Big(1 + \frac{x}{L}\Big),$$

$$-\epsilon\nabla^2\widetilde{V} = -\epsilon\nabla^2 V.$$

Prescribe the lifted values on all four sides and solve again using the same charge samples. The difference of the computed solutions should be the linear lift, which a centred second difference represents exactly on a uniform grid. A discrepancy points first to boundary-value handling or to the way known values enter the right-hand side. This comparison isolates that error from the truncation error on the original sine modes.

Solve the discrete problem on paper as a second reference

Let $h = L/n$. The standard negative Laplacian has centre coefficient $4/h^2$ and four neighbouring coefficients $-1/h^2$. Substitute the samples of $s_m(x)s_1(y)$. In the x direction the sum of forward and backward neighbours is twice the centre value times $\cos(m\pi/n)$. Collect the corresponding y term. The resulting positive grid eigenvalue is

$$\lambda_{h,m} = \frac{4}{h^2}\Big[\sin^2\frac{m\pi}{2n} + \sin^2\frac{\pi}{2n}\Big].$$

The continuum eigenvalue is $\lambda_m = (m^2+1)(\pi/L)^2$. If the source is the continuous source sampled at the nodes, the discrete voltage coefficient is multiplied by $\lambda_m/\lambda_{h,m}$. Since $\sin z < z$ for positive z, each factor exceeds one on a resolving grid. The faster spatial oscillation has the larger relative bias. The resulting total pointwise error can still change sign because the spatial modes themselves have signs.

This provides three independent calculations: a continuous closed-form field, an exact solution of the chosen difference equations and a numerical linear solve. Agreement of the latter two isolates assembly, indexing and algebraic errors. Comparing either with the continuous solution measures spatial discretisation error. Tightening an iterative tolerance indefinitely cannot remove a difference between $\lambda_{h,m}$ and λ_m.

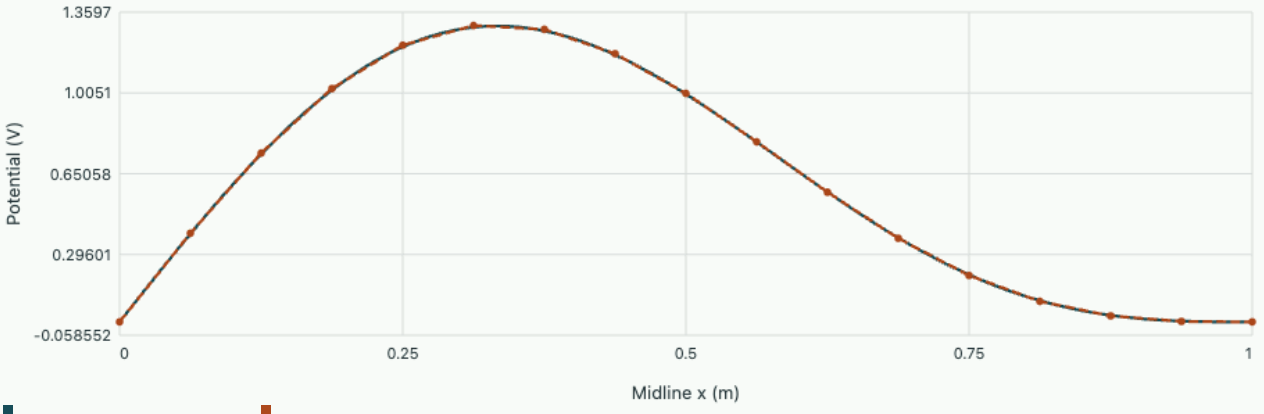
Run the full field and compare genuinely different observations

Open the live experiment

Try the calculation

Cells per side **16**

Signed second-mode amplitude **0.5**



Continuous field Solved grid field

Midline maximum error 0.0077708 V. Quarter-point difference 1.011 V. 2 CG step(s), relative residual 1.45e-13.

Download the OneSimulation model

In your OneSimulation environment, run the model with these inputs.

```
python field-laboratory.py modes --cells 16 --contrast 0.5
```

The browser constructs the right-hand side over the two-dimensional interior and applies conjugate gradients to the five-point operator from an all-zero initial field. The two-mode right-hand side belongs to a two-dimensional invariant subspace, so exact-arithmetic convergence takes at most two iterations. In the special case $g = 0$, only one mode is excited. This unusually small iteration count is a property of the test and says very little about a rough source on the same grid.

Download the model and run the displayed case. Its CSV contains the midline $y = L/2$, ordered by increasing x. For $n = 16,\ g = 2$ the value at $x = 0.1875$ m is about 2.42542 V on the grid and 2.40333 V analytically. If the centre is correct but the 0.1875 m reading is about 0.55736 V, check the orientation of the sampled data array. Keeping x fixed at its midpoint instead of y happens to annihilate the entire second mode.

A convergence report a second person can reproduce

Set $g = 0.5$. Record the centre voltage, the quarter-point voltage difference, the largest nodal error and the relative algebraic residual for $n = 8, 16, 32$. Use even grids divisible by four if reading exact quarter points. Estimate an order from two successive maximum errors using $p = \log(E_n/E_{2n})/\log 2$. For this smooth problem the leading difference between $\sin^2 z$ and z^2 is proportional to z^4, giving a second-order spatial error at fixed physical modes. If g or the physical source profile changes during refinement, the ratio no longer tests that claim.

An experiment that quotes only the centre voltage has poor sensitivity to g. An experiment quoting only the algebraic residual may perfectly solve the wrong mesh problem. An experiment that compares each plot to itself after smoothing lacks an independent reference. The three diagnostics above address these difficulties with quantities small enough to check by hand.

Extend the verification to a genuinely three-dimensional field

Ground all six faces of a cube with side length L=1 m. A potential that varies in every coordinate is

$$V_*(x,y,z) = A\sin(\pi x/L)\sin(\pi y/L)\sin(\pi z/L),$$

$$\rho = 3\epsilon(\pi/L)^2 V_*.$$

Each second derivative contributes −(π/L)² times the potential. Adding the three derivatives gives the source above; every boundary vanishes because one sine is zero. This is a manufactured source for checking a numerical method. In the supplied calculation ε=1 F/m and A=1 V; scaling both charge and permittivity together leaves the voltage unchanged.

On a uniform nodal grid, the seven-point negative Laplacian has centre coefficient 6/h² and six neighbour coefficients −1/h². The same sine addition identity used in two dimensions gives

$$\lambda_h = \frac{12}{h^2}\sin^2\!\left(\frac{\pi h}{2L}\right),$$

$$V_h = \frac{3(\pi/L)^2}{\lambda_h} V_*.$$

The browser evaluates this discrete eigenmode directly. The spatial view places a coloured slice inside the cube; its height is the physical z coordinate. Moving that slice reveals the interior field while the companion line plot compares the grid values with the exact three-dimensional solution. The numerical potential is slightly too large because λh is smaller than its continuum limit.

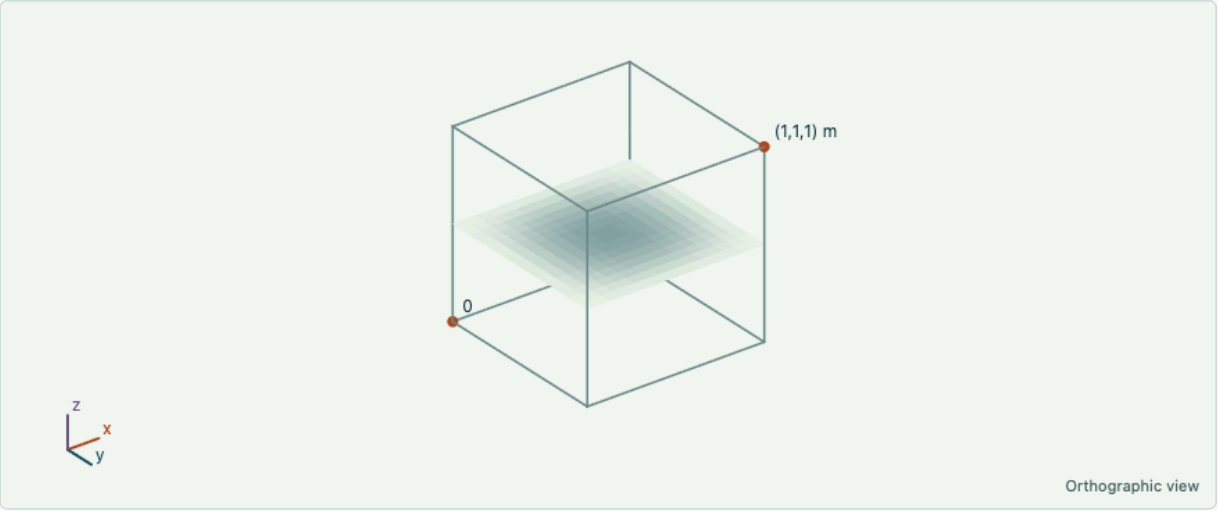
Try it · Rotate a 3D field, move its slice and refine its mesh

Move the plane from the centre to a grounded face and predict the change. Then compare 8 and 16 cells per direction: the maximum nodal error should decrease by approximately four. Doubling A doubles the absolute error while leaving the relative error almost unchanged. Rotation changes the viewpoint only.

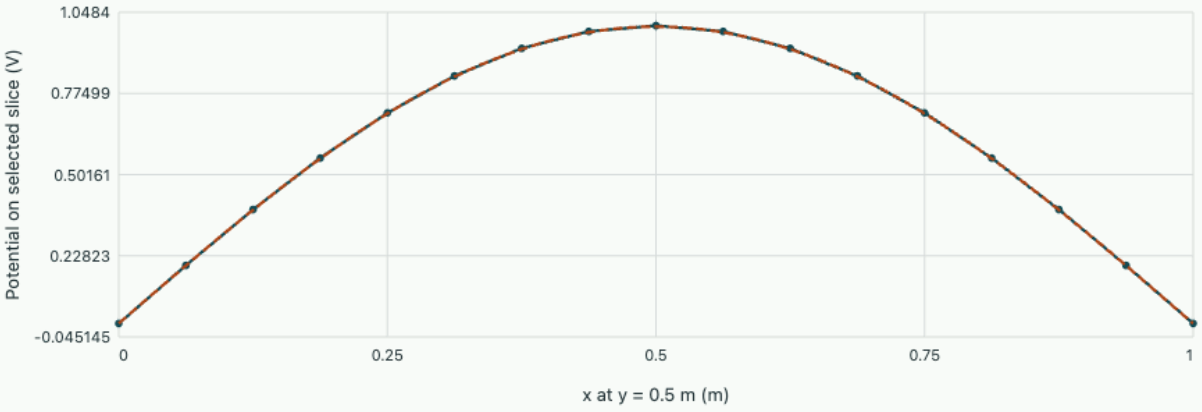
Open the live 3D experiment

Try the calculation

Cells per direction	16
Source amplitude (V)	1
Slice position (% of depth)	50



A slice through the three-dimensional potential, at z = 0.500 m. Colour is potential: pale = 0 V, dark green = 2.03 V; slice values range from 0 to 1.003 V. Grid-face colours average their four nodal values. Axes: x = position (m); y = position (m); z = position (m).



Seven-point FDM · nodal slice Exact 3D solution · same slice

16³ cells; selected grid plane z = 0.500 m. Maximum nodal error in the whole cube 3.219e-3 V (0.322%). All six boundary potentials are zero.

Download the OneSimulation model

In your OneSimulation environment, run the model with these inputs.

```
python poisson-cube-3d.py --cells 16 --amplitude 1 --slice 50
```

Download the 3D OneSimulation notebook · Download Python script

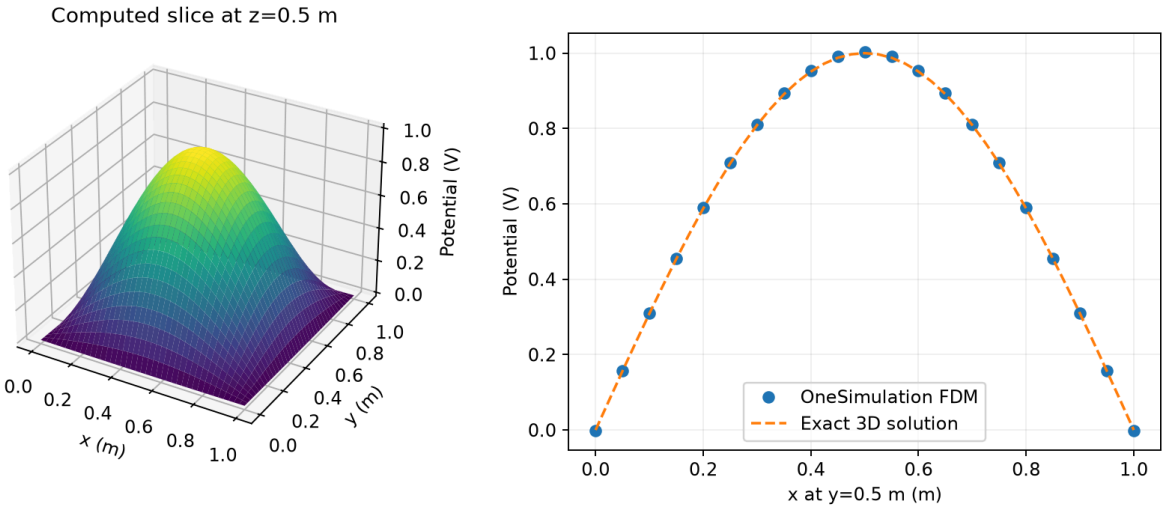
The notebook assembles and solves the full 3D OneSimulation finite-difference model. It reports an RMS error at cell centres; the browser reports a maximum error at nodes for its stated seven-point operator. These are different diagnostics. Compare each against its own continuous reference and repeat with a finer mesh.

```
"""Solve a grounded three-dimensional cube with OneSimulation and verify the field."""
import argparse
from dataclasses import dataclass
import numpy as np
import matplotlib.pyplot as plt
from onesim.cad.d3 import Box
from onesim.meshing import StructuredHexahedralMesher
from onesim.numerics import Numerics
from onesim.numerics.linear import SparseDirect
from onesim.numerics.fdm import FiniteDifference, CentralDifference, CentralLaplacian
from onesim.physics.electrostatics import FixedPotential, PoissonEquation
from onesim.simulation import Problem, Simulation

parser = argparse.ArgumentParser(description=__doc__)
parser.add_argument('--cells', type=int, default=16)
parser.add_argument('--amplitude', type=float, default=1.0, help='Reference potential amplitude in V')
parser.add_argument('--slice', type=float, default=50, help='Depth as a percentage of the cube')
args = parser.parse_args()
if not 8 <= args.cells <= 24 or not np.isfinite(args.amplitude) or args.amplitude <= 0 or not 0 <=
args.slice <= 100:
    parser.error('Use 8–24 cells, a finite positive amplitude, and a slice between 0 and 100 percent')

@dataclass(frozen=True)
class Permittivity:
    unit = 'F/m'
    def evaluate(self, **state):
        return 1.0

cube = Box(length=1, width=1, height=1)
exact = lambda x, y, z: args.amplitude * np.sin(np.pi*x) * np.sin(np.pi*y) * np.sin(np.pi*z)
field = PoissonEquation(domain=cube.volume, permittivity=Permittivity(),
                        charge_density=lambda x, y, z: 3*np.pi**2*exact(x,y,z))
problem = Problem(name='Three-dimensional manufactured potential', geometry=cube, physics=(field,),
                  boundary_conditions=(FixedPotential(physics=field, boundary=cube.boundaries.all,
value=0.0),))
mesh = StructuredHexahedralMesher(cells=(args.cells,)*3).generate(cube)
simulation = Simulation(problem=problem, mesh=mesh, numerics=Numerics(
    spatial=FiniteDifference(gradient=CentralDifference(), diffusion=CentralLaplacian()),
    linear_solver=SparseDirect()))
simulation.validate().require_valid()
result = simulation.solve_stationary()
potential = result.field(field.potential)
centres = mesh.coordinates[mesh.cells].mean(axis=1)
computed = potential.evaluate(centres)
reference = exact(*centres.T)
error = np.sqrt(np.mean((computed-reference)**2))
z = round(args.cells*args.slice/100)/args.cells
print(f'{args.cells}³ cells; source amplitude {args.amplitude:g} V; slice z={z:.3f} m')
print(f'Cell-centre RMS error against the exact 3D solution: {error:.6e} V')

x = np.linspace(0,1,41)
X,Y = np.meshgrid(x,x)
points = np.column_stack([X.ravel(),Y.ravel(),np.full(X.size,z)])
V = potential.evaluate(points).reshape(X.shape)
line = np.column_stack([x,np.full_like(x,.5),np.full_like(x,z)])
fig = plt.figure(figsize=(10,4),layout='constrained')
ax = fig.add_subplot(121,projection='3d')
```

```
surface = ax.plot_surface(X,Y,V,cmap='viridis',linewidth=0)
ax.set(xlabel='x (m)',ylabel='y (m)',zlabel='Potential (V)',title=f'Computed slice at z={z:g} m')
ax = fig.add_subplot(122)
ax.plot(x,potential.evaluate(line),'o',markevery=2,label='OneSimulation FDM')
ax.plot(x,exact(x,.5,z),'--',label='Exact 3D solution')
ax.set(xlabel='x at y=0.5 m (m)',ylabel='Potential (V)')
ax.grid(alpha=.2); ax.legend()
fig.savefig('poisson-cube-3d.png',dpi=170)
assert np.isfinite(error) and error/args.amplitude < .02
```

Measured output at the default settings:

```
16³ cells; source amplitude 1 V; slice z=0.500 m
Cell-centre RMS error against the exact 3D solution: 1.138076e-03 V
```



Actual OneSimulation output. Here the left panel plots potential as height above the selected x–y plane; the right panel compares the same field along y=0.5 m.

Tutorial

A colleague sets $g = -0.4$ and reports quarter-point voltages 1.107 V and 0.307 V on a very fine grid. Which change in the recorded setup would explain the observation?

▼ Work through the solution

For the stated sign, the expected pair in increasing x is 0.3071 V then 1.1071 V. Reversing the x axis interchanges the pair while preserving the centre and the symmetric mode. Reversing the sign of the second source mode produces the same pair. A field measurement with a labelled coordinate direction distinguishes these possibilities.

You prescribe the discrete eigenvalue as the source amplitude during every refinement. The returned solution agrees with your analytic curve at all nodes to machine precision. Which part of the verification has disappeared?

▼ **Work through the solution**

This tests the algebra of the discrete eigenmode and the linear solve. It changes the continuum source on every grid and cancels exactly the spatial discretisation error you intended to measure. Rerun with a wavelength and source amplitude specified in physical units independently of the grid.

A second probe has independent voltage noise of standard deviation 5 mV. Estimate the standard deviation of g inferred from the difference of the two quarter-point readings, if both probes have the same independent noise.

▼ **Work through the solution**

The difference has variance $2(0.005)^2$ V² and equals $2g$ volts. Dividing by two gives standard deviation $0.005/\sqrt{2} \simeq 0.00354$ in the dimensionless amplitude. A shared additive offset cancels; a gain error common to both readings does not.

Numerical field laboratory · 2. Measure the capacitance of a continuously graded dielectric

Find the exact solution of a variable-coefficient field equation, then estimate the same capacitance from charge, voltage drop and electrostatic energy.

Build the problem from what is continuous

Two very large parallel plates are separated by $L = 1$ m. Set their potentials to 0 and 1 V. Edge effects are negligible in the observation region. The relative permittivity of the material varies as $\epsilon_r(x) = 1 + gx/L$, where $g \geq 0$. There is no free volume charge. A common first guess is a linear potential between the plates, because the boundary voltages impose an average field. Test that guess using the local field law.

For a thin pillbox whose faces are perpendicular to x, zero enclosed free charge means equal normal displacement at both faces. Let $D_* > 0$ denote the magnitude of the constant displacement. Since potential increases with x, the electric field points towards negative x and $D_x = -D_*$. Hence $V' = D_*/[\epsilon_0(1 + gx/L)]$. Regions of low permittivity must sustain a larger electric field and account for a larger voltage drop per metre. We have a local constraint before solving a single matrix equation.

Obtain the voltage and both expressions for capacitance

Integrate V' from left to right and impose $V(L) - V(0) = \Delta V$. For $g > 0$ this gives

$$\frac{C}{A} = \frac{D_*}{\Delta V} = \frac{\epsilon_0}{L}\frac{g}{\log(1+g)},$$

$$\frac{V(x)}{\Delta V} = \frac{\log(1+gx/L)}{\log(1+g)}.$$

As g approaches zero, both expressions have removable singularities. Since $\log(1+g) = g - g^2/2 + O(g^3)$, the effective relative permittivity approaches $1 + g/2 + O(g^2)$. An arithmetic average can therefore look convincing at very small gradients despite having the wrong general derivation. For computation near zero, use a logarithm specifically evaluated as $\log(1+g)$ with a small-argument routine such as `log1p`.

Check the result through stored energy. Each volume element contributes $\epsilon E^2/2$, so

$$U/A = (D_*^2/2)\int_0^L dx/\epsilon(x).$$

Substituting the voltage relation gives

$$U/A = D_*\Delta V/2.$$

The same capacitance follows from $C/A = 2U/(A\Delta V^2)$. This equality requires a linear, lossless dielectric and a reversible charging path. A material whose permittivity changes with field needs the corresponding integral over the constitutive curve.

Work one case numerically before assembling the grid

At $g = 3$, the material goes from relative permittivity 1 to 4. The effective value is $3/\log 4 \simeq 2.1640$, smaller than the arithmetic mean 2.5. The midplane voltage is $\log 2.5/\log 4 \simeq 0.6610$ V, not 0.5 V. The electric field at the left is four times as large as at the right. A measurement of the midpoint voltage alone is enough to reject a uniform-field model if the plates and material are well characterised.

As a dimensional check, $C/A \simeq 19.16$ pF/m² for $L = 1$ m. For a geometrically similar layer one millimetre thick the same effective permittivity gives 19.16 nF/m². The voltage profile versus x/L stays fixed when g and the plate voltages stay fixed, while the field and free surface charge increase in inverse proportion to separation.

Assemble by balancing fluxes before dividing by a coefficient

Divide the gap into n cells. Across cell i approximate the displacement magnitude by $D_i = \epsilon_{i+1/2}(V_{i+1} - V_i)/h$. Enforcing $D_i = D_{i-1}$ at each internal node gives a tridiagonal linear system. An independent route views each cell as a capacitor with inverse capacitance per area $h/\epsilon_{i+1/2}$. These series contributions add, giving $D_*^{(h)} = \Delta V/\sum h/\epsilon_{i+1/2}$. Cumulative sums of the same drops then recover the entire grid potential. Both routes implement identical physics and the same discrete constitutive assumption.

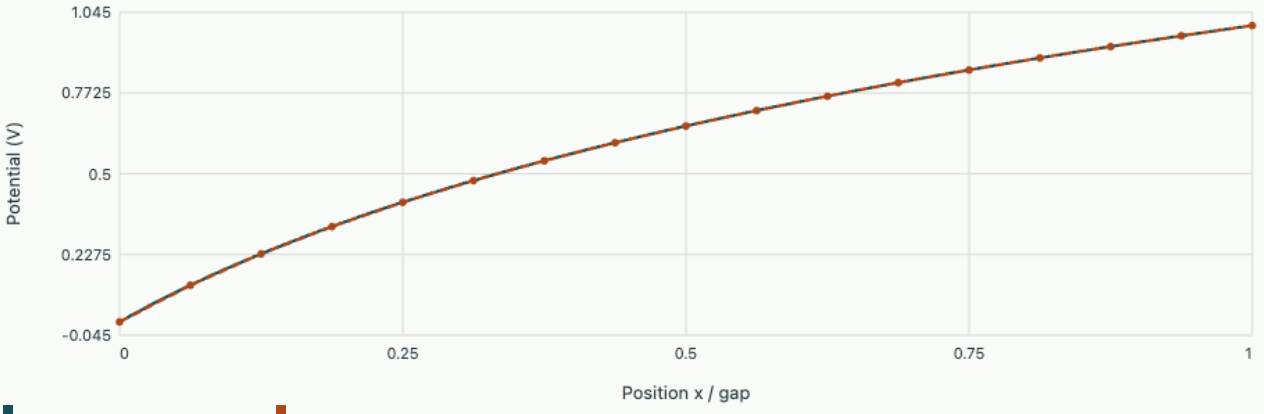
The browser uses this cumulative series construction. The OneSimulation input instead supplies the spatial coefficient $1 + gx/L$, zero source and the two exterior potential values to the coefficient-form PDE solver. Their agreement checks assembly against a distinct algorithm. In this example the integral of inverse permittivity also tells us what the midpoint discretisation approximates. Since $1/(1 + gx/L)$ is convex, midpoint quadrature underestimates its integral and therefore overestimates capacitance. This predicts the bias before looking at a convergence table.

Open the live experiment

Try the calculation

Cells across gap 16

Permittivity increase g 3



Analytical voltage Flux balance

Effective relative permittivity 2.1662 on the grid and 2.164 analytically. Midpoint voltage 0.66073 V.

Download the OneSimulation model

In your OneSimulation environment, run the model with these inputs.

```
python field-laboratory.py capacitor --cells 16 --contrast 3
```

Compare a local field measurement with an integral quantity

In the experiment, plot V across the dielectric and record the effective relative permittivity. At $g = 2$, $n = 16$, the computed potential at $x = 0.1875$ m is about 0.28974 V; the analytical reference is about 0.28987 V. The potential there is low while the computed capacitance is high. A voltage error at an arbitrary coordinate need not share the sign of the error in an integral quantity.

Now estimate the left electric field from the first-cell drop and from the analytical derivative at $x = 0$. The first-cell drop divided by h is an average across that cell, representing its midpoint to second order for a smooth field. Interpreting it as a value exactly at the wall introduces an $O(h)$ offset. Label sample positions when comparing a derivative against an analytic curve. The grid solution can converge quadratically in voltage while a carelessly placed flux diagnostic appears first order.

Tutorial

The gradient is reversed so $\epsilon_r = 1 + g(1 - x/L)$, but the plates remain at the same voltages. Which earlier observations change?

▼ **Work through the solution**

The integral of inverse permittivity is invariant under $x \mapsto L - x$, so the capacitance and stored energy stay the same. The new potential obeys $V_{\mathrm{rev}}(x) = \Delta V - V(L - x)$. The largest field is now at the right wall, and the midpoint voltage at $g = 3$ falls from about 0.661 to 0.339 V.

A sensor measures an apparently linear voltage profile and C/A greater than the vacuum value. Is the inferred dielectric necessarily uniform?

▼ **Work through the solution**

Under the one-dimensional, source-free and linear assumptions, exactly constant nonzero field and constant normal displacement require constant permittivity. A sparse set of voltage probes may simply undersample a graded profile. Lateral variation, free volume charge or uncertain boundary voltages would break the inference.

Derive bounds on the effective permittivity if only its material range $\epsilon_{\min}$ to $\epsilon_{\max}$ is trusted.

▼ **Work through the solution**

Integrate $1/\epsilon_{\max} \le 1/\epsilon(x) \le 1/\epsilon_{\min}$, multiply by $1/L$, and invert the positive quantities. The effective value lies between these two limits. A simulation reporting an effective permittivity outside that interval in this passive one-dimensional model fails this check regardless of mesh resolution.

Numerical field laboratory · 3. Is a computed screening layer resolved?

Use a measured decay length to select a mesh, derive the exact discrete penetration constant and decide whether a discrepancy belongs to the physical screening law or the numerical grid.

Derive the screened field and state its range

A dilute sample adjacent to an electrode can often be approximated, for small applied potentials, by a relation $\rho(x) = -\epsilon V(x)/\lambda^2$. It says that an excess positive potential attracts net negative charge. The parameter λ is independently calibrated for the ambient material. Combining this relation with Poisson's equation yields

$$V'' = V/\lambda^2.$$

A plus sign is essential: the physical disturbance decays into the sample. Changing the sign would turn a decaying layer into an oscillatory boundary-value problem.

Set $V(0) = V_s$ and $V(L) = 0$. Scale the voltage by V_s, position by L, and write $g = L/\lambda$. The dimensionless problem is

$$-u'' + g^2 u = 0,\ u(0) = 1,\ u(1) = 0.$$

The program solves for unit boundary voltage as a normalisation. For a dilute monovalent electrolyte at room temperature the linearised microscopic model requires the physical $|V_s|$ to be small compared with $k_B T/e$, roughly 26 mV. A computed unit-amplitude solution can represent a small physical voltage by multiplication. Applying 1 physical volt to the linearised electrolyte would generally exceed this assumption.

Solve the finite slab before taking the half-space limit

The general solution contains e^{gx} and e^{-gx}. Impose the right voltage first, using $\sinh[g(1-x)]$, then divide by $\sinh g$ to set the left voltage. Thus $u = \sinh[g(1-x)]/\sinh g$. For $g = 0$ take the limit to get $u = 1 - x$. For a thick slab and fixed physical distance from the left, use $e^{-gx}(1 - e^{-2g(1-x)})/(1 - e^{-2g})$. Close to the right wall the subtracted exponential is needed to satisfy grounding. In the interior many screening lengths away, even a big relative error can correspond to a negligible physical voltage.

At $g = 4$, the midpoint voltage fraction is $1/(2\cosh 2) \simeq 0.1329$. The half-space estimate is

$$e^{-2} = 0.1353.$$

A sample only four screening lengths across is already close to the half-space answer at its middle, though its right edge still has exactly zero voltage rather than e^{-4}. Determine where your instrument can distinguish the profiles before treating one as a better physical model.

Obtain the grid behaviour without running a solver

With grid spacing $h = 1/n$, the interior stencil is

$$-u_{i-1} + (2 + g^2 h^2) u_i - u_{i+1} = 0.$$

Try a geometric progression $u_i = e^{-\gamma_h i h}$. Divide through by the trial value. The sum of its neighbours gives

$$2\cosh(\gamma_h h) = 2 + (g h)^2.$$

Since $\cosh(2z) = 1 + 2\sinh^2 z$, the effective decay constant is

$$\gamma_h = \frac{2}{h}\operatorname{asinh}(g h/2) = g - \frac{g^3 h^2}{24} + O(g^5 h^4).$$

Thus the numerical layer is slightly too broad for a coarse mesh. The relevant resolution is $g h = h/\lambda$ in consistent units, not the cell count alone. Keeping n fixed while increasing g makes a visually unchanged mesh physically coarser. The finite-slab grid reference follows by replacing g with γ_h in the hyperbolic-sine ratio at the nodes.

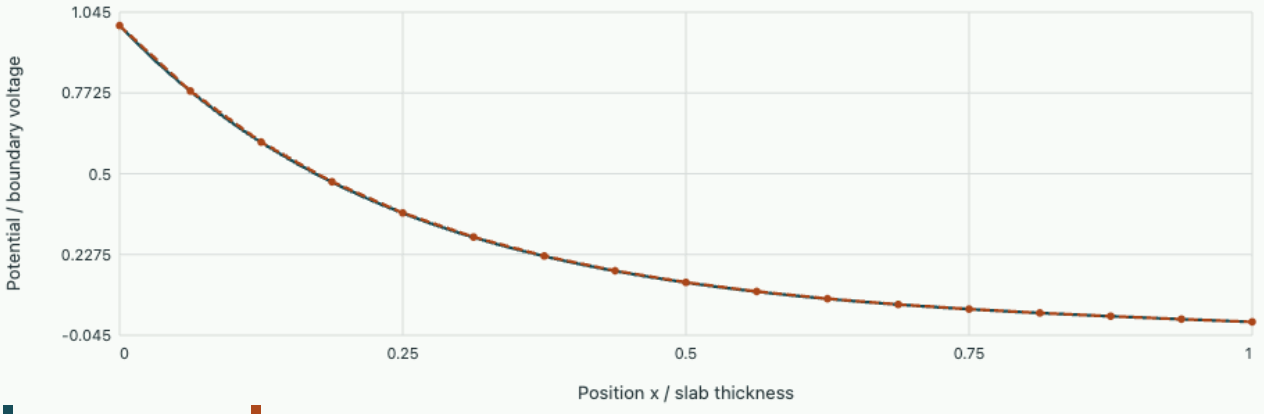
Separate the physical and computational experiments

Open the live experiment

Try the calculation

Cells across slab **16**

Gap / screening length **4**



Continuum slab Solved grid values

Midpoint 0.13357, exact 0.1329. Discrete decay constant × thickness = 3.9897.

Download the OneSimulation model

In your OneSimulation environment, run the model with these inputs.

```
python field-laboratory.py screen --cells 16 --contrast 4
```

The browser solves the tridiagonal boundary equations by forward elimination and backward substitution. The downloadable OneSimulation model supplies an absorption term $g^2 u$ and the same two Dirichlet values. For $g = 2,\ n = 16$, the potential fraction at $x = 0.1875$ is about 0.67311 on the grid and 0.67297 in the continuum. Reverse the grid spacing in the analytical expression for γ_h and independently predict whether doubling the number of cells raises or lowers the estimate there.

A typical measurement infers a length from the logarithmic slope,

$$1/\lambda_{\mathrm{fit}} = -d \log V / dx.$$

For two nearby positive voltages use the difference of their logarithms divided by probe separation. As V approaches the instrument floor, small absolute noise has a large effect on this ratio. At the grounded right wall the logarithm does not exist. Exclude that point and justify the usable region using both boundary physics and noise. Refining the field mesh cannot improve a voltage measurement below its resolution.

Tutorial

A grid has 16 cells and $g = 16$. Predict the fractional error in the effective decay rate, using the exact discrete expression rather than only the expansion.

▼ **Work through the solution**

Here $gh = 1$, so

$$\gamma_h / g = 2\,\mathrm{asinh}(0.5) \simeq 0.9624.$$

The numerical decay is slow by about 3.76%. The quadratic estimate gives 4.17%, reasonable but visibly imperfect at one cell per screening length.

For a desired relative decay-rate error below 0.1%, estimate the largest acceptable gh in the small-step regime.

▼ **Work through the solution**

The leading relative error is $(gh)^2/24$. Requiring it below 0.001 gives $gh < 0.155$, approximately 6.5 cells per screening length. This condition concerns the exponential rate. The derivative at the electrode may need a separate boundary stencil and its own check.

A calibration reports a longer decay length as electrode voltage increases, even after mesh refinement. Can the numerical dispersion formula alone explain this trend?

▼ **Work through the solution**

In the linear model the normalised profile is independent of the voltage amplitude. The discretisation error depends on g and the mesh. At fixed g, n it cannot create an amplitude-dependent normalised length. Recheck the linear-response regime, electrode boundary condition and instrument response before fitting a new λ.

Numerical field laboratory · 4. What can a mesh measure about a point charge?

Distinguish conserved charge and measurable exterior fields from a numerical peak that diverges as the source becomes more localised.

Ask three detectors different questions

Imagine three instruments looking at a small charged body in vacuum. One measures the voltage on a spherical probe outside it, one measures the integrated electric flux through a closed surface, and one estimates the largest electric field somewhere near the source. If the body is idealised as a point, which readings have finite limits? The distinction matters when a calculation displays a dramatic change in maximum field after a perfectly reasonable mesh refinement.

For a positive point charge q, the exterior potential $V = kq/r$ tends to infinity as the observation point approaches the source. Its radial field magnitude kq/r^2 also diverges. Gauss's law gives a different answer for the flux through any sphere of positive radius:

$$4\pi r^2 E = q/\epsilon_0.$$

A finite integral and an unbounded maximum are compatible. Knowing what the instrument measures determines which numerical output should be expected to converge.

Use a smooth source with a known answer

Replace the potential by $V_a = kq/\sqrt{r^2 + a^2}$, with $a > 0$ a distance. This is a particular physical regularisation. Differentiating before sampling gives

$$E_a = kqr/(r^2 + a^2)^{3/2}.$$

Divide the flux by q/ϵ_0. The putative enclosed-charge fraction becomes

$$F_a(r) = \frac{r^3}{(r^2 + a^2)^{3/2}}.$$

To recover the corresponding density, differentiate $qF_a(r)$ with respect to radius and divide by shell area. The terms involving r^4 cancel, leaving $\rho_a(r) = 3qa^2/[4\pi(r^2 + a^2)^{5/2}]$. This density is smooth, nonnegative and integrates over all space to q, since $F_a(0) = 0$ and $F_a(\infty) = 1$. It has an extended tail. The parameter a is a smoothing scale rather than an outer radius where charge abruptly stops.

At $r = a$ only $1/(2\sqrt{2})$, about 35.36%, of the charge is enclosed. At $r = 2a$ the fraction is about 71.55%. A common intuition that two smoothing distances already contain essentially everything fails for this distribution. Measure the probe radius in units of the source size before blaming the flux calculation.

Hold the observable fixed while refining

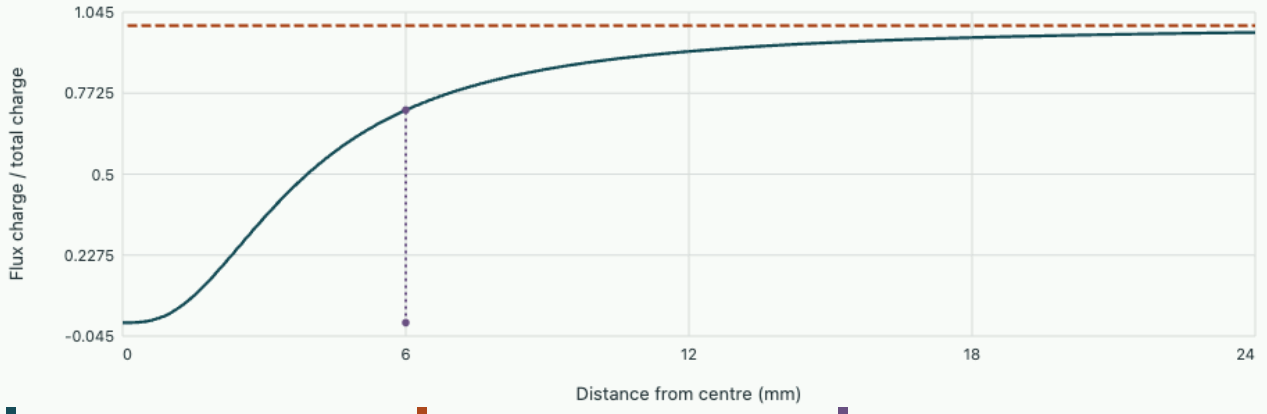
The exterior-field ratio relative to the true point charge is $[1+(a/r)^2]^{-3/2}$. At fixed physical r, expanding the bracket gives $1-\frac{3}{2}(a/r)^2+O((a/r)^4)$. Thus a probe far enough away sees a second-order regularisation error. Locating the maximum of E_a, however, requires setting $a^2-2r^2=0$, giving $r=a/\sqrt{2}$ and a peak proportional to $1/a^2$. Refine a and move the observer with the maximum: the reported peak grows even as a fixed exterior probe reading converges.

Open the live experiment

Try the calculation

Charge-cloud radius a (mm) 3

Probe radius r (mm) 6



The probe encloses 71.554% of the charge. Reducing a by half gives 91.308% at the same physical probe.

Keep the probe at 6 mm and reduce a from 3 to 1.5 mm. The flux rises from about 71.55% to 91.31% of the point-charge value. Now instead halve both the probe radius and a. The measured fraction stays the same. The two procedures have distinct limiting observations; writing “halve the mesh” without recording the source representation and probe definition would make the experiment irreproducible.

Use continuity as a numerical diagnostic

Suppose a code deposits a prescribed charge into nearby cells and then solves for a nodal potential. Summing deposited cell charges tests normalisation of the source. Summing outward face fluxes over a larger box tests a separate constraint on the computed field. A correct finite-volume implementation cancels interior face pairs. Interpolating a nodal field to a smooth plotting surface with an unrelated quadrature need not conserve the same discrete quantity. A box-flux test using the solver's own numerical face fluxes and a smooth-surface test using the reconstructed field answer complementary questions.

A second check concerns energy. For a true point charge the field-energy integral near the origin scales as $\int_0 r^2 (r^{-2})^2 dr$, which diverges. For a body of radius a this short-distance contribution typically scales like $1/a$. Comparing two source regularisations through their total self-energy without specifying a physical charge radius is therefore an ill-defined validation exercise. Interaction energy with a smooth distant field, in contrast, can have a well-defined point-particle limit.

Tutorial

How far from this smooth charge should a flux probe be placed to enclose 99%?

▼ Work through the solution

Solve $[1 + (a/r)^2]^{-3/2} = 0.99$. This gives

$$r/a = [0.99^{-2/3} - 1]^{-1/2} \simeq 12.2.$$

The answer depends on the chosen source shape. A uniform finite ball would have a different exterior behaviour.

A code uses $a = 2h$ and reports the field one cell away, $r = h$, on every refined mesh. What do you expect for its flux fraction and pointwise field?

▼ Work through the solution

The fixed ratio $a/r = 2$ gives $F = 5^{-3/2} \simeq 0.0894$ at every resolution. The field still scales as r^{-2} times that factor and grows like h^{-2}. The observer is moving towards the singular source.

If total deposited charge differs from q by 2% independently of mesh size, can adjusting a cure the exterior result?

▼ Work through the solution

At large radius the flux follows the total charge that was actually deposited. Changing width cannot repair an incorrect integral. Fix the source normalisation and cell-volume weights before interpreting a convergence slope.

Numerical field laboratory · 5. Measure the phase speed of a numerical electromagnetic wave

Start with a lossless Maxwell wave, build a second-order time update and compare its actual travel time with what the continuum wave equation predicts.

A pulse can arrive late without losing energy

In a uniform lossless one-dimensional material take electric field along y, magnetic field along z and propagation along x. Maxwell's curl laws reduce to

$$\epsilon E_t = -H_x,$$
$$\mu H_t = -E_x.$$

Differentiate the first in time and substitute the second to obtain

$$E_{tt} = c^2 E_{xx},$$
$$c = 1/\sqrt{\mu\epsilon}.$$

A real laboratory signal is a combination of spatial frequencies. Different errors in travel time for different components can distort its shape even when the scheme does not attenuate any resolved Fourier mode.

Use periodic boundaries so that reflection from an end cannot be confused with this effect. Scale distance by one wavelength and time by one exact period. The periodic field $E(x,t) = \sin[2\pi(x-t)]$ has unit amplitude and returns to its initial shape after every integer period. A probe comparing just its envelope would miss a discrepancy in phase.

Derive the discrete dispersion relation

Take points separated by h and times separated by Δt. Centred second differences give

$$u_j^{n+1} = 2u_j^n - u_j^{n-1} + C^2(u_{j+1}^n - 2u_j^n + u_{j-1}^n),$$

with $C = c\Delta t/h$. Insert the complex trial field $e^{i(j\theta - n\Omega)}$. The time pair contributes $2\cos\Omega - 2$, the spatial pair $2\cos\theta - 2$. The physical positive-frequency branch obeys

$$\sin(\Omega/2) = C\sin(\theta/2),$$
$$\frac{c_{\text{phase,grid}}}{c} = \frac{\Omega}{C\theta}.$$

If $0 < C \le 1$ a real Ω exists for every grid phase, and each Fourier mode has constant discrete amplitude in exact arithmetic when started on one travelling branch. For $C > 1$ sufficiently rapid spatial oscillations have no real frequency. One time factor then grows. Testing stability on a single very smooth wave can miss this unstable part of the supported spectrum.

For long waves expand $\sin$ and $\arcsin$:

$$\Omega = C\theta + C(C^2-1)\theta^3/24 + O(\theta^5).$$

The relative phase-speed error is $(C^2-1)\theta^2/24$ at leading order. Here increasing C towards one improves wave propagation. This special cancellation depends on a uniform one-dimensional grid, constant wave speed and the selected difference operators.

Initialise a clean numerical experiment

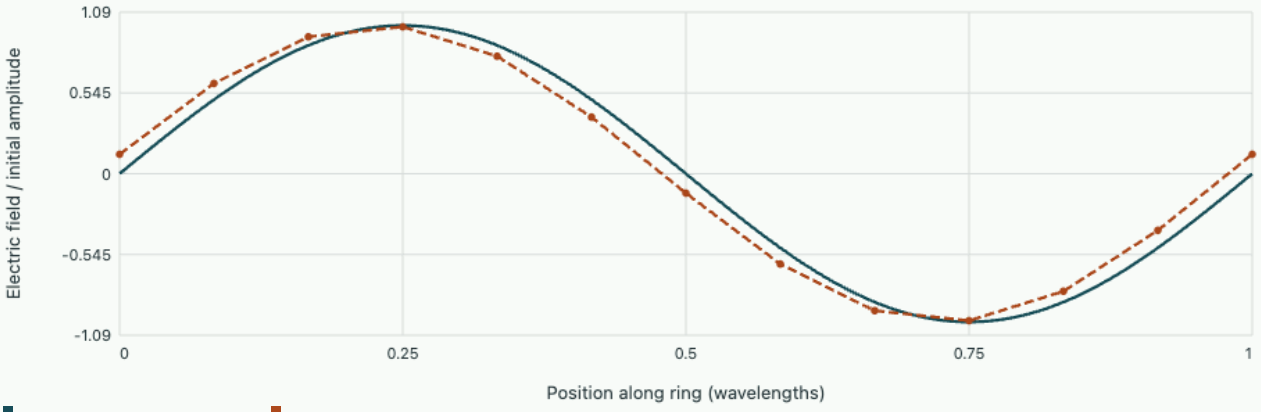
A second-order recurrence needs two time levels. Taking $u_j^0=\sin(j\theta)$ and $u_j^{-1}=\sin(j\theta+\Omega)$ excites a single right-going numerical mode. Using the exact continuum phase $C\theta$ at the second level instead introduces a small mixture of right- and left-travelling grid modes. That is a valid test of initialisation, but it mixes errors if the quantity of interest is solely the numerical dispersion relation. State in your notes how the second level was produced.

Open the live experiment

Try the calculation

Points per wavelength **12**

Courant number **0.8**



75 steps, 5 physical periods. Phase speed ratio 0.99582. Accumulated phase lag 7.5244°.

The browser actually advances the nodal difference equations. It then reports the dispersion prediction as an independent check. The time is chosen as an integer number of numerical steps close to five exact periods and is stated with the result. At twelve points per wavelength and $C=0.8$ there are fifteen steps per period. Both exact and discrete solutions have the same amplitude to roundoff. At the end of five periods examine how far their zero crossings have separated.

From a phase test to a pulse test

For a narrow packet, envelope motion is set by group velocity $d\omega/dk$. Differentiate the discrete relation at fixed h, Δt. The result is $v_g/c=\cos(\theta/2)/\sqrt{1-C^2\sin^2(\theta/2)}$ on the positive branch for resolved

nonendpoint modes. At $C = 0.8$ a mode close to the grid Nyquist limit has small group speed. A wave carrier and its envelope can show different speed errors. In a real dielectric wavelength dependence may also come from the material itself; a grid-refinement test with unchanged material law separates these contributions.

As an independent benchmark, set $C = 1$. Then $\Omega = \theta$ for $0 \le \theta \le \pi$, and the right-going recurrence translates the samples exactly by one grid cell at each time step. The lab should align both curves even on a surprisingly coarse grid for this particular smooth periodic mode. This checks the sign and indexing of propagation without implying exact accuracy for arbitrary geometry and material interfaces.

Tutorial

A signal must accumulate less than 1 degree of phase error over ten periods. Estimate a required points-per-wavelength count for $C = 0.5$ using the leading error.

▼ **Work through the solution**

The total continuum phase advance is 20π rad. Requiring $20\pi(1 - C^2)(2\pi/N)^2/24 < \pi/180$ gives $N^2 > 450\pi^2$, or $N \gtrsim 67$. Evaluate the exact dispersion relation to check the chosen integer.

Why is measuring only the field at its crest a poor detector of small phase lag?

▼ **Work through the solution**

The derivative of a sinusoid vanishes at its crest. A small phase shift changes that value only quadratically; at a zero crossing the change is linear. A phase fit using the entire ring is more robust than either single sample.

The numerical envelope changes after a wall is introduced. What additional tests would separate dispersion from reflection errors?

▼ **Work through the solution**

First validate travel time on a uniform periodic domain. Next use a single known interface with an analytical reflection coefficient. Move the observation plane and refine the mesh separately. Compare reflected phase and amplitude with an energy-flux balance so a boundary error cannot masquerade as bulk dispersion.

Numerical field laboratory · 6. Move the reference plane of a wave measurement

Recover a load reflection coefficient from a noncontact voltage measurement and explain why identical amplitude data can represent physically different terminations.

Specify which plane defines the load

A uniform lossless transmission line has real characteristic impedance $Z_0 = 50\ \Omega$. Drive it sinusoidally at one frequency. Choose peak phasors with time factor $e^{i\omega t}$ and measure distance d from the load towards the source. Incident and reflected waves move in opposite directions. Taking the load as phase origin, the ratio of the reflected to incident voltage at the load is

$$\Gamma_L = (Z_L - Z_0)/(Z_L + Z_0).$$

The ratio at a sensor plane a distance d away is $\Gamma_{\text{in}} = \Gamma_L e^{-2i\beta d}$. The factor two is the phase difference of two opposite travellers. Moving the sensor does not change a physical resistor.

For a real resistive load, Γ_L is real and ranges from –1 for a short to +1 for an open. A 75 Ω load gives +0.2 and a $100/3$ Ω load gives –0.2. These loads have the same reflected power fraction of four per cent. Distinguishing them requires phase information or positioning the sensor at more than one known location.

Derive the envelope and the ambiguity

The local voltage is the sum of incident and reflected phasors. Factor out the incident phase and take the magnitude:

$$\frac{|V(d)|^2}{|V^+|^2} = 1 + |\Gamma_L|^2 + 2\,\mathrm{Re}\{\Gamma_L e^{-2i\beta d}\}.$$

Its maxima and minima have amplitude ratios $1 + |\Gamma|$ and $1 - |\Gamma|$ for a passive load, giving $\mathrm{VSWR} = (1 + |\Gamma|)/(1 - |\Gamma|)$. The distance between successive maxima is $\lambda/2$, because the measured intensity involves a relative wave number 2β. A measured ratio 1.5 yields $|\Gamma| = 0.2$ but leaves its sign and, for general impedances, its phase undetermined. Extremum positions contain information that a ratio discards.

At one eighth of a wavelength from a 75 Ω load the phase rotation is $-\pi/2$, giving $\Gamma_{\text{in}} = -0.2i$. If a researcher plugs that complex coefficient directly into $Z = Z_0(1 + \Gamma)/(1 - \Gamma)$, the inferred impedance is about $46.15 - 19.23i\ \Omega$. That is the correct input impedance at the sensor plane. To infer the actual element multiply back by $e^{2i\beta d}$ before the inversion. The recovered load is then 75 Ω with zero imaginary part.

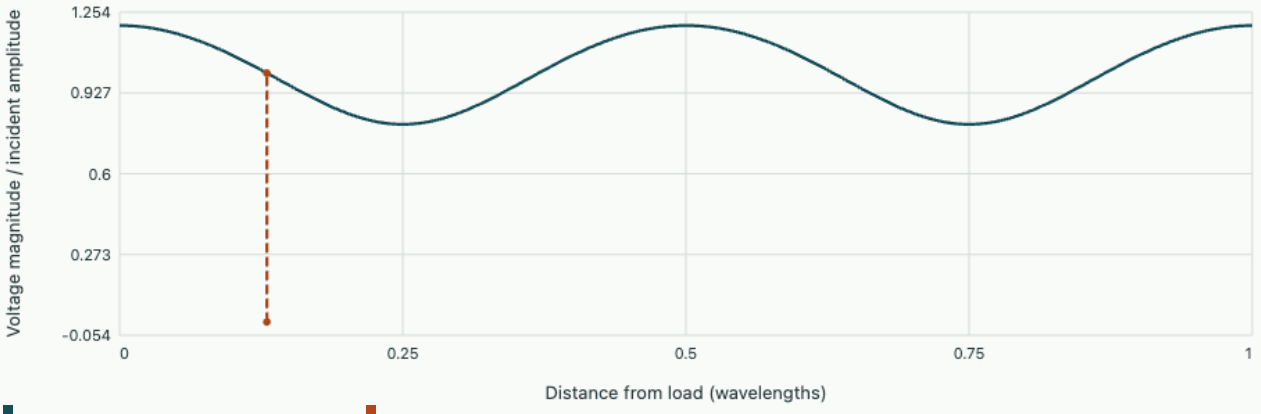
Put a probe between voltage extrema

Open the live experiment

Try the calculation

Load resistance (Ω) **75**

Probe plane from load (wavelengths) **0.13**



Standing-wave envelope Selected reference plane

Load Γ = 0.2. Probe-plane phase -93.6°. Voltage standing-wave ratio 1.5.

Set the load to 75 Ω and move the sensor from the load through a quarter wavelength. The envelope changes from 1.2 to 0.8 times the incident amplitude. At $\lambda/8$ it is $\sqrt{1+0.2^2}$. Centre your first trial on a voltage extremum and estimate how much the reading changes for a small positioning error. Repeat around the steepest part of the envelope. A high-slope position is sensitive to phase; a stationary amplitude extremum is useful when estimating the reflectance magnitude.

Use the matched case next. The field envelope becomes flat. The program reports that reflection phase is undefined. A zero complex amplitude has no observable phase angle. An instrument that prints a quickly varying phase for a nearly matched load can simply be reporting the angle of noise around the origin. Propagate uncertainty in real and imaginary parts before interpreting such a number.

Recover a complex reflection from several amplitude measurements

If $|V^+|$ is independently calibrated, sample squared voltage at $d = 0, \lambda/8, \lambda/4, 3\lambda/8$. Denote normalised intensity by I_d. Writing $\Gamma = x + iy$ gives $I_0 - I_{\lambda/4} = 4x$ and $I_{\lambda/8} - I_{3\lambda/8} = 4y$ with the convention above. Constant detector offsets cancel in these differences. A multiplicative gain error remains and must be calibrated. A mismatch in the assumed wavelength rotates or distorts the reconstructed coefficient rather than merely adding random noise.

For $\Gamma = 0.2$, the four intensities are 1.44, 1.04, 0.64 and 1.04. The reconstruction returns $x = (1.44 - 0.64)/4 = 0.2$, $y = 0$. The observation probes have provided both quadratures without a phase-sensitive receiver. For a lossy line the round-trip propagation factor also attenuates magnitude; the lossless envelope formula and phase-only de-embedding then require modification.

Tutorial

A reference plane is misplaced by $\lambda/100$. Estimate the resulting phase error of the reconstructed load.

▼ **Work through the solution**

The round-trip phase is $4\pi d/\lambda$. A placement error of $\lambda/100$ therefore rotates the result by 0.04π rad, or 7.2°. Its sign depends on whether distance was over- or underestimated.

Why is recovering a very large resistance from reflection poorly conditioned?

▼ **Work through the solution**

$$dZ/d\Gamma = 2Z_0/(1-\Gamma)^2.$$

Near the open-circuit value +1 this derivative becomes large. A small error in the complex coefficient can imply a large change in estimated resistance while reflected power barely changes.

A standing-wave ratio is measured with a detector displaying power rather than amplitude. Its largest-to-smallest reading is 2.25. What is the correct reflection magnitude?

▼ **Work through the solution**

Take the square root to get a voltage ratio of 1.5, then

$$|\Gamma| = (1.5-1)/(1.5+1) = 0.2.$$

Substituting 2.25 directly into the voltage formula would yield a biased result.

Numerical field laboratory · 7. When a film is thinner than a skin depth

Solve a finite conducting layer in complex notation and independently account for its losses using a surface impedance and a volume integral.

Which field does the back face support?

A wide conducting film carries an alternating sheet current in the x direction. The field varies only with depth z across thickness t. The surrounding configuration is arranged so the tangential magnetic field at the back face is negligible; at the front its peak phasor is H_s. This back condition describes one particular excitation. A film illuminated from both sides, a nearby magnetic backing or a transmitted wave in a substrate need different conditions. A skin-depth formula by itself does not determine a film impedance.

Use the time convention $e^{i\omega t}$, a real scalar conductivity σ, real permeability μ, and the good-conductor approximation $\sigma \gg \omega\epsilon$. Write $H_y = H(z)$ and $J_x = -H'(z)$. Faraday's law then gives

$$H'' = i\omega\mu\sigma H.$$

Define $\delta = \sqrt{2/(\omega\mu\sigma)}$ and $\gamma = (1+i)/\delta$, so $\gamma^2 = i\omega\mu\sigma$. H and its derivative can be complex even though the observable field is real.

Impose both surfaces before computing power

A solution satisfying the rear condition is

$$H(z) = H_s \sinh[\gamma(t-z)] / \sinh(\gamma t).$$

The electric field at the front is $-H'(0)/\sigma$. Dividing it by the applied field gives the sheet impedance,

$$Z_s = \frac{\gamma}{\sigma}\coth(\gamma t).$$

The thick limit gives $Z_s = (1+i)/(\sigma\delta)$, a resistive and reactive surface contribution of equal size. To find the thin limit expand $\coth w = 1/w + w/3 + O(w^3)$. Substitution gives $Z_s \simeq 1/(\sigma t) + i\omega\mu t/3$. The first term is the ordinary DC resistance per square. The second is the internal inductive reactance for the chosen one-sided magnetic boundary condition. Its factor one third can be checked by integrating the magnetic energy of a linear field profile through the thickness.

Take $\sigma = 4.0 \times 10^7$ S/m and a frequency giving $\delta = 0.2$ mm. The thick-film resistance scale is $1/(\sigma\delta) = 0.125$ mΩ per square. A $20\ \mu$m film has $t/\delta = 0.1$; its DC value is 1.25 mΩ per square, already ten times the thick-conductor scale. Treating a very thin film as a half-space can severely underestimate heat production for a specified sheet current.

Derive two independent loss measurements

The inward time-averaged Poynting flux per area equals $\frac{1}{2}\mathrm{Re}(E_s H_s^*) = |H_s|^2 \,\mathrm{Re}\, Z_s/2$ for peak phasors in this geometry. Independently the local cycle-averaged Joule power density is $|H'|^2/(2\sigma)$. Integrate it from front to back. A quick proof of equality avoids pages of hyperbolic algebra: integrate $(H'H^*)' = H''H^* + |H'|^2$ across the film. The term $H''H^* = i\omega\mu\sigma|H|^2$ is purely imaginary. Take real parts, insert $H(t) = 0$, and the boundary term at zero reproduces the surface expression. An unexpected minus sign here often means the outward normal has been used as though it pointed into the sheet.

The imaginary part of the same identity measures reactive magnetic storage. This separation helps diagnose a phase-sign mistake: conjugating every phasor reverses the sign assigned to reactive quantities for the opposite time convention, while average Joule heating stays positive. An output that predicts negative dissipation for a passive film warrants checking definitions before interpreting a material resonance.

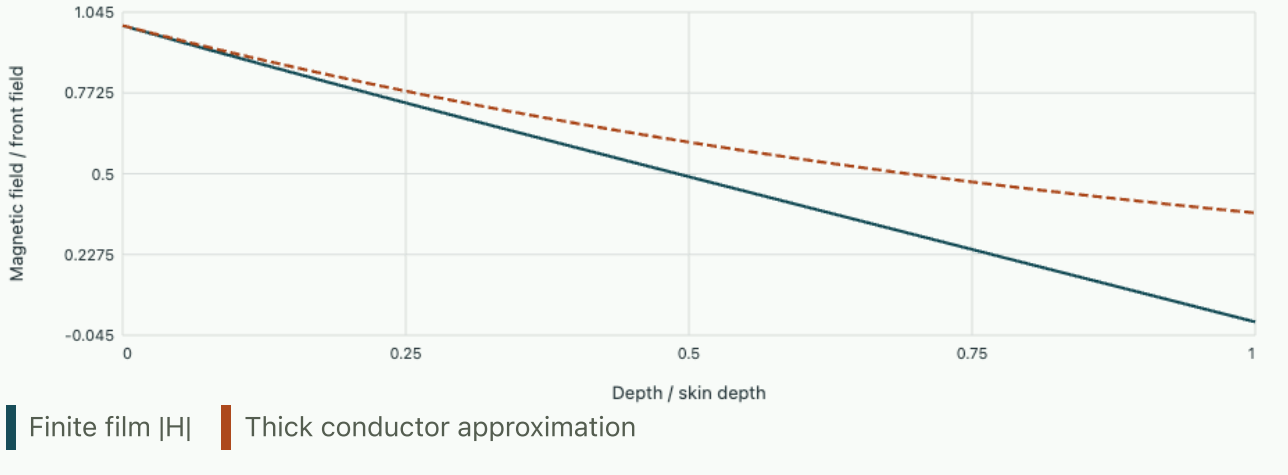
Inspect how the back boundary enters a measured field

Open the live experiment

Try the calculation

Thickness / skin depth **1**

Surface current (A/m, peak) **10**



Normalised surface impedance 1.0856 + i 0.65039. Mean power per area divided by the thick-conductor sheet resistance is 54.282 A²/m².

The finite-layer plot uses $|\sinh[(1+i)x]|^2 = \sinh^2 x + \sin^2 x$, with depth in units of δ. The half-space magnitude is simply $e^{-z/\delta}$. At the front both curves are one because the same field is imposed. At the back the finite film must reach zero. Increase t/δ until the front half of the two curves becomes experimentally indistinguishable. Determine whether thickness or the absolute amplitude changes the normalised profile. Then inspect the reported power, which depends quadratically on the actual surface current.

For $u = t/\delta$, the reported real part in thick-conductor resistance units is $[\sinh(2u) + \sin(2u)]/[\cosh(2u) - \cos(2u)]$. The imaginary part changes the plus to minus in the numerator. Substituting $u = 1$ gives a real part about 1.086. A nominal thickness of one skin depth is quite close to the thick-limit resistance for this boundary problem, even though the magnetic field is forced to zero at the back. Agreement on one integral measurement does not establish that the interior distribution has been measured correctly.

Tutorial

An instrument reports RMS surface fields instead of peak fields. What should change in the Poynting and Joule expressions?

▼ Work through the solution

The explicit factor one half is omitted when both phasors are RMS amplitudes. The impedance and the normalised field profile stay unchanged. If an experiment is wrong by a uniform factor two in power, check the amplitude convention before changing material parameters.

At fixed H_s, σ and very small positive frequency, what happens to film loss and to its stored magnetic energy?

▼ Work through the solution

In the thin limit the sheet current is nearly uniform, giving mean loss per area $|H_s|^2/(2\sigma t)$ for peak sinusoidal phasors. The peak field decreases approximately linearly from H_s to zero. Its cycle-mean magnetic energy per area is $\mu|H_s|^2 t/12$. The reactive voltage vanishes with frequency although a finite field can store energy.

Could the same Z_s formula be used for an insulating rear surface that imposes zero tangential electric field?

▼ Work through the solution

Zero tangential electric field would give $H'(t) = 0$, not $H(t) = 0$. The hyperbolic solution switches to a cosh profile and the sheet impedance uses $\tanh(\gamma t)$. Identify the electromagnetic boundary condition, rather than inferring it just from the word “backing”.

Numerical field laboratory · 8. Separate diffraction from a finite observation window

Recover the far-field angular spectrum of a smooth aperture, then isolate the errors from truncating and sampling the aperture field.

Begin with a transform that can be checked analytically

Consider a long slit system whose transmitted scalar field varies along one transverse coordinate x. At a distant observation point in direction θ, a small aperture element accumulates an approximate relative phase $-kx\sin\theta$. The Fraunhofer amplitude is proportional to $\int A(x)e^{-2\pi i f x}\,dx$ with spatial frequency $f=\sin\theta/\lambda$. The approximation neglects the quadratic phase across the aperture; either a suitable lens or sufficient propagation distance is needed. A numerical Fourier transform cannot cure violation of this optical assumption.

Take the ideal aperture field $A(x)=e^{-x^2/w^2}$ and normalise far-field amplitude by its on-axis value. Integrating by parts or completing the square gives

$$F(f)/F(0)=e^{-(\pi w f)^2}.$$

One derivation avoids any complex contour: differentiate F with respect to f and replace $xA=-(w^2/2)A'$. Integrating by parts gives

$$F'=-2\pi^2 w^2 f F.$$

Solve this first-order equation with the measured $F(0)$. Thus a wider beam on the aperture narrows the far-field distribution. Squaring the amplitude gives the intensity measured by an ordinary detector.

At $w=1$ mm, the $1/e$ amplitude spatial frequency is $1/(\pi w)\simeq 0.3183$ cycles/mm. For 500 nm light this is a small angle of 1.59×10^{-4} radians. An instrument calibrated in radians per millimetre rather than cycles per millimetre would need a factor 2π; stating the Fourier convention prevents that frequent scale discrepancy.

What the instrument has really sampled

A scanner covers $-W\le x\le W$ and gives one amplitude at each of N equally spaced cell midpoints. The numerical transform is the weighted sum $\Delta x\sum_j A(x_j)e^{-2\pi i f x_j}$, where $\Delta x=2W/N$. By symmetry its imaginary part vanishes for a perfectly centred real Gaussian; the program sums the cosine projection. It normalises by the numerical on-axis amplitude, allowing us to compare shapes on a common vertical scale.

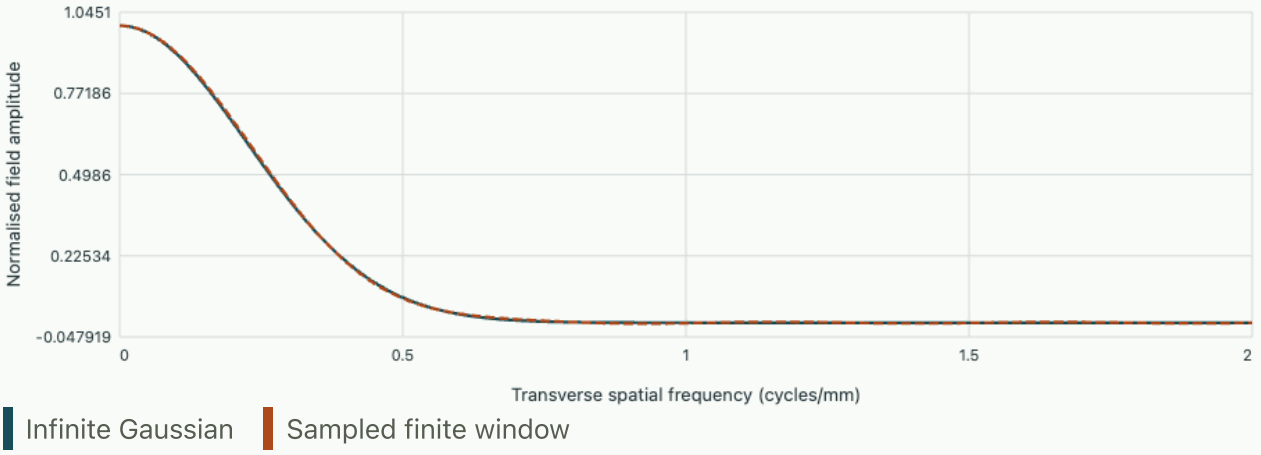
Two separate approximations have entered. Truncating the aperture multiplies the field by a rectangle. In frequency space this convolves the Gaussian spectrum with a sinc and can produce signed sidelobes. Sampling at regular intervals makes the spectrum repeat up to a phase convention, with alias copies separated by $1/\Delta x$. Increasing W at fixed N reduces truncation but worsens sampling. Widening the field of view and adding more detector pixels are not interchangeable changes.

Design a convergence test with two axes

Open the live experiment

Try the calculation

Gaussian aperture radius w (mm)	1
Window half-width / w	2
Samples across window	30



Sample spacing 0.13333 mm, Nyquist frequency 3.75 cycles/mm. Measured amplitude at 1 cycle/mm: -0.0014811.

Keep $w = 1$ mm and select 40 samples. Compare $W = w$ and $W = 3w$. The aperture field at the boundary falls from e^{-1} to e^{-9}, so an apparent sidelobe associated with the abrupt window should decrease strongly. Next fix $W = 3w$ and lower the number of samples. The displayed Nyquist frequency falls as Δx grows. When a plotted point lies above it, a smooth-looking value is an aliased sample of the periodic discrete spectrum.

At $W = 2w$, a centred Gaussian retains the on-axis amplitude fraction $\operatorname{erf}(2) \simeq 0.9953$ if integrated exactly. The normalised plot hides this 0.47% loss at zero frequency. If absolute power calibration matters, repeat the experiment without normalisation and independently integrate $|A|^2$ across the aperture. A shape comparison and a throughput comparison are different measurements, both meaningful.

Recover a size from a far-field measurement

Suppose a detector measures the intensity falling to one half at $f_{1/2}$. Since $I/I_0 = e^{-2(\pi w f)^2}$, the inferred w is $\sqrt{\ln 2/2}/(\pi f_{1/2})$. For $f_{1/2} = 0.1874$ cycles/mm the radius is about 1 mm. If the instrument actually reported amplitude half-width the corresponding formula would have $\sqrt{\ln 2}$ in the numerator; the radius would differ by a factor $\sqrt{2}$. The raw measured quantity and its threshold should accompany any quoted fitted radius.

A positive intensity cannot distinguish a negative sidelobe amplitude from a positive amplitude of the same magnitude. To observe the sign change shown by the finite-window amplitude curve, interfere the field with a phase reference. Conversely, when comparing with a camera image, square the signed numerical amplitude once and apply the detector pixel average. Comparing a real-amplitude plot directly with a recorded intensity array is an avoidable model mismatch.

Tutorial

How could one distinguish an intrinsic physical sidelobe from ringing produced by the scanner window without a known reference aperture?

▼ **Work through the solution**

Increase the scanned width while keeping Δx fixed by adding samples. Window artefacts should change as the omitted tails shrink. Repeat with a different smooth weighting and report its effect on resolution. An intrinsic feature that survives these tests can then be compared against an independent physical model.

At $W = 3$ mm and $N = 24$, what spatial frequencies can be represented without aliasing, and what does adding zero padding do?

▼ **Work through the solution**

$\Delta x = 0.25$ mm, so the Nyquist limit is 2 cycles/mm. Zero padding supplies more displayed frequency samples of the same discrete transform. It adds no measured aperture detail and cannot resolve aliasing already present in the original samples.

A beam shifted by x_0 within a sufficiently wide window has the same far-field intensity. What information changes?

▼ **Work through the solution**

Substituting $x' = x - x_0$ in the Fourier integral multiplies its amplitude by $e^{-2\pi i f x_0}$. The phase ramp records the displacement; the intensity is invariant. A truncated window may break this invariance by clipping different amounts of field.

Numerical field laboratory · 9. Estimate a cavity linewidth from a finite, biased record

Connect the time and frequency descriptions of one damped mode, then diagnose why a weak detector background systematically biases the inferred quality factor.

Know whether the detector reads field or energy

After a drive switches off, let a single cavity mode have complex analytic amplitude $a(t) = a_0 e^{-\kappa t/2} e^{-i\omega_0 t}$ for $t \geq 0$. The energy is proportional to $|a|^2$, giving $U = U_0 e^{-\kappa t}$. Its instantaneous decay rate $-U'/U = \kappa$ implies

$$Q = \omega_0 U / P_{\text{loss}} = \omega_0 / \kappa.$$

A phase-sensitive receiver measuring field amplitude sees a time constant twice as long as the energy lifetime. Before fitting the data, ask how the measurement transducer converts the field into a reported number.

Fourier transforming the causal decaying amplitude gives an integral proportional to $1/[\kappa/2 - i(\omega - \omega_0)]$, up to the time convention. Its squared magnitude has half-power points at $|\omega - \omega_0| = \kappa/2$. Therefore the full power linewidth in hertz is $\Delta f = \kappa/(2\pi)$. A measurement of either the energy decay time $\tau_U = 1/\kappa$ or the spectral width can test the other without having to calibrate the absolute detector gain.

At $f_0 = 2.0$ GHz and $\tau_U = 0.50\ \mu$s, we obtain $\kappa = 2.0 \times 10^6$ s^{-1}, $\Delta f \simeq 318.3$ kHz and $Q \simeq 6283$. A field receiver would report an amplitude time constant of 1.0 µs. Substituting that time directly into $Q = 2\pi f_0 \tau_U$ would double the apparent quality.

Derive the two-point fit with a detector offset

Let a power detector report $S(t) = Ae^{-\kappa t} + B$, with a constant background $B > 0$. In the absence of background two samples separated by Δt give

$$\widehat{\kappa} = \log[S(t)/S(t+\Delta t)]/\Delta t.$$

With background, the logarithmic derivative is

$$-\frac{d}{dt}\log S = \kappa \frac{Ae^{-\kappa t}}{Ae^{-\kappa t} + B}.$$

It starts below the true rate and tends to zero in the tail. A straight line fitted to late log data can therefore look impressively precise while overestimating the lifetime. Increasing the acquisition duration can worsen this bias if the entire record is forced into a straight-line model.

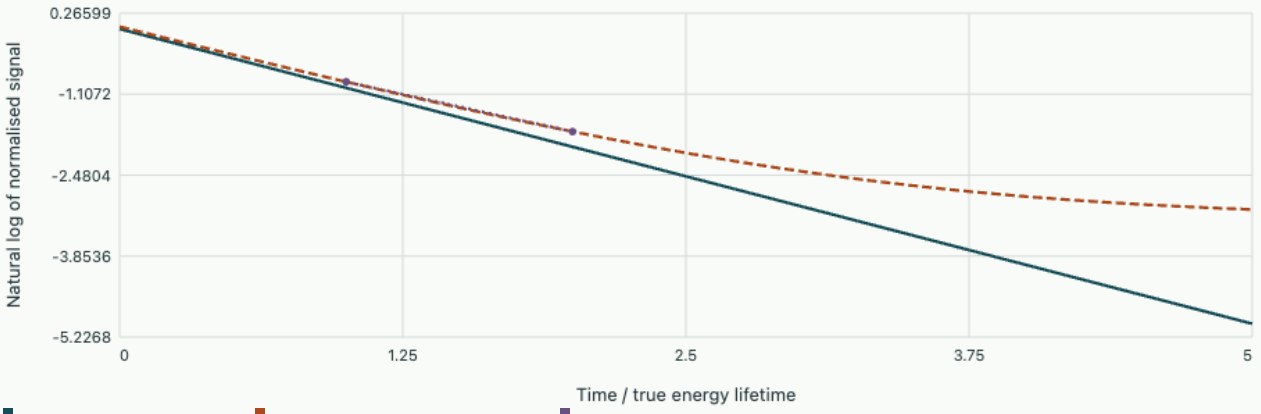
In the experiment set $b = B/A = 0.04$. A two-point fit from one true lifetime to two has dimensionless rate $\log[(e^{-1} + 0.04)/(e^{-2} + 0.04)]$, approximately 0.844. Its apparent lifetime is about 18.5% high. The source, walls and coupling have not changed. Only the interpretation of detector output has changed. Move the start of the fit towards three lifetimes and inspect the increasing gap from the ideal log line.

Open the live experiment

Try the calculation

Additive detector offset / initial energy **0.04**

Start of fitting window (lifetimes) **1**



Two-point inferred lifetime 1.1845 times the true lifetime. The same fractional bias transfers to Q for a known mode frequency.

Design a second independent inference

With three equally spaced, noiseless readings S_0, S_1, S_2 of one exponential plus a constant, subtract consecutive values before taking a logarithm. If $a = e^{-\kappa \Delta t}$, the differences are $A'(1-a)$ and $A'a(1-a)$, giving $a = (S_1 - S_2)/(S_0 - S_1)$. This cancels the constant background. At an interval of one lifetime, a first signal of unit energy and background 0.04 gives approximately $1.04, 0.407879, 0.175335$. Their ratio of differences recovers e^{-1}. In a real noise floor two nearly equal readings make this division unreliable. Use multiple windows or fit the raw signal with a measured noise model.

Finite spectral resolution creates a separate challenge. Multiplying the measured field by a rectangular window of duration T convolves its spectrum with a sinc-like response of width of order $1/T$. For the 318 kHz linewidth in the example, $T = 1$ µs gives frequency bins spaced by 1 MHz in a basic Fourier transform. Zero padding can interpolate its curve, but cannot restore the unmeasured late-time signal. A parametric fit may extract a sub-bin frequency or decay rate by using its model assumptions; its uncertainty then depends on signal-to-noise ratio and on whether the one-mode theory is adequate.

Check the single-mode model before refining its fit

If two nearby modes reach the same detector, their fields add before squaring. The result contains an interference term oscillating at their frequency difference on top of a decaying envelope. A scalar positive background cannot explain that alternating residual. Test by narrowing the excitation band, moving the pickup to a symmetry node of one mode, or examining a complex receiver's quadratures. A model with two

exponentials has more degrees of freedom and can fit noise as well as signal; require consistency across independent records.

An honest linewidth report specifies the measured quantity, acquisition window, background treatment, coupling state, reference frequency and uncertainty. A fit to stored energy measures total loss. Inferring internal absorption alone additionally requires a known port coupling or a measurement that changes it separately, as in the course's cavity project.

Tutorial

A detector background is independently measured but subtracted too aggressively. What feature can appear in the logarithmic tail?

▼ Work through the solution

The corrected signal can approach zero and cross below it. The logarithm then becomes undefined. Even before that point its apparent local slope can become arbitrarily large. Inspect the raw signed residual and its uncertainty instead of dropping nonpositive points without accounting for selection bias.

A second experiment doubles the peak stored energy while the absolute detector offset is fixed. What should happen to the shape of the normalised log tail?

▼ Work through the solution

$b = B/A$ halves. Background contamination becomes important later by about $\tau_U \ln 2$. A real change in the physical lifetime need not follow that shift when the response remains linear.

Can a passive resonator's intrinsic losses be determined from a single ring-down measured at one port?

▼ Work through the solution

The observed decay is the sum of internal and external rates. A single accurately measured lifetime identifies their sum. A calibrated change in coupling or an independent power budget is needed to separate them.

Numerical field laboratory · 10. A magnetic orbit that spirals outward in a perfect vacuum

Track a charged particle in a prescribed uniform field, derive a rotation update and separate numerical energy growth from a real magnetic force.

Use the work law before solving the trajectory

A nonrelativistic charged particle of mass m and charge q moves perpendicular to a uniform, time-independent magnetic field $B\hat{\mathbf{z}}$. There is no electric field and the dilute test particles do not affect the applied field. The Lorentz force gives $\dot{v}_x = \omega v_y$, $\dot{v}_y = -\omega v_x$ for $\omega = qB/m$. Dotting with $\mathbf{v}$ gives zero, so the speed should stay constant. This property holds for either sign of the charge; the direction of rotation changes. It is a better initial test of a trajectory solver than how circular a short segment looks on screen.

Define $J(v_x, v_y) = (v_y, -v_x)$, so $J^T = -J$ and $J^2 = -I$. The exact velocity over a time step h is $[\cos(\omega h)I + \sin(\omega h)J]\mathbf{v}$. Integration in time gives a circular orbit of radius $v_\perp/|\omega|$ once the guiding centre is chosen. A magnetic field alone bends a trajectory without changing its kinetic energy; a slowly growing orbit radius at fixed B and fixed initial speed demands an explanation.

Why forward Euler adds energy at every step

Euler gives $\mathbf{v}_+ = (I + dJ)\mathbf{v}$ with $d = \omega h$. Squaring the norm eliminates the cross term because $v^T J v = 0$, leaving $\|\mathbf{v}_+\|^2 = (1 + d^2)\|\mathbf{v}\|^2$. No nonzero step has exact unit gain. A geometric picture makes the bias tangible: a tangent step leaves the initial velocity circle on its outside, whatever the point of departure.

After N steps the speed factor is $(1 + d^2)^{N/2}$. With 24 steps per physical orbit, $d = 2\pi/24 \simeq 0.2618$, so a full orbit multiplies speed by roughly 2.2. At shorter times or smaller steps this can resemble weak physical acceleration. Reducing h at fixed physical duration reduces the growth. Testing one step at fixed count, rather than comparing over a fixed physical interval, underestimates the cumulative error.

Solve a centred magnetic step exactly

Use the time-centred velocity in the Lorentz force:

$$\mathbf{v}_+ - \mathbf{v} = dJ(\mathbf{v}_+ + \mathbf{v})/2.$$

Premultiply by the transpose of $\mathbf{v}_+ + \mathbf{v}$. The left side is the difference of squared speeds and the right side vanishes, so the update preserves speed for every d in exact arithmetic. Rearranging yields a two-dimensional linear solve; using $J^2 = -I$ turns it into a rotation,

$$\mathbf{v}_+ = \frac{1 - t^2}{1 + t^2}\mathbf{v} + \frac{2t}{1 + t^2}J\mathbf{v},$$

$$t = d/2.$$

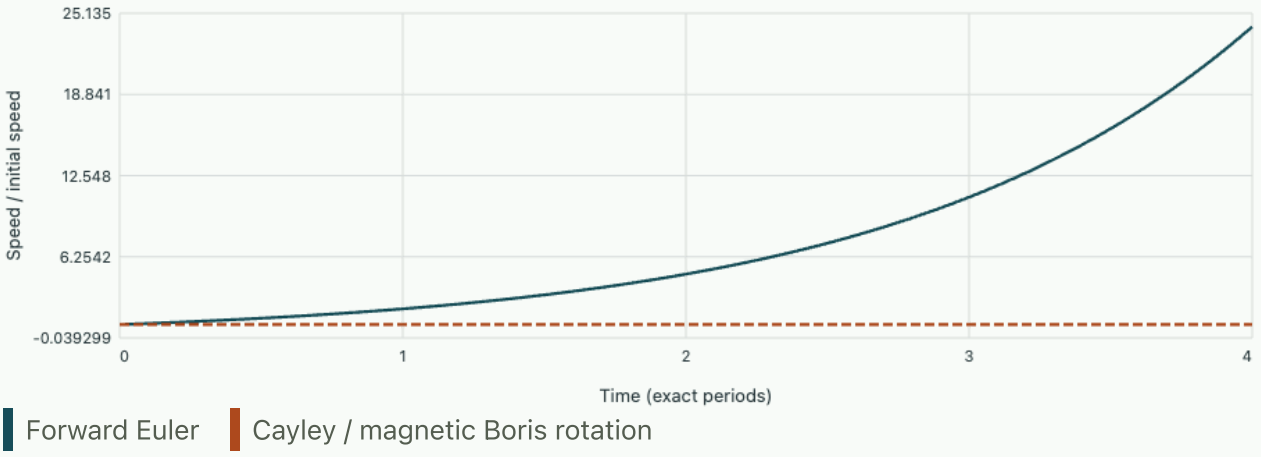
These coefficients are $\cos\alpha$ and $\sin\alpha$ for $\alpha = 2\arctan(d/2)$. This is the magnetic rotation used in the nonrelativistic Boris integrator. For simultaneous electric fields, the usual split step adds two half electric impulses around the magnetic rotation. The current experiment isolates the magnetic part; it makes no claim about a nonuniform electric field or the accuracy of interpolation from a field mesh to particles.

Open the live experiment

Try the calculation

Steps per exact orbit **24**

Number of orbits **4**



After 4 orbit(s), Euler speed 24.096, rotation speed 1. Rotation still lags the exact phase by 8.1411°.

Speed conservation does not fix arrival time

At small d,

$$\alpha = d - d^3/12 + O(d^5).$$

A speed-preserving orbit lags in gyrophase, even when its energy plot stays perfectly flat. After four true orbits at 24 steps per orbit the predicted lag is about 8.14 degrees. The browser advances the Euler and rotation velocities and reports both norm and lag. The y-axis rescales to fit both curves, so a sharply growing Euler error can flatten the appearance of the well-behaved curve. Inspect numerical values as well as the picture.

Suppose an optical pickup observes the time at which a dilute ion tracer first crosses a known line of sight after several orbits. A phase error can be an important measurement error even when the speed is exact. To use the trajectory as a reference, refine the step until the crossing time changes by less than the experimental timing resolution. If the required integration spans many revolutions, a uniform field permits exact substeps. In realistic fields one can retain a structure-preserving rotation while resolving spatial variation independently.

A separate diagnostic is time reversal. Starting from the final velocity and applying the rotation with $-d$ exactly inverts a centred step, up to rounding. Forward Euler with $-d$ after $+d$ gives

$$(I - dJ)(I + dJ) = (1 + d^2)I.$$

A forward/backward test detects the lost reversibility without knowing a single experimental constant.

Tutorial

Given an orbit with 48 steps per period, estimate the gyrophase error after ten periods using the leading asymptotic term.

▼ **Work through the solution**

There are 480 steps with $d = 2\pi/48$. The accumulated lag is about $480d^3/12 \simeq 0.0897$ rad, or 5.14 degrees. Doubling the number of steps per period reduces the leading accumulated error fourfold.

A computation preserves speed but reverses the sign of its orbit relative to observation. Would the magnetic work test detect the error?

▼ **Work through the solution**

Both field signs give zero work. Check a signed initial acceleration component $a_y = -qBv_x/m$ for a positive initial v_x, or observe the oriented trajectory. A scalar invariant cannot detect all vector-sign errors.

What changes in the relativistic case with appreciable kinetic energy?

▼ **Work through the solution**

Evolve momentum $\mathbf{p} = \gamma m\mathbf{v}$ and infer velocity from momentum. In a uniform magnetic field without electric work γ stays constant, and the orbital frequency is reduced to $|q|B/(\gamma m)$. An energy-preserving nonrelativistic integration at the wrong frequency would still give the wrong measured period.

Numerical field laboratory · 11. When differentiating a measured dipole invents radiated power

Infer dipole radiation from a sampled displacement and show quantitatively why position noise can dominate a seemingly well-resolved power estimate.

From trajectory to a physically meaningful quantity

An electrically small source carries an oscillating electric dipole moment $p(t) = p_0 \cos \omega t$ in a fixed direction. In nonrelativistic electric dipole radiation, the total instantaneous far-zone power associated with retarded time is

$$P = |\ddot{p}|^2/(6\pi\epsilon_0 c^3).$$

The average over a single sinusoidal cycle is $p_0^2\omega^4/(12\pi\epsilon_0 c^3)$. The fourth power of frequency expresses a physical sensitivity, but a numerical second derivative can introduce a separate sensitivity to high-frequency measurement errors.

A sensor supplies moment or displacement samples at times $t_j = jh$. If the source is a known rigid displacement $x(t)$ of charge q, then $p = qx$. The constant conversion from moment to displacement cancels when comparing a numerical power with a reference power. We therefore scale amplitude by p_0 and time by $1/\omega$. The true acceleration trace is $-\cos t$. Comparing it with the estimated acceleration before squaring reveals signs and phases that a power curve conceals.

Derive the frequency response of the differentiation formula

Use $D_2 p_j = (p_{j-1} - 2p_j + p_{j+1})/h^2$. On a complex exponential $e^{i\nu jh}$ the symbol is $-4\sin^2(\nu h/2)/h^2$. Dividing its magnitude by the exact ν^2 gives a squared sinc factor. At a fixed physical frequency the error is quadratic in h. For a signal oscillating in sign from one sample to the next, $\nu h = \pi$, and the numerical second derivative multiplies its amplitude by $4/h^2$. A numerical derivative smooths the very highest frequency relative to an exact high-frequency derivative yet still strongly amplifies the raw high-frequency component relative to a low-frequency physical acceleration.

Add a deliberately reproducible alternating error $\epsilon(-1)^j$ to the position readings. Its acceleration error is $-4\epsilon(-1)^j/h^2$. Squared and averaged over a complete even grid, the alternating error and the fundamental are orthogonal. Therefore the normalised inferred mean power becomes

$$\frac{\langle P_{\mathrm{est}} \rangle}{\langle P_{\mathrm{true}} \rangle}$$

=\left[\frac{\sin(h/2)}{h/2}\right]^4+2\left(\frac{4\epsilon}{h^2}\right)^2 .

The first term tends to one from below as sampling is refined. At a fixed per-sample alternating displacement amplitude the second grows as h^{-4}. The experiment holds a particular *sampling-related* error fixed; it does not claim that all real noise changes frequency when the sampling rate changes. A fixed-frequency physical disturbance obeys the symbol at its own fixed ν.

A numerical example in which refinement hurts

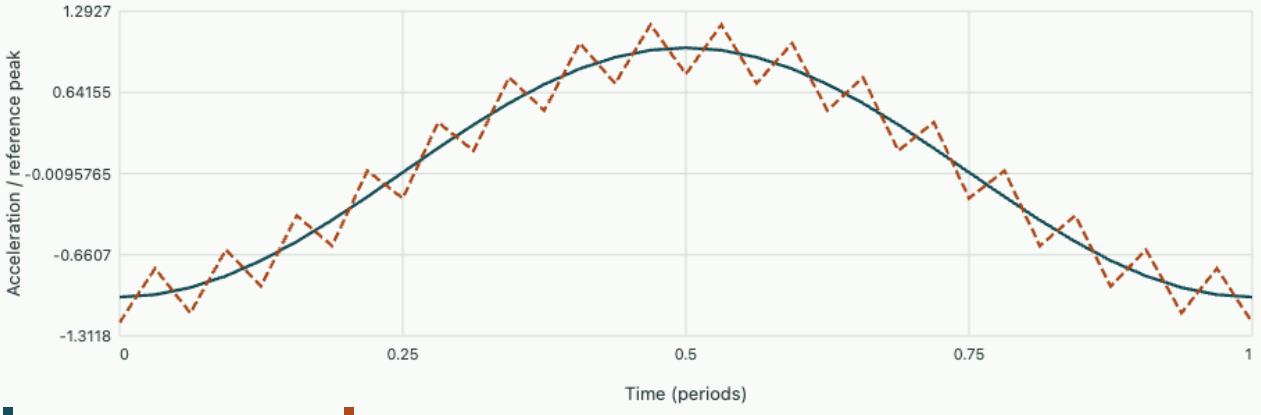
For 32 samples per period, $h = 2\pi/32 \simeq 0.19635$, and a positional disturbance two parts per thousand of full amplitude gives a spurious acceleration peak about 0.2075 of the true peak. Position data plotted at ordinary scale can look nearly perfect. The normalised mean-square acceleration is already about 1.080. At 64 samples per period with the same alternating error the spurious acceleration amplitude grows fourfold, and its contribution to mean power grows sixteenfold.

Open the live experiment

Try the calculation

Samples per physical period **32**

Position disturbance / amplitude **0.002**



Mean squared acceleration / exact value 1.0797. The alternating error contributes peak acceleration 0.20751 in the same units.

Set the perturbation to zero first and verify the approach to one as the physical sinusoid becomes increasingly well resolved. Restore the perturbation and determine which combination of resolution and disturbance first incurs a ten-per-cent error. The error can initially decrease and later increase with sample rate. Choosing the estimate that happens to lie closest to one at a single resolution can hide a cancellation between physical-mode discretisation error and positive noise power.

Recover a physical model without claiming information you did not measure

If an independent measurement shows the source contains only one sinusoidal frequency, fit its sine and cosine coefficients to the positions rather than differentiating every neighbour. On a full uniformly sampled cycle these coefficients are orthogonal to the alternating offset. Multiplying the fitted phasor by $-\omega^2$ gives the model-based acceleration. This can be a much more precise estimate, but it assumes other harmonics are absent or irrelevant. A weak physical second harmonic has an acceleration amplification of four and

power weighting of sixteen relative to equal moment amplitude. Validate the assumed spectrum before suppressing it.

For broad-band signals one may fit a smooth basis with an explicit high-frequency cutoff or use a model of detector bandwidth. Record the chosen cutoff, its physical justification and the uncertainty propagated into power. Repeating the computation with nearby cutoffs is a test of robustness, not a derivation of the correct filter. Where possible, compare the inferred radiated energy with a distant calibrated flux measurement over the same retarded-time interval. These two measurements have different systematic errors.

Tutorial

Independent, zero-mean random sample errors have variance σ_p^2. What is the variance in one centred second derivative and why need adjacent derivative errors not be independent?

▼ **Work through the solution**

The weights are $1, -2, 1$ divided by h^2. Independence of the original samples gives variance $6\sigma_p^2/h^4$. Adjacent derivative estimates share two original samples with nonzero weights, so their covariance is generally nonzero.

A second physical harmonic has amplitude 5% of the fundamental. How much does it contribute to the total time-averaged electric-dipole radiation, assuming orthogonal harmonics over the record?

▼ **Work through the solution**

Its acceleration amplitude relative to the fundamental is

$$0.05 \times 2^2 = 0.2.$$

Its mean power is $0.2^2 = 0.04$ of the fundamental contribution, about 3.85% of their sum. A source spectrum small in displacement need not be equally small in radiated power.

Why must a measured power spectrum specify whether it represents one-sided or two-sided frequencies?

▼ **Work through the solution**

A real signal contains conjugate spectral amplitudes at opposite frequencies. A one-sided positive-frequency spectrum usually combines their powers. Counting both after applying such a convention doubles the inferred variance and radiation power.

Numerical field laboratory · 12. Choose a frequency that identifies a coil's stored and dissipated energy

Design a small-signal experiment to estimate resistance and inductance, propagate amplitude and phase uncertainty and decide when a distributed field model is needed.

A measured impedance is one complex number

A laboratory sensing coil is driven by a small sinusoidal current at angular frequency ω. In its simplest usable range, a series model predicts $Z = R + i\omega L$. The voltage across a port defines the sign of current entering the positive terminal. With peak phasors, complex power is

$$S = VI^{*}/2 = |I|^{2}(R + i\omega L)/2.$$

Thus the real part measures average irreversible loss. Magnetic energy at peak current is $L|I|^{2}/2$; reactive power equals ω times this peak energy. These equalities supply two routes to the same parameters provided parasitic capacitance and frequency-dependent losses remain negligible.

At 100 Hz, a coil with $R = 20\ \Omega$ and $L = 40$ mH has reactance $2\pi(100)(0.040) = 25.13\ \Omega$, magnitude 32.12 Ω and phase about 51.49 degrees. A 10 mA peak current produces about 0.321 V peak and 1 mW average resistive loss. The peak magnetic energy is only $2\ \mu\mathrm{J}$. Check that the reactive power obtained from the phase and current is about 1.26 mvar. Confusion between RMS and peak phasors easily introduces a factor two while still leaving the impedance ratio correct.

Derive what an amplitude error and a timing error do

Suppose the readout gives magnitude M and phase ϕ. Then $R = M\cos\phi$, $L = M\sin\phi/\omega$. For small independent uncertainties and known ω, differentiate both expressions. The relative first-order perturbations are

$$\frac{\delta R}{R} = \frac{\delta M}{M} - \tan\phi\,\delta\phi,$$

\frac{\delta L}{L}=\frac{\delta M}{M}+\cot\phi\,\delta\phi .

The phase uncertainty must be expressed in radians. For independent random errors their variances add with the squared coefficients; the signed equations above describe individual perturbations. Near zero phase even a very accurate amplitude tells little about inductance, while near 90 degrees inferring resistance requires exceptional phase accuracy. A near-45-degree measurement balances these two sensitivities when the uncertainty budget treats them similarly.

If the phase uncertainty is 0.5° and relative magnitude uncertainty is 1%, the 45-degree standard deviation in either recovered parameter is approximately $\sqrt{0.01^{2} + 0.00873^{2}} = 0.0133$, or 1.33%. At five degrees the relative inductance contribution from phase alone is $\cot 5^{\circ} \times 0.00873 \simeq 10\%$. Taking data at more digits does not improve phase calibration.

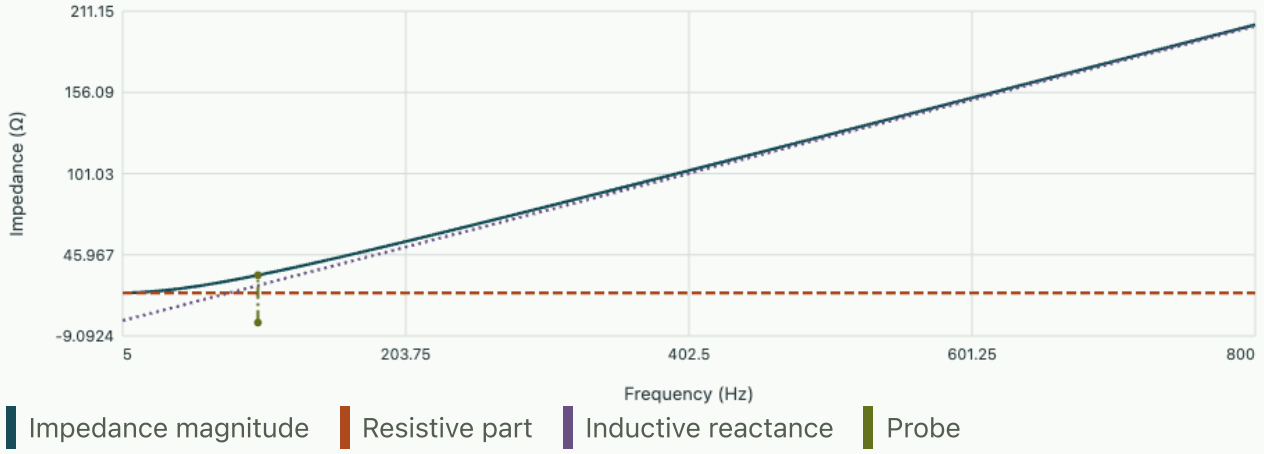
Choose the operating point, then challenge the model

Open the live experiment

Try the calculation

Probe frequency (Hz) 100

Coil inductance (mH) 40



At the probe: |Z| = 32.119 Ω, phase 51.488°. Equal resistive and reactive contributions at 79.577 Hz.

Vary frequency at fixed inductance and read where resistive and reactive contributions are equal. The crossover is $f_c = R/(2\pi L)$, about 79.58 Hz for the default coil. Now double the inductance without changing R; the crossover halves. The graph is a prediction with known parameters. In an actual inverse experiment you know only rough bounds initially. Begin with a small frequency sweep that brackets the crossover, then allocate repeated measurements where the target parameter is most sensitive.

Measure at two or more separated frequencies and reconstruct R and L independently at each. A frequency-dependent apparent R could be caused by skin effect or core loss. A falling apparent inductance near a self-resonance can signal stray capacitance. Simply allowing R, L to take arbitrary values at each frequency fits all data without explaining the mechanism. Add a candidate physical effect and check an independent observable such as loss under a different current amplitude, ring-down frequency or a change in lead geometry.

Close the loop from a fitted parameter to a field calculation

The lumped magnetic energy $LI^2/2$ should match the field integral $\int \mathbf{B} \cdot \mathbf{H}\, dV/2$ for a linear reversible system that includes the relevant surrounding field. This is why the exterior air region and its imposed boundary can affect a numerical estimate even when the measured coil is compact. Increase the domain and refine the mesh separately. A converged mesh in a small box can deliver a very repeatable value for the wrong boundary problem. The earlier chapter on mutual-inductance matrices generalises this check to cross terms between currents.

Another independent measurement uses a slowly executed current ramp with a known winding resistance. Subtract the measured resistive voltage from terminal voltage and integrate over time. Under the linear model,

$$\int (V - RI)\, dt = L\Delta I.$$

This test weights offsets very differently from sinusoidal phase detection: a constant voltmeter offset integrates into a bias proportional to record duration, while a high-frequency phase error drops out. Agreement of the two methods is stronger evidence than repeating a single instrument setup at the same frequency.

Tutorial

Only voltage and current magnitudes are available, but two frequencies are accurately known. Can a constant R and L be recovered, and what makes the inference fragile?

▼ **Work through the solution**

$$M_i^2 = R^2 + \omega_i^2 L^2.$$

Subtract the equations to get

$$L^2 = (M_2^2 - M_1^2)/(\omega_2^2 - \omega_1^2).$$

Insert this in either equation for R^2 and take positive roots. Nearly equal frequencies produce a small denominator; nearly equal magnitudes amplify measurement noise. Under a misspecified model the inferred square can even become negative.

At a fixed frequency the current reverses direction in a repeat measurement but its reference voltage polarity is held fixed. What measurement error could this create?

▼ **Work through the solution**

Reversing the defined port-current orientation multiplies the inferred impedance by –1 unless the voltage reference is reversed consistently. An apparent negative resistance in passive hardware may be a port convention issue. Check the direction of measured power flow with a known reference resistor.

What extra evidence would justify carrying a fitted inductance into a time-domain prediction?

▼ **Work through the solution**

Make the prediction for a current waveform whose spectrum lies within the validated range of the measurements. Test at another amplitude for nonlinearity or hysteresis and record the residual voltage

after subtracting $RI + L\,dI/dt$. A successful fit at one frequency alone does not establish a broadband constitutive law.

Numerical field laboratory · 13. Where should a numerical electrostatic domain end?

A finite mesh needs an outer boundary even when the physical field extends to infinity. Solve the artificial-boundary error exactly for a dipole and test whether enlarging the numerical domain or matching the exterior field is more effective.

Match a harmonic in an empty outer shell

At the surface of a source-containing sphere of radius a prescribe $V(a,\theta) = V_a \cos\theta$. Outside this sphere there is no charge. Insert $V = f(r)\cos\theta$ in Laplace's equation. The angular Laplacian supplies the eigenvalue -2, so $(r^2 f')' - 2f = 0$. A power-law trial gives

$$f = Ar + B/r^2.$$

Decay at infinity removes the growing term and the open-space reference is $f_\infty = V_a(a/r)^2$.

Terminate the calculation at R and ground the artificial wall. The two boundary equations instead give

$$f_R(r) = V_a\Big(\frac{a}{r}\Big)^2 \frac{1-(r/R)^3}{1-(a/R)^3}.$$

A well-converged interior solve can reproduce this answer perfectly while retaining a systematic exterior-domain bias. An alternative boundary condition is $f'(R) + 2f(R)/R = 0$. It admits the decaying dipole and annihilates the unwanted growing coefficient. A single local coefficient matches one spherical-harmonic order exactly.

Estimate the error before refining the mesh

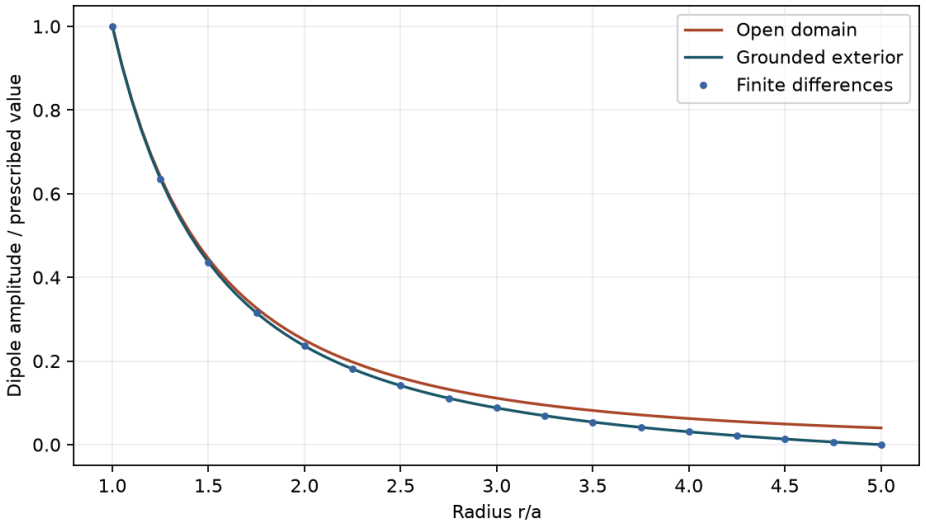
At $r = 2a$ with $R = 5a$, the grounded boundary gives

$$f_R/f_\infty = (1 - 8/125)/(1 - 1/125) = 117/124.$$

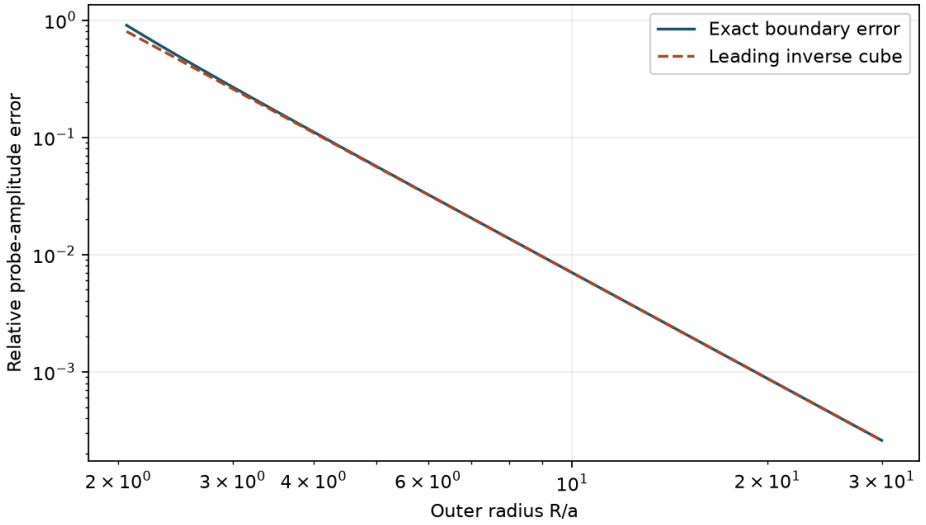
The amplitude error is about minus 5.65 percent. At a fixed observation radius, the leading bias decreases with the inverse cube of R. Refining the interior grid at fixed R cannot remove it. Compare a second-order radial difference solution with both analytic curves to measure the two errors independently.

To see what the linear solver actually controls, put a uniform grid on r/a. Approximate the outgoing flux $r^2 f'$ at the right face by $r_{i+1/2}^2(f_{i+1} - f_i)/h$ and form the same quantity at the left face. Subtract the fluxes, divide by h and equate to $2f_i$. The resulting tridiagonal system has positive diagonal $2 + (r_{i+1/2}^2 + r_{i-1/2}^2)/h^2$. Move the known inner and outer boundary terms to the right-hand side. For the grounded wall the outer contribution is zero; the inner value is the amplitude unit.

At $r/a = 2$ the open-domain amplitude is 0.2500. The grounded reference gives 0.2359 when $R/a = 5$. Repeating at $R/a = 10$ raises that prediction to 0.24825. Keep the physical grid spacing fixed when making this domain comparison. If the cell count stays fixed, increasing the radius also coarsens the mesh, and the two effects can mask each other. The first figure supplies both analytical references needed to separate them.



The finite-difference solution follows the field selected by the grounded exterior wall. Its agreement with that curve says nothing about whether the wall models open space.



Increasing the domain radius at fixed physical probe location shows the inverse-cube asymptote. A mesh study alone would miss this error.

Exercise · size the exterior shell

For a probe at $2a$, require an error below one percent. What radius is needed?

▼ Work through the solution

The exact relative bias magnitude is $7/[(R/a)^3 - 1]$. Hence $(R/a)^3 > 701$ or $R/a > 8.883$. A value of nine satisfies the condition before numerical discretisation is considered.

Exercise · change the angular dependence

What Robin coefficient would match a pure exterior quadrupole?

▼ Work through the solution

A decaying multipole of order ℓ behaves as $r^{-(\ell+1)}$. Its logarithmic slope at R is $-(\ell+1)/R$, so the quadrupole needs $3/R$. A superposition of orders requires angular information or an exterior operator that acts differently on each mode.

Try it · Open domain

Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Grounded outer radius / sphere radius 5

Relative boundary error at r/a=2 is 0.056452. This error remains after solving the discretised bounded-domain problem accurately.

Download the corresponding editable notebook. The browser recalculates the reduced equations directly; the notebook exposes the full calculation and its checks.

Open editable notebook Download .ipynb

Applied study

Derive the fields of a coaxial cable and follow how electromagnetic energy travels along it.

Where does energy travel in a cable?

A coaxial cable contains two conductors separated by an insulator. Current travels along the metal surfaces. To locate the energy delivered to a load, find the electric and magnetic fields in the space between them. Consider concentric, perfectly conducting cylinders of radii a and b, separated by a uniform lossless dielectric of permittivity ϵ and permeability μ. Treat a straight, uniform section far from connectors.

Place charge per unit length $+q'$ on the inner conductor and $-q'$ on the facing shield. A cylindrical Gaussian surface of radius r, with $a < r < b$, gives $2\pi r D_r = q'$ per unit length. Integrating $E_r = D_r/\epsilon$ from inner to outer conductor gives $V = q' \ln(b/a)/(2\pi\epsilon)$. Thus $C' = q'/V = 2\pi\epsilon/\ln(b/a)$ is capacitance per metre. More charge fits at the same voltage when the dielectric permittivity rises. Fields and voltage at a dielectric interface.

Find the second energy store

Let a current I flow along the inner conductor and return on the shield. Ampère's law gives a circumferential magnetic field $H_\phi = I/(2\pi r)$. Integrate its energy density $\mu H_\phi^2/2$ over the annular cross-section. The resulting magnetic energy per metre is $\mu I^2 \ln(b/a)/(4\pi)$. Equating this to $L' I^2/2$ gives $L' = \mu \ln(b/a)/(2\pi)$. The product $L'C' = \mu\epsilon$ is independent of the two radii for this homogeneous line. Magnetic fields from steady currents.

Couple the fields into a travelling wave

A slice of cable of length dz gains charge when more current enters than leaves. Its charge is $C'V\,dz$, giving $\partial_z I = -C'\partial_t V$. The changing magnetic flux gives the complementary voltage drop $\partial_z V = -L'\partial_t I$ by Faraday's law. Differentiate once more to obtain

$$\partial_{zz} V = L'C'\partial_{tt} V.$$

For a waveform travelling in the positive z direction, these equations require

$$v = \frac{1}{\sqrt{L'C'}},$$

$$\frac{V}{I} = Z_0 = \sqrt{\frac{L'}{C'}}.$$

Here Z_0 is the ratio for a single travelling wave. The radial electric field crossed with the circumferential magnetic field gives axial Poynting flux. Indeed $\int_a^b E_r H_\phi \, 2\pi r \, dr = VI$. The surrounding dielectric carries this electromagnetic power in the ideal model.

What crosses the gap in a charging capacitor?

A changing electric field can also generate a magnetic field where there are no moving free charges. Charge a parallel-plate capacitor through a wire at steady current I. A loop encircling the wire bounds a surface cut by the wire, but also a deformed surface passing between the plates. A consistent magnetic circulation for both surfaces requires the Maxwell term in Ampère’s law,

$$\oint \mathbf{H}\cdot d\boldsymbol{\ell} = I_{\mathrm{free}} + d\int \mathbf{D}\cdot d\mathbf{A}/dt.$$

Between wide circular plates the total electric displacement flux is the plate charge, so its derivative is exactly I.

Take plate radius $R = 5.0$ cm and $I = 1.0$ A. Inside the gap, neglect fringing and assume slow charging, so that D is uniform over the plate area. A centred loop of radius $r = 2.5$ cm encloses the fraction $r^2/R^2 = 1/4$ of the changing flux. The magnetic flux density there is $B = \mu_0 I r/(2\pi R^2) = 2.0$ µT. A loop just outside the plate radius encloses the full displacement current. Charge conservation, the electric field and the magnetic circulation thus give the same current whether the chosen surface cuts a wire or a dielectric.

A short pulse in polyethylene

Use $a = 0.450$ mm, $b = 1.475$ mm,

$$\epsilon = 2.25\epsilon_0,$$
$$\mu = \mu_0.$$

The logarithm of the radius ratio is 1.187. The line constants are 105.44 pF/m and 237.43 nH/m. They give $Z_0 = 47.45\ \Omega$ and a propagation speed of 2.00×10^8 m/s. A 1 V forward pulse has current $1/47.45 = 21.1$ mA and instantaneous power 21.1 mW while its flat top passes a cross-section. It reaches the end of a 10 m line about 50 ns after launch.

Exercises on transmission and reflection

1. Change voltage without changing geometry

Double the amplitude of the forward pulse. Compare electric energy per metre, magnetic energy per metre and transmitted power with their original values.

▼ **Work out the energy factors**

$I = V/Z_0$ doubles. Thus the two energies $C'V^2/2$ and $L'I^2/2$ both quadruple, and VI also quadruples. Their ratio is unity for a travelling wave because $L'/Z_0^2 = C'$. The pulse speed stays fixed since neither line constant changed.

2. Match the load

Terminate the cable in $R_L = 2Z_0$. At the load the forward and backward voltages add, while their currents have opposite propagation signs. Find the reflected fraction of incident power.

▼ **Apply the boundary condition**

Write $V_+ + V_- = R_L(V_+ - V_-)/Z_0$. This gives

$$V_-/V_+ = (R_L - Z_0)/(R_L + Z_0) = 1/3.$$

Each wave's power magnitude is proportional to its voltage squared, so $1/9$ returns and $8/9$ is absorbed. For $R_L = Z_0$, the boundary condition accepts the incident wave without a reflected wave.

ONESIMULATION LABORATORY

Recover cable impedance from a cross-section

Compare recorded OneSimulation runs on this page. To test your own inputs, download the notebook and model files and run the model in an activated OneSimulation environment.

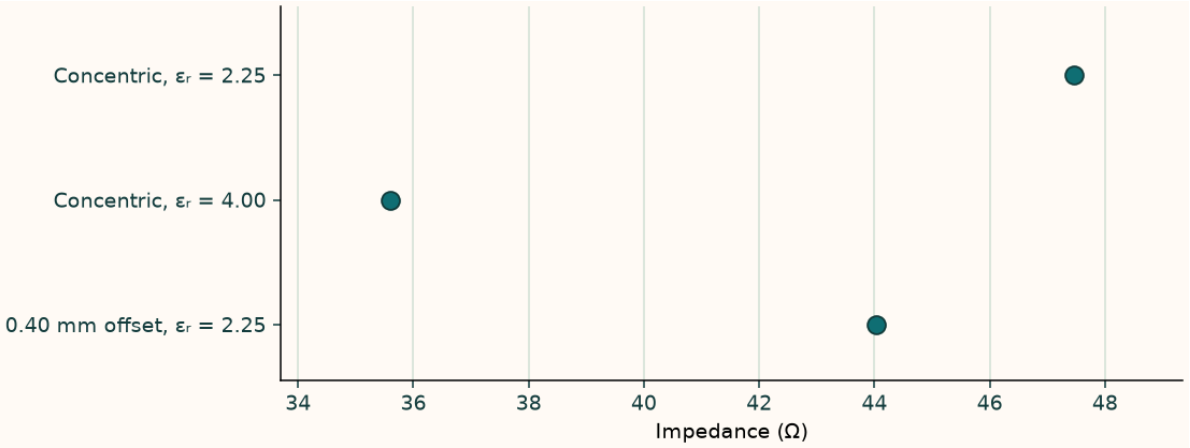
The cable notebook solves the electrostatic cross-section with the shield grounded, infers C' independently from charge and energy, and uses the homogeneous TEM relation for Z_0. These steady field solutions give the constants needed for the pulse calculation. Coaxial cable field model.

1. **Baseline.** Run both supplied STEP cross-sections and read the `concentric` row first. At 1 V, compare the two estimates of C' with 105.44 pF/m and Z_0 with 47.45 Ω. Use a 0.1% relative tolerance.
2. **Change the dielectric.** In the model cell set the multiplier of `VacuumPermittivity()` to 4.0 from 2.25 and rerun that cell and the line-constants cell. In the concentric result C' rises to 187.45 pF/m, L' stays near 237.43 nH/m and Z_0 falls to 35.59 Ω. Explain all three changes from the field integrals.
3. **Move the inner conductor.** Restore the original dielectric and compare the `eccentric` row with the baseline. Its centre is offset by 0.40 mm. On the close side the field concentrates. Expect about 113.61 pF/m and 44.04 Ω, within 0.1%. Check that changing the geometry alters impedance while leaving $1/\sqrt{L'C'}$ unchanged to rounding.

Line constants recovered from charge in the numerical cross-section

Cable cross-section	Capacitance (pF/m)	Inductance (nH/m)	Impedance (Ω)
Concentric, ε_r = 2.25	105.4389	237.433	47.4537
Concentric, ε_r = 4.00	187.4469	237.433	35.5902

Cable cross-section	Capacitance (pF/m)	Inductance (nH/m)	Impedance (Ω)
0.40 mm offset, $\varepsilon_r = 2.25$	113.6166	220.343	44.0381



Increasing the concentric cable's relative permittivity from 2.25 to 4 lowers the impedance from 47.45 to 35.59 Ω. The offset conductor lowers it to 44.04 Ω at the original permittivity.

Download experiment files (ZIP)

Extract the files together and open `coaxial_cable_impedance.ipynb` in a local OneSimulation environment.

Try it · Capacitance and impedance of the same cable

Keep the notebook radii at 0.450 mm and 1.475 mm. Set the permittivity to 4, or restore 2.25 and move the inner conductor by 0.40 mm, to reproduce the two table comparisons. The exact eccentric-cylinder TEM relation supplies a reference for the finite-element charge and energy measurements.

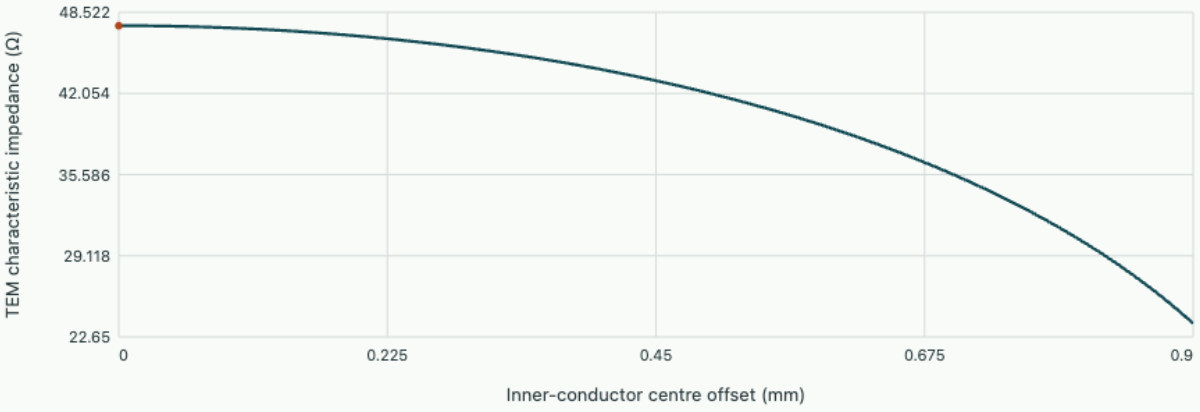
Open the corresponding OneSimulation notebook, code and output

Open this live calculation

Try the calculation

Relative permittivity **2.25**

Inner-conductor centre offset (mm) **0**



Eccentric cylindrical cable | Chosen geometry

C = 105.44 pF/m; L = 237.43 nH/m; Z_0 = 47.454 Ω. Exact TEM reference for the notebook geometry a = 0.450 mm, b = 1.475 mm; uniform nonmagnetic dielectric, lossless circular conductors. The OneSimulation charge/energy estimates include mesh error.

Problem sheet 6 · Can impedance identify a dielectric by itself?

Mathematical field guide to electrodynamics

Begin with a voltmeter and a small surface. A voltage difference measures work per charge; electric flux measures a component of the field summed over area. Following those two measurements leads to the field equations. Later entries show how to choose boundary data, recognise a harmonic convention and check a result in a different inertial frame.

Choose a quantity you could measure

Begin with the field on the line between two large parallel plates. Take the left plate at $+20$ V and the right at zero, separated by 2.0 cm. A positive test charge released between them accelerates to the right. Work per charge is $20\,\mathrm{J/C}$; divided by $0.020\,\mathrm{m}$ this gives

$$E = 1000\,\mathrm{V/m} = 1000\,\mathrm{N/C}.$$

On a $+3.0\,\mathrm{nC}$ bead the force is $3.0\,\mu\mathrm{N}$. A voltmeter attached to the two plates still reads 20 V if their separation doubles at fixed potentials, even though the field halves.

This book uses SI. Magnetic flux density B is in tesla, magnetic field strength H in A/m, electric displacement D in C/m², and electric field E in V/m. In vacuum $D = \epsilon_0 E,\ B = \mu_0 H$. For calculations here $\epsilon_0 \simeq 8.8542 \times 10^{-12}\,\mathrm{F/m}$, $\mu_0 \simeq 1.2566 \times 10^{-6}\,\mathrm{H/m}$ and $c = (\epsilon_0 \mu_0)^{-1/2}$. If this is your first encounter with field theory, work the plate example, then the gradient and flux examples immediately below, before starting the charge chapter. Start with charge and potential.

Take a directional derivative before assigning a force

For a smooth scalar potential ϕ, a small displacement produces

$$d\phi = \partial_x \phi\, dx + \partial_y \phi\, dy + \partial_z \phi\, dz = \nabla\phi \cdot d\boldsymbol{\ell}.$$

Work done by a static field on charge q is $q\boldsymbol{E} \cdot d\boldsymbol{\ell} = -q\, d\phi$. Equality for every displacement gives

$$\boldsymbol{E} = -\nabla\phi.$$

The minus sign follows from the decrease in potential energy $q\phi$ when the field does positive work.

In a patch with $\phi = (3\,\mathrm{V/m^2})xy$, at $(x, y) = (2, 1)\,\mathrm{m}$ the field is $(-3, -6, 0)\,\mathrm{V/m}$. A $1\,\mathrm{cm}$ displacement in $+x$ raises the potential by 0.030 V; the field does negative work on a positive bead moving that way. Following any curve between two points in this static patch gives the same integral $-\Delta\phi$. A time-varying magnetic flux changes that last measurement.

Attach the right area and volume to a coordinate

A radial increment is not itself a spherical volume. At radius r, angular increments sweep arc lengths $r\, d\theta$ and $r \sin\theta\, d\varphi$, where θ is the polar angle. Their product gives

$$dA = r^2 \sin\theta\, d\theta\, d\varphi = r^2 d\Omega.$$

Integrating over the sphere gives $4\pi r^2$, so $dV = r^2 dr\, d\Omega$. In cylindrical coordinates the analogous azimuthal arc is $s\, d\varphi$, giving $dV = s\, ds\, d\varphi\, dz$.

A uniformly charged sphere of radius $3\,\mathrm{cm}$ contains only $1/27$ of its charge inside $1\,\mathrm{cm}$, since the volume integral scales as r^3. A detector viewing a source through a polar cap of half-angle 60° subtends $2\pi(1-\cos 60^\circ) = \pi$ steradians. It collects one quarter of an isotropic source's total power. A dipole weights the same solid-angle elements unequally. Integrate the dipole radiation pattern.

Infer a local divergence from flux

Across the two opposite x-faces of a small rectangular box, the net outward flux of $\boldsymbol{F}$ is approximately $[F_x(x+dx) - F_x(x)]\, dy\, dz$. Dividing by the volume and adding the other face pairs gives $\nabla \cdot \boldsymbol{F}$. Internal faces cancel when boxes are joined. The resulting identity is $\int_V \nabla \cdot \boldsymbol{F}\, dV = \oint_{\partial V} \boldsymbol{F} \cdot \boldsymbol{n}\, dA$ with outward normal. Applied to $\boldsymbol{D}$, the flux counts free charge.

In a cube of side $0.10\,\mathrm{m}$ let $D_x = (2.0\,\mu\mathrm{C/m^3})x$ and other components vanish. The divergence is $2.0\,\mu\mathrm{C/m^3}$, so the cube contains 2.0 nC of free charge. Directly, the x-faces differ by $0.20\,\mu\mathrm{C/m^2}$ over $0.010\,\mathrm{m^2}$, giving the same result. Moving the cube does not change the difference here. A single point's field magnitude alone would not give the enclosed charge.

Follow a loop with an oriented normal

A small loop in the xy-plane, traversed counterclockwise as seen from $+z$, has circulation per area $\partial_x F_y - \partial_y F_x = (\nabla \times \boldsymbol{F})_z$. Summing loops that tile a surface cancels interior edges, leaving $\oint_{\partial S} \boldsymbol{F} \cdot d\boldsymbol{\ell} = \int_S (\nabla \times \boldsymbol{F}) \cdot \boldsymbol{n}\, dA$. Curl measures circulation density. A field can have a large magnitude parallel to a straight wire while having zero circulation around a distant closed loop.

Take a stationary loop of area $0.020\,\mathrm{m^2}$, with uniform normal field B_z growing at $0.30\,\mathrm{T/s}$. Faraday's law gives $\oint \boldsymbol{E} \cdot d\boldsymbol{\ell} = -6.0\,\mathrm{mV}$ in the stated counterclockwise orientation. Reverse the normal and path together; the reported circulation changes sign. A moving conducting loop also requires the magnetic force per charge $\boldsymbol{v} \times \boldsymbol{B}$ in its measured emf. Work through a moving bar.

Keep point and sheet sources when differentiating

For $r > 0$,

$$\nabla \cdot (\hat{\boldsymbol{r}}/r^2) = r^{-2}\partial_r (r^2/r^2) = 0.$$

But every sphere enclosing the origin has outward flux 4π. A volume source must preserve that finite integral even as the sphere shrinks. Distributionally,

$$\nabla \cdot (\hat{\boldsymbol{r}}/r^2) = 4\pi \delta^3(\boldsymbol{r}).$$

The delta has units $\mathrm{m^{-3}}$ and obeys $\int f(\boldsymbol{r})\delta^3(\boldsymbol{r}-\boldsymbol{a}) dV = f(\boldsymbol{a})$ for smooth test functions.

A sheet at $z = 0$ with density $\sigma = 2.0\,\mathrm{nC/m^2}$ instead has volume density $\rho = \sigma\delta(z)$; integrating through the sheet over $0.40\,\mathrm{m}^2$ gives 0.80 nC. A finite-volume model that smears the sheet over thickness h needs $\rho \simeq \sigma/h$. Refining h doubles the peak if the thickness halves, while its integral stays fixed. Derive the source term from a test function.

Choose the boundary condition that matches the apparatus

Gauss's law and electrostatics give $-\nabla \cdot (\epsilon \nabla \phi) = \rho_{\mathrm{free}}$ in a linear dielectric. If ϵ varies, leaving it inside the divergence keeps the correct flux. A voltage source supplies Dirichlet data $\phi = g$. Specified normal electric displacement supplies Neumann data $-\epsilon \partial_n \phi = D_n$ for scalar ϵ. Prescribing normal flux everywhere requires

$$\oint D_n dA = Q_{\mathrm{free}}.$$

A compatible pure Neumann problem fixes the potential up to a constant in a connected region. An isolated neutral conductor is equipotential at an unknown voltage, with its total charge constrained; a grounded conductor instead has a specified voltage and can exchange charge.

To test uniqueness with fixed electrode voltages, subtract two candidate solutions for the same source. Their difference u satisfies $\nabla \cdot (\epsilon \nabla u) = 0$ and vanishes on the electrodes. Integrating $u \nabla \cdot (\epsilon \nabla u)$ by parts gives $\int \epsilon |\nabla u|^2 dV = 0$ if the boundary term vanishes everywhere. For positive ϵ the difference is constant and at a fixed electrode it is zero. Test a grounded-domain solution on a grid. Compare a floating electrode with ground.

Turn a point response into an extended-source solution

Let G solve $-\nabla^2 G(\boldsymbol{r}, \boldsymbol{r}') = \delta^3(\boldsymbol{r} - \boldsymbol{r}')$ with homogeneous boundary conditions. In vacuum, integrating weighted responses gives

$$\phi(\boldsymbol{r}) = \epsilon_0^{-1} \int G(\boldsymbol{r}, \boldsymbol{r}') \rho(\boldsymbol{r}')\, dV'.$$

Applying $-\nabla^2$ under the integral recovers ρ/ϵ_0. Nonzero prescribed boundary voltages supply an additional source-free solution. In free space $G = 1/(4\pi|\boldsymbol{r} - \boldsymbol{r}'|)$, with units of inverse length.

Put a point charge at $z = a > 0$ above a grounded infinite plane. Subtract the free-space response of a reflected charge at $-a$. On the plane both distances agree, so the potential vanishes there; in the accessible half-space only the actual charge supplies a delta source. Uniqueness justifies the image construction. For $q = 1.0\,\mathrm{nC}$ and $a = 1.0\,\mathrm{cm}$, the force from the image has magnitude $q^2/[4\pi\epsilon_0(2a)^2] = 22.5\,\mu\mathrm{N}$, directed towards the plane. Construct the bounded Green function and its reciprocity.

Apply a pillbox and a ribbon at a material interface

Point a unit normal $\boldsymbol{n}$ from medium 1 into medium 2. A shrinking Gaussian pillbox leaves its two faces and gives

$$\boldsymbol{n} \cdot (\boldsymbol{D}_2 - \boldsymbol{D}_1) = \sigma_f,$$

$$\boldsymbol{n} \cdot (\boldsymbol{B}_2 - \boldsymbol{B}_1) = 0.$$

A narrow rectangular integration ribbon leaves the two tangential segments, giving $\boldsymbol{n} \times (\boldsymbol{E}_2 - \boldsymbol{E}_1) = 0$ for finite magnetic fields and $\boldsymbol{n} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{K}_f$ for a free sheet current and finite displacement field.

At a charge-free interface from vacuum to a linear dielectric with $\epsilon_r = 4$, suppose the normal and tangential components of $\boldsymbol{E}_1$ are 800 and 300 V/m. Then $E_{2n} = 200\,\mathrm{V/m}$, while $E_{2t} = 300\,\mathrm{V/m}$. The normal displacement is $7.08\,\mathrm{nC/m^2}$ on both sides; the total field angle changes. A sheet of bound charge is consistent with these conditions because the pillbox for $\boldsymbol{D}$ counts only free charge. Work through polarisation and free charge.

Use a conservation law to identify a missing current

If charge leaves a fixed region through its surface at a positive rate, the charge inside must decrease. This gives $\partial_t \rho + \nabla \cdot \boldsymbol{J} = 0$ after applying the divergence theorem. Taking the divergence of $\nabla \times \boldsymbol{H} = \boldsymbol{J}_f + \partial_t \boldsymbol{D}$ recovers the same statement: the divergence of a curl vanishes and $\nabla \cdot \boldsymbol{D} = \rho_f$. The displacement-current term makes the field equation compatible with charging an insulator-bounded region.

A current of $2.0\,\mathrm{mA}$ charges large capacitor plates of area $5.0 \times 10^{-3}\,\mathrm{m^2}$. With negligible fringing, $A\,\partial_t D_n = I$, hence $\partial_t D_n = 0.40\,\mathrm{A/m^2}$. In a vacuum gap this means $\partial_t E_n = 4.52 \times 10^{10}\,\mathrm{V/(m\,s)}$. The conduction current crosses a wire section and the displacement current crosses a surface through the gap. Either spanning surface predicts the same magnetic circulation around its common rim. Follow the charging capacitor through space and time.

Reconstruct fields from potentials and test a gauge change

Locally, $\nabla \cdot \boldsymbol{B} = 0$ allows $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. Faraday's law then makes $\boldsymbol{E} + \partial_t \boldsymbol{A}$ curl free and locally expressible as $-\nabla\phi$. In simply connected regions the scalar potential for this combination is single-valued. Under $\boldsymbol{A}' = \boldsymbol{A} + \nabla\chi,\ \phi' = \phi - \partial_t\chi$, the added terms in $\boldsymbol{E}$ cancel by equality of mixed derivatives; the curl of the added gradient in $\boldsymbol{B}$ is zero.

For a static uniform $B_0\hat{\boldsymbol{z}}$, use either $\boldsymbol{A}_1 = B_0 x\hat{\boldsymbol{y}}$ or $\boldsymbol{A}_2 = B_0(-y\hat{\boldsymbol{x}} + x\hat{\boldsymbol{y}})/2$. Their difference is the gradient of $\chi = -B_0 xy/2$. A rectangular loop of area $0.050\,\mathrm{m^2}$ normal to $B_0 = 0.20\,\mathrm{T}$ has $\oint \boldsymbol{A} \cdot d\boldsymbol{\ell} = 0.010\,\mathrm{Wb}$ in either gauge. Follow the gauge choice into time-dependent fields.

Use the double curl to choose between waves and diffusion

The identity $\nabla \times (\nabla \times \boldsymbol{E}) = \nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E}$ connects curl measurements to a spatial operator acting on each Cartesian component. In a uniform source-free dielectric with $\boldsymbol{J} = \sigma \boldsymbol{E}$ and $\nabla \cdot \boldsymbol{E} = 0$, combine Faraday and Ampère–Maxwell to obtain

$$\nabla^2 \boldsymbol{E} = \mu\sigma \partial_t \boldsymbol{E} + \mu\epsilon \partial_t^2 \boldsymbol{E}.$$

Setting $\sigma = 0$ gives wave speed $1/\sqrt{\mu\epsilon}$. For harmonic fields the ratio of displacement to conduction currents is $\omega\epsilon/\sigma$; when small, dropping the second time derivative gives magnetic diffusivity $1/(\mu\sigma)$.

At $1.0\,\mathrm{kHz}$, a vacuum wavelength is roughly 300 km. In a good conductor with $\mu \simeq \mu_0$ and $\sigma = 5.8 \times 10^7\,\mathrm{S/m}$, the harmonic skin depth is only $\sqrt{2/(\omega\mu\sigma)} = 2.09\,\mathrm{mm}$. Inserting the vacuum wavelength as a penetration length in the conductor would miss the dominant term in its field equation. Compare a field step with a harmonic field.

Close an energy balance over a chosen volume

Dot the Ampère–Maxwell equation with $\boldsymbol{E}$, dot Faraday's law with $\boldsymbol{H}$, and use $\nabla \cdot (\boldsymbol{E} \times \boldsymbol{H}) = \boldsymbol{H} \cdot \nabla \times \boldsymbol{E} - \boldsymbol{E} \cdot \nabla \times \boldsymbol{H}$. For linear, time-independent, nondispersive media with symmetric constitutive tensors this gives

$$\partial_t u + \nabla \cdot \boldsymbol{S} = -\boldsymbol{J}_f \cdot \boldsymbol{E},$$

with $u = (\boldsymbol{E} \cdot \boldsymbol{D} + \boldsymbol{B} \cdot \boldsymbol{H})/2$ and $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$. In vacuum the same formula holds directly. In dispersive matter the energy stored in the material's dynamical response must also be accounted for.

Suppose 3.0 W flows into a stationary resistor enclosure and 0.20 W leaves by other electromagnetic paths. During a transient the field energy inside rises at $0.30\,\mathrm{J/s}$. The remaining $2.50\,\mathrm{W}$ is delivered to matter. With a uniform Ohmic field $E = 10\,\mathrm{V/m}$, conductivity $5\,\mathrm{S/m}$ and resistor volume $0.005\,\mathrm{m}^3$, the independent volume integral $\sigma E^2 V$ gives that same 2.50 W. Derive energy transport from the fields.

Average a real product before interpreting a complex power

Choose peak phasors and write $\boldsymbol{E}(t) = \mathrm{Re}[\widetilde{\boldsymbol{E}} e^{-i\omega t}]$, so ∂_t acts as $-i\omega$. A forward attenuating wave can then be written $e^{ikz - i\omega t}$ with positive $\mathrm{Im}\, k$. Some chapters instead state $e^{+i\omega t}$; conjugate the phasors and the spatial exponential consistently when switching conventions. Time-averaged physical powers have the same sign. Multiplying the two real fields produces constant and double-frequency terms; averaging removes the latter, giving $\langle \boldsymbol{S} \rangle = \frac{1}{2}\mathrm{Re}(\widetilde{\boldsymbol{E}} \times \widetilde{\boldsymbol{H}}^*)$. Using rms phasors absorbs the factor one half.

With orthogonal transverse peak amplitudes $E = 10\,\mathrm{V/m}$, $H = 0.040\,\mathrm{A/m}$ and a relative temporal phase of 60°, the mean flux magnitude along their cross product is $0.5(10)(0.040)\cos 60^\circ = 0.100\,\mathrm{W/m^2}$. A phase difference of 90° gives zero mean power even when both local amplitudes are large. Compare average and reactive power in a conductor.

Check two observers' fields without transforming every component

Fix SI components and the metric $(+,-,-,-)$, with $x^0 = ct$. The field combinations $I_1 = B^2 - E^2/c^2$ and $I_2 = \boldsymbol{E}\cdot\boldsymbol{B}$ have the same values under proper Lorentz transformations. The first follows from the tensor contraction $F_{\mu\nu}F^{\mu\nu} = 2I_1$; the second is proportional to the contraction with the dual tensor and changes sign under a spatial reflection. The factor $1/c^2$ converts squared electric-field units into squared tesla. A plane wave has $E = cB$ and perpendicular fields, giving two zero invariants even though it carries energy.

Take $\boldsymbol{E} = (3.0\times10^{5}, 0, 0)\,\mathrm{V/m}$ and $\boldsymbol{B} = (0, 0, 2.0\times10^{-3})\,\mathrm{T}$. Then $I_1 \simeq 3.0\times10^{-6}\,\mathrm{T}^2 > 0,\ I_2 = 0$. For an observer moving at $\boldsymbol{v} = \boldsymbol{E}\times\boldsymbol{B}/B^2$, the perpendicular transformation gives $\boldsymbol{E}' = \gamma(\boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B}) = 0$. The required speed is $1.5\times10^{8}\,\mathrm{m/s} < c$, along $-y$. From the invariant, $B' = \sqrt{I_1} \simeq 1.73\,\mathrm{mT}$. Both observers can therefore check the same physical example with different field readings. Compare measurements at one event. Derive the tensor contractions.

Turn a derivation into a solver check

Choose a case with an exact result and use the same units, geometry, source and boundary data in the numerical problem. For a grounded square of side L, choose $\rho = \rho_0 \sin(\pi x/L)\sin(\pi y/L)$. Applying $-\nabla^2$ to the same sine product multiplies it by $2\pi^2/L^2$. Therefore $\phi_{\mathrm{centre}} = \rho_0 L^2/(2\pi^2\epsilon_0)$. With $L = 1\,\mathrm{m},\ \rho_0 = 2.0\,\mathrm{nC/m^3}$, the centre voltage is $11.4433\,\mathrm{V}$.

A square mesh with N intervals per side has $h = L/N$. The five-point negative Laplacian acting on that sampled sine has eigenvalue $8\sin^2[\pi/(2N)]/h^2$. For $N = 12$ it predicts $11.5089\,\mathrm{V}$; the discrete equations can be solved to roundoff and still differ from the continuum by $0.0656\,\mathrm{V}$. Doubling N should reduce this nodal error by close to four because $\sin x = x - x^3/6 + \cdots$. Compare the calculated value, the source scaling and the refinement rate. These three observations distinguish a wrong sign or unit from a converged solution on a coarse mesh. Run the square-domain experiment and its OneSimulation example.

Notation and assumptions

Read each defining equation with its assumptions and units, then follow the link to the argument that gives the equation its meaning.

Electric displacement

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$$

Its flux measures free charge. Polarisation is dipole moment per volume. A discontinuity of normal displacement depends on free surface charge.

Derivation · How microscopic dipoles produce macroscopic charge

Maxwell capacitance coefficient

$$Q_i = \sum_j C_{ij} V_j$$

Potentials are defined relative to a stated reference. A floating electrode has a prescribed total charge and a potential determined by the other electrodes.

Derivation · Capacitance as a matrix of electrostatic energies

Potential and electric field

$$\mathbf{E} = -\nabla V$$

A volt per metre is a newton per coulomb. This relation describes an electrostatic field. Time-dependent induction introduces a vector-potential term.

Derivation · Charge, potential and field energy

Poynting flux

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}$$

Integrating its outward normal component gives field power leaving a closed surface. Work on matter and changing stored energy complete the balance.

Derivation · Electromagnetic waves and energy flux

Reciprocity

$$\int \rho_1 V_2 \, dV = \int \rho_2 V_1 \, dV$$

The displayed form assumes compatible homogeneous boundary conditions and a reciprocal linear medium. Boundary work terms matter for separately driven electrodes.

Derivation · Building a potential from point responses

Retarded time

$$t - t_r = |\mathbf{r} - \mathbf{r}_s(t_r)|/c$$

The source is evaluated at emission. For a moving source its location on the right also depends on the unknown emission time.

Derivation · Potentials, gauge freedom and the arrival of a signal

Skin depth

$$\delta = \sqrt{2/(\mu\sigma\omega)}$$

In a good homogeneous conductor under sinusoidal excitation this is the amplitude decay length. Intensity decays twice as fast in the exponent.

Derivation · How a conductor admits a field step and a sinusoid

Wave impedance

$$Z = E_t/H_t$$

The ratio has units of ohms. The field direction and mode determine the relevant tangential components. In a guide it can differ from the intrinsic impedance of the filling material.

Derivation · Waves in a hollow guide

Further reading

- Follow sections 10.5–10.7 from the field laws to diffusion and skin depth. Check which timescale changes if you double the conductor thickness. MIT · Haus and Melcher, Magnetic diffusion
- Use the electrostatics and electromagnetism chapters to derive the interface conditions from the integral laws. Compare the roles of fields, sources and potentials in the same boundary-value problem. David Tong · Electromagnetism

Compare specific topics with the university sources