Condensed matter and band theory

Read a lattice from diffraction, derive its vibration and electronic bands, and use electronic transport, magnetic response and spectroscopy to identify the state of a solid.

Read the subject in depth

Each volume develops its arguments in sequence, with worked problems and links to experiments.

- ### Structure, excitations and probes of solids

  Connect microscopic symmetry and dynamics to a measurement of a real solid. Develop the derivation and identify which experimental assumptions are needed to interpret the signal.

Advanced study across the subject

Each sequence connects the course to detailed chapters in neighbouring subjects. The complete course PDF includes their derivations, examples and exercises.

1. ### Magnetic and electrical material measurements

   Follow a microscopic carrier or bound current into a measured Hall voltage, external field, magnetic response or penetration depth.

2. ### Fluctuations, occupation and collective order

   Connect an occupation probability to a heat capacity, a correlation length or a symmetry-breaking order parameter. Compare electrical measurements with a thermal fluctuation.

3. ### Periodic potentials, interfaces and experimental probes

   Follow a state across a boundary, a band gap or an avoided crossing. Use optical, tunnelling and calorimetric probes to distinguish an energy gap from a loss peak.

How the course fits together

1. **Chapters 1–4**

   Real-space periodicity selects reciprocal lattice vectors. Apply the same translation symmetry to displacements and electron states to find phonon modes, allowed electronic states and Bloch bands.

2. **Chapters 5–8**

   A chemical potential sets which bands carry charge. A periodic dielectric gives a second, optically measurable band problem. Scattering and magnetic fields then expose different properties of mobile electrons.

3. **Chapters 9–12**

   Move from individual magnetic moments to cooperative order. Screening and spectroscopic excitations offer distinct experimental ways of detecting the superconducting state and testing a

pairing model.

Contents

1. Chapter 1 · Translational symmetry and diffraction
2. Chapter 2 · Vibrations of monoatomic and diatomic chains
3. Problem sheet 1 · Identifying a lattice and its mechanical modes
4. Chapter 3 · Filling electronic states and measuring the Fermi surface
5. Chapter 4 · Bloch waves, band edges and effective mass
6. Problem sheet 2 · Occupied states, heat capacity and band curvature
7. Chapter 5 · Carriers, neutrality and transport in semiconductors
8. Chapter 6 · Predicting a gap and testing it in a periodic medium
9. Problem sheet 3 · Distinguishing carriers, reflection and attenuation
10. Chapter 7 · Transport, scattering and the mean free path
11. Chapter 8 · Cyclotron motion and quantised magnetic orbits
12. Problem sheet 4 · Transport lifetimes and magnetic orbits
13. Chapter 9 · Magnetic moments and susceptibility
14. Chapter 10 · Exchange, spontaneous magnetisation and critical response
15. Problem sheet 5 · Moments, susceptibility and collective order
16. Chapter 11 · Superconductivity, magnetic screening and flux quantisation
17. Chapter 12 · Pairing, excitation gaps and what a spectrum measures
18. Problem sheet 6 · Screening, spectra and the meaning of a gap
19. Chapter 13 · How alternating bonds split a one-dimensional band
20. Chapter 14 · Finding an end state one bond at a time
21. Chapter 15 · Why a honeycomb lattice has Dirac cones
22. Chapter 16 · Counting carriers and weighing a cyclotron orbit
23. Chapter 17 · Quantising the orbit and finding the zero Landau level
24. Chapter 18 · How the same electrons carry heat and charge
25. Chapter 19 · Why a polar crystal has two optical phonon frequencies
26. Chapter 20 · How one light isotope traps a lattice vibration
27. Problem sheet 7 · Inferring a chain termination from bulk bands and end-site data
28. Problem sheet 8 · Why a crystal with three sound speeds has one low-temperature heat coefficient
29. Structure, excitations and probes of solids
30. Magnetic and electrical material measurements
31. Fluctuations, occupation and collective order
32. Periodic potentials, interfaces and experimental probes
33. Applied study
34. Notation and assumptions
35. Further reading

Chapter 1 · Translational symmetry and diffraction

Work backwards from a scattering pattern to the repeating translations and the atoms within one cell. Learn which structural information an intensity measurement loses.

FROM FIRST PRINCIPLES

Recover a lattice spacing from phase addition

Illuminate a row of N equally spaced atoms with a coherent plane wave and observe weak elastic scattering far away. Each atom sees an incident phase and contributes an outgoing phase. Their difference depends on $\mathbf{q} = \mathbf{k}_{\text{out}} - \mathbf{k}_{\text{in}}$. With the origin at the first atom, let $\phi = \mathbf{q} \cdot \mathbf{a}$ be the phase difference between neighbours displaced by $\mathbf{a}$. Up to an overall sign convention the total amplitude is a geometric sum:

$$A = f \sum_{j=0}^{N-1} e^{ij\phi} = f \, \frac{1 - e^{iN\phi}}{1 - e^{i\phi}}.$$

Here f is the amplitude from one atom. Use $|1 - e^{ix}| = 2|\sin(x/2)|$ and take the squared magnitude. When $\phi = 2\pi m$, evaluate the original sum directly:

$$I \propto |f|^2 \frac{\sin^2(N\phi/2)}{\sin^2(\phi/2)},$$

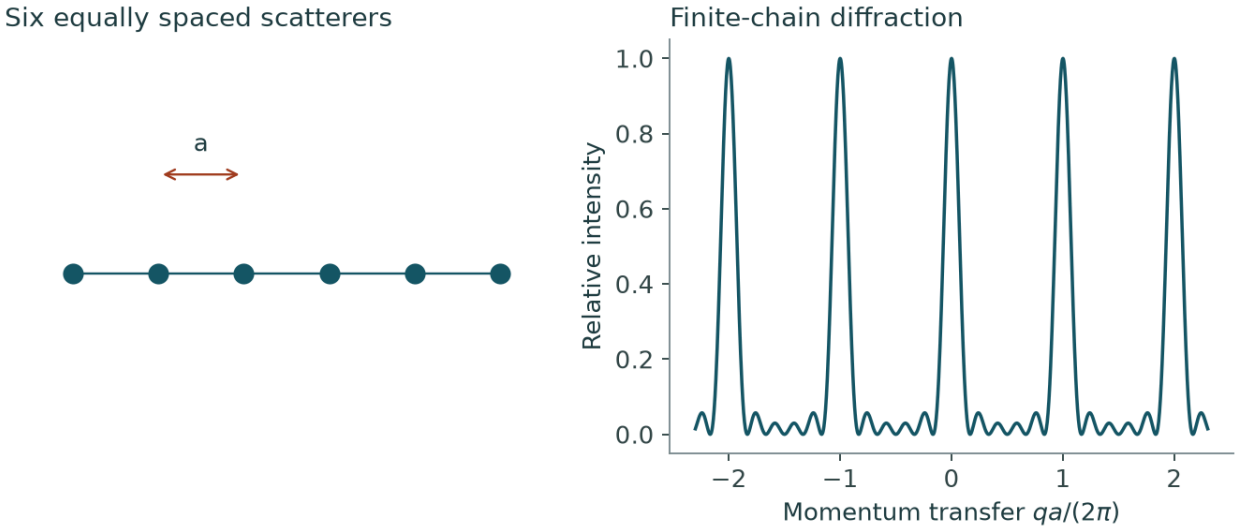
$$I_{\text{peak}} \propto N^2 |f|^2.$$

For a three-dimensional lattice every primitive translation must meet this phase condition. The resulting peak positions identify reciprocal-lattice vectors. For several atoms in each cell the single-cell factor becomes $\sum_s f_s e^{i\mathbf{q}\cdot\mathbf{r}_s}$. Its squared magnitude can vanish even at a permitted reciprocal vector. An absent peak can therefore reveal interference inside the cell.

Background · Reciprocal lattice

$$\mathbf{G} \cdot \mathbf{R} = 2\pi n$$

A reciprocal-lattice vector leaves the phase unchanged under every direct-lattice translation. A basis of different atoms can still suppress a reflection.



Why a finite crystal gives diffraction peaks of nonzero width

Successive sites acquire a relative phase qa. At qa = 2π every contribution returns in phase. The first zero on either side is offset by 2π/6 because six equally spaced phase vectors sum to zero. Doubling the number of sites halves this offset while leaving the peak spacing fixed.

Explore the related parameter study beside the computed figure.

DERIVATION AND APPLICATIONS

Extract atomic displacement from the loss of Bragg intensity

A lattice can keep the same mean spacing while its Bragg peaks lose intensity on heating. Consider a crystal of identical atoms whose displacements about their mean positions fluctuate during the exposure. Treat scattering in the single-scattering approximation. At a reciprocal vector of magnitude Q, every mean lattice position supplies the same phase. The instantaneous displacement projected along the scattering vector, u, adds a phase Qu. If this component is Gaussian with mean zero and variance s^2, averaging the complex amplitude gives

$$\left\langle e^{iQu} \right\rangle = \frac{1}{\sqrt{2\pi s^2}} \int_{-\infty}^{\infty} e^{-u^2/(2s^2)} e^{iQu}\, du = e^{-Q^2 s^2/2}.$$

The last equality follows by Fourier transformation of the normalized Gaussian. The coherent peak measures the square of this average amplitude. Squaring the instantaneous phase factor first would give unity and erase the effect of phase differences between sites. With $I_0(Q)$ the coherent intensity of the perfectly stationary array, the Debye–Waller attenuation is

$$I_{\mathrm{Bragg}}(Q) = I_0(Q) e^{-Q^2 s^2}.$$

The formula describes the sharp coherent contribution. The difference between the total averaged intensity and the coherent intensity appears in diffuse scattering. Stronger attenuation at large Q follows because an identical displacement spans a larger fraction of a scattering phase cycle. For corrected integrated intensities of the same reflection at two temperatures, taken at essentially the same Q, the unknown stationary intensity cancels,

$$s_2^2 - s_1^2 = -\frac{\ln[I_2(Q)/I_1(Q)]}{Q^2}.$$

This comparison assumes unchanged occupancies, instrument normalization and negligible changes in the mean structure. For isotropic displacement s^2 is one third of the mean squared three-dimensional displacement.

WORKED EXAMPLE

Recover a change in mean-square displacement

Change the inputs and investigate this derivation.

A reflection at $Q = 50.0\,\mathrm{nm^{-1}}$ falls from 10,000 to 8,000 corrected counts. Its projected variance increases by $-\ln(0.800)/2500 = 8.93 \times 10^{-5}\,\mathrm{nm^2}$. If its original projected root-mean-square displacement was $0.00600\,\mathrm{nm}$, the new value is

$$\sqrt{0.00600^2 + 0.0000893} = 0.0112\,\mathrm{nm}.$$

Adding the displacements directly would use the wrong measured quantity.

SECOND APPLICATION

Test the angular dependence with another reflection

Change the inputs and investigate this derivation.

A second specimen retains 90.0 percent of its low-temperature peak intensity at $Q = 30.0\,\mathrm{nm^{-1}}$. For the same displacement direction at twice that wavevector, the attenuation exponent quadruples. The predicted retention is

$$0.900^4 = 0.656.$$

A corrected retention near 0.90 for both reflections would instead suggest a common change in illuminated volume or flux that this displacement model cannot explain.

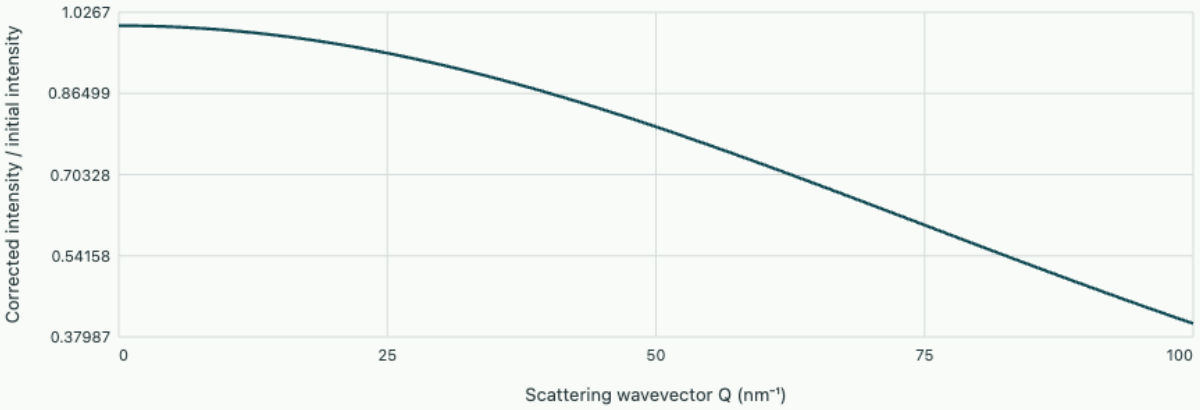
Try it · Debye waller

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

Projected variance increase (nm²) **9.00e-5**



Gaussian displacement attenuation

Intensity ratio is exp(−Q²Δ⟨u²⟩). Doubling Q quadruples the attenuation exponent for the same projected displacement distribution.

What repeats in a crystal?

A one-dimensional chain with identical atoms every a has translations $R = na$. Replacing every other atom by a different species doubles the primitive translation, even if the nearest-neighbour separation stays fixed. The set of equivalent cell origins is the Bravais lattice; the atoms placed relative to each origin form the basis. Choosing a larger unit cell is allowed but folds distinct waves onto extra branches. Count independent degrees of freedom per unit length as a check on any cell choice.

In three dimensions write $\mathbf{R} = \sum_i n_i \mathbf{a}_i$. Define reciprocal vectors by $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$. Then $\mathbf{G} = \sum_i h_i \mathbf{b}_i$ has $e^{i\mathbf{G}\cdot\mathbf{R}} = 1$ at every equivalent origin. These vectors have units of inverse length. For a rectangular lattice doubling a real-space period halves the corresponding reciprocal spacing. The first Brillouin zone consists of wavevectors closer to the reciprocal origin than to any other reciprocal point. Later we will label collective vibrations and electronic states inside this region.

Construct reciprocal space from the primitive cell

A diffraction experiment measures momentum transfer. To predict its possible values, ask which plane waves acquire the same phase at every translated lattice point. For a Bravais lattice with translations $\mathbf{R} = \sum_i n_i \mathbf{a}_i$, the condition is $\exp(i\mathbf{G}\cdot\mathbf{R}) = 1$ for every integer choice of n_i. It suffices to require

$$\mathbf{G}\cdot\mathbf{a}_i = 2\pi h_i.$$

Collect the primitive direct-lattice vectors as columns of A and their reciprocal vectors as columns of B. The defining dot products become one matrix equation. Solving it gives

$$\begin{aligned} A^{\mathsf{T}}B &= 2\pi I, \\ B &= 2\pi (A^{-1})^{\mathsf{T}}. \end{aligned}$$

Each column of B has units of inverse length. Its dot product with its own direct-lattice vector is 2π and with every other basis vector is zero. Integer combinations generate the allowed reciprocal vectors. The transpose matters whenever the direct axes are oblique.

For a triangular layer take $\mathbf{a}_1 = (a, 0)$ and $\mathbf{a}_2 = (a/2, \sqrt{3}a/2)$. The constraints on $\mathbf{b}_1$ first give $b_{1x} = 2\pi/a$ and then $b_{1y} = -2\pi/(\sqrt{3}a)$. Similarly $\mathbf{b}_2 = (0, 4\pi/(\sqrt{3}a))$. Verify that translating along $\mathbf{a}_2$ leaves the phase of $\mathbf{b}_1$ unchanged. Reciprocal vectors have units of inverse length.

How far can a wave vector move from the origin before an equivalent representative is nearer? The boundary shared with $\mathbf{G}$ solves $|\mathbf{k}| = |\mathbf{k} - \mathbf{G}|$, or $\mathbf{k}\cdot\mathbf{G} = |\mathbf{G}|^2/2$. The half-planes around the origin define the first Brillouin zone. Here the six shortest reciprocal vectors give a hexagon. With $a = 0.300\,\mathrm{nm}$ the shortest $|\mathbf{G}|$ is $4\pi/(\sqrt{3}a) = 24.18\,\mathrm{nm}^{-1}$, putting a face midpoint $12.09\,\mathrm{nm}^{-1}$ from the centre. The zone area is $(2\pi)^2/(\sqrt{3}a^2/2) = 506.5\,\mathrm{nm}^{-2}$, equal to one reciprocal primitive-cell area. Doubling all lattice lengths halves the distances in the diffraction pattern and quarters its zone area. These reciprocal positions follow from translation symmetry. The measured intensities also depend on what occupies each cell.

Practise on sheet 1 · Fold a wave vector into the first zone. Further reading · University of Cambridge · David Tong, Solid State Physics · Chapter 2, Band Structure, §2.2.2 (construction of the reciprocal lattice).

Add amplitudes before taking intensities

An elastic plane wave scattered from positions $na + d_j$ accumulates a phase equal to the momentum-transfer wavevector Q times position. For N identical cells the amplitude factorises:

$$A(Q) = \underbrace{\sum_j f_j(Q) e^{iQd_j}}_{F(Q)} \sum_{n=0}^{N-1} e^{iQna},$$

$$I(Q) \propto |A(Q)|^2.$$

The last sum is geometric. Its magnitude is $|\sin(NQa/2)/\sin(Qa/2)|$, with limiting value N at $Qa = 2\pi h$. Large finite crystals have narrow peaks of height proportional to N^2. The separation to the adjacent first zero is $2\pi/(Na)$; doubling coherent length halves this width. The basis factor F sets the relative integrated strengths for a given lattice. A finite instrument resolution also broadens measured peaks and should be characterised before inferring a sample size.

Background · Why are diffraction peaks from small crystals broad?

Along one direction take N equally spaced scattering planes at $x_n = na$. Near a reciprocal-lattice vector G, an additional wavevector mismatch Δq accumulates phase $na\Delta q$ at the nth plane. Summing the complex contributions gives a geometric series.

$$\frac{I(G+\Delta q)}{I(G)} = \left[\frac{\sin(Na\Delta q/2)}{N\sin(a\Delta q/2)}\right]^2.$$

The nearest zero lies at $2\pi/(Na)$. A coherently scattering length of 100 nm therefore produces a first-zero offset near 0.063 nm^{-1}. A finite width alone does not determine a lattice strain. Finite size sets a scale in absolute wavevector, whereas a distribution of fractional lattice spacings tends to give a width proportional to the reflection wavevector. Comparing several orders helps separate the two mechanisms.

An extinction is evidence about the basis

Place two sites at $d = 0, a/2$, with real scattering amplitudes $f_A = 3$ and $f_B = 1$ held constant over the angular range. At $G_h = 2\pi h/a$, $F_h = 3 + (-1)^h$. Even and odd orders have intensities proportional to 16 and 4 respectively, before angular or instrumental corrections. If both amplitudes become equal, the odd peaks vanish. If the two sites have identical chemistry as well as identical positions up to translation, a cell of length $a/2$ would describe the structure more economically. Equal scattering amplitudes in a particular probe alone cannot establish chemical identity.

For incident wavelength λ, with the total scattering angle 2θ,

$$|Q| = (4\pi/\lambda)\sin\theta.$$

At $\lambda = 0.154$ nm and spacing $d = 0.200$ nm the first Bragg peak therefore occurs at $\theta = 22.6^\circ$, or detector angle 45.3°. A factor-of-two error in reading the detector would corrupt any inferred lattice constant.

What survives a change of origin?

Translating the entire basis by d_0 multiplies $A(Q)$ by e^{iQd_0}. Intensity is unchanged, so diffraction cannot locate the absolute origin. More generally the phase of the amplitude is missing from a measurement of $|A|^2$. Peak positions suggest candidate lattices, corrected intensities test a proposed basis and peak

widths probe finite coherence or strain. These are different questions even when addressed with the same scan.

Derive the selection rules of a cubic powder pattern

A powdered crystal supplies many crystallite orientations, so a fixed detector angle can intercept every family of reciprocal planes with the appropriate spacing. For a cubic lattice with conventional cell parameter a, reciprocal vectors have magnitude $2\pi\sqrt{h^2+k^2+l^2}/a$. Bragg's condition then gives

$$\sin^2\theta = \lambda^2(h^2+k^2+l^2)/(4a^2).$$

The sequence of squared sines can be used to index the pattern, but only reciprocal points with nonzero structure amplitude produce peaks.

For a monatomic face-centred cubic crystal, describe the lattice with conventional cubic coordinates and basis positions $(0,0,0)$, $(0,\tfrac12,\tfrac12)$, $(\tfrac12,0,\tfrac12)$ and $(\tfrac12,\tfrac12,0)$. For identical atomic amplitude f their phase sum is

$$F_{hkl} = f[1+(-1)^{k+l}+(-1)^{h+l}+(-1)^{h+k}].$$

If all indices are even or all are odd, every phase factor is positive and $F = 4f$. Mixed parity makes the sum zero. Thus the first allowed squared-index sums are 3 for (111), 4 for (200), 8 for (220), 11 for (311) and 12 for (222). These extinctions encode the centring translations; adding an internal basis to each lattice site can introduce additional missing peaks. Powder intensity also depends on multiplicity and on angle-dependent scattering factors, so the strongest peak need not have the smallest indices.

Infer a lattice parameter and a strain

Suppose powder peaks under wavelength 0.1540 nm occur at detector angles $2\theta = 38.48^\circ$, 44.74° and 65.10°. Their squared half-angle sines are about 0.1086, 0.1448 and 0.2896. After division by the first they give $1 : 1.33 : 2.67$, consistent with $3 : 4 : 8$. Indexing the middle peak as (200) yields

$$a = \lambda/\sin 22.37^\circ = 0.405\,\mathrm{nm}.$$

The other two peaks independently give the same value to the displayed precision, unlike an arbitrary assignment to successive integers.

Now let the (200) detector angle fall by 0.100° under load. Differentiating $2d\sin\theta = \lambda$ gives

$$\Delta d/d \simeq -\cot\theta\,\Delta\theta.$$

Here $\Delta\theta = -0.0500^\circ = -8.73\times10^{-4}$ rad, hence

$$\Delta d/d = +2.12\times10^{-3}.$$

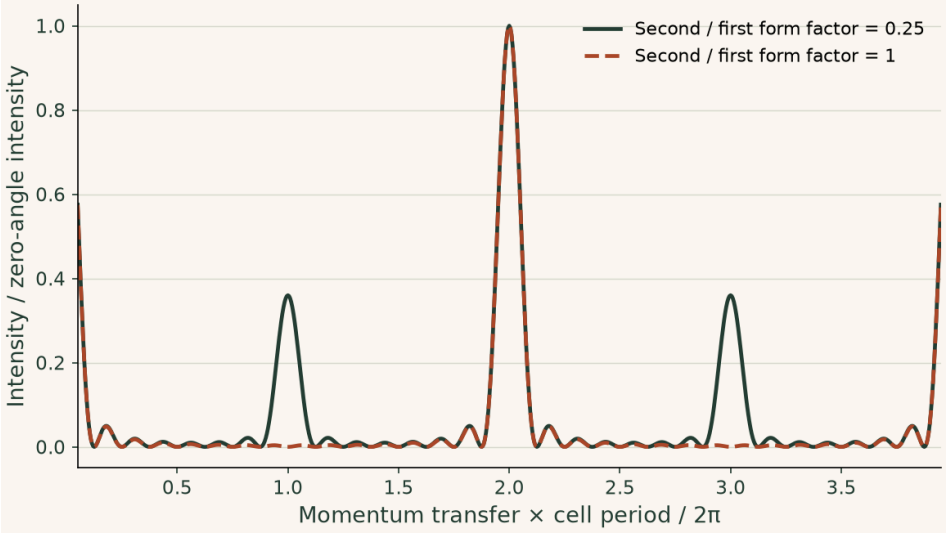
That is tensile strain along the scattering vector. A single peak shift cannot establish an isotropic volume strain without further orientations. Multiple peaks can also separate a uniform scale error in wavelength from a genuine directional strain. In the unstrained limit the lattice parameter extracted independently from every correctly indexed family should agree within instrumental uncertainty.

Peak position and width respond differently to imperfections. If all sampled crystallites undergo the same fractional change of their plane spacing, their peaks shift together by the differential Bragg relation. A distribution of tensile and compressive strains can instead broaden the peak around an almost unchanged mean angle. A finite crystallite size can also broaden it, but its angular dependence across several reflection families differs from the leading strain broadening. Comparing several indexed peaks therefore gives more information than treating a wide single line as direct evidence for a particular mean lattice parameter.

Instrumental angle calibration should be checked with a reference powder so that a detector zero shift does not masquerade as a specimen strain.

Tell a finite-size minimum from a basis extinction

Eight identical cells contain sites at 0 and half a lattice period. Sum the eight cell amplitudes coherently, then multiply by the basis amplitude $f_A + f_B e^{iQa/2}$. Here the form factors are real, positive and independent of angle.



For equal form factors the odd orders vanish. With fB/fA = 1/4 an odd order has 36% of the corresponding even-order peak intensity. The subsidiary narrow fringes arise from the finite number of cells and would narrow in a longer crystal.

Try it · Diffraction chain

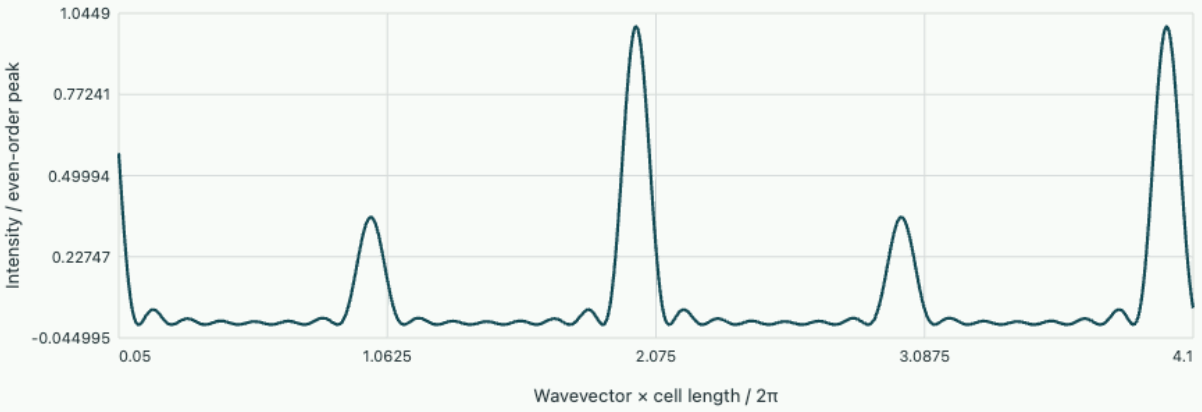
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Number of unit cells **8**

Second-site / first-site form factor **0.25**



Finite two-site crystal

Odd/even principal peak ratio = 0.36. Increasing cell count narrows the peaks; the basis determines systematic absences.

▶ **Read values from the plot**

Study the full derivation

Which reciprocal-lattice points produce a diffraction spot? · Structure, excitations and probes of solids

A bright spot depends on the wavelength and orientation of an incident beam, on the periodic repeat of a crystal, and on the atoms inside one repeat. Separate these three conditions so that a missing reflection becomes an interpretable observation. Then use the relative strength of two spots to estimate chemical order.

Chapter 2 · Vibrations of monoatomic and diatomic chains

Derive normal modes from a local force law and identify the physical motion behind the acoustic and optical branches.

FROM FIRST PRINCIPLES

Turn a local spring law into a phonon dispersion relation

Consider identical atoms of mass m spaced by a along a chain. Write u_j for the displacement from the equilibrium position of atom j. For small changes of bond length, expand the bond energy to second order about its minimum. Its first derivative vanishes and its positive second derivative K is the measured bond stiffness. The bond on the right pulls with force $K(u_{j+1}-u_j)$; the left bond contributes $K(u_{j-1}-u_j)$. Newton's law gives

$$m\ddot{u}_j = K(u_{j+1} - 2u_j + u_{j-1}).$$

Translation of a normal mode by one site can change its phase but not its shape. Insert $u_j = Ue^{i(kja-\omega t)}$, divide by the common amplitude, and simplify:

$$-m\omega^2 = K(e^{ika} + e^{-ika} - 2) = -4K\sin^2(ka/2),$$

$$\omega = 2\sqrt{K/m}\,|\sin(ka/2)|.$$

For wavelengths much larger than a,

$$\sin(ka/2) \simeq ka/2.$$

The slope gives the sound speed $a\sqrt{K/m}$. At $k=0$ all atoms move together and no bond stretches, which explains the zero frequency. More than one atom per primitive cell permits internal relative motion even at $k=0$, the ingredient needed for an optical branch.

DERIVATION AND APPLICATIONS

Count the modes that crowd against a chain's band edge

Why can a narrow frequency bin near the top of a vibration band contain a substantial share of all modes? Take a monoatomic chain of N equal masses m with one longitudinal displacement each. Impose periodic ends and use identical harmonic nearest-neighbour springs of stiffness K with spacing a. The frequency on the positive half of the first zone rises monotonically to $\omega_m = 2\sqrt{K/m}$. Inverting that branch gives the wavevector bounding the states below a trial frequency,

$$k(\omega) = \frac{2}{a}\arcsin\!\left(\frac{\omega}{\omega_m}\right), \qquad 0 \le \omega \le \omega_m.$$

Allowed wavevectors are spaced by $2\pi/(Na)$. Counting the interval of width $2k(\omega)$ includes both propagation directions. In a large chain the fractional count and its derivative per atom are therefore

$$F(\omega) = \frac{2k(\omega)}{2\pi/(Na)}\frac{1}{N} = \frac{2}{\pi}\arcsin\left(\frac{\omega}{\omega_m}\right),$$

$$g(\omega) = \frac{dF}{d\omega} = \frac{2}{\pi\omega_m\sqrt{1-(\omega/\omega_m)^2}}.$$

Here $g\,d\omega$ is a fraction of the modes. The normalization follows from $F(\omega_m) = 1$, as required by the count of independent displacements. Near the edge the slope of the dispersion tends to zero, so evenly spaced wavevectors map into progressively smaller frequency intervals. This causes an integrable singularity in the continuum density. A finite spectrometer bin always contains a finite count. A measured spectrum can also weight modes differently through a probe's coupling to atomic motion.

WORKED EXAMPLE

Count modes within the upper tenth of the band

Change the inputs and investigate this derivation.

For $N = 10{,}000$, the interval between $0.9\omega_m$ and the upper edge contains $N[1-(2/\pi)\arcsin(0.9)] \simeq 2871$ modes, allowing a mode or two for endpoints. A uniform distribution in frequency would give 1000. A nearly stationary wave packet near the upper edge and a large number of modes per unit frequency have the same origin in the small group velocity.

SECOND APPLICATION

Use the spectral median to infer a spring stiffness

Change the inputs and investigate this derivation.

An ideal chain with atomic mass $5.00 \times 10^{-26}\,\mathrm{kg}$ has half of its longitudinal modes below an ordinary frequency of $4.00\,\mathrm{THz}$. Setting $F = 1/2$ gives

$$\omega_{1/2}/\omega_m = \sin(\pi/4).$$

The upper frequency is $5.66\,\mathrm{THz}$, and $K = m(2\pi f_m)^2/4 = 15.8\,\mathrm{N/m}$. The frequency ratio is unchanged by conversion to angular frequency, but the final stiffness requires that conversion.

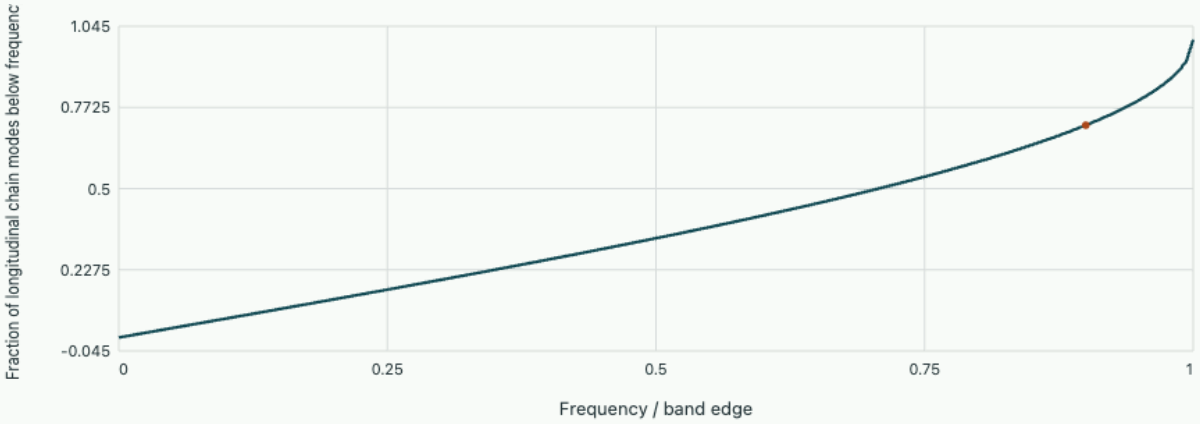
Try it · Chain count

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

Selected fraction of band-edge frequency **0.9**



Uniform wavevector count | Selected cutoff

Fraction above selected cutoff 0.28713. The spectral median is ωmax/√2; mode density diverges integrably at the band edge.

From bond extension to a travelling displacement

Let atoms of mass m on a line be coupled by springs of stiffness K, spaced a apart. The spring to the right pulls site n by $K(u_{n+1}-u_n)$; the left contributes $K(u_{n-1}-u_n)$. A uniform shift produces no force. Translation symmetry suggests $u_n = u e^{i(kna-\omega t)}$. Substitution gives

$$\omega^2(k) = \frac{2K}{m}(1-\cos ka) = \frac{4K}{m}\sin^2(ka/2).$$

The sound speed at long wavelength is $a\sqrt{K/m}$, consistent with a continuum rod whose one-dimensional stiffness is Ka and whose mass per length is m/a. At $k=\pi/a$ neighbouring atoms move in opposition and the slope of the dispersion vanishes. A wave packet made of nearby k's moves with the group velocity $d\omega/dk$, which can differ greatly from the phase velocity ω/k near the boundary.

Background · Why must one vibration have exactly zero frequency?

Give every atom the same displacement. If the potential energy depends only on separations, no bond changes length and no atom experiences a restoring force. Differentiate the force with respect to this uniform translation. For one displacement direction the resulting row-sum condition on the force constants is

$$\sum_j K_{ij} = 0,$$

$$D\,(\sqrt{m_1}, \ldots, \sqrt{m_N})^T = 0,$$

$$D = M^{-1/2} K M^{-1/2}.$$

With unequal masses the translational eigenvector is uniform in physical displacement, while its mass-weighted components vary. A nonzero frequency at zero wavevector may reveal missing neighbour interactions in a numerical model. A real spring to a substrate also breaks translation symmetry and produces a physical gap. The boundary conditions decide which explanation is plausible.

Introduce an internal coordinate

Place masses m, M alternately, with spacing $a/2$ between neighbours and cell period a. Let their displacements in cell n be Ue^{ikna} and Ve^{ikna} times the same time factor. The two force balances contain sums of neighbouring amplitudes. For example the A mass sees B in cell n and in cell $n-1$, giving $m\omega^2 U = 2KU - K(1 + e^{-ika})V$. Exchanging the masses and conjugating the phase gives the second row. Requiring a nonzero solution yields

$$mM\omega^4 - 2K(m+M)\omega^2 + 4K^2 \sin^2(ka/2) = 0.$$

There are two nonnegative solutions for ω^2 per k. With N cells under periodic boundary conditions this gives $2N$ longitudinal modes, exactly matching the number of independent displacements. At $k = 0$ the lower branch has $U = V$ and zero frequency. For the upper branch the centre of mass stays fixed, $mU + MV = 0$, with $\omega^2 = 2K(1/m + 1/M)$. An electric field can excite some optical modes in polar crystals; the branch name alone does not guarantee optical activity.

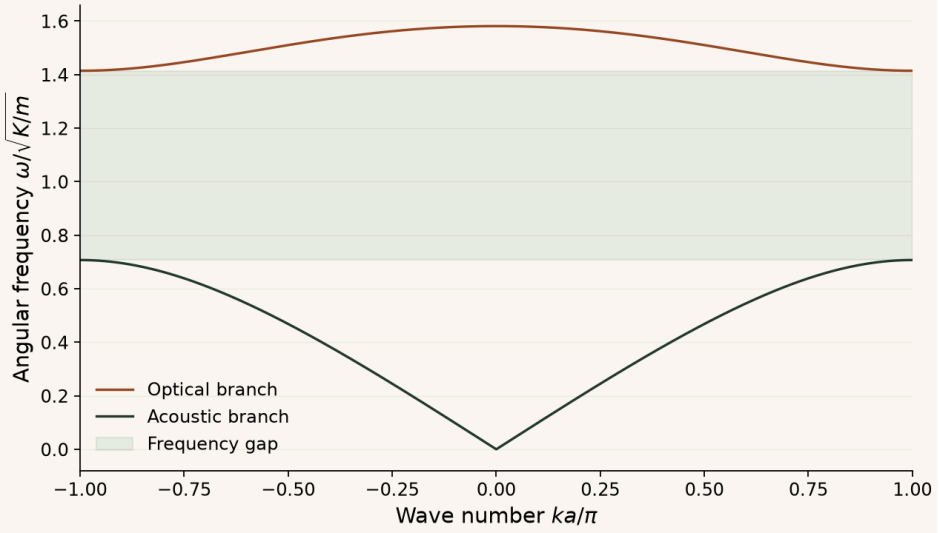
A gap and a sound speed answer different questions

At the zone boundary $1 + e^{ika} = 0$, so the sublattices decouple. If $M = 4m$, the two edge frequencies are $\sqrt{K/(2m)}$ and $\sqrt{2K/m}$, a factor of two apart. When the masses become identical the two folded branches meet, recovering the original monoatomic chain in a doubled cell. An apparent branch boundary created by plotting in a smaller zone need not be a physical frequency gap.

Expand the determinant for small k on the lower branch. Since ω^2 scales as k^2, discard the term of order k^4. This gives

$$v_s = a\sqrt{K/[2(m+M)]}.$$

The two springs in a cell act in series for long waves, while both masses contribute to inertia. The sound speed alone determines a stiffness-to-mass ratio; measuring the upper branch supplies independent information about the internal motion.



For a four-to-one mass ratio the zone-edge frequencies differ by a factor of two. The shaded interval contains no longitudinal normal mode of the infinite chain.

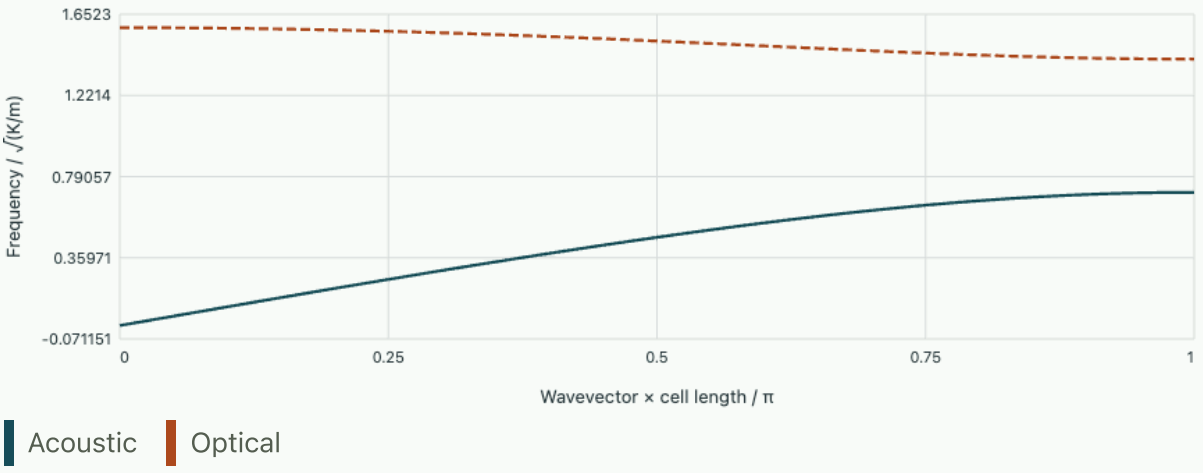
Try it · Diatomic

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

Heavy / light mass 4



Zone-edge frequencies 0.70711 and 1.4142. The gap closes at equal masses.

Convert a mode count into thermal predictions

Quantisation assigns a harmonic mode levels $\hbar\omega(j+1/2)$. The thermal part of its mean energy is $\hbar\omega/[\exp(\hbar\omega/k_BT)-1]$. Low-temperature excitations thus select long-wavelength acoustic modes while high-frequency optical modes remain scarcely populated. Counting states in a three-dimensional sphere gives a low-frequency density proportional to ω^2 for linear branches. Rescaling the energy integral by $\hbar\omega/k_BT$ gives thermal energy proportional to T^4 and heat capacity proportional to T^3. Dimensionality and linear dispersion are essential assumptions behind that power.

Count thermally generated vacancies

A mode calculation for a perfect crystal starts with one atom on every site. Heating a real sample can leave some atoms at a surface or reservoir and some lattice sites empty. How large should that defect population be even at thermal equilibrium? Consider N equivalent sites and n_v independent vacancies. At fixed temperature and pressure let g_f be the Gibbs cost of making one vacancy and transferring its atom

to the reservoir. It may contain local vibrational contributions. We count the separate entropy of choosing the missing sites explicitly.

There are $W = N!/[n_v!(N-n_v)!]$ such arrangements. With $x = n_v/N$ and macroscopic occupations, Stirling's formula gives

$$\frac{G - G_0}{N} = x g_f + k_B T\,[x \ln x + (1-x)\ln(1-x)].$$

Minimizing with respect to x gives

$$g_f + k_B T \ln[x/(1-x)] = 0.$$

The second derivative is positive. Thus $x = 1/(1 + e^{g_f/(k_B T)})$. For positive $g_f \gg k_B T$ there are few vacancies and $x \simeq e^{-g_f/(k_B T)}$. Each vacancy costs energy, but at tiny x the number of possible positions grows fast enough to favour a nonzero population. A finite microscopic specimen can still contain zero vacancies at a given instant.

Take $g_f = 0.800\,\mathrm{eV}$ as constant over the chosen range and $k_B = 8.6173 \times 10^{-5}\,\mathrm{eV/K}$. At $900\,\mathrm{K}$, $k_B T = 0.07756\,\mathrm{eV}$ and $x = 3.31 \times 10^{-5}$. A region with 10^{15} sites contains an expected 3.31×10^{10} empty sites. At $600\,\mathrm{K}$ the equilibrium fraction drops to 1.91×10^{-7}. Rapid cooling may leave an excess because vacancies must migrate to sinks before the concentration relaxes. The formation free energy in this calculation sets an equilibrium population. A migration barrier would enter the time to attain it.

Practise on sheet 1 · Infer a defect formation enthalpy from two temperatures. Further reading · Massachusetts Institute of Technology · 3.46 Photonic Materials and Devices · Notes on Defects in Solids, §3 Point Defect Equilibria.

A light substitution can localize a vibration

An infinite monatomic chain has a continuous vibration band, while a single altered atom can support a vibration spatially concentrated near one site. Let identical springs of stiffness K connect masses m on a chain of spacing a, except the central mass at site zero is M. Seek a symmetric localized solution with alternating neighbouring amplitudes $u_n = A(-q)^{|n|} e^{-i\omega t}$ for $0 < q < 1$. Alternation is motivated by the top of the perfect-chain band where adjacent atoms oscillate out of phase. Exponential decay supplies the localization.

At any ordinary site away from the impurity, the equation $-m\omega^2 u_n = K(u_{n+1} + u_{n-1} - 2u_n)$ gives

$$m\omega^2 = K(2 + q + q^{-1}).$$

Since $q + q^{-1} > 2$, this frequency lies above the perfect-chain band edge $\omega_{\max} = 2\sqrt{K/m}$. At site zero both neighbours have displacement $-qu_0$, giving $M\omega^2 = 2K(1+q)$. Eliminating ω and cancelling $1+q$ yields

$$q = \frac{M}{2m - M},$$

$$\omega^2 = \frac{K}{m}(2 + q + q^{-1}).$$

The physical condition $0 < q < 1$ requires a lighter central mass $0 < M < m$. A heavy isotope can still scatter travelling waves or produce resonant response within the band. This ansatz predicts no exponentially localized symmetric mode above the band for the heavier substitution. Its absence follows from the admissible decay factor rather than an assumption that a heavy impurity cannot move.

Predict a localized spectral peak

Take host atomic mass $m = 4.00 \times 10^{-26}$ kg, spring stiffness $K = 40.0$ N/m, lattice spacing $a = 0.300$ nm and impurity $M = m/2$. The host upper angular frequency is

$$2\sqrt{40.0/(4.00 \times 10^{-26})} = 6.32 \times 10^{13}\,\mathrm{rad/s}.$$

For the impurity $q = 1/3$, so $\omega^2 = (16/3)K/m$, giving $\omega = 7.30 \times 10^{13}$ rad/s or $f = 11.6$ THz. The ordinary chain ends near 10.1 THz.

The neighbour amplitude is minus one third of the central amplitude, and the second neighbour has plus one ninth. Expressing the absolute envelope as $e^{-|n|a/\xi}$ gives localization length $\xi = a/\ln 3 = 0.273$ nm. As M approaches m from below, q approaches one, localization length diverges and the extra frequency merges with the band edge. A finite-chain calculation should therefore vary chain length before claiming that a very weak impurity mode is spatially localized: a box shorter than its decay length cannot separate that broad tail from a standing host mode.

Spatially resolved displacement data can provide two independent estimates of impurity mass. Invert the decay-factor relation to obtain $M/m = 2q/(1 + q)$ from neighbouring displacement amplitudes. Separately insert the measured localized frequency and the same host force constant in the central equation of motion. Agreement tests the assumption that the impurity changes only inertia. A substituted atom may also change the two neighbouring springs: in that case the observed frequency and decay factor can no longer be fitted by altering mass alone. This inversion distinguishes an isotope-like substitution from a stronger local bonding perturbation.

A weak second impurity far away perturbs the localized mode only through its exponentially small tail. Bringing two such impurities within a few decay lengths allows the modes to mix and split in frequency, offering another test of their spatial localization.

Study the full derivation

Read atomic motion from Bragg attenuation and diffuse scattering · Structure, excitations and probes of solids

A crystal can retain its periodic mean positions while its diffraction pattern changes markedly with temperature. Start with the thermal average of a scattering amplitude, derive the displacement factor and learn which component of atomic motion a given reflection measures.

Problem sheet 1 · Identifying a lattice and its mechanical modes

Start from reciprocal-space or frequency observations and infer which properties of the crystal they determine.

1. Two peaks in a one-dimensional specimen

Successive reciprocal-lattice orders occur at 12.57 and 25.13 nm⁻¹. The corrected integrated intensity of the second order is nine times that of the first. (a) Find the cell length. (b) For sites at $0, a/2$ with real positive amplitudes $f_A > f_B$, determine their ratio. (c) What changes if the crystal is translated by one quarter of a cell?

▼ **Separate positions, amplitudes and phases**

The spacing gives $a = 2\pi/12.56 \simeq 0.500$ nm. Even orders have amplitude $f_A + f_B$ and odd ones $f_A - f_B$, so $(f_A + f_B)/(f_A - f_B) = 3$, hence $f_A/f_B = 2$. Translation adds the phase $e^{iG_h a/4}$ to every amplitude for order h. Peaks and intensities stay fixed. The ratio calculation assumes that instrumental and form-factor differences between orders have already been corrected.

---

2. Sound speed in two isotopes

A long monoatomic chain has sound speed 4000 m/s, spacing 0.25 nm and site mass 5.0×10^{-26} kg. (a) Find its spring constant. (b) Find its edge frequency in cycles per second. (c) If every mass increases by 10% while the spring constant and lattice spacing are fixed, predict the two fractional changes.

▼ **Carry the angular frequency through to the final step**

From $v_s = a\sqrt{K/m}$, $K = m(v_s/a)^2 = 12.8$ N/m. The edge angular frequency is $2v_s/a = 3.20 \times 10^{13}$ s⁻¹, giving 5.09 THz after division by 2π. Both the sound speed and edge frequency scale as $m^{-1/2}$; both fall by $1 - 1/\sqrt{1.1} = 4.65\%$.

---

3. A diatomic inverse measurement

A diatomic chain with equal nearest-neighbour springs shows angular frequencies 2.0×10^{13} and 3.0×10^{13} s⁻¹ at the zone boundary. (a) Infer the heavy-to-light mass ratio. (b) Predict the optical frequency at the zone centre. (c) Explain whether these two data determine the absolute masses and stiffness.

▼ **Use the boundary decoupling**

The smaller squared frequency is $2K/M$ and the larger is $2K/m$, hence

$$M/m = (3/2)^2 = 2.25.$$

At the centre $\omega_{\mathrm{op}}^2 = 2K(1/m + 1/M)$, the sum of the two squared edge frequencies. This gives 3.61×10^{13} s⁻¹. Multiplying both masses and the spring stiffness by the same factor leaves all frequencies invariant. One independently known mass or stiffness resolves that ambiguity.

Fold a wave vector into the first zone

A rectangular layer has primitive lengths 0.400 and $0.250\,\mathrm{nm}$ along perpendicular axes. A momentum-resolved measurement reports $\mathbf{k} = (9.00, -14.00)\,\mathrm{nm}^{-1}$. Determine the first-zone boundaries, fold the point, and give one real-space measurement for which the two labels yield the same phase.

▼ **Construct the zone and check a translation**

The reciprocal periods are

$$2\pi/0.400 = 15.708,$$

$$2\pi/0.250 = 25.133\,\mathrm{nm}^{-1}.$$

The first zone extends from -7.854 to 7.854 along x and from -12.566 to $12.566\,\mathrm{nm}^{-1}$ along y. Subtract one x period and add one y period to get

$$\mathbf{k}' = (-6.708, 11.133)\,\mathrm{nm}^{-1}.$$

Over one translation along x the phase difference is $(k_x - k_x')(0.400\,\mathrm{nm}) = 2\pi$. A phase comparison between equivalent lattice sites therefore cannot distinguish these labels.

Revisit chapter 1 · Construct reciprocal space from the primitive cell.

Infer a defect formation enthalpy from two temperatures

Equilibrated vacancy fractions are $x_1 = 1.00 \times 10^{-5}$ at $800\,\mathrm{K}$ and $x_2 = 1.00 \times 10^{-4}$ at $1000\,\mathrm{K}$. Suppose $g_f = h_f - T s_f$ with constant h_f, s_f. Eliminate the unknown entropy to infer h_f. Why would repeating the experiment on immediately quenched specimens require a different interpretation?

▼ **Eliminate the intercept and interpret the slope**

Taking log odds gives

$$\ln[x/(1-x)] = s_f/k_B - h_f/(k_B T).$$

Subtract the first measurement from the second. The result is

$$h_f = k_B \ln\{[x_2/(1-x_2)]/[x_1/(1-x_1)]\}/(1/T_1 - 1/T_2).$$

The log ratio is 2.3027 and the inverse-temperature difference $0.000250\,\mathrm{K}^{-1}$, yielding $0.794\,\mathrm{eV}$ per vacancy. After a quench, a measured fraction can retain memory of an earlier temperature. Fitting it against the final measurement temperature assumes equilibration that may not have occurred.

Revisit chapter 2 · Count thermally generated vacancies.

Explore these relations

- For the two diffraction peaks, vary the basis amplitude ratio; a ratio of one half gives an even-to-odd intensity ratio of nine. Open the live calculation.
- For the diatomic inverse measurement, vary the mass ratio and inspect both zone-edge frequencies. Open the live calculation.

Chapter 3 · Filling electronic states and measuring the Fermi surface

Use exclusion and state counting to predict carrier energies, thermal response and the effect of dimensionality.

FROM FIRST PRINCIPLES

Count the states in a filled Fermi sphere

Take independent spin-one-half electrons in a large cube of volume $V = L^3$ with periodic boundary conditions. The condition $e^{ik_x L} = 1$ permits $k_x = 2\pi n_x / L$, with the same spacing in the other directions. Thus each allowed wave vector occupies reciprocal volume $(2\pi)^3/V$. For the free-particle energy $E = \hbar^2 k^2/(2m)$, all states below energy E lie in a sphere. Counting its volume and two spin states gives

$$N(E) = 2\frac{V}{(2\pi)^3}\frac{4\pi}{3}\left(\frac{2mE}{\hbar^2}\right)^{3/2}.$$

At zero temperature the exclusion principle fills these single-particle states up to E_F. Writing $n = N/V$, solve the counting formula at that upper limit:

$$k_F = (3\pi^2 n)^{1/3},$$

$$E_F = \frac{\hbar^2}{2m}(3\pi^2 n)^{2/3}.$$

The number of available states in a thin energy shell is the derivative $D(E)dE$. In three dimensions $D(E) \propto \sqrt{E}$ because the shell area grows with k^2 while dk/dE falls as $1/k$. This counting assumes an isotropic parabolic band. Additional valleys or a different effective dimension change the result.

Background · Fermi level

$$f(E) = 1/[1 + \exp((E-\mu)/k_B T)]$$

At equilibrium a one-electron level at the chemical potential has mean occupancy one half per available state. Spin multiplicity enters the number of states separately.

Background · Density of states

$$g(E) = dN_{\text{states}}/dE$$

State density and state occupation are independent factors. Their product, integrated over a band, gives the mean particle number.

DERIVATION AND APPLICATIONS

Obtain the pressure of a filled Fermi sea by compressing its states

Even at zero temperature a box of mobile electrons exerts a pressure on its boundaries. Consider N noninteracting spin-one-half fermions in a three-dimensional volume V, with fixed isotropic mass m and quadratic energy. Take the bottom of the band as a volume-independent energy reference. Every state below E_F is occupied once per allowed spin state. Since the density of states including spin has the form $D(E) = C\sqrt{E}$, dividing two elementary integrals eliminates the normalization constant,

$$\frac{U}{N} = \frac{\int_0^{E_F} E\, C\sqrt{E}\, dE}{\int_0^{E_F} C\sqrt{E}\, dE} = \frac{3}{5} E_F.$$

State counting gives

$$E_F = (\hbar^2/2m)(3\pi^2 N/V)^{2/3}.$$

During a uniform compression at fixed N the upper occupied energy rises. At zero entropy the energy derivative with respect to box volume is the negative pressure, so

$$P = -\left(\frac{\partial U}{\partial V}\right)_N = \frac{2U}{3V} = \frac{2}{5} n E_F,$$

$$n = \frac{N}{V}.$$

Density appears both explicitly and through E_F. A second volume derivative gives the positive resistance to a small fractional compression,

$$B = -V\left(\frac{\partial P}{\partial V}\right)_N = \frac{5}{3} P = \frac{2}{3} n E_F.$$

Both P and the bulk modulus B have units of energy per volume. The factor $5/3$ distinguishes the change of pressure from pressure itself. An actual solid also has electrostatic and lattice contributions.

Extracting the electronic contribution from a measured solid modulus requires information about these other stresses and the volume dependence of the band parameters.

WORKED EXAMPLE

Estimate the electronic stiffness of a simple metal

Change the inputs and investigate this derivation.

Take $n = 5.00 \times 10^{28}\,\mathrm{m^{-3}}$ and $E_F = 5.00\,\mathrm{eV}$. With $1\,\mathrm{eV} = 1.602 \times 10^{-19}\,\mathrm{J}$, the product nE_F is $4.01 \times 10^{10}\,\mathrm{Pa}$. The pressure contribution is $16.0\,\mathrm{GPa}$ and the modulus is $26.7\,\mathrm{GPa}$. A one percent volume decrease would raise this pressure by approximately $0.267\,\mathrm{GPa}$ in the linear response regime.

SECOND APPLICATION

Predict a finite compression without linearizing

Change the inputs and investigate this derivation.

Suppose a separate trapped gas starts at pressure $12.0\,\mathrm{kPa}$ and is compressed slowly to 0.800 of its original volume at fixed particle number and negligible entropy. Its pressure becomes

$$12.0(1/0.800)^{5/3} = 17.4\,\mathrm{kPa}.$$

Its energy per particle rises by only $(1/0.800)^{2/3} = 1.16$ times. The larger pressure ratio also contains the increase in particles per unit volume.

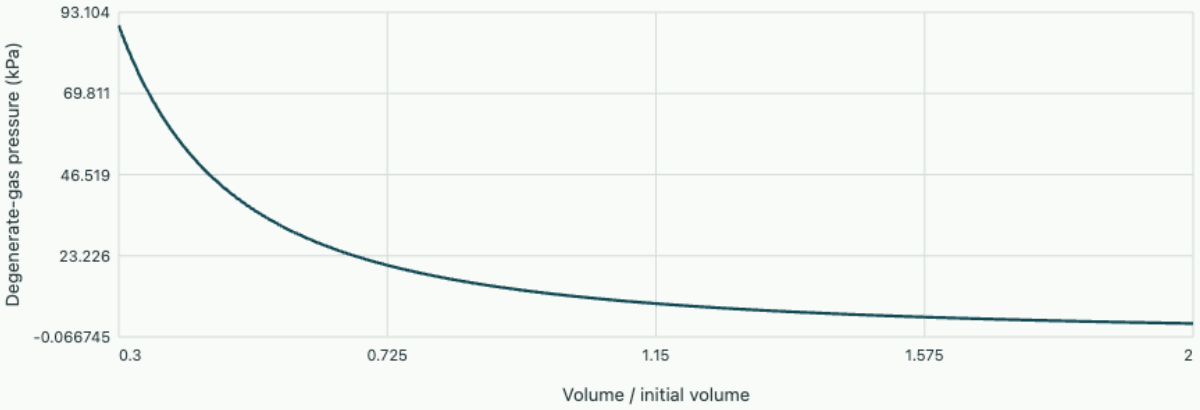
Try it · Fermi pressure

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

Initial pressure (kPa) **12**



Zero-temperature 3D ideal Fermi gas

At fixed particle number P∝V^(−5/3), and the bulk modulus is 5P/3. Nonrelativistic, noninteracting fermions; thermal corrections omitted.

Count occupied states at zero temperature

Consider a cubic box of volume V with periodic boundaries and free electrons of effective model mass m. Allowed wavevectors occupy a volume $(2\pi)^3/V$ per state for each spin. At zero temperature each state below the chemical potential is filled. For $\epsilon = \hbar^2 k^2/2m$ the filled region is a sphere and each k has two spin states. Hence

$$N = 2\frac{V}{(2\pi)^3}\frac{4\pi k_F^3}{3},$$

$$k_F = (3\pi^2 n)^{1/3},$$

$$E_F = \frac{\hbar^2 k_F^2}{2m}.$$

With $n = 8.5 \times 10^{28}$ m⁻³ and the free electron mass, k_F is about 1.36×10^{10} m⁻¹ and E_F about 7.0 eV. The corresponding temperature E_F/k_B is roughly 81,000 K. Ordinary room temperature therefore changes occupation only near the outer surface of this sea in a simple metal. This is a model estimate; anisotropic bands and multiple pockets change its interpretation.

The density of states is a derivative

Replacing k_F by $k(E) = \sqrt{2mE}/\hbar$ gives the integrated state count below energy E. Its derivative is

$$D(E) = \frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{E},$$

$$D(E_F) = \frac{3N}{2E_F}.$$

Check dimensions by integrating $D(E)\,dE$ to recover a count. In two dimensions the occupied area grows as k^2 and the density per energy is constant for a parabolic band. In one dimension the count grows as k and the density has an integrable $E^{-1/2}$ singularity at its bottom. A large density of states means many states fit in a narrow energy window. It does not imply a large group velocity: a flat dispersion may create the large density.

OneSimulation · What would a probe see at the last atom?

An open one-dimensional chain with one orbital per atom has standing-wave eigenvectors. Counting their energies gives a spectrum for the complete specimen. An ideal local probe couples to the amplitude on the atom below it. For a probe over site j, broaden each spectral line with a normalised Gaussian g_η of experimental width η.

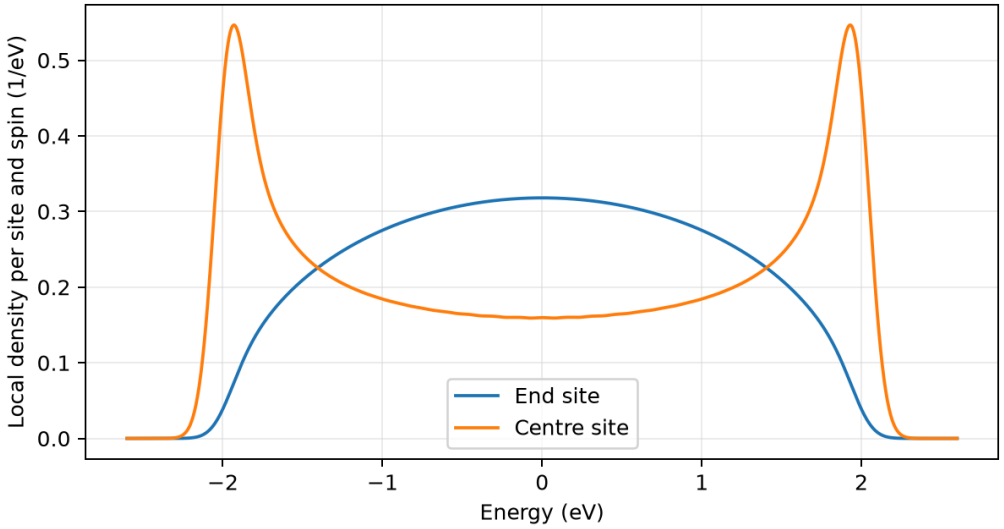
$$\rho_j(E) = \sum_n |\psi_n(j)|^2 g_\eta(E - E_n),$$

$$\int \rho_j(E)\,dE = 1.$$

The integral counts one orbital per spin at either site. Near a band edge the standing waves vary slowly and have a small amplitude next to the missing neighbour. Their large bulk density of states can coexist with a weak end signal. The OneSimulation eigensolve uses the same Hamiltonian for both curves. Only the projected site changes.

```
"""Local spectroscopy at an open chain end and in its interior."""
import matplotlib.pyplot as plt
import numpy as np
from onesim.numerics.linear.matrices.dense_linear_algebra import symmetric_eigensystem

sites = 81
hopping = 1.0  # eV
broadening = 0.09  # eV, standard deviation of the energy resolution
hamiltonian = -hopping * (np.eye(sites, k=1) + np.eye(sites, k=-1))
levels, vectors = symmetric_eigensystem(hamiltonian)
energy = np.linspace(-2.6 * hopping, 2.6 * hopping, 1001)
```

```
kernel = np.exp(-0.5 * ((energy[:, None] - levels) / broadening)**2)
kernel /= np.sqrt(2 * np.pi) * broadening
fig, ax = plt.subplots(figsize=(6.6, 3.5), layout="constrained")
for site, label in [(0, "End site"), (sites // 2, "Centre site")]:
    weights = np.abs(vectors[site])**2
    density = kernel @ weights
    assert np.isclose(np.trapezoid(density, energy), 1, atol=1e-6)
    print(f"{label}: local density at zero energy = {density[len(energy)//2]:.3f} /eV")
    ax.plot(energy, density, label=label)
ax.set(xlabel="Energy (eV)", ylabel="Local density per site and spin (1/eV)")
ax.grid(alpha=0.2)
ax.legend()
fig.savefig("condensed-matter.png", dpi=180)
```



The integrated areas agree. Near zero energy the end site carries about 0.318 states/eV per spin and the centre carries 0.160. A change in probe location redistributes spectral weight.

Try it · Local dos

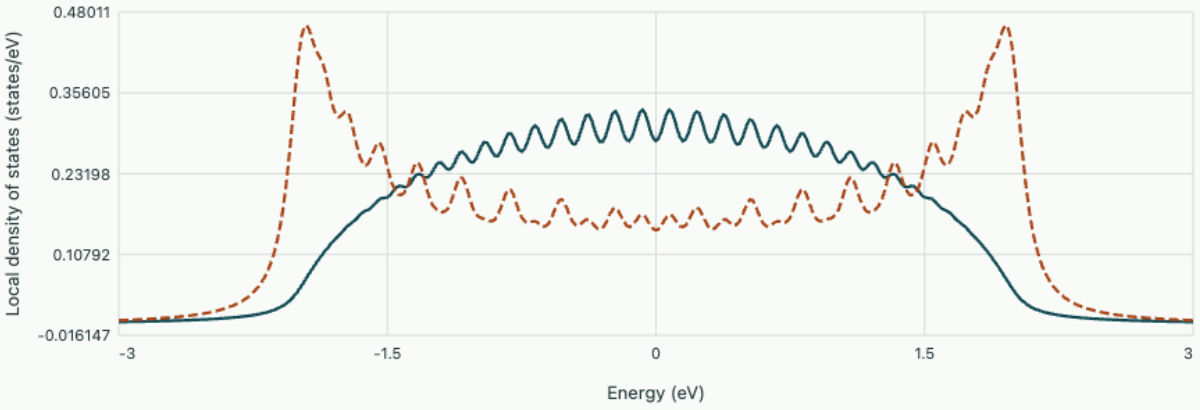
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Probe site in 40-site open chain **1**

Lorentzian broadening (eV) **0.08**



Selected site Central site

Nearest-neighbour hopping is 1 eV. Each full-energy spectral integral is one state per spin; finite plotted limits omit Lorentzian tails.

Download the Python example

Why only a shell contributes to heat capacity

At finite temperature the occupation is $f(E) = [1 + e^{(E-\mu)/k_B T}]^{-1}$. Heating can promote an electron only into a vacant state. Deep in the filled sea all nearby destinations are occupied. The active shell has width of order $k_B T$, involving about $D(E_F) k_B T$ states; the typical energy change is another factor $k_B T$. Thus the thermal energy increases quadratically and the heat capacity rises linearly with T.

A coefficient follows by expanding the smooth density of states across that shell. At fixed particle number, the relevant excess energy to leading order is $\int (E - E_F) D(E_F) [f(E) - \Theta(E_F - E)]\, dE$. The filled and empty sides contribute equally to this order. Substituting $x = |E - E_F|/k_B T$ and using $\int_0^\infty x/(e^x + 1)\, dx = \pi^2/12$ gives

$$C_{\mathrm{el}} = (\pi^2/3) D(E_F) k_B^2 T.$$

This expansion requires the density to be smooth on the thermal scale; a band edge at the chemical potential needs separate treatment.

Separate two low-temperature contributions

In a three-dimensional simple metal combine $C_{\mathrm{el}} = \gamma T$ with the phonon result $C_{\mathrm{ph}} = \beta T^3$. Plot C/T against T^2; the intercept estimates γ, and the slope estimates β. If two points are $(4, 3)$ and $(16, 6)$ in K² and mJ mol⁻¹ K⁻², respectively, $\beta = 0.25$ mJ mol⁻¹ K⁻⁴ and $\gamma = 2.0$ mJ mol⁻¹ K⁻². Curvature across the fitted interval would be evidence against this two-term description. To go beyond a free sphere we next retain the periodic potential of the ions.

Count charges and electronic compressibility in two dimensions

A two-dimensional electron layer illustrates how changing available spatial dimensions alters the density of states and an electrical measurement. Assume one parabolic subband $E = E_c + \hbar^2(k_x^2 + k_y^2)/(2m^*)$, spin degeneracy two and no valley degeneracy. A square sample of area A with periodic boundary conditions supplies one state per spin per reciprocal area $(2\pi)^2/A$. Counting the disk up to k_F gives areal density $n_s = k_F^2/(2\pi)$, hence

$$E_F - E_c = \frac{\pi \hbar^2 n_s}{m^*},$$

$$D_A(E) = \frac{dn_s}{dE_F} = \frac{m^*}{\pi \hbar^2}$$

above the subband edge at zero temperature. Unlike the square-root density in three dimensions, each occupied parabolic two-dimensional subband contributes a constant step. When an additional subband begins to fill it supplies another step, possibly with different degeneracy or mass. Temperature and disorder broaden the step in a real measurement.

An electrostatic gate must both create a field across the insulating spacer and raise the layer’s chemical potential. For areal capacitance C_g' of the spacer and increment of sheet charge magnitude $dQ' = e\,dn_s$, the required voltage is

$$dV = dQ'/C_g' + dE_F/e.$$

The second term defines the quantum capacitance $C_q' = e^2 D_A$. Their inverse capacitances add in series. For a small density of states the electronic contribution measurably lowers the total capacitance even when the geometric dimensions are known precisely.

Measure which capacitance limits a device

Take $m^* = 0.067 m_e$ and $n_s = 2.00 \times 10^{15}$ m⁻². Then $k_F = \sqrt{2\pi n_s} = 1.12 \times 10^{8}$ m⁻¹ and $E_F - E_c = 7.15$ meV. At 4 K the thermal scale $k_B T = 0.345$ meV is smaller than this separation, supporting the degenerate estimate. The electronic capacitance per area is 0.0448 F/m². For a spacer with $C_g' = 0.0100$ F/m², the measured series value is $(1/0.0100 + 1/0.0448)^{-1} = 0.00818$ F/m².

Over a gate area of 100 µm², the predicted capacitance is 0.818 pF rather than the purely geometric 1.00 pF. Treating that measured value as a change of dielectric thickness would misidentify an electronic

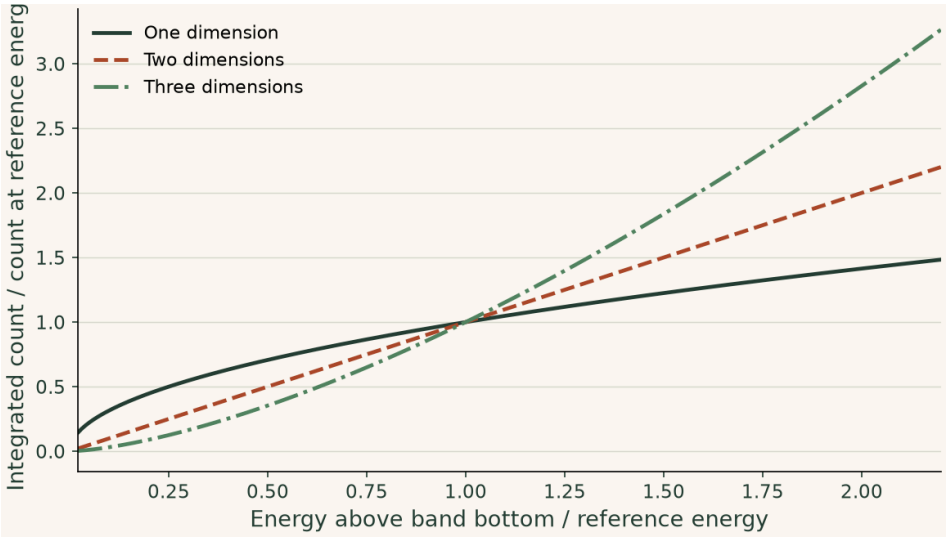
effect. In the limit of a very high density of states, the quantum capacitance becomes large and the geometric value is recovered. Near depletion the assumptions of a sharp occupied subband and uniform layer require scrutiny. Comparing capacitance with an independent sheet-density measurement can test whether a low reading reflects depleted electronic states or simply a different insulating spacer.

Using the series result, a small gate swing can be split into two measurable energy changes. A 10.0 mV swing at the stated operating point changes sheet charge magnitude by 8.18×10^{-5} C/m², or sheet density by 5.10×10^{14} m⁻². The geometric voltage increment is 8.18 mV and the chemical-potential contribution is 1.82 mV, corresponding to a Fermi-level shift of 1.82 meV. Their sum recovers the imposed swing. This decomposition assumes the layer remains in the same subband and equilibrates through its electrical contact on the capacitance measurement timescale.

A slow measurement with significant leakage instead mixes capacitance with conductance and can corrupt both estimated voltage contributions.

Count states added by raising the Fermi level

For a parabolic band, states within a radius k fill a length, area or volume in reciprocal space. Since k scales as $\sqrt{E}$, the integrated count varies as $E^{d/2}$. Normalise at one energy to compare exponents independently of the sample dimensions.



Doubling the Fermi energy raises the count by √2, 2 or 2√2. The local slope gives the density of states. Its behaviour at the band edge is quite different even when the three integrated counts agree at the reference energy.

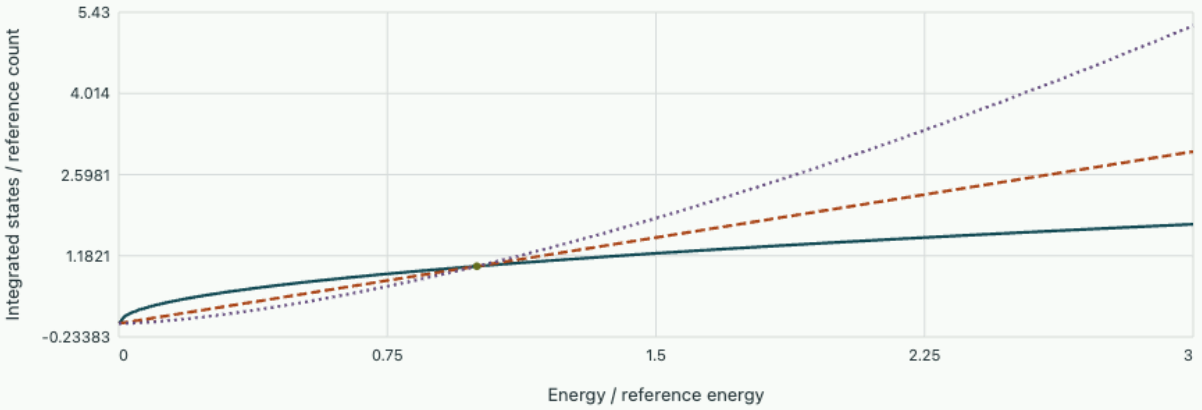
Try it · State count

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 energy / reference 1



At selected energy: 1D 1, 2D 1, 3D 1. Differentiating these curves gives the corresponding density of states.

▶ **Read values from the plot**

Study the full derivation

Infer a superconducting gap from a tunnelling spectrum · Structure, excitations and probes of solids

A tunnel junction reports differential conductance versus voltage. Extracting a superconducting energy gap requires connecting a condensate's quasiparticle dispersion to a density of states, then accounting for temperature and finite spectral resolution before comparing peaks with the data.

Chapter 4 · Bloch waves, band edges and effective mass

Derive an electronic band from both weak and strong binding, then use its slope and curvature to predict motion.

FROM FIRST PRINCIPLES

Why translating a lattice state produces an energy band

In a perfect periodic chain the Hamiltonian is unchanged when the site labels are shifted by one. A wave may therefore have site amplitudes of the form $c_j = Ce^{ikja}$, with k defined up to $2\pi/a$. To give that phase a physical energy, choose one localized orbital per cell of spacing a. Assume an orthonormal basis, on-site energy ϵ_0, and a real hopping matrix element $-t$ to either nearest neighbour. Projecting the stationary Schrödinger equation on orbital j gives

$$Ec_j = \epsilon_0 c_j - t(c_{j-1} + c_{j+1}).$$

Successive amplitudes differ only by $e^{\pm ika}$. Cancel c_j, then combine these phases into a cosine:

$$E(k) = \epsilon_0 - t(e^{ika} + e^{-ika}) = \epsilon_0 - 2t\cos(ka).$$

For $t > 0$ the in-phase combination has the lowest energy. Comparing $k = 0$ and $k = \pi/a$ gives bandwidth $4t$. Form a narrow wave packet from neighbouring k's. Its envelope speed is

$$d(E/\hbar)/dk = (2ta/\hbar)\sin(ka).$$

Translation symmetry determines the allowed phase labels, while the coupling strength determines the numerical dispersion. A real crystal can require several orbitals and longer-range couplings.

Background · Band velocity

$$\mathbf{v}_n(\mathbf{k}) = \hbar^{-1}\nabla_{\mathbf{k}} E_n$$

This is the group velocity of a band wave packet in the ordinary semiclassical approximation. The Fermi surface determines which states can respond at low temperature.

DERIVATION AND APPLICATIONS

Use band curvature to predict acceleration off the force axis

Near a smooth, isolated band minimum the first derivative of energy vanishes. In a low-symmetry crystal its second derivatives can depend on direction. Choose principal axes x, y in the plane near the minimum

$\mathbf{k}_0$, and write $\mathbf{q} = \mathbf{k} - \mathbf{k}_0$. Taylor expansion defines two positive effective masses through the curvatures,

$$E(\mathbf{k}_0 + \mathbf{q}) \simeq E_0 + \frac{\hbar^2 q_x^2}{2m_x} + \frac{\hbar^2 q_y^2}{2m_y}.$$

Treat a narrow wave packet semiclassically in a uniform force $\mathbf{F}$, over a time short enough for scattering to be negligible. Neglect interband transitions and any anomalous velocity. Differentiating energy component by component gives group velocity. Combining this with the crystal-momentum equation gives its acceleration,

$$v_i = \frac{1}{\hbar}\frac{\partial E}{\partial q_i} = \frac{\hbar q_i}{m_i},$$

$$\hbar \dot{q}_i = F_i,$$

$$a_i = \frac{F_i}{m_i}.$$

Thus a force rotated through the principal axes traces an ellipse of accelerations. To describe a longitudinal measurement, project the full vector onto the applied force. If the force makes angle θ to the x axis and ϕ is the direction of the resulting acceleration in the same quadrant, then

$$\frac{a_\parallel}{F} = \frac{\cos^2\theta}{m_x} + \frac{\sin^2\theta}{m_y},$$

$$\tan\phi = \frac{m_x}{m_y}\tan\theta.$$

The longitudinal inverse mass is a weighted average of inverse principal masses. The difference $\phi - \theta$ vanishes along principal axes and for an isotropic band. The expansion requires the whole accelerated packet to stay inside the region of roughly constant curvature.

WORKED EXAMPLE

Follow an oblique force pulse

Change the inputs and investigate this derivation.

Let $m_x = 0.20 m_e$, $m_y = 0.80 m_e$, where m_e is the free electron mass. A force pulse of magnitude $1.00 \times 10^{-17}\,\mathrm{N}$ acts at 45° for $50.0\,\mathrm{fs}$ on a packet initially at rest. Its velocity gains are $F\Delta t/(\sqrt{2}m_i)$, or $1.94\,\mathrm{m/s}$ along x and $0.485\,\mathrm{m/s}$ along y. The velocity points only 14.0° above the light-mass axis.

SECOND APPLICATION

Recover an unseen principal curvature

Change the inputs and investigate this derivation.

A second band has longitudinal mass $0.300 m_e$ for a force at 45°, and a measurement along its x axis gives

$$m_x = 0.200 m_e.$$

In units of $1/m_e$, inversion yields $m_e/m_y = 2/0.300 - 1/0.200 = 1.667$, so $m_y = 0.600 m_e$. A pulse along y should produce one third of the acceleration measured along x for the same force, providing an independent check of the principal-axis assignment.

Try it · Mass tensor

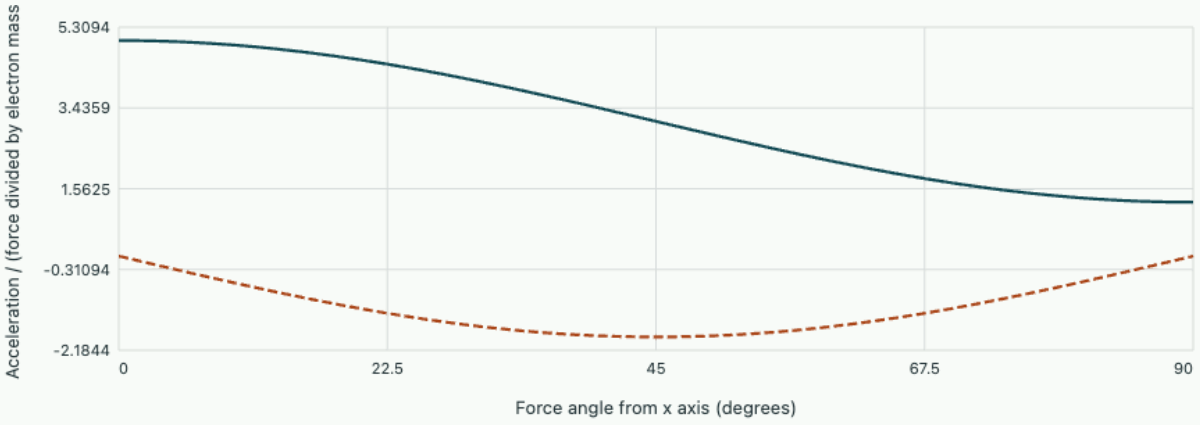
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Principal mass mx/me **0.2**

Principal mass my/me **0.8**



Acceleration need not align with force. The longitudinal inverse mass is cos²θ/mx + sin²θ/my.

Start from this chapter. At a Brillouin-zone edge, periodic modulation mixes counter-propagating waves. Test this symmetry argument using optical modes. Their eigenvalues are electromagnetic frequencies and illustrate the analogous band splitting.

OneSimulation laboratory · Open a band gap in a periodic dielectric

Why label a state by crystal momentum?

A lattice translation by a commutes with the Hamiltonian when $V(x+a)=V(x)$. Translation eigenvalues have modulus one because translation preserves norm. Write them e^{ika}; then $\psi_k(x)=e^{ikx}u_k(x)$ with $u_k(x+a)=u_k(x)$. Since $e^{i(k+G)a}=e^{ika}$ for $G=2\pi h/a$, a single Brillouin zone suffices. The index distinguishing states at the same k is the band index. Crystal momentum labels translation symmetry and can change by a reciprocal vector in scattering.

A weak potential couples a degenerate pair

Near $k=G/2$, free states $|k\rangle$ and $|k-G\rangle$ have nearly equal kinetic energies. The periodic potential has Fourier coefficients V_G that couple these waves. Absorbing the mean potential in both energies, the two-state characteristic equation is

$$(E-E_k)(E-E_{k-G})=|V_G|^2,$$

with solutions

$$E_\pm=\tfrac{1}{2}(E_k+E_{k-G})\pm\sqrt{\tfrac{1}{4}(E_k-E_{k-G})^2+|V_G|^2}.$$

At equality the splitting is $2|V_G|$. For a real coefficient the states are standing waves whose densities peak at different locations relative to the potential. Far from the degeneracy, the square root is dominated by the kinetic splitting and one obtains two nearly free waves again. A Fourier component that vanishes by symmetry cannot open a first-order gap at that crossing. Check other bands before interpreting a local splitting as a gap for the entire solid.

Build a band from a localised orbital

Start instead with one localised orbital $|n\rangle$ per cell, all of energy ϵ_0, with overlap neglected in the basis and nearest-neighbour hopping matrix element $-t$. A translation eigenstate has site amplitudes $c_n=e^{ikna}/\sqrt{N}$. Acting with H on one site amplitude gives $Ec_n=\epsilon_0 c_n-t(c_{n-1}+c_{n+1})$, and hence $E(k)=\epsilon_0-2t\cos ka$. The bandwidth $4|t|$ measures the energy spread produced by hopping. In the limit of vanishing hopping the local energy repeats for every k. More than one orbital per cell generates a matrix eigenproblem even in this simple construction.

The group velocity is $v=\hbar^{-1}dE/dk$. Under a weak force with $\hbar\dot{k}=F$, acceleration is $FE''(k)/\hbar^2$. Define $m^*=\hbar^2/E''(k)$, a local measure of curvature. Near the bottom of the cosine band with $t>0$,

$$E\simeq\epsilon_0-2t+t(ka)^2,$$

$$m^*=\hbar^2/(2ta^2).$$

Near the top the curvature changes sign. Describing the missing electrons in a nearly full band as positively charged holes with positive mass simplifies transport calculations.

Use filling to decide whether current can flow

For a complete time-reversal-symmetric band, the contributions of k and $-k$ to current cancel. More generally the integral of the derivative of a periodic band over a whole zone is zero. A weak field can redistribute carriers near a partially filled band edge, whereas a fully filled isolated band has no vacant neighbouring states. With one spin-degenerate orbital per primitive cell there are two states per cell. One electron per cell half fills this noninteracting band; two fill it. Symmetry breaking, interactions and additional bands can change the conclusion. A single measured zone-edge splitting therefore gives incomplete information about conduction.

For further work on three-dimensional bands, Fermi surfaces and electron motion, see David Tong's Solid State Physics lectures at Cambridge.

Open a gap by doubling a chain's period

A uniform one-dimensional chain with one spin-degenerate orbital and one electron per site has a half-filled band. A distortion can change both the spectrum and the periodicity used to count states. Let the sites be separated by a before the distortion. Group successive sites into A, B pairs and allow alternating hopping amplitudes $t_1, t_2 > 0$. The distorted unit cell has length $2a$. For equal on-site energies, chosen as zero, the amplitude on A couples to B inside its own cell and to B one cell away. Bloch translation supplies a phase e^{-2ika} for the latter coupling. The sublattice equations give a two-by-two matrix with zero diagonal and off-diagonal elements $-(t_1 + t_2 e^{\mp 2ika})$. Its characteristic equation yields

$$E_{\pm}(k) = \pm\sqrt{t_1^2 + t_2^2 + 2t_1 t_2 \cos(2ka)}, \qquad |k| \le \pi/(2a).$$

At the zone centre the hoppings add. At the new zone edge they subtract, opening a gap $2|t_1 - t_2|$. In the uniform limit the folded branches simply touch. Merely drawing a smaller zone did not change the crystal. It takes a physical alternation to split the states. With two electrons per doubled cell, the lower band is exactly filled when spin is counted. Its electronic energy falls on dimerization at fixed mean hopping: writing $t_{1,2} = t \pm \delta$ gives

$$E_{-} = -2\sqrt{t^2 \cos^2(ka) + \delta^2 \sin^2(ka)},$$

which is at most the undistorted lower energy. The lattice also pays an elastic cost. Quantitative distortion amplitudes require that competing energy.

For $t_1 = 1.20\,\mathrm{eV}$ and $t_2 = 0.80\,\mathrm{eV}$ the zone-centre energies remain $\pm 2.00\,\mathrm{eV}$, while the zone-edge energies become $\pm 0.40\,\mathrm{eV}$. The minimum direct excitation costs $0.80\,\mathrm{eV}$, corresponding to a vacuum wavelength of about $1.55\,\mu\mathrm{m}$ if the optical transition is allowed. The cause of this insulating filling is the pairing of bonds. Away from the commensurate filling the chemical potential can lie inside a band even if the same structural splitting remains visible.

Practise on sheet 2 · Recover alternating hoppings from spectroscopy. Further reading · University of Cambridge · David Tong, Solid State Physics · Chapter 4, Phonons, §4.1.3 Peierls Transition.

Follow a wave packet through the whole Brillouin zone

In a weak static field a band electron's crystal momentum changes even though its group velocity need not keep rising. Take one isolated tight-binding band $E(k) = E_0 - 2t\cos(ka)$, positive hopping energy t, lattice spacing a and electron charge $-e$. Let the wave packet be narrow in k compared with the zone and neglect scattering and transitions to other bands. The semiclassical force equation is $\hbar\dot{k} = -e\mathcal{E}$, while velocity follows from the band gradient:

$$v = (2ta/\hbar)\sin(ka).$$

Starting at $k = 0$ yields $ka = -\omega_B t_{\mathrm{lab}}$ with Bloch angular frequency $\omega_B = e\mathcal{E}a/\hbar$. Integrating velocity over laboratory time gives

$$x(t_{\mathrm{lab}}) - x(0) = \frac{2t}{e\mathcal{E}}[\cos(\omega_B t_{\mathrm{lab}}) - 1],$$

$$T_B = \frac{h}{e\mathcal{E}a}.$$

Although the field never changes sign, velocity reverses when the packet samples regions of opposite band slope. Over a complete zone traversal there is no net displacement because the band energy is periodic. Energy conservation offers another derivation: the electrostatic potential-energy change of a negatively charged electron in a uniform field is $+e\mathcal{E}\Delta x$, offsetting the change in band energy. The maximal excursion equals the band width $4t$ divided by the force magnitude.

Estimate an observable oscillation in a superlattice

Use a one-dimensional superlattice with $a = 8.00$ nm, hopping $t = 5.00$ meV and field $\mathcal{E} = 2.00 \times 10^5$ V/m. The potential-energy change per lattice period is $e\mathcal{E}a = 1.60$ meV. Thus $T_B = h/(1.60\,\mathrm{meV}) = 2.58$ ps and the peak-to-peak excursion is $20.0\,\mathrm{meV}/(e\mathcal{E}) = 100$ nm, about 12.5 lattice periods. The maximum speed from the band derivative is

$$2ta/\hbar = 1.22 \times 10^5\,\mathrm{m/s}.$$

At one quarter cycle velocity reaches its largest negative magnitude, while at half a cycle the electron reaches its furthest position and momentarily stops.

A momentum-relaxation time of 0.20 ps would interrupt most trajectories before a cycle completes, obscuring a Bloch oscillation despite the periodic dispersion. Conversely a scattering time several times longer than 2.58 ps would make coherent oscillatory transport plausible if the field does not tunnel carriers into other bands. In the weak-field limit the period and excursion both grow, so the mathematical limit of arbitrarily large excursions cannot be taken in a finite imperfect specimen. Recording the oscillation frequency as field is varied tests the predicted linear scaling separately from the much more sample-dependent amplitude.

The motion also clarifies the role of negative effective mass near a band top. The electron's crystal momentum continues to change in the direction of the applied force throughout the period. Its

acceleration is proportional to the local band curvature, however, and can point in the opposite direction once that curvature changes sign. This reverses packet acceleration without a sudden elastic collision with a particular lattice plane. A spectrum measured along the direction of motion can predict the velocity through its local slope, giving a direct cross-check of the acceleration inferred from a time-domain trace.

The probe must resolve the oscillations before relaxation for that comparison to be meaningful.

Study the full derivation

Construct allowed bands from a cell transfer matrix · Structure, excitations and probes of solids

In a finite well a bound level follows from matching its wavefunction at the two walls. A periodic lattice replaces those two external walls by a matching condition on every cell. Derive a band criterion that works without guessing a wave inside the forbidden gap and use it to predict transmission through a finite sequence of cells.

Problem sheet 2 · Occupied states, heat capacity and band curvature

Keep spin degeneracy, state counting and energy units consistent while comparing three electronic measurements.

1. Counting a two-dimensional electron gas

Electrons with spin degeneracy two occupy an isotropic parabolic band in an area A. (a) Derive k_F and $D(E)/A$. (b) Estimate E_F at $n = 10^{16}$ m⁻² and $m^* = 0.10 m_e$. Use $\hbar^2/(2m_e) = 0.0381$ eV nm². (c) State whether a zero-temperature step gives a close picture at 300 K.

▼ Count the disk with two spin states

The density is $n = 2\pi k_F^2/(2\pi)^2 = k_F^2/(2\pi)$, hence $k_F = \sqrt{2\pi n}$. Substituting $k^2 = 2m^*E/\hbar^2$ and differentiating gives

$$D(E)/A = m^*/(\pi\hbar^2).$$

Convert the density to 0.010 nm⁻²; the Fermi energy measured from the band minimum is

$$0.381(2\pi)(0.010) = 0.0239\,\mathrm{eV}.$$

At 300 K, $k_B T$ is about 0.0259 eV; the sharp-shell approximation is poor.

---

2. Calorimetry on a metal

At 2 K a metal has $C/T = 4.8$ mJ mol⁻¹ K⁻²; at 4 K it has $C/T = 7.2$ in the same units. (a) Fit the low-temperature electronic and phonon coefficients. (b) At what temperature do the two contributions become equal according to the fit? (c) What further measurement is needed to assess that extrapolation?

▼ Fit the straight line in squared temperature

The slope is $\beta = (7.2 - 4.8)/(16 - 4) = 0.20$ mJ mol⁻¹ K⁻⁴. Substituting either point gives $\gamma = 4.0$ mJ mol⁻¹ K⁻². Equality requires

$$T = \sqrt{\gamma/\beta} = \sqrt{20} = 4.47\,\mathrm{K}.$$

Two observations always determine a straight line. Measure several points across and beyond this temperature and inspect C/T versus T^2 for curvature or an anomaly.

---

3. A cosine band from spectroscopy

For a one-orbital chain the lower band edge at $k = 0$ is –1.2 eV and the upper edge at $k = \pi/a$ is 0.8 eV. Take $a = 0.30$ nm. (a) Find t, ϵ_0. (b) Find the bottom effective mass in units of m_e and the wavevector of maximum positive group velocity in the first zone. (c) How many electrons fit in this band for 100 cells?

▼ **Separate the centre and width of the band**

The centre is -0.2 eV and the width is 2.0 eV, giving $t = 0.50$ eV. Then

$$m^*/m_e = 0.0381/(0.50 \times 0.30^2) = 0.847.$$

The derivative of $-2t \cos ka$ is largest and positive at $ka = \pi/2$. There are 100 allowed k values and two spin states per k, so 200 electrons fill the band.

Recover alternating hoppings from spectroscopy

Two bands in a dimerized chain are separated by $3.20\,\mathrm{eV}$ at the zone centre and by $0.240\,\mathrm{eV}$ at the edge. Infer both hopping magnitudes in the equal-on-site-energy model. Can this energy spectrum identify which bond starts the chosen cell? Would the same gap make a chain with one electron per doubled cell insulating?

▼ **Invert the two splittings and count the electrons**

At the centre $2(t_1 + t_2) = 3.20\,\mathrm{eV}$. At the edge $2|t_1 - t_2| = 0.240\,\mathrm{eV}$. Choosing $t_1 \geq t_2$ gives $t_1 = (1.60 + 0.120)/2 = 0.860\,\mathrm{eV}$ and $t_2 = 0.740\,\mathrm{eV}$. Interchanging the hoppings preserves both bands, so these energies cannot fix the cell convention. With one electron per doubled cell the lower band is only half filled. In this noninteracting idealization empty states remain available arbitrarily near the Fermi level.

Revisit chapter 4 · Open a gap by doubling a chain's period.

Explore these relations

- For the electron gas, compare the dimension-dependent integrated state counts before differentiating to obtain the density of states. Open the live calculation.
- For the alternating-hopping spectrum, vary the two hoppings and compare the gap and band edges. Open the live calculation.

Chapter 5 · Carriers, neutrality and transport in semiconductors

Determine the chemical potential from particle balance and connect the same equilibrium to conductivity and a Hall measurement.

FROM FIRST PRINCIPLES

Determine electron and hole densities from one chemical potential

In thermal equilibrium fermions occupy a level E with probability $1/[e^{(E-\mu)/(k_BT)}+1]$. Take a nondegenerate semiconductor with chemical potential several k_BT inside the gap and parabolic edges E_c, E_v. For electrons in the conduction band the exponential dominates the denominator. Write energy as $E = E_c + \varepsilon$ and integrate the density of states against this occupation:

$$n = e^{-(E_c-\mu)/(k_BT)} \int_0^\infty D_c(\varepsilon) e^{-\varepsilon/(k_BT)}\, d\varepsilon = N_c e^{-(E_c-\mu)/(k_BT)}.$$

For holes count missing valence electrons. Taking $E = E_v - \varepsilon$ instead gives

$$p = N_v e^{-(\mu-E_v)/(k_BT)},$$

with N_v defined by the corresponding integral. At fixed temperature multiplication removes the chemical potential entirely:

$$np = N_c N_v e^{-(E_c-E_v)/(k_BT)} = n_i^2.$$

An intrinsic neutral crystal requires

$$n = p = n_i.$$

With ionized donors and acceptors it instead requires $n - p = N_D^+ - N_A^-$, because electron and acceptor charges are negative. This equation and the product law determine the two carrier densities. Heavy doping can move the chemical potential into a band, where the exponential approximation must be replaced by the full occupation integral.

DERIVATION AND APPLICATIONS

Measure minority-carrier lifetime from a spatial decay

Illuminate a narrow stripe at the edge of a long, uniformly doped sample. In an n-type crystal the illumination creates mobile minority holes, which diffuse away and recombine. Let $u(x) = p(x) - p_0$ be the excess hole density above its equilibrium value. Consider steady, weak injection into the half-space

$x > 0$. Assume a constant diffusivity D_p, linear recombination with lifetime τ_p, and negligible electric field. The stripe maintains a fixed excess at $x = 0$, and the bulk away from it has no further generation.

$$\Gamma_p = -D_p \frac{du}{dx},$$

$$\frac{\partial u}{\partial t} = -\frac{\partial \Gamma_p}{\partial x} - \frac{u}{\tau_p}.$$

The first equation gives flux of particles per area per time. The second counts particles in a thin slice. A declining flux with distance supplies the holes that recombine inside the slice. Setting the time derivative to zero gives an equation with two exponential solutions,

$$\frac{d^2 u}{dx^2} = \frac{u}{L_p^2},$$

$$L_p = \sqrt{D_p \tau_p},$$

$$u(x) = u(0) e^{-x/L_p}.$$

The exponentially increasing solution violates the equilibrium boundary condition far from the stripe. The remaining length L_p gives the depth over which the population falls by a factor e. Multiplying the surface flux by the positive charge e gives the hole current density into the bulk. Integrating the loss over the half-space independently recovers this flux,

$$\Gamma_p(0) = \frac{D_p u(0)}{L_p} = \int_0^\infty \frac{u(x)}{\tau_p}\, dx.$$

A second nearby surface, significant drift, or recombination changing with injection can all alter the spatial profile. Plotting the logarithm of excess signal against position tests the single-exponential assumption.

WORKED EXAMPLE

Predict the current required at an illuminated edge

Change the inputs and investigate this derivation.

For $D_p = 12.0\,\mathrm{cm}^2/\mathrm{s}$, $\tau_p = 3.00\,\mu\mathrm{s}$ and $u(0) = 2.00 \times 10^{13}\,\mathrm{cm}^{-3}$, the length is $0.00600\,\mathrm{cm}$, or $60.0\,\mu\mathrm{m}$. The surface flux is $4.00 \times 10^{16}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, giving $6.41\,\mathrm{mA}/\mathrm{cm}^2$. A stripe extending over $0.100\,\mathrm{cm}^2$ of injecting area supplies $0.641\,\mathrm{mA}$ to this bulk recombination channel.

SECOND APPLICATION

Infer a lifetime without calibrating the signal amplitude

Change the inputs and investigate this derivation.

In a second specimen, a signal proportional to excess density drops by a factor of four between positions separated by $90.0\,\mu\mathrm{m}$. Its diffusion length is

$$90.0/\ln 4 = 64.9\,\mu\mathrm{m}.$$

An independently measured $D_p = 8.00\,\mathrm{cm}^2/\mathrm{s}$ gives

$$\tau_p = (0.00649)^2/8.00 = 5.27\,\mu\mathrm{s}.$$

The ratio cancels a common detector gain, provided the equilibrium background was subtracted first.

Try it · Diffusion recombination

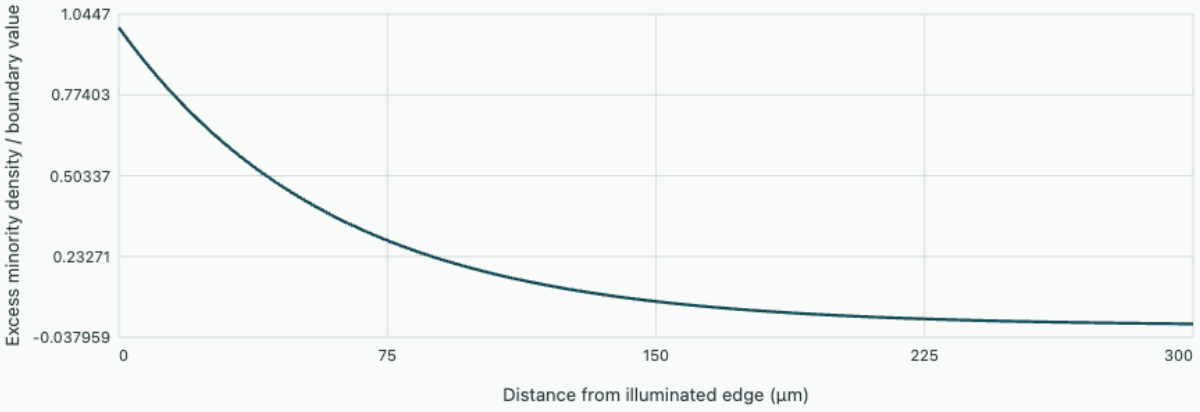
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Diffusivity (cm²/s) 12

Minority lifetime (µs) 3



Steady diffusion with recombination

Diffusion length 60 µm. The surface flux / boundary density equals D/L = 2000 cm/s. Constant diffusivity, low injection, semi-infinite sample.

Why an optically injected electron–hole packet spreads at one rate

A short light pulse creates equal excess densities of electrons and holes. Independent diffusion would let the faster carriers escape first, leaving a local excess charge. Derive the internal field and the shared packet width measured after the excitation.

Eliminate the internal field from the two currents

In a nondegenerate semiconductor let $q > 0$ be the elementary charge. The conventional charge currents along x are $J_n = q\mu_n n E + qD_n\,\partial_x n$ and $J_p = q\mu_p p E - qD_p\,\partial_x p$. The opposite signs of their diffusive terms follow from the opposite carrier charges. For a weak, slowly varying packet in a homogeneous region, quasineutrality sets $\delta n = \delta p = u$. In an open circuit away from the dielectric relaxation transient take $J_n + J_p = 0$. The spatial derivatives of the equilibrium densities vanish.

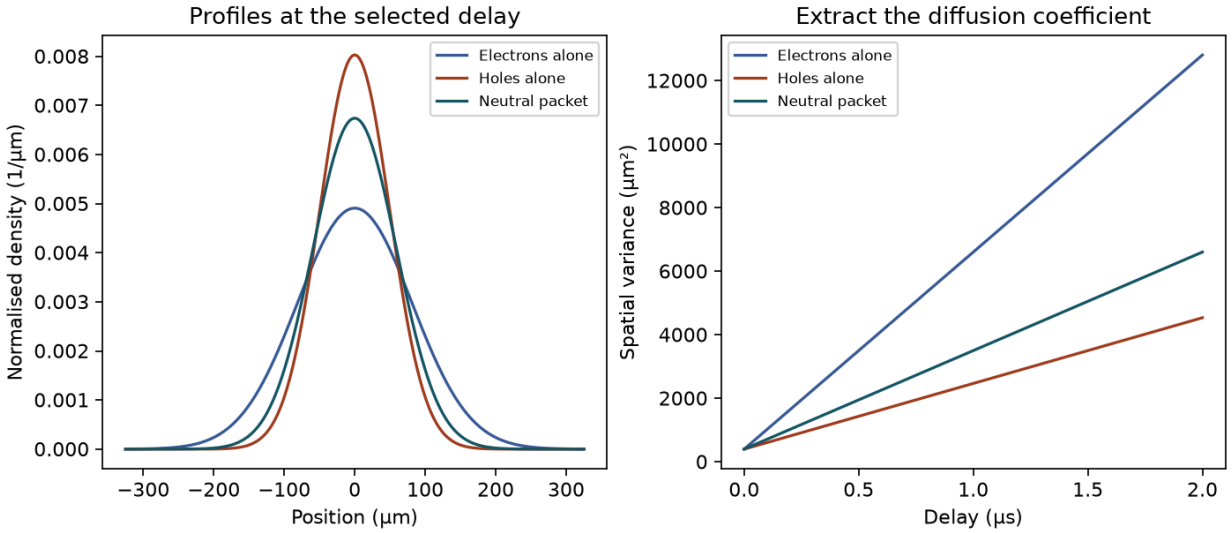
The resulting field is

$$E = -(D_n - D_p)\partial_x u/(\mu_n n + \mu_p p).$$

An electron particle flux is $-J_n/q$ and a hole particle flux is J_p/q. With total charge current zero the two are therefore equal, even while each species moves. Substitute the field back into either flux to obtain a common particle flux $-D_a\partial_x u$, where

$$D_a = \frac{\mu_n n D_p + \mu_p p D_n}{\mu_n n + \mu_p p}.$$

In a heavily electron-doped sample with a weak optical perturbation, $\mu_n n$ can dominate the denominator. The common coefficient approaches the minority-hole D_p. In an intrinsic sample with equal carrier densities and the Einstein relation $D_i = \mu_i k_B T/q$, it becomes the harmonic mean $2D_nD_p/(D_n + D_p)$. These are distinct physical limits of one current balance. A conductivity-weighted arithmetic average with the same diffusivity paired to its own mobility would predict the wrong limit.



All profiles carry the same integrated excess. The area is therefore conserved even though the peak gets lower as the packet widens.

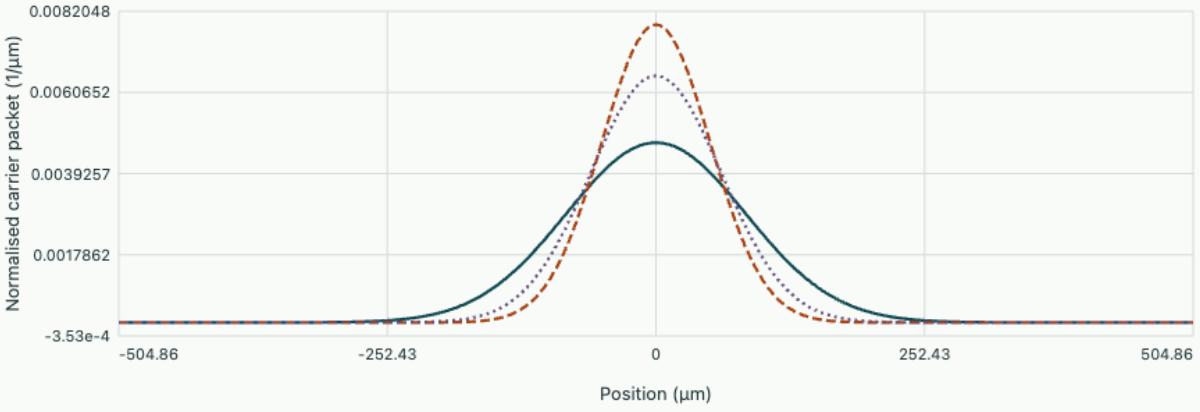
Try it · Ambipolar

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

Temperature (K)	300
Electron / hole background density	1
Delay (µs)	1



Ambipolar diffusivity 0.0018308 m²/s; packet rms width 61.332 µm. Mobilities 0.135 and 0.048 m²/(V s); initial rms width 10 µm. Integrated carrier excess is conserved.

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

For mobilities 0.12 and 0.04 m²/(V s) at 300 K, the independent diffusion coefficients are about 0.00310 and 0.00103 m²/s. A neutral packet at equal densities has $D_a = 0.00155$ m²/s. In one dimension, convolution with a diffusion kernel adds $2D_a t$ to the initial variance. An initial Gaussian rms width of 20 micrometres therefore becomes about 59.2 micrometres after one microsecond. The result is a width of the carrier concentration; an optical signal proportional to a product of populations can have a different width.

Change the background density ratio and inspect the slope of the variance against delay. This linear calculation treats D_a as approximately uniform. Strong spatial variation in the density ratio makes the transport law nonlinear. A short uniform recombination lifetime reduces the amplitude and integrated count while leaving the width law of an otherwise homogeneous weak packet intact. Surface recombination or a strongly density-dependent rate can change its shape as well.

A fit gives an rms width increase from 30 to 70 micrometres in two microseconds. Which diffusion coefficient would it imply, and why is dividing the width difference by the time unsuitable?

▼ **Work through the solution**

The increase in variance is $70^2 - 30^2 = 4000$ square micrometres. Dividing by $2t = 4$ microseconds gives 1000 square micrometres per microsecond, or 0.0010 m²/s. The difference of the widths divided by time has units of speed and depends on the initial spot size. A constant imaging blur adds a fixed variance that cancels in the difference.

Open editable notebook Download .ipynb

Integrate the occupation near a band edge

For a three-dimensional parabolic conduction band, the density of states per volume grows as $\sqrt{E - E_c}$. If $E_c - \mu$ is several $k_B T$, the Fermi function is well approximated by $e^{-(E-\mu)/k_B T}$. Substituting $x = (E - E_c)/k_B T$ in the carrier integral produces a temperature-dependent prefactor and an exponential separation from the band edge. Carrying out the corresponding hole integral at the valence-band maximum gives

$$n = N_c e^{-(E_c - \mu)/k_B T},$$

$$p = N_v e^{-(\mu - E_v)/k_B T},$$

$$np = N_c N_v e^{-E_g/k_B T} = n_i^2.$$

Here N_c, N_v include spin and any relevant valley degeneracies, $E_g = E_c - E_v$, and n_i is the intrinsic concentration. In a parabolic band the prefactors scale as $T^{3/2}$. The product relation assumes thermal equilibrium with a common chemical potential. Under illumination electrons and holes can have different quasi-Fermi levels; using the same mass-action product there requires additional justification.

Close the equations with charge neutrality

In an intrinsic uniform sample $n = p = n_i$. Equating the two exponentials gives

$$\mu = (E_c + E_v)/2 + (k_B T/2) \ln(N_v / N_c).$$

The intrinsic chemical potential is at midgap only if the effective state counts match. With fully ionised donors and acceptors, neutrality reads $n - p = N_D - N_A = D$. Combining it with mass action gives

$$n = (D + \sqrt{D^2 + 4 n_i^2})/2, \; p = n_i^2 / n.$$

For large negative D, computing p first from the corresponding positive root avoids subtracting nearby numbers.

For $D = 10^{16}$ cm⁻³ and $n_i = 10^{10}$ cm⁻³, the electron density is essentially 10^{16} cm⁻³ while p is about 10^{4} cm⁻³. Adding net donors moves the Fermi level upward and suppresses holes. At low temperature incomplete ionisation can matter, and at high donor concentration the nondegenerate approximation can fail. At sufficiently high temperature the intrinsic density can grow comparable to the dopant density. Use both balance equations when comparing data across such a crossover.

From a drift balance to a measured signal

In a single isotropic parabolic band a relaxation-time model balances electric acceleration against momentum relaxation. For electrons $m^*\dot{v} = -eE - m^*v/\tau$, so the steady drift velocity is $-e\tau E/m^*$. Define the positive mobility $u_n = e\tau/m^*$. Multiplying by electron charge and density gives current parallel to the field. With holes present, $\sigma = e(nu_n + pu_p)$. A conductivity alone supplies a product of concentration and mobility and cannot separate the two.

For one carrier type with charge q, zero transverse current implies

$$E_y = v_x B_z.$$

Substituting $j_x = nqv_x$ yields $R_H = E_y/(j_x B_z) = 1/(nq)$ in this simple model. Combining its sign and magnitude with σ gives a density and mobility. If both types contribute, in weak fields the two-band expression becomes

$$R_H = (pu_p^2 - nu_n^2)/[e(pu_p + nu_n)^2].$$

The mobility squared makes a highly mobile minority carrier disproportionately visible. Interpreting every measured Hall coefficient as the inverse density can therefore misidentify both carrier count and dominant mechanism.

Set up electrostatics at a junction between two dopings

A junction between p-type and n-type regions can sustain an electric field in thermal equilibrium. On contact, carriers diffuse down their concentration gradients. The uncovered, immobile dopant ions leave a negative region on the p side and a positive one on the n side. This space charge creates a field opposing further majority-carrier diffusion. Consider an abrupt, one-dimensional junction in one nondegenerate semiconductor. Assume complete ionization, uniform dopant densities N_A, N_D, no applied bias and neutral regions far from the junction.

The equilibrium Fermi level is constant. Equivalently, Boltzmann carrier densities on opposite sides and the mass-action relation $np = n_i^2$ imply a built-in voltage magnitude $V_{\mathrm{bi}} = (k_B T/e)\ln(N_A N_D/n_i^2)$. In the depletion approximation mobile carriers are neglected across widths x_p and x_n adjoining the junction. Requiring field zero at both outer edges enforces charge neutrality $N_A x_p = N_D x_n$. Inside either side $dE/dx = \rho/\epsilon$ is constant, so the field profile consists of two straight segments meeting at its maximum magnitude. Integrating the triangular field profile gives

$$V_{\mathrm{bi}} = \frac{1}{2} E_{\max} W,$$

$$W = x_p + x_n = \sqrt{\frac{2\epsilon V_{\mathrm{bi}}}{e}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)}.$$

Most of the width lies on the lightly doped side because more volume there is needed to uncover the same fixed charge per area. This relation follows from charge balance and need not be imposed as another boundary condition.

Predict a width and a measurable differential capacitance

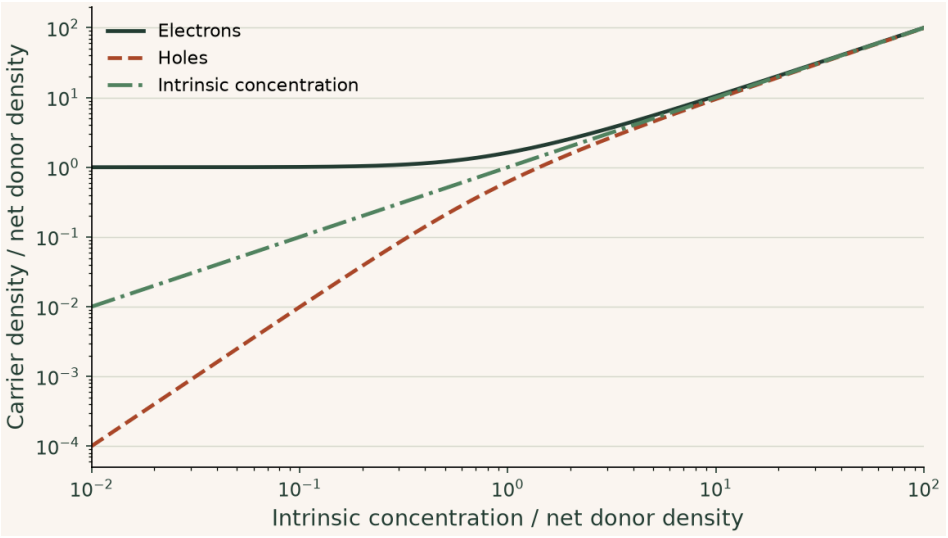
Take silicon-like $\epsilon = 11.7\epsilon_0$, $n_i = 1.00 \times 10^{16}$ m⁻³, $N_A = 1.00 \times 10^{23}$ m⁻³ and $N_D = 1.00 \times 10^{22}$ m⁻³ at 300 K. The logarithm is $\ln(10^{13}) = 29.93$, giving $V_{\mathrm{bi}} = 0.774$ V. The depletion width is 0.332 µm, partitioned as $x_n = 0.302$ µm and $x_p = 0.0302$ µm. The peak electric field is $2(0.774)/(0.332\,\mu\mathrm{m}) = 4.66 \times 10^6$ V/m. The built-in voltage is an internal electrostatic difference; equilibrium contact potentials complicate interpreting it as a terminal voltmeter reading.

A reverse bias of magnitude 0.500 V increases the voltage across the depletion layer to 1.274 V and multiplies its width by $\sqrt{1.274/0.774} = 1.28$. Differentiating the uncovered charge per area with respect to reverse bias gives a differential capacitance per area ϵ/W. It falls from about 0.312 to 0.244 mF/m² in this sweep. A plot of reciprocal squared capacitance against reverse voltage is therefore approximately linear for uniform doping. A curved graph can indicate a graded interface, traps or departure from the simple depletion approximation.

The abrupt-junction model predicts how the measured capacitance changes if the device area is doubled with doping and bias held fixed. Terminal capacitance doubles, while the inferred depletion width and field maximum stay fixed. Perimeter leakage and fringe capacitance instead scale differently with lateral device size, so comparing multiple diode areas can help isolate the interior junction contribution. Only after this separation is it reasonable to interpret a slope of the inverse squared capacitance as a dopant-density measurement. Frequency dependence provides another diagnostic since slowly responding traps can alter a low-frequency capacitance without changing the static depletion geometry.

Locate the crossover from doped to intrinsic conduction

Assume fully ionised net donors of density $D > 0$ and a nondegenerate equilibrium semiconductor. Set $r = n_i/D$; solve neutrality and mass action together rather than discarding minority carriers when the intrinsic density rises.



At r = 1 the electron and hole densities are about 1.618D and 0.618D. Their difference still equals D. When r is much larger than one the two curves nearly coincide in relative terms, but the charge-balance difference persists.

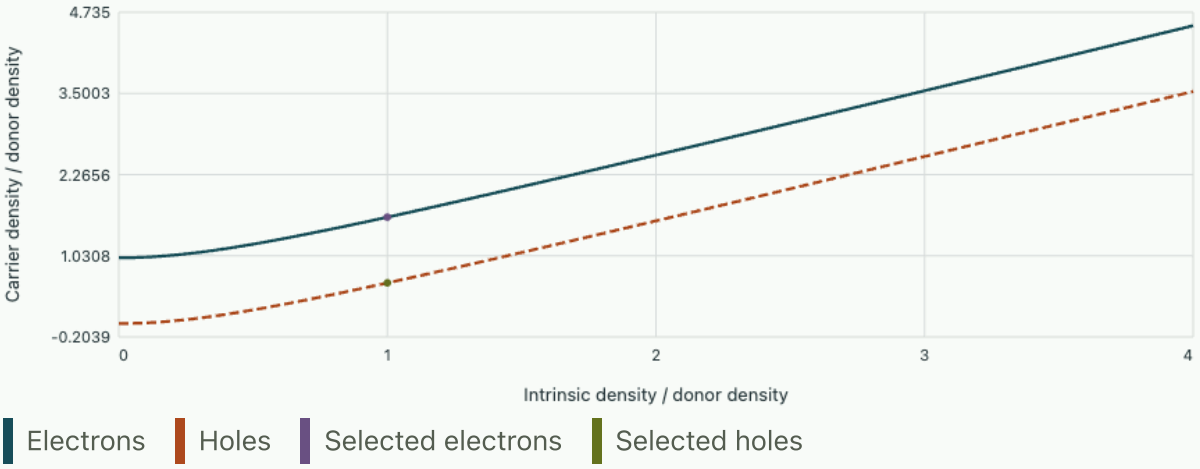
Try it · Carriers

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

Intrinsic density / donor density 1



n−p = 1 donor-density unit; np = 1. Complete donor ionisation and nondegenerate equilibrium statistics are assumed.

▶ **Read values from the plot**

Study the full derivation

Why a large joint density of states can leave a dark optical edge · Structure, excitations and probes of solids

A band diagram gives the energy an electron needs to move between bands. It does not by itself tell whether light couples those states. Construct a two-orbital model with an explicit velocity operator, calculate its allowed and suppressed transitions, and distinguish electronic band curvature from optical selection rules.

Chapter 6 · Predicting a gap and testing it in a periodic medium

Combine the translation argument, a measurable band edge and an independent numerical calculation to distinguish a gap from a folded crossing.

FROM FIRST PRINCIPLES

Infer pass bands and stop bands from a repeated cell

For a normally incident monochromatic plane wave in a lossless layered dielectric, record its tangential electric and magnetic fields in a two-component state. Interface conditions transmit these same tangential fields between adjacent layers. Propagation through one complete cell of period a is a linear map represented by a matrix $M(\omega)$. Forward and backward propagation in each lossless layer have reciprocal phase factors, so the cell transfer matrix has determinant one. A Bloch state reproduces its field after one cell up to a multiplier:

$$Mv = \lambda v,$$
$$\lambda = e^{ika},$$
$$\lambda^2 - (\operatorname{tr} M)\lambda + 1 = 0.$$

The second root is $1/\lambda$. Adding the two roots and using the phase relation gives the band condition

$$\cos(ka) = \tfrac{1}{2}\operatorname{tr} M(\omega).$$

For a real lossless layered system the trace is real. Where its half lies between –1 and 1 there are real wave numbers. Outside this interval the two multipliers have different magnitudes and one solution decays with each successive cell. In a uniform material the trace simply reproduces homogeneous propagation: folding the dispersion into the first zone can make branches touch without opening a gap. Index contrast allows repeated reflections and can create finite stop bands.

DERIVATION AND APPLICATIONS

Find the long-wavelength optical index of a layered cell

A finite wavelength can resolve two different dielectrics in a repeating stack. Far below the first band edge, what uniform medium would reproduce the phase accumulated per cell? Consider lossless, nonmagnetic layers of thickness d_1, d_2, real refractive indices n_1, n_2, and cell length $a = d_1 + d_2$. At normal incidence the field is tangential to the interfaces. Define $k_0 = \omega/c$ and the phase thickness $\delta_j = k_0 n_j d_j$. For Bloch wavevector K, the exact trace relation is

$$\cos(Ka) = \cos\delta_1 \cos\delta_2 - \frac{1}{2}\left(\frac{n_1}{n_2} + \frac{n_2}{n_1}\right)\sin\delta_1 \sin\delta_2.$$

On the acoustic-like branch through zero frequency, all three phase angles are small. Keep the quadratic terms in each cosine and the product of the linear terms in the sines. Constant terms on both sides cancel. Substitution of the optical phases gives

$$(Ka)^2 \simeq k_0^2 \left[n_1^2 d_1^2 + n_2^2 d_2^2 + (n_1^2 + n_2^2) d_1 d_2 \right].$$

The mixed term records the interfaces even at long wavelengths. Factor the bracket as $a(n_1^2 d_1 + n_2^2 d_2)$ and take the forward-propagating root. Writing the first layer's thickness fraction as $f = d_1/a$ yields

$$n_{\mathrm{eff}} = \lim_{\omega \to 0} \frac{K}{k_0} = \sqrt{f n_1^2 + (1-f) n_2^2}.$$

This is the thickness-weighted average of the relative permittivities under a square root. It agrees with a tangential electric field that is nearly uniform across a small cell. As a thickness fraction tends to unity the remaining material's index is recovered. Near a band edge the omitted fourth-order terms in the trace matter, so a constant effective index cannot predict the gap.

WORKED EXAMPLE

Measure the phase of a deeply subwavelength stack

Change the inputs and investigate this derivation.

A cell contains $10.0\,\mathrm{nm}$ of index 2.00 and $30.0\,\mathrm{nm}$ of index 1.50. Its effective index is

$$\sqrt{0.25(4) + 0.75(2.25)} = 1.64.$$

At vacuum wavelength $4.00\,\mu\mathrm{m}$, ten cells accumulate Bloch phase $2\pi n_{\mathrm{eff}}(0.400/4.00) = 1.03$ radians. The larger layer phase per cell is just 0.071 radians, consistent with the expansion.

SECOND APPLICATION

Infer a layer fraction from a low-frequency slope

Change the inputs and investigate this derivation.

Another stack uses indices 3.00 and 1.50. Its long-wavelength measured phase velocity is $c/2.10$. Solving the squared-index average gives

$$f = (2.10^2 - 1.50^2)/(3.00^2 - 1.50^2) = 0.320.$$

A $50.0\,\mathrm{nm}$ cell should contain $16.0\,\mathrm{nm}$ of the high-index material. A microscope can check that inference independently of the optical phase measurement.

Try it · Homogenised stack

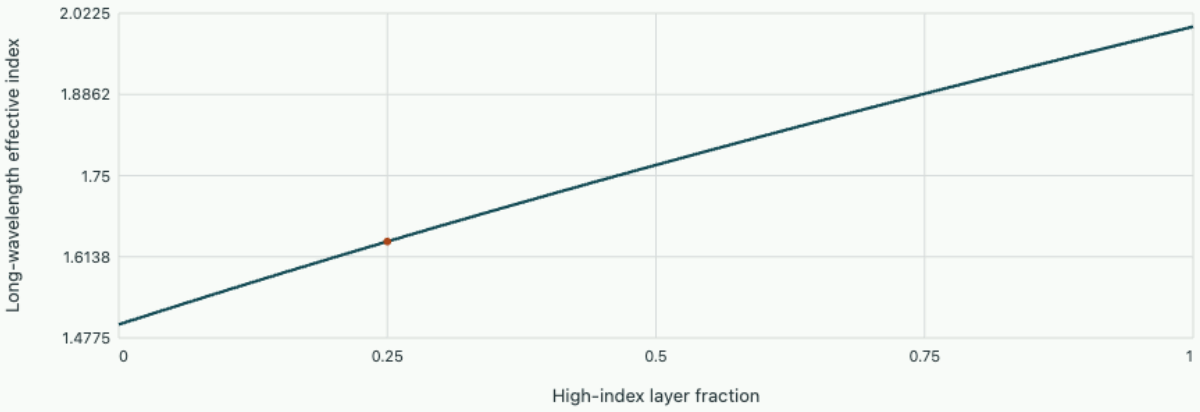
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

High-index layer fraction **0.25**

High refractive index **2**



Normal-incidence dielectric stack | Selected fraction

Low index is 1.5. The squared indices, rather than indices, average in this long-wavelength transverse-field limit. The cell must be much shorter than the wavelength.

Start from this chapter. With no index contrast the folded branches meet. Predict what changing only the second dielectric layer does to the two zone-edge frequencies, then compare the finite-element and transfer-matrix results.

OneSimulation laboratory · Open a band gap in a periodic dielectric

Transfer a field by one period

Consider a lossless stack of two nonmagnetic dielectrics at normal incidence. Layer j has real refractive index n_j, thickness d_j and optical phase $\delta_j = 2\pi f n_j d_j / c$. Within a layer the electric field is a sum of forward and backward waves; the corresponding magnetic field has the opposite sign for the backward wave. Eliminating the wave amplitudes between the two faces gives the layer transfer matrix for electric field and magnetic field scaled by the vacuum impedance (one time-phase convention):

$$M_j = \begin{pmatrix} \cos\delta_j & i\sin\delta_j/n_j \\ i n_j \sin\delta_j & \cos\delta_j \end{pmatrix}, \qquad \det M_j = 1.$$

The tangential fields are continuous at a simple dielectric interface. Thus the matrix of one complete cell of size $a = d_1 + d_2$ is the product of the two layer matrices. Its eigenvalues multiply to one. For propagation in an infinite periodic stack the Bloch multipliers must be $e^{\pm ika}$. Their sum fixes the half-trace:

$$\cos(ka) = \cos\delta_1\cos\delta_2 - \tfrac{1}{2}(n_1/n_2 + n_2/n_1)\sin\delta_1\sin\delta_2.$$

If the right side lies outside $[-1, 1]$, no real k exists. One solution decays and the reciprocal one grows along the chosen direction. A semi-infinite passive medium selects the decaying solution inside a stop band. A finite sample can still transmit an evanescent field, especially when it contains few cells.

Derive a check that the numerical solver could fail

Set $n_1 = n_2 = n$. The right side reduces to $\cos(\delta_1 + \delta_2)$ by the addition formula. The first two folded branches meet where the combined phase is π, at $fa/c = 1/(2n)$. There is no interface reflection and no forbidden interval. With $n = \sqrt{13}$ both branches should meet at about 0.138675. Two nearly equal eigenvalues on a finite mesh must be compared with the numerical resolution before reporting a narrow gap.

As another analytic limit, take each layer a quarter wavelength thick at the same frequency, so

$$\delta_1 = \delta_2 = \pi/2.$$

The half-trace is $-\tfrac{1}{2}(r + r^{-1})$, $r = n_1/n_2$. Its magnitude exceeds one whenever $r \neq 1$. For $r = 2$ the decaying multiplier has magnitude $1/2$ per cell. The attenuation here arises from coherent reflections in a lossless model. Small transmission need not mean that electromagnetic energy was dissipated.

Read the finite-element eigenvalues against the calculation

In the OneSimulation band experiment, keep the layer fractions fixed and change only the second permittivity. For the pair (13,12) the two reported zone-edge frequencies are 0.1396 and 0.1433 in reduced units; the gap is about 2.6% of their mean. Compare this separation with the zero-contrast result and with the two roots of the half-trace equation at -1. For the much larger contrast (13,1), using a weak-perturbation approximation would be hard to justify. The transfer matrix still includes repeated scattering exactly for these ideal layers.

When investigating an unfamiliar contrast, halve the spatial mesh size and repeat the edge calculation. Track modes by their fields as well as by sorted eigenvalue if two branches nearly meet. For the scalar one-dimensional normal-incidence problem the observed stop band has a precise meaning. It is an interval of optical frequencies for that propagation geometry. Inferring a complete three-dimensional gap would require other directions and polarisations. The link to electronic bands is the shared translation

eigenvalue and coherent coupling at a crossing; electronic filling and the optical transfer-matrix test answer different physical questions.

Turn an infinite stop band into finite-stack reflectance

The forbidden Bloch interval of an infinite periodic medium does not imply perfect reflection from a finite sample. Its termination and number of periods matter. Consider normal incidence on alternating lossless dielectric layers with refractive indices n_H, n_L, each of optical thickness $\lambda_0/4$ at the selected vacuum wavelength. Treat all permeabilities as μ_0 and use normalized optical admittance $Y = n$. Across a quarter-wave layer the characteristic matrix mapping the fields on its far face to its entrance has zero diagonal and off-diagonal entries i/Y and iY, with an overall phase convention fixed consistently.

For a terminating medium of admittance Y_{load} the resulting input admittance is $Y_{\mathrm{layer}}^2/Y_{\mathrm{load}}$. Applying this inversion first to the rear low-index layer and then to the front high-index layer multiplies load admittance by $(n_H/n_L)^2$. After N identical pairs, with the high-index layer first as seen by incident light,

$$Y_{\mathrm{in}} = n_s (n_H/n_L)^{2N},$$

$$R = \left| \frac{n_0 - Y_{\mathrm{in}}}{n_0 + Y_{\mathrm{in}}} \right|^2 .$$

Here n_s is the substrate index and n_0 the incident medium index. This design-frequency result follows from phase accumulation in each layer. For a general wavelength the matrices are no longer simple inversions and both diagonal entries must be retained. Interchanging the order of high and low layers in a finite stack also changes the input admittance and reflection phase.

Select a period count for a required mirror

Let $n_0 = n_s = 1.00$, $n_H = 2.00$ and $n_L = 1.50$ at $\lambda_0 = 600$ nm. Quarter-wave physical thicknesses are

$$600/(4 \times 2.00) = 75.0\,\mathrm{nm},$$

$$600/(4 \times 1.50) = 100\,\mathrm{nm}.$$

Four pairs give

$$Y_{\mathrm{in}} = (4/3)^8 = 9.99,$$

$$R = [(9.99 - 1)/(9.99 + 1)]^2 = 0.669.$$

Eight pairs give $Y_{\mathrm{in}} \simeq 99.8$ and $R = 0.961$. Thus a fairly clear infinite-medium stop band can coexist with substantial transmission through a short specimen.

For at least 99.0 percent reflectance at the design wavelength the required admittance ratio is $(1 + \sqrt{0.990})/(1 - \sqrt{0.990}) \simeq 398$. Solving $(4/3)^{2N} \geq 398$ gives $N \geq 10.4$, hence eleven complete pairs. If the indices become equal and match both external media, $Y_{\mathrm{in}} = n_0$ for any period count and the reflection vanishes. This finite-stack result tests a different observable from an infinite cell

eigenfrequency. It connects the gap calculation to a transmission or reflectance measurement while exposing the additional data, particularly termination and thickness, needed for that prediction.

At the chosen wavelength the reflection amplitude for high-index-first ordering is negative because the effective admittance exceeds the incident admittance. Measuring reflected intensity alone loses this phase sign. Interference with a calibrated reference reflector can recover the sign and identify termination differences even when two samples have similar reflectances. This matters when the periodic mirror forms one end of a resonant cavity: the reflection phase contributes to the total round-trip phase and can move the resonance frequency. A pair of finite mirrors can therefore share a stop-band range while giving different cavity phase conditions.

Derive the quarter-wave stop-band width

The finite mirror is easy to evaluate at the design wavelength. The same stack admits an analytic expression for the width of its infinite first stop band. In both layers the optical phase thickness is $\delta = (\pi/2)(\omega/\omega_0)$ when refractive indices are treated as nondispersive. The one-period transfer matrix has unit determinant and half-trace $\cos(Ka)$, where K is the Bloch wavenumber and a the physical period. Multiplication of the two layer matrices gives

$$\cos(Ka) = \cos^2\delta - \frac{1}{2}\left(\frac{n_H}{n_L} + \frac{n_L}{n_H}\right)\sin^2\delta.$$

At band centre the half-trace is less than minus one for unequal indices, so K has a nonzero imaginary component. At either neighbouring edge set the half-trace equal to minus one and solve for δ. The two edge phases are $\pi/2 \pm \arcsin[(n_H - n_L)/(n_H + n_L)]$ for $n_H > n_L$. Their separation gives fractional angular-frequency width

$$\frac{\Delta\omega}{\omega_0} = \frac{4}{\pi}\arcsin\left(\frac{n_H - n_L}{n_H + n_L}\right).$$

With indices 2.00 and 1.50 this is 0.183. A central frequency of 500 THz thus has edges near 454 and 546 THz. The frequency interval is symmetric about the centre in this ideal quarter-wave construction; the corresponding wavelength interval is asymmetric because wavelength is reciprocal to frequency. As the two indices merge, the edge separation vanishes even though an artificially repeated unit cell still has a Brillouin-zone boundary.

At centre the decaying Bloch amplitude falls by $n_L/n_H = 0.750$ per period. A physical period of 175 nm therefore gives a field-amplitude decay length $175/\ln(2/1.5) = 608$ nm. This infinite-medium length explains the strong dependence of the finite mirror on period count. The external-medium matching factors are still needed to turn that internal decay into the actual reflected intensity.

Study the full derivation

Construct allowed bands from a cell transfer matrix · Structure, excitations and probes of solids

In a finite well a bound level follows from matching its wavefunction at the two walls. A periodic lattice replaces those two external walls by a matching condition on every cell. Derive a band criterion that works without guessing a wave inside the forbidden gap and use it to predict transmission through a finite sequence of cells.

Problem sheet 3 · Distinguishing carriers, reflection and attenuation

Test an inference about a material with an independent balance, a limiting case or a second method.

1. Conductivity and Hall data

A film has $R_H = -6.25 \times 10^{-4}$ m³/C and $\sigma = 160$ S/m. Assume one isotropic carrier band and $e = 1.60 \times 10^{-19}$ C. (a) Determine carrier sign and density. (b) Determine the mobility and drift speed at 50 V/m. (c) Explain what becomes ambiguous if a second band is populated.

▼ **Recover the two independent quantities**

The sign indicates electrons in this model; $n = 1/(e|R_H|) = 1.00 \times 10^{22}$ m⁻³. The mobility is $\sigma|R_H| = 0.100$ m²/(V·s), giving drift speed 5.0 m/s opposite the electric field. With two carrier types the conductivity is a mobility-weighted sum while the Hall numerator contains a signed mobility-squared combination. The two observations alone cannot identify four concentrations and mobilities.

---

2. Compensation and the mass-action law

At a given temperature $n_i = 10^{10}$ cm⁻³. (a) Solve neutrality for net donor density $D = 1.5 \times 10^{10}$ cm⁻³. (b) Reverse the sign of D and recompute. (c) If $N_c = N_v$, infer the shift of the chemical potential from its intrinsic value in units of $k_B T$.

▼ **Use the positive concentration root**

For positive D, $\sqrt{D^2 + 4n_i^2} = 2.5 \times 10^{10}$ cm⁻³, giving $n = 2.0 \times 10^{10}$ and $p = 0.50 \times 10^{10}$ cm⁻³. Reversing the compensation interchanges the two concentrations. From $n/n_i = e^{(\mu-\mu_i)/k_B T}$ the shifts are $+\ln 2$ and $-\ln 2$, respectively. In this example dropping minority carriers would overestimate the imbalance needed to obtain a given majority concentration.

---

3. Does the computed gap survive two controls?

Read the optical band-edge data. (a) Calculate the relative gap for (13,1) from the displayed edges. (b) Give an absolute-frequency prediction for a uniform (13,13) cell of length 2 mm. (c) Propose one numerical and one physical test to distinguish a real narrow gap from a mesh artefact.

▼ **Compare scale, contrast and discretisation**

The relative gap is $2(0.2566 - 0.1509)/(0.2566 + 0.1509) = 0.519$, subject to the rounded edges. In a uniform cell $f = c/(2na) \simeq 20.8$ GHz at the folded crossing, with no gap. Refine the mesh until the separation is insensitive to further refinement at the claimed precision. Then reduce contrast to zero at

fixed geometry: the splitting must close. An independent layer-matrix computation at the same nonzero contrast can check the two edges without spatial meshing.

Explore these relations

- For compensation and mass action, compare electron and hole populations as intrinsic density increases; the live model assumes complete donor ionisation. Open the live calculation.
- For the optical-gap controls, compare the actual equal-width layer stack at permittivities 13/13, 13/12 and 13/1. Open the live calculation.

Reconstruct the edges of a photonic gap

A periodic dielectric has a midgap reduced frequency $f_{\mathrm{mid}}a/c = 0.20375$ and relative width 51.89%. Recover the two edges and compare them with the high-contrast run. A geometrically similar cell has period 2 mm. Estimate the corresponding frequencies in GHz using $c = 3.00 \times 10^{8}$ m/s.

▼ **Work through the solution**

Writing the edges x_{-}, x_{+} in reduced units gives

$$(x_{+} + x_{-})/2 = 0.20375,$$

$$x_{+} - x_{-} = 0.5189(0.20375).$$

Hence $x_{\pm} = 0.20375(1 \pm 0.5189/2)$, or 0.1509 and 0.2566. Multiplying by $c/a = 150$ GHz gives about 22.6 and 38.5 GHz. The same reduced spectrum can occupy different frequency bands in hertz when all lengths scale together and the permittivities are frequency-independent.

Use the OneSimulation laboratory to check your answer

Set the second permittivity to 1 and the period to 2 mm. Compare both reduced gap edges and their physical frequencies with the reconstruction. Open the corresponding live laboratory reference.

Notebook practice · Why an optically injected electron–hole packet spreads at one rate

Chapter 7 · Transport, scattering and the mean free path

Connect an equilibrium band structure to a current-carrying distribution and determine what a measured resistivity says about scattering.

FROM FIRST PRINCIPLES

Derive a transport coefficient from momentum balance

Measure an electrical current in a homogeneous, isotropic conductor with one type of mobile carrier. Approximate the band by a parabolic effective mass m^*, and suppose collisions relax the mean drift momentum with a constant time τ. The random equilibrium velocities average to zero. A weak spatially uniform electric field gives a carrier of signed charge q a systematic force $q\mathbf{E}$, so the drift equation is

$$m^* \frac{d\mathbf{v}_d}{dt} = q\mathbf{E} - \frac{m^*}{\tau}\mathbf{v}_d.$$

For constant field, set the time derivative to zero after the transient has decayed. Multiplying the resulting drift by charge and by the number density n converts it into current per area:

$$\mathbf{j} = nq\mathbf{v}_d = \frac{nq^2\tau}{m^*}\mathbf{E},$$

$$\rho_{\mathrm{el}} = \frac{m^*}{nq^2\tau}.$$

The current is along the field for either sign of charge because both the force and the transported charge reverse. A measured resistivity determines a combination of density, mass and scattering time. Holding n and m^* fixed makes changes in resistivity a probe of scattering. An activated carrier density in a semiconductor can change the same measurement through a different mechanism.

DERIVATION AND APPLICATIONS

Separate collision time from carrier density with alternating current

A steady conductivity determines a product of carrier density and relaxation time. A field that reverses during a collision time puts an additional timescale into the measurement. Model a spatially uniform field in a single isotropic parabolic band. Take carrier charge $-e$, mass m^*, density n, and a constant momentum relaxation time τ. Neglect changes of the equilibrium density, interband absorption and internal field gradients. For a harmonic drive use $E(t) = \mathrm{Re}[E_0 e^{-i\omega t}]$. The mean drift velocity obeys

$$m^* \frac{dv}{dt} = -eE - \frac{m^* v}{\tau}, \qquad (1/\tau - i\omega)v_0 = -\frac{e}{m^*}E_0.$$

Multiplying v_0 by the mobile charge density gives the current density amplitude. The complex conductivity relates it to the applied field,

$$\sigma(\omega) = \frac{J_0}{E_0} = \frac{\sigma_0}{1 - i\omega\tau},$$

$$\sigma_0 = \frac{ne^2\tau}{m^*}.$$

Multiplication by the complex conjugate of the denominator separates the response that dissipates work from the response a quarter cycle out of phase,

$$\sigma' = \frac{\sigma_0}{1 + (\omega\tau)^2},$$

$$\frac{\sigma''}{\sigma'} = \omega\tau,$$

$$\langle JE \rangle = \tfrac{1}{2}\sigma' |E_0|^2.$$

The power expression uses a peak electric field amplitude. Dissipation drops at high frequency because carriers have less time to develop drift before the driving field reverses. For the time convention above a positive conductivity phase corresponds to a delayed current maximum. The magnitude and phase can test the single-time assumption independently when measured over a range of frequencies.

WORKED EXAMPLE

Read a collision time from a phase measurement

Change the inputs and investigate this derivation.

At $f = 2.00\,\mathrm{THz}$ a film gives

$$\sigma''/\sigma' = 0.500.$$

Thus

$$\tau = 0.500/(2\pi f) = 39.8\,\mathrm{fs}.$$

If $\sigma' = 8.00 \times 10^4\,\mathrm{S/m}$, the inferred steady conductivity is $1.00 \times 10^5\,\mathrm{S/m}$. With $m^* = 0.250 m_e$, the implied density is

$$\sigma_0 m^*/(e^2\tau) = 2.23 \times 10^{25}\,\mathrm{m^{-3}}.$$

SECOND APPLICATION

Predict the heating under a fixed peak field

Change the inputs and investigate this derivation.

A second material has $\sigma_0 = 2.50 \times 10^5\,\mathrm{S/m}$ and $\tau = 0.200\,\mathrm{ps}$. At $f = 1/(2\pi\tau) = 0.796\,\mathrm{THz}$ its dissipative conductivity is half its steady value. For a peak field of $20.0\,\mathrm{V/m}$ it absorbs $0.5(1.25 \times 10^5)(20.0)^2 = 2.50 \times 10^7\,\mathrm{W/m^3}$. The simultaneous quadrature current equals the in-phase current in amplitude but contributes zero cycle-averaged heating.

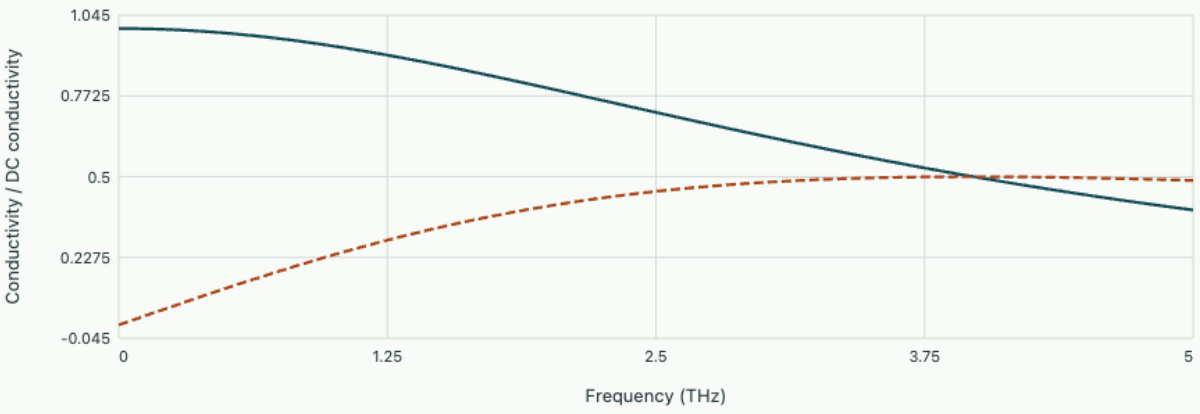
Try it · Drude

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

Collision time (fs) 40



Imaginary/real ratio equals ωτ. Heating uses the real conductivity and half the squared peak electric field.

Drive an occupation out of equilibrium

In a uniform metal the equilibrium occupation $f_0(E)$ is symmetric in velocity and has zero current. An electric field changes the crystal momentum through $\hbar\dot{\mathbf{k}} = -e\mathbf{E}$. A simple collision model relaxes the occupation toward equilibrium with time τ. To first order in a steady uniform field, the Boltzmann equation gives

$$\delta f = e\tau(\mathbf{E}\cdot\mathbf{v})\,\partial f_0/\partial E,$$

where $e > 0$. Since the energy derivative is negative, states moving opposite the field become more occupied. Multiplying by electron charge gives a conventional current along the applied field.

Collecting velocities in that current yields, for an isotropic band including both spins,

$$\sigma = \frac{e^2}{3}\int v(E)^2\tau(E)D_V(E)\left(-\frac{\partial f_0}{\partial E}\right)dE.$$

Here D_V is density of states per unit volume. The factor one third averages a squared velocity component in three dimensions. The derivative of the occupation restricts conduction in a degenerate metal to energies near E_F. A high density of states at one energy is insufficient on its own to predict high conductivity; velocities and transport lifetimes also enter.

Recover the Drude result as a special case

At low temperature replace $-\partial f_0/\partial E$ by a sharp weight at E_F. With a parabolic band and constant τ, $v_F^2 = 2E_F/m^*$ and $D_V(E_F) = 3n/(2E_F)$, giving $\sigma = ne^2\tau/m^*$. For $n = 8.5\times10^{28}$ m⁻³, $m^* = m_e$ and resistivity $1.7\times10^{-8}\ \Omega\,\mathrm{m}$, the inferred lifetime is 2.46×10^{-14} s. With $v_F = 1.57\times10^{6}$ m/s the transport mean free path is about 39 nm. This is much larger than an atomic spacing yet small relative to many bulk specimen dimensions.

Do not multiply τ by the much smaller average drift speed to estimate the mean free path. Scattering changes the momentum of already fast electrons; the tiny net imbalance in their occupations gives the much slower average drift. Likewise a transport lifetime weights deflections by how much they relax a directed current. Small-angle scattering can broaden a quantum state considerably while relaxing directed momentum relatively slowly.

Interpret a temperature sweep

Static impurities can give an approximately temperature-independent residual resistivity, while phonons add temperature-dependent scattering. If independent mechanisms have approximately additive transport rates,

$$\rho \simeq \rho_{\mathrm{imp}} + \rho_{\mathrm{ph}}(T).$$

This form is a model with a roughly common transport mass and carrier density. A temperature sweep through a carrier-density crossover or a change in band population calls for more information, such as Hall data. At low temperatures boundary scattering can become important once the nominal bulk mean free path approaches the sample width.

Compare heat and charge transport

Within a degenerate, nearly elastic model the same carriers also move excess energy. Integrating the squared energy relative to the chemical potential across the narrow thermal shell gives the Wiedemann–Franz ratio $\kappa_{\mathrm{el}}/(\sigma T) \simeq \pi^2 k_B^2/(3e^2)$. In SI units the ratio is about 2.44×10^{-8} W Ω K⁻². A measurement of total thermal conductivity includes phonons as well; subtracting this estimated electronic contribution is model dependent. A discrepancy can indicate inelastic scattering, lattice heat conduction or a breakdown of the simple band description.

Resolve electrical transport by carrier energy

A temperature gradient drives particles above and below the chemical potential in different directions. If their contributions to electrical current cancel exactly, no thermoelectric field is needed. Any asymmetry in the transport of the two groups leaves a field at open circuit. Introduce the energy-resolved transport function $\Xi(E)$, proportional to the density of states times mean squared velocity along the measured direction times relaxation time. For ordinary electron charge $-e$ with $e > 0$, define transport moments

$$L_j = \int dE\, \Xi(E)(E-\mu)^j[-\partial f/\partial E],$$

$$S = -\frac{L_1}{eTL_0}.$$

This sign convention makes a simple electron band with increasing transport function have negative Seebeck coefficient. The derivative of the Fermi function is sharply centred on μ in a degenerate material, with width of order $k_B T$. Expand $\Xi(E) = \Xi(\mu) + (E-\mu)\Xi'(\mu) + \cdots$. Odd moments of the symmetric thermal window vanish. Then $L_0 \simeq \Xi(\mu)$, while the first nonzero contribution to L_1 is $\Xi'(\mu)(\pi^2/3)(k_B T)^2$. This yields the low-temperature Mott relation

$$S \simeq -\frac{\pi^2 k_B^2 T}{3e}\left.\frac{d\ln \Xi}{dE}\right|_{E_F}.$$

A large conductivity by itself does not determine thermopower. Multiplying the whole transport function by a constant doubles conductivity but leaves its logarithmic slope and the leading Seebeck coefficient unchanged. The result assumes an electronic diffusive regime with sufficiently smooth energy dependence over the thermal window. Drag by nonequilibrium phonons or an abrupt band edge can violate these assumptions.

Infer the energy dependence of scattering

For a three-dimensional parabolic electron band let E be measured upward from its bottom. The density of states scales as $E^{1/2}$ and the squared speed as E. If $\tau(E) \propto E^r$, the transport function scales as $E^{3/2+r}$. At $T = 100$ K and $E_F = 0.200$ eV the ratio $k_B T/E_F = 0.0431$. With energy-independent scattering time the relation gives $S = -(\pi^2/3)(k_B/e)(0.0431)(1.5) = -18.3$ μV/K. A 5.00 K

temperature difference therefore produces a thermovoltage of magnitude about 91.5 μV in a measurement correcting for the reference leads.

If an independent band measurement supports the same Fermi energy but the observed coefficient is -24.4 μV/K, the inferred exponent is $3/2 + r \simeq 2.00$ or $r \simeq 0.50$ in this model. This inference uses a derivative and is sensitive to additional bands. For a particle–hole symmetric transport function around μ the odd moment vanishes and the predicted diffusive thermopower is zero even when electrical conductivity is large.

The Seebeck coefficient is usually measured relative to metallic leads, and equilibrium contact potentials are present even at zero temperature difference. Reversing the thermal gradient reverses the sought thermovoltage while leaving a fixed voltage offset unchanged. Taking half the difference of two reversed readings removes such an offset. A temperature-dependent reference lead contribution still has to be calibrated or subtracted. Without that measurement, extracting a scattering exponent from a few tens of microvolts can assign properties of the lead material to the sample even if the transport-function derivation itself is correct.

Taking readings at several gradient magnitudes further tests linearity of the fitted Seebeck response.

Study the full derivation

Count vacancies and infer an atomic migration barrier · Structure, excitations and probes of solids

Tracer atoms spread through a crystal even when the average lattice positions remain fixed. The rate can depend both on the number of vacant sites and on how readily an atom jumps into one. Derive the two factors separately, then compare an equilibrated specimen with a rapidly quenched one.

Chapter 8 · Cyclotron motion and quantised magnetic orbits

Relate a magnetic-field measurement to band curvature, level spacing and a cross-section of the Fermi surface.

FROM FIRST PRINCIPLES

Find the orbit and energy spacing imposed by a magnetic field

Consider an electron in an isotropic parabolic band with effective mass m^*, no electric field and a uniform magnetic induction $B\hat{\mathbf{z}}$. The Lorentz force is perpendicular to velocity, so it changes the direction of the transverse velocity without changing its magnitude. Taking absolute values in the radial force balance gives

$$\frac{m^* v_\perp^2}{r_c} = e v_\perp B,$$

$$r_c = \frac{m^* v_\perp}{eB},$$

$$\omega_c = \frac{eB}{m^*}.$$

To obtain energy levels one also needs quantum mechanics. Choose $\mathbf{A} = (0, Bx, 0)$ and electron charge $-e$ with $e > 0$. The Hamiltonian $(\mathbf{p} + e\mathbf{A})^2/(2m^*)$ is independent of y, z, so their canonical momenta can be treated as fixed in a plane-wave state. Complete the square:

$$H = \frac{p_x^2}{2m^*} + \frac{m^* \omega_c^2}{2}\left(x + \frac{p_y}{eB}\right)^2 + \frac{p_z^2}{2m^*}.$$

This is a harmonic oscillator about a centre set by p_y. Its transverse levels have energies $\hbar\omega_c(n + 1/2)$; longitudinal translation remains free. Spin splitting, anisotropic band curvature and collision broadening require further information. The same frequency occurs in the classical orbit and the quantum level spacing for distinct reasons made explicit by these two constructions.

DERIVATION AND APPLICATIONS

Determine which mass a cyclotron resonance measures

An orbit on an anisotropic band need not be circular. Consider a weak uniform magnetic induction B along z, and motion in a single band with elliptical constant-energy contours in the xy plane. Neglect scattering over an orbit, electric fields, and spin splitting. Write the local dispersion as $E = \hbar^2 k_x^2/(2m_x) + \hbar^2 k_y^2/(2m_y)$, measured from its minimum. For charge $-e$, the semiclassical force law and the velocity components give

$$\dot{k}_x = -\frac{eB}{m_y} k_y,$$

$$\dot{k}_y = \frac{eB}{m_x} k_x.$$

Taking a second derivative cancels the other component. Both coordinates oscillate at the same fundamental angular frequency even though their amplitudes differ,

$$\ddot{k}_x = -\frac{e^2 B^2}{m_x m_y} k_x,$$

$$\omega_c = \frac{eB}{m_c},$$

$$m_c = \sqrt{m_x m_y}.$$

A geometric measurement leads to the same mass. The two semiaxes of the contour at energy E are $\sqrt{2 m_i E}/\hbar$. Its reciprocal-space area $A(E)$ and energy derivative are

$$A(E) = \frac{2\pi E}{\hbar^2} \sqrt{m_x m_y},$$

$$\frac{\hbar^2}{2\pi} \frac{dA}{dE} = m_c.$$

The derivative explains how a measurement of orbit area over a small energy range tests the resonance result. Two bands with the same area at one energy can have different derivatives there. For a circle, both principal masses and the cyclotron mass coincide. A single resonance frequency for this orientation measures their geometric mean and leaves their ratio undetermined.

WORKED EXAMPLE

Predict the resonance of an elliptical pocket

Change the inputs and investigate this derivation.

A layer has $m_x = 0.100 m_e$, $m_y = 0.900 m_e$ and is placed in $B = 1.50\,\mathrm{T}$. The cyclotron mass is $0.300 m_e$, so the ordinary frequency is

$$f_c = eB/(2\pi m_c) = 140\,\mathrm{GHz}.$$

The period is $7.14\,\mathrm{ps}$. The long reciprocal-space semiaxis lies along y and is three times the short one, even though both components complete their cycle in exactly the same time.

SECOND APPLICATION

Use two nearby contour areas as a mass measurement

Change the inputs and investigate this derivation.

In a different sample the orbit area rises by $0.200\,\mathrm{nm}^{-2}$ when energy rises by $20.0\,\mathrm{meV}$. Treat the measured slope as locally constant. In SI units $dA/dE = 2.00 \times 10^{17}/(3.204 \times 10^{-21})\,\mathrm{m}^{-2}\mathrm{J}^{-1}$. The area formula gives

$$m_c = 1.10 \times 10^{-31}\,\mathrm{kg} = 0.121 m_e.$$

An independent determination of $m_x = 0.0600 m_e$ would then imply $m_y = m_c^2/m_x = 0.245 m_e$ at the displayed precision.

Try it · Cyclotron mass

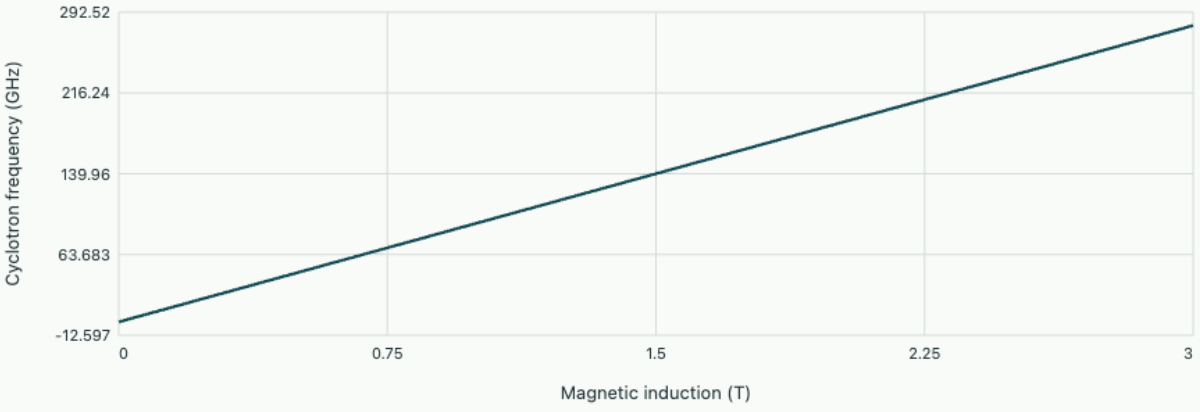
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Principal mass mx/me **0.1**

Principal mass my/me **0.9**



Elliptical parabolic pocket

Cyclotron mass / me = √(mx my) = 0.3. This geometric mean differs from the longitudinal effective mass.

Follow a wave packet in a band

A narrow wave packet in a well-isolated band moves with $\mathbf{v} = \hbar^{-1}\nabla_{\mathbf{k}}E$. In a static uniform magnetic field the semiclassical force law is

$$\hbar\dot{\mathbf{k}} = -e\,\mathbf{v}\times\mathbf{B}.$$

Thus $\nabla E\cdot\dot{\mathbf{k}} = 0$: the energy is constant because a magnetic field does no work. The wavevector component along $\mathbf{B}$ is also constant. The orbit in reciprocal space lies along the intersection of a constant-energy surface and a plane perpendicular to the field. Tilting the field samples different sections of a nonspherical Fermi surface.

For an isotropic parabolic band, motion transverse to the field is circular at $\omega_c = eB/m^*$. The real-space orbit radius is $v_\perp/\omega_c$. One can measure an effective mass from the resonant angular frequency without separately knowing the elastic scattering time. Whether the motion survives long enough to be resolved depends on $\omega_c\tau$. A very short scattering time smears the resonance and suppresses oscillatory magnetotransport.

Derive the quantised spacing

For a two-dimensional electron choose $A_y = Bx$ and $A_x = 0$. Since the Hamiltonian for charge $-e$ is $H = [p_x^2 + (p_y + eBx)^2]/(2m^*)$, a state of definite $p_y = \hbar k_y$ leaves a one-dimensional harmonic oscillator in x, centred at $-\hbar k_y/(eB)$. Its levels are $E_j = (j + 1/2)\hbar\omega_c$, neglecting spin. At fixed level, changing k_y moves the orbit centre without changing its energy. The centres allowed inside a rectangle of area A count to eBA/h per spin, up to edge corrections. In three dimensions motion along the field adds a dispersing kinetic-energy term.

Take $B = 5$ T and $m^* = 0.10\,m_e$. The orbital spacing is about 5.79 meV, equivalent to k_BT at 67 K. At 4 K the thermal scale is about 0.345 meV, smaller than the spacing. An observation still requires sufficiently weak disorder broadening and appropriate carrier filling. As the field changes, both the level spacing and the number of states carried by each level change.

Extract a reciprocal-space area from a measurement

For a closed semiclassical orbit of area $A_k(E)$, quantisation gives

$$A_k = 2\pi eB(j + \gamma)/\hbar,$$

where the phase offset γ depends on the orbit. Successive levels crossing the Fermi energy give oscillations periodic in reciprocal field with frequency $F = \hbar A_k(E_F)/(2\pi e)$, expressed in tesla. For $F = 100$ T, $A_k \simeq 9.55\times10^{17}$ m⁻². A circular section of that area has radius 5.51×10^{8} m⁻¹. In a three-dimensional metal dominant oscillations come from extremal sections. Inferring a full Fermi volume from one frequency requires additional assumptions about shape and symmetry; measurements at several orientations can test them.

Count orbit centres and Hall filling

Landau level spacing determines the energy cost of changing orbital index. A second question is how many independent states fit into each level. Take a rectangular two-dimensional layer of area $L_x L_y$, magnetic field B normal to it and Landau gauge $A_y = Bx$. Translation symmetry along y gives wavevectors $k_y = 2\pi j / L_y$. In the transverse Hamiltonian the term $(\hbar k_y + eBx)^2/(2m^*)$ describes a harmonic oscillator with centre $x_0 = -\hbar k_y/(eB)$.

To place an orbit centre inside the sample, k_y must range over an interval of length $eBL_x/\hbar$. Dividing by the allowed spacing $2\pi/L_y$ gives $N_\phi = eBL_xL_y/h$ bulk states per spin-resolved orbital level, neglecting edge corrections. Each magnetic flux quantum h/e adds one state per level. For sheet density n_s define filling factor $\nu = n_s h/(eB)$. This counts occupied spin-resolved levels when it is an integer; if a residual spin degeneracy is unresolved, levels arrive in pairs.

In an ideal low-temperature integer quantum Hall state, with the chemical potential in a mobility gap, longitudinal dissipation can vanish and Hall conductance magnitude is $\nu e^2/h$. The classical expression $n_s e/B$ has the same value at exact integer filling, but alone cannot explain persistence of plateaus over a finite interval of field. Localization and edge transport are essential for that observation.

Locate an integer filling in a measurement

A two-dimensional electron system has $n_s = 3.00 \times 10^{15}$ m⁻² and effective mass $0.067 m_e$. The field for total filling $\nu = 2$ is

$$B = n_s h/(2e) = 6.20\,\mathrm{T}.$$

The number of states per area per spin-resolved level is $eB/h = 1.50 \times 10^{15}$ m⁻², half the measured sheet density. The orbital spacing at this field is

$$\hbar eB/m^* = 10.7\,\mathrm{meV}.$$

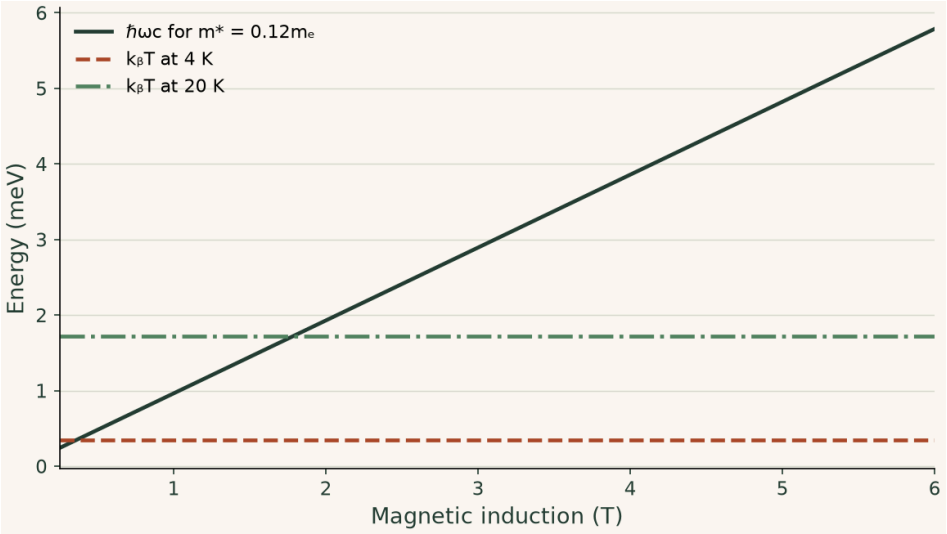
At 4.2 K the thermal energy is 0.362 meV, so thermal smearing is smaller than the orbital gap; disorder and spin structure still matter.

The ideal Hall resistance magnitude on the $\nu = 2$ plateau is $h/(2e^2) = 12.9$ kΩ. A weak current of 0.500 μA would produce a transverse voltage magnitude of 6.45 mV. Increasing the field at fixed density decreases filling and can move the Fermi level into another broadened level. A simultaneous longitudinal voltage measurement is indispensable: a Hall voltage of the right numerical size at a single field can also occur in a classical conductor. A broad flat Hall trace together with a longitudinal-resistance minimum gives stronger evidence for quantized transport.

Changing the sample area at constant sheet density and magnetic field increases the number of occupied particles and the number of orbit centres in the same proportion. Filling factor and the ideal Hall resistance therefore remain unchanged, even though total particle count can be very different. In contrast, a gate that varies sheet density at fixed field moves the layer through successive fillings. The expected density spacing per spin-resolved level is eB/h. A gate-calibrated scan can thus test Landau degeneracy independently of the cyclotron energy spacing inferred from optical or thermal observations.

When can a Landau-level spacing clear thermal broadening?

For a two-dimensional band with isotropic mass $0.12m_e$ and negligible spin splitting, calculate $\hbar eB/m^*$. Compare the result with the thermal energy for two temperatures. This gives a scale for designing a sweep, rather than a sharp visibility criterion.



At 20 K the two energy scales match near 1.79 T. Lowering temperature to 4 K moves the equality to about 0.36 T. Level broadening from scattering and the measurement response can demand a stronger field than this scale alone suggests.

Try it · Landau scale

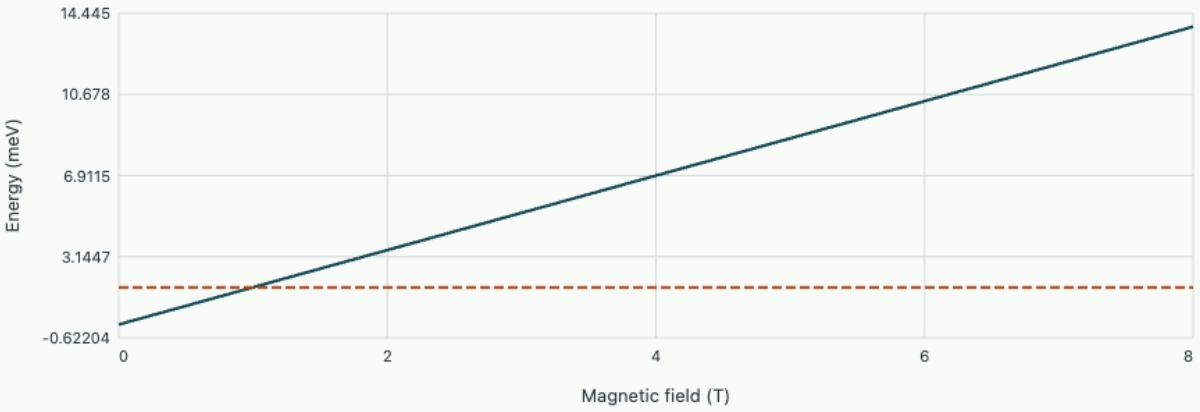
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Temperature (K) **20**

Effective mass / electron mass **0.067**



Cyclotron level spacing | Thermal energy

Thermal and cyclotron scales match at 0.99745 T. Disorder broadening is a separate experimental scale.

▶ **Read values from the plot**

Problem sheet 4 · Transport lifetimes and magnetic orbits

Connect three different observations to scattering, thermal transport and Fermi-surface geometry.

1. A metal wire through a temperature sweep

A simple metal has $n = 5.0 \times 10^{28}$ m⁻³, $m^* = m_e$, $v_F = 1.2 \times 10^6$ m/s and $\rho = 2.0 \times 10^{-8}$ Ω m. (a) Infer its transport lifetime and mean free path. (b) If resistivity falls by a factor ten on cooling, predict the new length at fixed n, m^*. (c) What extra limitation could arise in a 100 nm wire?

▼ **Carry the conductivity estimate into a length**

Use $\tau = m_e/(ne^2\rho)$ with $e = 1.602 \times 10^{-19}$ C to get 3.55×10^{-14} s; multiplying by v_F gives 42.6 nm. At ten times smaller resistivity the nominal path is 426 nm. Carriers may reach boundaries before losing directed momentum through bulk scattering. Surface specularity and wire size then matter, so assigning all the resistance to a bulk lifetime can mislead.

---

2. How much heat could electrons carry?

A sample at 100 K has $\rho = 5.0 \times 10^{-8}$ Ω m and total $\kappa = 72$ W/(m K). (a) Estimate electronic and lattice contributions using $L_0 = 2.44 \times 10^{-8}$ W Ω K⁻². (b) Compare the measured ratio $\kappa\rho/T$ with L_0. (c) Can the difference alone establish a failure of the electronic Wiedemann–Franz law?

▼ **Evaluate the residual contribution conditionally**

The electronic estimate $L_0 T/\rho = 48.8$ W/(m K) leaves 23.2 W/(m K) for other contributions under this model. The total measured ratio is 3.60×10^{-8} W Ω K⁻². A positive lattice heat current alone could account for the excess. Independent evidence for the electronic fraction or scattering mechanism is needed to test the electronic law.

---

3. What does one oscillation frequency measure?

A two-dimensional electron gas with a circular pocket and two spin states shows one reciprocal-field oscillation frequency $F = 50$ T. (a) Find the pocket area and radius from the quantisation rule. (b) Estimate sheet density. (c) What changes if there are two identical valleys?

▼ **Count states in reciprocal space**

$A_k = 2\pi eF/\hbar = 4.77 \times 10^{17}$ m⁻² and $k_F = \sqrt{A_k/\pi} = 3.90 \times 10^8$ m⁻¹. Counting both spins gives $n_{2D} = 2A_k/(2\pi)^2 = 2.42 \times 10^{16}$ m⁻². Two identical valleys double the total density but do not generate a second distinct frequency. A frequency alone therefore cannot always establish degeneracy.

Explore these relations

- For the metal lifetime, vary scattering in the Drude response and connect the inferred relaxation time to mean free path. Open the live calculation.
- For magnetic-orbit data, distinguish the measured orbit area from its energy derivative and cyclotron mass. Open the live calculation.

Chapter 9 · Magnetic moments and susceptibility

Derive thermal magnetisation for a two-state moment and distinguish its signature from the response of a filled electronic sea.

FROM FIRST PRINCIPLES

Obtain the Curie law by counting two magnetic orientations

Take many identical, independent localized moments per volume n, each with component $+\mu$ or $-\mu$ along an applied induction B. The magnetic energies are $-\mu B$ and $+\mu B$. At temperature T their probabilities are proportional to their Boltzmann factors. With $x = \mu B/(k_B T)$, normalize the two weights and subtract their contributions to the mean:

$$Z = e^{x} + e^{-x},$$

$$\langle m_z \rangle = \mu \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} = \mu \tanh x.$$

Multiply by number density to get magnetization. In weak fields or at high temperatures $x \ll 1$, so the leading term in $\tanh x$ gives

$$M \simeq \frac{n\mu^2 B}{k_B T},$$

$$\chi = \frac{\partial M}{\partial H} \simeq \frac{\mu_0 n \mu^2}{k_B T}.$$

The susceptibility expression uses a weakly magnetic sample with $B \simeq \mu_0 H$ and negligible demagnetizing correction. At low temperature and fixed nonzero field both the expansion and the $1/T$ extrapolation fail. The exact two-state average saturates at $n\mu$ because no moment can contribute more than μ. Exchange between neighbours can also invalidate the independent weights.

DERIVATION AND APPLICATIONS

Distinguish the saturation moment from the Curie moment

An ion with angular momentum quantum number J can have more than two projections on a magnetic field. Consider independent identical ions of number density n in thermal equilibrium, with isotropic Landé factor g and negligible splitting at zero field. Label the levels $m = -J, \ldots, J$ so that their moments along the applied induction B are $g\mu_B m$. Adjacent Zeeman levels are then separated by $g\mu_B B$. Introduce $x = g\mu_B B/(k_B T)$. Weighing each level with its Boltzmann factor gives

$$Z = \sum_{m=-J}^{J} e^{mx},$$

$$M = ng\mu_B \frac{\partial \ln Z}{\partial x}.$$

In weak fields each zero-field state is almost equally likely. The zeroth-order sum of m vanishes by symmetry. The next term in the numerator is $x \sum m^2$, while the denominator is $2J + 1$ to this order. Using the sum of squares of the allowed projections gives

$$\langle m \rangle \simeq x \frac{\sum_{m=-J}^{J} m^2}{2J+1} = \frac{J(J+1)}{3} x.$$

For a dilute specimen take $B \simeq \mu_0 H$. The dimensionless volume susceptibility is the derivative of M with respect to magnetic field strength H. In strong fields the highest projection alone survives. These limits identify two different combinations of the microscopic parameters,

$$\chi = \frac{\mu_0 n (g\mu_B)^2 J(J+1)}{3 k_B T},$$

$$M_{\mathrm{sat}} = ng\mu_B J.$$

The Curie-law effective moment is $g\mu_B \sqrt{J(J+1)}$, whereas the aligned projection is $g\mu_B J$. Their ratio approaches one for large J. A susceptibility fit requires $Jx \ll 1$; a saturation measurement requires the opposite tendency and sufficiently weak interactions between moments.

WORKED EXAMPLE

Predict the response of dilute spin-one ions

Change the inputs and investigate this derivation.

Take $J = 1$, $g = 2.00$,

$$n = 1.00 \times 10^{26}\,\mathrm{m}^{-3},$$

$$T = 10.0\,\mathrm{K}.$$

The aligned magnetization is $1.85 \times 10^{3}\,\mathrm{A/m}$. The Curie formula gives $\chi = 2.09 \times 10^{-4}$. At $B = 0.100\,\mathrm{T}$, $Jx = 0.0134$, and the predicted magnetization is only $\chi B/\mu_0 = 16.6\,\mathrm{A/m}$, consistent with weak polarization.

SECOND APPLICATION

Identify a moment from two calibrated measurements

Change the inputs and investigate this derivation.

Suppose independent concentration measurements in another material yield a saturation moment of $3.00\mu_B$ per ion and an effective Curie moment of $3.87\mu_B$. The squared ratio is $1 + 1/J = (3.87/3.00)^2 = 1.66$, giving $J \simeq 1.50$. The saturation value then gives $g \simeq 2.00$. Susceptibility alone fixes only a product of the factor g and the size of the angular momentum.

Try it · Spin magnetisation

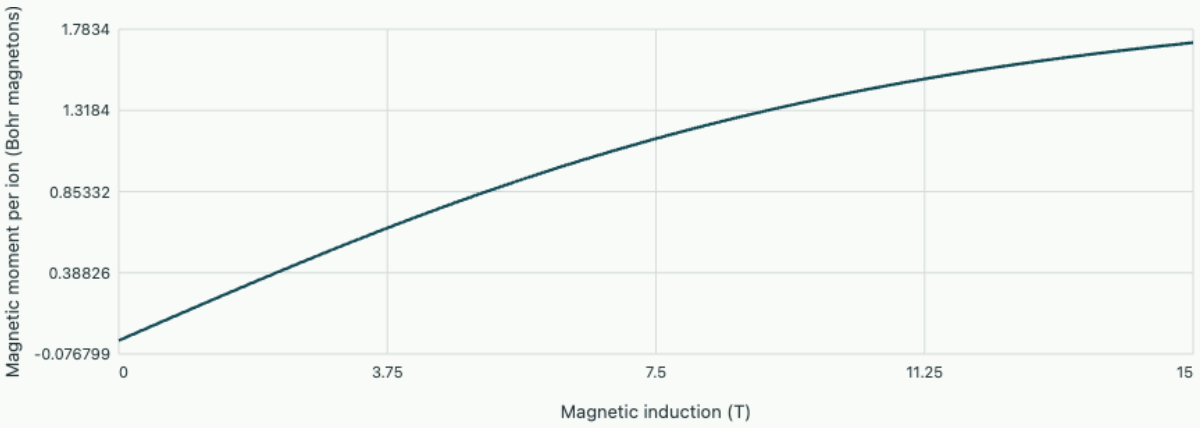
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Spin quantum number J **1**

Temperature (K) **10**



Exact spin partition sum, g=2

Saturation moment 2J=2 μB; Curie effective moment 2√(J(J+1))=2.8284 μB. Independent local moments.

A moment has an energy as well as a direction

Consider N independent identical magnetic moments in volume V, each with allowed projection $\pm\mu$ along a uniform applied induction B. The energies are $\mp\mu B$; the state aligned with the field has lower energy. For $\beta = 1/(k_B T)$ the one-moment partition function is $Z_1 = 2\cosh(\beta\mu B)$. From $\langle m_z \rangle = k_B T\, \partial_B \ln Z_1$ obtain

$$M = n\mu \tanh(\mu B / k_B T),$$

$$n = N/V.$$

The full curve approaches $n\mu$ at high field or low temperature. In the weak-field regime $\mu B \ll k_B T$, expansion of the hyperbolic tangent gives

$$M \simeq n\mu^2 B/(k_B T).$$

Under weak magnetisation and negligible demagnetising corrections use $B \simeq \mu_0 H$; the dimensionless SI susceptibility $\partial M/\partial H$ is then

$$\chi = \mu_0 n\mu^2/(k_B T).$$

Be explicit about differentiating with respect to B or H, which carry different units.

Work a moment count from two measurements

Suppose a dilute two-state sample exhibits saturation magnetisation 9.27×10^3 A/m and the moments are known to have $\mu = \mu_B = 9.27 \times 10^{-24}$ J/T. The implied density is 1.00×10^{27} m⁻³. At 10 K and 1 T the argument $\mu B/k_B T$ is 0.0672, safely in the linear range for a rough estimate: $M \simeq 622$ A/m. At the same temperature a field of about 14.9 T gives argument one and $M/M_{\mathrm{sat}} = 0.762$. Fitting a straight line up to that field would underestimate the zero-field slope.

If the moment magnitude is unknown, saturation determines $n\mu$ and the small-field slope supplies $n\mu^2$. Taken together they can separate density from moment within this model. A mixture of moment species, orbital contributions or interactions can defeat this two-equation inference. Plotting curves at several temperatures against B/T tests the independent-moment scaling without needing to know the overall density beforehand.

Compare with itinerant spin response

For a weakly perturbed degenerate metal, the two spin species shift by $\mp\mu B$. Almost all states below E_F remain doubly occupied, so only a thin shell at the Fermi surface changes its net spin. At low temperature and nearly constant total density of states $D_V(E_F)$, the spin imbalance is approximately $D_V(E_F)\mu B$, giving $M \simeq \mu^2 D_V(E_F) B$. The leading Pauli response is nearly temperature independent, in contrast to the $1/T$ scaling of independent localised moments. This argument addresses spin susceptibility; orbital diamagnetism also contributes to a real measured susceptibility.

Diagnose a small impurity contribution

If bulk data show a roughly constant offset plus a low-temperature upturn, test a simple additive model $\chi = \chi_0 + C/T$ over a weak-field interval. Plot χ against $1/T$ rather than taking an inverse before subtracting the offset. An upturn can arise from a dilute impurity population even when the host electronic sea dominates heat capacity. If scaling fails systematically, examine interactions between moments; the next chapter shows how cooperative ordering changes the small-field susceptibility.

Cool local moments by adiabatic demagnetization

Independent moments have an entropy that depends on the ratio of magnetic field to temperature. For a two-state moment of magnitude μ, energies are $\mp\mu B$. Define $x = \mu B/(k_B T)$. The one-moment partition function is $2\cosh x$, average energy is $-\mu B \tanh x$, and the magnetic entropy follows from $S = k_B \ln Z + U/T$,

$$\frac{S_{\mathrm{spin}}}{N k_B} = \ln(2\cosh x) - x\tanh x.$$

At zero field for fixed positive temperature the moment has two equiprobable orientations and contributes $k_B \ln 2$. At large x nearly all moments favour the lower state and the spin entropy tends to zero. If field is increased at fixed temperature, this reduction of spin entropy expels heat to a bath. Now isolate the material and reduce the field quasistatically. In the ideal spin-dominated limit entropy remains fixed, so x remains fixed and T/B is constant.

Thermal isolation and sufficiently slow variation are both required. The first prevents external heat from entering, and the second allows internal degrees of freedom to follow equilibrium. Other stores of entropy, particularly lattice vibrations or mobile electrons, participate in an actual specimen. The pure spin scaling can then be modified even when the process as a whole is perfectly adiabatic.

Estimate temperature and cooling capacity

Take $N = 1.00 \times 10^{21}$ independent moments each of magnitude μ_B, initially at 2.00 K. Magnetize isothermally from negligible field to 3.00 T, using $\mu_B/k_B = 0.6717$ K/T. The final magnetic parameter is $x = 1.01$, giving spin entropy per particle about $0.362 k_B$ compared with the initial $0.693 k_B$. The heat passed to the bath in that reversible isothermal step is $T N k_B (0.693 - 0.362) = 9.14$ mJ. Work is supplied by the field source during preparation.

Isolate and reduce field from 3.00 T to 0.300 T. Holding x fixed predicts

$$T_f = 2.00(0.300/3.00) = 0.200\,\mathrm{K}.$$

Further field reduction in the ideal noninteracting equations would approach zero temperature. Real moments interact, however, and internal effective fields or ordering lift their near-zero-field degeneracy. Thermal leaks and finite relaxation rates also become increasingly important. A practical test records temperature against field on several ramp rates; dependence on the rate can distinguish irreversible heating from a reproducible correction to the equilibrium entropy model. The final cold sample can absorb only finite heat before its moments become disordered, even if the predicted endpoint temperature is very low.

A second check comes from magnetization along the ideal isolated sweep. Since $M = N\mu\tanh x$ at fixed volume and x is conserved, the spin magnetization should stay almost constant while both the external field and the temperature fall. In contrast, an isothermal field decrease reduces x and magnetization. Measuring temperature alone cannot establish that the sample really followed an isentropic spin path. A simultaneous magnetic reading tests the fixed-population interpretation.

Deviations concentrated at the weakest fields can signal interactions between moments rather than a general failure of the high-field entropy estimate.

Chapter 10 · Exchange, spontaneous magnetisation and critical response

Use a mean-field model and a free-energy expansion to explain how a macroscopic moment can arise in zero applied field.

FROM FIRST PRINCIPLES

How an interaction permits spontaneous magnetization

Put an Ising spin $s_i = \pm 1$ on each lattice site and use a ferromagnetic bond energy $-J s_i s_j$ with $J > 0$. If every site has z neighbours, replacing their actual spins by the common mean $m = \langle s \rangle$ assigns a trial spin energy $-s(zJm + \mu B)$. Thermal averaging this two-level problem must reproduce the same mean used to construct the local field:

$$m = \frac{\sum_{s=\pm 1} s\, e^{s(zJm+\mu B)/(k_B T)}}{\sum_{s=\pm 1} e^{s(zJm+\mu B)/(k_B T)}} = \tanh\!\left(\frac{zJm + \mu B}{k_B T}\right).$$

At zero applied field the right-hand side is odd, so $m = 0$ always solves the equation. To decide when two nearby nonzero solutions appear, compare the slope at the origin with one, then retain the cubic correction. Write $a = zJ/(k_B T)$:

$$m = am - \tfrac{1}{3}a^3 m^3 + \cdots,$$
$$T_c^{\mathrm{MF}} = zJ/k_B, \qquad m^2 \simeq 3(a-1)/a^3.$$

Just below this mean-field critical temperature either sign of m is possible even in zero applied field. The approximation replaces correlated neighbours by an average. In short-range one-dimensional systems that omission can even change whether a finite-temperature transition exists.

DERIVATION AND APPLICATIONS

Calculate a correlation length without a mean-field transition

Mean-field theory replaces the neighbours of each spin by a common average. A chain gives a case in which the energy of a local reversal can instead be counted exactly. Take an open chain of Ising spins $s_i = \pm 1$, zero external field, and ferromagnetic nearest-neighbour interaction $J > 0$. The energy is $-J \sum_i s_i s_{i+1}$. Specify the first spin and all bond variables $b_i = s_i s_{i+1}$. This is a one-to-one description of every configuration. Since the energy is a sum over these independent binary bonds, each bond has weights e^{K} and e^{-K}, where $K = J/(k_B T)$.

$$p_{\mathrm{wall}} = P(b_i = -1) = \frac{1}{1 + e^{2K}},$$
$$\langle b_i \rangle = 1 - 2p_{\mathrm{wall}} = \tanh K.$$

A wall costs $2J$ compared with an aligned bond. To compare two spins separated by r bonds, multiply the intervening b_i. Every interior spin then occurs twice and drops out. Bond independence makes the average of this product particularly simple,

$$\langle s_i s_{i+r} \rangle = \left\langle \prod_{j=i}^{i+r-1} b_j \right\rangle = (\tanh K)^r .$$

Writing the result as an exponential in the physical separation ra, where a is the site spacing, gives

$$\xi = -\frac{a}{\ln(\tanh K)},$$

$$\xi \simeq \frac{a}{2} e^{2K} \quad (K \gg 1).$$

At low temperature the last expression follows by expanding $\tanh K \simeq 1 - 2e^{-2K}$. At every positive temperature the probability of a wall remains positive. Thus widely separated spins lose correlation in this short-range one-dimensional model. A finite sample shorter than ξ can still look ordered during a measurement.

WORKED EXAMPLE

Estimate decorrelation over a microscope field

Change the inputs and investigate this derivation.

Let $J/k_B = 20.0\,\mathrm{K}$, $T = 10.0\,\mathrm{K}$ and $a = 0.400\,\mathrm{nm}$. The wall probability is $1/(1+e^{4}) = 0.0180$, giving about 18 walls per 1000 bonds on average. The exact correlation length is $10.9\,\mathrm{nm}$. Across $20.0\,\mathrm{nm}$, or 50 bonds, the measured spin correlation should be approximately $(\tanh 2)^{50} = 0.160$.

SECOND APPLICATION

Infer exchange from a measured spatial decay

Change the inputs and investigate this derivation.

In a different chain the correlation drops to one half over eight bonds at $T = 30.0\,\mathrm{K}$. Each step has correlation $2^{-1/8} = 0.9170$. Inverting $K = \tfrac{1}{2}\ln[(1+\langle b\rangle)/(1-\langle b\rangle)]$ gives

$$J/k_B = 47.1\,\mathrm{K}.$$

A local image counting opposite neighbours should then find a fraction $(1 - 0.9170)/2 = 0.0415$, an independent test using adjacent spins.

Try it · Ph ising ring

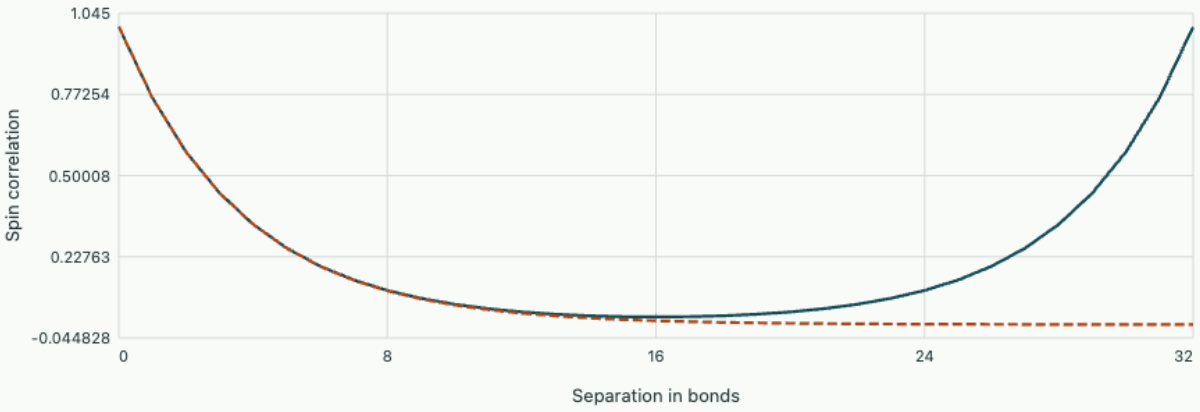
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Temperature $k_\beta T/J$ **1**

Sites around the ring **32**



Closed finite ring Infinite chain

Infinite-chain correlation length 3.672 lattice spacings. The opposite site on the 32-site ring has correlation 0.02562.

Replace neighbouring moments by their average

An independent moment has zero mean at zero field for any positive temperature. In a ferromagnet aligned neighbours can lower the exchange energy. Consider Ising variables $s_i = \pm 1$ with energy $-J\sum_{\langle ij\rangle} s_i s_j - \mu B \sum_i s_i$, $J > 0$. Each distinct nearest-neighbour bond is counted once. In a uniform mean-field approximation replace a neighbour's spin by $m = \langle s \rangle$. For coordination number z, an individual spin then feels an effective energy bias $zJm + \mu B$. Applying the two-state result to that spin requires self-consistency:

$$m = \tanh[(zJm + \mu B)/(k_B T)].$$

At $B = 0$, $m = 0$ always solves the equation. The slope of the right-hand side at the origin is $zJ/(k_B T)$. Above the mean-field transition temperature $T_c = zJ/k_B$ only the disordered solution exists. Below it

two symmetry-related nonzero stable solutions appear. The sign selected in a macroscopic region can depend on a small field and on its history.

Approach the transition from either side

For small field above T_c, expand $\tanh x \simeq x$. This gives $m \simeq \mu B/[k_B(T - T_c)]$, a Curie–Weiss susceptibility. A positive intercept of an inverse-susceptibility plot can indicate dominant ferromagnetic exchange over a suitable high-temperature interval. It is not an exact measurement of the true ordering temperature, particularly when fluctuations, geometry or competing interactions matter.

Just below T_c, at zero field, expand through the cubic term. Writing $t = T_c/T$, the nonzero solution has $m^2 \simeq 3(t - 1)/t^3$, approaching $3(T_c - T)/T_c$ sufficiently near the transition. The square-root onset is a mean-field prediction. A finite-temperature transition of this type does not exist for a one-dimensional short-range Ising chain in the thermodynamic limit, despite the same simple self-consistency calculation predicting one. Neglecting the cost and entropy of fluctuations can qualitatively change a prediction.

Interpret the curvature of a free energy

A local free energy even under $m \mapsto -m$ at zero field may be expanded as $f = \frac{1}{2}a(T - T_c)m^2 + \frac{1}{4}bm^4 - hm$, $a, b > 0$. At $h = 0$, the minima sit at 0 above the transition and at $m = \pm\sqrt{a(T_c - T)/b}$ below it. A positive second derivative indicates local stability, and the differential response $dm/dh = 1/f''$ grows as the curvature weakens. For $a(T - T_c) = -1, b = 4$ in scaled units, the minima are at ± 0.5, have curvature 2 and are lower than the intervening maximum by 0.0625 in free-energy density units.

From a local moment to a macroscopic measurement

Different regions can select opposite stable magnetisations. Domain walls cost exchange and anisotropy energy, while domain formation can reduce the magnetic field energy outside a specimen. A small measured zero-field net moment can therefore coexist with substantial local order. A static uniform free-energy graph alone cannot predict coercivity: wall pinning, nucleation, defects and measurement times affect a hysteresis loop. Temperature-dependent susceptibility, local probes and field-cycling measurements answer complementary questions about the same ordered sample.

Derive the profile and energy of a domain wall

Two domains magnetized in opposite easy-axis directions can coexist, but rotating the magnetization in between costs exchange and anisotropy energy. Consider a uniaxial ferromagnet with magnetization magnitude fixed and angle $\theta(x)$ relative to the easy axis varying along one spatial coordinate. Assume a planar 180-degree wall in a geometry where additional magnetostatic costs can be neglected for this profile. With exchange stiffness A in J/m and anisotropy density K in J/m³, the wall energy per unit area is

$$\sigma_w = \int_{-\infty}^{\infty} [A(\partial_x \theta)^2 + K \sin^2 \theta]\, dx.$$

A narrow wall saves anisotropy energy but makes the spatial gradient steep. A broad wall saves exchange energy but places more material away from the preferred axis. Variation gives

$$2A\theta'' = 2K \sin\theta \cos\theta.$$

Multiply by θ' and integrate once. Since both gradient and anisotropy term vanish deep inside either domain, the integration constant is zero, yielding $A(\theta')^2 = K \sin^2 \theta$. Exchange and anisotropy costs are locally equal on the minimum-energy profile.

Take increasing θ between zero and π and define $\Delta = \sqrt{A/K}$. Separation gives

$$d\theta / \sin\theta = dx / \Delta.$$

With the wall centred at zero,

$$\tan(\theta/2) = \exp(x/\Delta).$$

Inserting the first integral back into the wall energy and changing variable from x to θ gives

$$\sigma_w = 2K\Delta \int_0^{\pi} \sin\theta\, d\theta = 4\sqrt{AK}.$$

The profile has long tails, so a quoted wall width must state its convention; $\pi\Delta$ is a frequently used slope-based width rather than a sharp boundary.

Compare two anisotropies at fixed exchange stiffness

Use $A = 1.00 \times 10^{-11}$ J/m and $K = 1.00 \times 10^{5}$ J/m³. The scale Δ is 10.0 nm and conventional width $\pi\Delta$ is 31.4 nm. Wall energy is 4.00 mJ/m². A planar patch of area 1.00 µm² costs 4.00×10^{-15} J, even though the two far domains have the same zero-field bulk energy.

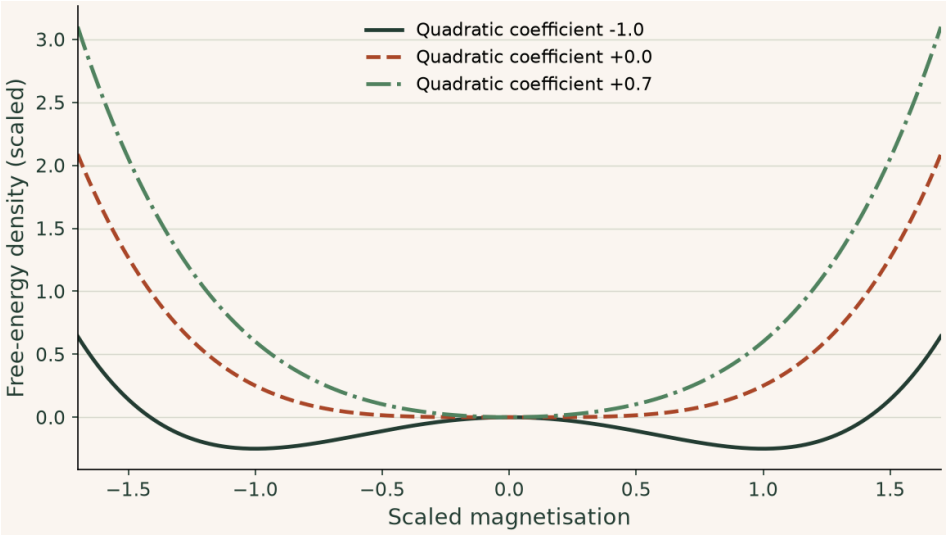
If a second specimen at the same temperature has four times the anisotropy and the same exchange stiffness, the predicted width halves while the cost per area doubles. The combination of width and energy could therefore independently estimate A and K, whereas either measurement by itself fixes only a ratio or a product. In the limit $K \to 0$ the preferred axis loses meaning and the wall spreads indefinitely in this model. Finite sample size, magnetostatic fields, boundaries and defects then set spatial scales omitted from the one-dimensional calculation.

One can also express the measured magnetization components explicitly. With the local moment in a fixed rotation plane, the easy-axis component is proportional to $\cos\theta = -\tanh(x/\Delta)$ and the transverse component to $\sin\theta = 1/\cosh(x/\Delta)$. The easy-axis profile passes smoothly through zero even though the moment magnitude remains fixed throughout the wall. A probe sensitive only to that component might incorrectly interpret the central zero as suppression of ferromagnetism. Measuring a

transverse component or a quantity sensitive to the full moment tests the distinction between a rotating order parameter and a vanishing one.

Distinguish a soft magnetic response from a loss of stability

In zero applied field the lowest stabilising even expansion is $f = am^2/2 + m^4/4$ in these units. Vary the quadratic coefficient through zero. Equilibrium requires both stationarity and positive curvature at a minimum.



For a $= -1$ the stable minima lie at ± 1 with energy density -0.25. At a $= 0$ the quartic term stabilises a flat minimum. This bulk potential cannot by itself predict the field at which an inhomogeneous sample will reverse its magnetisation.

Try it · Landau potential

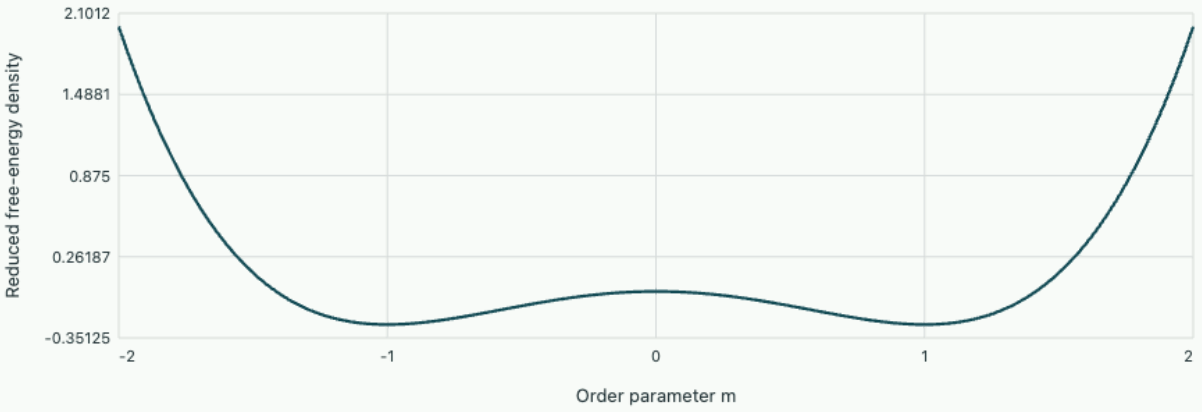
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Quadratic coefficient a **-1**

Applied dimensionless field **0**



Landau potential

The quartic coefficient is one. At zero field, minima lie at ±1. A bulk potential does not predict a real specimen's coercive field.

▶ **Read values from the plot**

Problem sheet 5 · Moments, susceptibility and collective order

Derive a moment scale from data and test the stability of a proposed magnetic state.

1. Independent moments at two temperatures

A two-state sample has $M_{\mathrm{sat}} = 2.0 \times 10^{4}$ A/m and $\mu = \mu_B$. (a) Find the density of moments. (b) At 2 K find the field at which $M = M_{\mathrm{sat}}/2$, using $\mu_B/k_B = 0.672$ K/T. (c) Repeat at 8 K and identify what could break this scaling.

▼ **Invert the hyperbolic tangent**

The density is $M_{\mathrm{sat}}/\mu_B = 2.16 \times 10^{27}$ m⁻³. The required argument is $\operatorname{artanh}(1/2) = 0.5493$, so the fields are $0.5493(2)/0.672 = 1.63$ T and 6.54 T. Interacting moments, additional low-energy levels or a background linear susceptibility can spoil scaling with B/T.

---

2. A Curie tail on a constant background

A sample has $\chi(10\,\mathrm{K}) = 5.0 \times 10^{-4}$ and $\chi(30\,\mathrm{K}) = 3.0 \times 10^{-4}$. (a) Fit $\chi = \chi_0 + C/T$. (b) Predict its value at 20 K. (c) Would two points alone distinguish this fit from cooperative Curie–Weiss behaviour?

▼ **Plot against inverse temperature**

The slope is $C = (2.0 \times 10^{-4})/(1/10 - 1/30) = 3.0 \times 10^{-3}$ K, and $\chi_0 = 2.0 \times 10^{-4}$. The prediction at 20 K is 3.5×10^{-4}. Two measured points determine two constants but cannot validate the functional form. Measure intermediate and lower temperatures and test field dependence in the weak-field regime.

---

3. Curvature and a spontaneous moment

For $f(m) = \tfrac{1}{2}tm^2 + \tfrac{1}{4}m^4 - hm$ in scaled units, (a) find stable equilibria at $h = 0$ for $t = 1$ and $t = -1$. (b) Compute the small-field slope on each stable branch. (c) Find the zero-field barrier height below the transition and state why this is insufficient to predict a real sample's coercive field.

▼ **Inspect the stationary equation and its derivative**

The equation is

$$tm + m^3 = h.$$

At $t = 1$ the stable solution is zero with slope $dm/dh = 1/(t + 3m^2) = 1$. At $t = -1$, $m = \pm 1$ are stable and have slope $1/2$; zero has negative curvature. The difference between $f(0) = 0$ and either

minimum $f(\pm 1) = -1/4$ is 0.25 per unit volume in the chosen scales. Domain-wall motion, nucleation and defects can provide reversal routes absent from the homogeneous model.

Explore these relations

- For the independent moments, select spin one half and compare the saturation curve at different temperatures. Open the live calculation.
- For spontaneous order, vary free-energy curvature and field and inspect the changing minima. Open the live calculation.

Chapter 11 · Superconductivity, magnetic screening and flux quantisation

Separate persistent current from magnetic flux expulsion and derive two measurable length and flux scales of a superconducting state.

FROM FIRST PRINCIPLES

Connect a material current law to magnetic screening

A vanishing electrical resistance by itself allows an already existing current to persist. To predict the equilibrium field profile in a thick superconductor one needs additional physical input. In a uniform isotropic superconductor with no vortices the second London relation connects the curl of the supercurrent to the induction. Here n_s counts carriers of mass m and signed charge q. Combine this constitutive law with static Ampère's law:

$$\nabla \times \mathbf{j}_s = -\frac{n_s q^2}{m}\mathbf{B},$$
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}_s.$$

Take a second curl of the field law and use $\nabla \cdot \mathbf{B} = 0$. The sign from the vector identity cancels that in the London relation:

$$\nabla^2 \mathbf{B} = \frac{\mathbf{B}}{\lambda_L^2},$$
$$\lambda_L = \sqrt{\frac{m}{\mu_0 n_s q^2}}.$$

Fill $x > 0$ with the material and impose a parallel surface field $B(0) = B_0$. With no field growth deep in the bulk, only one exponential survives:

$$B(x) = B_0 e^{-x/\lambda_L}.$$

Thus the single measured length constrains a ratio $m/(n_s q^2)$. This construction assumes linear local response and a field low enough for the vortex-free state and its prescribed surface condition to apply.

Background · London length

$$B(x) = B(0) e^{-x/\lambda_L}$$

This is a field profile in a thick homogeneous superconductor in a linear screening regime. A thin slab requires both boundary conditions.

DERIVATION AND APPLICATIONS

Determine how deeply a parallel field penetrates a finite slab

The exponential London solution assumes a sample extending far beyond a single exposed surface. A film has two surfaces whose screening currents can overlap. Consider an infinite slab of total thickness $2d$, occupying $-d < x < d$. Apply a static induction B_a parallel to its faces, along z, and assume a uniform London penetration length λ. Work in the vortex-free Meissner state with fields small enough for linear response. Translational invariance parallel to the slab gives the boundary problem

$$\frac{d^2 B_z}{dx^2} = \frac{B_z}{\lambda^2},$$
$$B_z(-d) = B_z(d) = B_a.$$

A general solution has a hyperbolic cosine and a hyperbolic sine. Reflection about the midplane makes the field even, which removes the sine term. The surface field fixes the remaining coefficient. Ampère's law then determines the screening current, whose sign reverses at the midplane,

$$B_z(x) = B_a \frac{\cosh(x/\lambda)}{\cosh(d/\lambda)},$$
$$J_y(x) = -\frac{B_a}{\mu_0 \lambda} \frac{\sinh(x/\lambda)}{\cosh(d/\lambda)}.$$

A probe at the centre and a probe integrating flux over the cross-section read different quantities. The latter reports the thickness average obtained by integrating the field profile,

$$\frac{\overline{B}}{B_a} = \frac{1}{2dB_a} \int_{-d}^{d} B_z(x)\, dx = \frac{\lambda}{d} \tanh(d/\lambda).$$

For a thick slab the mean fraction tends to λ/d, consistent with two active surface layers. For a thin slab it is $1 - (d/\lambda)^2/3$ to leading order. Thin samples can therefore exhibit weak parallel-field exclusion even when the material is fully superconducting throughout their volume.

WORKED EXAMPLE

Compare central and average field in a thick slab

Change the inputs and investigate this derivation.

A sample of thickness $600\,\mathrm{nm}$ with $\lambda = 100\,\mathrm{nm}$ has $d/\lambda = 3$. In $B_a = 2.00\,\mathrm{mT}$ its central field is

$$2.00/\cosh 3 = 0.199\,\mathrm{mT}.$$

The mean induction is

$$2.00\ \tanh(3)/3 = 0.663\,\mathrm{mT}.$$

The higher mean follows directly from the substantial contribution of material close to both exposed faces.

SECOND APPLICATION

Extract a penetration length using a local probe

Change the inputs and investigate this derivation.

A probe halfway through a $120\,\mathrm{nm}$ film measures $0.800B_a$. The central-field solution gives $d/\lambda = \operatorname{arcosh}(1.25) = 0.693$, so $\lambda = 86.6\,\mathrm{nm}$. The same film should transmit a mean fraction $\tanh(0.693)/0.693 = 0.866$. Comparing this value with an independent flux integral tests the assumptions of uniform thickness and a single penetration length.

Try it · Solid optics methods london slab

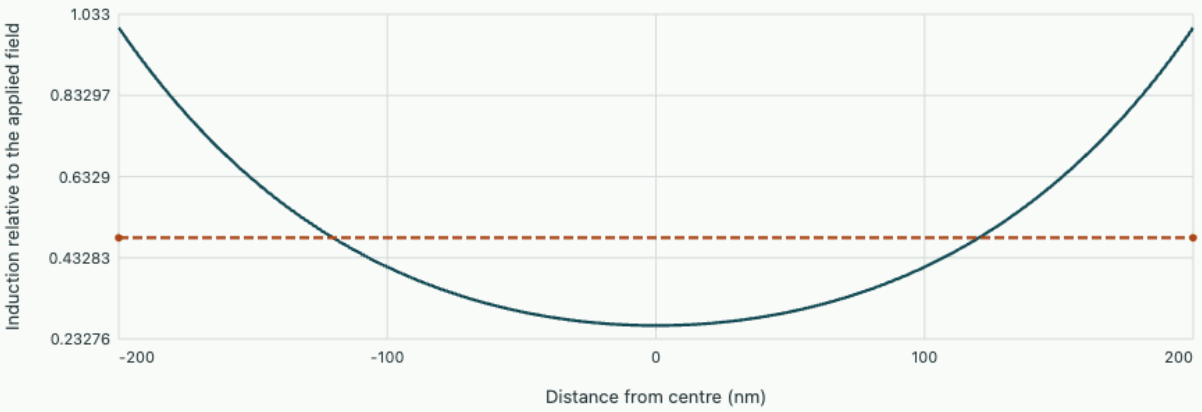
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Film thickness (nm) 400

London length (nm) 100



Local induction Cross-section mean

At the centre B/Ba = 0.2658, whereas the cross-section mean is 0.4820. These apply to a parallel field in a vortex-free slab.

Compare two magnetic histories

A normal conductor dissipates energy from a sustained current. A superconductor can support persistent current and, in an appropriate low-field equilibrium state, expels magnetic flux from its bulk. Imagine cooling a specimen through its transition in a field and subsequently measuring its interior. Perfect conductivity alone would imply that a previously existing flux could remain frozen. Equilibrium flux expulsion supplies additional evidence about the state; sample shape, vortices and pinning affect what a real experiment detects. A resistance measurement and a magnetisation measurement therefore probe different properties.

Derive the London screening length

Let mobile condensate carriers have density n_s, mass m_s and charge q_s, using a consistent convention for particles or pairs. In a bulk region without vortices the London relation is

$$\nabla \times \mathbf{j}_s = -(n_s q_s^2/m_s)\mathbf{B}.$$

Combine this with magnetostatic Ampère's law $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}_s$ and $\nabla \cdot \mathbf{B} = 0$. Taking a curl gives

$$\nabla^2 \mathbf{B} = \mathbf{B}/\lambda_L^2,$$

$$\lambda_L = \sqrt{m_s/(\mu_0 n_s q_s^2)}.$$

For a half-space $x > 0$ with surface field parallel to the surface and deep interior field zero,

$$B(x) = B(0)e^{-x/\lambda_L}.$$

At depth $3\lambda_L$, 4.98% of the surface amplitude remains. Taking $n_s = 2.0 \times 10^{28}$ m⁻³ electrons and electron mass in the simple carrier convention gives $\lambda_L \approx 37.6$ nm. Real effective masses, temperature-dependent superfluid density and nonlocality can change this value. Mixing pair density with single-electron charge would introduce a systematic factor error.

Quantise a circulating phase

A coherent condensate has a phase ϕ whose change around a closed path must be $2\pi n$ for integer n if the complex order parameter is single-valued and nonzero along the path. The gauge-invariant momentum is $m_s \mathbf{v}_s = \hbar \nabla \phi - q_s \mathbf{A}$. Integrating around a loop relates circulation and flux. If a path lies sufficiently deep in a thick ring that its current density is negligible, the velocity term vanishes and enclosed flux comes in units $\Phi_0 = h/|q_s|$. For paired carriers $|q_s| = 2e$ the unit is 2.068×10^{-15} Wb. In thin rings where the kinetic term matters, it is the full fluxoid combination that is quantised.

Use length scales to interpret the experiment

The penetration depth characterises field screening. A separate coherence length ξ measures the distance over which the condensate amplitude recovers after a perturbation. In bulk Ginzburg–Landau theory near the transition, the ratio λ/ξ distinguishes type I and type II at $1/\sqrt{2}$. In a type-II material between the lower and upper critical fields, flux can penetrate in quantised vortices with normal-like cores surrounded by screening currents. A specimen in this mixed state can retain superconducting regions while admitting magnetic flux. Stating field, temperature and cooling history is essential when interpreting the observed signal.

Locate the upper critical field from the linearized order parameter

The London equation assumes an established superconducting density. Near the loss of superconductivity at a high applied field, the order parameter itself becomes small and its spatial variation matters. For a bulk type-II material near its transition use Ginzburg–Landau free-energy density terms $\alpha|\psi|^2 + \beta|\psi|^4/2 + |(-i\hbar\nabla - q_s\mathbf{A})\psi|^2/(2m_s)$ with $\alpha < 0$, $\beta > 0$ and pair charge magnitude $|q_s| = 2e$. Very close to the upper critical field the density is small, field screening is weak and a uniform induction B is a suitable leading approximation.

Minimizing the free energy with respect to ψ^* and discarding the cubic term gives a charged-particle magnetic eigenproblem. The lowest transverse kinetic eigenvalue is the lowest Landau energy $\hbar|q_s|B/(2m_s)$, achieved with no longitudinal variation. A nontrivial small solution is possible at the onset when this cancels α. Defining $\xi^2 = \hbar^2/(2m_s|\alpha|)$ and $\Phi_0 = h/|q_s|$ gives

$$B_{c2} = \frac{\Phi_0}{2\pi\xi^2}.$$

Thus a coherence length inferred from upper critical induction measures the scale of condensate recovery. Penetration depth instead controls field screening in an already developed condensate. The GL derivation is controlled near the critical temperature and for orbital suppression. Strong paramagnetic pair breaking or pronounced anisotropy can complicate the interpretation of a measured high-field transition.

Connect critical induction to vortex spacing

At a specified temperature close to the transition a bulk specimen loses its superconducting response at $B_{c2} = 1.20$ T. The inferred coherence length is

$$\sqrt{2.068 \times 10^{-15}/(2\pi \times 1.20)} = 16.6\,\mathrm{nm}.$$

An independent penetration-depth measurement at the same temperature yields $\lambda = 120$ nm, giving $\lambda/\xi = 7.24$, above the type-II boundary. Using lengths measured at unrelated temperatures would weaken that check.

At mean induction 0.200 T within the mixed state each singly quantized vortex carries one flux quantum. Its share of sample area is

$$\Phi_0/B = 1.03 \times 10^{-14}\,\mathrm{m}^2.$$

For a triangular lattice that area also equals $\sqrt{3}\,a_v^2/2$, giving vortex spacing $a_v = 109$ nm. Increasing induction fourfold halves the spacing if the array remains in the mixed state and retains its geometry. Near B_{c2} cores overlap and an isolated-vortex picture becomes unreliable even though the linearized onset scale remains useful. Comparing a measured spacing with Φ_0/B tests flux counting independently of the inferred coherence length and of the magnetic field profile around any single vortex.

A useful unit check compares the superconducting flux quantum with the electron flux scale from ordinary Landau-level counting. A condensate pair has twice the charge magnitude of one electron, giving half the flux quantum for a phase winding. At 0.200 T a square micrometre should contain about $BA/\Phi_0 = 96.7$ singly quantized vortices away from edges. Counting vortices in an image therefore gives a bulk average induction estimate independent of a local peak-field calibration. A surface barrier, pinning or an insufficiently large field of view can account for a local deviation from the average.

Imaging successively larger regions tests how well such local counts represent the bulk.

Study the full derivation

A field profile that distinguishes a superconductor from a perfect conductor · Structure, excitations and probes of solids

Zero dc resistance tells how a current can persist. Field expulsion tells how a specimen responds when it enters a superconducting equilibrium state. Derive the London length, calculate what a finite slab and a thin film actually screen, and connect magnetic and electrical measurements to the same superfluid density.

Chapter 12 · Pairing, excitation gaps and what a spectrum measures

Use a two-component excitation model to predict spectroscopic signatures, then distinguish three physically different gaps encountered across the course.

FROM FIRST PRINCIPLES

See how pairing opens a spectroscopic gap

Begin with a single normal-state band and let ξ be its energy measured from the chemical potential. A minimal mean-field model for a uniform paired state mixes an electron amplitude with a hole amplitude through a complex pairing energy Δ. The two diagonal energies have opposite signs. To find its stationary excitations write the eigenvalue condition:

$$H = \begin{pmatrix} \xi & \Delta \\ \Delta^* & -\xi \end{pmatrix}, \qquad \det(H - EI) = (\xi - E)(-\xi - E) - |\Delta|^2 = 0.$$

It follows that the two branches are $E_\pm = \pm\sqrt{\xi^2 + |\Delta|^2}$. At the former Fermi surface $\xi = 0$, even the cheapest positive excitation costs $|\Delta|$. To predict a tunnelling spectrum for constant isotropic gap and nearly constant normal density of states, count the normal states mapping into an energy interval dE. The Jacobian from $\xi = \pm\sqrt{E^2 - |\Delta|^2}$ gives the normalized density, with the electron and hole weights combined:

$$\frac{D_s(E)}{D_n(0)} = \frac{|E|}{\sqrt{E^2 - |\Delta|^2}} \quad (|E| > |\Delta|).$$

There are no states within the ideal gap at this level of description. Finite temperature, lifetime broadening and instrumental resolution change the measured conductance near its edges. The model assumes a nonzero Δ; it predicts consequences of pairing without deriving the attractive interaction that supports it.

DERIVATION AND APPLICATIONS

Predict the weights of a particle peak and its partner

A pairing gap determines excitation energies, but a momentum-resolved spectrum also measures how strongly an electron couples to each peak. Consider a clean, uniform superconductor with a real positive gap Δ, at temperature low enough to neglect thermal broadening. Measure the normal-state energy ξ at one wavevector from the chemical potential. In the electron and hole basis the positive-energy eigenvector has components u, v. The first row of its eigenvalue equation requires

$$(\xi - E)u + \Delta v = 0,$$

$$\frac{v}{u} = \frac{E - \xi}{\Delta},$$

$$E = \sqrt{\xi^2 + \Delta^2}.$$

Normalize the vector to unit probability. Since $(E - \xi)(E + \xi) = \Delta^2$, the squared amplitude ratio can be written without subtracting two nearly equal energies. The resulting coherence factors are

$$|u|^2 = \frac{1}{1 + (E - \xi)/(E + \xi)} = \frac{1}{2}\left(1 + \frac{\xi}{E}\right),$$

$$|v|^2 = \frac{1}{2}\left(1 - \frac{\xi}{E}\right).$$

The negative-energy eigenvector is orthogonal to the positive one and consequently has electron weight $|v|^2$. In the ideal single-particle spectral function at this momentum the two sharp peaks have unit total weight,

$$A(\epsilon) = |u|^2 \delta(\epsilon - E) + |v|^2 \delta(\epsilon + E),$$

$$\int A(\epsilon)\, d\epsilon = 1.$$

Thus the weights exchange as the wavevector passes through the normal Fermi surface. At $\xi = 0$ the two weights are equal. Well above the surface the positive-energy peak carries almost all the weight. To compare both sides quantitatively an experiment must account for occupation and matrix elements; ordinary photoemission alone samples the occupied side most directly.

WORKED EXAMPLE

Resolve a weak partner peak

Change the inputs and investigate this derivation.

At a selected momentum let $\xi = 3.00\,\mathrm{meV}$ and $\Delta = 4.00\,\mathrm{meV}$. The spectral energies are $+5.00$ and $-5.00\,\mathrm{meV}$, with respective electron weights 0.800 and 0.200. At a momentum with the opposite ξ and the same gap the two peak energies stay fixed while their weights reverse. Such a weight exchange tests particle–hole mixing more specifically than the peak separation alone.

SECOND APPLICATION

Reconstruct a gap using corrected peak areas

Change the inputs and investigate this derivation.

In a different sample the peaks are separated by $6.00\,\mathrm{meV}$, and corrected electron spectral areas have positive-to-negative ratio $3:2$. Their normalized weights are 0.600 and 0.400, so

$$\xi/E = 0.200.$$

With $E = 3.00\,\mathrm{meV}$ this gives $\xi = 0.600\,\mathrm{meV}$ and $\Delta = \sqrt{3.00^2 - 0.600^2} = 2.94\,\mathrm{meV}$. Half the separation overestimates the gap here because the probed momentum lies away from the normal Fermi surface.

Try it · Bcs weights

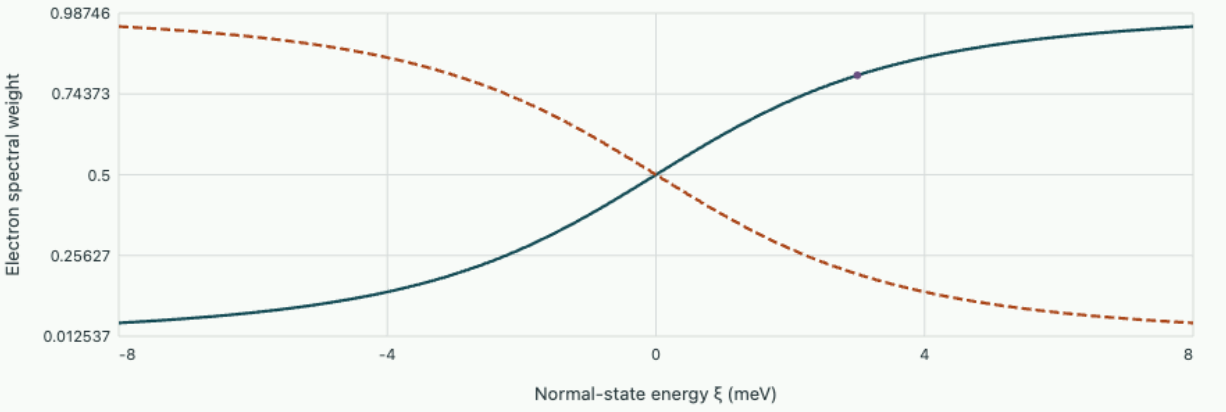
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Gap (meV) 4

Normal-state energy ξ (meV) 3



Peak energies ±5 meV; weights sum to one. The measured electron spectrum also needs occupation and instrumental factors.

Build excitations on a paired state

For a conventional fully gapped superconductor, a mean-field description mixes particle and hole amplitudes near the Fermi surface. With $\xi_k = \epsilon_k - \mu$ and gap parameter Δ, a simplified matrix for a time-reversed pair of states is

$$\begin{pmatrix} \xi_k & \Delta \\ \Delta^* & -\xi_k \end{pmatrix} \begin{pmatrix} u_k \\ v_k \end{pmatrix} = E \begin{pmatrix} u_k \\ v_k \end{pmatrix}, \qquad E_\pm = \pm\sqrt{\xi_k^2 + |\Delta|^2}.$$

A positive-energy quasiparticle has minimum energy $|\Delta|$ where the normal band crossed the chemical potential. Breaking a pair into two excitations can cost at least $2|\Delta|$. These are different thresholds, so the probe and transition being discussed must be specified. The matrix displays level repulsion, reminiscent of the nearly free electron calculation, but its basis states and the physical origin of the coupling are different.

Transform a density of states

Suppose the normal-state density is approximately constant over the gap scale and the gap is isotropic. For $E > |\Delta|$,

$$|\xi| = \sqrt{E^2 - |\Delta|^2},$$

$$d|\xi|/dE = E/\sqrt{E^2 - |\Delta|^2}.$$

Counting the paired branches relative to the normal-state convention yields $D_s(E)/D_n = |E|/\sqrt{E^2 - |\Delta|^2}$ outside the gap and zero inside in the ideal zero-broadening model. The edge singularity is integrable. States removed from the gap reappear near its edges. Thermal broadening, disorder, nodes or an anisotropic gap can change the measured shape.

A low-temperature tunnel junction with a normal counterelectrode of nearly constant density can use differential conductance as a probe of this spectrum, subject to matrix-element and broadening effects. Peaks at biases ± 1.5 mV suggest $|\Delta| \simeq 1.5$ meV in this geometry. Peak separation on the voltage axis is 3.0 mV. Using the weak-coupling BCS ratio $2|\Delta(0)|/(k_B T_c) \simeq 3.53$ gives $T_c \simeq 9.86$ K. Measure T_c independently; agreement with this estimate supports the simple model, while deviation calls for closer investigation.

Select a second observation that distinguishes the mechanism

A nearly free electron band gap separates eigenvalues near a zone boundary. For conduction it also matters where the chemical potential sits. A superconducting gap opens at the Fermi surface on entering an ordered state, and has magnetic screening and a characteristic temperature dependence. A photonic stop band instead excludes real propagating wavevectors in a frequency range for the specified periodic optical medium. All three can produce missing spectral weight or reduced transmission. Their energies, frequency units, occupation rules and control parameters differ.

Return to the numerical optical stack with a prediction: making both layers identical closes its first stop band even though the artificial zone boundary remains. For a proposed superconducting gap, repeat the tunnelling and magnetisation measurements as temperature crosses the transition. For an electronic band edge, relate spectroscopy to filling, dispersion and transport. This combination of an analytic limit, a

controlled change and an independent observable is what makes an interpretation of a spectrum persuasive.

For continued study, see the MIT lectures on Theory of Solids, including quasiparticles and superconducting electrodynamics.

Read the phase difference across a superconducting weak link

A single-particle tunnelling spectrum can reveal an excitation gap. A weak junction between two superconductors can also carry a coherent pair current that responds directly to their relative condensate phase. Use the gauge-invariant phase difference φ, including the vector-potential line integral across the junction. For a simple tunnel junction in the low-transparency regime, its leading coupling energy is

$$U_J = -E_J \cos\varphi.$$

Its 2π periodicity reflects invariance under an integer phase winding and the energy minimum at aligned phase for a conventional junction.

A potential difference V between the condensates causes their relative phase to advance at $\dot{\varphi} = 2eV/\hbar$, choosing a consistent orientation of current and voltage. Combining electrical power IV with the rate of change of junction energy gives $IV = E_J \sin\varphi\,\dot{\varphi}$, hence

$$I = I_c \sin\varphi,$$

$$I_c = \frac{2eE_J}{\hbar}.$$

Under a steady subcritical current the phase can remain constant while $V = 0$. At constant nonzero voltage the phase winds and the ideal coherent current oscillates with frequency $f_J = 2e|V|/h$. These two statements distinguish a DC pair current with no terminal voltage from an AC pair response under bias. A real junction includes quasiparticle conduction and capacitance, so its observed total current can contain additional components.

Convert a microwave frequency into a voltage

A junction biased at 5.00 μV has Josephson oscillation frequency $f_J = (2e/h)V = 2.42$ GHz. Illuminate the junction with a stable microwave reference of frequency 10.0 GHz. If its mean phase advance locks to one full winding per reference cycle, then $\langle\dot{\varphi}\rangle = 2\pi f_{\mathrm{ref}}$, requiring mean voltage $hf_{\mathrm{ref}}/(2e) = 20.7$ μV. Locking to two windings per cycle gives 41.4 μV. Such voltage steps probe phase dynamics and the charge associated with coherent transfer.

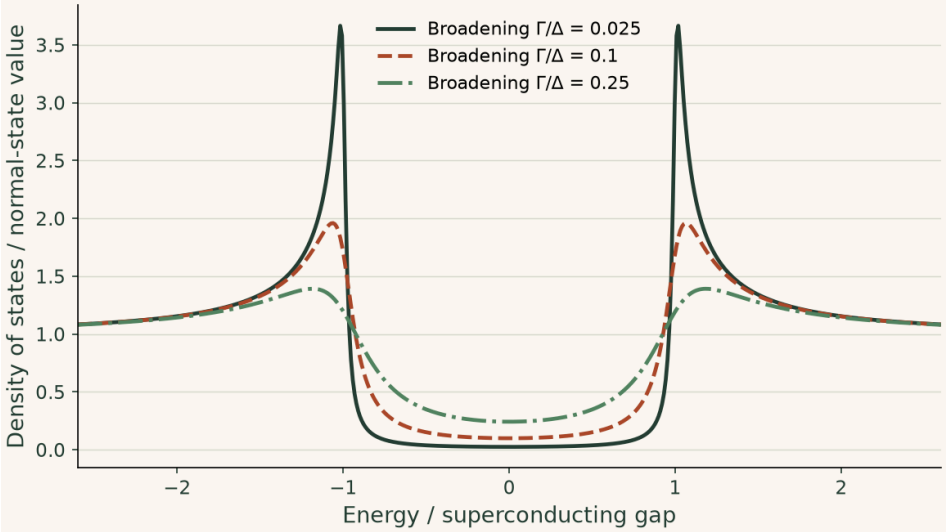
Suppose the same device has measured critical current $I_c = 30.0$ nA at the operating temperature. Its coupling energy is $\hbar I_c/(2e) = 9.87 \times 10^{-24}$ J, corresponding to $E_J/k_B = 0.715$ K. At a bath temperature of 50 mK thermal energy is appreciably smaller, although phase fluctuations can still depend on capacitance and the surrounding circuit. If either electrode is driven normal, the equilibrium phase-link energy disappears and the ideal supercurrent relation ceases to apply. Observing quantized phase

locking and an excitation gap therefore accesses two different signatures of the superconducting state; their characteristic frequencies need not be numerically identical.

The two Josephson relations yield a further calibration test that does not require measuring energy directly. At fixed microwave frequency, the voltage separation of successive locking steps is $hf_{\mathrm{ref}}/(2e)$ independently of junction area in the ideal phase model. Increasing junction area can change total critical current and the visibility or width of current intervals over which locking persists, yet should leave the quantized voltage spacing fixed. Looking at spacing and current width separately distinguishes uncertainty in the reference frequency from a change in the junction coupling or its electromagnetic environment.

Estimate a spectral gap when peaks have finite width

Add a phenomenological lifetime broadening to the isotropic quasiparticle density of states by evaluating $z/\sqrt{z^2-1}$ at $z=(E+i\Gamma)/\Delta$ and taking the physical nonnegative real density. Hold the gap itself fixed in all three runs.



A broader peak and finite central density are possible even with the same gap parameter. At E = 0 the ratio is γ/√(1+γ²), where γ = Γ/Δ. Inferring the gap from a threshold in conductance also requires the thermal and instrumental resolution.

Try it · Dynes

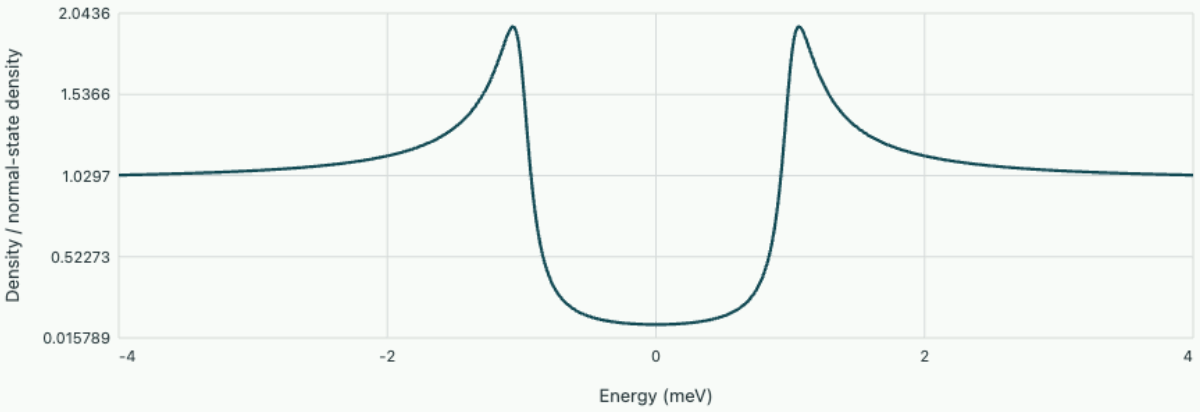
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Lifetime broadening Γ/Δ **0.1**

Gap Δ (meV) **1**



Broadened superconducting spectrum

Zero-energy density ratio 0.099504. This density-of-states model precedes thermal and instrument convolution.

▶ **Read values from the plot**

Study the full derivation

Infer a superconducting gap from a tunnelling spectrum · Structure, excitations and probes of solids

A tunnel junction reports differential conductance versus voltage. Extracting a superconducting energy gap requires connecting a condensate's quasiparticle dispersion to a density of states, then accounting for temperature and finite spectral resolution before comparing peaks with the data.

Problem sheet 6 · Screening, spectra and the meaning of a gap

Work out the length and energy scales before deciding which measurement would identify the underlying order.

1. A magnetic profile below a transition

A thick superconducting half-space has $B(120\ \mathrm{nm}) = 0.20 B(0)$ in a low parallel field. (a) Infer the penetration depth and predict the ratio at 240 nm. (b) If superfluid density falls by a factor four, how does the profile change at 120 nm, assuming constant mass? (c) Why could this estimate fail in a mixed state?

▼ **Use a length inferred from an amplitude**

For an exponential $\lambda = 120/\ln 5 = 74.6$ nm. At double depth the ratio is

$$0.20^2 = 0.040.$$

A factor-four decrease in density doubles λ; at the original depth the ratio becomes

$$\sqrt{0.20} = 0.447.$$

Vortex penetration creates spatially varying fields and the measured signal may no longer describe screening from one flat surface.

---

2. Flux and a tunnelling spectrum

A thick ring encloses 6.20×10^{-15} Wb on a path with negligible current. A normal-metal–superconductor junction on the same material shows peaks at ± 0.90 mV. (a) Estimate the flux integer for charge $2e$. (b) Infer the gap and weak-coupling estimate of T_c. (c) Identify two independent quantities a measurement could check.

▼ **Keep flux quantisation independent of the spectroscopic estimate**

Dividing by 2.068×10^{-15} Wb gives approximately three flux quanta. The peak magnitude suggests $\Delta = 0.90$ meV. Using $k_B = 0.08617$ meV/K yields

$$T_c = 1.80/(3.53 \times 0.08617) = 5.92\,\mathrm{K}.$$

A magnetisation or heat-capacity measurement locates the transition independently; a scan over ring flux states tests the period without using the tunnelling interpretation.

---

3. Two missing signals with different mechanisms

A dielectric multilayer has a narrow optical transmission minimum. A separate metal film develops a low-temperature gap in tunnelling conductance. (a) Give a controlled change and a quantitative limit for the optical structure. (b) Give two independent tests of the superconducting interpretation. (c) Explain why small transmission alone cannot establish dissipation.

▼ **Test the mechanism that produced the missing signal**

For the lossless two-layer stack, take both indices to n at fixed cell size. The first folded branches then meet at $fa/c = 1/(2n)$, with no forbidden interval. For the film, follow the tunnelling feature with temperature and check for low-field magnetic screening at the same transition. A lossless stack can reject transmission by reflection; testing absorption requires an energy balance including reflected power and other output channels.

Infer a Fermi surface from a transport measurement

A three-dimensional metal at low temperature has Hall coefficient -6.24×10^{-11} m³/C and resistivity 2.00×10^{-8} Ω m. Assume a single isotropic electron band with free-electron mass and spin degeneracy two. Estimate the carrier density, Fermi wavevector, scattering time and mean free path. Explain what a second mobile carrier species would change.

▼ **Worked solution**

In the one-band Drude limit $R_H = -1/(ne)$, giving $n \approx 1.00 \times 10^{29}$ m^{-3}. Counting occupied states in a sphere with two spin states gives $n = k_F^3/(3\pi^2)$, so $k_F \approx 1.44 \times 10^{10}$ m^{-1}. The associated Fermi speed $\hbar k_F/m_e$ is 1.66×10^6 m/s.

The conductivity is $1/\rho = ne^2\tau/m_e$. Therefore $\tau \approx 1.77 \times 10^{-14}$ s and $\ell = v_F \tau \approx 29.5$ nm. Multiplying k_F by ℓ gives about 423, comfortably larger than unity. The lifetime only becomes a distance after choosing the velocity of carriers at the Fermi surface.

For simultaneous electron and hole conduction, the weak-field Hall signal weights the densities with squared mobilities and with opposite signs. The longitudinal conductivity instead adds their positive mobility-weighted contributions. A small Hall coefficient can then arise from compensation. Inserting it into the single-band density formula would give a misleadingly large density.

Explore these relations

- For magnetic penetration, compare finite-slab profiles with the thick-sample exponential limit used in the question. Open the live calculation.
- For the tunnelling spectrum, vary gap and lifetime broadening and distinguish spectral filling from the gap scale. Open the live calculation.

Chapter 13 · How alternating bonds split a one-dimensional band

Two periodic chains contain precisely the same number of atoms at the same spacing. One has equal bonds and the other alternates strong and weak bonds. The difference changes which electron energies can propagate, even without changing the on-site energy of any atom.

Choose a primitive cell that repeats the bonds

Let a cell of length a contain orbitals A and B. An electron hops between A and B in the same cell with real amplitude $t_1 > 0$, and between B and A in the next cell with amplitude $t_2 > 0$. Use one orbital per site, neglect interactions and set both on-site energies to zero. We can absorb an overall hopping sign into the B basis state. Translation by one cell multiplies each component by e^{ika}, although the relative amplitudes on A and B remain unknown. Writing those amplitudes as A, B turns the real-space nearest-neighbour equations into

$$EA = (t_1 + t_2 e^{-ika})B,$$

$$EB = (t_1 + t_2 e^{ika})A.$$

The cell boundary crossed by the second hopping accounts for its phase. The two coefficients are complex conjugates as required by Hermiticity. Substitute the second relation into the first without dividing by energy. A nonzero component then requires

$$E^2 = (t_1 + t_2 e^{ika})(t_1 + t_2 e^{-ika}).$$

The two energies at each wavevector are

$$E_{\pm}(k) = \pm\sqrt{t_1^2 + t_2^2 + 2t_1 t_2 \cos ka}.$$

For positive hoppings their magnitude is largest at $k = 0$, namely $t_1 + t_2$, and smallest at $ka = \pi$, namely $|t_1 - t_2|$. The gap is $2|t_1 - t_2|$. At one spinless electron per cell the lower band is full, so this is also the independent-particle excitation gap in a large periodic chain.

Read the band formation from two limits

Take $t_1 = 1.0\,\mathrm{eV}$, $t_2 = 0.6\,\mathrm{eV}$. The lower band runs from -1.6 to $-0.4\,\mathrm{eV}$, the upper from 0.4 to $1.6\,\mathrm{eV}$. The gap is $0.8\,\mathrm{eV}$, while each bandwidth is $1.2\,\mathrm{eV}$. If t_2 tends to zero, the cell becomes an isolated pair with two discrete bonding energies $\pm t_1$. Bands flatten because cells can no longer exchange probability. At $t_1 = t_2 = t$, the spectrum becomes $\pm 2t|\cos(ka/2)|$ in the reduced zone. The two pieces meet at its edge. The two-site construction then folds the simple uniform chain into a smaller wavevector interval.

One can also read off a velocity. Away from a crossing, differentiate $E_{\pm}$ to obtain

$$v_{\pm} = \mp a t_1 t_2 \sin(ka)/[\hbar |E_{\pm}|].$$

At extrema of a gapped band it vanishes. At the equal-hopping crossing the derivative written as a ratio has a removable directional ambiguity, because the ordered upper and lower bands meet with opposite slopes.

Exchange the two bonds

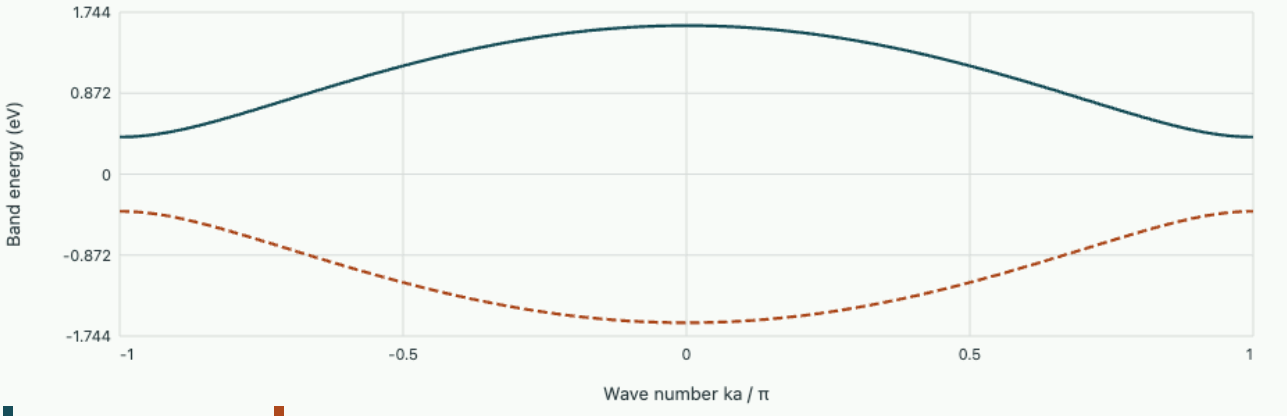
Keep t_1 fixed and move t_2 through it. Compare both the gap and the bandwidth. Interchanging the numerical values of the two hoppings exactly preserves the bulk energies. The experiment in the next chapter distinguishes these patterns by opening the chain at a specified bond. A termination supplies physical information that the unordered hopping pair and the energy graph alone cannot encode.

Open this experiment online

Try the calculation

Intracell hopping t_1 (eV) **1**

Intercell hopping t_2 (eV) **0.6**



Direct gap 0.8 eV. Each band spans 1.2 eV.

Tutorial on inferring the hopping strengths

A measured dimer band has gap $0.50\,\mathrm{eV}$ and reaches $1.45\,\mathrm{eV}$ above the symmetry centre at $k = 0$. Find the two hopping amplitudes and predict the upper-band width.

▼ **Work through the solution**

The sum is 1.45 eV and the absolute difference is 0.25 eV. Adding and subtracting gives the unordered pair 0.85 and 0.60 eV. The upper bandwidth is

$$1.45 - 0.25 = 1.20\,\mathrm{eV}.$$

Bulk energies do not determine which hopping is interrupted by a particular surface cut.

Expand the squared band energy near the Brillouin-zone boundary for unequal hoppings. Identify the small-q dispersion and its range.

▼ Work through the solution

Set $ka = \pi + qa$. Since $\cos(\pi + qa) \simeq -1 + (qa)^2/2$,

$$E^2 \simeq (t_1 - t_2)^2 + t_1 t_2 (qa)^2.$$

For sufficiently small q the upper edge has quadratic curvature $E_+ \simeq |t_1 - t_2| + t_1 t_2 (qa)^2/[2|t_1 - t_2|]$. The last expansion fails when its second term in the squared energy competes with the gap term.

If the same constant on-site energy U is added to every site, which inferred quantities would move in an experiment measuring energies against a fixed reference?

▼ Work through the solution

The Hamiltonian gains UI, moving both bands and their centre by U. Their gap, widths, group velocities, and the hopping estimates from differences remain unchanged. The half-filling chemical potential can still lie inside the shifted gap.

Further reading · Asbóth, Oroszlány and Pályi, SSH model and bulk bands.

Chapter 14 · Finding an end state one bond at a time

The bulk spectrum of an alternating chain is unchanged when strong and weak bonds exchange places. A boundary with a named first site breaks this equivalence. The zero-energy equations provide its wavefunction without diagonalising a large matrix.

Start at an exposed A site

Use the alternating chain with intracell link t_1 and intercell link t_2, both positive. End the left edge at A_1. The first B site connects A_1 and A_2. Seek an eigenstate at $E = 0$ supported only on A sites. Each A-site equation then holds because all neighbouring B amplitudes vanish. Each B-site equation requires cancellation between its two neighbouring A amplitudes. Hence

$$0 = t_1 A_j + t_2 A_{j+1},$$

$$A_j = (-t_1/t_2)^{j-1} A_1.$$

Relative signs matter in the cancellation, even though a local probability measurement sees only squared magnitudes. On a semi-infinite chain the geometric series for the norm converges only for $r = t_1/t_2 < 1$. Normalising gives

$$|A_1|^2 = 1 - r^2.$$

The amplitude decay length satisfies $e^{-a/\ell} = r$, so $\ell = a/\ln(t_2/t_1)$. The probability decay length is half that. As the bulk gap closes, the decay length diverges and the bound state spreads into the interior.

Impose the far boundary too

A finite chain can have M A sites and $M - 1$ B sites, ending on A at both edges. Every B site still imposes exactly the recurrence above, leaving one overall free amplitude. The resulting state has exactly zero energy for any positive hoppings. Its finite norm is the sum of M successive powers of r^2. At $r = 1$ the apparent zero over zero in the closed formula is resolved by summing M unit terms. The state then has probability $1/M$ at every A site.

When $r > 1$ its norm remains finite in this finite sample and the weight accumulates at the right edge. An equal-sublattice finite chain ending on B instead has an additional boundary equation. Two exponentially localised end solutions can then hybridise, shifting their energies slightly above and below zero. Equal sublattice counts alone provide no exact zero eigenvalue. A change of boundary conditions can therefore alter the interpretation of a measured near-zero peak.

Connect the recurrence to the bulk phase

The off-diagonal matrix element $h = t_1 + t_2 e^{ika}$ traces a circle centred at t_1 with radius t_2 as k crosses the Brillouin zone. If $t_2 > t_1$ it winds once around zero. If $t_2 < t_1$ it misses zero. Passing between the two with fixed cell choice requires the gap $2|h|$ to close. A diagonal potential on just one sublattice would

break the off-diagonal symmetry used to pin energies to zero. The winding statement and the boundary argument require a consistent assignment of the unit cell and termination.

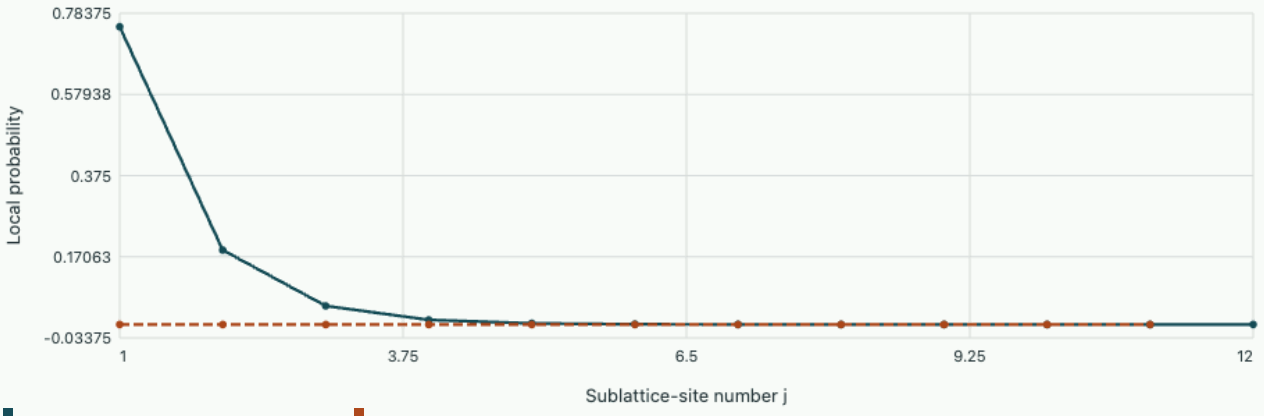
For $r = 0.5$ the first three semi-infinite A probabilities are 0.75, 0.1875 and 0.046875. For $r = 0.90$, the first site has only 0.19 and the amplitude decay length increases from $1.44a$ to $9.49a$. Use the finite odd chain below to move the density from one end to the other. At reciprocal hopping ratios the two distributions must be mirror images. Check this using the endpoint readouts.

Open this experiment online

Try the calculation

Hopping ratio t_1/t_2 **0.5**

Number of A sites **12**



Probabilities at the first and last A sites are 0.75 and 1.788e-7. The amplitude decay length from the favoured edge is 1.443 cell spacings.

Tutorial on inspecting a local measurement

Four successive A sites have relative intensities $1, 0.36, 0.1296, 0.0467$ near a left edge. Infer the hopping ratio and amplitude decay length. Would a local intensity probe detect the alternating signs of the state?

▼ **Work through the solution**

Ratios of neighbouring probabilities equal r^2, hence $r = 0.60$. The amplitude length is

$$-a/\ln(0.60) = 1.96a.$$

A measurement proportional to the absolute square loses the sign. Interference with a phase reference or a measurement sensitive to the amplitude across a bond can access the phase relation.

A finite odd chain has five A sites and hopping ratio $r = 0.8$. What fraction of zero-mode probability lies on its first two A sites?

▼ **Work through the solution**

Sum the finite geometric sequence. The requested fraction is

$$(1 + r^2) / \sum_{j=0}^{4} r^{2j} = (1 - r^4)/(1 - r^{10}) = 0.661.$$

Taking the semi-infinite formula would give

$$1 - r^4 = 0.590.$$

The far boundary changes the normalisation substantially for this short array.

Can bond disorder that preserves positive nonzero nearest-neighbour hoppings destroy the exact zero solution of a finite chain ending on A at both ends?

▼ **Work through the solution**

At each B site retain the local recursion $A_{j+1} = -(t_{1j}/t_{2j})A_j$. It produces a finite nonzero solution after choosing A_1, with all B amplitudes zero. The profile is a product of local ratios and may cease to look like a single exponential. A generic A-site potential would invalidate the zero-energy equation.

Further reading · Asbóth, Oroszlány and Pályi, edge states and bulk-boundary correspondence.

Chapter 15 · Why a honeycomb lattice has Dirac cones

In an alternating chain the two paths from A to B can cancel at a special wavevector. The honeycomb lattice adds a third path. Their cancellation now occurs at isolated points in a two-dimensional Brillouin zone and produces a linear low-energy dispersion.

Add the three Bloch phases

Each honeycomb A site has three nearest B neighbours at displacements $d(0,1)$, $d(\sqrt{3}/2,-1/2)$, and $d(-\sqrt{3}/2,-1/2)$, where d is the bond length. For equal real hopping magnitude t, sum the translation phase for each displacement. The common off-diagonal factor is

$$f(\boldsymbol{k}) = e^{ik_y d} + 2e^{-ik_y d/2}\cos(\sqrt{3}k_x d/2),$$

$$E_\pm = \pm t|f|.$$

This uses a basis phase convention that includes the orbitals' intra-cell positions. A rephasing can change the phase of f but leaves its magnitude invariant. Along $k_y = 0$, cancellation occurs where $1 + 2\cos(\sqrt{3}k_x d/2) = 0$. Choose $K = (4\pi/(3\sqrt{3}d), 0)$. The counterpart at $-K$ is inequivalent up to primitive reciprocal vectors. Near either valley many wavevectors represent distinct states even when their energies almost coincide.

Linearise at a zero, then test the approximation

Write $\boldsymbol{k} = K + \boldsymbol{q}$ with $qd \ll 1$. At K the cosine equals $-1/2$ and the sine $\sqrt{3}/2$. Differentiating the cosine term with respect to k_x consequently gives $-3d/2$. Differentiating the two exponential factors with respect to k_y gives

$$id + id/2 = 3id/2.$$

The constant terms cancel. Therefore

$$f(K+\boldsymbol{q}) \simeq \tfrac{3d}{2}(-q_x + iq_y),$$

$$E_\pm \simeq \pm\hbar v_F|\boldsymbol{q}|,$$

$$v_F = \frac{3td}{2\hbar}.$$

With $t = 2.7\,\mathrm{eV}$, $d = 0.142\,\mathrm{nm}$ and $\hbar = 0.6582\,\mathrm{eV\,fs}$, the speed is

$$0.874\,\mathrm{nm/fs} = 8.74 \times 10^5\,\mathrm{m/s}.$$

At $qd = 0.05$ the leading energy magnitude is $0.2025\,\mathrm{eV}$. The energy varies linearly with momentum displacement even though the microscopic Hamiltonian contains only hopping between neighbours. The two components describe sublattice amplitude, distinct from the real electron spin.

Break the equivalence of the sublattices

Assign on-site energies $+\Delta$ to A and $-\Delta$ to B. The characteristic equation of the two-by-two Hamiltonian becomes

$$E^2 = \Delta^2 + t^2|f|^2.$$

A gap $2|\Delta|$ opens at K. For $\Delta = 0.10\,\mathrm{eV}$, the edge sits at $0.10\,\mathrm{eV}$ and the predicted conduction energy at $qd = 0.05$ becomes about $0.226\,\mathrm{eV}$. Far enough from the origin the slopes recover their nearly linear values. Close to it, expansion of the square root gives

$$E_+ \simeq |\Delta| + \hbar^2 v_F^2 q^2/(2|\Delta|).$$

A band-edge mass $|\Delta|/v_F^2$ has emerged. The expansion requires $\hbar v_F q \ll |\Delta|$.

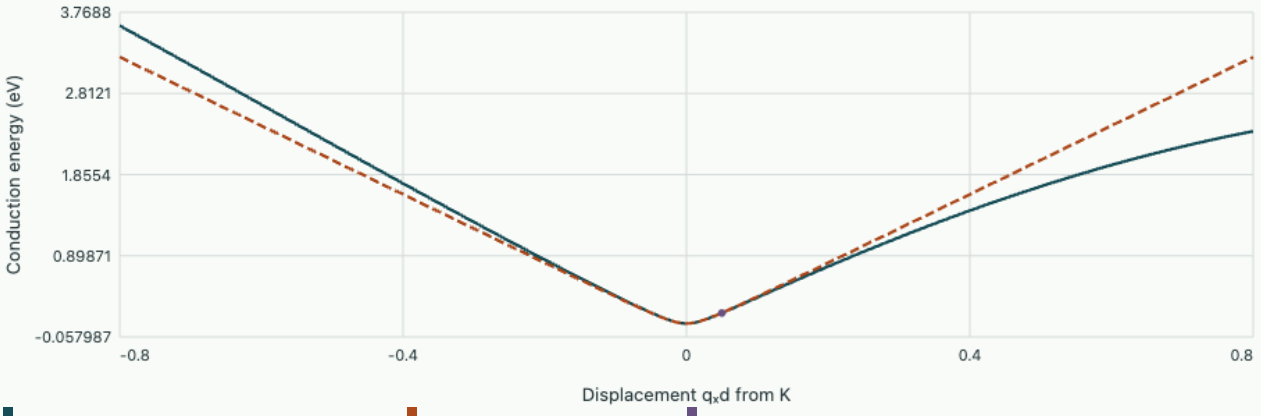
Compare the exact positive-energy cut along q_x with the local cone in the graph. Unequal deviations to the right and left of K are allowed by the lattice and cannot be captured by the lowest-order isotropic cone. The reflected cut at the other valley restores the time-reversal relation.

Open this experiment online

Try the calculation

Sublattice energy offset Δ (eV) **0.1**

Selected displacement $q_x d$ **0.05**



Direct gap 0.2 eV. At the selected displacement the lattice gives 0.2235 eV and the local model 0.2258 eV.

Tutorial on a measured slope

A near-valley spectrum has slope $0.55\,\mathrm{eV\,nm}$. Infer v_F and the nearest-neighbour hopping for a bond length $d = 0.142\,\mathrm{nm}$. A gap of $80\,\mathrm{meV}$ then appears with the same asymptotic slope. What is $|\Delta|$?

▼ **Work through the solution**

The slope equals $\hbar v_F$. Dividing by $0.6582\,\mathrm{eV\,fs}$ gives $0.836\,\mathrm{nm/fs}$. Since the slope also equals $3td/2$,

$$t = 2(0.55)/(3 \times 0.142) = 2.58\,\mathrm{eV}.$$

The two edges shift symmetrically by half the full gap, giving $|\Delta| = 40\,\mathrm{meV}$.

Consider a gapless honeycomb cone exactly at the valley centre. What does the tight-binding Hamiltonian predict for the energy and eigenvectors, and what information comes from approaching along a direction?

▼ **Work through the solution**

At $f = 0$ the matrix vanishes, so every two-component vector in the degenerate subspace has energy zero. For a nearby nonzero f, a chosen eigenbranch fixes a relative phase between its A and B components according to the argument of f. The limiting eigenvector therefore depends on the direction of approach even though the limiting energy is always zero.

At $q = 0$, repeat the eigenvector argument with $\Delta > 0$. Predict a consequence for sublattice-resolved spectroscopy at the two band edges.

▼ **Work through the solution**

At the valley centre the Hamiltonian is diagonal with entries $+\Delta, -\Delta$. The upper state lives on A and the lower on B in this basis. A probe selective to A should see more weight at the conduction edge than at the valence edge. Farther from the valley centre the hopping mixes the sublattices again.

Further reading · Delft Topology in Condensed Matter, honeycomb Hamiltonian and Dirac points.

Chapter 16 · Counting carriers and weighing a cyclotron orbit

A conical band has no constant quadratic band mass at its touching point. A carrier on a finite-energy orbit still has a measurable cyclotron mass. Counting states fixes the Fermi circle and determines how this mass changes when a gate adds charge.

Count states in a two-dimensional disc

Put a two-dimensional sheet of area $A = L_x L_y$ under periodic boundary conditions. Neighbouring allowed wavevectors differ by $2\pi/L_x$ and $2\pi/L_y$, so each state occupies reciprocal-space area $(2\pi)^2/A$. At zero temperature fill all conduction states up to radius k_F, measured from a Dirac point. There are two spin states and two inequivalent valleys in ideal monolayer graphene, giving total degeneracy $g = 4$. Density relative to charge neutrality is therefore

$$n = g\frac{\pi k_F^2}{(2\pi)^2} = \frac{k_F^2}{\pi},$$

$$E_F = \hbar v_F \sqrt{\pi n}.$$

For electron and hole densities use $|n|$ under the square root and give E_F the sign of n relative to the touching point. Differentiating the state count with respect to energy gives the density of states per area $D(E) = g|E|/[2\pi(\hbar v_F)^2]$. It tends to zero at an ideal neutral cone. Thermal broadening, disorder and a gap all change what a finite-resolution probe observes near neutrality.

Background · Cyclotron mass

$$m_c = (\hbar^2/2\pi)\,\partial A/\partial E$$

The area is that of a closed constant-energy orbit in reciprocal space normal to the magnetic field. This is a measured orbital mass rather than a universal free-particle parameter.

Let a perpendicular field bend the orbit

For a wavepacket in a band with weak magnetic field,

$$\hbar\dot{\boldsymbol{k}} = -e\,\boldsymbol{v} \times \boldsymbol{B},$$

where e denotes the positive elementary charge. An isotropic band has velocity normal to a circular constant-energy contour. The orbit is traversed with speed $|\dot{\boldsymbol{k}}| = ev_F B/\hbar$. Dividing its circumference $2\pi k_F$ by that speed gives the period. Express its angular frequency as eB/m_c to identify

$$m_c = \frac{\hbar k_F}{v_F} = \frac{|E_F|}{v_F^2}.$$

The mass measures the time needed to travel around a contour. It depends on the radius even when the magnitude of the group velocity stays fixed. For a general band area $\mathcal{A}(E)$, neighbouring contours separated perpendicular to themselves by $dk_\perp$ differ in energy by $\hbar v\, dk_\perp$. Thus $|d\mathcal{A}/dE| = \oint dl/(\hbar v)$, which gives $m_c = \hbar^2 |d\mathcal{A}/dE|/(2\pi)$ by integrating the same equation of motion.

Convert a gate density into an energy and a mass

At $n = 10^{12}\,\mathrm{cm}^{-2} = 10^{16}\,\mathrm{m}^{-2}$ and $v_F = 10^6\,\mathrm{m/s}$, one obtains $k_F = 1.772 \times 10^8\,\mathrm{m}^{-1}$, $E_F = 0.117\,\mathrm{eV}$ and $m_c = 1.87 \times 10^{-32}\,\mathrm{kg} = 0.0205 m_e$. Quadrupling the density doubles the Fermi energy and cyclotron mass. By contrast a single isotropic parabolic band $E = \hbar^2 k^2/(2m)$ has the same counting geometry but constant $m_c = m$. Transport mobility or disorder scattering time does not enter the state count, although scattering can obscure the cyclotron motion.

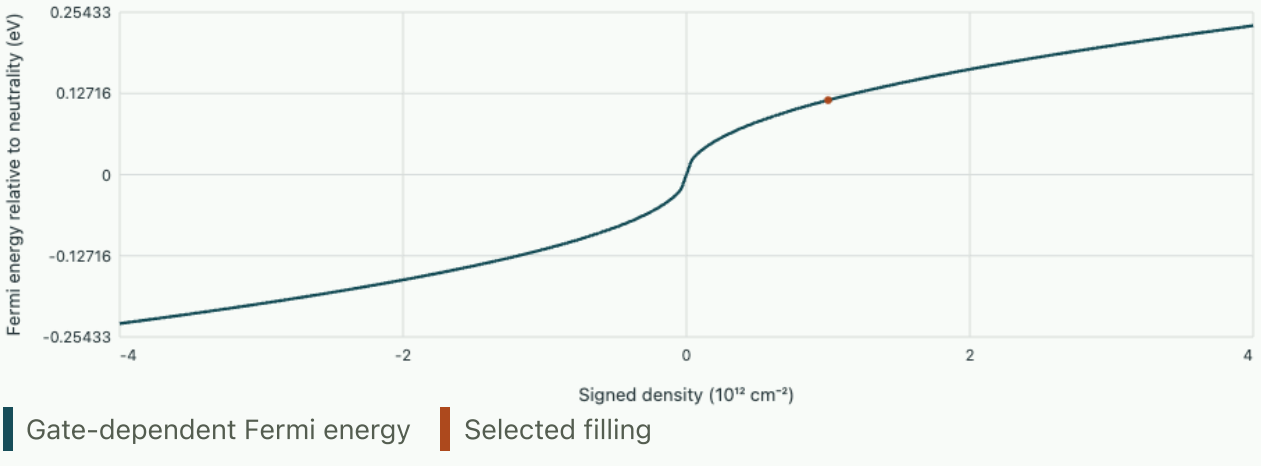
Change velocity and signed density. Raising the velocity at fixed density increases $|E_F|$ while reducing the mass because the orbit radius stays fixed and the wavepacket moves faster. Exactly at charge neutrality, a single classical Fermi orbit no longer describes the carriers. Finite-field quantum levels replace the continuum picture in the next chapter.

Open this experiment online

Try the calculation

Signed density (10^{12} cm^{-2}) 1

Cone velocity (10^6 m/s) 1



Fermi energy 0.1167 eV, wave number 0.1772 nm^{-1}, cyclotron mass 0.02052 electron masses.

Tutorial on a density sweep

Two cyclotron measurements on the same gapless sheet return masses $0.024 m_e$ and $0.036 m_e$. Assume the velocity is independent of density. Infer the ratio of absolute carrier densities and the change in the weak-field cyclotron frequency at fixed field.

▼ **Work through the solution**

Since m_c scales as $\sqrt{|n|}$, the second density is $(0.036/0.024)^2 = 2.25$ times the first. The frequency eB/m_c drops to two thirds of its first value. The masses alone cannot reveal whether either measurement used electrons or holes.

Find the weak-field cyclotron radius at the worked density $10^{12}\,\mathrm{cm}^{-2}$ and field $0.10\,\mathrm{T}$. How does raising the velocity at fixed density affect it?

▼ **Work through the solution**

The semiclassical radius is

$$r_c = v_F m_c/(eB) = \hbar k_F/(eB).$$

With $k_F = 1.772 \times 10^{8}\,\mathrm{m}^{-1}$,

$$r_c = 1.17\,\mu\mathrm{m}.$$

At fixed density and field the radius stays fixed when velocity changes. The orbit period changes because the speed around that circle changes.

A measurement doubles positive E_F at constant temperature sufficiently close to zero for the step filling approximation. How do state density per area and density of states per energy change?

▼ **Work through the solution**

For fixed v_F, the occupied density scales as E_F^2, so it rises fourfold. The derivative $D(E_F)$ scales as $|E_F|$ and doubles. Confusing the count with its derivative would give the wrong charge change inferred from spectroscopy.

Further reading · Novoselov and colleagues, measured energy dependence of the cyclotron mass.

Chapter 17 · Quantising the orbit and finding the zero Landau level

In a magnetic field the two kinetic momentum components cannot be specified simultaneously. In a conical band their commutator turns the two sublattice equations into coupled oscillator equations. Their spectra differ by one state, leaving an unpaired level at zero energy.

Construct the oscillator from kinetic momentum

In one valley of a gapless monolayer use the long-wavelength Hamiltonian $H = v_F(\sigma_x \Pi_x + \sigma_y \Pi_y)$, with Pauli matrices acting on sublattices. For an electron of charge $-e$,

$$\boldsymbol{\Pi} = \boldsymbol{p} + e\boldsymbol{A}.$$

In a uniform perpendicular field $B > 0$, applying $[p_j, A_k] = -i\hbar \partial_j A_k$ gives

$$[\Pi_x, \Pi_y] = -ie\hbar B.$$

Introduce $a = (\Pi_x - i\Pi_y)/\sqrt{2e\hbar B}$. Its adjoint uses the opposite sign, giving $[a, a^\dagger] = 1$. The magnetic length $\ell_B = \sqrt{\hbar/(eB)}$ sets the spatial size of the oscillator.

$$H = v_F\sqrt{2e\hbar B}\begin{pmatrix} 0 & a \\ a^\dagger & 0 \end{pmatrix}, \qquad H^2 = 2e\hbar B v_F^2 \begin{pmatrix} aa^\dagger & 0 \\ 0 & a^\dagger a \end{pmatrix}.$$

In the B component the squared Hamiltonian has oscillator numbers $0, 1, 2, \ldots$. In the A component $aa^\dagger = a^\dagger a + 1$, so its sequence starts at one. For integer $n \geq 1$ combine $|n-1\rangle$ on A with $\pm|n\rangle$ on B in equal magnitudes. Acting with H confirms energies $E_{n,\pm} = \pm v_F\sqrt{2e\hbar B n}$. For $n = 0$, the B-only spinor with $|0\rangle$ is annihilated by a. Its energy vanishes. Valley and field conventions can exchange which sublattice carries that state.

Count independent centres at the same energy

Choose gauge $A_y = Bx$ and retain translation along y. Quantisation gives

$$k_y = 2\pi j / L_y.$$

The kinetic momentum $\Pi_y = \hbar k_y + eBx$ places the orbit centre at $x_0 = -\hbar k_y/(eB)$. Successive centres differ by $2\pi\ell_B^2/L_y$. Counting centres in a width L_x gives $L_x L_y/(2\pi\ell_B^2)$, or eB/h states per area for each spin and valley, apart from edge corrections. This degeneracy is independent of level index. Spin splitting, lattice effects and interactions are set aside here.

At $B = 1\,\mathrm{T}$, $v_F = 10^6\,\mathrm{m/s}$, $\ell_B = 25.7\,\mathrm{nm}$ and the first positive level is $36.3\,\mathrm{meV}$ above zero. The next sits at $51.3\,\mathrm{meV}$. The spacing between the first two positive levels is consequently only 15.0 meV. At 4 T all these energies double while the number of orbit centres per unit area quadruples. A parabolic band would instead give equally spaced levels at fixed field.

Identify the measurable dependence

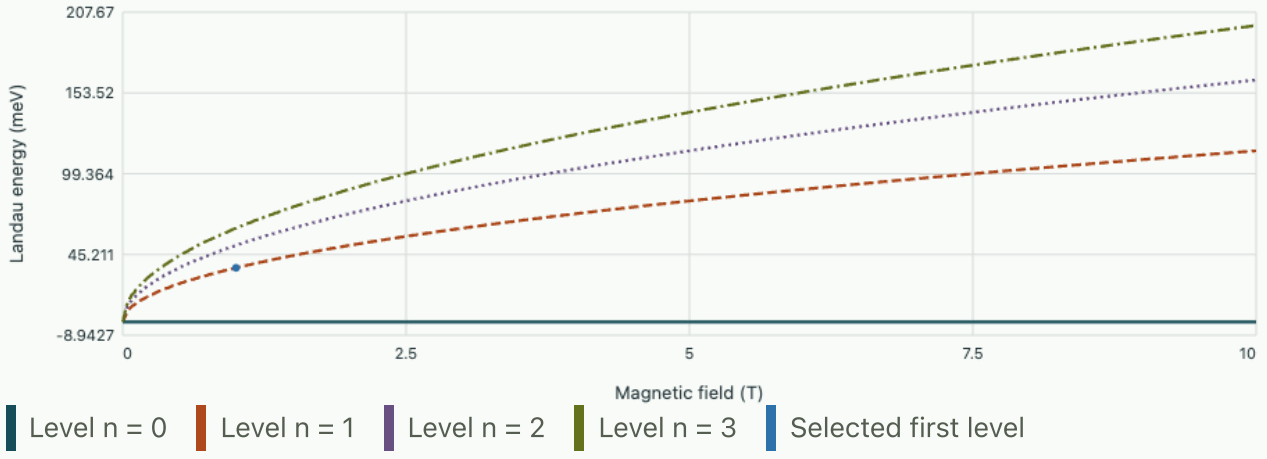
Compare the first few positive levels as field changes. Dividing E_n by $\sqrt{Bn}$ should return one material-dependent coefficient. An optical measurement sees transitions between occupied and available states, subject to matrix-element selection rules. A measured transition energy therefore need not coincide with the energy of an individual level. At zero field this discrete-level construction loses its normalisable magnetic oscillator; the limiting energies on the graph describe the approach from finite field.

Open this experiment online

Try the calculation

Perpendicular field (T) 1

Cone velocity (10⁶ m/s) 1



First level 36.28 meV, first-to-second spacing 15.03 meV. Each spin and valley has 2.418 × 10¹⁰ orbit centres per cm².

Tutorial on two spectral lines

At $B = 2\,\mathrm{T}$ an ideal sample has a zero-to-first level spacing of $48\,\mathrm{meV}$. Infer its velocity relative to $10^6\,\mathrm{m/s}$. Predict the first-to-second positive-level spacing at $8\,\mathrm{T}$.

▼ **Work through the solution**

The reference first level at 2 T is $36.3\sqrt{2} = 51.3\,\mathrm{meV}$, so $v_F \simeq 0.935 \times 10^6\,\mathrm{m/s}$. Raising the field by a factor of four doubles all energies. The first level at 8 T is 96 meV, and $E_2 - E_1 = 96(\sqrt{2} - 1) = 39.8\,\mathrm{meV}$. This inference assumes an ungapped cone without interaction shifts.

At $B = 2\,\mathrm{T}$, estimate the number of orbit centres in area $1\,\mu\mathrm{m}^2$ for a single level. Give both the per-flavour count and the full spin–valley count.

▼ **Work through the solution**

Use eBA/h, with $A = 10^{-12}\,\mathrm{m}^2$. This gives roughly 484 centres per flavour and 1934 across four flavours. The numbers are approximate because a finite sample admits edge corrections, whereas the area formula is a bulk degeneracy density.

At high positive level index show how the level spacing regains the semiclassical cyclotron frequency eB/m_c.

▼ **Work through the solution**

Write $E_n^2 = 2e\hbar B v_F^2 n$. At large n the difference $E_{n+1} - E_n \simeq dE/dn = e\hbar B v_F^2 / E_n$. Dividing by $\hbar$ gives eB/m_c with $m_c = E_n / v_F^2$. This connection uses many occupied levels, while the unpaired zero state is outside the semiclassical limit.

Further reading · Jiang and colleagues, infrared spectroscopy of Landau levels in monolayer graphene · Li and colleagues, spatially resolved graphene Landau levels.

Chapter 18 · How the same electrons carry heat and charge

A metal with high electrical conductivity often also conducts heat well. The quantitative relation involves an open-circuit measurement, the thermal width of the Fermi surface, and the energy dependence of the states able to carry current.

Weigh the states that can carry a current

Use an isotropic band of quasiparticles with energy E, velocity magnitude $v(E)$, density of states per volume $D(E)$ and elastic relaxation time $\tau(E)$. In a weak electric field or temperature gradient the departure from local equilibrium is proportional to the relaxation time times the applied driving term. For a given transport direction the angular average gives $\langle v_x^2 \rangle = v^2/3$ in three dimensions. Define $\Sigma(E) = D(E)v(E)^2\tau(E)/3$. Electrical current weights a particle by its charge $q = -e$, while heat current weights it by $E - \mu$. Three moments consequently suffice for the two linear responses:

$$K_j = \int \Sigma(E)(E-\mu)^j[-\partial_E f(E)]\,dE,$$

$$j = 0, 1, 2.$$

Let $\mathcal{E}$ denote the electric/electrochemical field with any chemical-potential gradient properly included. The resulting currents along one direction are $j_e = q^2K_0\mathcal{E} - qK_1\nabla T/T$ and $j_Q = qK_1\mathcal{E} - K_2\nabla T/T$. The factor $E - \mu$ in the thermal drive follows from differentiating the Fermi function with respect to temperature. The repeated first moment is a manifestation of reciprocal coupling of electrical and thermal transport in this zero-field model.

Impose the actual electrical boundary condition

A voltage probe draws negligible charge current. Set $j_e = 0$ to obtain $\mathcal{E} = S\nabla T$ with $S = K_1/(qTK_0)$. Substitute this field into the heat current to get

$$\kappa_e = (K_2 - K_1^2/K_0)/T.$$

Under the positive measure $\Sigma(-f')dE/K_0$, the bracket divided by K_0 is the variance of $E - \mu$. Hence the open-circuit electronic thermal conductivity is nonnegative. Measuring with zero imposed electric field would retain a thermally driven charge current and a different heat-flow coefficient.

For a degenerate metal with smooth Σ, the weight $-f'$ is a narrow even peak. Replacing Σ by $\Sigma(\mu)$ gives $K_1 = 0$, $K_0 = \Sigma(\mu)$, and $K_2 = (\pi^2/3)(k_BT)^2\Sigma(\mu)$. The dimensionless second moment follows by integrating $x^2/[4\cosh^2(x/2)]$ over the real axis. With $\sigma = q^2K_0$,

$$\frac{\kappa_e}{\sigma T} = L_0 = \frac{\pi^2 k_B^2}{3e^2} = 2.44 \times 10^{-8}\,\mathrm{W\,\Omega/K^2}.$$

At $T = 100\,\mathrm{K}$, a metal with $\sigma = 2.0 \times 10^7\,\mathrm{S/m}$ has estimated electronic $\kappa_e = 48.8\,\mathrm{W/(m\,K)}$. A total measured conductivity of 60 would leave about $11\,\mathrm{W/(m\,K)}$ for other carriers if the assumptions hold. Phonons, inelastic scattering or strong energy dependence may change that inference.

Select a narrow range of transport energies

A model energy filter $\Sigma \propto \exp[-(x-c)^2/(2w^2)]$, with $x = (E-\mu)/(k_B T)$, c, w dimensionless, introduces a tunable departure from this metallic limit. The Seebeck coefficient is $(k_B/q)\langle x \rangle$ and the normalised Lorenz ratio is $\mathrm{var}(x)/(\pi^2/3)$. A very narrow filter centred above the chemical potential produces strong signed thermopower but little heat transport at zero charge current, because current-carrying particles have almost identical excess energies. Those statements refer to the ratio and thermal moments; the absolute conductivity also depends on the overall transmission.

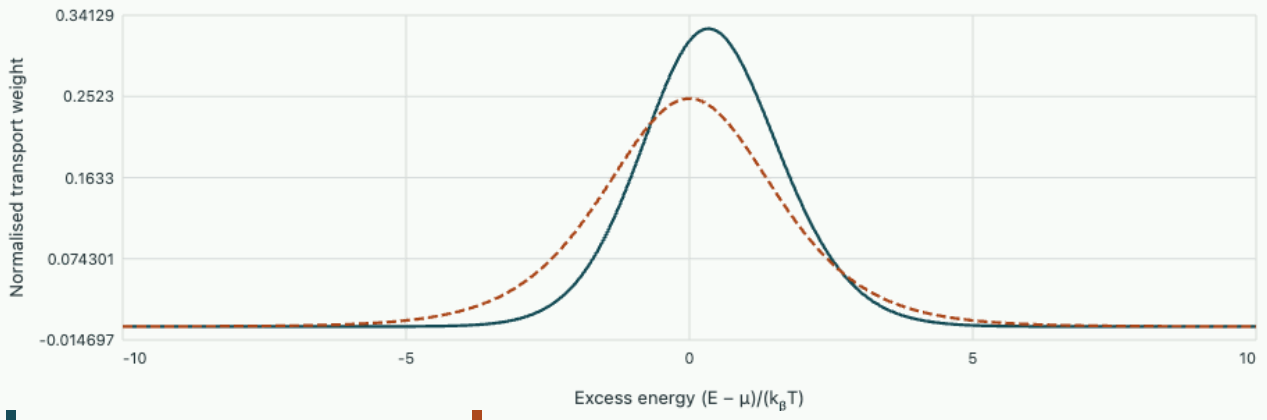
The experiment integrates these moments numerically and plots their normalised weight against the reference Fermi derivative. Vary c from positive to negative at fixed width. Predict which observable reverses sign and which is invariant before reading the numerical results.

Open this experiment online

Try the calculation

Filter centre (E – μ)/(k$_\beta$T) 1

Filter width / (k$_\beta$T) 2



Mean excess energy 0.4018 k$_\beta$T; electron Seebeck coefficient -34.63 μV/K. Open-circuit Lorenz ratio L/L$_0$ = 0.4937.

Tutorial on open-circuit transport

For a fitted transport weight with $\langle x \rangle = 0.50$, $\langle x^2 \rangle = 2.25$, determine the Seebeck coefficient and L/L_0. Use $k_B/e = 86.17\,\mu\mathrm{V/K}$.

▼ **Work through the solution**

For electrons $S = -43.1\,\mu\mathrm{V/K}$. The variance is

$$2.25 - 0.50^2 = 2.00.$$

Hence

$$L/L_0 = 2.00/(\pi^2/3) = 0.608.$$

Using the second moment without subtracting the square of the first would represent the wrong electrical boundary condition.

Expand a smooth transport function to first order about μ. What controls the low-temperature sign of S?

▼ **Work through the solution**

The constant term contributes nothing to K_1 by symmetry. The slope gives

$$K_1 \simeq \Sigma'(\mu)(\pi^2/3)(k_B T)^2.$$

Dividing by qTK_0 yields

$$S \simeq [\pi^2 k_B^2 T/(3q)]\partial_E \ln \Sigma|_\mu.$$

For electrons a rising transport function gives negative S in the stated field convention.

Further reading · ETH Zürich, electronic transport coefficients · University of Pisa, transport integrals and Wiedemann–Franz law.

Study the full derivation

Measure a phonon mean free path without confusing it with a sample length · Structure, excitations and probes of solids

A cold insulating bar can carry heat primarily through lattice vibrations. Heat capacity tells how much energy the vibrations can carry; conductivity also depends on how far and in which direction they carry it. Derive the connection, test its limiting regimes, and identify when a measured thermal resistance belongs to a boundary.

Chapter 19 · Why a polar crystal has two optical phonon frequencies

A uniform relative displacement of positive and negative ions creates an electric polarisation. A longitudinal vibration induces a macroscopic restoring field while a transverse displacement has a different electrical boundary condition. An infrared spectrum can reveal both frequencies.

Separate the ionic and electronic contributions

For one infrared-active optical mode let u be relative displacement per primitive cell, M the effective reduced mass, Z^* the effective charge in coulombs, and N the number of cells per volume. The oscillator equation driven by the macroscopic field is

$$M(\ddot{u} + \gamma\dot{u} + \omega_{\mathrm{TO}}^2 u) = Z^* E.$$

The zero-macroscopic-field stiffness defines the transverse optical frequency in this oscillator model. Bound electronic response at these frequencies supplies a real relative background permittivity ϵ_∞. The ionic polarisation is $P_{\mathrm{ion}} = NZ^*u$, so $D = \epsilon_0\epsilon_\infty E + P_{\mathrm{ion}}$.

With time convention $e^{-i\omega t}$, solving for u gives

$$\epsilon_r(\omega) = \epsilon_\infty + \frac{S}{\omega_{\mathrm{TO}}^2 - \omega^2 - i\gamma\omega},$$

$$S = \frac{N(Z^*)^2}{\epsilon_0 M}.$$

The signed effective charge appears twice, once as drive and once in the dipole density. Its square therefore sets oscillator strength. At zero frequency the ions can follow the field and $\epsilon_s = \epsilon_\infty + S/\omega_{\mathrm{TO}}^2$. Far above the phonon resonance but below electronic absorption the ionic term is small, making ϵ_∞ the observed plateau.

Apply Gauss's law to a longitudinal vibration

Consider a long-wavelength, nonzero wavevector $\boldsymbol{k}$ with no external free charge. Gauss's law gives

$$\boldsymbol{k} \cdot \boldsymbol{D} = 0.$$

For a longitudinal field $D_L = 0$, so the ions create $E_L = -NZ^*u/(\epsilon_0\epsilon_\infty)$. Substituting this opposing field into the undamped oscillator equation adds a restoring term. Thus

$$\omega_{\mathrm{LO}}^2 = \omega_{\mathrm{TO}}^2 + S/\epsilon_\infty,$$

$$\frac{\omega_{\mathrm{LO}}^2}{\omega_{\mathrm{TO}}^2} = \frac{\epsilon_s}{\epsilon_\infty}.$$

Equivalently the longitudinal collective mode is the zero of the lossless dielectric function. The transverse optical frequency is its pole. For a single isotropic mode this is the Lyddane–Sachs–Teller relation. With

several infrared modes, anisotropy, damping or mobile free carriers, assigning one measured peak and one dielectric ratio to a single oscillator needs further justification.

Connect a reflectance band with the two modes

In the lossless limit rewrite $\epsilon_r = \epsilon_\infty(\omega_{\mathrm{LO}}^2 - \omega^2)/(\omega_{\mathrm{TO}}^2 - \omega^2)$. Between the two frequencies this is negative. A transverse electromagnetic wave satisfies $k^2 = \epsilon_r\omega^2/c^2$, so its wavenumber becomes imaginary there in a nonmagnetic lossless crystal. At a planar vacuum boundary a semi-infinite crystal therefore has unit ideal reflectance over this interval. Finite damping makes the refractive index complex and allows absorption, smoothing the edges. The two mode frequencies have been deduced from mechanical response and electric boundary conditions; an electromagnetic wave interacting with the lattice is a coupled phonon–polariton excitation.

For $f_{\mathrm{TO}} = 6.0\,\mathrm{THz}$, $\epsilon_\infty = 4$ and $\epsilon_s = 9$, $f_{\mathrm{LO}} = 9.0\,\mathrm{THz}$. At 7.5 THz in the lossless limit $\epsilon_r = 4(81 - 56.25)/(36 - 56.25) = -4.89$. Below the mode, normal-incidence static-index reflectance is

$$[(\sqrt{9} - 1)/(\sqrt{9} + 1)]^2 = 0.25.$$

Above the phonon plateau the corresponding background reflectance approaches $1/9$. The interactive plot evaluates the passive complex refractive index and the exact single-interface reflectance. Change the static permittivity and damping; compare the high-reflection interval with the lossless zero and pole.

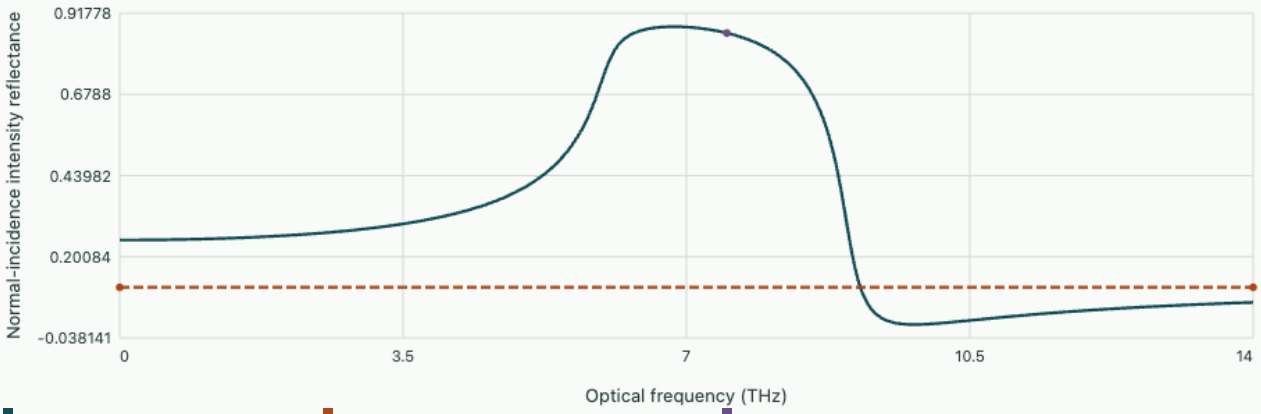
Open this experiment online

Try the calculation

Static relative permittivity **9**

Damping ratio γ/ωTO **0.05**

Probe frequency (THz) **7.5**



Damped ionic crystal Electronic background only Selected frequency

Lossless TO pole 6 THz and LO zero 9 THz. At 7.5 THz the chosen damping gives reflectance 0.8596.

Tutorial on recovering mode parameters

Optical data give $f_{\mathrm{TO}} = 8.0$ THz and $f_{\mathrm{LO}} = 10.0$ THz for a single weakly damped mode. With $\epsilon_\infty = 5.0$, estimate the static permittivity and dielectric increment.

▼ **Work through the solution**

$$\epsilon_s = 5(10/8)^2 = 7.8125.$$

The increment attributed to the optical mode is

$$7.8125 - 5 = 2.8125.$$

Angular and cyclic frequency ratios are equal, although an absolute oscillator strength S requires angular frequencies in the equation of motion.

At a fixed wavevector a probe sees displacement perpendicular to the wavevector. Why is the longitudinal restoring term absent in the electrostatic argument, and what additional coupling can still occur?

▼ **Work through the solution**

Its ionic polarisation has $\boldsymbol{k} \cdot \boldsymbol{P}_{\text{ion}} = 0$, hence it creates no longitudinal bulk bound charge $-i\boldsymbol{k} \cdot \boldsymbol{P}_{\text{ion}}$. There is no obligatory depolarising longitudinal field. A transverse propagating electromagnetic field can still drive the infrared-active vibration, as encoded in $\epsilon_r(\omega)$.

Further reading · MIT solid state physics, optical phonons and dielectric response.

Study the full derivation

Why a large joint density of states can leave a dark optical edge · Structure, excitations and probes of solids

A band diagram gives the energy an electron needs to move between bands. It does not by itself tell whether light couples those states. Construct a two-orbital model with an explicit velocity operator, calculate its allowed and suppressed transitions, and distinguish electronic band curvature from optical selection rules.

Chapter 20 · How one light isotope traps a lattice vibration

A perfect chain supports travelling lattice vibrations. Replace one atom by a lighter isotope while keeping every spring unchanged. Determine whether a vibration can now be confined to a few sites, calculate its frequency and decay length, and find out from a finite simulation when the word localised is justified.

What is permitted in the unperturbed crystal?

Let each atom of an infinite one-dimensional monatomic chain have mass M, nearest-neighbour spring constant K and equilibrium spacing a. A small longitudinal displacement u_j stretches the spring on the right by $u_{j+1} - u_j$ and that on the left by $u_j - u_{j-1}$. The restoring force is their difference multiplied by K. Newton's equation conserves total momentum in the absence of external forces and makes uniform translation an exact zero-frequency motion.

For a travelling wave $u_j(t) = Ue^{i(kja-\omega t)}$, the spatial difference contributes $e^{ika} + e^{-ika} - 2$. Thus

$$\omega^2(k) = \frac{4K}{M}\sin^2(ka/2).$$

This defines a continuous band from zero to $\omega_{\max} = 2\sqrt{K/M}$. At the upper edge adjacent sites move oppositely; for a fixed amplitude the springs experience their largest extension. A measured peak above that bound could not be a propagating plane wave in this model. It might indicate an incorrect force constant, another branch, an anharmonic response or a local structural change. We shall test one specific explanation using a spatial prediction as well as a spectral one.

The second difference still permits an evanescent displacement

Replace the centre atom by mass $m = qM$. All equations except that at $j = 0$ remain the same. A vibration spatially confined to the defect should decay in both directions without any boundary driver replenishing it. The upper-band-edge sign pattern suggests $u_j = U(-r)^{|j|}$, with $0 < r < 1$. The sign alternation is crucial: using a positive decaying exponential in a monatomic acoustic branch would give a negative ω^2 for a freely oscillating stable chain. The ideal symmetric mass defect couples to an even displacement about the origin. An antisymmetric displacement would vanish there and would not feel a change of its inertia.

For $j \neq 0$, substitute the ratio $u_{j+1}/u_j = -r$ on the right half chain into the equation of motion at frequency ω. At the centre the neighbours on either side both equal $-rU$. These two equations must determine the same frequency:

$$\frac{M\omega^2}{K} = 2 + r + \frac{1}{r},$$

$$\frac{m\omega^2}{K} = 2(1 + r).$$

Eliminating ω gives $q(1+r)^2 = 2r(1+r)$, or $r = q/(2-q)$ when the decaying solution exists. The physical restriction $0 < r < 1$ then requires $0 < q < 1$. The complete result is

$$\omega_{\mathrm{loc}} = \frac{2\sqrt{K/M}}{\sqrt{q(2-q)}},$$

$$\frac{\ell}{a} = -\frac{1}{\ln r}.$$

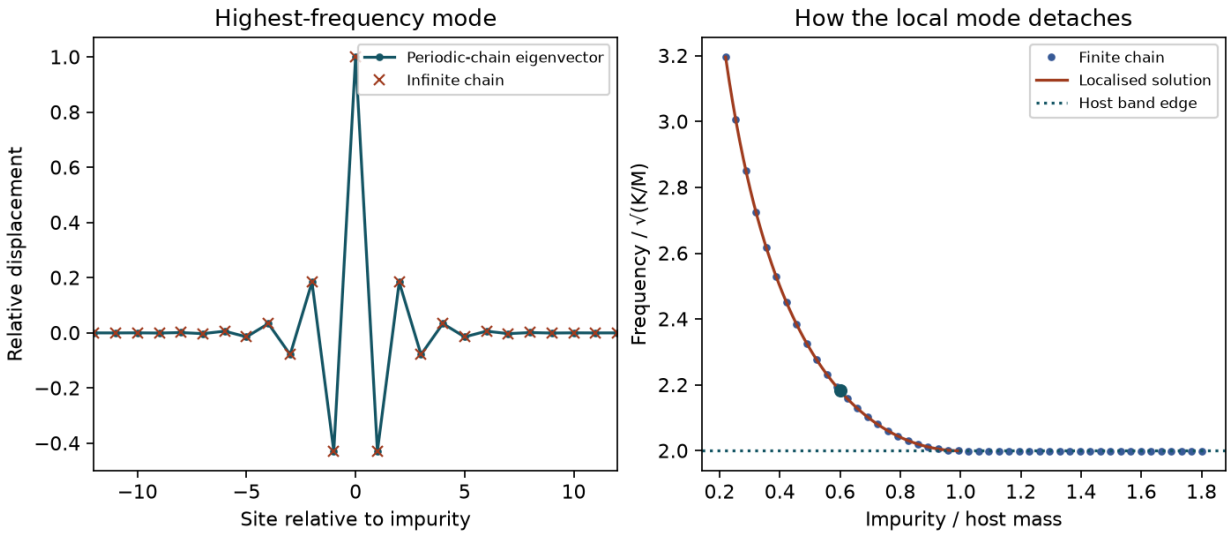
The second expression measures the decay length of displacement amplitude. The envelope of intensity or squared amplitude decays twice as fast and has half that length. A group velocity evaluated by inserting a real wave number into the ordinary band cannot describe energy transport away from this mode. Its spatial decay replaces a real propagation phase by a complex lattice wave number with real part π/a.

Open editable notebook Download .ipynb

Look at what happens when the defect barely differs from its host

Write $q = 1 - \eta$ with a small positive mass reduction η. The frequency ratio to the band edge is $1/\sqrt{1-\eta^2} \simeq 1 + \eta^2/2$. The displacement ratio is $(1-\eta)/(1+\eta)$, giving $\ell/a \simeq 1/(2\eta)$. Two consequences go in opposite directions. A small mass mismatch always gives a mathematical above-band solution in this infinite one-dimensional model, but its frequency is hard to separate from the band and its spatial extent becomes large. An instrument with finite spectral width or a simulation with few atoms may therefore see little evidence of confinement.

Take $q = 0.60$. Here $r = 3/7 \simeq 0.429$; after three sites the signed amplitude relative to the centre is about -0.0787. The decay length is about $1.18a$. In frequency units $\sqrt{K/M}$, the mode sits at 2.18218, clearly above the edge at 2. For $q = 0.98$, by contrast, the excess frequency is only about 0.020 per cent and the decay length is approximately $25a$. An 80-site ring in the latter case lets the two tails meet at a distance comparable to their decay length. Interpreting the finite eigenvalue as an isolated impurity in an infinite crystal requires a size check.



The left panel compares physical displacements normalised to the largest absolute value. The right panel tracks a frequency and contains no eigenvector information. Near the band edge, use both panels and repeat with a larger ring.

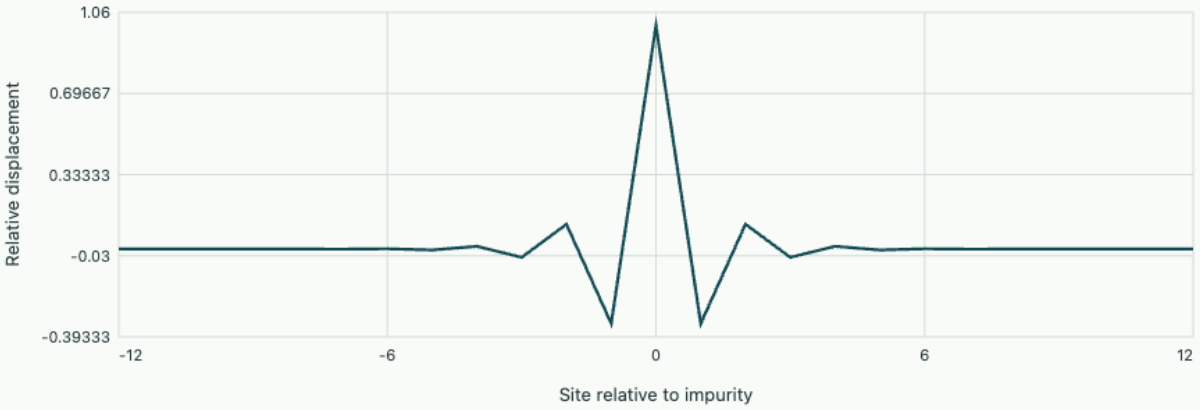
Try it · Impurity

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

Impurity / host mass **0.5**



Infinite-chain local mode

Frequency 2.3094 × √(K/M), above host band edge 2. Decay length 0.91024 sites. Near equal masses use the finite-ring notebook to test finite-size effects.

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

For an ideal isotope change the force constants are held fixed. A chemical substitution can change a bond's stiffness along with its mass, and a crystal with more than one atom per unit cell has additional bands. A heavy impurity can have other local or resonant behaviour in more general systems. The result here concerns an above-band normalisable mode of one mass defect in this harmonic acoustic chain. Simply plugging a value $q > 1$ into the algebraic frequency expression discards the decay condition and can give a spurious interpretation.

Construct the finite eigenproblem from kinetic and elastic energies

For a periodic ring write $\mathcal{M}$ for the positive diagonal matrix of masses and $\mathcal{K}$ for the cyclic second-difference stiffness. Elastic energy is $u^{\mathsf{T}}\mathcal{K}u/2$, the sum of positive spring energies. Kinetic energy is $\dot{u}^{\mathsf{T}}\mathcal{M}\dot{u}/2$. The variational equations give

$$\mathcal{K}u = \omega^2 \mathcal{M}u.$$

Directly forming $\mathcal{M}^{-1}\mathcal{K}$ yields a nonsymmetric matrix in ordinary Euclidean coordinates. It has the right eigenvalues mathematically, but feeding it to a symmetric eigensolver without transforming it would ask the solver to process a different matrix.

Set $w = \mathcal{M}^{1/2}u$. The equivalent symmetric eigenproblem has operator $\mathcal{M}^{-1/2}\mathcal{K}\mathcal{M}^{-1/2}$. An orthonormal eigenvector w corresponds to a physical displacement $u = \mathcal{M}^{-1/2}w$. Consequently a detector measuring displacement at one site and a map of kinetic participation do not plot the same quantity. At fixed mode frequency the mean kinetic-energy weight at site j is proportional to $m_j|u_j|^2$. The notch in the mass at the centre matters even after the spatial shape of displacement has been found.

The notebook forms the ring stiffness from two shifted identity matrices and uses an even site count so that the unperturbed zone edge is an allowed wave number. It reports both the frequency and the impurity's kinetic weight. Validate the result by substituting the recovered displacement into the original force balance at every site, including the two periodic endpoints. Check also the zero-frequency translation mode. A small defect residual and a correct band edge test different aspects of the assembly. The optional OneSimulation cell independently invokes its symmetric eigenproblem interface on the mass-weighted matrix and checks the residual of the highest mode.

Which experiment distinguishes localisation from a stiff crystal?

Suppose spectroscopy resolves an extra high-frequency line. Increasing every spring constant would shift the entire host band by the same square-root scaling. Changing a small set of isotopic masses makes an extra feature with large amplitude near those sites and leaves much of a large sample nearly unperturbed. A local displacement measurement moving away from the known defect therefore discriminates between two explanations that can be adjusted to match one peak frequency. The sign reversal is accessible to a phase-sensitive measurement; an intensity-only experiment determines the envelope but cannot directly establish that successive displacements are out of phase.

One can estimate the weight concentrated at the impurity analytically. With central displacement one, the mass-weighted norm divided by M is $q + 2\sum_{j\ge 1} r^{2j} = q + 2r^2/(1-r^2)$. The centre fraction is its first term divided by this sum. Even a pronounced peak need not contain all the energy, and changing the arbitrary normalisation of an eigenvector cannot change such a fraction. This gives a second scalar check more sensitive to errors in the eigenvector than the frequency alone.

Tutorial on localisation and finite rings

A measured displacement envelope falls by a factor of four over two lattice spacings. Within the one-defect model, infer the impurity mass and the frequency relative to the band edge.

▼ **Work through the solution**

Two spacings multiply amplitude by $r^2 = 1/4$, hence $r = 1/2$. The inverse formula gives

$$q = 2r/(1+r) = 2/3.$$

The frequency ratio is

$$1/\sqrt{(2/3)(4/3)} = 3/\sqrt{8} \simeq 1.061.$$

If the reported quantity were intensity rather than amplitude the inference would change, so the physical meaning of the recorded ordinate belongs with the number.

For a mass ratio of 0.6, compute the fraction of time-averaged kinetic energy at the impurity. Why is it different from a simple count of squared displacement?

▼ **Work through the solution**

With $r = 3/7$, the sum over both sides is

$$2r^2/(1 - r^2) = 0.45.$$

The denominator of the mass-weighted norm is $0.6 + 0.45 = 1.05$, giving a centre fraction of $0.6/1.05 = 4/7$. Counting squared displacements alone would give $1/1.45$, which overestimates the central kinetic contribution because that atom carries less mass.

The measured light defect has $q = 0.98$. Estimate a ring size for which the amplitude of one infinite-chain tail at the antipodal site is below one per cent of its central value.

▼ **Work through the solution**

Require $\exp[-Na/(2\ell)] < 0.01$, hence $N > 2(\ell/a)\ln 100$. With about 25 spacings per decay length this gives a little over 230 sites. This is a geometric tail criterion, not an error bound on every eigenvalue. For higher accuracy compare two larger rings and test the residual against the infinite-chain formula.

Further reading · Qiang Luo · Local vibrational mode of an impurity in a monatomic chain with free and periodic boundaries

Problem sheet 7 · Inferring a chain termination from bulk bands and end-site data

A fabricated one-dimensional array has two sites per primitive cell, real positive nearest-neighbour hoppings and a reproducible way of cleaving its left edge. Use energy and intensity measurements to decide which link is exposed. Check which conclusions survive when the array contains only a few cells.

Extract the unordered bonds from a bulk energy scan

The upper band has energy $1.60\,\mathrm{eV}$ at $k = 0$ and $0.40\,\mathrm{eV}$ at $ka = \pi$, measured from the symmetry centre. Find the unordered hopping strengths, the direct gap and the energy at $ka = \pi/2$. Explain what the bulk measurements leave ambiguous.

▼ **Work through the solution**

The two extrema give $t_1 + t_2 = 1.60$ and $|t_1 - t_2| = 0.40$ eV. The positive hoppings are 1.00 and 0.60 eV. The full direct gap is 0.80 eV and the intermediate energy is

$$\sqrt{1^2 + 0.6^2} = 1.166\,\mathrm{eV}.$$

The characteristic equation is unchanged when the hoppings are exchanged, so even a perfect positive-energy measurement over the full Brillouin zone cannot label the first physical bond.

Swap the two bonds and compare the two bulk energy scans

Use a boundary measurement to label the bonds

The leftmost orbital is A_1. In an odd chain ending on an A site, local spectroscopy near zero energy gives relative intensities $1.00, 0.36, 0.130$ on the first three A sites. Determine t_1/t_2, label the hoppings and predict the first intensity outside the measured set.

▼ **Work through the solution**

At zero energy the neighbouring B-site equations impose $A_{j+1} = -(t_1/t_2)A_j$. Intensities fall by r^2, giving $r = 0.60$. Therefore the first physical link A_1B_1 is the weaker 0.60 eV hopping. The fourth A site has relative intensity $0.36^3 = 0.04666$. The signs of successive amplitudes alternate although the intensities carry no sign.

Check the odd-chain end state, including its finite normalisation

Estimate finite-size contamination

Take $M = 8$ A sites and seven B sites. What fraction of the normalised state lies at A_1? How much is at A_8? Repeat those endpoint predictions after exchanging the intra-cell and intercell hoppings while

preserving the cut.

▼ **Work through the solution**

With ratio $r = 0.60$, finite normalisation gives

$$P_1 = (1 - r^2)/(1 - r^{2M}) = 0.64018.$$

Eight A sites contain seven steps from one end to the other. The far-end probability is

$$P_8 = P_1 r^{14} = 0.000502.$$

Replacing r by $1/r$ reflects the normalised geometric sequence. The two endpoint probabilities exchange, while the bulk energy spectra from stage one coincide. In a short sample near $r = 1$, finite-size effects would become much larger.

Test whether a zero-energy claim depends on the model

Add on-site energy $+\Delta$ to every A site and $-\Delta$ to every B site, with $\Delta = 0.10\,\mathrm{eV}$. Given the original odd-chain state has amplitude only on A, find its new energy. Obtain the new bulk gap and say which observables still identify the weaker boundary bond.

▼ **Work through the solution**

The diagonal perturbation acts as $+\Delta$ times the original state. The off-diagonal Hamiltonian still annihilates that state because the B-site cancellations survive. Its energy consequently becomes $+0.10\,\mathrm{eV}$ with unchanged profile. The bulk edges move to $\pm\sqrt{(t_1 - t_2)^2 + \Delta^2}$, giving a gap $2\sqrt{0.4^2 + 0.1^2} = 0.825\,\mathrm{eV}$. Exact zero energy depended on the off-diagonal symmetry. The measured decay ratio and the specified termination still identify the hopping order for this uniform sublattice shift.

Problem sheet 8 · Why a crystal with three sound speeds has one low-temperature heat coefficient

Calorimetry measures a slope in a plot of heat capacity divided by temperature against temperature squared. Determine what it says about the elastic modes of a three-dimensional crystal and how an unnoticed electronic term changes an estimate of the sound speed.

Count states of one acoustic branch

Take an isotropic, periodic cube of volume V. Each allowed wavevector occupies reciprocal-space volume $(2\pi)^3/V$. Derive the number of acoustic modes per unit volume and per unit angular frequency for a branch of speed v_s. Decide whether an extra factor of two for waves travelling in opposite directions is appropriate.

▼ **Work through the solution**

The number of vectors in a shell equals $4\pi k^2\,dk$ divided by the reciprocal-space volume per state. The shell already includes opposite wavevectors. With $\omega = v_s k$ the result for a single polarisation is

$$g_s(\omega)/V = \omega^2/(2\pi^2 v_s^3).$$

There are two transverse polarisations and one longitudinal polarisation in the isotropic solid. The weight of a slow transverse branch is large because many wavevectors can carry the same small energy.

Obtain the calorimetric coefficient without a fitted cutoff

At temperatures far below the acoustic cutoff, change variables to $x = \hbar\omega/(k_B T)$ and replace the upper limit with infinity. Derive c_{lat}/T^3, where lower-case c is heat capacity per volume. Why can the zero-point energy be omitted here?

▼ **Work through the solution**

Thermal energy per mode is $\hbar\omega/(e^x-1)$; its zero-point contribution is independent of temperature at fixed spectrum. Summing over the three branches and using $\int_0^\infty x^3/(e^x-1)\,dx = \pi^4/15$ gives energy density proportional to T^4. Differentiation produces

$$\frac{c_{\mathrm{lat}}}{T^3} = \beta = \frac{2\pi^2 k_B^4}{15\hbar^3}\left(\frac{1}{v_L^3}+\frac{2}{v_T^3}\right).$$

The reciprocal cubes must be summed before taking a mean. The effective speed defined by $3/v_D^3 = 1/v_L^3 + 2/v_T^3$ has units of velocity. With $v_L = 6000$ and $v_T = 3000\,\mathrm{m\,s^{-1}}$ it is about $3365\,\mathrm{m\,s^{-1}}$. A direct arithmetic average would weight spectral phase space incorrectly.

Extract the lattice and electronic parts from two measurements

An isotropic metal has $c/T = 8.0\,\mathrm{J\,m^{-3}K^{-2}}$ at 2 K and 20.0 in the same units at 4 K. Assuming $c = \gamma T + \beta T^3$, determine γ, β. Predict c at 3 K and give a clear experimental sign of a missing contribution.

▼ **Work through the solution**

Fit the straight line $c/T = \gamma + \beta T^2$. The slope is $(20 - 8)/(16 - 4) = 1.0\,\mathrm{J\,m^{-3}K^{-4}}$ and its intercept is $4.0\,\mathrm{J\,m^{-3}K^{-2}}$. At 3 K predict

$$3(4 + 9) = 39\,\mathrm{J\,m^{-3}K^{-1}}.$$

A reproducibly curved plot, or a slope depending systematically on the selected temperature window, invalidates at least one assumption. Magnetic excitations, a low optical branch or a thermal background can each introduce curvature. Two measurements alone cannot discriminate those possibilities.

Decide which sound measurement is worth improving

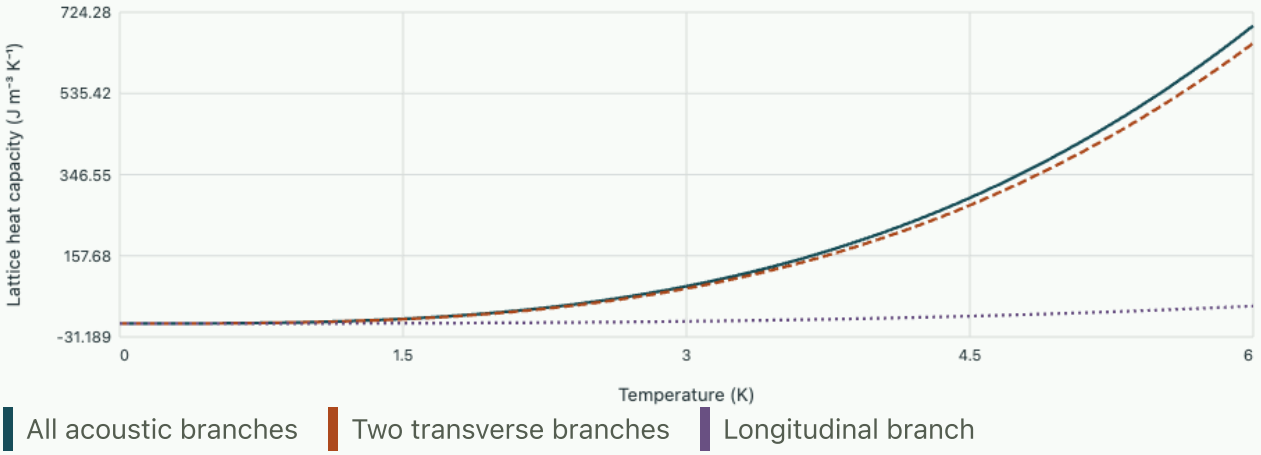
Change one sound speed at a time. Inspect the separate longitudinal and transverse contributions before inspecting their sum. If the transverse speed has a small relative bias, how does that bias propagate into the total lattice coefficient? How could an independent longitudinal measurement help?

Open this experiment online

Try the calculation

Longitudinal sound speed (m/s) **6000**

Transverse sound speed (m/s) **3000**



β = 3.2088 J m⁻³ K⁻⁴. Transverse branches supply 94.118%.

▼ **Work through the solution**

Differentiating v^{-3} gives a fractional contribution change of $-3\,dv/v$. At the example speeds the transverse fraction of β is $16/17$. A positive one-percent bias in v_T biases the total coefficient low by about 2.82% at first order. A positive one-percent bias in v_L changes it by only 0.18%. One calorimetric slope can identify a weighted inverse cube, not two independent velocities. A measured v_L makes it possible to infer v_T, provided that isotropy, the two-branch degeneracy and the temperature window are justified.

The independent mode-counting construction in the Cambridge solid-state course connects the acoustic approximation to a full lattice dispersion.

Notebook practice · How one light isotope traps a lattice vibration

Structure, excitations and probes of solids

Connect microscopic symmetry and dynamics to a measurement of a real solid. Develop the derivation and identify which experimental assumptions are needed to interpret the signal.

1. ## Which reciprocal-lattice points produce a diffraction spot?

   A bright spot depends on the wavelength and orientation of an incident beam, on the periodic repeat of a crystal, and on the atoms inside one repeat. Separate these three conditions so that a missing reflection becomes an interpretable observation. Then use the relative strength of two spots to estimate chemical order.

2. ## Read atomic motion from Bragg attenuation and diffuse scattering

   A crystal can retain its periodic mean positions while its diffraction pattern changes markedly with temperature. Start with the thermal average of a scattering amplitude, derive the displacement factor and learn which component of atomic motion a given reflection measures.

3. ## Construct allowed bands from a cell transfer matrix

   In a finite well a bound level follows from matching its wavefunction at the two walls. A periodic lattice replaces those two external walls by a matching condition on every cell. Derive a band criterion that works without guessing a wave inside the forbidden gap and use it to predict transmission through a finite sequence of cells.

4. ## Why a large joint density of states can leave a dark optical edge

   A band diagram gives the energy an electron needs to move between bands. It does not by itself tell whether light couples those states. Construct a two-orbital model with an explicit velocity operator, calculate its allowed and suppressed transitions, and distinguish electronic band curvature from optical selection rules.

5. ## Count vacancies and infer an atomic migration barrier

   Tracer atoms spread through a crystal even when the average lattice positions remain fixed. The rate can depend both on the number of vacant sites and on how readily an atom jumps into one. Derive the two factors separately, then compare an equilibrated specimen with a rapidly quenched one.

6. ## Measure a phonon mean free path without confusing it with a sample length

   A cold insulating bar can carry heat primarily through lattice vibrations. Heat capacity tells how much energy the vibrations can carry; conductivity also depends on how far and in which direction they carry it. Derive the connection, test its limiting regimes, and identify when a measured thermal resistance belongs to a boundary.

7. ## A field profile that distinguishes a superconductor from a perfect conductor

Zero dc resistance tells how a current can persist. Field expulsion tells how a specimen responds when it enters a superconducting equilibrium state. Derive the London length, calculate what a finite slab and a thin film actually screen, and connect magnetic and electrical measurements to the same superfluid density.

8. Infer a superconducting gap from a tunnelling spectrum

A tunnel junction reports differential conductance versus voltage. Extracting a superconducting energy gap requires connecting a condensate's quasiparticle dispersion to a density of states, then accounting for temperature and finite spectral resolution before comparing peaks with the data.

9. Read twist and mismatch from the moiré of two lattices

Stack two square grids that differ slightly in orientation or pitch. The slowly varying registry map is visible over many unit cells. Derive its scale and direction in reciprocal space, and find which combinations of layer parameters a single period measurement can resolve.

10. From bonding energy to lattice spacing and sound speeds

Diffraction gives a spacing and ultrasound gives a stiffness. One measures where an energy curve is stationary, the other measures its curvature. Building the simplest energy of an ionic solid gives both observables and shows which quantities a spring model has already assumed.

11. Collective charge oscillations, bound excitons and lattice-dressed electrons

A loss peak in an electron spectrum, a line below an optical gap and an enhanced carrier mass call for different kinds of excitations. Start each model with its moving charges and its restoring interaction. That identifies what a measurement can determine.

12. Ferromagnetic resonance from angular momentum and free-energy curvature

A uniformly magnetised thin specimen can absorb microwave power at a frequency far from the bare spin Larmor frequency. The precessing magnetic moment creates its own field. Its geometry and the anisotropy determine which restoring torques appear in the measured resonance.

13. Lattice dislocations and structure without long-range order

A missing atom, a slipped lattice plane and a glass can all broaden a measurement, through very different mechanisms. Geometry distinguishes a line defect from a local vacancy. Correlation functions make the distinction between short-range order and crystallinity quantitative.

14. Why some alloys demix while others order on two sublattices

A binary alloy can cease to look random by separating into compositionally distinct regions or by arranging unlike atoms on alternating sites. A free energy for composition answers the first question. A second order parameter is needed for the second.

15. What extremal Fermi-surface orbits reveal about a metal

A single magnetic oscillation frequency gives one area in reciprocal space. Many different surfaces can share that area. Rotating the field and following how quantum oscillations and conductivity change turns the measurement into a three-dimensional inference problem.

Structure, excitations and probes of solids · 1. Which reciprocal-lattice points produce a diffraction spot?

A bright spot depends on the wavelength and orientation of an incident beam, on the periodic repeat of a crystal, and on the atoms inside one repeat. Separate these three conditions so that a missing reflection becomes an interpretable observation. Then use the relative strength of two spots to estimate chemical order.

Add amplitudes before squaring

Let an incident plane wave have wavevector $\mathbf{k}_i$ and the detected far-field direction select $\mathbf{k}_f$. Elastic scattering gives

$$|\mathbf{k}_i| = |\mathbf{k}_f| = 2\pi/\lambda.$$

In the single-scattering approximation a point at $\mathbf{r}$ contributes phase $(\mathbf{k}_i - \mathbf{k}_f)\cdot\mathbf{r}$. Define $\mathbf{Q} = \mathbf{k}_f - \mathbf{k}_i$; reversing the sign of all phases leaves the usual real-scatterer intensity invariant. With site positions $\mathbf{R} + \mathbf{d}_j$ and atomic scattering factors f_j, the total amplitude factors into a lattice sum over $\mathbf{R}$ and a basis structure factor $F(\mathbf{Q}) = \sum_j f_j(\mathbf{Q}) e^{-i\mathbf{Q}\cdot\mathbf{d}_j}$. Intensity is proportional to the squared magnitude of their product.

In a row of N sites of spacing a the lattice sum is a finite geometric series. Its intensity factor is $\sin^2(NQ_x a/2)/\sin^2(Q_x a/2)$, with value N^2 at an exact reciprocal point. The first zeros lie $2\pi/(Na)$ away from the maximum, so increasing the illuminated coherent length sharpens the spots. The area under an individual one-dimensional peak grows roughly linearly with N; the increase of peak height is accompanied by decreasing width. Comparing the height of an unresolved peak with an integrated intensity from another instrument can confuse finite domain size with a structural change.

Impose the Ewald condition and identify systematic absences

A reciprocal vector $\mathbf{G}$ produces a geometric reflection if $\mathbf{k}_f = \mathbf{k}_i + \mathbf{G}$ also lies on the elastic sphere. Squaring gives

$$2\mathbf{k}_i \cdot \mathbf{G} + G^2 = 0.$$

The reciprocal lattice point must lie on a sphere of radius k centred at $-\mathbf{k}_i$ in reciprocal coordinates. Rotating the crystal changes which points cross the sphere without changing their lengths. For plane spacing $d = 2\pi/G$, this condition gives

$$2d\sin\theta = \lambda,$$

with 2θ the angle between incident and outgoing beams. Geometric access requires $G \le 2k$; extending detector angle cannot violate the wavelength limit.

Now take a simple cubic repeat of side a with one site at $(0,0,0)$ and a second at $(a/2, a/2, a/2)$. For index (h,k,l),

$$F = f_A + (-1)^{h+k+l} f_B.$$

Even sums add, odd sums subtract. If the scatterers are identical the odd family vanishes: the apparent cubic description represented a body-centred Bravais lattice. If A and B are chemically distinct a corresponding odd spot can be visible. Its small intensity need not indicate a small lattice or poor alignment. The structural question is why a reciprocal point that satisfies the Ewald condition has small basis amplitude.

For $a = 0.400\,\mathrm{nm}$ and $\lambda = 0.154\,\mathrm{nm}$, a powder permits the (100) and (110) orientations. Their predicted 2θ positions are about 22.2° and 31.6°. If one replaces the scattering factors by constants $f_A = 28$, $f_B = 14$ for an illustrative comparison, odd-to-even squared amplitudes give

$$14^2/42^2 = 1/9.$$

The raw integrated powder peaks need not obey that ratio: multiplicity, angular geometry, polarisation, absorption, thermal attenuation and Q variation of atomic form factors also enter. Use corrected structure factors or compare symmetry-related measurements controlling those factors before converting a ratio into atomic contrast.

Infer partial chemical order and distinguish noise from disorder

Let half the atoms be species A and half B. On sublattice one, let probabilities for A, B be $(1+S)/2, (1-S)/2$; reverse them on sublattice two. $S = 1$ denotes complete order in the chosen assignment, $S = 0$ random mean occupancies. The average factors are $\bar{f}_1 = (f_A + f_B)/2 + S(f_A - f_B)/2$ and $\bar{f}_2 = (f_A + f_B)/2 - S(f_A - f_B)/2$. Thus the coherent odd Bragg amplitude is $S(f_A - f_B)$, while the even amplitude stays $f_A + f_B$. For equal domains and corrected geometry the ratio falls as S^2. Disorder may generate diffuse intensity away from Bragg peaks; its angular correlations contain information absent from the average occupancies.

If the corrected odd-to-even intensity ratio is 0.0400 for the stated $28 : 14$ contrast, $|S| = 0.600$. The magnitude alone cannot distinguish exchange of the two sublattices: S and $-S$ differ by an origin choice in this ideal case and lead to the same intensity. If relative intensity uncertainty is ten percent, the relative uncertainty of $|S|$ from that measurement is about five percent before scattering-factor and correction uncertainty. Near $S = 0$ a background subtraction or weak multiple scattering can dominate the inferred value. Report an upper limit when a corrected odd peak is consistent with the background.

Predict spots before changing the wavelength

The live calculation holds the cubic spacing and two scattering factors fixed, varies order and wavelength, and shows the resulting allowed powder-line positions and approximate strengths. The curves use a common artificial instrumental broadening to make discrete lines visible. Set $S = 0$ and identify which set survives. Restore order and alter wavelength; visible lines move because the Ewald geometry changes, whereas the parity rule itself remains. At a large wavelength some high-index lines

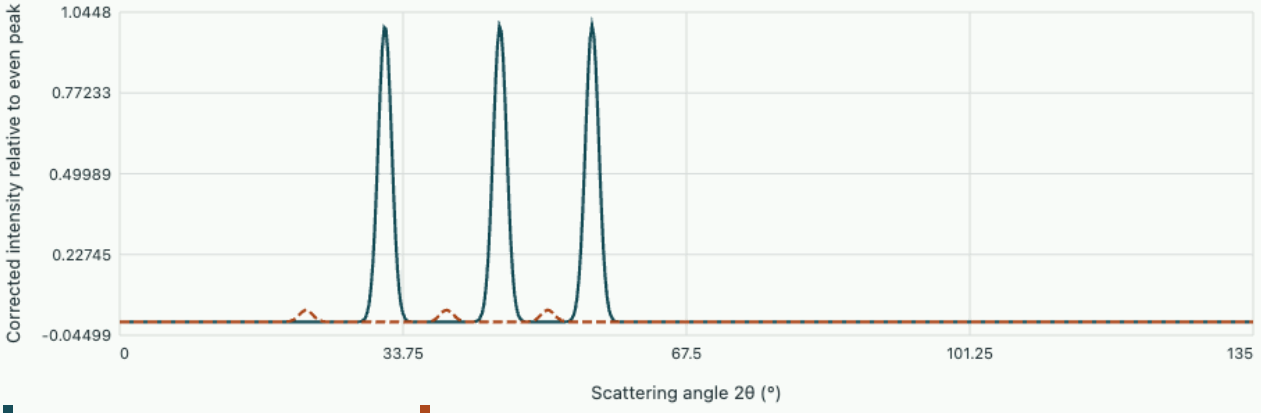
leave the accessible angular interval entirely. A disappearing peak caused by this scan is physically distinct from a basis extinction at the same index.

Open this experiment online

Try the calculation

Chemical order S **0.6**

Vacuum wavelength (nm) **0.154**



Even sum (selected families) Odd sum (selected families)

The corrected odd/even line ratio is 0.0400. The (100) and (110) geometric positions are 22.20° and 31.59°. Multiplicity and instrument response are excluded.

Compare a single-crystal rocking scan with a powder ring

An oriented single crystal at fixed wavelength generally places only a few reciprocal points near the Ewald sphere. Rotate it by a small angle about an axis transverse to the scattering plane and record intensity as one selected point sweeps through the sphere. Finite beam divergence, wavelength spread, mosaic block orientations and finite sample dimensions all broaden this rocking curve. A powder already contains crystallites with many orientations, so a matched spacing creates a ring; its integrated population is determined by the orientation distribution and multiplicity. A texture that aligns particular crystallites can strengthen or weaken one powder family without changing the underlying atomic structure factor.

An instructive control compares two samples cut from the same single crystal, one thick and one thin in the direction of the measured reciprocal vector. If the smaller coherent size is the main broadening mechanism, the reciprocal width should grow approximately inversely with thickness while the centre remains fixed. A distribution of lattice spacing produces a relative radial width that often increases with the magnitude of $\mathbf{G}$. Map both rocking and radial scans where possible, and measure an instrument standard. These independent directions distinguish a spread of orientations from a spread of spacings more effectively than fitting one width with an assumed coherent-domain formula.

Tutorial on geometric and structural tests

A corrected (100) intensity falls by a factor four after heating, while its corrected (110) intensity stays unchanged in the basic order model. What follows and what independent measurement should accompany it?

▼ **Work through the solution**

The absolute value of the order parameter halves because the odd line scales as S^2. Check more than one odd and even line and measure their angular widths and positions. Thermal motion generally attenuates high-Q intensities and a phase transition may change cell dimensions or domain size. A single ratio alone cannot separate every one of those effects.

For a cubic lattice $a = 0.400\,\mathrm{nm}$, can a wavelength $0.300\,\mathrm{nm}$ access the $(3, 0, 0)$ reflection by rotating the sample?

▼ **Work through the solution**

Its plane spacing is $0.1333\,\mathrm{nm}$, so even at backscattering $2d = 0.2667\,\mathrm{nm} < \lambda$. No rotation satisfies the elastic Ewald condition. Changing composition to increase the odd structure factor will not create a reflection at this wavelength. Shorten the wavelength or select a longer-spacing family.

Two samples have the same peak positions; one has a broader but lower central maximum. Does that establish a different lattice spacing?

▼ **Work through the solution**

Peak-centre position is what fixes spacing for a known wavelength. A finite coherent domain broadens peaks with width of order the inverse domain size. Microstrain or instrumental broadening can also broaden peaks. Measure the integrated line shape, a standard of known width, and several reciprocal vectors. Distinct causes of reduced peak height can then be separated from a change in mean lattice spacing.

David Tong derives real and reciprocal lattices in his Cambridge band-structure course.

Structure, excitations and probes of solids · 2. Read atomic motion from Bragg attenuation and diffuse scattering

A crystal can retain its periodic mean positions while its diffraction pattern changes markedly with temperature. Start with the thermal average of a scattering amplitude, derive the displacement factor and learn which component of atomic motion a given reflection measures.

Average a phase associated with a fluctuating position

Let an atom have position $\mathbf{R} + \mathbf{u}$, where $\mathbf{R}$ is its mean position and $\langle \mathbf{u} \rangle = 0$. At momentum transfer $\mathbf{Q}$, the coherent elastic amplitude is proportional to $e^{i\mathbf{Q}\cdot\mathbf{R}}\langle e^{i\mathbf{Q}\cdot\mathbf{u}} \rangle$. For a Gaussian displacement distribution with covariance $U_{ij} = \langle u_i u_j \rangle$, the scalar $X = \mathbf{Q}\cdot\mathbf{u}$ is Gaussian with zero mean and variance $Q_i U_{ij} Q_j$. Completing the square in its characteristic-function integral gives

$$\langle e^{iX} \rangle = e^{-\langle X^2 \rangle / 2}.$$

Thus the amplitude is multiplied by $e^{-Q_i U_{ij} Q_j / 2}$; its coherent Bragg intensity receives the factor $e^{-Q_i U_{ij} Q_j}$ for equivalent atoms. Writing amplitude and intensity separately avoids the common factor of two error.

For isotropic motion the mean square displacement along one specified axis is U, with total three-dimensional mean square displacement $3U$. The intensity factor is $e^{-Q^2 U}$, not $e^{-3Q^2 U}$. Crystallographic displacement notation often uses $B_{\mathrm{iso}} = 8\pi^2 U$, a quantity of dimension length squared unrelated to magnetic induction. With $Q = 4\pi \sin\theta / \lambda$, the amplitude factor can be written $e^{-B_{\mathrm{iso}} \sin^2\theta / \lambda^2}$. Different sources quote the exponent for amplitude or intensity, so determine which observable is being fitted before comparing reported displacement parameters.

Derive the thermal motion of one harmonic mode

For a one-dimensional harmonic displacement of mass M and angular frequency ω, quantisation gives

$$\hat{u} = \sqrt{\hbar / (2M\omega)}\,(a + a^\dagger).$$

In a number state the terms a^2 and $(a^\dagger)^2$ have zero expectation, while $a a^\dagger + a^\dagger a = 2n + 1$. Thermal averaging with $\bar{n} = (e^{\hbar\omega / k_B T} - 1)^{-1}$ yields

$$\langle u^2 \rangle = \frac{\hbar}{2M\omega} \coth \frac{\hbar\omega}{2 k_B T}.$$

At low temperature the mode retains a zero-point contribution independent of temperature. At high temperature $\coth x \simeq 1/x$ gives $k_B T / (M\omega^2)$, the equipartition result for the quadratic potential. Light atoms and soft modes can therefore show substantial positional fluctuations. A real crystal contains modes with different polarisations, frequencies and atomic amplitudes; a single fitted Einstein frequency is a model of their weighted contribution, not automatically the frequency of one identifiable phonon.

For illustrative $M = 40$ atomic mass units and ordinary frequency $f = 5.0\,\mathrm{THz}$, $\hbar\omega / k_B$ is about $240\,\mathrm{K}$. The one-mode zero-point variance is about $2.53 \times 10^{-23}\,\mathrm{m}^2 = 0.00253\,\text{Å}^2$. At $300\,\mathrm{K}$ the multiplier

$\coth(0.40)$ is about 2.63, so the rms displacement grows from $0.0503\ \text{Å}$ to $0.0816\ \text{Å}$. At $Q = 8.0\ \text{Å}^{-1}$ the corresponding intensity factors are about 0.850 and 0.654. Atomic motion still much smaller than an interatomic spacing can therefore noticeably attenuate a high-angle peak.

Infer a covariance change and interpret missing intensity

For the same reflection at two temperatures, after scale, form factor, absorption, structural order and other changes have been corrected,

$$\ln[I_2/I_1] = -Q_i \Delta U_{ij} Q_j.$$

In an isotropic crystal plot this logarithm against Q^2. The slope estimates negative ΔU; a nonzero intercept can indicate a relative monitor or normalisation change. With $I_2/I_1 = 0.80$ at $Q = 6.0\ \text{Å}^{-1}$,

$$\Delta U = 0.00620\ \text{Å}^2.$$

This determines a variance difference, not the rms at either temperature individually. Simply taking the square root of a variance difference and calling it the rms increase generally gives the wrong quantity.

For anisotropic fluctuations the measurement probes a quadratic form. Reflections along three crystallographic axes constrain the diagonal elements, while an off-axis vector also depends on cross terms such as $2Q_xQ_yU_{xy}$. A full covariance must be positive semidefinite. A refinement returning a negative directional variance may indicate insufficient angular coverage, a wrong model or strongly correlated fitting parameters. Symmetry of site occupation can reduce the number of independent covariance components; an analysis with more unconstrained parameters than informative reflections risks interpreting noise as a preferred vibration axis.

Diffuse scattering clarifies what has happened to the lost coherent intensity. For a given configuration the measured intensity is the squared sum of amplitudes. Its thermal expectation can be split into the squared mean amplitude plus the variance of fluctuations about that mean. Correlated displacements between atoms contribute structured diffuse intensity, especially near Bragg positions. An instrument integrating a wide region around a peak can inadvertently include temperature-dependent diffuse background and underestimate the attenuation of the coherent peak. Energy-resolved scattering can separate elastic peaks from inelastic phonon scattering more directly than a simple static intensity ratio.

Reconstruct an off-axis displacement correlation

Suppose calibrated temperature-difference measurements in an orthogonal x, y plane give

$$\Delta U_{xx} = 0.0040\ \text{Å}^2,$$

$$\Delta U_{yy} = 0.0080\ \text{Å}^2.$$

Along a direction halfway between the axes the inferred projection variance is $0.0075\ \text{Å}^2$. For $\hat{n} = (1,1)/\sqrt{2}$, the directional variance is $(\Delta U_{xx} + \Delta U_{yy})/2 + \Delta U_{xy}$. Thus the cross covariance is

$+0.0015\ \text{Å}^2$. The direction $(1,-1)/\sqrt{2}$ should give $0.0045\ \text{Å}^2$. A scan of both diagonals cancels the trace term in their difference and isolates twice the off-diagonal covariance directly.

At $Q = 8\ \text{Å}^{-1}$ these two diagonal directions predict hot-to-cold ratios $e^{-64(0.0075)} = 0.619$ and $e^{-64(0.0045)} = 0.750$. An isotropic fit to their average would conceal this significant angular difference. If only one orientation is measurable due to sample mounting, symmetry constraints or a second mounting angle can supply missing information. The two eigenvalues of the illustrative covariance difference are 0.0035 and $0.0085\ \text{Å}^2$, both positive. In general the absolute covariance at every temperature must be positive semidefinite, while a difference of two positive covariances need not be positive along every axis. A structural change on heating can reduce fluctuations along one direction while increasing them along another.

Compare a temperature difference with an absolute structure refinement

An unknown constant site displacement and an unknown absolute intensity scale make a single-temperature fit less direct than a temperature ratio. On heating, measure both a small-Q reflection and a high-Q reflection in rapid alternation with a fixed incident monitor. Slow detector gain variation then affects the two less differently. If both corrected ratios fall equally, a scale change is more plausible than an added Gaussian displacement factor, which predicts a stronger effect at larger Q^2. Repeating over several independent directions tests that conclusion against anisotropic motion.

Tutorial on a temperature scan

Two equivalent high-symmetry reflections have $Q = 4$ and $8\ \text{Å}^{-1}$. Their corrected hot-to-cold ratios are 0.92 and 0.72. Are these consistent with a common isotropic variance change?

▼ **Work through the solution**

Individually they give

$$-\ln 0.92/16 = 0.00521,$$

$$-\ln 0.72/64 = 0.00513\ \text{Å}^2.$$

The close results are consistent within modest intensity uncertainties. Agreement across two points does not establish isotropy in all directions; add noncollinear reflections and check the fitted scale offset.

A fit reports an rms displacement of $0.10\ \text{Å}$ along each of three independent axes. What is the expected coherent intensity fraction at $Q = 10\ \text{Å}^{-1}$, and what is the total rms distance from the mean site?

▼ **Work through the solution**

Each directional variance is $0.010\ \text{Å}^2$. The fraction is

$$e^{-1} = 0.368.$$

The rms length of the three-dimensional displacement vector is

$$\sqrt{0.030} = 0.173\ \text{Å}.$$

The projection onto the scattering vector uses one directional variance, while total rms length sums all three.

An isotropic displacement fit stays large at the lowest measured temperature. What measurements distinguish static disorder from zero-point motion?

▼ Work through the solution

A constant displacement contribution can arise from both an unresolved distribution of mean site positions and quantum fluctuations. Extend the temperature range, compare isotopes or mode frequencies where feasible, and measure diffuse scattering with energy resolution. A harmonic calculation with independently measured phonon frequencies predicts a zero-point scale. Absolute Bragg attenuation alone gives a distribution width without uniquely identifying its microscopic source.

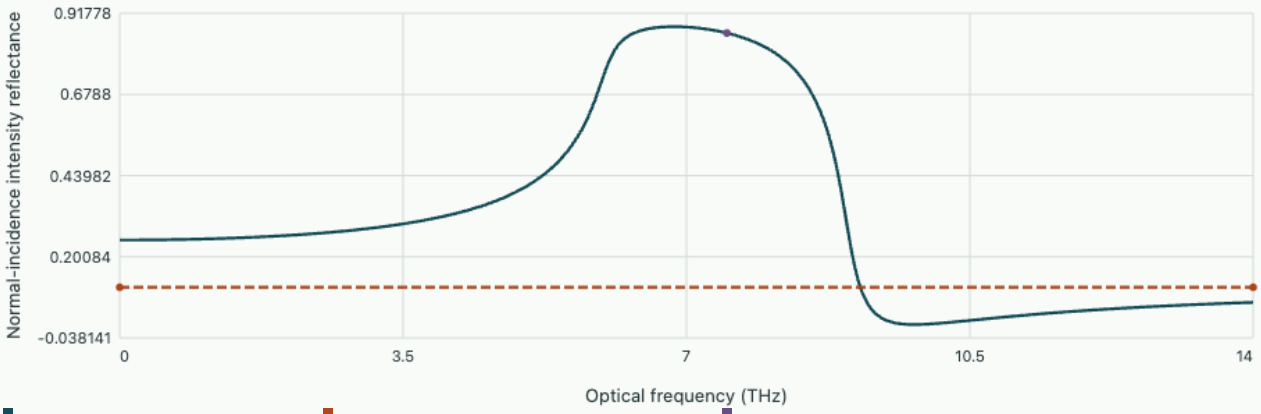
The International Union of Crystallography examines the displacement factors entering a diffraction intensity.

Compare the dynamics of a polar lattice mode with a static displacement measurement.

Open this experiment online

Try the calculation

Static relative permittivity	9
Damping ratio γ/ωTO	0.05
Probe frequency (THz)	7.5



Damped ionic crystal | Electronic background only | Selected frequency

Lossless TO pole 6 THz and LO zero 9 THz. At 7.5 THz the chosen damping gives reflectance 0.8596.

Structure, excitations and probes of solids · 3. Construct allowed bands from a cell transfer matrix

In a finite well a bound level follows from matching its wavefunction at the two walls. A periodic lattice replaces those two external walls by a matching condition on every cell. Derive a band criterion that works without guessing a wave inside the forbidden gap and use it to predict transmission through a finite sequence of cells.

Transfer a wave and its derivative through one interval

Consider a particle of mass m in a one-dimensional array of repulsive delta barriers $V(x) = g\sum_n \delta(x - na)$, with $g > 0$ of units energy times length. Between barriers $E > 0$, so the Schrödinger equation becomes $\psi'' + k^2\psi = 0$, $k = \sqrt{2mE}/\hbar$. A solution over one empty interval maps the column (ψ, ψ') by a matrix with rows $(\cos ka, \sin ka/k)$ and $(-k\sin ka, \cos ka)$. Both value and derivative propagate; their units differ, as the factors of k show.

Integrate the Schrödinger equation across one narrow interval containing a delta barrier. The derivative jump is $\psi'_+ - \psi'_- = (2mg/\hbar^2)\psi$, while the wavefunction remains continuous. A barrier map therefore has unit diagonal, zero upper off-diagonal term and lower off-diagonal value $2mg/\hbar^2$. The determinant of each propagation map is one, reflecting conservation of the Wronskian. Multiplying barrier and free matrices gives the one-cell map M. Starting at a different point of the same periodic cell changes that matrix by a similarity transformation and leaves its trace and eigenvalues unchanged.

Derive the Bloch condition and identify a forbidden interval

A Bloch solution satisfies

$$(\psi, \psi')(x + a) = e^{iqa}(\psi, \psi')(x).$$

Therefore e^{iqa} is an eigenvalue of M; determinant one gives its partner e^{-iqa}. Half the trace is their average. Multiplying the two elementary matrices and introducing $x = ka$ and dimensionless barrier strength $P = mga/\hbar^2$ gives

$$\cos(qa) = F(x) = \cos x + P\frac{\sin x}{x}.$$

The free propagation wave number k defines the energy, while the Bloch wave number q specifies phase advance between equivalent cells. They coincide modulo reciprocal lattice vectors only in the unperturbed limit. A real q exists exactly when $|F| \le 1$. If $|F| > 1$, q must be complex; field magnitude changes exponentially from one cell to the next. A perfect infinite lattice then has no extended state at that energy, though a finite sample can support evanescent penetration from a contact.

For $P = 1$, $x = \pi/2$ gives

$$F = 2/\pi = 0.63662,$$

$$qa = \arccos(2/\pi) \simeq 0.881.$$

At $x = \pi$, $F = -1$ independent of P, a band edge. At $x = 3.40$, $F \approx -1.042$, inside a gap. Writing $q = \pi/a + i\kappa$ there gives $\cos qa = -\cosh(\kappa a)$, or $\kappa a = \operatorname{arcosh}|F| \approx 0.29$. A pure decaying solution falls by roughly $e^{-2.9}$ in amplitude over ten cells. A two-contact transmission measurement also depends on interface matching and admixtures of growing and decaying solutions; this envelope alone does not give the exact transmission probability.

Connect band edges to standing waves and curvature

At a zone edge the unperturbed right- and left-going states are degenerate. Their symmetric and antisymmetric combinations place maxima or nodes on the barriers. A sine standing wave with nodes at every barrier vanishes where the delta potential acts and retains its corresponding free energy. The complementary standing wave has nonzero barrier amplitude and shifts up. For small P, expand near $x = \pi + \delta$:

$$F \simeq -1 + \delta^2/2 - P\delta/\pi.$$

The two band edges with $F = -1$ are $\delta = 0$ and $\delta \simeq 2P/\pi$. Their energy separation is $\Delta E \simeq 2g/a$, since $E = \hbar^2 x^2/(2ma^2)$. This equals twice the magnitude of the first reciprocal-space Fourier amplitude of the periodic delta array, connecting the transfer-matrix and nearly-free-electron pictures.

Band curvature controls the acceleration of a slowly varying wave packet. Implicit differentiation gives $-a\sin(qa)\,dq/dE = dF/dE$ and hence

$$v_g = \hbar^{-1} dE/dq = -a\sin(qa)/[\hbar F'(E)].$$

At an ordinary edge $\sin qa = 0$, group velocity vanishes. Expanding both sides quadratically in the displacement of q shows the local effective mass can be positive at a lower minimum and negative at an upper maximum. Direct differentiation of a numerically sampled inverse cosine near its endpoint can amplify rounding error. Fitting $E(q)$ locally to a parabola and checking the interval used is generally more reliable.

Reconstruct a band diagram from the trace

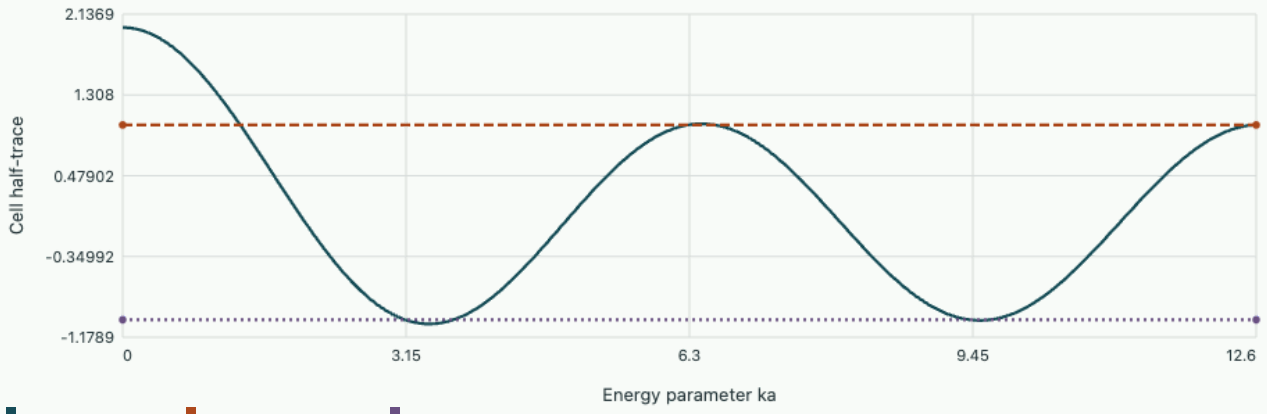
The live plot displays the half-trace as a function of energy parameter ka together with the boundaries at plus and minus one. Start with the free system and then raise P; identify where a previously tangential contact with $|F| = 1$ opens into an interval outside the allowed strip. Choose a test ka and read either the Bloch phase or the penetration length. Changing a at fixed P, ka rescales the physical energy as a^{-2}. A real scan of lattice spacing at fixed barrier strength g also changes P, so it cannot generally be represented by shifting only the horizontal energy labels.

Open this experiment online

Try the calculation

Barrier strength P **1**

Selected ka **3.4**



Half-trace at the selected energy is -1.0420. The decaying Bloch amplitude falls by 1/e in 3.46 cells.

Determine how a finite pass band produces resonant transmission

The trace controls a finite periodic sample as well as its infinite limit. For a determinant-one cell matrix the Cayley–Hamilton identity is

$$M^2 - 2FM + 1 = 0.$$

Repeatedly applying it gives

$$M^N = U_{N-1}(F)M - U_{N-2}(F)1,$$

with U_j the Chebyshev polynomials of the second kind. Inside a band $F = \cos qa$, so $U_{N-1} = \sin(Nqa)/\sin(qa)$. When $Nqa = j\pi$ strictly inside a band, the coefficient of M vanishes and the remaining total matrix is $(-1)^j 1$. With identical homogeneous leads on both sides this matches a travelling wave without reflection, giving a finite-stack transmission resonance even when one cell alone reflects substantially.

For six identical cells the candidate phases $qa = \pi/6, \ldots, 5\pi/6$ lie within each uncomplicated pass band. Their energies are found by solving the half-trace relation at each prescribed cosine; they are generally not equally spaced in energy. At a band edge $\sin qa = 0$ the divided expression appears singular, but the limiting polynomial remains finite. Exact transparency at the interior phases also assumes coherent propagation and accurately repeated barriers. Variations in cell strength, unequal external leads and inelastic dephasing alter the heights or positions. Comparing how the resonance spacing changes with cell count distinguishes finite-stack resonances from a fundamental bulk band gap.

Check dimensions against a measured energy scale

For $a = 0.50\,\mathrm{nm}$ and bare electron mass, the zone-edge free energy $\hbar^2\pi^2/(2ma^2)$ is about $1.50\,\mathrm{eV}$. Doubling the repeat distance at fixed mass lowers this reference by a factor four. Thus a plotted gap width measured in ka cannot be quoted directly in electronvolts without the lattice spacing, mass and selected location on the nonlinear energy axis. Using a fitted effective mass already incorporating the same periodic potential can also double count its effect. State the degrees of freedom represented by the unperturbed kinetic term.

Tutorial on allowed and forbidden propagation

Why is there no extended zero-energy state for any positive barrier strength in this model?

▼ **Work through the solution**

As $x \to 0$, $\sin x/x \to 1$ and $F(0) = 1 + P > 1$. The bottom of the first allowed band is shifted upward from zero. A constant wavefunction would have nonzero values at the barriers but zero derivative jump, violating the jump condition. Removing the barriers gives $P = 0$ and restores the constant zero-energy state.

At an energy with $F = -1.10$, how many cells attenuate a purely evanescent amplitude by a factor of ten?

▼ **Work through the solution**

The inverse decay length in units of cells is $\operatorname{arcosh}(1.10) = 0.444$. Setting $e^{-N\cdot 0.444} = 0.10$ gives $N = 5.19$, so six complete cells reduce it by more than ten. Intensity decays with twice the exponent. The alternating Bloch phase associated with a negative trace does not reverse the exponential attenuation.

A numerically multiplied matrix after several thousand gap cells overflows. Does this invalidate the band calculation?

▼ **Work through the solution**

A lossless cell has reciprocal eigenvalues; in a gap one grows while the physical decaying solution shrinks. Direct multiplication amplifies roundoff in the growing direction. Recover the bulk exponent from the one-cell trace or use a stable scattering-matrix or logarithmic-derivative formulation for a finite sample. Check conservation in an allowed band and compare the extracted decay exponent as the finite sample grows.

David Tong connects one-dimensional periodic-potential matching to Bloch bands.

Structure, excitations and probes of solids · 4. Why a large joint density of states can leave a dark optical edge

A band diagram gives the energy an electron needs to move between bands. It does not by itself tell whether light couples those states. Construct a two-orbital model with an explicit velocity operator, calculate its allowed and suppressed transitions, and distinguish electronic band curvature from optical selection rules.

Build a two-orbital Bloch Hamiltonian from parity

Consider one even and one odd orbital centred on each site of a one-dimensional chain of period a. Their onsite energies are $+\Delta$ and $-\Delta$. Suppose the only hopping retained mixes opposite orbital types on adjacent sites and has opposite signs to the left and right. Fourier transforming that difference gives a term proportional to $e^{ika} - e^{-ika} = 2i \sin ka$. Choosing orbital phases appropriately gives

$$H(k) = \Delta\sigma_z + 2t \sin(ka)\sigma_y.$$

This is a deliberately simple spinless two-band model with orthonormal, co-centred orbitals and a fixed basis. Physical multiorbital models can require additional intra-atomic dipole or position-matrix elements; here the optical coupling is defined through a uniform vector potential acting on the hopping phases.

Diagonalising a traceless Pauli-vector Hamiltonian gives

$$E_\pm = \pm E(k),$$

$$E(k) = \sqrt{\Delta^2 + 4t^2 \sin^2 ka}.$$

At half filling and zero temperature, the lower band is occupied and the upper is empty. A photon of visible-light wavelength usually has very small momentum compared with the size of a Brillouin zone in an atomic crystal. Neglecting its crystal momentum gives a vertical transition at fixed k with required energy $2E(k)$. This energetic condition defines a joint density of candidate states. A matrix element still determines their optical weight.

Derive the velocity matrix element before summing transitions

A uniform vector potential shifts the crystal wavevector in this tight-binding model, yielding velocity $v = \hbar^{-1}\partial_k H = (2ta/\hbar)\cos(ka)\,\sigma_y$. Let $\hat{\mathbf{d}} = (0, 2t\sin ka, \Delta)/E$. The upper eigenstate has $\langle +|\boldsymbol{\sigma}|+\rangle = \hat{\mathbf{d}}$. Insert the completeness relation $1 = |+\rangle\langle +| + |-\rangle\langle -|$ into $\langle +|\sigma_y^2|+\rangle = 1$. This gives

$$|\langle +|\sigma_y|-\rangle|^2 = 1 - \hat{d}_y^2 = \Delta^2/E^2.$$

Thus

$$|v_{+-}(k)|^2 = \left(\frac{2ta}{\hbar}\right)^2 \cos^2(ka)\,\frac{\Delta^2}{\Delta^2 + 4t^2\sin^2(ka)}.$$

At $ka = 0$ or π the interband velocity is maximal, even though the diagonal group velocity of either band is zero at these extrema. The optical matrix element connects different states; it is not the derivative of one eigenvalue. At $ka = \pi/2$ the interband velocity vanishes because the Hamiltonian has zero first derivative. The energy separation there is maximal and has zero slope as well. A state-counting argument alone would emphasise that band edge while the optical coupling suppresses it.

For $\Delta = 0.50\,\mathrm{eV}$, $t = 0.40\,\mathrm{eV}$, and $a = 0.40\,\mathrm{nm}$, transition energies range from 1.00 to $1.89\,\mathrm{eV}$. The velocity scale $2ta/\hbar$ is about $4.86 \times 10^5\,\mathrm{m/s}$. At $ka = \pi/4$, $E^2 = 0.25 + 0.32 = 0.57\,\mathrm{eV}^2$, so the squared matrix element in units of its zone-centre value is

$$0.5(0.25/0.57) = 0.219.$$

States at that wavevector contribute substantially less per state than at threshold even before differences in available state density or occupation are included.

Derive the line shape near the upper and lower edges

In a weak-field independent-electron model an interband spectral weight contains an integral $\int dk\,[f_-(k) - f_+(k)]|v_{+-}(k)|^2\delta[\hbar\omega - 2E(k)]$, together with known optical prefactors. For a simple root of $2E(k) = \hbar\omega$, the delta function contributes an inverse derivative $1/|2E'(k)|$. Near a quadratic one-dimensional extremum this factor grows as the inverse square root of the energy distance from the edge. That is the familiar van Hove feature in the unweighted joint density. One must multiply by the matrix element at precisely the same wavevector before interpreting it as absorption.

Near the lower edge write $ka = x \ll 1$. Then $2E \simeq 2|\Delta| + 4t^2x^2/|\Delta|$, and the squared velocity tends to a finite constant. The independent-particle absorption weight has an inverse-square-root onset in this ideal one-dimensional model. Near the upper edge put $ka = \pi/2 + s$ with small s. The distance from the maximum transition energy is proportional to s^2, but $|v_{+-}|^2 \propto s^2$. Multiplying this by the $1/|s|$ joint-density factor produces a weight that vanishes as a square root. The same quadratic band-edge geometry thereby creates entirely different optical signatures at the two edges.

Finite temperature modifies the occupations; moving the chemical potential into the upper band can block some low-energy transitions by Pauli exclusion. Lifetime and instrument broadening smear each edge. Electron–hole attraction can also create bound exciton peaks below the independent-particle onset. A measured absorption line below a band gap inferred from one-electron energies is therefore not automatically an error in the band calculation. Test how the feature changes with carrier density, temperature, polarisation and resolution before assigning its microscopic origin.

Distinguish a vertical transition from a phonon-assisted onset

If the valence maximum and conduction minimum occur at different crystal wavevectors, a direct vertical photon transition may require a larger energy than the minimum electronic gap. A lattice vibration can supply or remove the missing crystal momentum. Let the indirect gap be E_g, and suppose a phonon mode of energy $\hbar\omega_{\mathrm{ph}}$ connects the extrema. Phonon absorption provides part of the required energy, so photon onset can occur near $E_g - \hbar\omega_{\mathrm{ph}}$, weighted by the occupation n_{ph}. Phonon emission costs

energy and starts near $E_g + \hbar\omega_{\mathrm{ph}}$, weighted by $n_{\mathrm{ph}} + 1$. At low temperature the emission channel can remain while the phonon absorption channel weakens.

For $E_g = 1.10\,\mathrm{eV}$ and phonon energy $0.040\,\mathrm{eV}$, the simple channel edges are 1.06 and $1.14\,\mathrm{eV}$. At $300\,\mathrm{K}$ the Bose occupation is about 0.270, so the ratio of absorption to emission statistical factors is $n/(n+1) \approx 0.213$. Matrix elements and available electronic states also affect measured intensity, so the intensity ratio need not equal that number directly. Compare temperature and momentum dependence with an independently mapped electronic dispersion. A weak low-energy feature can reflect an indirect process rather than a symmetry-forbidden direct transition, two mechanisms that the direct one-dimensional model intentionally separates.

Check what reversing the hopping sign can reveal

Changing t to $-t$ in this restricted Hamiltonian leaves both band energies and the squared interband matrix element unchanged. A redefinition of one local orbital by a phase of π produces exactly that sign change. A fit to one absorption spectrum therefore cannot identify an absolute sign for this lone hopping. Relative phases among several independent paths can become observable through interference in a richer orbital model.

Tutorial on interpreting optical data

At equal density of candidate states for two selected small k intervals, the second interval has $ka = \pi/4$ and the first $ka = 0$. Which optical signal ratio does the model predict if occupancies and other prefactors are matched?

▼ **Work through the solution**

The answer is the squared velocity ratio 0.219 for the numerical parameters. A ratio of velocity magnitudes would give the square root of that value and would be inappropriate for a transition probability. In a direct frequency sweep the state-density factors at those two energies generally differ, so this particular comparison presupposes selected or separately corrected state intervals.

A numerical spectrum represents each delta function by a box of width Γ. How should a peak comparison be checked as Γ decreases?

▼ **Work through the solution**

Increase the k-mesh density so each bin contains enough states, especially near extrema where roots merge. Compare integrated weight over a fixed finite energy window and use broadening normalised to unit area. Then vary Γ separately from the mesh. Otherwise a disappearing or sharpening feature may reflect changing sample counts per bin instead of the true matrix element.

Can the model predict a strong direct transition simply because $2E(k)$ equals the photon energy at a band maximum?

▼ Work through the solution

The maximum at $ka = \pi/2$ is an explicit counterexample: its velocity matrix element vanishes. Additional allowed optical couplings or symmetry breaking in a real material can restore weight, so identify which perturbation the model includes. The band energies alone supply a necessary energetic condition and the matrix element supplies an independent condition on how the external field drives the transition.

Vary a two-band hopping model and distinguish its dispersion from the optical weights.

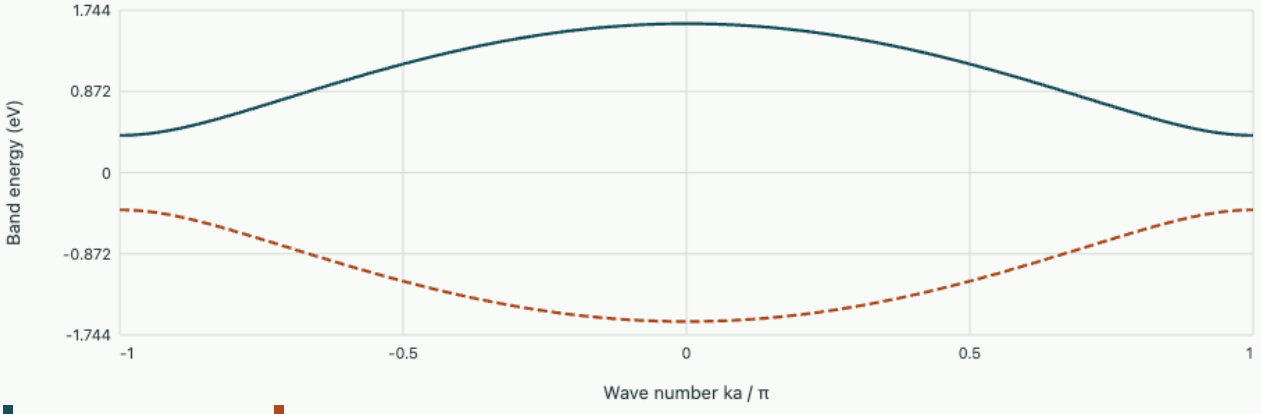
David Tong derives tight-binding Hamiltonians and electronic dynamics in solids.

Open this experiment online

Try the calculation

Intracell hopping t_1 (eV) **1**

Intercell hopping t_2 (eV) **0.6**



Direct gap 0.8 eV. Each band spans 1.2 eV.

Structure, excitations and probes of solids · 5. Count vacancies and infer an atomic migration barrier

Tracer atoms spread through a crystal even when the average lattice positions remain fixed. The rate can depend both on the number of vacant sites and on how readily an atom jumps into one. Derive the two factors separately, then compare an equilibrated specimen with a rapidly quenched one.

Minimise the free energy of a dilute vacancy population

Take N lattice sites and n equivalent vacancies, ignoring their interactions. The microscopic formation Gibbs free energy $G_f = H_f - TS_f$ includes local vibrational and other nonconfigurational changes associated with creating one defect and transferring the removed atom to an appropriate reservoir. The configurational multiplicity is $\binom{N}{n}$. Thus $G - G_0 = nG_f - k_BT \ln \binom{N}{n}$. Applying Stirling's approximation and differentiating with respect to n gives $G_f + k_BT \ln[n/(N-n)] = 0$, or $c_v/(1-c_v) = e^{-G_f/k_BT}$, where $c_v = n/N$. In the dilute limit $c_v \simeq e^{S_f/k_B} e^{-H_f/k_BT}$.

A positive formation enthalpy does not imply zero vacancies in thermal equilibrium. The number of ways to place a few defects contributes entropy. The dilute independent-site argument ceases to apply if vacancies cluster, interact strongly, or if charge neutrality ties their population to other charged defects. Specifying the atomic reservoir is especially important in compounds: removing a metal atom and removing a nonmetal atom are different chemical operations with potentially different energies and charge states. The present simple model describes one kind of neutral vacancy under fixed thermodynamic conditions.

For $H_f = 1.00\,\mathrm{eV}$, $S_f = k_B$ and $T = 800\,\mathrm{K}$, $k_BT = 0.06894\,\mathrm{eV}$ and $c_v \simeq 1.36 \times 10^{-6}$. A cubic micrometre containing an illustrative 5×10^{10} sites then contains about 6.8×10^{4} vacancies on average. The fraction is tiny while the count in a region accessible to a diffusion experiment can be large. In a nanocrystal with far fewer sites the mean can be much less than one; a single defect entering or leaving may then change the observed motion intermittently.

Derive a tracer diffusivity from correlated jumps

Let the coordination number be z, jump length d, and the conditional rate of one atom jumping into one specified neighbouring vacancy be $\Gamma = \nu_0 e^{-H_m/k_BT}$. For dilute, randomly located vacancies a tracer has on average zc_v vacant nearest neighbours and jump rate $zc_v\Gamma$. For an uncorrelated isotropic three-dimensional random walk $\langle |\Delta \mathbf{r}|^2 \rangle = zc_v\Gamma t\, d^2 = 6Dt$. Successive tracer jumps can be correlated because the vacancy just exchanged with it is now directly behind it. Introduce the cumulative correlation factor f to obtain

$$D = fzd^2\Gamma c_v/6.$$

The definition of Γ per specified neighbour matters: a tabulated total jump frequency already summed over neighbours must not acquire a second factor z.

Under the stated equilibrium assumptions the apparent Arrhenius activation energy of the tracer is $H_f + H_m$. The first term controls whether an opportunity to jump exists, the second controls the rate at which an available opportunity is taken. At fixed excess vacancy concentration only the migration

contribution enters to leading order if f and the prefactor stay nearly constant. This offers a way to separate the two barriers, provided the excess concentration persists during the measurement. A fitted straight line through data from specimens with different thermal histories is not automatically the property of one microscopic jump.

For $H_m = 0.50\,\mathrm{eV}$, $\nu_0 = 10^{13}\,\mathrm{s^{-1}}$, $z = 12$, $d = 0.25\,\mathrm{nm}$, $f = 0.78$ and the equilibrium vacancy fraction at $800\,\mathrm{K}$, one obtains $\Gamma \simeq 7.1 \times 10^{9}\,\mathrm{s^{-1}}$ and $D \simeq 9.4 \times 10^{-16}\,\mathrm{m^2/s}$. The apparent discrepancy between a gigahertz conditional jump rate and the slow macroscopic spreading comes largely from the small vacancy fraction. After one hour the rms spread in one Cartesian direction is $\sqrt{2Dt} \simeq 2.6\,\mu\mathrm{m}$; the three-dimensional rms displacement is larger by $\sqrt{3}$.

Read a diffusion measurement and identify a changing population

For a thin tracer layer deposited at the origin of a long homogeneous specimen the one-dimensional conservative diffusion equation gives

$$c(x,t) = M (4\pi D t)^{-1/2} e^{-x^2/(4Dt)},$$

where M is the deposited amount per area. Its variance is $2Dt$. An initial Gaussian width or an independent Gaussian instrument resolution adds its own variance. Measure the growth of variance with time to estimate D, instead of attributing the entire final width to diffusion. An asymmetric tail can signal surface or grain-boundary transport, which may not be described by a uniform bulk D.

Now quench a specimen rapidly from a higher temperature. Vacancies may initially remain near their high-temperature population while their mobility changes immediately. If their excess decays to sinks as $c_v(t) = c_{\mathrm{eq}} + c_{\mathrm{exc}} e^{-t/\tau_s}$, tracer spreading at fixed temperature follows $\sigma_x^2(t) - \sigma_x^2(0) = 2\int_0^t D(t')\,dt'$. If D is proportional to c_v, this consists of a linear equilibrium contribution and a saturating excess term proportional to $\tau_s(1 - e^{-t/\tau_s})$. A single long-time apparent diffusivity conceals early rapid motion. Measuring several dwell times and different specimen sizes can reveal the role of sinks.

Use unequal dwell times to test a quenched vacancy excess

Take a quenched specimen at fixed measuring temperature and suppose the excess initially raises the tracer diffusivity from its equilibrium value D_{eq} to $D_{\mathrm{eq}} + 3D_{\mathrm{eq}}$. Let excess vacancies decay with time constant $600\,\mathrm{s}$. A variance measurement over the first $300\,\mathrm{s}$ infers a time-averaged coefficient $D_{\mathrm{app}}/D_{\mathrm{eq}} = 1 + 3(600/300)[1 - e^{-300/600}] = 3.36$. The same calculation over the first $3600\,\mathrm{s}$ gives about 1.50. Neither averaged coefficient equals the initial fourfold value or the long-time equilibrium value. A collection of samples analysed after unequal dwell times can therefore seem to have different activation energies even at equal measurement temperatures.

Independent reheating or quenching measurements can help identify this bias. First anneal for several estimated sink times, then perform a fresh short tracer measurement at the same temperature. Agreement of late-time slopes for different initial quench histories supports a common equilibrium diffusion coefficient. Changes in specimen thickness that shorten the distance to a vacancy sink should affect the excess decay stronger than the elementary jump barrier. Verify that changing the thickness has

not simultaneously changed grain size or surface diffusion paths. Combining the full time-dependent variance with an independent vacancy-sensitive observable is more informative than treating a series of single final widths as separate equilibria.

Tutorial on activation energies

The slope of $\ln D$ versus $1/T$ is $-1.74 \times 10^{4}\,\mathrm{K}$ in an equilibrated specimen. A short quenched measurement with a fixed vacancy population gives $-5.80 \times 10^{3}\,\mathrm{K}$. Estimate both enthalpies.

▼ **Work through the solution**

Multiplying slopes in magnitude by $k_B = 8.6173 \times 10^{-5}\,\mathrm{eV/K}$ gives approximately $1.50\,\mathrm{eV}$ total and $0.50\,\mathrm{eV}$ migration. Their difference gives $1.00\,\mathrm{eV}$ formation. This inference assumes the same dominant vacancy mechanism and no significant changes of the prefactor or defect population over the quenched measurement.

A tracer profile grows in rms width from 1.5 to $3.0\,\mu\mathrm{m}$ in $3600\,\mathrm{s}$. Estimate the one-dimensional diffusion coefficient.

▼ **Work through the solution**

Subtract the initial variance rather than the initial width:

$$D = [(3.0)^2 - (1.5)^2]10^{-12}/7200 = 9.38 \times 10^{-16}\,\mathrm{m^2/s}.$$

If the instrument resolution changes between runs its variance change must also be estimated. A constant independent resolution cancels in the subtraction.

What would distinguish fast grain-boundary transport from an unexpectedly low bulk migration barrier?

▼ **Work through the solution**

Compare single crystals with specimens of varying grain size at the same thermal history and fit the full profile rather than only its second moment. Enhanced fast tails that grow with boundary area support an additional path. An independent defect-population measurement helps separate a low jump barrier from an enhanced concentration. A single Arrhenius slope represents the dominant observed transport, not a unique proof of its mechanism.

The University of Zurich solid-state course connects microscopic arrangement and macroscopic material response.

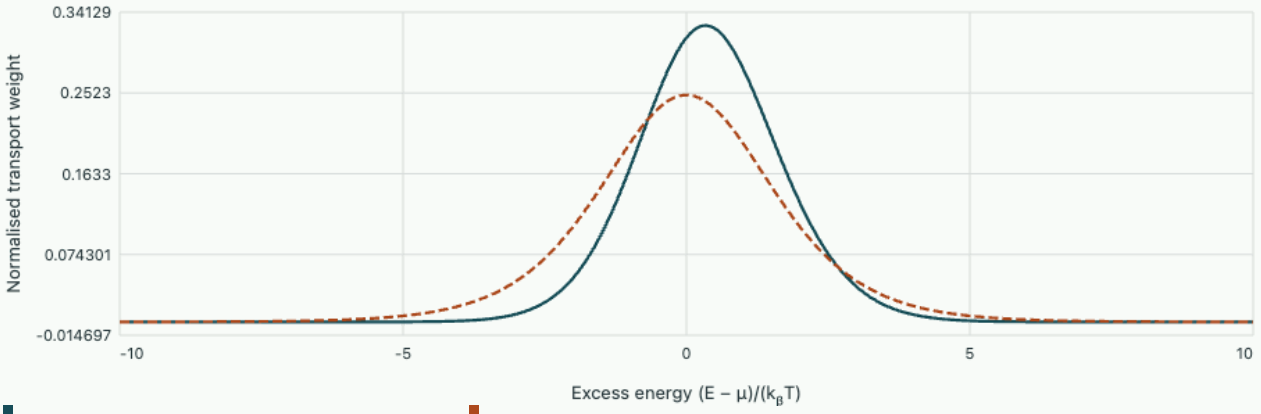
Compare a different inferred transport coefficient with the kinetic factors behind it.

Open this experiment online

Try the calculation

Filter centre $(E - \mu)/(k_\beta T)$ **1**

Filter width / $(k_\beta T)$ **2**



Energy-selected current carriers Smooth metallic transport

Mean excess energy 0.4018 $k_\beta T$; electron Seebeck coefficient -34.63 µV/K. Open-circuit Lorenz ratio L/L_o = 0.4937.

Structure, excitations and probes of solids · 6. Measure a phonon mean free path without confusing it with a sample length

A cold insulating bar can carry heat primarily through lattice vibrations. Heat capacity tells how much energy the vibrations can carry; conductivity also depends on how far and in which direction they carry it. Derive the connection, test its limiting regimes, and identify when a measured thermal resistance belongs to a boundary.

Derive the kinetic expression from a displaced occupation

Consider phonon modes labelled by wavevector $\mathbf{k}$ and branch s, with angular frequency $\omega_{\mathbf{k}s}$ and group velocity $\mathbf{v}_{\mathbf{k}s} = \nabla_{\mathbf{k}}\omega_{\mathbf{k}s}$. A mode transports energy $\hbar\omega$ per quantum. At equilibrium states with opposite velocities cancel in the heat current. A small slowly varying temperature gradient perturbs the Bose occupation $n_0(T)$. In the relaxation-time approximation the stationary transport equation is

$$\mathbf{v} \cdot \nabla T \, \partial_T n_0 = -\delta n / \tau.$$

This describes weak near-equilibrium transport when a mode can be assigned a characteristic relaxation time.

Insert the correction into the heat current density $\mathbf{j}_Q = V^{-1} \sum \hbar\omega \, \mathbf{v} \, \delta n$. Comparing with $j_{Q,i} = -\kappa_{ij} \partial_j T$ gives

$$\kappa_{ij} = V^{-1} \sum c_{\mathbf{k}s} v_i v_j \tau_{\mathbf{k}s},$$

where $c_{\mathbf{k}s} = \hbar\omega \, \partial_T n_0$ is the heat capacity contribution of one mode. In an isotropic solid the angular average of v_x^2 is $v^2/3$. An effective common speed and path $\ell = v\tau$ then yield $\kappa \simeq C_V v \ell / 3$ for heat capacity per volume C_V. The factor one third counts directional projection, not the three separate phonon branches.

Connect the Debye heat capacity to low-temperature transport

For a three-dimensional isotropic acoustic branch with $\omega = v_s k$, the number of states in a wavevector shell is proportional to $k^2 dk$. Thus its density of states per volume is proportional to ω^2 / v_s^3. In the thermal energy $\int d\omega \, g(\omega) \hbar\omega / (e^{\hbar\omega/k_B T} - 1)$, substitution $x = \hbar\omega / k_B T$ produces a factor T^4 when occupied frequencies remain below the high-frequency cutoff. Differentiation gives

$$C_V \propto T^3.$$

Different sound speeds of longitudinal and transverse branches enter the coefficient. An optical mode with a nonzero minimum frequency contributes much less at sufficiently low temperature.

If rough sample boundaries dominate scattering and give an approximately temperature-independent mean path, the kinetic expression predicts $\kappa \propto T^3$. A more specular boundary can increase the effective path of long-wavelength modes and modify that exponent. Conversely once many modes are thermally

populated, heat capacity tends toward saturation and increasing phonon–phonon or defect scattering can reduce conductivity. Adding rates as $1/\tau = 1/\tau_{\mathrm{b}} + 1/\tau_{\mathrm{defect}} + 1/\tau_{\mathrm{anh}}$ assumes independent effective relaxation mechanisms. Normal phonon collisions conserve crystal momentum and do not always relax heat current as efficiently as a naive single-rate sum implies.

Take an insulating crystal at one temperature with measured $C_V = 40\,\mathrm{J/(m^3K)}$, representative $v = 4000\,\mathrm{m/s}$, and $\kappa = 0.80\,\mathrm{W/(mK)}$. The inferred effective path is $3\kappa/(C_V v) = 15\,\mu\mathrm{m}$, corresponding to $3.75\,\mathrm{ns}$. In a $1.0\,\mathrm{mm}$-wide specimen this is much shorter than its width. An assertion that only diffuse sidewall collisions limit transport would need further evidence. Measuring a second transverse dimension and checking whether the inferred path scales with it is more discriminating than matching a single temperature exponent.

Background · Which temperature does a finite-gradient measurement report?

Suppose the conductivity obeys $\kappa(T) = bT^3$ across a homogeneous bar with a diffusive local temperature field. In a steady state the heat flux is uniform. Integrating Fourier's law shows that dividing the heat flux times length by the end-to-end temperature difference returns the temperature average of κ. Let T_m be the arithmetic mean and $\Delta T = T_h - T_c$.

$$\kappa_{\mathrm{app}} = \frac{1}{\Delta T}\int_{T_c}^{T_h} bT^3\,dT = bT_m^3\left(1 + \frac{\Delta T^2}{4T_m^2}\right).$$

A 1 K drop centred on 2 K raises the apparent conductivity by 6.25% relative to the local value at 2 K. Halving the temperature difference reduces the excess by a factor of four. That dependence gives an experimental check before attributing a deviation to boundary scattering.

Distinguish sample conductivity from contact and ballistic resistance

A steady one-dimensional bar of area A and length L carries power $P = \kappa A \Delta T_{\mathrm{bulk}}/L$ in the diffusive regime. If the two contacts have thermal resistances R_1, R_2, an externally measured temperature difference gives

$$R_{\mathrm{tot}} = \Delta T/P = R_1 + R_2 + L/(\kappa A).$$

A plot of total resistance against length for otherwise equivalent specimens can separate intercept from bulk slope. A contact-limited device fitted as pure bulk appears to have a length-dependent conductivity. At very small lengths, ballistic transport itself supplies a length-independent resistance, which a length sweep alone cannot always distinguish from ordinary contacts.

For area $A = 1.0\,\mathrm{mm}^2$, suppose resistance rises from $3600\,\mathrm{K/W}$ for a $2.0\,\mathrm{mm}$ bar to $8600\,\mathrm{K/W}$ for a $6.0\,\mathrm{mm}$ bar under matched contacts. The slope gives $\kappa = 0.800\,\mathrm{W/(mK)}$ and the intercept $1100\,\mathrm{K/W}$. At $P = 20\,\mu\mathrm{W}$ the shorter device has an external temperature difference $0.072\,\mathrm{K}$, of which $0.022\,\mathrm{K}$ is

attributable to the contacts. Failing to subtract the intercept gives an apparent conductivity only $0.556\,\mathrm{W/(mK)}$. Temperature probes located inside the bar can measure the bulk gradient more directly.

When the path is comparable to a sample dimension the local Fourier relation with one well-defined temperature per point needs justification. In a simple isotropic ballistic picture a reservoir emits an energy flux on one side with angular factor $1/4$, leading to a conductance per area of order $C_V v/4$ for perfectly transmitting contacts. This differs from the $1/3$ in the diffusive expression; the angular integrals describe different boundary-value problems. A finite transmission probability, mode mismatch and partial equilibration near reservoirs can further reduce conductance. A stated mean free path inferred using Fourier's law beyond this crossover is an effective fit, not an independently measured microscopic length.

Estimate when a heater measurement leaves linear response

A cold crystal can have a temperature-dependent conductivity over even a small absolute thermometer difference. In a uniform bar at steady state, constant heat current obeys $PL/A = \int_{T_c}^{T_h} \kappa(T)\,dT$. If $\kappa = \beta T^3$, then $P = \beta A(T_h^4 - T_c^4)/(4L)$. Dividing measured flux by a nominal gradient gives the interval average of conductivity, rather than its exact value at either thermometer temperature. For $T_c = 2\,\mathrm{K}, T_h = 4\,\mathrm{K}$, the averaged conductivity is 30β, whereas κ at the mean temperature $3\,\mathrm{K}$ is 27β. The eleven-percent difference is a consequence of nonlinear temperature dependence, even with perfect contacts and sensors.

Shrink the heater excursion at the same mean temperature to test the small-gradient approximation. A contact bottleneck can itself become nonlinear, in which case subtracting a constant contact resistance found at a different mean temperature is also inappropriate. Modulate a small heating power around a stable bias and analyse the differential thermal resistance, or measure temperatures at internal probes for several heat currents. These protocols probe a particular operating point and can distinguish the curvature of the intrinsic material response from the curvature of the two contact heat-transfer laws.

Tutorial on a thermal measurement

A bar with boundary-limited $C_V \propto T^3$ is cooled from 4 K to 2 K with unchanged diffuse boundary path. How should its thermal conductivity and its diffusivity change?

▼ **Work through the solution**

Conductivity falls by a factor eight. Thermal diffusivity $\kappa/C_V \simeq v\ell/3$ stays approximately constant. An experiment measuring the relaxation time of a spatial sinusoidal temperature disturbance probes diffusivity, while a steady gradient at given heat flux probes conductivity. The two measurements need not have the same temperature dependence.

Using the worked $15\,\mu\mathrm{m}$ path, predict the thermal diffusivity and the decay time of a long-wavelength sinusoidal temperature modulation with wavelength $1.0\,\mathrm{mm}$.

▼ **Work through the solution**

The diffusivity is

$$0.80/40 = 0.020\,\mathrm{m}^2/\mathrm{s}.$$

The diffusive decay time is

$$1/[D_T(2\pi/\lambda)^2] = 1.27\,\mu\mathrm{s}.$$

Here $k\ell \approx 0.094$, reasonably small. At a wavelength comparable to the inferred path a diffusive exponential should be tested rather than assumed.

A sample's conductivity increases after isotopic purification while its measured heat capacity hardly changes. What microscopic inference is supported?

▼ Work through the solution

In the kinetic picture similar heat capacity and sound speed leave a longer weighted phonon path as the natural interpretation. Isotope disorder can scatter vibrations while leaving the gross mode count similar. Check geometric dimensions and contact resistances before attributing the full change to bulk scattering. A mean path deduced this way is weighted by mode heat capacities and velocities, not a single universal length for every phonon.

David Tong counts acoustic and optical phonons and derives their thermodynamics.

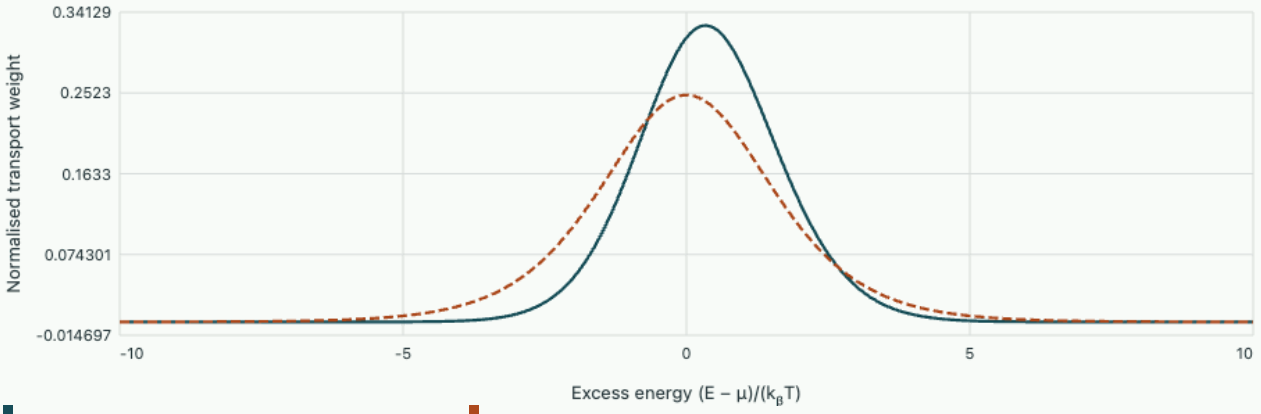
Compare lattice heat flow with heat carried by electronic excitations.

Open this experiment online

Try the calculation

Filter centre $(E - \mu)/(k_\beta T)$ **1**

Filter width / $(k_\beta T)$ **2**



Energy-selected current carriers Smooth metallic transport

Mean excess energy 0.4018 $k_\beta T$; electron Seebeck coefficient -34.63 µV/K. Open-circuit Lorenz ratio L/L_o = 0.4937.

Structure, excitations and probes of solids · 7. A field profile that distinguishes a superconductor from a perfect conductor

Zero dc resistance tells how a current can persist. Field expulsion tells how a specimen responds when it enters a superconducting equilibrium state. Derive the London length, calculate what a finite slab and a thin film actually screen, and connect magnetic and electrical measurements to the same superfluid density.

Separate frozen flux from equilibrium screening

For mobile carriers of mass m, charge q and number density n_s, a simple inertial equation $m\dot{\mathbf{v}} = q\mathbf{E}$ gives

$$\partial_t \mathbf{j}_s = n_s q^2 \mathbf{E}/m.$$

This first London equation explains dissipation-free acceleration by an electric field in the ideal response. Take its curl and use Faraday's law: $\partial_t[\nabla \times \mathbf{j}_s + (n_s q^2/m)\mathbf{B}] = 0$ at fixed uniform n_s. The expression in brackets is time independent. This derivative equation by itself allows an arbitrary initial curl–field combination, so it does not force a previously existing static induction to leave a conductor.

In a simply connected, vortex-free equilibrium superconductor the second London relation supplies the missing condition,

$$\nabla \times \mathbf{j}_s = -(n_s q^2/m)\mathbf{B}.$$

A condensate phase argument gives the same curl relation locally: mechanical pair momentum is $\hbar\nabla\varphi - q\mathbf{A}$; on a patch without a vortex $\nabla \times \nabla\varphi = 0$. With $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}_s$ and $\nabla \cdot \mathbf{B} = 0$, take the curl once more to obtain

$$\nabla^2 \mathbf{B} = \mathbf{B}/\lambda_L^2,$$
$$\lambda_L^2 = m/(\mu_0 n_s q^2).$$

The pair of signs yields exponential decay, not sinusoidal oscillation into a bulk specimen.

The effective mass, charge and density must describe the same carriers. For Cooper pairs of mass about $2m_e$, charge magnitude $2e$, and pair density $n_{\mathrm{electron}}/2$, the combination $n_s q^2/m$ agrees with the expression in terms of participating electron density, charge e and mass m_e. Mixing a pair charge with an electron density and electron mass overpredicts the stiffness by a factor of four. A material effective mass or nonlocal response may alter an inferred density, so a penetration-depth measurement fixes the stiffness combination rather than each microscopic parameter separately.

Solve a finite slab in a parallel applied field

A slab fills $-d/2 < x < d/2$ and a uniform applied induction $B_a\hat{z}$ is imposed parallel to both faces. Assume the Meissner regime, negligible demagnetising fields, and a thickness large compared with

microscopic dimensions needed for continuum electrodynamics. The even London solution is

$$B_z(x) = B_a \cosh(x/\lambda_L)/\cosh[d/(2\lambda_L)].$$

Its current is $j_y = -(1/\mu_0)\partial_x B_z$, of opposite signs on the two faces. The mean induction is

$$\frac{\bar{B}}{B_a} = \frac{2\lambda_L}{d}\tanh\frac{d}{2\lambda_L}.$$

In a thick slab the centre field is exponentially small while the average falls only as $2\lambda_L/d$, because a surface layer from each face still contains field. A probe at the centre and one measuring total flux through the cross section thus report different thickness dependence. In a thin slab the average is $1 - d^2/(12\lambda_L^2) + \cdots$, giving weak screening of a parallel field. This is not the perpendicular-field screening problem for a thin film, where fields outside the sample and lateral geometry lead to another length scale.

For $\lambda_L = 100\,\mathrm{nm}$, $d = 400\,\mathrm{nm}$, the centre ratio is $\operatorname{sech} 2 = 0.266$ and the mean ratio is

$$\tanh 2/2 = 0.482.$$

With $B_a = 0.50\,\mathrm{mT}$ the current magnitude just inside a face is

$$B_a \tanh 2/(\mu_0\lambda_L) = 3.84 \times 10^9\ \mathrm{A/m^2}.$$

It flows only over a thin scale. If a material's critical current or critical field is exceeded, this linear vortex-free solution ceases to describe the experiment. Driven hysteresis or trapped vortices can then make the same external field produce different internal profiles after different preparation histories.

Connect kinetic energy to a sheet inductance

A thin uniform strip of length l, width w, thickness $d \ll \lambda_L$ carries a current nearly uniformly through its thickness in a suitable transport geometry. The carrier kinetic energy per volume is

$$mn_s v^2/2 = mj_s^2/(2n_s q^2) = \mu_0\lambda_L^2 j_s^2/2.$$

Integrating over volume and using $j_s = I/(wd)$ gives $E_{\mathrm{kin}} = L_k I^2/2$ with $L_k = \mu_0\lambda_L^2 l/(wd)$. Thus the sheet kinetic inductance is $L_{k,\square} = \mu_0\lambda_L^2/d$, and the number of squares is l/w. It is distinct from geometric inductance storing energy in the magnetic field around the wire.

For $d = 20\,\mathrm{nm}$, $\lambda_L = 100\,\mathrm{nm}$ the sheet value is about $0.628\,\mathrm{pH}$ per square. A strip containing two thousand squares gives $1.26\,\mathrm{nH}$ of kinetic inductance. Its resonance with a capacitor probes the sum of this inductance and geometric or stray contributions. If the superfluid density falls while the effective mass stays approximately fixed, λ_L^2 increases and the resonant frequency tends to decrease. Measuring several lengths and widths can distinguish a distributed kinetic contribution from an approximately constant lead contribution. At frequencies approaching the gap scale a purely nondissipative London description is insufficient.

Determine what a thickness series can identify

A thin-film measurement sometimes fits the sheet kinetic inductance at a single thickness and quotes a bulk condensate density. In the uniform-film limit the measured product is $L_{k,\square} d = \mu_0 \lambda_L^2$. A thickness calibration error therefore rescales the inferred stiffness even if the resonance frequency is measured extremely accurately. For the $20\,\mathrm{nm}$ example, a systematically inactive layer of $2.0\,\mathrm{nm}$ leaves only $18\,\mathrm{nm}$ contributing. Substituting the nominal thickness into a stiffness inference underestimates the conducting carrier density by ten percent if the other model assumptions hold. A series of films and an independent estimate of the active thickness can test this interpretation.

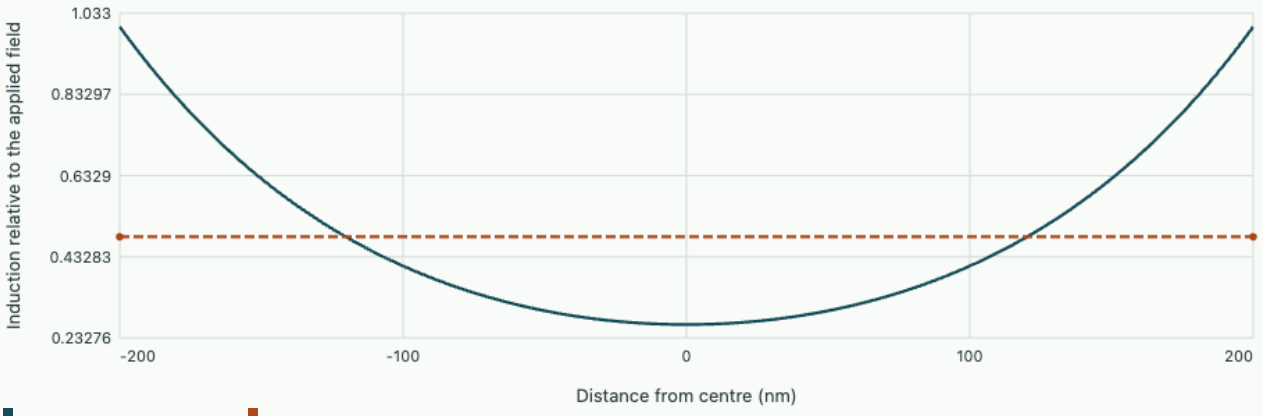
Compare with the parallel-field slab calculation by changing the ratio of thickness to penetration depth in the live profile. At a ratio of one, the centre and surface inductions differ only modestly. At a ratio of ten the interior is nearly field free while a substantial fraction of the averaged flux still resides near the two faces. The programme graphs B/B_a across the physical thickness and reports both centre and mean, which are different experimental weightings of the same field. Compare those numbers with an independently measured kinetic inductance before attributing a temperature shift entirely to an altered carrier density. Higher kinetic inductance can also reflect an increased effective mass or a reduced active cross section.

Open this experiment online

Try the calculation

Film thickness (nm) **400**

London length (nm) **100**



At the centre B/Ba = 0.2658, whereas the cross-section mean is 0.4820. These apply to a parallel field in a vortex-free slab.

Compare parallel screening with a perpendicular-field measurement

For a wide film in perpendicular field the magnetic flux lines bend around the edges, and the one-dimensional parallel-field cosh solution no longer applies. Sample width, aspect ratio and the field outside the superconducting material enter the response. Rotate the same thin film through ninety degrees and an apparent change in susceptibility may therefore be largely geometrical. Verify orientation and demagnetising conditions before combining the two datasets into one estimate of penetration depth.

Tutorial on comparing probes

A thick slab shows a centre field less than one percent of the applied parallel field. What lower bound on d/λ_L follows from the local London solution?

▼ Work through the solution

Require $\cosh[d/(2\lambda_L)] > 100$, hence $d/\lambda_L > 2\,\mathrm{arcosh}(100) \approx 10.60$. The mean field at this threshold would still be about $2/10.60 = 0.189$ of the applied field. A centre measurement at one percent therefore does not imply a one-percent cross-section average.

If the superfluid charge stiffness $n_s q^2/m$ falls by half, how do penetration depth and sheet kinetic inductance change for the same film?

▼ Work through the solution

The penetration depth increases by $\sqrt{2}$, and the sheet kinetic inductance doubles because it scales as λ_L^2. A resonance in which kinetic inductance is only one quarter of total inductance will shift less than the $1/\sqrt{2}$ expected if all inductance doubled. The initial circuit partition must be estimated.

A ring retains flux after cooling in a field. Does this disprove the equilibrium slab result?

▼ Work through the solution

The slab derivation assumed simple connectivity, a vortex-free state and prescribed fields on two faces. A ring permits nontrivial phase winding and fluxoid quantisation. A multiply connected circuit can maintain trapped flux even with local London screening in its material. Compare with a simply connected specimen under controlled zero-field cooling and field-cooling protocols before interpreting the ring as a direct slab test.

The University of Zurich solid-state course includes superconductivity and its magnetic signatures.

Compare ordinary resistive field diffusion with an equilibrium London profile.

Structure, excitations and probes of solids · 8. Infer a superconducting gap from a tunnelling spectrum

A tunnel junction reports differential conductance versus voltage. Extracting a superconducting energy gap requires connecting a condensate's quasiparticle dispersion to a density of states, then accounting for temperature and finite spectral resolution before comparing peaks with the data.

Diagonalise the paired-state Hamiltonian

In a simple isotropic weak-coupling description, pair a state of normal-state energy $\xi_k = \epsilon_k - \mu$ with the time-reversed hole amplitude. Take a real gap $\Delta > 0$ by a phase choice. The two-component mean-field matrix has rows (ξ_k, Δ) and $(\Delta, -\xi_k)$, giving eigenvalues $\pm E_k$ with $E_k = \sqrt{\xi_k^2 + \Delta^2}$. The positive-energy quasiparticle mixes particle-like and hole-like amplitudes with squared coefficients $(1 + \xi_k/E_k)/2$ and $(1 - \xi_k/E_k)/2$. Their sum is one. Near the normal-state Fermi surface they are equal, while far above it the excitation becomes predominantly particle like.

The minimum positive excitation energy is Δ. Creating two independent quasiparticles requires at least 2Δ in the ideal isotropic spectrum. These two thresholds refer to different probes. In an ideal normal-metal–insulator–superconductor tunnel junction, a bias can transfer one electron into an allowed quasiparticle state; the zero-temperature threshold is $|eV| = \Delta$. In a junction between two identical gapped electrodes, lowest-order quasiparticle tunnelling at zero temperature instead begins near 2Δ, aside from pair tunnelling and other processes. Confusing the physical junction type can double the reported gap.

Transform the normal-state density of levels

Assume approximately constant normal electronic density near the Fermi level over an energy range much larger than Δ. The quasiparticle dispersion flattens near $\xi = 0$. Mapping the two normal-state branches $\xi = \pm\sqrt{E^2 - \Delta^2}$ onto a positive energy and accounting for their electron spectral weights gives the normalised superconducting tunnelling density

$$\frac{N_s(E)}{N_n(0)} = \begin{cases} 0, & |E| < \Delta, \\ |E|/\sqrt{E^2 - \Delta^2}, & |E| > \Delta. \end{cases}$$

The divergence at the gap edge is integrable. Indeed integrating from Δ to $A > \Delta$ gives $\sqrt{A^2 - \Delta^2}$, tending toward the corresponding normal integral at large cutoff. The edge feature represents redistribution of states from the gap, rather than creation of an infinite number of states. In a finite normal band the constant-density approximation is only local, so statements about total spectral weight should use a consistent large-energy model.

For $\Delta = 1.50\,\mathrm{meV}$, the ideal density at $E = 2.00\,\mathrm{meV}$ is $2/\sqrt{4 - 2.25} = 1.51$ times the normal level. At $5.0\,\mathrm{meV}$ it is approximately 1.048, illustrating the slow approach to the normal baseline. A normalising window too close to the edge can bias the whole inferred curve downward. A comparison with a normal-state spectrum taken with the same junction can control a varying normal density or voltage-dependent transmission better than a single distant conductance point.

Derive the thermal convolution measured by a normal probe

In lowest-order tunnelling with a featureless normal tip and approximately constant matrix elements, the current is proportional to $\int N_s(E)[f(E-eV)-f(E)]\,dE$. The two Fermi factors count occupied initial and available final states. Differentiating in bias gives normalised conductance $G(V)/G_n = \int [N_s(E)/N_n][-f'(E-eV)]\,dE$. Here $-f'(x) = \operatorname{sech}^2[x/(2k_BT)]/(4k_BT)$ has unit integral. At zero temperature it becomes a delta function and the conductance reproduces the density. At finite temperature it samples a window: conductance can be nonzero inside the underlying hard gap without any actual subgap states.

The full width at half height of this derivative is $4k_BT\,\operatorname{arcosh}\sqrt{2} \approx 3.53k_BT$. At $T = 2.0\,\mathrm{K}$ this is $0.608\,\mathrm{meV}$, appreciable compared with a $1.50\,\mathrm{meV}$ gap. The electron temperature at the tunnel contact can exceed the temperature reading of a nearby bath. Voltage noise and a finite lock-in excitation add other convolution kernels. A fit that varies gap, thermal width and lifetime broadening all freely against a single narrow window can have strongly correlated parameters. Measure an independent thermal reference or repeat with deliberately varied excitation to separate instrumental from material broadening.

Compute the ideal convolution with an edge-safe quadrature

Numerically sampling the divergent density at one grid point exactly on the edge gives an artificial infinity. A variable transformation removes it. For $E > \Delta$ set $\xi = \sqrt{E^2 - \Delta^2}$. Then $(E/\sqrt{E^2-\Delta^2})\,dE = d\xi$. Integrate over positive ξ and include the negative-energy mirror:

$$G/G_n = \int_0^\infty [K_T(\sqrt{\xi^2+\Delta^2} - eV) + K_T(-\sqrt{\xi^2+\Delta^2} - eV)]\,d\xi.$$

Both terms are smooth at zero for positive temperature. This numerical reformulation follows directly from state counting.

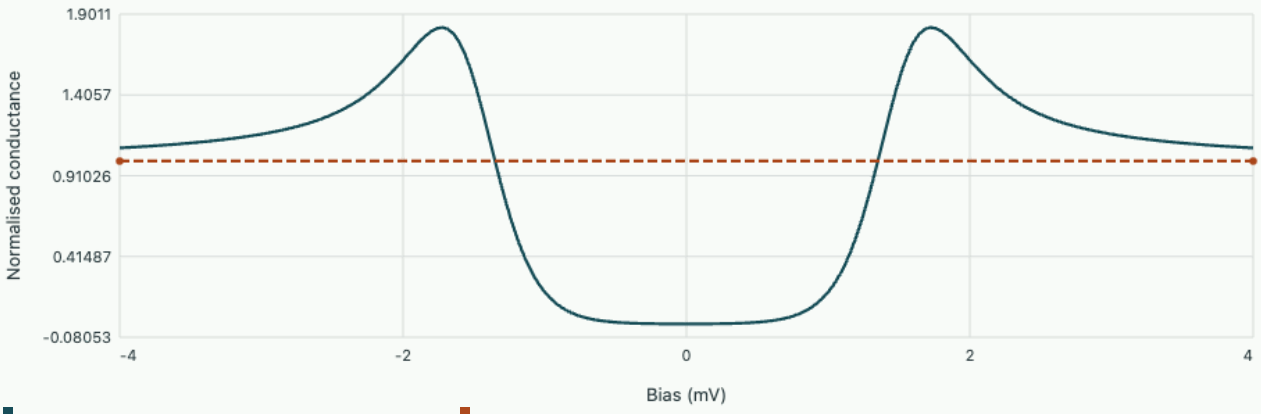
The live spectrum uses this transformed integral with energies in meV and voltages in mV for a single electron. First choose a low but finite temperature and locate the two coherence features. Increase temperature at fixed Δ to isolate the effect of the measurement kernel. A physically heated sample would generally also have a temperature-dependent gap; this controlled calculation keeps the intrinsic spectrum explicit. Compare the zero-bias conductance, peak height and high-bias baseline. A change in one feature alone seldom establishes the source of broadening.

Open this experiment online

Try the calculation

Gap Δ (meV) 1.5

Electron temperature (K) 2



Thermally broadened tunnelling Normal-state reference

Thermal kernel FWHM 0.608 meV. Zero-bias conductance is 0.0013 of the normal value. Gap and probe temperature vary independently.

Identify a spectral asymmetry that the symmetric model cannot fit

For a featureless normal tip and a particle–hole symmetric superconducting spectrum, the convolution with the even thermal kernel gives

$$G(V) = G(-V).$$

The average of opposite bias measurements estimates the even component, while half their difference isolates an odd component. A small bias-offset calibration error V_0 produces an apparent odd part approximately $-V_0\, dG_{\mathrm{even}}/dV$ for small shifts. It is therefore largest on steep gap-edge flanks and changes sign when the bias offset reverses. An energy-dependent normal background can generate a smoother intrinsic asymmetry; a temperature sweep through the transition and an independent voltage-zero calibration help distinguish these mechanisms.

Suppose the spectral flanks have slope $1.8 G_n/\mathrm{mV}$ and the measured odd part peaks near $0.036 G_n$. A voltage offset of only $0.020\,\mathrm{mV}$ could generate that size, much smaller than the illustrative gap. Comparing the sign pattern on both flanks tests the derivative shape rather than fitting one data point. A single shifted peak position can reflect offset, thermal broadening or changing gap amplitude. Record both bias directions, both coherence peaks and their separation as well as their midpoint. A common shift of both peaks changes the midpoint but ideally leaves their separation unchanged.

The model also assumes a normal electrode with roughly flat density. If the tip itself is superconducting, the current is a convolution of both electrode densities and their occupation difference. In that case symmetric features can shift even when the sample gap is unchanged, requiring an independent tip calibration or a different junction model.

Tutorial on tunnelling data

An ideal normal-tip spectrum at very low temperature has gap edges separated by $3.20\,\mathrm{mV}$. Estimate Δ, and compare with a weak-coupling transition temperature using $2\Delta/(k_B T_c) \approx 3.53$.

▼ **Work through the solution**

The symmetric single-particle edges lie near $\pm\Delta/e$, hence $\Delta \approx 1.60\,\mathrm{meV}$. Inserting $k_B = 0.08617\,\mathrm{meV/K}$ gives $T_c \approx 10.5\,\mathrm{K}$. This ratio is a weak-coupling benchmark for the isotropic model, not an exact conversion for every material. Verify the actual transition and junction type independently.

An experiment at $4.2\,\mathrm{K}$ tries to distinguish two gaps differing by $0.10\,\mathrm{meV}$ using peak positions alone. Which scale should be compared first?

▼ **Work through the solution**

The thermal derivative width is $3.53(0.08617)(4.2) \approx 1.28\,\mathrm{meV}$, much larger than that gap difference. High signal quality and a constrained model can sometimes infer parameters finer than a raw linewidth, but separate visible peaks are unlikely from thermal resolution alone. Lower electron temperature and calibrated voltage modulation add stronger discriminating evidence.

What check distinguishes a quadrature error from true thermal subgap conductance?

▼ **Work through the solution**

Refine the transformed integration grid, extend its energy cutoff and confirm the conductance is even in voltage for this symmetric model. A normal reference with $\Delta = 0$ should yield unity at every bias when the kernel is fully integrated. The computed zero-bias value for positive gap should fall as temperature decreases. A subgap signal sensitive to whether an energy grid happens to land on an edge is a numerical artifact.

The University of Zurich solid-state course develops superconducting excitations and their physical consequences.

Structure, excitations and probes of solids · 9. Read twist and mismatch from the moiré of two lattices

Stack two square grids that differ slightly in orientation or pitch. The slowly varying registry map is visible over many unit cells. Derive its scale and direction in reciprocal space, and find which combinations of layer parameters a single period measurement can resolve.

Subtract two spatial phases

One family of planes in layer one carries a density modulation proportional to $\cos(\mathbf{G}_1 \cdot \mathbf{r})$, with $|\mathbf{G}_1| = 2\pi/a$. Layer two has spacing $a(1+\delta)$ and is rotated by θ. Multiplying their modulations, as in a local registry correlation, produces half the sum of two cosines. The rapidly varying term contains $\mathbf{G}_1 + \mathbf{G}_2$; averaging over several original cells leaves the term with $\Delta\mathbf{G} = \mathbf{G}_1 - \mathbf{G}_2$. A distance parallel to $\Delta\mathbf{G}$ that advances the resulting phase by 2π spans one full registry period.

Taking the dot product of the rotated reciprocal vectors gives the exact geometric scale

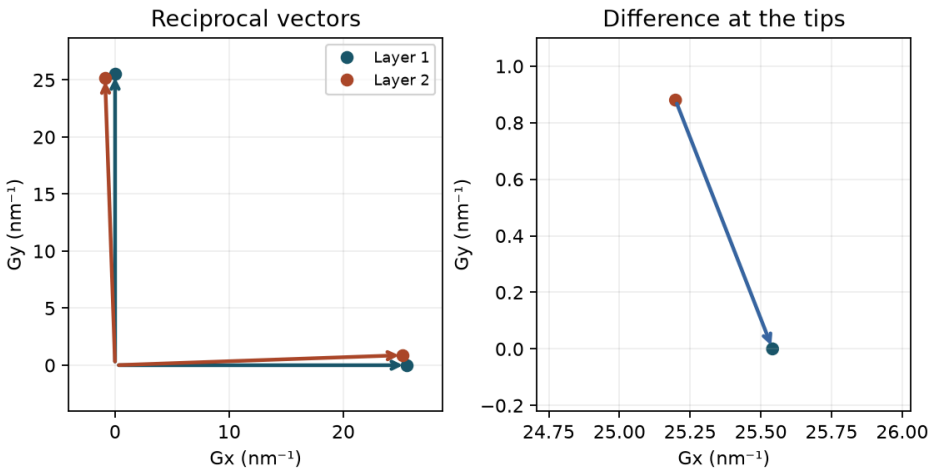
$$\frac{\Lambda}{a} = \left[1 + \frac{1}{(1+\delta)^2} - \frac{2\cos\theta}{1+\delta}\right]^{-1/2}.$$

The second perpendicular family of a square lattice rotates and rescales the same way. Its difference vector is perpendicular to the first and has the same magnitude. Their two slow cosines generate a square registry pattern. For small mismatch and small twist measured in radians the period is approximately $a/\sqrt{\delta^2+\theta^2}$. The interference construction specifies a local continuum period. An exact common atomistic supercell would require additional commensurability conditions.

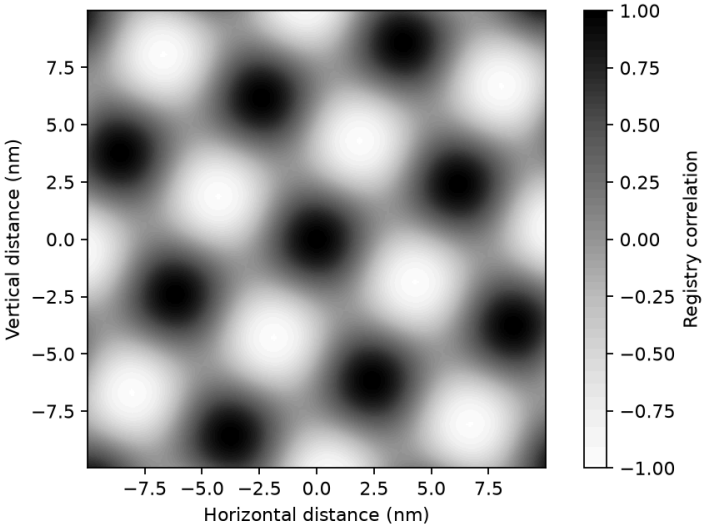
Interpret a long measured period

Take $a = 0.246$ nm, mismatch 1.3 percent and twist two degrees. In units of $2\pi/a$ the horizontal family's difference vector is $(1-\cos\theta/(1+\delta), -\sin\theta/(1+\delta))$, approximately $(0.0134, -0.0345)$. Its length is about 0.0370, so the registry period reaches about 6.65 nm. Each layer still has a spacing near a quarter nanometre. The larger scale is the distance over which small phase differences accumulate.

If an image measures only this period and the first layer spacing, it supplies a single scalar for the two unknowns δ and θ. A reciprocal-space measurement referenced to the first lattice supplies direction as well. In the simple isotropic model this resolves the rotated and rescaled vector by $\mathbf{G}_2 = \mathbf{G}_1 - \Delta\mathbf{G}$. Additional distortions or strain require more directions before an inversion is possible.



The right panel enlarges the displacement between reciprocal-vector tips. Its direction is the normal to one family of registry fringes.



The map sums the two perpendicular slow modulations; the colour bar measures their dimensionless average. At perfect alignment with equal spacing the registry is constant everywhere.

Exercise · determine a twist from pure rotation

The lattice spacings have been verified equal at 0.246 nm. What is the magnitude of the twist if the registry spacing is 10.0 nm?

▼ **Work through the solution**

With $\delta = 0$, the reciprocal vectors form an isosceles triangle and $\Lambda = a/[2|\sin(\theta/2)|]$. The small positive solution is $2\arcsin(0.246/20.0) = 0.02460$ rad or 1.41 degrees. Changing the twist sign reflects the pattern without changing its period.

Exercise · make the same spacing with no twist

Use the same first-layer spacing and measured 10.0 nm period. What positive mismatch would explain the observation for aligned layers?

▼ **Work through the solution**

At $\theta = 0$ and positive mismatch,

$$\Lambda = a(1+\delta)/\delta.$$

Solving gives $\delta = a/(\Lambda - a) = 0.02522$, about 2.52 percent. Comparing this result with the previous exercise shows explicitly why period alone cannot identify a twist.

Try it · Moire

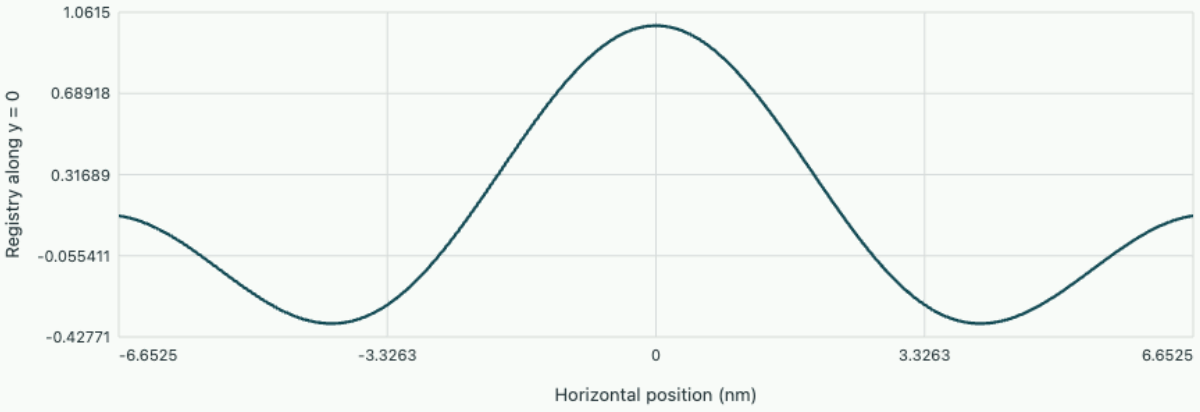
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Lattice mismatch **0.013**

Relative twist (degrees) **2**



Two perpendicular registry modulations

Registry period 6.6525 nm; reciprocal mismatch magnitude 0.94448 nm⁻¹.

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

Structure, excitations and probes of solids · 10. From bonding energy to lattice spacing and sound speeds

Diffraction gives a spacing and ultrasound gives a stiffness. One measures where an energy curve is stationary, the other measures its curvature. Building the simplest energy of an ionic solid gives both observables and shows which quantities a spring model has already assumed.

Decide which degrees of freedom contribute to binding

In an ionic crystal the leading energy is the electrostatic interaction of charged sublattices. In a covalent solid electron sharing favours specific bond directions, as in a tetrahedrally coordinated network. In a metal the mobile electronic states extend across the lattice and contribute an energy that depends on occupation and interatomic spacing. Closed-shell atoms can attract through fluctuating dipoles even when they have no permanent dipole. At small separations overlapping electronic clouds incur a rapidly rising energy. These mechanisms can coexist. No single pair potential reproduces every elastic or optical property.

For an alternating line of charges separated by r, the electrostatic potential at a positive ion receives two contributions at each distance jr. Summing nearest positive and negative shells in neutral groups gives the dimensionless lattice sum $2(1 - 1/2 + 1/3 - \cdots) = 2\ln 2$. In three dimensions the corresponding Madelung constant depends on structure. Neutral summation or a convergent Ewald construction is needed for a Coulomb lattice. A naive rapidly growing box of uncompensated charges can give a spurious result.

Find the energy minimum and infer the short-range repulsion

For one formula unit of a 1:1 ionic lattice take $\Phi(r) = -A/r + B/r^n$ with $A = \alpha e^2/(4\pi\epsilon_0)$ and $n > 1$. Counting pair energy once gives this coefficient per monovalent ion pair. The stationarity condition at measured nearest-neighbour spacing r_0 eliminates B. The cohesion and curvature at the same point become

$$B = \frac{A r_0^{n-1}}{n},$$

$$\Phi(r_0) = -\frac{A}{r_0}\left(1 - \frac{1}{n}\right),$$

$$\Phi''(r_0) = \frac{(n-1)A}{r_0^3}.$$

The energy zero here is separated ions. Getting the cohesive energy relative to separated neutral atoms additionally requires the ionisation and electron-affinity terms. For a rock-salt structure $\alpha = 1.7476$ and the volume per formula unit is $v = 2r^3$. At zero pressure $K = v\,d^2\Phi/dv^2 = v\Phi''/(dv/dr)^2$. Substitution gives

$$K = \frac{\Phi''(r_0)}{18 r_0} = \frac{(n-1)A}{18 r_0^4}.$$

For $r_0 = 0.282$ nm, $n = 9$ and $e^2/(4\pi\epsilon_0) = 1.440$ eV nm, the model gives $A = 2.517$ eV nm, binding relative to separated ions of 7.93 eV per pair, and $K \simeq 28.3$ GPa. Inverting the curvature using an independently measured modulus tests the assumed repulsive exponent.

Recall · Radial curvature and a bulk modulus have different units

The curvature of one pair energy has units N/m. An elastic modulus has units N/m² and refers to energy per volume. The volume per formula unit and the density of bonds determine the conversion.

Resolve the stiffness tensor using three acoustic polarisations

For a crystal write $\sigma_{ij} = C_{ijkl}\varepsilon_{kl}$. Substitute a displacement $\mathbf{u} = \mathbf{e}\, e^{i(k\mathbf{n}\cdot\mathbf{x}-\omega t)}$ into $\rho\ddot{u}_i = \partial_j\sigma_{ij}$. The strain symmetry gives a symmetric three-dimensional eigenproblem. Its eigenvectors give mode polarisations and its eigenvalues give measurable phase speeds.

$$\Gamma_{ik}e_k = \rho c^2 e_i,$$
$$\Gamma_{ik} = C_{ijkl}n_j n_l.$$

Cubic symmetry leaves C_{11}, C_{12}, C_{44}. Along [100] the eigenvalues are C_{11}, C_{44}, C_{44}. Along [110] they are $(C_{11}+C_{12}+2C_{44})/2$, $(C_{11}-C_{12})/2$, and C_{44}. Under hydrostatic load $K = (C_{11}+2C_{12})/3$. A bulk modulus from ionic compression supplies just one combination. Shear-wave data probe independent curvature directions.

OneSimulation experiment · Three acoustic branches in a cubic crystal

Try a three-dimensional acoustic-speed surface

Rotate the 3D surface to compare propagation along an edge, a face diagonal and a body diagonal. Its radius in a direction is the selected phase speed in that direction. The 2D plot shows the three sorted eigenvalues in the (001) plane; the branch selector changes which one is drawn in 3D.

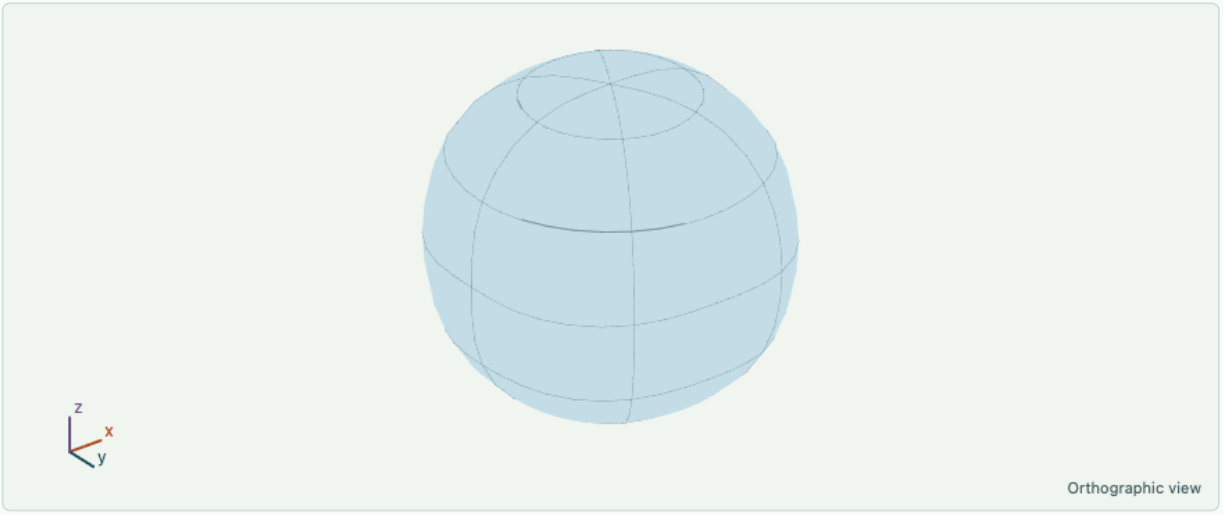
Set C44 to 51 GPa to satisfy C11 − C12 = 2C44. Predict the shape of the surface before moving the slider. Then multiply density by four and check the predicted factor of two in speed. C11 = 166 GPa and C12 = 64 GPa stay fixed. Every surface point solves the symmetric 3 × 3 Christoffel eigenproblem with a browser Jacobi iteration. The full-direction notebook below independently computes the entire surface with OneSimulation, including its [100], [110] and [111] cuts and an isotropic-limit check.

The surface gives phase velocity. In an anisotropic crystal the energy-flow direction can differ from the wave-normal direction, so this surface is not a point-source wavefront.

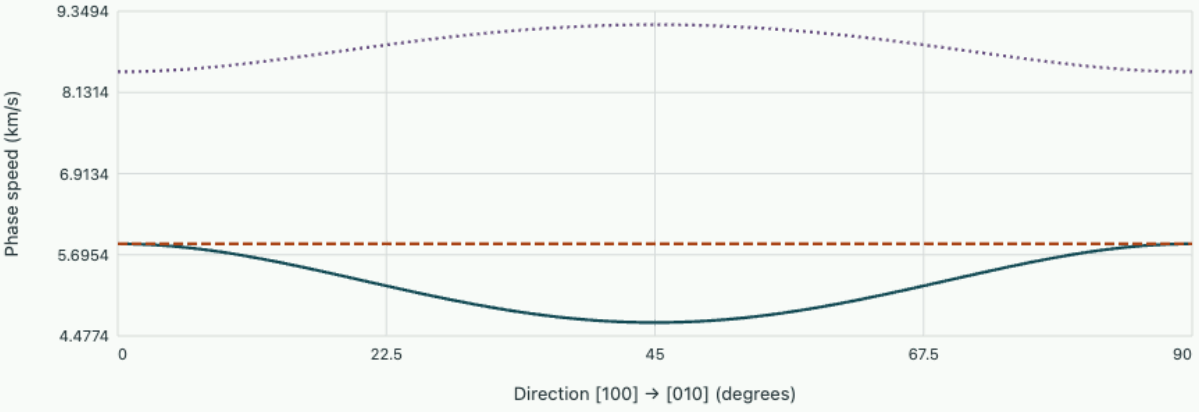
Open this experiment online

Try the calculation

Shear stiffness C44 (GPa)	**80**
Density (kg/m³)	**2330**
3D branch: 0 slow, 1 middle, 2 fast	**2**



Three-dimensional phase-speed surface: in each propagation direction n, the radius is the selected eigenvalue's speed v(n). Coordinates are v(n)n in km/s. This is a material property surface, not a deformed crystal or a group-velocity wavefront. Axes: x = vx (km/s); y = vy (km/s); z = vz (km/s).



C11 = 166 GPa, C12 = 64 GPa. [100] speeds: 5.85959, 5.85959, 8.44065 km/s; [110]: 4.67851, 5.85959, 9.14828 km/s. C44 = 51 GPa makes the material isotropic. The 3D view shows the fast branch.

Download the OneSimulation model

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

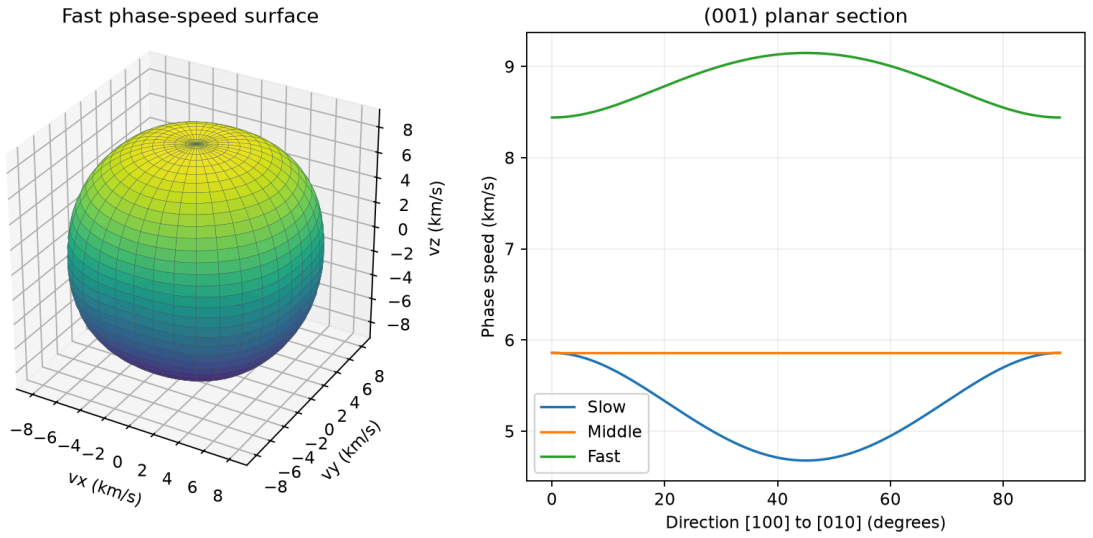
```
python complete-physics-cubic-surface.py --c44 80 --density 2330 --branch 2
```

Full three-dimensional OneSimulation result

Download the executed 3D notebook · Download the parameterised 3D script. The script command above follows all three controls. Every surface vertex solves the Christoffel eigenproblem in its own propagation direction.

```
from onesim.numerics.linear.matrices.dense_linear_algebra import symmetric_eigenvalues
# n is a unit propagation direction; stiffnesses are in GPa.
gamma = (64. + c44) * np.outer(n, n)
gamma += np.diag(c44 + (102. - 2*c44) * n**2)
velocity = np.sqrt(symmetric_eigenvalues(gamma) * 1e9 / density) / 1000
surface_point = velocity[branch] * n
```

```
[100] 5.859587 5.859587 8.440654 km/s
[110] 4.678505 5.859587 9.148278 km/s
[111] 5.102666 5.102666 9.372287 km/s
Isotropic directional invariance: verified
```



Executed OneSimulation output for C44=80 GPa and density 2330 kg/m³. The surface samples all polar and azimuthal directions; the planar curves provide a reproducible section through it.

Reproduce with OneSimulation

Rotate the propagation direction without changing density or stiffness. Predict which two modes coincide along a cube edge and compare the computed [110] speeds with the three analytic eigenvalues. Then impose the isotropic condition.

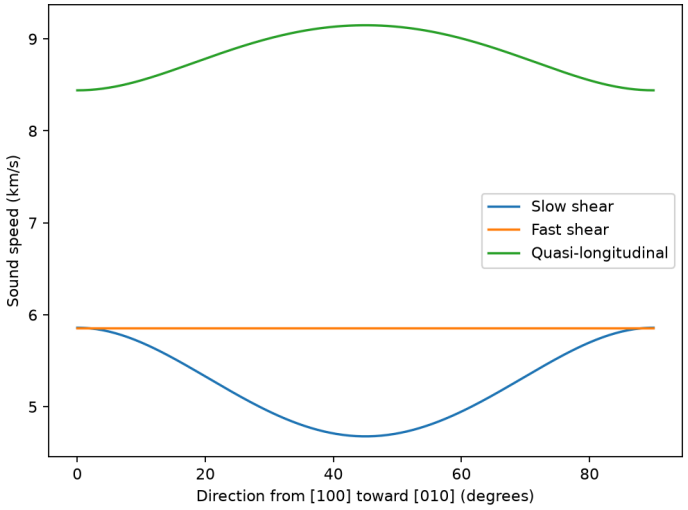
Run the downloaded notebook with OneSimulation, NumPy and Matplotlib installed. The output below was recorded with the displayed calculation.

```
"""Sound velocities when a wave rotates through the (001) plane of a cubic solid."""
import numpy as np
import matplotlib.pyplot as plt
from onesim.numerics.linear.matrices.dense_linear_algebra import symmetric_eigenvalues

c11, c12, c44 = np.array([166, 64, 80]) * 1e9  # Pa
density = 2330.0  # kg/m^3
```

```
def speeds(angle):
    n = np.array([np.cos(angle), np.sin(angle), 0])
    gamma = (c12+c44) * np.outer(n, n)
    gamma += np.diag(c44 + (c11-c12-2*c44)*n**2)
    return np.sqrt(symmetric_eigenvalues(gamma)/density)

angle = np.linspace(0, np.pi/2, 181)
velocity = np.array([speeds(a) for a in angle]) / 1000
print("Wave speeds in km/s, sorted by magnitude")
for label, a in [("[100]", 0), ("[110]", np.pi/4)]:
    print(label, " ".join(f"{v/1000:.3f}" for v in speeds(a)))
for branch, label in enumerate(["Slow shear", "Fast shear", "Quasi-longitudinal"]):
    plt.plot(np.degrees(angle), velocity[:,branch], label=label)
plt.xlabel("Direction from [100] toward [010] (degrees)")
plt.ylabel("Sound speed (km/s)")
plt.legend()
plt.tight_layout()
plt.savefig("physics-cubic-sound.png", dpi=160)
```

Computed output

```
Wave speeds in km/s, sorted by magnitude
[100] 5.860 5.860 8.441
[110] 4.679 5.860 9.148
```



Three acoustic branches in a cubic crystal.

Download editable notebook · Download Python script

Tutorial on identifying a binding law

1 · Recognise a false equilibrium

▼ **Work through the solution**

With only $-A/r$ the energy falls monotonically as ions approach. A measured lattice spacing and a real sound speed require an energy minimum with positive curvature. Attraction alone supplies neither.

---

2 · Recover one modulus from a transverse mode

▼ **Work through the solution**

At density 2330 kg/m³ an out-of-plane shear mode along [110] travels at 5.86 km/s. Hence $C_{44} = \rho c^2 \simeq 80.0$ GPa. The slower in-plane shear mode instead measures $(C_{11} - C_{12})/2$.

---

3 · Decide whether an isotropic model could fit both shear waves

▼ **Work through the solution**

Cubic shear degeneracy for all directions requires $C_{11} - C_{12} = 2C_{44}$. The notebook uses 102 versus 160 GPa, so a scalar shear modulus cannot fit both [110] speeds. Alter C_{44} to 51 GPa and predict the angular plot before rerunning.

For further study see Kittel, Introduction to Solid State Physics, 8th edition, chapters 1–5 on crystal structure, binding, elastic constants and acoustic lattice modes.

Structure, excitations and probes of solids · 11. Collective charge oscillations, bound excitons and lattice-dressed electrons

A loss peak in an electron spectrum, a line below an optical gap and an enhanced carrier mass call for different kinds of excitations. Start each model with its moving charges and its restoring interaction. That identifies what a measurement can determine.

Displace an electron fluid against its background

Consider a wide slab with carrier number density n, band mass m^* and immobile neutralising positive background. Displace the electron cloud uniformly through a small distance u normal to the faces. The opposite surface charges have magnitude neu per area and generate a restoring field $E = neu/(\epsilon_0\epsilon_\infty)$ inside. The factor ϵ_∞ accounts for faster electronic degrees of freedom excluded from the carrier model.

$$m^*\ddot{u} = -\frac{ne^2}{\epsilon_0\epsilon_\infty}u,$$

$$\omega_p^2 = \frac{ne^2}{m^*\epsilon_0\epsilon_\infty}.$$

In a homogeneous bulk the same long-wavelength mode follows from continuity, carrier momentum and Gauss's law for a longitudinal density fluctuation. A longitudinal self-sustaining oscillation occurs at a zero of the dielectric function. With damping time τ the relative permittivity in this model is $\epsilon(\omega) = \epsilon_\infty[1 - \omega_p^2/(\omega(\omega + i/\tau))]$ for $e^{-i\omega t}$. Electron energy loss is sensitive to $-\mathrm{Im}(1/\epsilon)$, while transmission also depends on interfaces.

For $n = 10^{24}$ m⁻³, $m^* = 0.20m_e$ and $\epsilon_\infty = 12$, the model gives $\omega_p \simeq 3.64 \times 10^{13}$ rad/s, or 5.80 THz. Four times the density doubles this frequency if mass and background screening are unchanged.

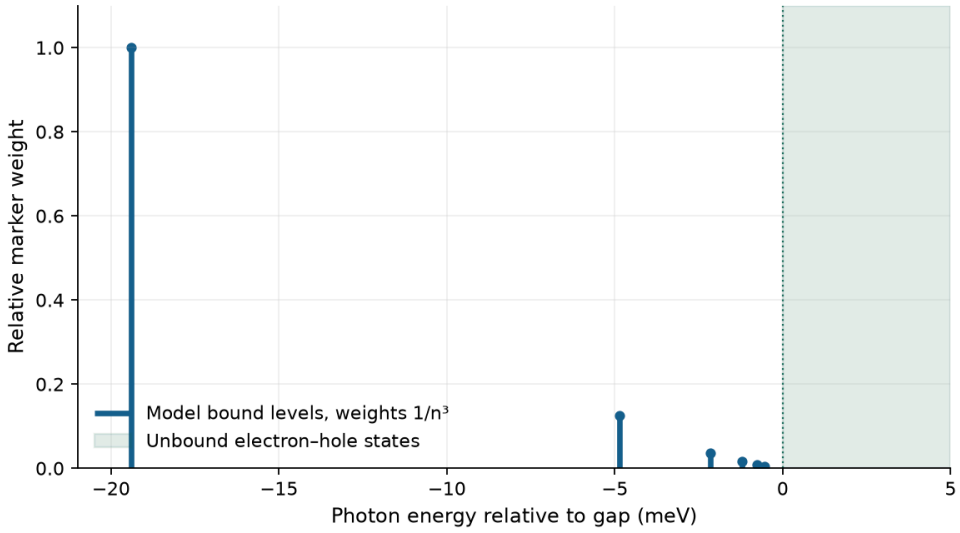
Solve an effective electron–hole relative-coordinate problem

In a three-dimensional direct-gap semiconductor a photon can create a conduction electron and a valence hole. Separate their centre of mass from their relative coordinate. Near parabolic band edges the relative mass obeys $\mu^{-1} = 1/m_e^* + 1/m_h^*$. Approximate screening by a constant ϵ_r over the length and time scales of the pair. The relative Hamiltonian is then hydrogenic.

$$H_{\mathrm{rel}} = -\frac{\hbar^2\nabla^2}{2\mu} - \frac{e^2}{4\pi\epsilon_0\epsilon_r r},$$

$$E_n = E_g - \frac{R^*}{n^2}.$$

Minimising the expectation for an exponential radius a gives kinetic energy $\hbar^2/(2\mu a^2)$ and potential $-e^2/(4\pi\epsilon_0\epsilon_r a)$. The minimum yields $a^* = (\epsilon_r m_e/\mu)a_0$ and $R^* = (\mu/m_e)13.606\,\mathrm{eV}/\epsilon_r^2$. With $m_e^* = 0.20m_e$, $m_h^* = 0.50m_e$, $\epsilon_r = 10$, these give $\mu = 0.1429m_e$, $a^* = 3.70$ nm and $R^* = 19.4$ meV. The most strongly bound line lies this distance below the unbound continuum. A large radius relative to the unit cell supports the effective-mass model. A sub-gap peak alone can also come from a defect.



A hydrogenic bound series with a 19.4 meV exciton Rydberg approaches the continuum from below. Marker weights use the illustrative direct-transition envelope law proportional to the inverse cube of the level index. Broadening is omitted.

Try it · Exciton

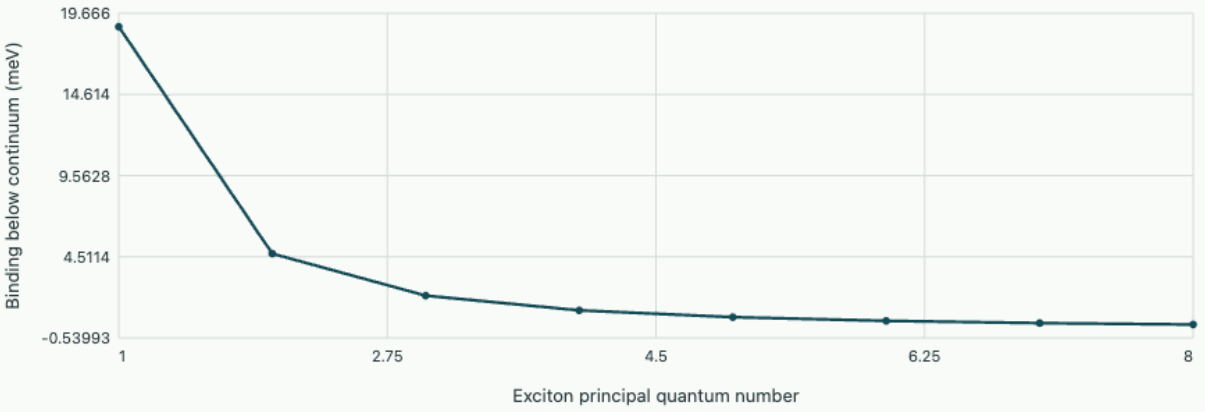
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Reduced electron-hole mass / electron mass **0.1**

Relative permittivity **8.5**



Hydrogenic exciton series

Effective Rydberg 18.831 meV; effective Bohr radius 4.498 nm. Effective-mass and dielectric screening approximations require a sufficiently extended exciton.

Let a single carrier relax a lattice coordinate

Take a local displacement Q_i, stiffness K, and local occupation n_i. Each site's lattice energy and coupling are $KQ_i^2/2 - gQ_i n_i$. At a fixed occupation distribution minimise over each displacement. In the adiabatic classical distortion model,

$$Q_i = g n_i / K,$$

$$E_{\mathrm{relax}} = -\frac{g^2}{2K} \sum_i n_i^2.$$

A carrier with all its weight on one site gains $g^2/(2K)$. With uniform weight across N sites the gain is smaller by N. Hopping competes with this localisation energy. For $g = 2.0$ eV/nm and $K = 20$ eV/nm² the one-site relaxation is 0.10 eV and the optimum displacement 0.10 nm. Compare with the bandwidth

before concluding that a small polaron will form. Quantised phonons, tunnelling of the distortion and long-range polar coupling are additional physics.

> **Recall · Screening has a frequency scale**
>
> A bound pair and a fast plasma oscillation need not see the same dielectric constant. Identify which lattice and electronic modes can follow each perturbation before importing a tabulated static permittivity.

Tutorial on assigning a measured feature

1 · Extract an exciton binding scale from two lines

▼ **Work through the solution**

In the hydrogenic series $E_2 - E_1 = 3R^*/4$. A splitting of 12 meV gives $R^* = 16$ meV and continuum $E_g = E_1 + 16$ meV. Check another member or the absorption onset to distinguish an accidental pair of defect peaks.

---

2 · Separate two density dependencies

▼ **Work through the solution**

A mode increasing from 5.8 to 8.2 THz as the carrier density doubles follows the plasmon square-root trend. The simple isolated exciton model has no such density law. At high density screening or many-body effects can also modify the exciton spectrum.

---

3 · Test the localisation argument

▼ **Work through the solution**

For equal occupation of two sites the distortion energy gain is half the one-site gain. A two-site bonding state can lower the electronic kinetic energy by a hopping energy t. If t is large compared with the difference in relaxation energies, localisation cannot be inferred from the static distortion gain alone.

For further study see Kittel, Introduction to Solid State Physics, 8th edition, chapters 14–15 on collective excitations, polarons and optical response.

Structure, excitations and probes of solids · 12. Ferromagnetic resonance from angular momentum and free-energy curvature

A uniformly magnetised thin specimen can absorb microwave power at a frequency far from the bare spin Larmor frequency. The precessing magnetic moment creates its own field. Its geometry and the anisotropy determine which restoring torques appear in the measured resonance.

Linearise the torque around the equilibrium orientation

Consider a saturated, uniformly magnetised ellipsoid with equilibrium along x. Write the dimensionless unit magnetisation as $\mathbf{m} \simeq (1, m_y, m_z)$. To second order use a free energy density $F = F_0 + M_s(B_y m_y^2 + B_z m_z^2)/2$ with both curvature fields positive in tesla. They include bias, demagnetising fields and anisotropy. A small applied microwave field $b_y(t)$ adds $-M_s b_y m_y$. Define the effective field by $-\partial F/\partial \mathbf{m}$ divided by M_s.

$$\dot{\mathbf{m}} = -\gamma\, \mathbf{m} \times \mathbf{B}_{\mathrm{eff}},$$

$$\dot{m}_y = -\gamma B_z m_z,$$

$$\dot{m}_z = \gamma B_y m_y - \gamma b_y.$$

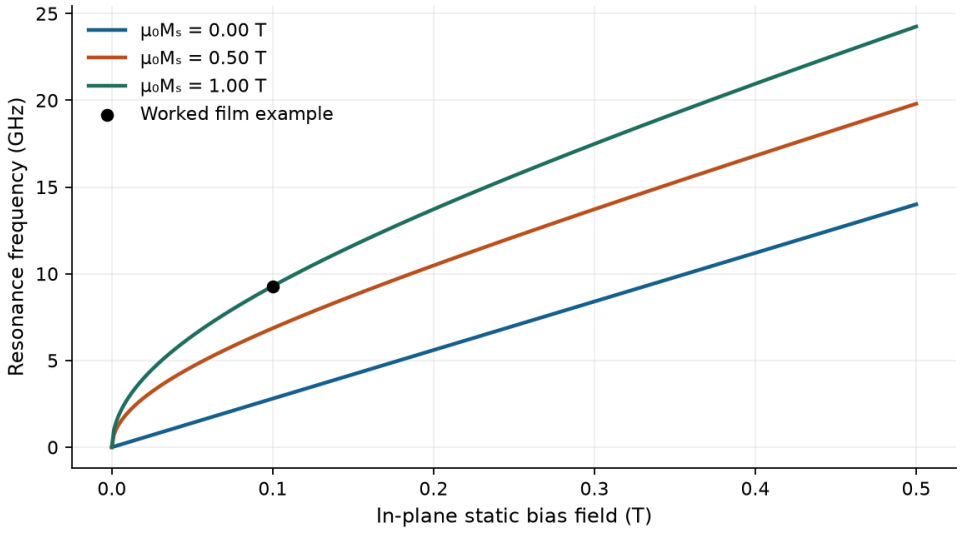
Differentiating the first equation and eliminating m_z gives $\ddot{m}_y + \gamma^2 B_y B_z m_y = \gamma^2 B_z b_y$. The free precession frequency is $\omega_0 = \gamma\sqrt{B_y B_z}$. The orbit generally has unequal semiaxes. In the limit of equal curvature fields the circular bare-spin result is recovered.

Obtain the thin-film frequency in an in-plane field

In a wide thin film, uniform in-plane rotation creates little demagnetising field. An out-of-plane tilt costs approximately $\mu_0 M_s^2 m_z^2/2$. A bias B_0 contributes $-M_s B_0 m_x \simeq -M_s B_0 + M_s B_0(m_y^2 + m_z^2)/2$. Neglect crystalline anisotropy and edge corrections. Thus $B_y = B_0$ and $B_z = B_0 + \mu_0 M_s$. With $\gamma/(2\pi) = 28$ GHz/T, $\mu_0 M_s = 1.00$ T and $B_0 = 0.10$ T we predict $f_0 = 9.29$ GHz. Using $28 B_0$ alone would give 2.80 GHz.

Recall · Compare B-like and H-like parameters consistently

In these formulas all stiffness fields are in tesla and γ is in radians per second per tesla. If a field is in A/m it must first be converted where appropriate with μ_0. The static demagnetising field depends on shape as well as material.



The bare-spin straight line is recovered when the magnetisation contribution to the restoring field vanishes. An in-plane saturated film develops a curved frequency response.

Try it · Fmr

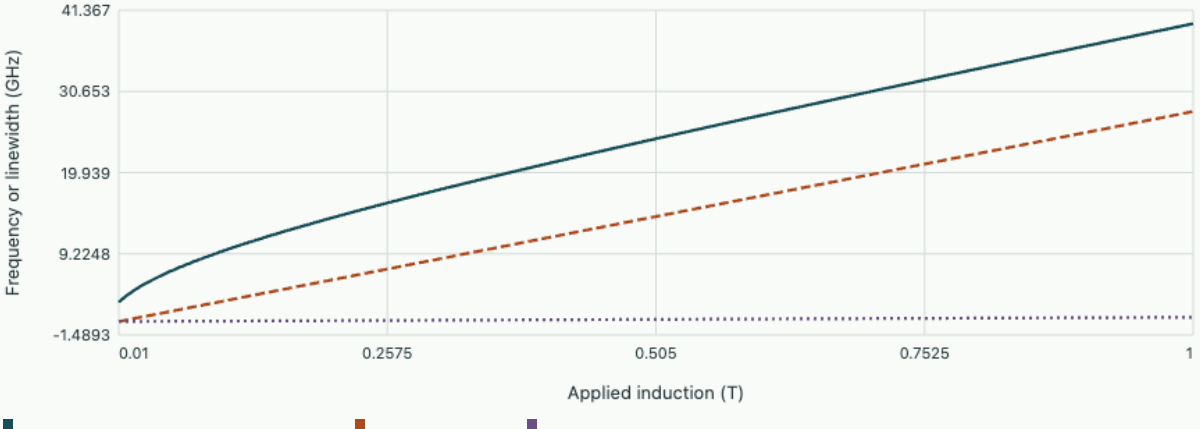
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Saturation induction μ0M (T) **1**

Gilbert damping α **0.01**



In-plane film resonance · Bare spin · Weak-damping FWHM

Gyromagnetic ratio / 2π = 28 GHz/T. Uniform saturated film, in-plane field, no extra anisotropy; linewidth is the weak-damping estimate.

Derive the absorbed-power peak and its width

Gilbert damping adds $+\alpha \mathbf{m} \times \dot{\mathbf{m}}$ to the torque equation. To first order in small α one obtains the damped denominator $\omega_0^2 - \omega^2 - i\alpha\gamma(B_y + B_z)\omega$. With convention $e^{-i\omega t}$ the average loss from a real drive of complex amplitude b is proportional to $\omega\ \mathrm{Im}(\chi)|b|^2/2$. Near resonance the absorptive susceptibility has a Lorentzian line shape. Its angular-frequency full width at half maximum to this order is

$$\Delta\omega_{\mathrm{FWHM}} \simeq \alpha\gamma(B_y + B_z).$$

For the model film at $B_0 = 0.10$ T and $\alpha = 0.010$ the frequency-swept width is about 0.336 GHz. A field-swept experiment must divide the frequency width by $|df_0/dB_0|$ and distinguish the width of the absorption peak from a derivative peak-to-peak linewidth. Inhomogeneous fields and mode overlap broaden real lines as well.

Tutorial on extracting magnetic parameters

1 · Recover a saturation magnetisation

▼ Work through the solution

For an in-plane film at 0.12 T, a uniform resonance at 10.0 GHz gives $\mu_0 M_s = (10/28)^2/0.12 - 0.12 = 0.943$ T in this approximation. Repeating the measurement at several fields tests the assumed geometry.

---

2 · Infer an intrinsic damping estimate

▼ Work through the solution

For curvature fields 0.20 and 1.20 T and a 0.392 GHz full width, $\alpha \simeq 0.392/[28(1.40)] = 0.0100$. A field-independent broadening that also contributes to the observed width would bias this single-frequency estimate upward.

---

3 · Explain a soft mode

▼ Work through the solution

As one curvature approaches zero at a loss of equilibrium stability, the predicted restoring frequency approaches zero while the other curvature can remain finite. A broad or shifted low-frequency peak calls for checking domains and the validity of the uniform-magnetisation assumption.

For further study see Kittel, Introduction to Solid State Physics, 8th edition, chapters 11–13 on collective magnetic response and magnetic resonance.

Structure, excitations and probes of solids · 13. Lattice dislocations and structure without long-range order

A missing atom, a slipped lattice plane and a glass can all broaden a measurement, through very different mechanisms. Geometry distinguishes a line defect from a local vacancy. Correlation functions make the distinction between short-range order and crystallinity quantitative.

Construct a Burgers circuit and the field of a screw dislocation

Follow a closed sequence of lattice steps around a suspected defect, well outside its core. Repeat the same sequence in a perfect reference lattice. The closure failure defines the Burgers vector $\mathbf{b}$ up to the agreed sense of the circuit and line orientation. Its component parallel to the dislocation line is screw character and its perpendicular component is edge character. Smooth elastic deformation without a dislocation gives no persistent closure failure.

For a straight screw line along z, with Burgers magnitude b along z, an elastic displacement outside the core is $u_z = b\phi/(2\pi)$ in cylindrical coordinates. After one revolution the displacement changes by b. The engineering shear is $\gamma_{\phi z} = r^{-1}\partial_\phi u_z = b/(2\pi r)$, giving $\sigma_{\phi z} = \mu b/(2\pi r)$. Integrating the energy density $\mu\gamma^2/2$ in an annulus gives an energy per unit line length

$$\frac{E}{L} = \int_{r_c}^{R} \frac{\mu}{2}\left(\frac{b}{2\pi r}\right)^2 2\pi r\, dr = \frac{\mu b^2}{4\pi}\ln(R/r_c).$$

The core radius limits continuum elasticity at small radius. Nearby defects or surfaces set the outer cutoff. For $\mu = 30$ GPa, $b = 0.25$ nm and $R/r_c = 10^4$, the elastic term is 1.37×10^{-9} J/m. A core energy must be added to infer formation energies.

Determine how an applied shear does work on a line

A dislocation of length L moving a distance dx sweeps area $L dx$. Across the swept area the material slips by b. A resolved shear stress τ supplies work $\tau b L dx$, hence the glide force per unit length is τb. At $\tau = 20$ MPa and $b = 0.25$ nm this gives 5.0×10^{-3} N/m. Obstacles and other dislocations oppose this driving force. Diffusion-assisted climb exchanges atoms or vacancies and depends on temperature in a different way.

> **Recall · Edge dislocations affect the number of planes**
>
> In an edge geometry the terminating extra half-plane ends at the dislocation line. A vacancy has no Burgers vector around a sufficiently large circuit. Concentration alone cannot describe a line's orientation, connectivity and mobility.

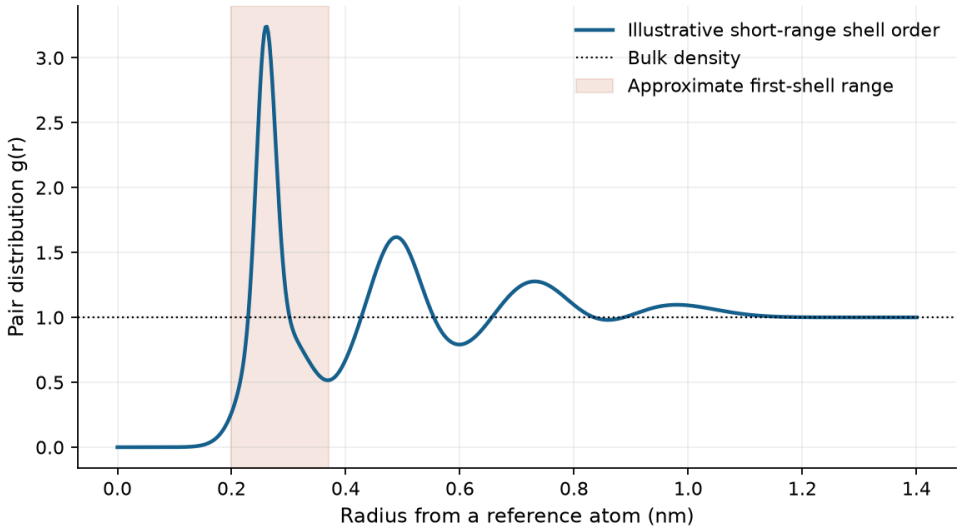
Recover pair order from scattering without indexing Bragg peaks

For particles of mean number density ρ, define $g(r)$ by the conditional count $dN = 4\pi r^2 \rho g(r)\, dr$ about an atom in an isotropic bulk sample. The first-shell coordination is the integral from the first-neighbour peak to its first minimum, including the correct r^2 weight. Scattering measures amplitudes before averaging the squared total. After averaging orientations the pair phase factor becomes $\sin(qr)/(qr)$.

$$S(q) = 1 + 4\pi\rho \int_0^\infty [g(r) - 1] \, \frac{\sin(qr)}{qr} \, r^2 dr.$$

A glass may have a sharp first-neighbour peak and further damped shells with $g(r) \to 1$. This permits strong local coordination without periodic translational order. An ideal crystal produces Bragg peaks in reciprocal space. A limited sample size or instrumental resolution also broadens crystalline peaks. Compare several peak widths and a real-space correlation length before assigning amorphous structure.

A glass transition is observed when a structural relaxation time becomes comparable to a measurement or cooling time. A slower scan permits structural rearrangement to lower temperatures. The measured transition temperature may therefore shift with rate. The sharp equilibrium melting point of a pure crystal and the observed glass-transition interval answer different experimental questions.



Illustrative pair-correlation profile for an isotropic disordered solid. The shaded first shell gives a coordination number after weighting the curve by the spherical shell volume and density.

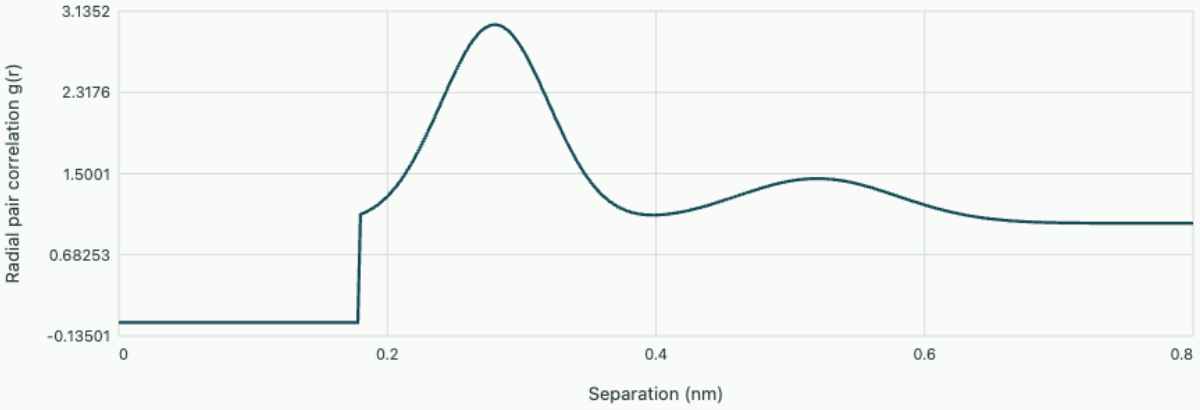
Try it · Pair order

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

First-shell peak height	2
First-shell width (nm)	0.04
Number density (nm^{-3})	40



Illustrative disordered-solid pair correlation

Coordination integrated over 0.18–0.40 nm = 17.88. This adjustable phenomenological profile demonstrates shell-volume weighting; it is not a fitted material structure.

Tutorial on disorder measurements

1 · Identify a dislocation character

▼ Work through the solution

A line runs along z while the closure failure points along x. Its screw projection is zero, so the line is pure edge. An inclined segment on a curved line generally has mixed character even with the same conserved Burgers vector.

2 · Estimate a coordination number

▼ **Work through the solution**

Approximate a first shell by $g = 3$ between 0.23 and 0.27 nm with $\rho = 60$ nm⁻³ and zero elsewhere within that shell. Then $N = 4\pi\rho[0.27^3 - 0.23^3] \simeq 5.67$. Truncating below the first minimum or failing to subtract scattering backgrounds changes the estimate.

---

3 · Separate line broadening from amorphous order

▼ **Work through the solution**

A single broad powder maximum is insufficient. Examine whether higher peaks can be indexed on a consistent lattice, vary instrumental resolution, and estimate the range of oscillatory pair correlations. Compare thermal evolution and cooling-rate dependence for independent evidence.

For further study see Kittel, Introduction to Solid State Physics, 8th edition, chapters 19–21 on amorphous structure, point defects and dislocations.

Structure, excitations and probes of solids · 14. Why some alloys demix while others order on two sublattices

A binary alloy can cease to look random by separating into compositionally distinct regions or by arranging unlike atoms on alternating sites. A free energy for composition answers the first question. A second order parameter is needed for the second.

Count configurations and derive an energy of mixing

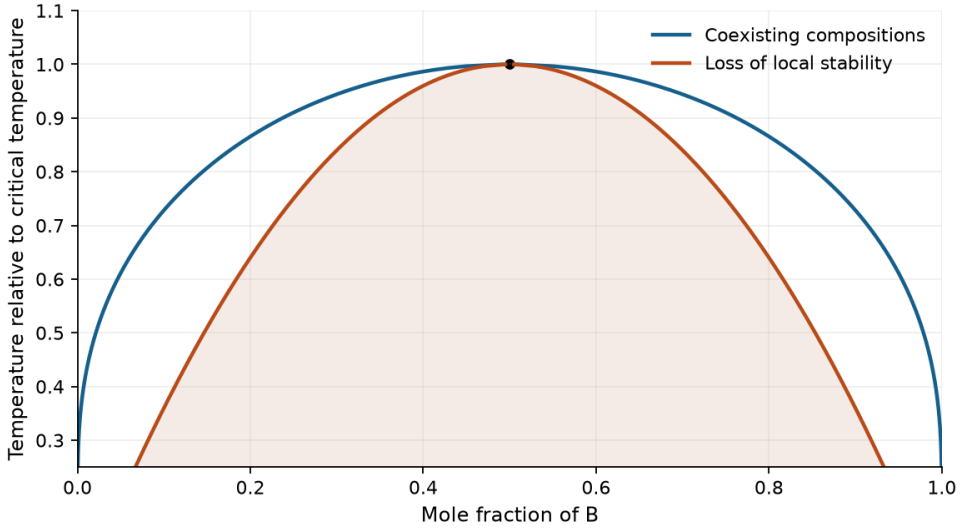
Let N lattice sites contain Nx atoms of B. The number of random arrangements is $N!/[(Nx)!(N(1-x))!]$. Stirling's approximation gives a configurational entropy per site $-k_B[x\ln x + (1-x)\ln(1-x)]$. In a random alloy each of $zN/2$ nearest-neighbour bonds is unlike with probability $2x(1-x)$. With one bond energy ϵ_{ij} per pair, subtract the energies of separated pure phases at the same overall composition.

$$\Omega = z\,[\epsilon_{AB} - (\epsilon_{AA}+\epsilon_{BB})/2],$$

$$\Delta g = \Omega x(1-x) + k_B T[x\ln x + (1-x)\ln(1-x)].$$

For $\Omega > 0$, unlike neighbours cost energy. The curvature at fixed temperature is $g'' = -2\Omega + k_B T/[x(1-x)]$. A homogeneous state is locally unstable when this curvature is negative. Equal chemical potentials across coexisting phases give a common tangent to $g(x)$. For the symmetric model $x_2 = 1 - x_1$, the tangent is horizontal and the binodal obeys $\Omega(1-2x) + k_B T \ln[x/(1-x)] = 0$.

At $x = 1/2$ the spinodal reaches its highest temperature $T_c = \Omega/(2k_B)$. For $\Omega = 0.10$ eV the mean-field T_c is 580 K. At 400 K the spinodal satisfies $x(1-x) = 0.1723$, or $x \simeq 0.221$ and 0.779. A composition between the binodal and spinodal can still have a positive local curvature while phase separation would lower the total energy after nucleation.



Coexistence compositions and local stability limits for the symmetric regular-solution model. A homogeneous mixture between these curves needs an initial fluctuation large enough to nucleate the competing phases.

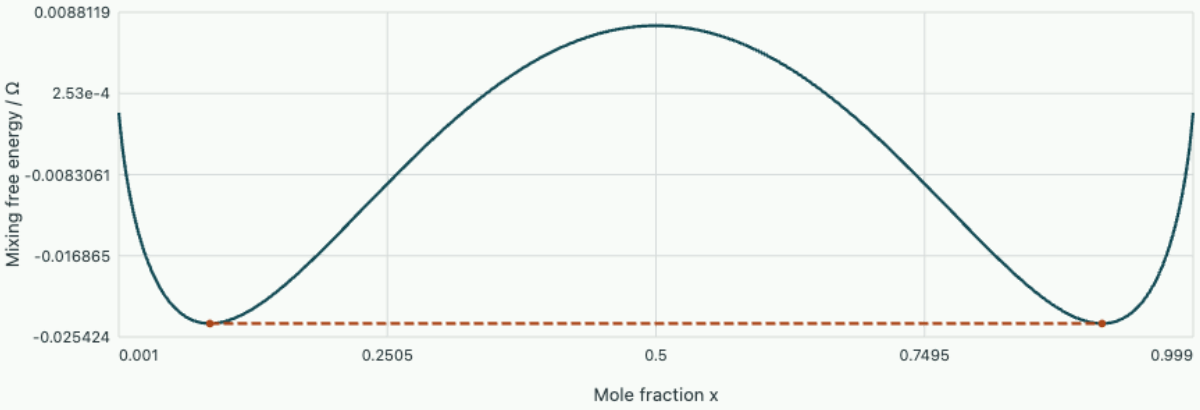
Try it · Regular solution

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

Reduced temperature T/Tc **0.7**



Binodal compositions 0.085683, 0.91432; spinodal 0.22614, 0.77386.

Construct an order parameter for preferential unlike neighbours

For an equiatomic alloy on a bipartite lattice let the A fractions of the two sublattices be $(1+\eta)/2$ and $(1-\eta)/2$. Then $\eta = 0$ is a random occupation and $|\eta| = 1$ is perfect order. Assuming independent occupation within each sublattice, the probability of an unlike nearest-neighbour bond is $(1+\eta^2)/2$. Put $w = \epsilon_{AB} - (\epsilon_{AA} + \epsilon_{BB})/2$. The energy change from the random state is $zw\eta^2/4$. Adding both sublattice mixing entropies gives the leading free-energy terms

$$g(\eta) - g(0) = \left(\frac{zw}{4} + \frac{k_B T}{2}\right)\eta^2 + \frac{k_B T}{12}\eta^4 + O(\eta^6).$$

For $w < 0$ the energetic gain favours order below $T_{\mathrm{ord}} = -zw/(2k_B)$. Domains can choose opposite signs and have the same composition. A superlattice diffraction reflection weights the difference of sublattice scattering amplitudes. Its amplitude is proportional to $(f_A - f_B)\eta$, so its ideal intensity scales as η^2. A measurement of the average composition alone overlooks the transition.

Recall · What makes this a mean-field model?

The product of single-site probabilities omits short-range correlations between neighbours. Clustering can therefore precede macroscopic separation and diffuse scattering can precede a sharp superlattice peak. Long-range elastic effects, unequal sublattice sizes or composition-dependent strain modify these predictions.

Distinguish an equilibrium phase boundary from slow atomic exchange

Composition changes by transport, which conserves the total fraction of each species. In a weak composition gradient a mobility $M > 0$ gives a flux proportional to $-\nabla\mu$ with $\mu = \partial g/\partial x$ in a locally uniform approximation. Negative curvature can drive atoms uphill in composition, amplifying fluctuations. A positive interfacial penalty proportional to $|\nabla x|^2$ prevents indefinitely fine segregation. In ordering at fixed composition, local exchanges can develop alternating occupation without transporting a macroscopic excess from one end of a specimen to the other.

Tutorial on alloy diffraction and stability

1 · Determine the sign of the interaction

▼ **Work through the solution**

If unlike bonds are 10 meV lower than the average like bond and coordination is eight, $w = -0.010$ eV gives a mean-field ordering temperature of $0.080/(2k_B) = 464$ K on a suitable bipartite lattice. Demixing is disfavoured in this nearest-neighbour calculation.

---

2 · Infer an order parameter

▼ **Work through the solution**

A superlattice reflection has one quarter of its intensity in a perfectly ordered single domain after correcting for geometry and thermal motion. The inferred $|\eta|$ is 0.5. Unequal domain volumes or poor scattering contrast demand a more complete intensity model.

---

3 · Diagnose two microstructures with similar average composition

▼ **Work through the solution**

In an equiatomic sample, composition mapping that resolves B-rich and A-rich regions indicates separation. Alternating-site superlattice peaks at essentially uniform local composition indicate ordering. Simultaneous data help when both mechanisms occur on different length scales.

For further study see Kittel, Introduction to Solid State Physics, 8th edition, chapter 22 on alloys and ordered structures.

Structure, excitations and probes of solids · 15. What extremal Fermi-surface orbits reveal about a metal

A single magnetic oscillation frequency gives one area in reciprocal space. Many different surfaces can share that area. Rotating the field and following how quantum oscillations and conductivity change turns the measurement into a three-dimensional inference problem.

Derive the plane of motion and the measured cross section

For semiclassical electrons of band energy $\epsilon(\mathbf{k})$ in a uniform static magnetic field with zero applied electric field, $\dot{\mathbf{r}} = \nabla_{\mathbf{k}}\epsilon/\hbar$ and $\hbar\dot{\mathbf{k}} = -e\dot{\mathbf{r}} \times \mathbf{B}$. Dotting the second equation with $\mathbf{B}$ shows that the orbit stays in a fixed plane perpendicular to the field. Dotting with $\nabla_{\mathbf{k}}\epsilon$ also shows that it remains on a constant-energy surface. A degenerate metal probes the intersection of these two surfaces near the Fermi level.

For a closed orbit of area $A(E, k_{\parallel})$ the semiclassical area quantisation condition is $A = 2\pi eB(n+\gamma)/\hbar$. At fixed Fermi energy adjacent crossings at increasing $1/B$ then have spacing $2\pi e/(\hbar A)$. Bulk signals sum oscillations from many neighbouring planes. Rapid phase variation cancels most contributions on averaging, leaving extrema of A along $k_{\parallel}$ particularly visible.

$$F = \frac{\hbar}{2\pi e} A_{\mathrm{ext}},$$

$$m_c = \frac{\hbar^2}{2\pi} \left.\frac{\partial A}{\partial E}\right|_{\mathrm{orbit}}.$$

The second relation follows by differentiating the area quantisation with respect to energy and comparing successive-level spacing to $\hbar eB/m_c$. A measured area does not by itself give its energy derivative. Temperature damping or an independent mass measurement contains additional information.

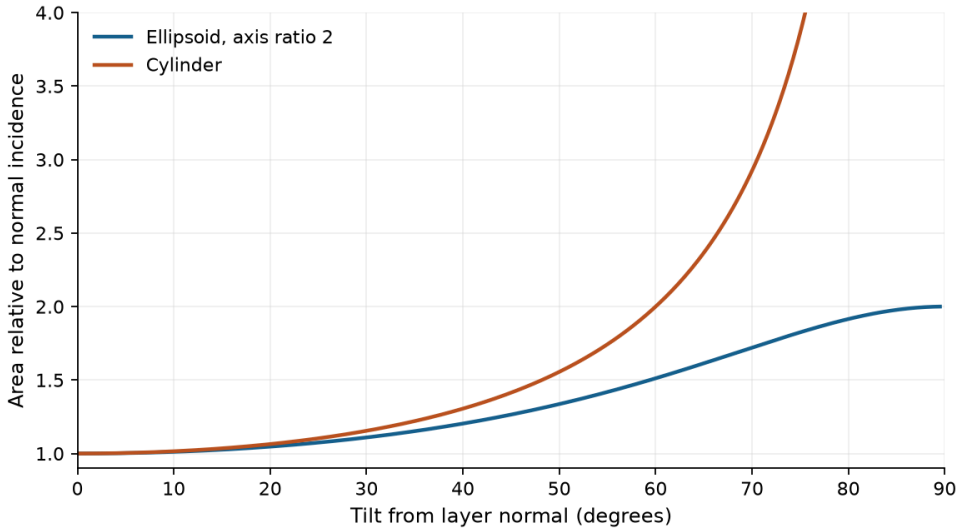
Predict the angular dependence for two competing surfaces

For an ellipsoid of semiaxes a, b, c in wave-number units, a linear scaling from a unit sphere gives a central section normal to $\mathbf{n}$ of area $\pi abc/\sqrt{a^2n_x^2 + b^2n_y^2 + c^2n_z^2}$. One way to see the denominator is to take a thin slab bounded by parallel planes. Under the inverse scaling the distance between the transformed planes changes by the inverse magnitude of the scaled normal. Equating volume before and after the transformation fixes the area factor.

For an axially symmetric ellipsoid take $a = b$ and tilt the field through angle θ from z. A circular cylinder of radius a instead has section area $\pi a^2/|\cos\theta|$ so long as a closed orbit exists. At normal incidence both predict the same frequency. For $c = 2a$, the ellipsoid predicts $F(60^\circ)/F(0) = 2/\sqrt{0.75+1} = 1.512$, while the cylinder predicts 2. An ellipsoid closes with a finite side-on area. A perfectly unwarped cylinder has open trajectories for a field in its plane.

Recall · Open orbits in a periodic crystal

In a reduced-zone drawing a trajectory can reappear at the opposite zone edge. Unfold the reciprocal lattice to see whether it closes or runs through successive zones. Such open orbits do not give an ordinary closed-area oscillation frequency and can strongly affect anisotropic transport.



An elongated ellipsoid and a cylinder agree for a field normal to the layers. Their extremal areas separate markedly as the field rotates toward the layer plane.

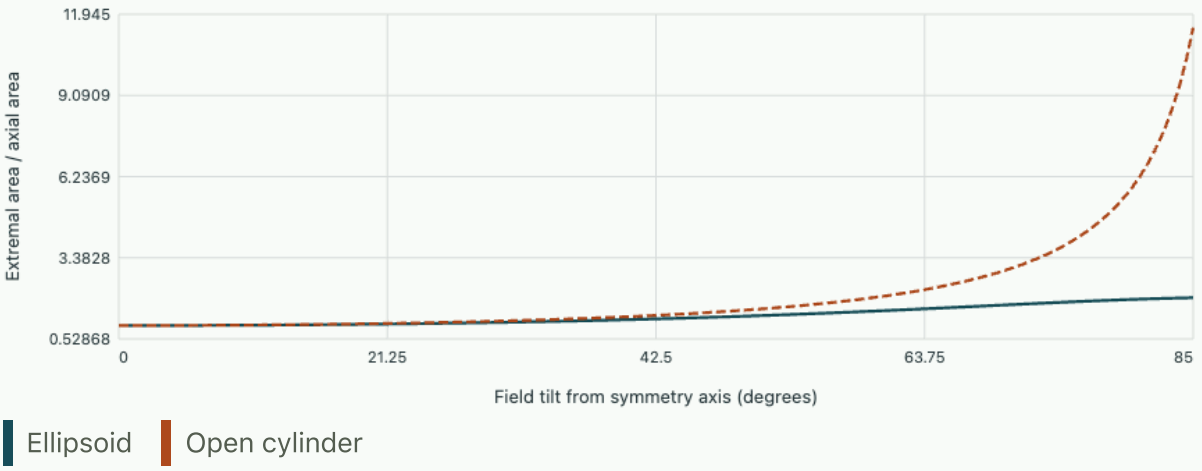
Try it · Fermi area

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

Ellipsoid polar / equatorial radius 2



Onsager frequency scales with the plotted area. An ellipsoid remains bounded at 90°; a perfect cylinder diverges in this ideal geometry.

Allow necks, multiple sheets and unequal visibility

A weakly warped quasi-two-dimensional sheet can have a belly and a neck. Model their areas as $A_0(1+\eta)$ and $A_0(1-\eta)$ for field along the layer normal. Then $F_0 = (F_b + F_n)/2$ and $\eta = (F_b - F_n)/(F_b + F_n)$. For 500 and 440 tesla, these give 470 tesla and a relative corrugation of 0.0638. With similar amplitudes the sum of the two oscillations contains an envelope $\cos[\pi(F_b - F_n)/B]$. The spacing of successive envelope nodes in inverse field is $1/|F_b - F_n|$.

Several frequencies can instead come from different electron and hole pockets. Measure the angular branches continuously and compare their effective masses, crystal symmetry and independent Hall and transport data. A heavy or strongly scattered orbit may be too weak to see. Assigning the entire carrier density to the most visible frequency can miss whole sheets. For an isotropic band with a spherical spin-degenerate pocket, the additional geometric assumption gives $n = k_F^3/(3\pi^2)$. Its boundary should be stated when reporting a density.

Tutorial on competing metal models

1 · Convert an oscillation frequency to an area

▼ **Work through the solution**

At $F = 500$ T, $A = (2\pi e/\hbar)F = 4.77 \times 10^{18}$ m⁻². If the orbit is circular the radius is $\sqrt{A/\pi} = 1.23 \times 10^{9}$ m⁻¹. The first number follows without assuming a circular shape.

---

2 · Estimate a beat period

▼ **Work through the solution**

With the two measured frequencies separated by 60 T, successive envelope zeros are separated by 0.0167 T⁻¹. A recording window of only 0.005 T⁻¹ cannot resolve the splitting convincingly. Vary the inverse-field window and check harmonic peaks.

---

3 · Infer an ellipsoid axis ratio from a side-on rotation

▼ **Work through the solution**

For $a = b$ the side-on to normal frequency ratio of a closed ellipsoid is c/a. A ratio 1.7 implies $c/a = 1.7$. If the signal disappears before reaching side-on, the missing end point leaves this inference underdetermined. Seek independent sections or a second rotation plane.

For further study see Kittel, Introduction to Solid State Physics, 8th edition, chapters 6–9 on electrons in metals, band structure and Fermi surfaces.

Magnetic and electrical material measurements

Follow a microscopic carrier or bound current into a measured Hall voltage, external field, magnetic response or penetration depth.

1. From microscopic moments to measured magnetisation
2. Interfaces, demagnetising fields and the work of magnetisation
3. Measure how quickly a dielectric follows an electric field
4. Identify mobile charge with a transverse field
5. When does a polar crystal change its preferred state?
6. Distinguish magnetic screening from electrical diffusion
7. Why rotating a layered dielectric changes a capacitance measurement
8. How a magnetic field rotates the polarisation of transmitted light
9. Adding angular momentum and predicting which lines can appear
10. When a magnetic field changes the quantum numbers of an atom

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.

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 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

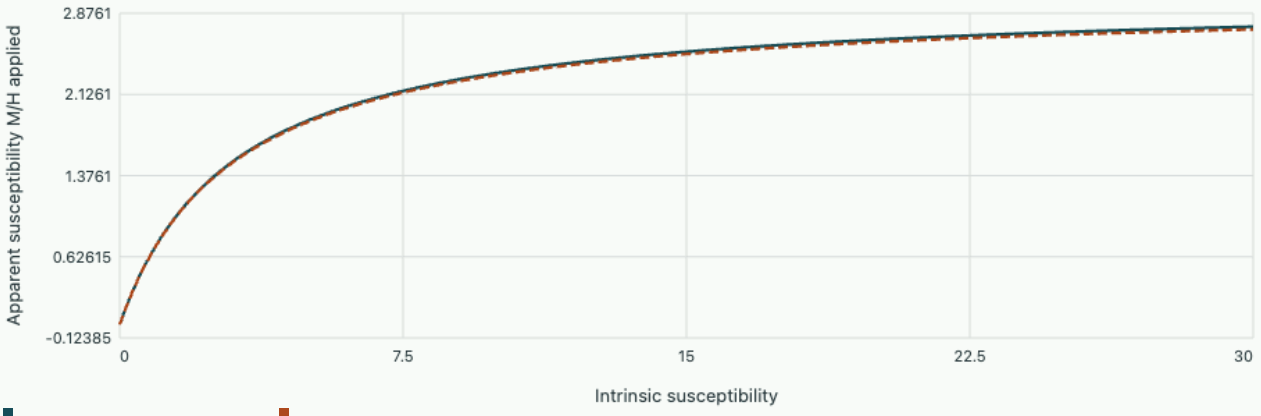
Change the magnetisation and compare interior and exterior fields at the surface. Check which tangential or normal field condition is relevant.

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 $2N\mu_0 w a M_0(0)$. 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 $2\pi a M_0$ 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.

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.

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

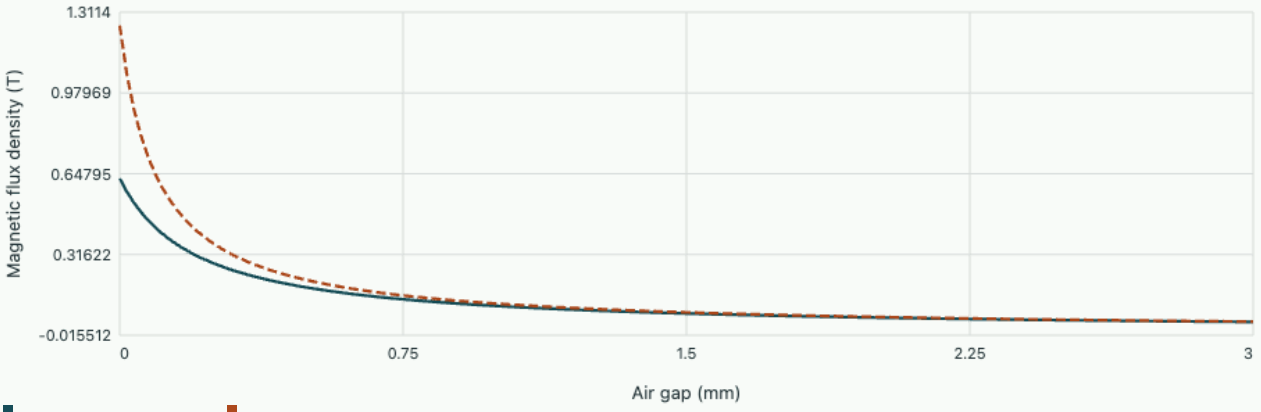
Change the reluctance of a narrow air gap and inspect which section controls flux. Check the assumptions needed before replacing the complete field with a circuit.

Open this experiment online

Try the calculation

Relative permeability **1000**

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



Core and gap Core with twice the permeability

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.

Measure how quickly a dielectric follows an electric field

Fields and the response of matter

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.

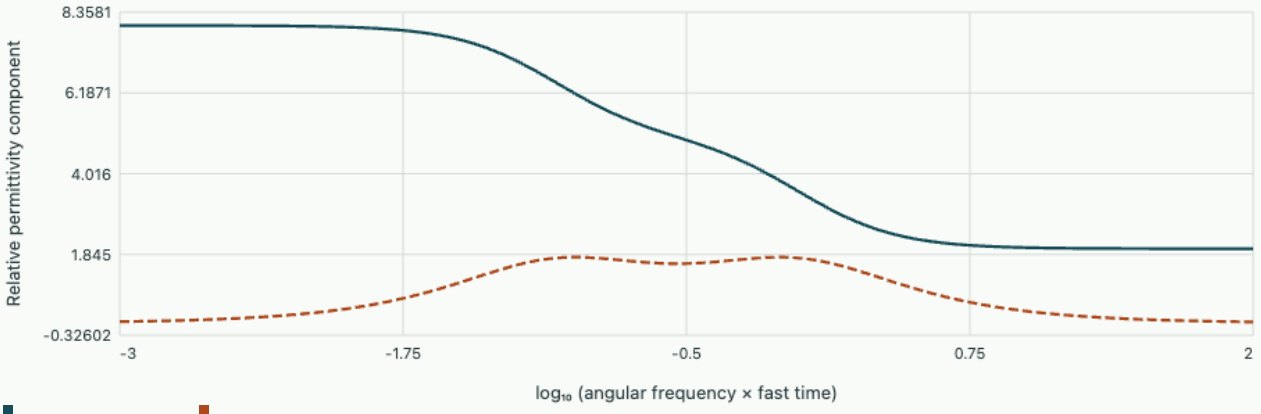
Resolve a second process only when the data require it

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.

Identify mobile charge with a transverse field

Fields and the response of matter

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

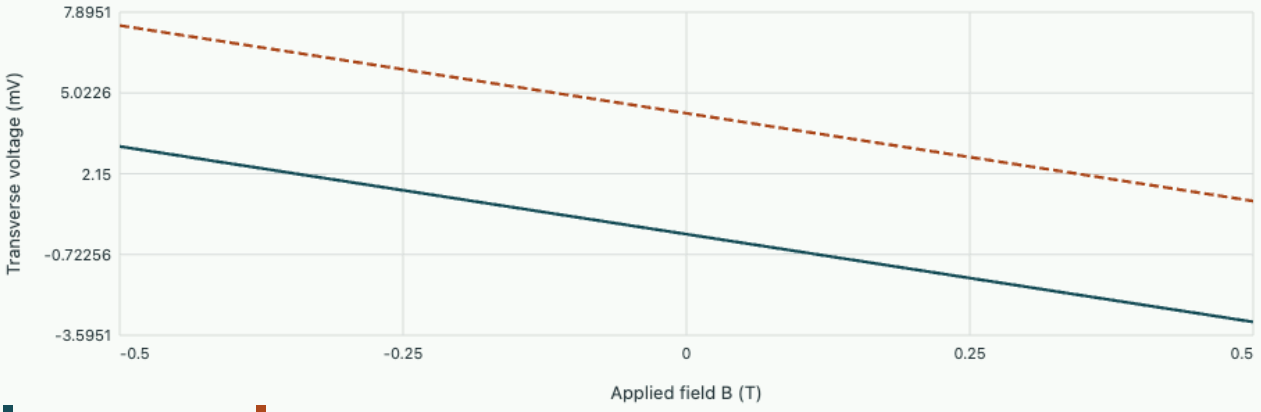
Practical side contacts seldom sit at exactly the same x. A fraction of the longitudinal voltage can leak into the transverse pair. Write the readout at a fixed current as $V_\mathrm{m}(B) = V_0 + \alpha V_{xx} + V_H(B)$. A background that is even in field cancels in $[V_\mathrm{m}(B) - V_\mathrm{m}(-B)]/2$. 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.

When does a polar crystal change its preferred state?

Fields and the response of matter

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.

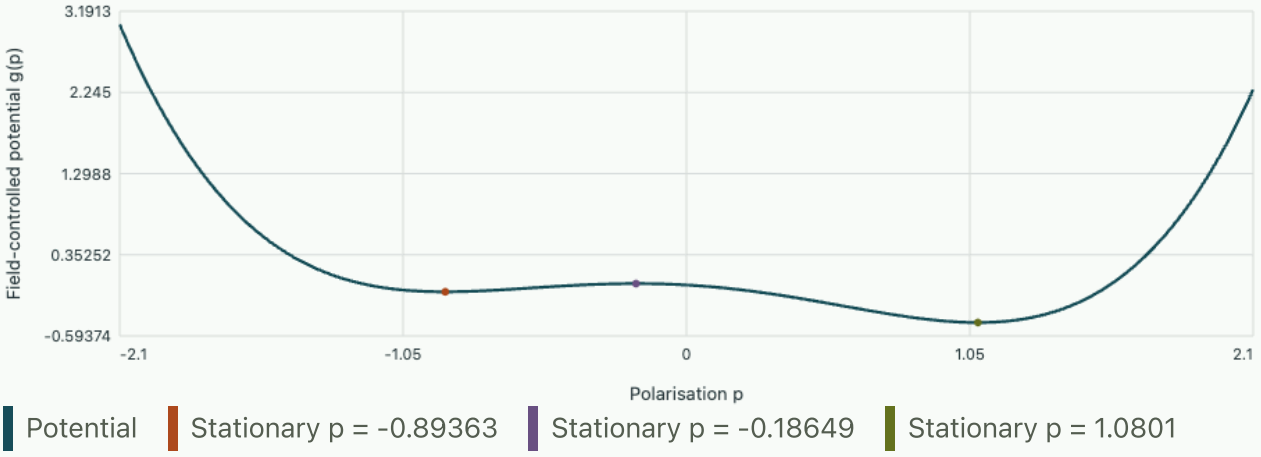
Read a landscape rather than tracing an unlabelled loop

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.

Distinguish magnetic screening from electrical diffusion

Fields and the response of matter

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.

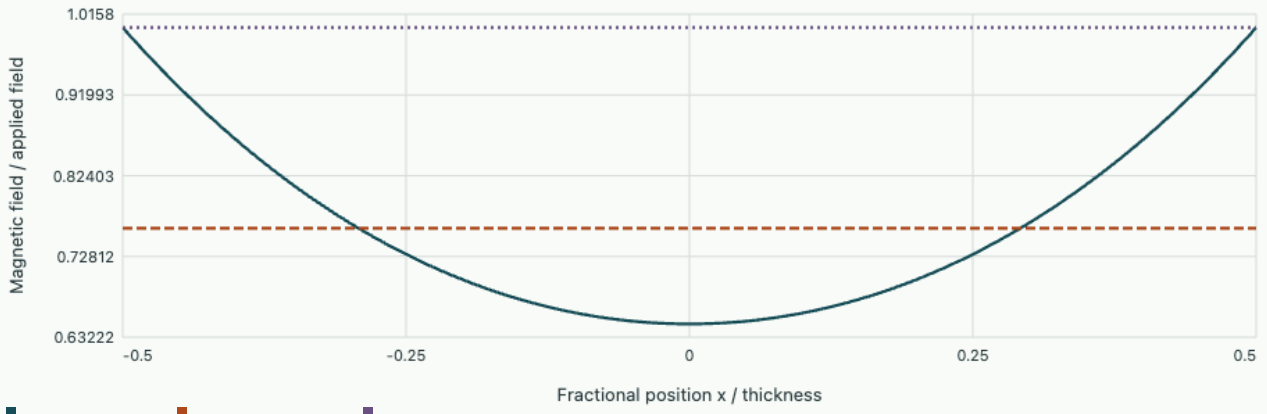
Compare two experimental controls

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.

Why rotating a layered dielectric changes a capacitance measurement

Fields and the response of matter

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.

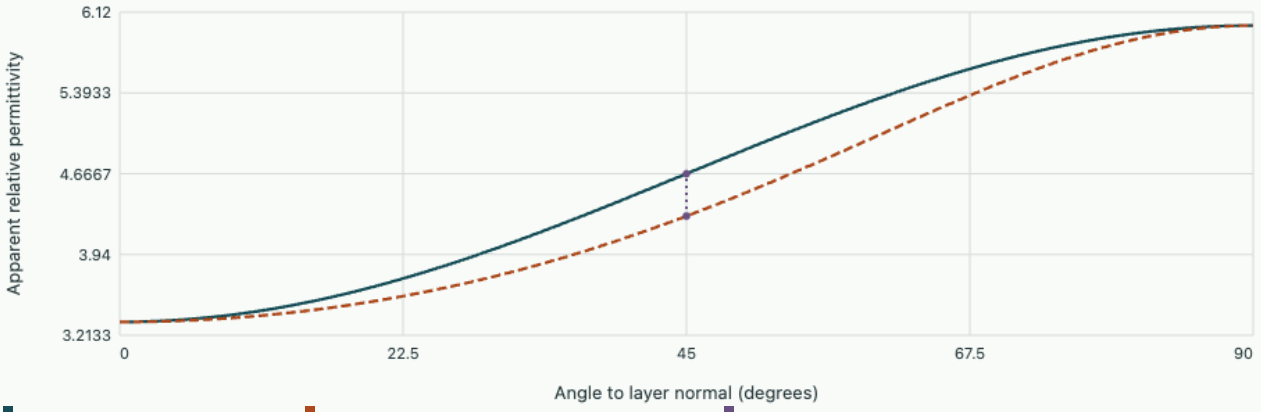
Test the measurement against the unit-cell model

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.

How a magnetic field rotates the polarisation of transmitted light

Fields and the response of matter

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 $\left(\begin{smallmatrix} \epsilon & ig \\ -ig & \epsilon \end{smallmatrix}\right)$ 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 $I/I_{\mathrm{in}} = \cos^2(\alpha - \theta)$. 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 $180/\pi$.

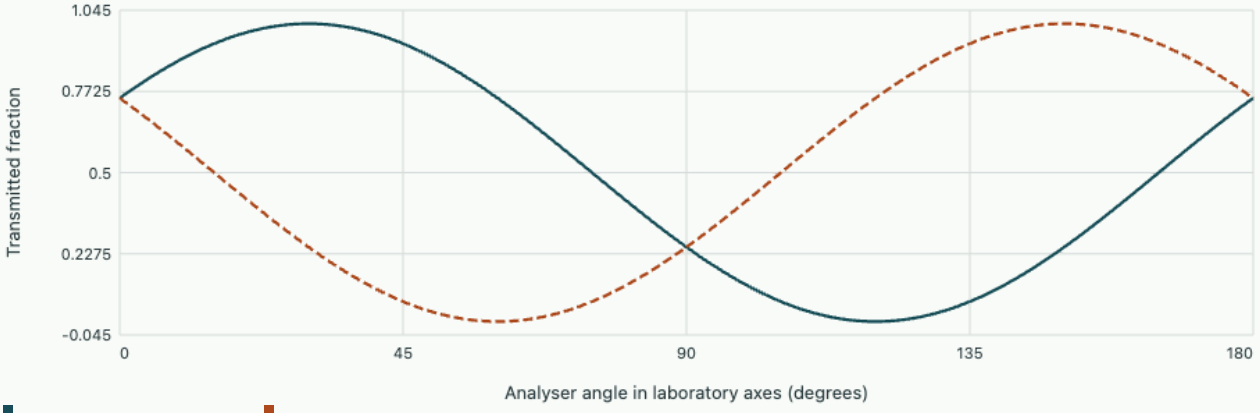
Use field reversal to reject alignment errors

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.

Adding angular momentum and predicting which lines can appear

Quantum states, dynamics and measurement

A three-component orbital wavefunction combines with a two-component spin. Derive the allowed total angular momenta, construct their amplitudes explicitly and predict both forbidden transitions and relative strengths in a polarisation-resolved spectrum.

Read the angular momentum algebra as a statement about rotations

Rotate a prepared quantum state through a small angle $\delta\theta$ about a unit laboratory direction $\mathbf{n}$. A continuous norm-preserving transformation has form $U = I - i\delta\theta\,\mathbf{n}\cdot\mathbf{J}/\hbar + O(\delta\theta^2)$, with Hermitian generators J_i. Rotations about different axes need not commute. Comparing the commutator of two infinitesimal rotations with the rotation around the third axis gives

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k.$$

The sum $J^2 = J_x^2 + J_y^2 + J_z^2$ commutes with each component. One can therefore label eigenstates of J^2 and one chosen component J_z simultaneously, while generally leaving the other components uncertain.

Define $J_\pm = J_x \pm iJ_y$. Their commutators with J_z are $\pm\hbar J_\pm$; with J^2 they vanish. If $J_z|m\rangle = \hbar m|m\rangle$, the two ladder operations change m by one without changing total angular momentum. To determine their coefficients, use $J_-J_+ = J^2 - J_z^2 - \hbar J_z$ and $J_+J_- = J^2 - J_z^2 + \hbar J_z$. Nonnegative squared norms bound the ladder above and below. Let j be its maximum magnetic quantum number. Vanishing raising norm there fixes the J^2 eigenvalue at $\hbar^2 j(j+1)$; vanishing lowering norm at the bottom gives

$$m_{\min} = -j.$$

The number of unit steps between the two endpoints is $2j$, so j is an integer or half-integer. With a consistent phase convention

$$J_\pm|j,m\rangle = \hbar\sqrt{j(j+1) - m(m\pm1)}\,|j,m\pm1\rangle.$$

Spatial orbital eigenfunctions on an ordinary single-valued scalar wavefunction carry integer orbital l. Spin-$\frac{1}{2}$ has two states that cannot be modelled by adding another spatial coordinate angle to a scalar orbital. The operator algebra describes both. In $|j,j\rangle$, J_z is definite but J_x, J_y fluctuate. Their variance sum is $\hbar^2 j(j+1) - \hbar^2 j^2 = \hbar^2 j$; rotational symmetry around z splits it evenly. Drawing a classical vector with precisely fixed components perpendicular to the axis would miss these fluctuations.

Construct a coupled basis from the stretched state

For independent angular momenta $\mathbf{L}, \mathbf{S}$, the total generator is $\mathbf{J} = \mathbf{L} + \mathbf{S}$. Operators on different factors commute. In the product basis $|m_l, m_s\rangle$, J_z is diagonal with eigenvalue $\hbar(m_l + m_s)$. The Hamiltonian

may instead depend on J^2, for example through $\mathbf{L}\cdot\mathbf{S}$. Then states with the same $m_l + m_s$ can mix. Resolving that degeneracy requires a unitary change of basis, whose entries are Clebsch–Gordan coefficients.

Take $l = 1, s = 1/2$. The unique state $m = 3/2$ is $|1,\uparrow\rangle$, so it must be $|3/2, 3/2\rangle$. Lower once with $J_- = L_- + S_-$. The left side is $\hbar\sqrt{3}|3/2, 1/2\rangle$; the right side is $\hbar(\sqrt{2}|0,\uparrow\rangle + |1,\downarrow\rangle)$. Normalising and constructing an orthogonal state in the same m plane gives

$$|3/2, 1/2\rangle = \sqrt{2/3}|0,\uparrow\rangle + \sqrt{1/3}|1,\downarrow\rangle,$$

$$|1/2, 1/2\rangle = \sqrt{1/3}|0,\uparrow\rangle - \sqrt{2/3}|1,\downarrow\rangle.$$

The sign of an entire coupled vector is conventional. The relative sign of its two components matters whenever a later measurement recombines amplitudes. Applying J_+ to the second combination gives zero, which verifies the top of the $j = 1/2$ multiplet independently of orthogonality. Lowering produces four states for $j = 3/2$ and two for $j = 1/2$, exhausting the six-dimensional product space. Merely listing both possible j's does not supply the amplitude with which a prepared product state belongs to each, which is what spectroscopy needs.

There is a counting derivation for arbitrary j_1, j_2. Start with one stretched state at $m = j_1 + j_2$ and remove its entire ladder. The next highest unused m supplies the next multiplet and so on down to $|j_1 - j_2|$. The identity $\sum_{j=|j_1-j_2|}^{j_1+j_2}(2j+1) = (2j_1+1)(2j_2+1)$ checks completeness. Distinct coupling schemes for three angular momenta are different bases of a common space. An intermediate angular momentum may be a reliable state label when a corresponding interaction dominates, yet cease to be conserved if an external field favours another axis or coupling.

Derive magnetic and parity selection rules

Consider a weak optical field coupled by the electric dipole interaction $-\mathbf{d}\cdot\mathbf{E}$, where $\mathbf{d}$ transforms as a vector and has odd parity. Its spherical components relative to z are

$$d_0 = d_z,\; d_{\pm1} = \mp(d_x \pm i d_y)/\sqrt{2}.$$

For a definite tensor component $[J_z, d_q] = q\hbar d_q$. Sandwiching the commutator between initial and final J_z eigenstates yields

$$(m_f - m_i - q)\langle f|d_q|i\rangle = 0.$$

Thus nonzero amplitudes require

$$m_f = m_i + q.$$

Coupling j_i with a rank-one tensor also requires $|j_i - 1| \le j_f \le j_i + 1$, and $0 \to 0$ is excluded. Parity gives a separate test: under inversion the matrix element gains factor $-\pi_f\pi_i$, so it vanishes for equal

initial and final parity in a parity-symmetric Hamiltonian. For purely orbital motion with parity $(-1)^l$, the triangle condition plus odd parity gives

$$\Delta l = \pm 1.$$

A direction label on a photon depends on propagation and polarisation conventions. Stating the driven operator component q is less ambiguous than calling a beam "right circular" without defining the viewing direction.

For a spin-independent dipole transition the spin factor is $\langle m_s' | m_s \rangle$. Begin in $l_i = 0, m_l = 0, m_s = +1/2$ and drive $q = 0$ to $l_f = 1, m_l = 0$. In the product basis there is one equally normalised orbital amplitude D. Its projections on the two coupled final levels have squared amplitudes $2|D|^2/3$ and $|D|^2/3$, giving a 2:1 line-strength ratio under equal field intensity, comparable detection and negligible population saturation. For $q = +1$ the product state is $|1, \uparrow\rangle$. The transition to final $j = 1/2$ vanishes because no $m = 3/2$ belongs to that multiplet. These statements predict relative angular factors. Actual count rates can also depend on frequencies, radial matrix elements, initial populations and how fluorescence is collected.

As an explicit orbital check,

$$Y_{00} = 1/\sqrt{4\pi},$$

$$Y_{10} = \sqrt{3/(4\pi)}\cos\theta.$$

The angular factor of $d_0 \propto r\cos\theta$ is $\int Y_{10}^* \cos\theta Y_{00}\, d\Omega = 1/\sqrt{3}$, while for $Y_{1,\pm1}$ the azimuthal integral vanishes. A radial integral can still vanish accidentally even when all angular tests permit the transition. Symmetry rules predict zeros robust to changes preserving that symmetry; a permitted channel has an amplitude that must still be calculated.

Separate atomic structure from analyser orientation with a reduced matrix element

The angular selection rule from the J_z commutator identifies which transitions are forced to zero but does not alone give their relative strengths. A rotationally covariant way to connect the allowed amplitudes is to view an irreducible tensor component $T_q^{(k)}$ as carrying angular momentum rank k. Under rotations its $2k+1$ components transform among themselves with the same representation matrices as the magnetic substates of an angular momentum k. Coupling $T_q^{(k)}$ to an initial $|jm\rangle$ therefore furnishes the same product representation as coupling $|jm\rangle$ to $|kq\rangle$. The component projecting onto a final j' must transform as j'. When this representation occurs once in the product, its relative magnetic amplitudes are fixed by Clebsch–Gordan coefficients up to one common factor independent of m, q, m'.

In one standard convention the Wigner–Eckart theorem writes

$$\langle \alpha' j' m' | T_q^{(k)} | \alpha j m \rangle = \frac{\langle jm; kq | j' m' \rangle}{\sqrt{2j'+1}} \langle \alpha' j' \| T^{(k)} \| \alpha j \rangle.$$

Additional labels α, α' specify the radial, spin or configuration information absent from the angular labels. Conventions placing a different square-root factor in the reduced element are equally common, so a table of reduced strengths should be checked before combining it with an external coefficient table. The theorem predicts ratios within a fixed j, j' multiplet. It does not remove differences in radial integrals between unrelated configurations, nor does it assert the reduced element is nonzero. This factorisation explains why angular data alone may fix a branching ratio while offering no absolute radiative lifetime.

For a vector operator connecting j to the same j, there is an especially transparent example. Within a given irreducible j space, any vector operator has matrix elements proportional to the corresponding J_q matrix elements, provided additional degeneracies are separated. For the conventional spherical components $J_0 = J_z$, $J_{+1} = -J_+/\sqrt{2}$, ask for the ratio of line strengths out of $j = 2, m = 1$ into $m' = 2$ via $q = +1$ and into $m' = 1$ via $q = 0$, with equal field amplitudes in the specified components. The squared matrix elements of J_q are $\hbar^2[6 - 1(2)]/2 = 2\hbar^2$ and $\hbar^2$, hence the ratio is two. The same numerical angular ratio can apply to any appropriate nonzero reduced vector coupling between these specified manifolds. For an electric dipole transition parity must still allow the initial and final configurations; using J_q here computes angular ratios and does not turn angular momentum itself into an electric dipole operator.

Summing over a complete set of final magnetic substates and all tensor components supplies a useful normalisation audit. For a fixed initial m, rotational invariance of the scalar sum $\sum_q T_q^\dagger P_{j'} T_q$ makes its diagonal expectation independent of m. A detector with a preferred polarisation or direction samples only a subset and may exhibit strong dependence on m. This difference is essential in interpreting “unpolarised” line strengths: an unpolarised pump that populates all m equally, an unanalysed detector that sums final polarisations, and a source with a definite m viewed along one axis are three different experimental ensembles.

Compare successive Stern–Gerlach measurements with addition of spin

Two spin-$1/2$ particles provide a minimal addition experiment with no spatial angular wavefunction needed. The coupled triplet basis is $|1,1\rangle = |\uparrow\uparrow\rangle$,

$$|1,0\rangle = (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)/\sqrt{2},$$
$$|1,-1\rangle = |\downarrow\downarrow\rangle;$$

the singlet is $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$. An apparatus measuring the total z component alone cannot distinguish the two $m = 0$ states. Measuring J^2 along with J_z can. Their local one-spin reduced matrices are both $I/2$, so even complete tomography of either spin alone fails to reveal the difference. The distinction lies in a joint correlation.

The singlet projector can be obtained as a polynomial in $\mathbf{S}_1 \cdot \mathbf{S}_2$. That scalar has eigenvalues $-3\hbar^2/4, +\hbar^2/4$, so $P_s = I/4 - \mathbf{S}_1 \cdot \mathbf{S}_2/\hbar^2$. Expressing $\mathbf{S}_i = (\hbar/2)\boldsymbol{\sigma}_i$ gives

$$P_s = [I - \sum_i \sigma_i \otimes \sigma_i]/4.$$

If three independently prepared ensembles yield measured same-axis correlations $(-0.80, -0.78, -0.82)$, their inferred singlet fraction is

$$(1 + 2.40)/4 = 0.85.$$

This result uses all three axes. Perfect z-anticorrelation alone constrains the population to a two-dimensional $m = 0$ space but says nothing about the relative phase within it. If each measurement has uncertainty 0.03 and the estimates are independent, the resulting statistical uncertainty in singlet fraction is $\sqrt{3}(0.03)/4 = 0.013$, before calibration uncertainties.

For a coherent initial product $|\uparrow\downarrow\rangle$, a Hamiltonian $H = J_{\mathrm{ex}}\mathbf{S}_1 \cdot \mathbf{S}_2/\hbar^2$ evolves the singlet and triplet amplitudes with distinct phases. Writing the initial state as $(|1, 0\rangle + |0, 0\rangle)/\sqrt{2}$, the probability to find spins reversed later is $\sin^2(J_{\mathrm{ex}}t/(2\hbar))$. The total m remains exactly zero at all times. If $J_{\mathrm{ex}}/h = 5.00$ MHz, the first full reversal occurs at $t = 1/(2 \times 5\,\mathrm{MHz}) = 100$ ns. An incoherent mixture with the same singlet and triplet populations has a constant reversal probability one half. This calculation makes the signs in the coupled basis observable and distinguishes a true exchange evolution from random spin flips inferred solely from individual populations.

Rotate the quantisation axis without changing the atomic Hamiltonian

A transverse magnetic or polarisation analyser can interrogate a state prepared relative to z along a different laboratory axis. For spin $1/2$, $e^{-i\theta S_y/\hbar} = \cos(\theta/2)I - i\sin(\theta/2)\sigma_y$, from $\sigma_y^2 = I$. To obtain the rotation probabilities for a spin-one state without memorising a rotation matrix, represent its stretched state as the symmetric product of two up spinors. Under a common rotation each spinor transforms as $\cos(\theta/2)|\uparrow\rangle + \sin(\theta/2)|\downarrow\rangle$. Expanding the two-spin product gives the triplet amplitudes $\cos^2(\theta/2), \sqrt{2}\cos(\theta/2)\sin(\theta/2), \sin^2(\theta/2)$ for the three m values. This construction exploits the representation and applies to any angular momentum one, regardless of its microscopic origin.

At $\theta = \pi/2$, a stretched $m = 1$ state projected on the rotated axis gives probabilities $(1/4, 1/2, 1/4)$. The expectation of its rotated component vanishes, but outcomes $+\hbar$ and $-\hbar$ still occur and the variance is $\hbar^2/2$. A source prepared in $m = 0$ along z gives a different test. Rotate its symmetrised up–down form on both spins and project on $m = 0$; the amplitude is $\cos\theta$ and the probability $\cos^2\theta$. At a right angle the “zero” outcome is absent, although the mean along either axis is zero. A single mean value cannot distinguish alignment from an unpolarised mixture. Measuring the full resolved substate distribution, or its quadrupole moment, can.

The general stretched state $|j, j\rangle$ can be represented by $2j$ symmetrised spinors all up. Its rotated m outcome distribution is binomial:

$$P(m) = \binom{2j}{j-m} \cos^{2(j+m)}(\theta/2)\sin^{2(j-m)}(\theta/2).$$

Its mean angular momentum component is $\hbar j\cos\theta$, with variance $\hbar^2(j/2)\sin^2\theta$. At fixed nonzero $\cos\theta$, the relative fluctuation falls like $j^{-1/2}$ as j grows. This supplies a precise sense in which a large stretched

state can approximate a classical oriented angular momentum while retaining quantum projection noise. A state with the same large j but $m = 0$ need not resemble an oriented classical arrow. The representation dimension alone does not determine the suitability of a classical approximation.

Distinguish an angularly forbidden line from coherent cancellation

Suppose two degenerate lower magnetic states $|g_1\rangle$, $|g_2\rangle$ both couple to one excited state $|e\rangle$ under controlled coherent drives. In a rotating-frame model the raising part is $|e\rangle(\Omega_1\langle g_1| + \Omega_2\langle g_2|)\hbar/2$, with the lowering term its adjoint. A normalised superposition proportional to $\Omega_2|g_1\rangle - \Omega_1|g_2\rangle$ has zero excitation amplitude because the two terms cancel. Each angular channel can be individually allowed and strong, yet this particular coherent combination is dark. Its orthogonal bright state is proportional to $\Omega_1^*|g_1\rangle + \Omega_2^*|g_2\rangle$ and has effective Rabi coupling $\sqrt{|\Omega_1|^2 + |\Omega_2|^2}$.

For equal real drives and an initially balanced antisymmetric superposition, the transition ideally vanishes. A randomly prepared mixture of the same two basis states populates the bright and dark subspaces equally. If a resonant pulse transfers the bright component completely, that mixture produces 50% excitation. Changing the relative optical phase changes which superposition is dark but does not change the selection rule for either individual transition. Preparing a calibrated ground-state phase and scanning the drive phase can therefore distinguish coherent cancellation from a forbidden angular matrix element or a disconnected drive channel. Two-photon detuning, magnetic splitting of the lower states and dephasing can spoil the destructive interference even if every angular coefficient is exactly known.

Track a spin while rotating the preparation axis

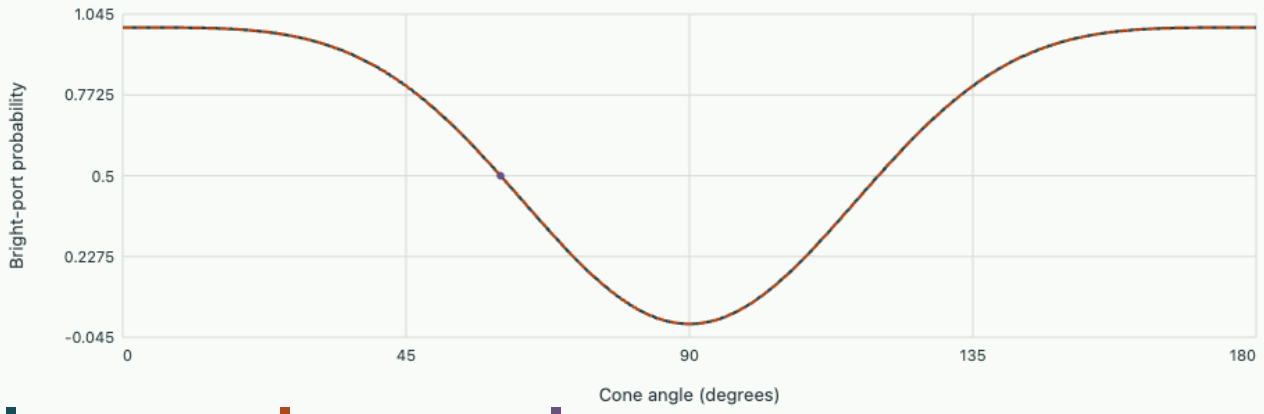
Change the path of a spin direction and compare a change in its geometric phase with an ordinary dynamical phase.

Open this experiment online

Try the calculation

Cone angle (degrees) 60

Reference phase (degrees) 0



Forward traversal Reverse traversal Selected forward probability

Forward geometric phase -0.5π radians. Forward port probability 0.5, reverse 0.5.

Predict which perturbation can make a dark transition visible

Electric dipole selection rules use the approximation that the driving field is essentially uniform across the atom. Expand a monochromatic plane wave $e^{i\mathbf{k}\cdot\mathbf{r}} = 1 + i\mathbf{k}\cdot\mathbf{r} + \cdots$. For a source of size a the next spatial term is suppressed in amplitude by a factor of order ka when $ka \ll 1$. Consistently expanding the light–matter Hamiltonian yields higher electric multipoles and magnetic couplings with their own symmetry. The electric quadrupole is a rank-two even-parity tensor. Thus it can connect same-parity states with $|j_i - 2| \le j_f \le j_i + 2$. A magnetic dipole is a rank-one operator of even spatial parity, permitting some same-parity transitions denied to an electric dipole. A line forbidden to one specified multipole need not vanish in all experimental conditions.

Take a model atom with a lower $j = 0$ even state and an upper $j = 2$ even state. Electric dipole fails both parity and angular-momentum tests. Magnetic dipole has the right parity but still cannot supply total angular rank two. Electric quadrupole has the correct parity and satisfies the triangle condition; its actual reduced matrix element and the applied field gradient still determine strength. In a $j = 0 \to j' = 2$ transition with $m_i = 0$, each spherical quadrupole component q excites only $m_f = q$. A simple claim that changing circular to linear polarisation opens an electric dipole line in this field-free atom would be inconsistent with both global selection rules. A static field that mixes opposite parity or different angular momentum can change the states themselves and thereby modify that conclusion.

For a 600 nm wavelength and characteristic radius $a = 0.10$ nm,

$$ka = 2\pi a/\lambda = 1.05 \times 10^{-3}.$$

The squared geometric factor is about 1.1×10^{-6} relative to a typical allowed electric dipole strength in comparable units, before distinct radial and angular factors. This is an order estimate, not a quantitative

branching ratio for two arbitrary lines. In particular, dividing a finite quadrupole amplitude by a strictly zero dipole amplitude of the same ideal transition is meaningless. Compare with a separate reference line whose dipole matrix element and detection efficiencies are known if a rate ratio is needed.

An independent way to brighten an otherwise forbidden line is mixing through a perturbation W. Suppose an upper state $|u\rangle$ has zero electric dipole amplitude from $|g\rangle$, while a nearby state $|v\rangle$ allows amplitude D. If $|\langle v|W|u\rangle| \ll |E_u - E_v|$, the perturbed upper state contains coefficient $\epsilon = \langle v|W|u\rangle/(E_u - E_v)$ of $|v\rangle$, and its borrowed dipole intensity scales as $|\epsilon|^2|D|^2$. Where W is proportional to a reversible static field, borrowed intensity typically begins quadratically in that field for a previously dark isolated line. A native higher-multipole contribution may already be present at zero field and can interfere if it feeds the same final channel with coherent angular dependence.

Use polarisation and bias sweeps to identify borrowed intensity

In a two-state model the bright level is 2.00 GHz away and a perturbation mixes it with the dark state at off-diagonal $W/h = 20$ MHz. Predict relative borrowed intensity, assess the leading approximation, and identify an independent spectroscopic consequence.

▼ **Derive and check the result**

The small parameter is

$$20/2000 = 0.010.$$

The leading borrowed intensity is 10^{-4} of the corresponding bright reference, with corrections of relative order 10^{-4} in the ideal two-level mixing expansion. The dark state's energy shifts by approximately $W^2/(E_u - E_v)$, whose magnitude expressed as ordinary frequency is

$$20^2/2000 = 0.20\,\mathrm{MHz}.$$

Its sign is away from the neighbouring bright level. Measuring this level repulsion together with the same polarisation pattern as the borrowed bright transition tests the mixing explanation. Changing the perturbing field by a factor of two should increase both the leading borrowed intensity and the magnitude of the second-order level shift by a factor of four while the coupling remains weak. A field-independent residual or a different polarisation dependence motivates examining other multipoles or additional states.

Tutorial · infer the state that a spectrum actually selects

Measure the orbital label after total angular momentum preparation

Prepare $|j = 1/2, m = 1/2\rangle$ in the coupled p manifold. Predict m_l outcomes, the spin expectation and the sum $\langle L_z + S_z\rangle$. Is either product component a state with definite J^2?

▼ **Work through the solution**

The probabilities are $1/3$ for $m_l = 0$ with spin up, $2/3$ for $m_l = 1$ with spin down. Thus $\langle L_z \rangle = 2\hbar/3$, $\langle S_z \rangle = -\hbar/6$ and the sum is $\hbar/2$. The product vectors in this two-dimensional plane are each superpositions of both j's, as can be checked by inverting the two orthogonal combinations. A joint measurement of orbital and spin labels can therefore disturb a previously exact J^2. The total m remains fixed in every outcome, stronger than merely having the correct mean.

Determine a spin–orbit interval without an angular integral

Within $l = 1, s = 1/2$, take $H' = A\,\mathbf{L}\cdot\mathbf{S}/\hbar^2$ with $A/h = 8$ GHz. Find level shifts, degeneracies and their degeneracy-weighted centre. Check the result from a product-basis trace.

▼ **Work through the solution**

Use $2\mathbf{L}\cdot\mathbf{S} = J^2 - L^2 - S^2$. For $j = 3/2$ the factor is $+1/2$; for $j = 1/2$ it is -1. Shifts are $+4$ and -8 GHz, with degeneracies four and two; their interval is 12 GHz and their weighted centre is zero. The product-basis trace of each $L_i S_i$ is $\mathrm{tr}\, L_i \;\mathrm{tr}\, S_i = 0$, which predicts that centre independently. If a measured centroid is weighted by optical intensity instead of counting every substate once, it need not equal this trace centre.

Distinguish resolved lines from a coherent pulse

A short $q = 0$ pulse prepares the product final state $|0, \uparrow\rangle$ above. If its spectrum covers both j levels coherently, find the later return probability to that same product state for a level angular-frequency splitting Δ. What would an incoherent population mixture give?

▼ **Work through the solution**

The return amplitude is $(2/3)e^{-iE_{3/2}t/\hbar} + (1/3)e^{-iE_{1/2}t/\hbar}$. Its squared modulus is $5/9 + (4/9)\cos\Delta t$, between $1/9$ and 1. The corresponding incoherent mixture gives a constant $5/9$. At $t = 0$ both constructions share the same energy-measurement probabilities, yet only the coherent preparation returns with certainty to the original product state. Temporal resolution, controlled pulse phase and bandwidth sufficient to cover both lines are part of the proposal. If the emitted or absorbed field records distinguishable frequencies, the observed beat can be suppressed.

Further reading

- Fließbach · Operatorenmethode, angular momentum algebra
- Cambridge · atoms in electromagnetic fields and electric dipole selection rules
- Course investigation · angular momentum and atoms

When a magnetic field changes the quantum numbers of an atom

Quantum states, dynamics and measurement

Weak-field Zeeman slopes label total-angular-momentum multiplets. At stronger fields those slopes curve and intensities migrate. Solve an entire orbital-plus-spin manifold and infer coupling constants from measurable line positions and polarisation.

Choose the Hamiltonian appropriate to a resolved orbital manifold

Suppose an isolated atomic p manifold with $l = 1, s = 1/2$ is sufficiently far from other orbital or principal levels to treat only its six substates. Take z along a uniform external magnetic field B. Work to first order in the electronic Zeeman operator and model spin–orbit coupling by the constant $A > 0$. With $b = \mu_B B$ and electron spin factor approximated by two,

$$H = \frac{A}{\hbar^2}\mathbf{L}\cdot\mathbf{S} + \frac{b}{\hbar}(L_z + 2S_z).$$

Both A and b are energies. The sign of the second term is positive for the electron because magnetic moment is opposite to angular momentum and field energy is minus moment dot field. Here $\mu_B = e\hbar/(2m_e) > 0$. The orbital manifold approximation requires magnetic splittings small compared with gaps to omitted levels. Neglected terms can include the difference of the actual spin factor from two, hyperfine interactions and the diamagnetic B^2 term. An “intermediate field” in this discussion means b comparable to A; it need not exceed any other electronic binding scale.

At zero field the coupled levels $j = 1/2, 3/2$ have energies $-A, +A/2$. At nonzero field the projection $m = m_l + m_s$ is still conserved because $[H, J_z] = 0$. For intermediate fields J^2 need not commute with the Zeeman term. Consequently the six-state problem reduces to two one-dimensional and two two-dimensional invariant blocks, even when the label j loses its exact meaning. Following a state by magnetic quantum number and overlap with a neighbouring field value is generally more reliable than sorting all computed energies from low to high. Crossings of different m can reorder a sorted list without exchanging the underlying states.

Derive the Landé slope from angular momentum projection

In a state with definite j, l, s, m, expectation of any vector operator proportional in direction to $\mathbf{J}$ within its irreducible multiplet can be obtained from its scalar projection on $\mathbf{J}$. For spin,

$$\frac{\langle S_z\rangle}{\hbar} = m\,\frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}.$$

One way to obtain the projection numerator is $2\mathbf{S}\cdot\mathbf{J} = J^2 + S^2 - L^2$. Since $L_z + 2S_z = J_z + S_z$, the energy slope for weak fields is $dE/db = g_j m$ with $g_j = 1 + [j(j+1) + s(s+1) - l(l+1)]/[2j(j+1)]$. A state $j = 0$ has vanishing linear vector moment and should be treated separately instead of dividing by zero. Within the p manifold the $j = 1/2$ factor is $2/3$ and the $j = 3/2$ factor $4/3$. For $m = +1/2$ these correspond to slopes $1/3$ and $2/3$ in

energy units of b. Their sum is exactly one, a trace constraint that remains valid for the sum of the two eigenvalues in that block at every field.

If $A/h = 8.00$ GHz as in the previous angular exercise, the zero-field interval is 12.00 GHz. At $B = 20.0$ mT, $b/h \simeq 0.280$ GHz using $\mu_B/h = 13.996$ GHz/T. The two positive-m shifts to linear order are $+0.0933$ and $+0.1866$ GHz. Actual optical frequencies measure final minus initial energies; an initial $s_{1/2}$ level has its own Zeeman slope. Omitting it is a common reason a slope inferred from a line scan seems inconsistent with the expected Landé factor.

Diagonalise the mixed levels across the entire crossover

In the $m = +1/2$ product basis $(|0,\uparrow\rangle, |1,\downarrow\rangle)$, expand $\mathbf{L}\cdot\mathbf{S} = L_zS_z + (L_+S_- + L_-S_+)/2$. The diagonal spin–orbit elements are $0, -A/2$; the off-diagonal element is $A/\sqrt{2}$. The Zeeman term is diagonal with entries $b, 0$. Hence

$$H_+(b) = \begin{pmatrix} b & A/\sqrt{2} \\ A/\sqrt{2} & -A/2 \end{pmatrix}, \qquad E_\pm(b) = \frac{b - A/2}{2} \pm \frac{1}{2}\sqrt{(b + A/2)^2 + 2A^2}.$$

At $b = 0$ the square root is $3A/2$, recovering the coupled levels. For the $m = -1/2$ basis $(|0,\downarrow\rangle, |-1,\uparrow\rangle)$, replace b by $-b$. The stretched levels $|1,\uparrow\rangle$ and $|{-1},\downarrow\rangle$ remain product states at every field with energies $A/2 \pm 2b$. Counting all levels and their traces provides an independent check: spin–orbit and Zeeman each have zero trace in the full six-dimensional space. At zero field both slopes in the $+$ block are finite and positive. At large positive b the upper slope tends to one and the lower to zero, approaching their uncoupled product labels.

In the negative-m block the two uncoupled energies would cross at $b = A/2$. The coupling opens a minimum gap $\sqrt{2}A$. Within the model, levels of the same m avoid one another whenever this matrix element is nonzero. Crossings between different m can remain exact because the uniform z-directed Hamiltonian does not mix them. A tilted perturbation could turn some of these into avoided crossings. Such symmetry dependence matters when trying to infer a coupling from the smallest visible distance between any two lines.

Obtain an eigenvector weight from an energy derivative

Suppose an initial s state has spin up. A weak $q = 0$ electric-dipole drive prepares the product basis state $|0,\uparrow\rangle$ in the $m = +1/2$ block. Its squared overlap with an exact eigenstate determines the angular part of that optical line's strength. Rather than computing a mixing angle with a branch ambiguity, differentiate the exact energies. Since $\partial H_+/\partial b = |0,\uparrow\rangle\langle 0,\uparrow|$, the eigenvalue derivative equals precisely that population:

$$w_\pm = \frac{dE_\pm}{db} = \frac{1}{2}\left[1 \pm \frac{b + A/2}{\sqrt{(b + A/2)^2 + 2A^2}}\right].$$

The weights sum to one as required by completeness. At zero field they are $2/3$ and $1/3$. At large positive field the upper line takes almost all this polarisation strength. In the negative block an analogous derivative with respect to $-b$ gives the population of $|0,\downarrow\rangle$. At its avoided crossing each eigenstate has equal weights and the two corresponding $q = 0$ intensities are equal for a selected initial spin down. The same experiment can test slopes measured from frequencies and overlaps measured through intensities, provided detection and radial coupling are common to both lines.

Set $b = A$ in the positive block. The level separation becomes $A\sqrt{4.25} = 2.06155A$, while the upper product weight is 0.8638. A naive extrapolation of the weak-field $g_j m$ slopes would give an interval $1.5A + A/3 = 1.833A$. The discrepancy is governed by the spin–orbit mixing scale rather than failure of the magnetic Hamiltonian itself. If the frequency analyser resolves only a single intensity-weighted centroid, completeness still constrains the first moment:

$$\sum_{\pm} w_{\pm} E_{\pm} = \langle 0,\uparrow | H_{+} | 0,\uparrow \rangle = b.$$

Thus the centroid of unresolved lines in this ideal selected excitation can move linearly even while each constituent bends.

Resolve a second angular-momentum hierarchy with the same algebra

The p-manifold example coupled orbital and electronic spin. In an s-state atom the dominant remaining weak angular coupling may be between an electronic $J = 1/2$ and a nuclear $I = 1/2$. This is a different scale and a different physical pair of degrees of freedom. To illustrate how a state label changes with field, take

$$H = \frac{A_h}{\hbar^2} \mathbf{I}\cdot\mathbf{J} + \frac{B}{\hbar}(g_J \mu_B J_z - g_I \mu_N I_z)$$

with positive hyperfine constant A_h and a positive model nuclear g_I. The minus before the nuclear-magneton term reflects the assumed positive nuclear magnetic moment. Actual sign and g-factor depend on the isotope. Here both A_h and the products of magnetic moments and field have energy units. At zero field the electronic and nuclear angular momenta combine to $F = 0, 1$. Their scalar-product shifts are $-3A_h/4$ and $+A_h/4$, respectively. A rotation of the entire coupled system leaves this zero-field Hamiltonian invariant, but independent rotations of just one constituent generally do not.

Within $m_F = 0$ order the product basis by electronic magnetic sign as $|\uparrow_J,\downarrow_I\rangle, |\downarrow_J,\uparrow_I\rangle$. Both diagonal hyperfine matrix elements are $-A_h/4$, off-diagonal exchange is $A_h/2$, and field changes the diagonal difference by $\delta = (g_J\mu_B + g_I\mu_N)B$. Hence the central pair of energies is

$$E_{\pm} = -A_h/4 \pm \tfrac{1}{2}\sqrt{A_h^2 + \delta^2}.$$

The two stretched $m_F = \pm 1$ levels sit at $+A_h/4 \pm (g_J\mu_B - g_I\mu_N)B/2$. Their slopes contain a difference of electronic and nuclear moments, while the central splitting depends on their sum. This sign

distinction can be determined directly by comparing the relevant product states: in the anti-aligned pair the sign change reverses one angular momentum along with its opposite magnetic sign. A sign convention imported from electron spin alone would give an erroneous nuclear contribution.

The transition between the two central states has gap $\sqrt{A_h^2 + \delta^2}$ and zero linear slope at $B = 0$. Its quadratic shift is $\delta^2/(2A_h)$ for $|\delta| \ll A_h$. Measuring at a point with zero first derivative reduces sensitivity to slow field fluctuations without eliminating it. Suppose $A_h/h = 1.00$ GHz and $(g_J\mu_B + g_I\mu_N)/h = 28.0$ GHz/T in an illustrative species model. At a 1.0 mT bias the central gap rises by about 0.392 MHz. A 1 µT field error at that bias gives a first-order frequency error of about 784 Hz using the derivative of the quadratic expansion. At zero mean bias, fluctuations with rms 1 µT instead give an average quadratic shift about 0.392 Hz. Stabilising the field mean and reducing its variance address different terms in the measured spectral position.

The magnetic drive operator for a small oscillating field along z is itself proportional in this block to σ_z. At zero bias the hyperfine eigenstates are the symmetric and antisymmetric superpositions, so this drive couples them off-diagonally despite the central transition frequency being stationary against a small static bias. The diagonal derivative and the off-diagonal transition matrix element are distinct quantities. At large bias the eigenstates approach product states and the same longitudinal drive becomes mostly diagonal, reducing its ability to flip between them. Thus a transition can have both a vanishing first-order shift and a nonzero resonant driving strength, consistent with the same Hamiltonian. Measuring both properties probes the field dependence of eigenvectors as well as eigenvalues.

Infer a field distribution from widths and centres

Spectra of spatially extended samples often contain a distribution of static magnetic fields. Consider a single weak-field transition with frequency $\nu(B) = \nu_0 + sB$, an intrinsic Gaussian line standard deviation w_0, and a sample field distributed approximately Gaussian with mean B_0 and standard deviation σ_B. A sum over independent source regions convolves the intrinsic line with the frequency distribution generated by sB. The observed centre is $\nu_0 + sB_0$, and variance is $w_0^2 + s^2\sigma_B^2$. A single linewidth cannot determine both an intrinsic width and a field inhomogeneity. Two lines with different known magnetic slopes or one calibrated zero-slope reference can separate the contributions.

For example two transitions with slopes 7.0 and 14.0 GHz/T have measured standard deviations 0.250 and 0.400 MHz. Subtract the squared widths to eliminate w_0^2:

$$\sigma_B^2 = \frac{(0.400)^2 - (0.250)^2}{(14000)^2 - (7000)^2}\ \mathrm{T}^2,$$

using MHz/T in the denominator. The result is $\sigma_B = 25.8\,\mu\mathrm{T}$, and the inferred intrinsic width is

$$\sqrt{0.250^2 - (7000\sigma_B)^2} = 0.173\,\mathrm{MHz}.$$

A third line with slope near zero supplies an out-of-sample check. If lines have significantly different lifetimes or power broadening, the assumption of a common w_0 should be relaxed; otherwise one can spuriously infer a magnetic gradient or even a negative intrinsic variance.

For a curved avoided-crossing line expand $\nu(B) \simeq \nu(B_0) + s(B - B_0) + q(B - B_0)^2/2$. A symmetric distribution shifts the measured mean by $q\sigma_B^2/2$ even when the mean field is known exactly. Near a stationary point the leading field broadening can be strongly asymmetric because a squared Gaussian variable has a one-sided distribution around its minimum. Fitting a symmetric line at such a point can shift its apparent peak or centre according to the fitting convention. This effect is especially relevant when locating a minimum gap from spectra at several nominal fields. A model integrating intensities over a measured spatial field distribution can be better constrained than independent fits of Gaussian centres, because all lines share the same underlying field profile.

An independent geometric test tilts the weak sensing light relative to the magnetic axis. For equal populations, a truly scalar average over all polarisations and emitted directions is insensitive to this coordinate change, but a selected q transition changes strength according to the decomposition of the laboratory polarisation. If a peak attributed to frequency splitting mainly changes its fitted centre while unresolved component intensities redistribute, some of its apparent shift may be an intensity-weighted centroid effect. Record spectra at two orthogonal polarisations, fit the same physical energy parameters and allow the predicted angular strengths to vary. A common energy correction that survives the intensity change carries a different interpretation from a change in the mixture of unresolved substates.

Predict state transfer during a field sweep

The negative-m avoided crossing of the p manifold also poses a time-dependent question. Does a changing field keep the atom in the instantaneous energy eigenstate, or in approximately the same product state? In the local two-state description subtract the trace and write $H = (\Delta(t)/2)\sigma_z + g\sigma_x$, where the difference of diabatic diagonal energies is Δ and the minimal adiabatic gap is $2|g|$. Introduce a mixing angle $\tan(2\vartheta) = 2g/\Delta$, chosen continuously through the crossing. Differentiating at fixed g gives

$$|\dot{\vartheta}| = |g\dot{\Delta}|/(\Delta^2 + 4g^2).$$

The moving eigenbasis itself couples the two instantaneous energy states with an energy scale $\hbar|\dot{\vartheta}|$.

At closest approach the order-of-magnitude adiabatic criterion becomes $\hbar|\dot{\Delta}|/(8g^2) \ll 1$. A slow sweep keeps occupation on an energy branch, which changes its product character across the avoided crossing. A fast sweep can leave the original product state approximately unchanged, moving population from one instantaneous energy branch to the other. Both descriptions refer to actual state transfer; the basis in which “remaining in the same state” is assessed is part of the observable. The criterion here follows directly by comparing the nonadiabatic derivative coupling with the instantaneous level separation. It estimates a regime but does not by itself give a transition probability.

For a strictly linear detuning passing from large negative to large positive values with constant g, the Landau–Zener result gives the diabatic survival probability $P_D = \exp[-2\pi g^2/(\hbar|\dot{\Delta}|)]$. Its applicability requires asymptotic start and end detunings well outside the coupling region, coherent passage and the indicated linearity over the crossing. In ordinary-frequency units take $g/h = 20$ kHz and $|d(\Delta/h)/dt| = 1.0 \times 10^{10}\ \mathrm{Hz/s}$. The exponent is $4\pi^2(2.0 \times 10^4)^2/10^{10} = 1.58$, giving $P_D \simeq 0.206$. Doubling the sweep rate halves the exponent and raises survival to about 0.454. Confusing angular with

ordinary frequency in this conversion loses a factor 2π, large enough to change the predicted population qualitatively.

A double sweep forms a useful control. On entering and later leaving the avoided crossing, the two amplitudes propagated along different energy branches can accumulate different phases and interfere. Treating the two passages as independent classical jumps would miss oscillations versus waiting time between them. Strong dephasing or averaging over a broad field distribution can remove the oscillations while retaining the one-pass populations. Distinguish these mechanisms by calibrating single-passage transfer first and then varying the waiting time over several periods at multiple sweep speeds. This extends a static magnetospectroscopy model to a specified dynamical experiment without inferring the full time dependence from a static level diagram alone.

Resolve the crossover from coupled to uncoupled angular momenta

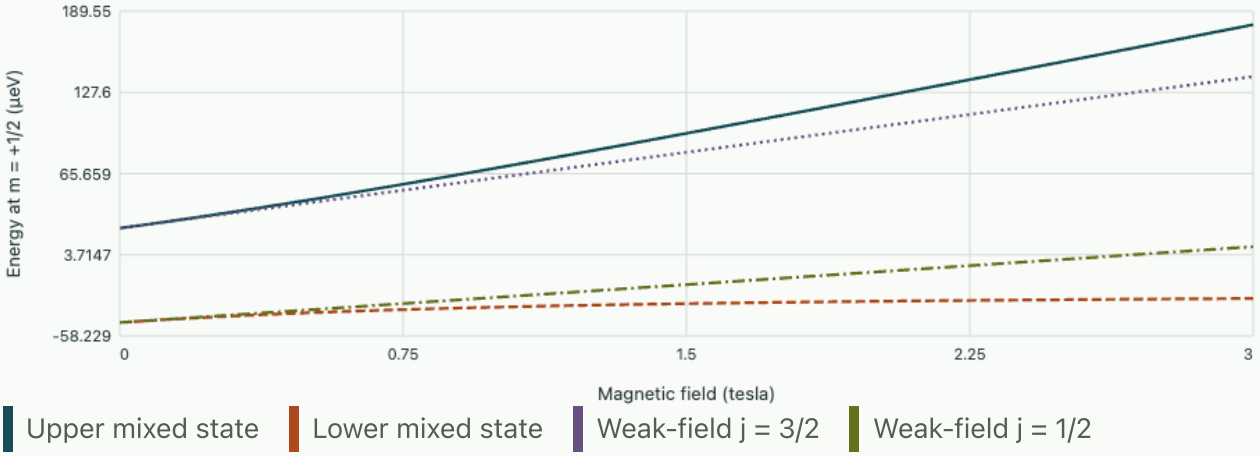
Scan the magnetic field across the spin–orbit scale. Compare the exact two-state spectrum at fixed magnetic quantum number with the weak-field shifts.

Open this experiment online

Try the calculation

Spin–orbit coefficient A (μeV) **48**

Selected magnetic field (T) **0.6**



Selected energies -39.516 and 50.246 μeV. The ratio μB B/A is 0.72355.

Test the orientation of a magnetic sensor from resolved spin populations

The scalar energy spectrum of an isotropic spin-$1/2$ Zeeman Hamiltonian measures field magnitude. A transition energy by itself cannot determine the signed orientation of that field. Let $H = \hbar\gamma\mathbf{B}\cdot\boldsymbol{\sigma}/2$, with signed angular gyromagnetic ratio γ defined in the experiment. Its eigenvalues are $\pm\hbar|\gamma||\mathbf{B}|/2$. If instead one prepares spin up along a known laboratory z direction and observes its population after evolution time t, the rotation axis appears in the dynamics. Exponentiating $U = \cos(\omega t/2)I - i\sin(\omega t/2)\hat{\mathbf{B}}\cdot\boldsymbol{\sigma}$ for a consistent choice of $\omega, \hat{\mathbf{B}}$ gives

$$P_{\downarrow z}(t) = (1 - \hat{B}_z^2)\sin^2(\gamma|\mathbf{B}|t/2).$$

The oscillation frequency gives magnitude and its contrast gives the squared transverse fraction. It is insensitive to whether the field points slightly above or below a perpendicular plane. Incomplete preparation or pulse contrast can also reduce the observed oscillation amplitude. A calibration with a known transverse field tests that ambiguity before assigning every missing contrast to geometric tilt.

Suppose a resolved precession has angular frequency corresponding to 0.50 mT and a calibrated flip contrast of 0.36 for initial $+z$. Then $\sqrt{B_x^2 + B_y^2} = 0.30$ mT and $|B_z| = 0.40$ mT. Repeating for an initial $+x$ state with calibrated contrast 0.64 yields $|B_x| = 0.30$ mT and thus $B_y = 0$ for these numbers. Neither experiment identifies the signs of B_x and B_z. To resolve them one can measure an early-time transverse quadrature with a phase reference. In Bloch notation $\dot{\mathbf{r}} = \gamma\mathbf{B}\times\mathbf{r}$, so the initially zero slope $\dot{r}_y = -\gamma B_x$ for $+z$ is sign sensitive. This derivative is independent of B_z at the initial instant, giving a direct check complementary to the energy splitting.

A second strategy adds a calibrated bias $\mathbf{b}_0$ in opposite directions and examines squared frequencies. Since $|\mathbf{B}+\mathbf{b}_0|^2 - |\mathbf{B}-\mathbf{b}_0|^2 = 4\mathbf{B}\cdot\mathbf{b}_0$, the difference isolates a signed projection even when the zero-field splitting reveals only magnitude. For a 0.20 mT known bias along z and a 0.40 mT signed z component the squared-field difference should be $0.32\,\mathrm{mT}^2$. Alternating the two bias signs across the measurement reduces slow field drift. A quadratic diamagnetic response, anisotropic g tensor or an uncalibrated bias axis can break this simple inversion; repeating in three independent directions tests the vector model and reveals which assumptions are being used.

Separate an anisotropic magnetic response from an unknown field direction

A crystal-fixed effective spin has diagonal $g = \mathrm{diag}(2.0, 2.0, 2.4)$. Derive the transition energy for a field rotated in the x, z plane by angle θ from z, and explain what an isotropic fit would infer.

▼ **Derive and check the result**

The Hamiltonian is $\mu_B\boldsymbol{\sigma}\cdot(g\mathbf{B})/2$. The splitting is $\mu_B B\sqrt{2.0^2\sin^2\theta + 2.4^2\cos^2\theta}$. At 45° the effective g is $\sqrt{4.88} = 2.209$; interpreting the measured splitting with constant $g = 2$ would overestimate field magnitude by about 10.5%. The spin also precesses around the effective axis $g\mathbf{B}$, whose polar angle satisfies

$$\tan\theta_{\mathrm{eff}} = (2.0/2.4)\tan\theta.$$

Measuring contrast from a calibrated spin preparation and rotating a field of known magnitude can therefore distinguish material anisotropy from a simple error in magnet alignment. One scalar spectral line at an unknown orientation cannot separate these effects.

Tutorial · reconstruct a magnetically mixed spectrum

Infer the coupling from an anticrossing

The two negative-m lines exhibit a minimum frequency separation of 11.31 GHz at 0.286 T. Infer A/h from each measurement. What calibration would be needed for a comparison of absolute optical line frequencies?

▼ Work through the solution

Minimum gap divided by h equals $\sqrt{2}A/h$, giving $A/h = 8.00$ GHz. The crossing condition $b = A/2$ gives

$$A/h = 2(13.996)(0.286) = 8.01\,\mathrm{GHz}.$$

Their agreement checks field calibration independently of a common frequency reference. Both final states must be viewed from the same initial state when using their frequency difference. Comparing absolute optical frequencies also requires the zero-field transition energy and the field-dependent initial-level shift. Two lines originating in different initial states would have a field-dependent separation even if their final levels had a fixed interval.

Extract the polarisation weight using a finite difference

Near $b = A$, estimate the slope of the upper positive-m level using its exact values at $b/A = 0.95, 1.05$. Compare with the exact weight and predict the second line's strength.

▼ Work through the solution

In units of A the upper level is $(t - 0.5 + \sqrt{(t+0.5)^2 + 2})/2$ with $t = b/A$. Subtracting the values and dividing by 0.10 gives about 0.864. The analytical derivative is

$$[1 + 1.5/\sqrt{4.25}]/2 = 0.8638.$$

The second line carries 0.1362 of the total angular strength. Halving the centred step should reduce the derivative's truncation error by about four for this smooth curve. In an experiment frequency noise per

point is instead amplified as inverse step. Compare several field steps to establish a practical slope uncertainty.

Use an unresolved centroid to detect a calibration problem

At $b = A$ a weak $q = 0$ experiment driven from a selected spin-up preparation records both positive-m final states. Its excitation-weighted centroid after restoring the initial-state energy is $0.78A$. Could the exact two-level mixing alone produce that centroid?

▼ **Work through the solution**

The sum rule requires $b = A$. Direct calculation gives energies approximately $1.2808A$ and $-0.7808A$; weights 0.8638 and 0.1362 give $1.000A$ within rounding. In this model mixing alone cannot explain $0.78A$. A missing weak line, saturation, polarisation contamination, different collected fluorescence efficiencies or a mistaken correction for the initial state can shift a measured centroid. Compare both individual integrated line areas at low intensity and reverse the field with the corresponding spin preparation. These tests use distinct symmetry and strength predictions rather than interpreting one centroid as a complete spectrum.

Further reading

- Cambridge · magnetic fields, Zeeman structures and optical transitions
- Course investigation · angular momentum and atoms
- Fließbach · Näherungsmethoden

Fluctuations, occupation and collective order

Connect an occupation probability to a heat capacity, a correlation length or a symmetry-breaking order parameter. Compare electrical measurements with a thermal fluctuation.

1. Indistinguishable particles and quantum gases
2. Vibrations and the Debye heat capacity
3. Extracting a length from thermal spin fluctuations
4. Why only a narrow shell of a Fermi gas carries heat
5. Mean-field ordering, response and the critical point
6. Measure equilibrium field fluctuations with a capacitor
7. From mobile ions to screening lengths and differential double-layer capacitance
8. Rotational and vibrational populations in a molecular gas
9. Bose condensation, dimensionality and a finite trap
10. Degenerate fermions, heat and magnetic response
11. Finite-size criticality and the evidence for a phase transition
12. How many lattice sites remember the colour of a neighbour?

Indistinguishable particles and quantum gases

Thermodynamics and statistical physics

Use single-state occupancy to derive Fermi–Dirac and Bose–Einstein distributions and estimate when classical gas statistics cease to apply.

FROM FIRST PRINCIPLES

Derive the mean population of one quantum state

Consider a single-particle state of energy ϵ exchanging particles and energy with a large reservoir at temperature T and chemical potential μ. In a noninteracting gas, occupation n of this state changes $E - \mu N$ by $n(\epsilon - \mu)$. Its statistical weight is q^n, where $q = e^{-(\epsilon-\mu)/(k_B T)}$. For one specified state including its spin label, fermions allow occupations 0 and 1. Bosons allow every nonnegative integer,

$$Z_F = 1 + q,$$

$$Z_B = 1 + q + q^2 + \cdots = \frac{1}{1-q}.$$

The Bose sum requires $q < 1$. Differentiating a weight with respect to q brings down its occupancy, so the normalized average is $q\,\partial_q \ln Z$. Applying this to the two sums gives

$$\bar{n}_F = \frac{1}{e^{(\epsilon-\mu)/(k_B T)} + 1},$$

$$\bar{n}_B = \frac{1}{e^{(\epsilon-\mu)/(k_B T)} - 1}.$$

For $q \ll 1$, both means approach q; visits by additional particles to the same state are then rare. The difference between the distributions follows entirely from which occupation numbers the quantum statistics allow.

DERIVATION AND APPLICATIONS

Derive the first exchange correction to ideal-gas pressure

Before quantum degeneracy becomes strong, exchange statistics already affect the relation between pressure and density. Consider noninteracting particles of mass m in a large three-dimensional box. All g internal states have the same energy and chemical potential. Introduce fugacity $z = e^{\mu/(k_B T)}$, thermal wavelength $\lambda_T = h/\sqrt{2\pi m k_B T}$, and a sign $s = +1$ for bosons or -1 for fermions. At small fugacity the occupation of a one-particle energy ϵ is a geometric expansion,

$$\bar{n}(\epsilon) = z e^{-\beta\epsilon} + s z^2 e^{-2\beta\epsilon} + O(z^3),$$

$$\beta = (k_B T)^{-1}.$$

Integrating the first term over momentum gives g/λ_T^3 times z. In the second term, doubling the inverse temperature narrows each Cartesian Gaussian by $\sqrt{2}$. The density integral therefore gains a factor $2^{-3/2}$. At fixed temperature the Gibbs–Duhem relation $dp = n\,d\mu$ gives

$$d[p/(k_B T)] = n\,dz/z.$$

Integration from the empty-gas limit gives the pressure series,

$$\frac{n\lambda_T^3}{g} = z + \frac{sz^2}{2^{3/2}} + O(z^3),$$
$$\frac{p\lambda_T^3}{g k_B T} = z + \frac{sz^2}{2^{5/2}} + O(z^3).$$

Experiments often fix density. Put $y = n\lambda_T^3/g$ and invert the density series once to get

$$z = y - sy^2/2^{3/2} + O(y^3).$$

Substituting into the pressure eliminates the unmeasured chemical potential,

$$\frac{p}{nk_B T} = 1 - \frac{s}{2^{5/2}}\frac{n\lambda_T^3}{g} + O(y^2).$$

At fixed density exchange raises the pressure of fermions and lowers that of bosons. The magnitude grows on cooling because the thermal wavelength increases. Real interactions can also alter the pressure, and the truncated expansion requires small y.

WORKED EXAMPLE

Estimate a pressure excess for a dilute Fermi gas

Change the inputs and investigate this derivation.

A spin-balanced ideal Fermi gas has $g = 2$, $n = 2.00 \times 10^{18}\,\mathrm{m}^{-3}$ and $\lambda_T = 400\,\mathrm{nm}$. It has $y = 0.0640$ and $p/(nk_B T) \simeq 1.0113$. A measured classical pressure scale $nk_B T = 20.0\,\mathrm{pPa}$ would therefore acquire an exchange correction of about $0.226\,\mathrm{pPa}$.

SECOND APPLICATION

Recover density from a bosonic pressure deficit

Change the inputs and investigate this derivation.

A single-state Bose gas at known $\lambda_T = 500\,\mathrm{nm}$ gives

$$p/(nk_B T) = 0.9950.$$

To leading order the deficit fixes $y = 0.0050 \times 2^{5/2} = 0.0283$. The inferred density is

$$y/\lambda_T^3 = 2.26 \times 10^{17}\,\mathrm{m}^{-3}.$$

At unchanged temperature halving this density halves the fractional pressure deficit, whereas the absolute pressure correction falls by a factor of four.

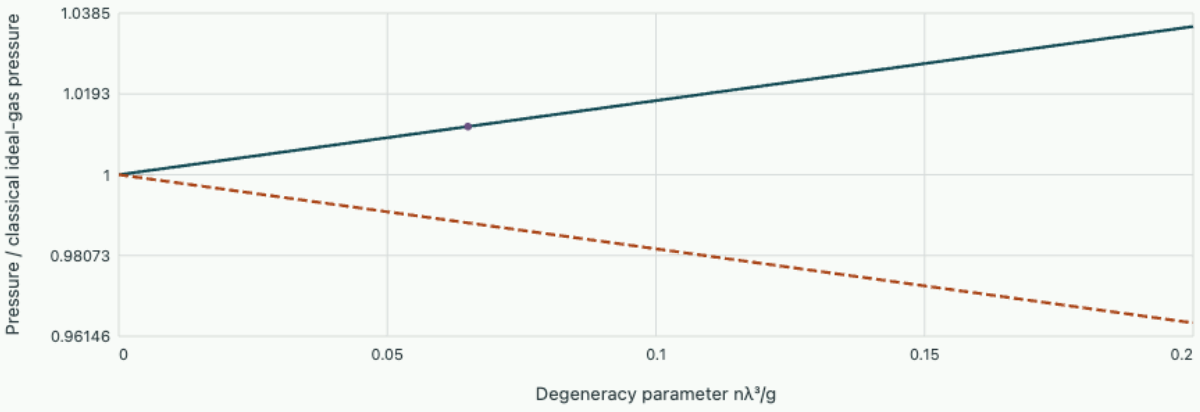
Try it · Quantum virial

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

Selected degeneracy nλ³/g **0.065**



These are first quantum virial corrections, valid for nλ³/g≪1. Exchange raises the dilute fermion pressure and lowers the dilute boson pressure.

Decide which exchanges leave the state unchanged

An ideal classical gas can often be treated with Maxwell–Boltzmann statistics. Cooling the gas or increasing its density eventually makes wave packets overlap appreciably. For indistinguishable quantum particles, merely assigning names to exchanged particles must not create extra states. In an ideal gas one can instead specify the occupation of each single-particle state. Identical fermions occupy a given quantum state at most once, including its spin label. Identical bosons can accumulate in the same state.

Use a reservoir that exchanges both energy and particles with the system. Its temperature T and chemical potential μ are fixed. A microstate with energy E and particle number N has grand-canonical weight proportional to $\exp[-\beta(E-\mu N)]$. This follows by applying the thermodynamic entropy differential to a reservoir that loses both E and N. The chemical potential sets the mean particle number through these statistical weights.

Sum the allowed occupancies of one state

For a single-particle state of energy ϵ, define $z_\epsilon = \exp[-\beta(\epsilon-\mu)]$. An occupancy of n contributes weight z_ϵ^n. For fermions the single-state grand partition sum is $1+z_\epsilon$, so the mean occupancy is $z_\epsilon/(1+z_\epsilon)$. For bosons the sum runs over all nonnegative integers and equals $1/(1-z_\epsilon)$, provided $z_\epsilon<1$. Differentiating the logarithm of each sum with respect to $\ln z_\epsilon$ gives

$$\bar{n}_F = \frac{1}{e^{\beta(\epsilon-\mu)}+1},$$

$$\bar{n}_B = \frac{1}{e^{\beta(\epsilon-\mu)}-1}.$$

These expressions describe one state, not the total population within an energy interval. To get the latter, multiply by the density of states and integrate. If the exponential is large then both occupancies reduce to $e^{-\beta(\epsilon-\mu)}$. The upper sign suppresses double occupancy while the lower sign enhances occupation near the lowest available energy.

Count the occupied sphere of a cold electron gas

In a large cubic volume V with periodic boundaries, allowed wave vectors fill reciprocal space with one state per volume $(2\pi)^3/V$ for each spin. At zero temperature the noninteracting electron gas fills all states inside a sphere of radius k_F and leaves higher ones empty. Including two spin states per wave vector gives $N = 2V(4\pi k_F^3/3)/(2\pi)^3$, so $k_F = (3\pi^2 N/V)^{1/3}$. For electrons of mass m_e, the Fermi energy is $\hbar^2 k_F^2/(2m_e)$.

With density $n = 8.0\times10^{28}$ m⁻³, $k_F = 1.33\times10^{10}$ m⁻¹ and $E_F \simeq 6.77$ eV. The associated temperature $T_F = E_F/k_B$ is approximately 7.9×10^4 K. At room temperature k_BT is small compared with this energy, so only a narrow region around the Fermi surface changes occupation noticeably. A classical equipartition estimate of the electron contribution to heat capacity would ignore this restriction.

Test the dilute limit with a cold atomic gas

A useful scale for translational wave packets is the thermal wavelength $\lambda_T = h/\sqrt{2\pi m k_B T}$. For a single internal state the dimensionless product $n\lambda_T^3$ compares the volume per particle with the thermal wave volume. Small values support the dilute classical approximation. Near unity the occupancy of low-energy states can become appreciable and the statistics of indistinguishability matter.

For particles of mass 6.6×10^{-27} kg at 1.0 K the thermal wavelength is approximately 0.88 nm. At density 10^{25} m⁻³, $n\lambda_T^3 \simeq 0.0068$. Reducing temperature to 0.040 K at unchanged density multiplies that product by $25^{3/2} = 125$, giving about 0.85. This is a degeneracy estimate, not a prediction of the phase of a real interacting gas. Once quantum statistics matter, particle interactions, dimensionality and confinement can also affect which phase the experiment actually finds.

Internal degeneracy enters these comparisons through the density per available internal state. Suppose the same total particle density is spread equally across two spin states with identical spatial dispersion. Each can fill the same momentum levels independently. In a cold ideal Fermi gas the Fermi wavevector at fixed total density is then smaller than in a fully spin-polarized gas. A measured pressure change after altering spin populations can therefore contain a statistical contribution even if the mass, volume and total particle count remain fixed.

Determine the capacity of excited Bose states

Ideal bosons can accumulate macroscopically in one spatial state. Consider a uniform three-dimensional noninteracting gas of identical spinless particles with fixed particle number in a large box. Set the lowest energy to zero. The excited-state occupation at energy ϵ is $1/[\exp((\epsilon-\mu)/(k_B T))-1]$ and requires $\mu \le 0$. In the thermodynamic limit, counting momentum states in a sphere gives an excited-state density per volume proportional to $\sqrt{\epsilon}$. Consequently, as the density rises, taking μ toward zero from below can increase the excited population only up to a finite value.

Explicitly integrate $d^3p/(2\pi\hbar)^3$ over positive energies $\epsilon = p^2/(2m)$. At $\mu = 0$, setting $x = \epsilon/(k_B T)$ reduces the integral to $\int_0^\infty x^{1/2}/(e^x-1)dx = \Gamma(3/2)\zeta(3/2)$. The result is

$$n_{\mathrm{ex,max}} = \frac{\zeta(3/2)}{\lambda_T^3},$$

$$\lambda_T = \frac{h}{\sqrt{2\pi m k_B T}},$$

$$\zeta(3/2) = 2.6124\ldots.$$

Below the temperature where this capacity equals the total density, additional particles occupy the lowest state. Solving gives

$$T_c = (2\pi\hbar^2/m k_B)[n/\zeta(3/2)]^{2/3},$$

$$N_0/N = 1 - (T/T_c)^{3/2}.$$

The state count explains why a high phase-space density matters. Reaching this regime does not require individual particles to lose their identity at a sharply defined minimum separation.

Estimate the onset temperature in a box

Take $n = 1.00 \times 10^{20}$ m⁻³ and particle mass 1.44×10^{-25} kg. At the ideal transition the thermal wavelength is

$$(2.6124/n)^{1/3} = 297\,\mathrm{nm}.$$

Inverting its definition gives $T_c = h^2/(2\pi m k_B \lambda_T^2) = 0.398$ µK. At $T = 0.250$ µK, the excited fraction is $(0.250/0.398)^{3/2} = 0.498$ and the condensate fraction is approximately 0.502. Doubling the density raises T_c by $2^{2/3} = 1.59$ at fixed particle mass.

The predicted fraction tends to zero continuously as T approaches T_c from below, and tends to one as temperature approaches zero for the ideal gas. A harmonic trapping potential has a different density of states and hence a different temperature exponent from a uniform box. Finite size smooths the mathematical onset; interactions and inaccurate thermometry can shift its inferred location. Comparing the number in the narrow lowest-momentum component with the thermal-cloud population requires that the analysis resolve these two contributions rather than simply count all particles near the centre of a spatial image.

The dimensionality of the gas is as essential as Bose exchange itself. In a uniform two-dimensional ideal box the single-particle density of states tends to a constant at low energy. At chemical potential approaching zero, Bose occupation varies as $k_B T/\epsilon$ there. Its energy integral consequently diverges logarithmically rather than approaching a finite excited-state capacity. The three-dimensional derivation of a nonzero transition temperature cannot then be carried across unchanged. A finite box, a confining trap or interparticle interactions can all alter the observed low-energy behaviour and must be identified when comparing different experimental geometries.

Repeat the counting in a harmonic trap

A harmonic trap changes the number of available excited states even for particles with the same mass and statistics. For three oscillator frequencies $\omega_x, \omega_y, \omega_z$, measure energy above the lowest level and suppose $k_B T$ greatly exceeds each $\hbar\omega_i$. The quantum labels n_i are nonnegative integers with energy $\sum_i \hbar\omega_i n_i$. The count below energy E is approximately the volume of the positive-coordinate tetrahedron bounded by this plane, namely $E^3/(6\hbar^3\omega_x\omega_y\omega_z)$. Differentiating gives an energy density of states proportional to E^2.

At the onset of condensation integrate this density against $1/(e^{E/k_B T} - 1)$. The identity $\int_0^\infty x^2/(e^x - 1)\,dx = 2\zeta(3)$ then gives

$$N_{\mathrm{ex,max}} = \zeta(3)\left(\frac{k_B T}{\hbar\bar{\omega}}\right)^3,$$

$$\bar{\omega} = (\omega_x\omega_y\omega_z)^{1/3}.$$

For $N = 1.00 \times 10^5$ particles in an isotropic 100 Hz trap, $\hbar\bar{\omega}/k_B = 4.80$ nK and $[N/\zeta(3)]^{1/3} = 43.7$, so the predicted transition temperature is 0.210 µK. At 0.150 µK the condensate fraction is

$$1 - (0.150/0.210)^3 = 0.636.$$

An analysis using the uniform-box exponent at the same temperature ratio would instead give 0.396. The difference comes entirely from the confinement-dependent state count.

Reducing all three trap frequencies at fixed particle number lowers the onset temperature in direct proportion. This differs from enlarging a uniform box at fixed mean density, which leaves its bulk critical temperature unchanged. Specifying how particle number, volume and external confinement are controlled is therefore essential when interpreting a measured condensate fraction.

Vibrations and the Debye heat capacity

Thermodynamics and statistical physics

Count vibrational modes in a solid and explain why a crystal's low-temperature heat capacity follows a cube law.

FROM FIRST PRINCIPLES

Derive the low-temperature heat capacity from acoustic modes

In a large three-dimensional crystal, long-wavelength acoustic vibrations have approximately linear dispersion $\omega = vk$. Count allowed wavevectors inside a sphere of radius k. Its volume grows as k^3, so the number of modes below ω grows as ω^3. The frequency density is its derivative. In a single-speed model normalized to the $3N$ displacements of N atoms,

$$g(\omega) = \frac{9N\omega^2}{\omega_D^3},$$

$$\int_0^{\omega_D} g(\omega)\, d\omega = 3N.$$

Each mode stores mean thermal energy $\hbar\omega/[e^{\hbar\omega/(k_BT)} - 1]$. Multiply by the mode count and substitute $x = \hbar\omega/(k_BT)$ to expose the temperature dependence,

$$U_{\mathrm{th}} = \frac{9N(k_BT)^4}{\hbar^3\omega_D^3} \int_0^{\hbar\omega_D/(k_BT)} \frac{x^3}{e^x - 1}\, dx.$$

When $T \ll \hbar\omega_D/k_B$, the omitted tail above the upper limit is exponentially small. The integral approaches a constant, giving $U_{\mathrm{th}} \propto T^4$ and $C_V = (\partial U/\partial T)_V \propto T^3$. The cube law relies on both three-dimensional mode counting and linear low-frequency dispersion.

DERIVATION AND APPLICATIONS

Recover the Debye velocity from distinct sound branches

A three-dimensional solid usually carries one longitudinal and two transverse acoustic polarizations. Their speeds can differ substantially. Consider a large elastically isotropic crystal of volume V, with linear dispersion $\omega = v_s k$ for each polarization s in the thermally populated range. Impose periodic boundaries; one wavevector state occupies reciprocal volume $(2\pi)^3/V$. For each polarization count all points in a sphere of radius ω/v_s, which includes opposite propagation directions. Differentiating gives the total density of modes,

$$N_s(<\omega) = \frac{V\omega^3}{6\pi^2 v_s^3},$$

$$g(\omega) = \frac{V\omega^2}{2\pi^2}\left(v_L^{-3} + 2v_T^{-3}\right).$$

A slower sound branch has more low-frequency states because the same frequency reaches a larger wavevector sphere. Define one effective velocity v_D to preserve this acoustic density of states. Requiring an isotropic monatomic Debye model to contain $3N$ vibrations, with number density $n = N/V$, fixes its frequency cutoff and temperature scale,

$$\frac{3}{v_D^3} = \frac{1}{v_L^3} + \frac{2}{v_T^3},$$

$$\omega_D = v_D(6\pi^2 n)^{1/3},$$

$$\Theta_D = \frac{\hbar\omega_D}{k_B}.$$

At low temperature the cutoff can be moved to infinity in the thermal energy integral. Changing variables to $x = \hbar\omega/(k_B T)$ uses $\int_0^\infty x^3/(e^x - 1)\,dx = \pi^4/15$. A temperature derivative yields the directly measurable coefficient

$$\frac{C_V}{VT^3} = \frac{2\pi^2 k_B^4}{15\hbar^3}\left(v_L^{-3} + 2v_T^{-3}\right).$$

The inverse cubes explain why replacing the three velocities by their arithmetic mean can noticeably underestimate the heat capacity. Optical or electronic contributions must be identified separately in specimens where they are thermally active.

WORKED EXAMPLE

Predict the acoustic temperature scale

Change the inputs and investigate this derivation.

For $v_L = 6000\,\mathrm{m/s}$, $v_T = 3000\,\mathrm{m/s}$, the longitudinal branch contributes only $1/17$ of the low-temperature coefficient. The effective velocity is

$$3000(24/17)^{1/3} = 3365\,\mathrm{m/s}.$$

With $n = 5.00 \times 10^{28}\,\mathrm{m^{-3}}$, the cutoff wavevector is $1.44 \times 10^{10}\,\mathrm{m^{-1}}$, giving $\Theta_D \simeq 369\,\mathrm{K}$. The long-wave estimates apply well below that scale.

SECOND APPLICATION

Extract the transverse speed from calorimetry

Change the inputs and investigate this derivation.

Suppose a separate specimen gives $v_D = 2500\,\mathrm{m/s}$ from its low-temperature heat capacity and $v_L = 5000\,\mathrm{m/s}$ from a longitudinal pulse. Solving for the missing speed gives

$$v_T = [2/(3/v_D^3 - 1/v_L^3)]^{1/3} = 2.22 \times 10^3\,\mathrm{m/s}.$$

A shear-wave transit across $10.0\,\mathrm{mm}$ should consequently take $4.51\,\mu\mathrm{s}$. This timing test uses a distinct mechanical measurement of the same states counted by calorimetry.

Try it · Debye speed

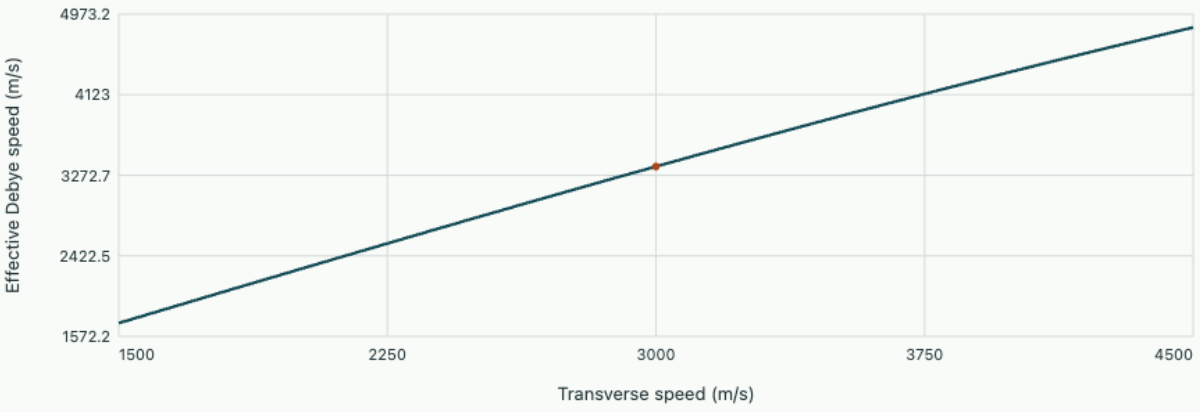
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Longitudinal speed (m/s) **6000**

Transverse speed (m/s) **3000**



Three acoustic polarisations Selected material

1/vD³=(1/vL³+2/vT³)/3. Longitudinal fraction of the low-temperature heat coefficient 0.058824.

What a solid can store

Consider an electrically insulating three-dimensional crystal of N atoms, volume V, sufficiently cold that its structure remains fixed and normal-mode frequencies vary weakly with temperature. In a harmonic approximation displacements about equilibrium separate into $3N$ independent oscillators. Their excitations are called phonons. Phonon number is not conserved when the solid exchanges heat, so each mode has a Bose mean occupation with zero chemical potential. The modes exist because the crystal can deform, including at very long wavelengths.

A model assigning the same frequency to every oscillator predicts an exponentially small low-temperature heat capacity, since exciting any mode would require a fixed nonzero energy. Acoustic vibrations instead approach zero frequency as their wavenumber approaches zero. Even in a cold macroscopic crystal, long-wavelength modes can remain thermally accessible. The key quantity is how many have energies of order $k_B T$ or less.

Test the single-frequency model before introducing a spectrum

The premise of the Einstein solid is deliberately restrictive. All $3N$ vibrational coordinates of a crystal are independent harmonic oscillators with one frequency ω_E. For one coordinate the levels are $E_n = \hbar\omega_E(n + 1/2)$. Write $x = \hbar\omega_E/(k_B T) = \Theta_E/T$. Summing the geometric series gives

$$Z_1 = e^{-x/2}/(1 - e^{-x}).$$

The partition function of the crystal is Z_1^{3N}, because each normal coordinate can be occupied independently and phonon number is free to change.

Apply $-\partial \ln Z/\partial\beta$, with $\beta = 1/(k_B T)$, to obtain

$$U = 3N\hbar\omega_E[1/2 + 1/(e^x - 1)].$$

The zero-point energy contributes no heat capacity at fixed volume. Differentiating the occupation term and using $dx/dT = -x/T$ gives

$$C_{V,E} = 3Nk_B \frac{x^2 e^{-x}}{(1 - e^{-x})^2}.$$

At high temperature $1 - e^{-x} \simeq x$, so the heat capacity per mole of atoms approaches $3R$. At low temperature the denominator tends to one, and the surviving factor $x^2 e^{-x}$ measures the cost of the lowest available excitation. For $\Theta_E = 240$ K the same mole would have $C_V = 18.06$ J/(mol K) at 120 K, 7.585 at 60 K and 0.536 at 30 K. Halving temperature does not give a fixed ratio. The curve crosses from a classical plateau to thermally rare excitations.

A real three-dimensional macroscopic crystal supports collective vibrations whose frequencies approach zero with wavelength increasing. Replacing the single frequency by a distribution puts modes below any fixed nonzero thermal scale. The next calculation finds the density of those modes, and therefore a low-

temperature power law. High-temperature agreement alone cannot distinguish the two models since both count the same total number of coordinates.

Count to a cutoff

Use a continuum approximation with three acoustic polarizations and a common sound speed v. The modes in a shell of wavevectors grow as $k^2 dk$, as for photons. Real crystals generally have distinct longitudinal and transverse speeds. An effective Debye speed averages their inverse cubes. Converting via $\omega = vk$ gives density $g(\omega) = 3V\omega^2/(2\pi^2 v^3)$. The continuum would supply infinitely many modes, whereas a crystal with N atoms has exactly $3N$. Choose an upper cutoff ω_D to match that count,

$$\int_0^{\omega_D} g(\omega)\, d\omega = 3N,$$

$$g(\omega) = 9N\omega^2/\omega_D^3.$$

This cutoff is an approximate representation of the lattice at high wavenumber. The low-temperature prediction depends on the long-wave acoustic spectrum and is insensitive to precise details near the cutoff. In a multi-atom cell additional optical branches change the intermediate-temperature behaviour, which cannot generally be reconstructed with a single effective speed.

Two limits and a measurement

The thermal energy is the integral of $g(\omega)\hbar\omega/[\exp(\hbar\omega/k_B T) - 1]$ over the band. Define $\Theta_D = \hbar\omega_D/k_B$. When $T \ll \Theta_D$, extend the upper limit of the dimensionless integral to infinity. The resulting energy scales as T^4 and differentiation yields

$$C_V = \frac{12\pi^4}{5} N k_B \left(\frac{T}{\Theta_D}\right)^3.$$

At high temperature each mode instead approaches $k_B T$ of thermal energy and the heat capacity tends to $3Nk_B$. The acoustic model has the correct count by construction. The constant zero-point energy never enters either heat capacity because it disappears upon differentiation.

For one mole of atoms with $\Theta_D = 300$ K, the low-temperature prediction at 10.0 K is 0.0720 J mol⁻¹ K⁻¹, far below the classical $3R = 24.9$ J mol⁻¹ K⁻¹. At 20.0 K the cubic asymptote predicts 0.576 J mol⁻¹ K⁻¹. Measuring several temperatures and plotting C/T^3 tests the constant coefficient directly. In a conducting solid an electronic term proportional to T can dominate at low temperatures. Separating that term through a plot of C/T against T^2 allows the intercept and slope to reveal the two physically distinct energy stores.

Finite size eventually supplies a lower frequency cutoff related to the crystal dimensions, so the continuum counting argument presupposes a specimen large enough to have many wavelengths populated throughout the temperature interval.

Connect vibrations to thermal expansion

A purely harmonic crystal with force constants independent of volume can store vibrational energy yet has no vibrational mechanism for thermal expansion. Real vibrational frequencies usually vary as the lattice spacing changes. The quasiharmonic approximation treats each mode as harmonic at each volume while allowing its angular frequency $\omega_j(V)$ to change smoothly. Define the mode Grüneisen parameter $\gamma_j = -\partial \ln \omega_j / \partial \ln V$. A positive value means expansion softens that mode.

For one oscillator the Helmholtz free energy is

$$F_j = \hbar\omega_j/2 + k_B T \ln[1 - e^{-\hbar\omega_j/(k_B T)}].$$

Differentiation at fixed T gives

$$\partial F_j / \partial \omega_j = U_j / \omega_j,$$

where U_j includes the zero-point energy and the mean thermal excitation. Since vibrational pressure is $-\partial F / \partial V$, the chain rule gives

$$p_{\mathrm{vib}} = V^{-1} \sum_j \gamma_j U_j.$$

A temperature derivative at fixed volume suppresses the constant zero-point contribution. The mechanical relation $(\partial p / \partial T)_V = \alpha_V B_T$ finally yields

$$\alpha_V = \frac{\sum_j \gamma_j C_{V,j}}{B_T V}.$$

Here B_T is the isothermal bulk modulus and α_V the volume expansion coefficient. The mode sum weights each Grüneisen parameter by its thermally active heat capacity. A low-frequency mode that softens under compression can contribute negatively and can dominate in a cold crystal while high-frequency modes remain frozen. A single positive average parameter is therefore a material model rather than a thermodynamic necessity.

Infer a mode average from dilatometry

For a cubic crystal take measured molar heat capacity $C_{V,m} = 12.0$ J/(mol K), molar volume $V_m = 1.20 \times 10^{-5}$ m³/mol and bulk modulus $B_T = 100$ GPa. A measured linear expansion coefficient 6.00×10^{-6} K⁻¹ corresponds to $\alpha_V \simeq 3\alpha_L = 1.80 \times 10^{-5}$ K⁻¹. The effective heat-capacity-weighted parameter is then

$$\bar{\gamma} = \alpha_V B_T V_m / C_{V,m} = 1.80.$$

The factor of three comes from differentiating the cube of a length, not from counting acoustic polarizations again.

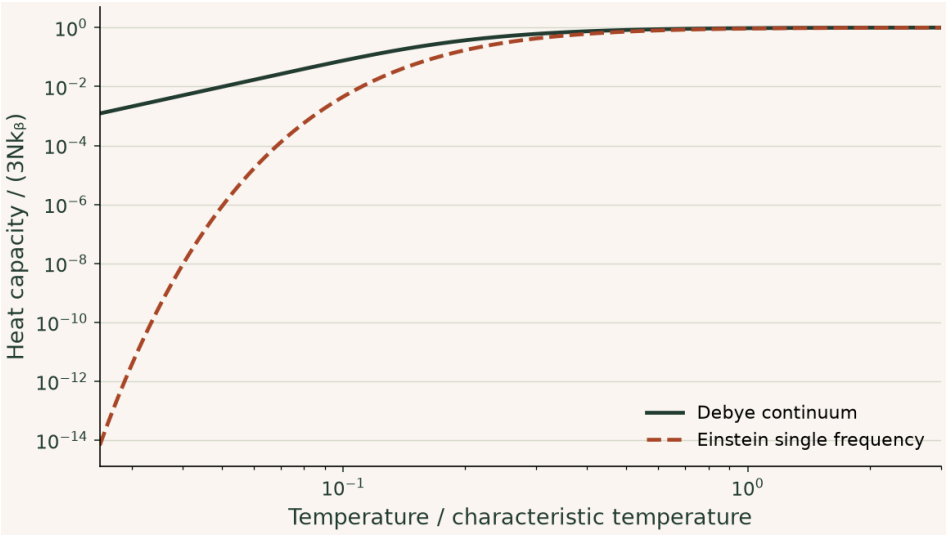
If the same positive average and bulk modulus remain approximately constant down to a regime where phonons give $C_V \propto T^3$, the volume expansion coefficient should also fall as T^3. Doubling temperature there increases the coefficient by about eight, although the accumulated length change requires integrating it over temperature. A comparison of α_V / C_V at several temperatures can test whether the active modes really share a common volume sensitivity. Molar heat capacity and molar volume must be taken on the same chemical counting basis before inferring any dimensionless Grüneisen parameter.

Hydrostatic pressure spectroscopy gives an independent way to measure a mode parameter. By the chain rule, the definition yields $\gamma_j = B_T (\partial \ln \omega_j / \partial p)_T$ within the quasiharmonic description. For the numerical bulk modulus, a mode with parameter 1.80 would increase its frequency by approximately 1.8 percent under a 1.00 GPa hydrostatic pressure increment. Agreement between this slope and the value inferred from expansion requires that the probed mode be representative of the thermally populated spectrum. A strong optical mode visible to a spectrometer need not dominate low-temperature heat capacity or expansion.

Quoted pressure derivatives must therefore specify the mode, the temperature and whether hydrostatic or directional stress was applied in the measurement.

Which lattice heat capacities survive at low temperature?

Compare a linear acoustic spectrum with a single oscillator frequency. Scale each temperature by its model’s characteristic temperature. The Debye prediction uses numerical quadrature of $3t^3 \int_0^{1/t} x^4 e^x / (e^x - 1)^2 \, dx$; both expressions approach unity at high temperature.



The single-frequency solid loses thermal excitations exponentially. The acoustic continuum still has low-frequency modes and its low-temperature curve follows a power law. Equal characteristic temperatures here set a comparison scale; fitting real data can give different values for the two models.

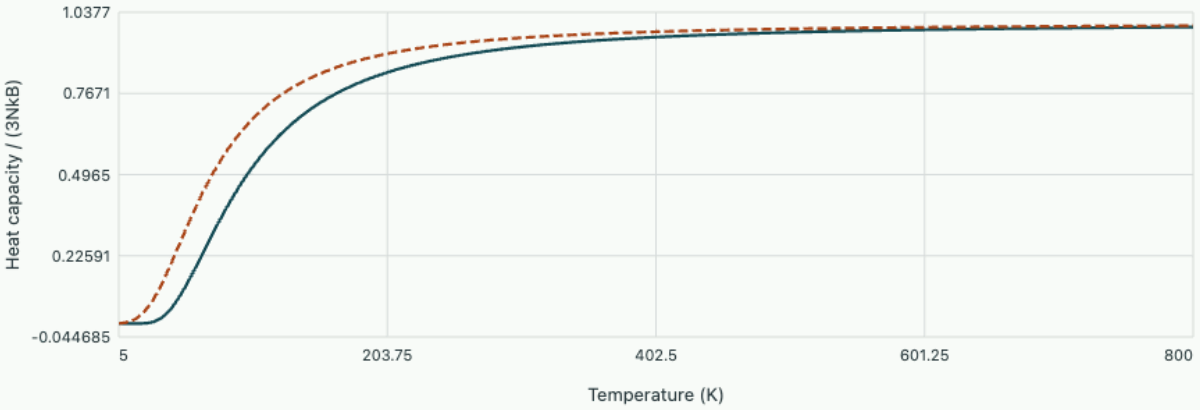
Try it · Einstein debye

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

Characteristic temperature (K) **300**



Einstein single-frequency model Debye acoustic continuum

Both approach the Dulong–Petit value one at high temperature. Debye retains low-frequency acoustic modes and gives a T^3 law at low temperature.

▶ **Read values from the plot**

Extracting a length from thermal spin fluctuations

Thermodynamics and statistical physics

A one-dimensional row of magnetic moments favours local alignment. At low temperature a snapshot can contain a long row of identical spins. An exact calculation distinguishes a large correlated domain from a thermodynamic phase with permanent order.

Sum all configurations by multiplying a small matrix

Consider N classical spins $s_j = \pm 1$ on a ring with $s_{N+1} = s_1$. At zero applied field the energy is $-J \sum_j s_j s_{j+1}$ with $J > 0$. The Boltzmann weight factors into nearest-neighbour terms $e^{K s_j s_{j+1}}$, with $K = J/(k_B T)$. Construct the matrix $T_{ss'} = e^{K s s'}$, whose diagonal elements are e^{K} and off-diagonal elements e^{-K}. Multiplying T's automatically sums each intermediate spin over its two choices. Closing the first and last indices gives

$$Z_N = \operatorname{tr}(T^N).$$

The symmetric and antisymmetric vectors $(1, 1)/\sqrt{2}$ and $(1, -1)/\sqrt{2}$ have eigenvalues $\lambda_+ = 2\cosh K$ and $\lambda_- = 2\sinh K$. Thus $Z_N = \lambda_+^N + \lambda_-^N$. At every finite positive temperature the ratio of eigenvalues is strictly below one. In the large-system limit the free energy per spin is $-k_B T \ln(2\cosh K)$. Differentiating with respect to β yields the energy per spin $-J \tanh K$.

Background · Connected correlation

$$C(r) = \langle s_i s_{i+r} \rangle - \langle s_i \rangle \langle s_{i+r} \rangle$$

Subtract the independent contribution before identifying a decay length. A periodic sample introduces a second propagation distance around the ring.

Insert two local measurements

Multiplying an individual configuration by a measured spin corresponds to inserting $S = \operatorname{diag}(1, -1)$ in the matrix trace. To measure two sites separated by r bonds use $\operatorname{tr}(S T^r S T^{N-r})/Z_N$. In the eigenbasis of T, S exchanges the symmetric and antisymmetric vectors. The numerator therefore contains one factor $\lambda_-^r \lambda_+^{N-r}$ and one with the eigenvalues exchanged. With $q = \lambda_-/\lambda_+ = \tanh K$,

$$\langle s_0 s_r \rangle_N = \frac{q^r + q^{N-r}}{1 + q^N},$$

$$\langle s_0 s_r \rangle_\infty = q^r = e^{-ra/\xi},$$

$$\xi = -\frac{a}{\ln q}.$$

Here a is the lattice spacing and the infinite-chain limit holds for fixed r. The second finite-ring term comes from the other way around the circle. At equal sites the full ratio equals one. In zero field spin

reversal makes the mean spin zero in the thermal distribution. Thus this correlation already equals the connected correlation.

Compare a chain at two temperatures

At $k_BT = J$, $q = 0.7616$ and $\xi = 3.67a$. Spins ten bonds apart have correlation only 0.0657. At $k_BT = 0.5J$, $\xi = 27.30a$ and the corresponding correlation is 0.693. Lower temperature suppresses individual domain walls. Since $s_j s_{j+1}$ is +1 for alignment and –1 for a wall, its average equals $1 - 2P_{\text{wall}} = q$. At low temperature $P_{\text{wall}} \simeq e^{-2K}$ and $\xi/a \simeq e^{2K}/2$. Even rare walls eventually destroy alignment in a sufficiently long chain.

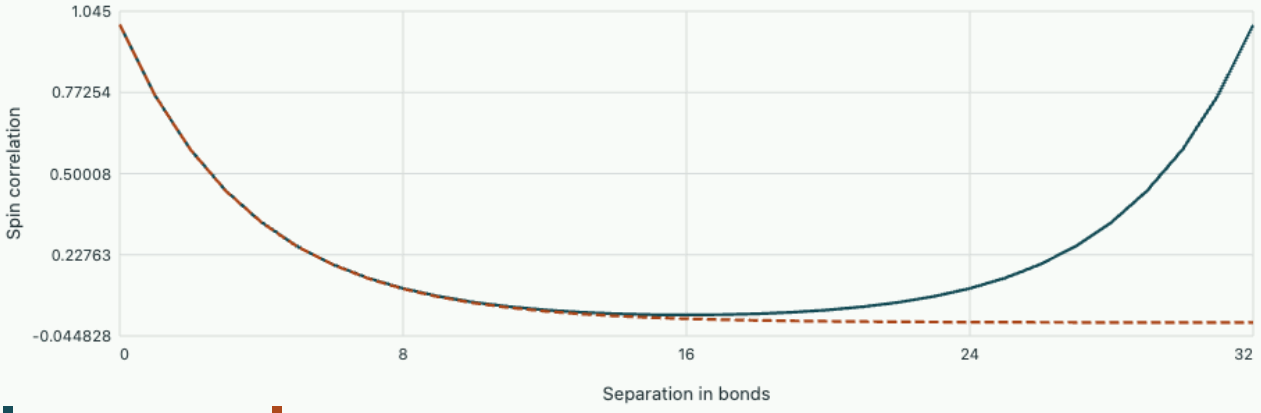
Compare the finite ring with the infinite-chain result. Change the ring length while leaving the temperature fixed. When the correlation length becomes comparable to the circumference, the long route contributes appreciably. Returning to the starting site must restore correlation one. A straight-line fit to the back half of a semilog finite-ring plot would consequently give a misleading decay length.

Open this experiment online

Try the calculation

Temperature $k_\beta T/J$ **1**

Sites around the ring **32**



Infinite-chain correlation length 3.672 lattice spacings. The opposite site on the 32-site ring has correlation 0.02562.

Tutorial on estimating an interaction

In a very long zero-field chain the nearest-neighbour correlation is 0.6. Determine $J/(k_BT)$, the probability of a wall on a given bond and the separation at which correlation first falls below 0.1.

▼ **Work through the solution**

Invert $q = \tanh K$ to obtain

$$K = \tfrac{1}{2}\ln[(1+q)/(1-q)] = 0.6931.$$

The wall probability is

$$(1-q)/2 = 0.20.$$

Solve $r > \ln(0.1)/\ln(0.6) = 4.51$. At five bonds the correlation is 0.07776. An even number of walls can restore the original spin direction, so this correlation counts more than the probability of finding a completely wall-free interval.

At $k_B T = J$ a nearest-neighbour bond is aligned with probability p. Find p and its relation to the mean energy per bond.

▼ Work through the solution

The correlation is $2p - 1 = \tanh 1 = 0.7616$, so $p = 0.8808$. Multiplying $2p - 1$ by $-J$ gives $-0.7616J$. This directly connects counts of local defects in snapshots with the thermodynamic energy obtained from the partition function.

Why can an odd number of walls never occur on a closed ring? Would dropping the finite-ring correction give the correct wall count at arbitrary temperature?

▼ Work through the solution

Each wall reverses the spin when following the ring. Returning to the starting site requires an even number of reversals. The exact nearest-neighbour correlation is $(q + q^{N-1})/(1 + q^N)$. Thus the mean wall count is $N(1 - C_1)/2$. Approximating C_1 by q works when N substantially exceeds the correlation length; otherwise ring closure changes the result.

Further reading · Cambridge statistical physics, Ising chain and correlation functions · Cambridge statistical mechanics, phase transitions and transfer matrices.

Why only a narrow shell of a Fermi gas carries heat

Thermodynamics and statistical physics

Heating a filled Fermi sea does not give every particle a classical thermal energy. Deep below the chemical potential, nearby states are already occupied. Quantifying the accessible states predicts both the low-temperature heat capacity and the shift of the chemical potential at fixed particle number.

Convert a sharp Fermi surface into a narrow weight

For independent fermions with occupancy $f(E) = 1/[1 + \exp((E-\mu)/(k_B T))]$, consider integrals $I = \int_0^\infty H(E) f(E)\, dE$. Choose $F(E) = \int_0^E H(s) ds$. Integration by parts gives $I = \int_0^\infty F(E)[-\partial_E f(E)] dE$, since the boundary terms vanish. The derivative is a peak of width $k_B T$ around μ. With $x = (E-\mu)/(k_B T)$ the weight becomes

$$w(x) = 1/[4\cosh^2(x/2)].$$

For μ many thermal units above the bottom and H smooth over the thermal shell, extend the lower limit to $-\infty$ when expanding locally around μ. The weight has unit integral and is even, so the linear term in $F(\mu + k_B T x)$ integrates to zero. Its second moment is $\pi^2/3$, obtained by expanding $w(x) = \sum_{\ell \geq 1} (-1)^{\ell+1} \ell e^{-\ell x}$ on $x > 0$ and integrating x^2 termwise. Since $F'' = H'$,

$$I = \int_0^\mu H(E) dE + \frac{\pi^2}{6} (k_B T)^2 H'(\mu) + O(T^4).$$

The first correction samples the slope of the smooth weighting function. If a band edge or singularity lies within the shell, replacing the lower limit and Taylor expanding are poor approximations.

Conserve the number of particles

For a three-dimensional parabolic dispersion the density of states is $D(E) = A\sqrt{E}$. At $T = 0$, $N = (2A/3) E_F^{3/2}$. At small T, use the expansion with $H = D$ and write the chemical potential as $E_F + \delta\mu$, where $\delta\mu = O(T^2)$. Subtract the zero-temperature count. To quadratic order $0 = D(E_F)\delta\mu + (\pi^2/6)(k_B T)^2 D'(E_F)$. Because $D'/D = 1/(2E_F)$,

$$\mu/E_F = 1 - \frac{\pi^2}{12} (T/T_F)^2 + O[(T/T_F)^4],$$

$$T_F = E_F/k_B.$$

The increasing density of states offers more available levels just above the Fermi energy than just below it. The chemical potential shifts down to preserve the count. At $T/T_F = 0.10$ the predicted shift is $-0.00822 E_F$. For an ideal gas with $E_F = 5.0\,\mathrm{eV}$, this particular reduced temperature corresponds to about 5800 K, so reduced temperature and ordinary laboratory temperature should be distinguished.

Differentiate the energy at fixed number

Use $H = ED$ in the same integral, including the just-derived shift of μ. The boundary shift contributes $E_F D(E_F)\delta\mu$. In the Sommerfeld term the derivative is $D(E_F) + E_F D'(E_F)$. The terms involving the slope of D cancel after fixing number. Thus $U - U_0 = (\pi^2/6)D(E_F)(k_B T)^2$, and $C_V = (\pi^2/3)D(E_F)k_B^2 T$. In three dimensions $D(E_F) = 3N/(2E_F)$, giving $C_V/(Nk_B) = (\pi^2/2)T/T_F$. This is an electronic contribution for independent particles. Other degrees of freedom such as phonons can dominate a measured material heat capacity.

The experiment reconstructs μ by numerical integration and bisection at fixed number. It overlays the asymptotic prediction, allowing the fractional error to be checked before using it at larger T/T_F. The quadrature uses the momentum-like substitution $E/E_F = y^2$ to remove the square-root endpoint singularity.

Open this experiment online

Try the calculation

Reduced temperature T/TF **0.1**



Fixed-number μ/EF = 0.9916, quadratic prediction 0.9918. Leading heat capacity $C_v/(Nk_B)$ = 0.4935.

Tutorial on interpreting heat-capacity data

For a gas with $E_F = 4.0\,\mathrm{eV}$ at $T = 300\,\mathrm{K}$, estimate $C_V/(Nk_B)$ and the chemical-potential change in eV. Use $k_B T = 0.02585\,\mathrm{eV}$.

▼ **Work through the solution**

The reduced temperature is 0.00646. The heat capacity ratio is approximately 0.0319, and $\delta\mu = -(\pi^2/12)(0.02585)^2/4.0 = -1.37 \times 10^{-4}\,\mathrm{eV}$.

A two-dimensional parabolic band has constant density of states above its edge. What does the quadratic expansion predict for the fixed-number chemical potential and energy?

▼ **Work through the solution**

$D' = 0$ makes the quadratic shift of μ vanish as long as the edge is distant. The energy integral still uses $(ED)' = D$, so it has a positive T^2 correction. Effects of the distant band edge can give exponentially small departures in μ.

Further reading · Cambridge statistical physics, degenerate Fermi gas and Sommerfeld expansion.

Mean-field ordering, response and the critical point

Thermodynamics and statistical physics

Interactions can make a magnet favour a nonzero moment even when the applied field vanishes. Averaging the local exchange field gives a solvable model whose critical predictions can be compared with the exact one-dimensional chain.

Build the free energy from the count of spin states

Take N spins $s_i = \pm 1$, nearest-neighbour energy $-J\sum_{\langle ij\rangle} s_i s_j$ and Zeeman energy $-h\sum_i s_i$, where h has energy units. Denote mean spin by m. The fractions of positive and negative spins are $p_\pm = (1 \pm m)/2$. The binomial count and Stirling's approximation give entropy per spin $-k_B(p_+ \ln p_+ + p_- \ln p_-)$. Independent-neighbour mean field replaces each pair average by m^2. With z neighbours and each bond counted once, interaction energy per spin becomes $-zJm^2/2$.

$$f(m) = -\tfrac{1}{2} zJm^2 - hm + k_B T[p_+ \ln p_+ + p_- \ln p_-].$$

Differentiate at fixed T, h. The entropy contribution has derivative $k_B T \operatorname{atanh} m$, giving $m = \tanh[(zJm + h)/(k_B T)]$. This equation can have three real roots, so an iterative solver returning a root alone is insufficient. The equilibrium phase minimises f; the second derivative $k_B T/(1 - m^2) - zJ$ distinguishes local minima and maxima.

Expand about the loss of stability

Near zero magnetisation,

$$\operatorname{atanh} m = m + m^3/3 + \cdots.$$

Set $T_c = zJ/k_B$ and integrate the expanded derivative to obtain

$$f(m) - f(0) \simeq \tfrac{k_B}{2}(T - T_c)m^2 + \tfrac{k_B T}{12} m^4 - hm.$$

At zero field and $T > T_c$, $m = 0$ is stable. Below T_c, symmetry gives two minima whose leading magnitudes satisfy $m^2 \simeq 3(T_c - T)/T$. Thus the order parameter vanishes like $(T_c - T)^{1/2}$ close to T_c. At $T = T_c$ a small nonzero field instead gives

$$m \simeq [3h/(k_B T_c)]^{1/3}.$$

These powers describe this model's limiting behaviour. For a finite sample symmetry makes the unbiased ensemble average zero if it samples both minima. Spontaneous magnetisation refers to the thermodynamic limit with an infinitesimal field selecting a sign.

Compute a response, then compare with an exact counterexample

Above T_c the small-field stationary equation gives

$$\chi = \partial m/\partial h = 1/[k_B(T - T_c)].$$

Below T_c, differentiate the exact stationarity equation on the positive branch, obtaining $\chi = [k_B T/(1 - m^2) - zJ]^{-1}$. Expanding near T_c there gives $\chi \simeq [2k_B(T_c - T)]^{-1}$, half the above-transition amplitude at equal distance. At $T = 0.8T_c$, the full zero-field solution is about $m = 0.7104$; the small-m estimate $\sqrt{3(1 - 0.8)/0.8} = 0.866$ is already visibly inaccurate.

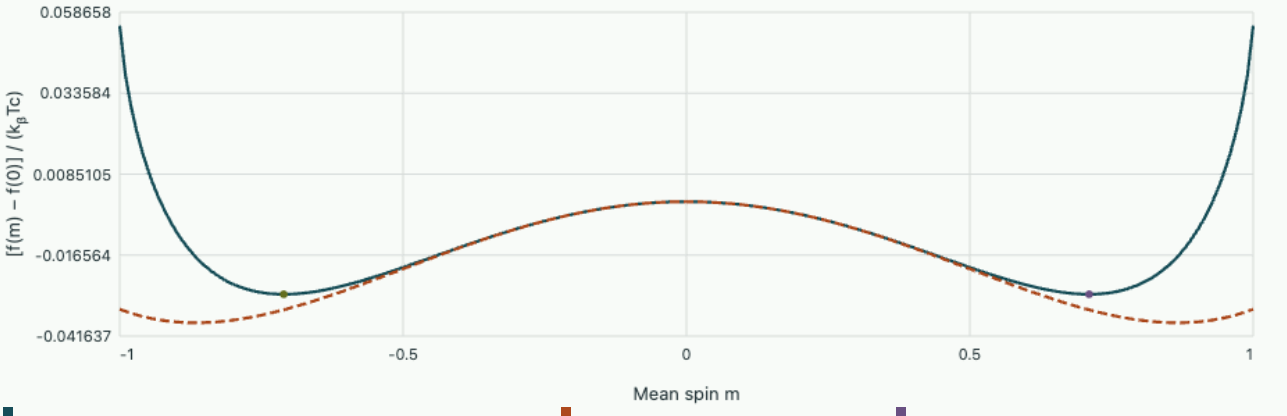
Mean field predicts an ordering temperature even for a one-dimensional chain with $z = 2$. The exact finite-temperature correlation length in the earlier transfer-matrix solution is finite at all $T > 0$. Spatial fluctuations and domain walls invalidate the independent-neighbour closure in one dimension. The experiment plots the full mean-field free energy and its quartic approximation for the same normalised T, h. Find the minima before inspecting the calculated order parameter. Compare how a weak bias changes their relative depths.

Open this experiment online

Try the calculation

Temperature T/Tc **0.8**

Field h/(k$_\beta$Tc) **0**



Full independent-neighbour free energy · Quartic approximation · Equilibrium minimum · Symmetry-related minimum

Equilibrium mean spin ±0.7104. Curvature at the minimum 0.6151 in units of k$_\beta$Tc.

Tutorial on critical response

At $T = 1.10T_c$, what is the linear-response mean spin expected at $h/(k_B T_c) = 0.002$? At the critical point use the cubic equation to estimate m for the same field.

▼ **Work through the solution**

Above the transition $m \simeq 0.002/0.10 = 0.020$. At criticality $m \simeq (0.006)^{1/3} = 0.182$. The larger response signals the loss of the quadratic restoring term; linear extrapolation through T_c would give a divergent susceptibility.

At $h = 0$ and $T = 0.95T_c$, is $m = 0$ a possible output from a fixed-point solver? How should an equilibrium calculation prevent a wrong interpretation?

▼ **Work through the solution**

Starting the symmetric iteration at exactly zero returns zero repeatedly, even below T_c. Its free-energy curvature is $k_B(T - T_c) < 0$; it is a local maximum. Compare the free energies or bracket the two symmetry-related nonzero stationary points before reporting an equilibrium moment.

Further reading · Cambridge statistical physics, mean-field Ising and Landau theory.

Measure equilibrium field fluctuations with a capacitor

Fields and the response of matter

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 $S_C(f)|H_{\mathrm{det}}(f)|^2$.

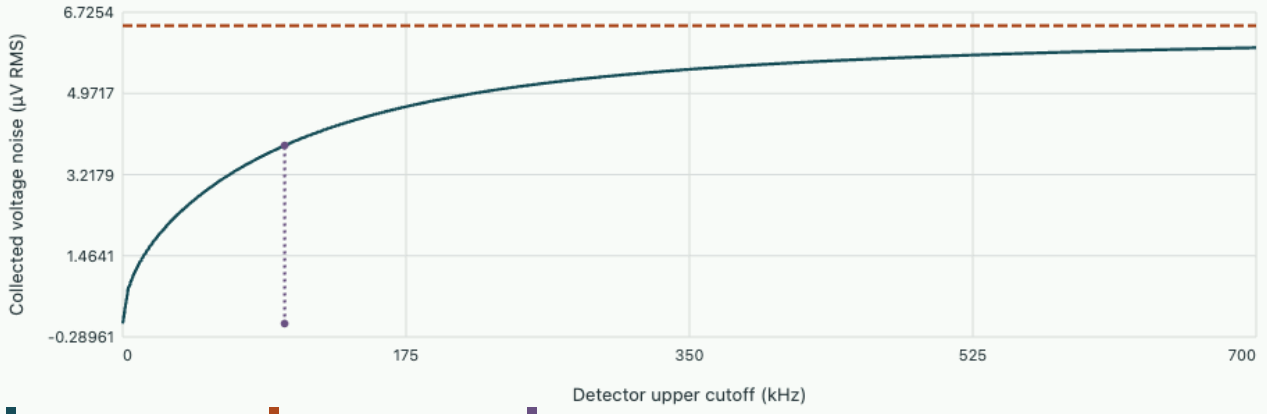
Open the live experiment

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.

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.

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

$$\sigma(V_s) = \frac{2\epsilon\kappa k_B T}{e}\sinh\left(\frac{eV_s}{2k_B T}\right),$$

$$\frac{C_{\mathrm{diff}}}{A} = \epsilon\kappa\cosh\left(\frac{eV_s}{2k_B T}\right).$$

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,

$$V_{\mathrm{app}} = V_d + \sigma(V_d)/c_s.$$

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 $\sigma_{\mathrm{ion}} = 2e^2 n_b D_i/(k_B T)$. 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 $R = L/(\sigma_{\mathrm{ion}} A)$ 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

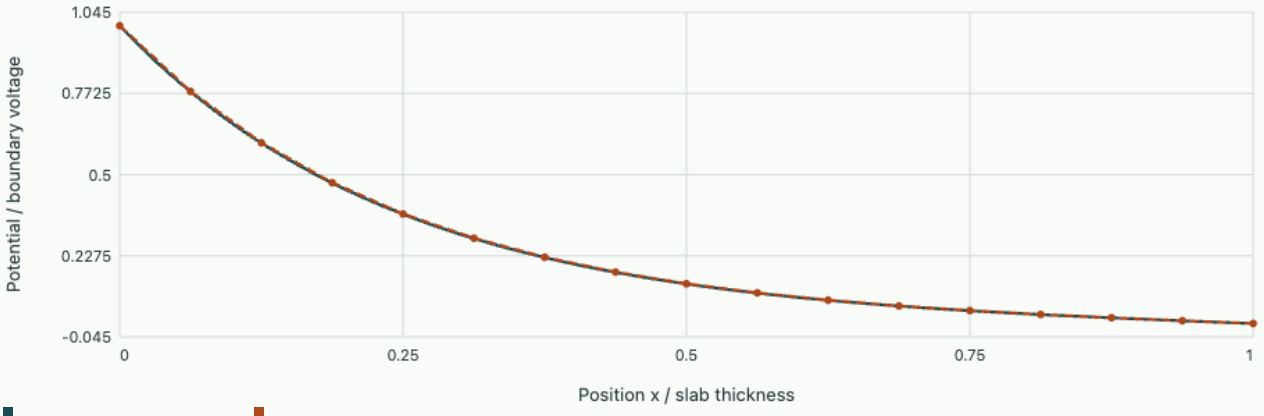
Compare the linear screened profile with a grid solution as the ratio of thickness to screening length grows. Separate failure to resolve the layer from failure of the linear material law.

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

Rotational and vibrational populations in a molecular gas

Statistical mechanics from microscopic counting to material response

A pressure thermometer and a calorimeter do not count molecular degrees of freedom in the same way. A dilute diatomic gas can retain nearly ideal pressure while its heat capacity changes appreciably as discrete internal levels become populated. Build the molecular partition function, derive the rotor and oscillator contributions, and infer temperatures from line populations without treating spectral intensity as a direct state count.

When a molecule's partition function factorises

Assume a dilute gas of stable diatomic molecules with weak interparticle interactions. Approximate one molecule's Hamiltonian by independent translational, rotational, vibrational and electronic parts. This requires a bond sufficiently stiff that rotation does not greatly alter vibrational energy and a temperature well below electronic excitations or dissociation when the electronic factor is treated as constant. If energies add and allowed states form a product, Boltzmann factors multiply. Thus $q = q_{\mathrm{tr}} q_{\mathrm{rot}} q_{\mathrm{vib}} q_{\mathrm{el}}$, $Z_N \simeq q^N / N!$ in the dilute classical translational regime. The logarithm converts this product into additive free energies, energies and heat capacities.

The translational factor V/λ_T^3 provides the volume derivative and hence ideal pressure $N k_B T / V$. A volume-independent internal factor changes neither that derivative nor the associated translational equation of state. It does change entropy, chemical potential and calorimetry. The molar heat capacities satisfy $C_p - C_V = R$ for a fixed-composition ideal gas, even when both vary strongly with temperature. Dissociation, chemical equilibration or significant intermolecular interactions can change the conclusion for the measured sample.

One danger is double-counting states. Angular momentum orientations belong to rotational degeneracy, oscillator quanta to vibration, electronic spin and nuclear exchange symmetry may constrain one another. A constant electronic degeneracy contributes an additive entropy $N k_B \log g$ but no internal energy or heat capacity at fixed level energies. A large heat capacity is evidence for energetically separated states becoming populated, not just for a large number of exactly degenerate states.

Derive the quantum rigid-rotor contribution

For a molecule of moment of inertia I about its centre of mass, rotational Hamiltonian is $\hat{L}^2/(2I)$. Angular momentum eigenvalues are $\hbar^2 J(J+1)$, with $2J+1$ projections at fixed $J = 0, 1, \ldots$. Define rotational temperature $\theta_r = \hbar^2/(2Ik_B)$. For distinguishable nuclei and a common nuclear spin factor omitted, the rotational sum and populations are

$$q_r = \sum_{J=0}^{\infty} (2J+1) e^{-\theta_r J(J+1)/T},$$

$$p_J = (2J+1) e^{-\theta_r J(J+1)/T} / q_r.$$

At high temperature the populated range contains many adjacent J values. Replacing the sum with an integral and substituting $u = J(J+1)$ gives

$$q_r \simeq T/\theta_r.$$

Hence $U_r \simeq Nk_BT$, $C_{V,r} \simeq Nk_B$, the classical result for two independent rotational kinetic energies of a linear object. At low temperature only $J = 0$ and $J = 1$ matter. Their degeneracies and gap $2k_B\theta_r$ give $q_r \simeq 1 + 3e^{-2\theta_r/T}$ and an exponentially vanishing heat capacity. Keeping the degeneracy three is essential for both its magnitude and the temperature of its turn-on.

For an illustrative heteronuclear rotor with $\theta_r = 85$ K, direct summation at 20 K gives $q_r = 1.00061$ and $C_{V,r} = 0.0440k_B$ per molecule. At 90 K the corresponding values are 1.471 and 1.063. At 300 K the heat capacity is 1.002. A small overshoot above the classical rotational value is possible during the crossover. A model that simply switches from zero to one classical degree contribution at $T = \theta_r$ misses the entire smooth population redistribution and cannot predict line ratios.

The most populated J need not be zero, despite increasing energies, because the projection multiplicity increases with J. Viewing the numerator as a smooth function of J gives $J_{\mathrm{peak}} \simeq [\sqrt{2T/\theta_r} - 1]/2$, to be checked against neighbouring integer values. The estimate is unreliable if it predicts only one or two populated levels. Computing a short discrete sum costs little and reveals its own truncation error through the final omitted weights.

Derive oscillator energy and heat capacity from a geometric series

Near its bond-length minimum approximate the interaction by a spring. With reduced mass m_r and angular vibration frequency ω_v, quantum energies are $E_n = \hbar\omega_v(n + 1/2)$. Define $\theta_v = \hbar\omega_v/k_B$. Summing the resulting geometric series and differentiating gives

$$q_v = \frac{e^{-\theta_v/(2T)}}{1 - e^{-\theta_v/T}},$$

$$\frac{U_v}{N} = k_B\theta_v\left(\frac{1}{2} + \frac{1}{e^{\theta_v/T} - 1}\right),$$

$$\frac{C_{V,v}}{Nk_B} = \frac{x^2e^{-x}}{(1 - e^{-x})^2},$$

$$x = \theta_v/T.$$

The zero-point term contributes energy but no heat capacity at fixed frequency. For $x \ll 1$ the heat capacity approaches k_B per independent vibrational mode, comprising kinetic and potential contributions. For $x \gg 1$ excitation is too costly and the heat capacity behaves as $k_Bx^2e^{-x}$. A molecule can still vibrate quantum mechanically in its ground state while contributing virtually no vibrational heat capacity. A calorimeter measures sensitivity of mean energy to temperature and need not see the ground-state motion.

Take $\theta_v = 3000$ K. At $T = 300$ K the vibrational contribution is only $0.00454k_B$, while at 1200 K it reaches $0.609k_B$. Combining with the earlier rotor gives molar C_V approximately $2.507R$ at 300 K and $3.109R$ at 1200 K, assuming stable composition and frozen electronic structure. For acoustic measurements the ideal adiabatic ratio $\gamma = (C_V + R)/C_V$ then decreases from about 1.399 to 1.322. An

acoustic drive faster than vibrational relaxation can measure a frequency-dependent effective ratio because internal populations fail to remain equilibrated through a compression cycle.

Compare this single oscillator with the frequency distribution in the lattice heat-capacity plot

Nuclear exchange and the symmetry number

For a homonuclear molecule, rotating its classical orientation by half a turn end for end does not create a second distinct arrangement. In a high-temperature rotational count a symmetry number two therefore divides the distinguishable-nuclei rotor partition function once the compatible spin states are accounted for. Since this factor is constant with temperature it shifts entropy and chemical potential, but not the classical rotational energy or heat capacity. At low temperature it is important to treat actual quantum parity, not divide every level by two.

A homonuclear diatomic rotational wavefunction acquires a factor $(-1)^J$ under exchange of nuclei. The complete state also includes nuclear spin, electronic and vibrational parts. For ordinary molecular hydrogen in its electronic and vibrational ground state the protons are spin-half fermions. Antisymmetry pairs even J with the single antisymmetric spin singlet and odd J with the three symmetric triplet spin states. In true spin-conversion equilibrium the rotor sum uses weights one or three in addition to $2J+1$. At temperatures large compared with rotational spacing, even and odd J sectors have nearly equal rotational sums and the equilibrium odd-to-even population ratio approaches three.

Conversion between spin species can be slow. A rapidly cooled sample may preserve its prepared fractions rather than attain spin equilibrium. Then compute separate parity-restricted partition functions for fixed species numbers, and average their thermal observables with the frozen proportions. Allowing the total odd fraction to relax yields a different heat capacity, including the energy liberated by conversion. Two experiments on chemically identical hydrogen can consequently report different heating curves because their preparation constrains different internal sectors.

Use intensities only after accounting for the instrument

An absorption line originating from level J has a strength depending on more than its population. Transition moment, polarisation, transition frequency, stimulated emission correction, instrument response and path length can all enter. Suppose line strengths and detector factors have been divided out to yield quantities proportional to level populations N_J. Then a straight-line fit of $\log[N_J/(2J+1)]$ against $J(J+1)$ has slope $-\theta_r/T$ for the simple heteronuclear model. Its intercept depends on total particle number, partition function and residual instrumental gain.

For $\theta_r = 10$ K, a measured corrected ratio $N_2/N_0 = 2.00$ implies $2 = 5e^{-60/T}$ and $T = 65.5$ K. Omitting the factor of five would make an upper population larger than the lower one appear incompatible with positive temperature. The inferred temperature assumes the rotational populations are equilibrated at one temperature; an optically pumped or rapidly expanding sample may have different rotational and translational distributions.

A plot over several J values can test whether one slope is adequate. Curvature may come from centrifugal distortion, line saturation, unresolved overlap, multiple temperatures or a wrong nuclear-spin

degeneracy. Alternating residuals for even and odd J indicate a different physical problem from a smooth quadratic residual against $J(J+1)$. Repeat with lower optical depth to isolate saturation, and use independently measured transition frequencies to fix level energies before fitting populations.

Recover molecular translation separately from internal structure

In a low-density molecular beam, the translational speed distribution is often easier to measure than the internal populations. For an equilibrium gas leaving a small opening into vacuum, the molecular flux is weighted by the normal velocity. Fast molecules reach and cross the aperture more frequently, so the distribution of speeds in the beam differs from the Maxwell speed distribution inside the reservoir. The mean translational energy of molecules in an effusive beam is $2k_B T$, whereas the bulk mean is $3k_B T/2$. This selectivity occurs even when rotational and vibrational populations remain canonical.

To derive the difference, start with the bulk velocity density proportional to $e^{-mv^2/(2k_B T)}$. Write the outward normal speed as $v\cos\theta$, integrate over the outward hemisphere and include the velocity-space shell factor $v^2 dv$. The resulting beam speed density is proportional to $v^3 e^{-mv^2/(2k_B T)} dv$, while the original bulk speed density has v^2. Taking a ratio of Gaussian moments gives the stated energy. Comparing a beam energy to the wrong bulk prefactor would yield a temperature one third too high.

If the aperture is small compared with the mean free path and transition-detection efficiency is independent of speed after correction, internal state probabilities factor out of this flux weighting. A rotational population ratio can then report the reservoir temperature while the observed mean translational energy follows the effusive weighting. In a dense nozzle undergoing many collisions, expansion can cool translation and rotation differently. An inferred rotational and translational temperature disagreement is therefore not automatically bad calibration; first identify whether the beam is effusive, collisionally expanding or selectively detected.

As a bench example, a beam mean translational energy of 0.060 eV corresponds to approximately 348 K under effusion, using $k_B = 8.6173 \times 10^{-5}$ eV/K. Treating it as the bulk mean would give about 464 K. Compare against a population ratio from two corrected lines spanning $80k_B$ kelvin and equal degeneracies. At the two temperatures the predicted ratios are about 0.795 and 0.842. Resolving this difference demands a relative line calibration appreciably better than six percent even though the two erroneous temperature interpretations differ by more than one hundred kelvin.

An isotope mixture reveals which spectral changes affect entropy

Compare two isotopic forms with the same approximate electronic potential surface. Bond force constant and equilibrium bond length then remain nearly unchanged while reduced nuclear mass changes. The harmonic vibrational frequency scales as $m_r^{-1/2}$, while the rotational temperature scales as m_r^{-1}. Consequently vibration and rotation respond differently to the same isotope substitution. Their shifts can jointly test whether two sets of spectral features genuinely belong to the same molecular carrier.

Suppose the reduced mass rises by a factor 1.21. A vibrational line at 2000 inverse centimetres moves approximately to 1818 inverse centimetres, and a rotational constant initially 2.00 inverse centimetres moves to 1.65 inverse centimetres. The vibrational characteristic temperature follows from $hc\tilde{\nu}/k_B$,

about 1.4388 K per inverse centimetre. At room temperature the two vibrational gaps still greatly exceed thermal energy, so the vibrational heat capacity can remain small even though the spectral isotope shift is easy to resolve.

The zero-point energies nevertheless differ. For this vibrational example their difference is $hc(2000 - 1818)/2$ per molecule, corresponding to about 131 kelvin in thermal units or 1.09 kJ/mol. Such offsets cancel within a fixed-species heat-capacity derivative but can matter to comparisons of chemical equilibrium or isotopic exchange. An analysis that discards all constant energy offsets independently in each molecular species can corrupt a reaction free energy unless it restores a consistent reference across reactants and products.

For a thermal entropy contribution from this one oscillator use $S_v = k_B[(n+1)\log(n+1) - n\log n]$ with $n = 1/(e^x - 1)$. This follows by inserting U_v and $\log q_v$ into $S = k_B \log q + U/T$; the zero-point contributions cancel exactly. Thus the measured isotope shift can change vibrational entropy via occupancy even while the zero-point offset itself contributes no entropy. It is helpful to evaluate all three quantities—energy difference, entropy difference and heat-capacity difference—rather than presume that a line shift must have the same thermodynamic role in each.

An internal relaxation time appears in frequency-dependent calorimetry

The equilibrium partition function assumes internal mode populations re-equilibrate on times short compared with the measurement. Model one slower internal degree by an internal energy U_s relaxing toward its equilibrium value at the instantaneous translational temperature T, with $\dot{U}_s = [U_{s,\mathrm{eq}}(T) - U_s]/\tau_s$. For a small oscillation around T_0,

$$\delta U_{s,\mathrm{eq}} = C_s \delta T.$$

In the $e^{-i\omega t}$ convention, the complex internal heat capacity and the total measured capacity including an instantaneous background C_f are

$$C^*(\omega) = C_f + \frac{C_s}{1 - i\omega\tau_s}.$$

At low frequency the capacity is $C_f + C_s$. At high frequency the slow internal energy cannot follow and the apparent in-phase capacity approaches C_f. Its out-of-phase component peaks at $\omega\tau_s = 1$, where real and imaginary increments equal $C_s/2$. The structure resembles dielectric relaxation because both arise from one relaxing state variable driven by a controlled sinusoid. It follows here from an explicit energy-exchange model and its assumptions.

Take a molar fast contribution $2.50R$, slow equilibrium contribution $0.60R$ and relaxation time 2 ms. The full equilibrium value is $3.10R$. At about 79.6 Hz the real measured capacity is $2.80R$ and the imaginary increment magnitude $0.30R$. A heat pulse much shorter than two milliseconds initially raises the temperature as if the capacity were close to $2.50R$, then energy flows into the internal mode. A slow calorimetric ramp can be perfectly consistent with a faster acoustic measurement that infers a smaller active capacity.

A sample container connected to a thermal bath produces its own complex temperature response even if molecular populations are always equilibrated. Disentangle those effects by calibrating the empty cell or an inert reference gas across the same frequency range, changing pressure to alter collision-mediated internal-relaxation time, and checking the predicted equilibrium limit from independently known level spacings. A mechanical resonance in the apparatus can be separated by varying cell geometry while keeping gas composition fixed. Fitting one imaginary peak to a molecular relaxation model without these controls does not uniquely determine the microscopic relaxation time.

How a line ratio and a calorimeter constrain degeneracies differently

Two groups report a low-energy excited manifold with dimensionless population ratio $r = N_1/N_0$ at two temperatures. A single ratio determines only $r = (g_1/g_0)e^{-\Delta/(k_B T)}$, so changing the assigned degeneracy changes the inferred gap. Taking the log ratio at two known temperatures removes a common degeneracy factor. The slope of $\log r$ against $1/T$ is minus the gap in kelvin, and the intercept is the log degeneracy ratio if the line-strength calibration and equilibrium assumptions hold.

For illustrative values $r = 0.300$ at 100 K and $r = 0.949$ at 200 K, the log change is approximately 1.152. Dividing by the inverse-temperature change magnitude 0.005 per kelvin gives $\Delta/k_B \simeq 230$ K. Inserting the first ratio gives $g_1/g_0 \simeq 3.00$. The same two-manifold model predicts a dimensionless heat capacity $x^2 r/(1+r)^2$ per molecule. At 100 K it is about 0.94, a separate calorimetric prediction independent of the absolute number of detected photons.

A third temperature materially improves the test. An unseen nearby manifold or a line strength that changes with instrumental settings may fit two exact population ratios but produce curvature and excess heat capacity elsewhere. Spectroscopy resolves ordered energy differences while calorimetry sums energy variances over all accessible states. Their combination offers more information than adding further precision to one population measurement under the same ambiguous degeneracy assumption.

Compress a gas without pretending its heat capacity stays constant

A reversible adiabatic compression is often summarised by a fixed exponent γ. That power law assumes temperature-independent heat capacities along the path. Molecular populations give a direct case where the exponent changes. For an ideal gas at fixed composition,

$$dS = C_V(T)\,dT/T + N k_B\,dV/V.$$

Along a reversible adiabatic path set the sum to zero and integrate. The volume ratio is determined by an integral of measured or calculated $C_V(T)/T$, not by a universal constant exponent across every temperature.

As an illustrative calculation approximate translation and fully active rotation by $C_V = 2.5 N k_B$ plus one harmonic vibrational mode with $\theta_v = 3000$ K. Heating from 300 to 1200 K at constant entropy gives a

volume contraction factor from the fast degrees alone of $4^{2.5} = 32$. The vibrational entropy difference adds another factor $e^{\Delta S_v/(Nk_B)}$. At 300 K the oscillator entropy is only about $0.00050 k_B$, while at 1200 K it is about $0.309 k_B$. The combined estimate requires $V_i/V_f \simeq 43.6$, a substantially larger compression than the frozen-vibration approximation predicts.

This comparison assumes ideal gas throughout and no dissociation, conditions to be checked in a real large compression. It nevertheless identifies the mechanism: part of the compression work goes into levels that were effectively inactive initially. Using the final high-temperature heat capacity across the whole path would err in the opposite direction by assigning that energy storage to the low-temperature portion as well. The entropy integral correctly weights each temperature interval by its actual internal population response.

A numerically stable solution can avoid differentiating noisy calorimetry. Integrate the entropy increment from tabulated heat capacities in the log-temperature variable. Compare an adaptive integral with a refined straightforward grid, paying particular attention to the rotational or vibrational crossover scale. Alternatively compute endpoint entropies from partition functions when an independently justified spectral model exists. Agreement between calorimetric integration and spectroscopy-derived endpoint entropy checks both assignments without demanding exact agreement of noisy pointwise derivatives.

Experimental pressure also checks the endpoint calculation independently. After equilibration, the ideal equation of state predicts $p_f/p_i = (T_f/T_i)(V_i/V_f)$, approximately 174 for this idealised path. Finite heat leakage or a pressure-dependent nonideality can change that value. A rapid adiabatic compression can be irreversible even if little heat leaves the chamber, and then the measured entropy can increase. Distinguish a reversible reference calculation from an actual rapid piston protocol before using a mismatch to infer missing molecular levels.

An independent isotropic acoustic test at a fixed initial temperature infers γ from the small-amplitude sound speed $c_s^2 = \gamma RT/M$, where M is molar mass. At fixed molecular mass a decrease of γ from 1.40 to 1.32 at the same temperature would lower sound speed by the square root of their ratio, roughly three percent. That comparison must truly hold temperature fixed: the large direct thermal dependence of sound speed can easily exceed the internal-mode effect. Frequency-resolved acoustics can see part of a vibrational mode freeze dynamically while equilibrium spectroscopy still predicts its population. Record both phase velocity and acoustic attenuation if the aim is to identify a relaxation time. Across the relaxation band, dispersion without the corresponding dissipative response should prompt a check of transducer phase calibration and acoustic boundary conditions.

Read the rotational level spacing from thermal energy storage

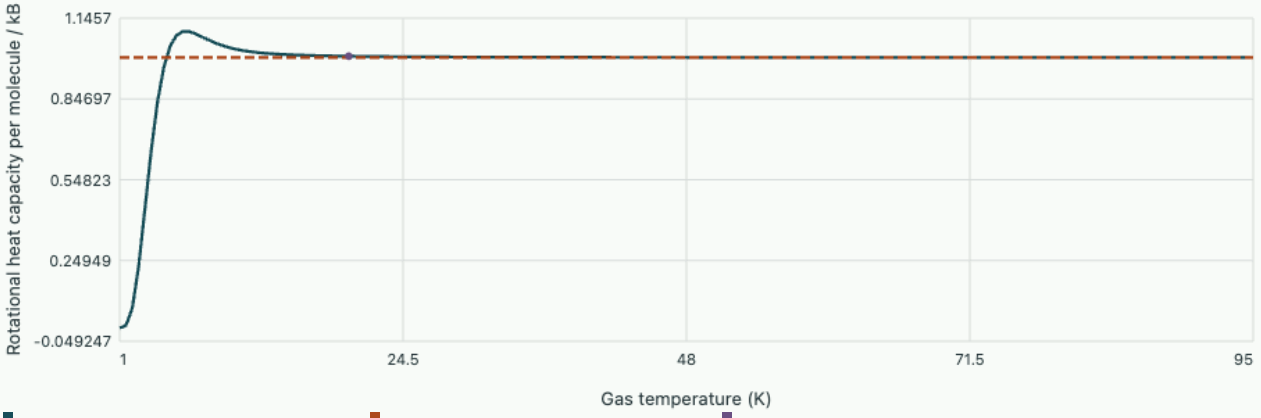
Sum the discrete levels of a linear rotor and compare the heat capacity with equipartition as temperature rises.

Open this experiment online

Try the calculation

Rotational temperature θrot (K) 8

Chosen gas temperature (K) 20



Heteronuclear rigid rotor | Two quadratic rotations | Selected temperature

First excited rotational level has E_1/k_B = 16 K. At the selected temperature C_{rot}/k_B = 1.0051. Vibrational and translational contributions are excluded.

Tutorial · Identify a rotor inertia from spectral spacing

For a heteronuclear linear rotor the lowest pure rotational transition has ordinary frequency 120 GHz. Infer θ_r, the moment of inertia and the expected $J = 1$ to $J = 0$ population ratio at 10 K. How could an isotope substitution test the interpretation?

▼ **Use the transition gap and keep angular and ordinary frequency distinct**

The gap is $2k_B\theta_r = h\nu$, so $\theta_r = 2.88$ K and $I = \hbar^2/(2k_B\theta_r) = 1.40 \times 10^{-46}$ kg square metres. The population ratio is

$$3e^{-5.76/10} = 1.69.$$

The upper manifold contains more molecules even though each of its levels has lower individual Boltzmann weight. Increasing reduced mass at nearly fixed bond length increases I, reducing the transition frequency approximately inversely with reduced mass. A vibrational frequency from the same spring would scale instead with inverse square root of reduced mass.

Tutorial · Infer an excitation temperature from calorimetry

A gas of linear molecules has molar $C_V = 2.91R$ at a temperature where translation and rotation contribute $2.50R$. Neglect electronic excitation and estimate the characteristic temperature of a single

harmonic vibration in units of measurement temperature. Would a measured C_p be usable unchanged?

▼ **Invert the oscillator contribution and verify its feasible range**

The remaining $0.41R$ lies between zero and the one-mode classical limit R. Solve $x^2e^{-x}/(1-e^{-x})^2 = 0.41$ on positive x. At 3.0 it is 0.496 and at 3.5 it is 0.393, placing the root near 3.41. Thus θ_v is about 3.41 times measurement temperature. Use several temperatures to verify the characteristic value and frozen-composition model. For ideal gas at fixed composition, subtract R from C_p before applying this constant-volume balance. Latent processes, reaction and internal relaxation would require an extended model.

Tutorial · Distinguish equilibrium hydrogen from a frozen spin mixture

At low temperature only $J = 0$ in the even sector and $J = 1$ in the odd sector contribute. Use $\theta_r = 85$ K and $T = 30$ K. Estimate the equilibrium odd fraction and compare the energy stored in a sample frozen at a three-quarter odd fraction, measured relative to the common lowest even state.

▼ **Normalise both sectors and identify the energy released on conversion**

The relative weight of $J = 1$ is $3(2J+1)e^{-170/30} = 0.0311$; the even ground state has weight one. The equilibrium odd fraction is approximately 0.0302. A frozen sample at odd fraction 0.75 stores about $0.75(170)k_B = 127.5k_B$ kelvin of rotational energy per molecule, versus about $5.1k_B$ kelvin for equilibrium. The excess expressed per mole is roughly $R(122.4\,\mathrm{K}) = 1.02$ kJ/mol.

The frozen low-temperature heat capacity can be small even while substantial energy remains trapped in the odd sector. A calorimeter monitoring a slowly converting sample measures time-dependent heat release, not simply the derivative of a stationary restricted partition function. Measure the spin fraction independently or vary equilibration time before comparing its heat curve with an equilibrium prediction.

Cambridge Statistical Physics · molecular partition functions and internal degrees of freedom

Bose condensation, dimensionality and a finite trap

Statistical mechanics from microscopic counting to material response

Cooling a gas of identical bosons can populate a single orbital macroscopically. The temperature at which that occurs depends on the low-energy density of states, not simply on the energy needed to put one particle in the first excited level. Compare a uniform box with a harmonic trap, estimate a condensate fraction and design an imaging test that separates occupation of a mode from a narrow but entirely thermal cloud.

Derive the occupation before approximating the density of states

A single noninteracting bosonic orbital of energy ϵ_j may contain $n_j = 0, 1, 2, \ldots$ particles. In contact with a reservoir of temperature T and particle chemical potential μ, its grand partition factor is the geometric series $\sum_n e^{-\beta n(\epsilon_j - \mu)}$. Convergence requires $\mu < \epsilon_0$, where ϵ_0 is the ground energy. Differentiation gives mean occupancy $1/[e^{\beta(\epsilon_j - \mu)} - 1]$. Shift the energy origin to $\epsilon_0 = 0$, so fugacity $z = e^{\beta\mu} < 1$. For a fixed mean particle number solve $\sum_j \langle n_j \rangle = N$.

As temperature falls, the distribution narrows toward low energies and μ generally rises toward the ground energy. It does not cross that energy in an equilibrium ideal finite gas. A large mean ground-state population corresponds to $1 - z$ of order $1/N_0$. Once N_0 is comparable to the total, replacing that single low-lying orbital by a continuum density of states can erase the very population of interest. Separate it before approximating the excited states.

A bosonic occupancy of one per quantum state is not an upper limit. At classical high temperature typical occupation of each relevant orbital is small and $n_j \simeq z e^{-\beta\epsilon_j}$. The statistical difference from a dilute classical gas becomes appreciable when the thermal de Broglie volume is comparable to the volume per particle. This criterion indicates growing exchange effects; the existence of a finite-temperature condensation transition additionally depends on dimension, dispersion and trapping geometry.

Count the three-dimensional excited particles

Use periodic boundary conditions in a large cubic box with side L. Allowed wavevectors occupy cells $(2\pi/L)^3$ in reciprocal space. For one spin species, counting the spherical shell and substituting $\epsilon = \hbar^2 k^2/(2m)$ gives an excited-state density proportional to $V\sqrt{\epsilon}$. Expand $1/(z^{-1}e^{x} - 1) = \sum_{\ell \geq 1} z^\ell e^{-\ell x}$ before evaluating the radial Gaussian integrals. In terms of the thermal wavelength and Bose functions $g_s(z) = \sum_{\ell \geq 1} z^\ell/\ell^s$, the result is

$$\frac{N_{\mathrm{ex}}}{V} = \lambda_T^{-3} g_{3/2}(z),$$

$$p = k_B T \lambda_T^{-3} g_{5/2}(z),$$

$$U_{\mathrm{ex}} = \frac{3}{2} p V.$$

Pressure follows from the grand potential or the kinetic virial identity in a uniform nonrelativistic gas. The ground state at zero momentum contributes no extensive kinetic energy or pressure in the thermodynamic continuum expression. It can nevertheless carry a macroscopic particle number. These

formulas describe an ideal gas: interactions can add pressure and shift excitations even at very low temperature.

The series $g_{3/2}(1) = \zeta(3/2) = 2.6124$ converges. At a fixed temperature the excited population per volume therefore has a maximum. At fixed total density n, once this maximum falls below n, the remaining particles accumulate in the ground orbital. Equating the two numbers defines the bulk critical temperature, and the below-transition condensate fraction follows:

$$T_c = \frac{2\pi\hbar^2}{m k_B}\left[\frac{n}{\zeta(3/2)}\right]^{2/3},$$

$$\frac{N_0}{N} = 1 - (T/T_c)^{3/2}, \quad T < T_c.$$

For a single-state gas of atoms with mass 87 unified atomic mass units at density 10^{20} per cubic metre, $T_c \simeq 398$ nK in this ideal box model. At $0.70 T_c$ the predicted condensate fraction is 0.414. Doubling density raises T_c by a factor $2^{2/3} = 1.587$. If imaging calibration rescales density but the reported temperature is independent, such a shift can move an inferred reduced temperature significantly. A finite trap or several spin components must be modelled before turning these example numbers into a laboratory calibration.

Why a uniform two-dimensional ideal gas is different

For free quadratic dispersion in d dimensions the density of states per volume scales as $\epsilon^{d/2-1}$. As $\mu \to 0^-$, occupancy behaves like $k_B T/\epsilon$ at low energy. The excited-population integral near its lower limit therefore behaves as $\int_0 \epsilon^{d/2-2} d\epsilon$. It converges only for $d > 2$. A uniform three-dimensional ideal gas saturates; in two dimensions the integral diverges logarithmically and in one dimension even more strongly. In the infinite uniform two-dimensional ideal gas at fixed density a finite positive temperature does not require a macroscopic ground occupation.

The two-dimensional integral can be evaluated explicitly. Per unit area $n_{2D} = -\lambda_T^{-2}\log(1-z)$. At phase-space density $n_{2D}\lambda_T^2 = 5$ this gives

$$z = 1 - e^{-5} = 0.99326.$$

In the formal grand description the mean ground occupation would be about 147 for this chemical potential, large on an atomic scale but still independent of total area at fixed density in the thermodynamic limit. Enlarging a system is how one distinguishes a finite prominent occupation from an extensive one.

A strictly finite box has discrete energies, so even a nominally two-dimensional sample can show a finite-size crossover with a narrow central state. Gapped transverse excitations may also make a nominally planar cloud effectively three-dimensional at sufficiently high temperature. Interacting two-dimensional fluids can exhibit a vortex-mediated transition involving long-range phase correlations of a different kind. The ideal uniform density-of-states argument neither derives nor rules out that interacting transition.

Count a harmonic trap in phase space

In a three-dimensional harmonic trap with frequencies $\omega_x, \omega_y, \omega_z$, subtract the zero-point ground energy. Excited energies are $\hbar(\omega_x n_x + \omega_y n_y + \omega_z n_z)$ with nonnegative integers. When $k_B T$ exceeds all the level spacings by a large factor, approximate the number of states below energy ϵ by the volume of a positive octant tetrahedron in the three occupation-number coordinates. Its derivative yields

$$D_{\mathrm{trap}}(\epsilon) = \frac{\epsilon^2}{2\hbar^3 \omega_x \omega_y \omega_z},$$

$$N_{\mathrm{ex}} = \left(\frac{k_B T}{\hbar \bar{\omega}}\right)^3 g_3(z),$$

$$\bar{\omega} = (\omega_x \omega_y \omega_z)^{1/3}.$$

Condensation now gives

$$k_B T_c = \hbar \bar{\omega} [N/\zeta(3)]^{1/3},$$

$$N_0/N = 1 - (T/T_c)^3.$$

In a trap with geometric mean frequency 100 Hz and $N = 10^5$, the ideal estimate is 210 nK. At 0.70 of this trap-specific critical temperature the condensate fraction is 0.657, appreciably greater than the 0.414 box result at the same reduced temperature. Mixing a trap density of states with a homogeneous critical formula cannot be repaired by inserting one arbitrarily chosen cloud volume.

For a harmonic trap in d dimensions the low-energy density behaves as ϵ^{d-1}. Excited-state saturation in this semiclassical trapped model already occurs for $d > 1$. The changed criterion arises because a higher-energy particle also explores more physical space in the trap. The thermodynamic limit must be specified: holding all trap frequencies fixed while adding arbitrarily many particles changes the central density. A conventional trapped limit sends the trap frequencies to zero with a suitable combination such as $N\bar{\omega}^3$ fixed.

Connect a narrow peak to occupation and spatial coherence

After an ideal noninteracting cloud is released, a particle initially at $\mathbf{r}_0$ and momentum $\mathbf{p}$ reaches $\mathbf{r}_0 + \mathbf{p}t/m$. At long expansion times the position distribution maps approximately to momentum distribution after accounting for initial cloud size. A condensate occupying one trap orbital can therefore produce an extended but narrow component in a momentum image superposed on a broader thermal distribution. Ballistic expansion length, resolution and interaction-driven expansion all affect the observed width.

For the 87-unit-mass example at 200 nK a one-axis thermal velocity scale is

$$\sqrt{k_B T/m} = 4.37\,\mathrm{mm/s}.$$

After 20 ms its thermal contribution to one-axis rms width is about 87 micrometres. A point-spread width of ten micrometres is a small correction for this broad part but can dominate a genuinely much narrower central component. Fit integrated particle number and the spatial response of each component. A Gaussian peak height alone increases when the same number of atoms occupies a narrower instrument function.

A more fundamental criterion uses the one-particle density matrix. Its eigenvalues are orbital occupations; a condensate has an eigenvalue proportional to the total number as the appropriate system size grows. Interference measurements of first-order coherence or occupation scaling over different atom numbers can test this picture. Condensation is not identical to zero viscosity or a critical flow speed. An ideal uniform quadratic gas has low-energy excitations with arbitrarily small energy per momentum and does not by itself predict the finite critical speed of an interacting superfluid.

Solve the discrete number equation before passing to a continuum

A small cold gas in a box is an excellent numerical test because the discrete energy levels, their degeneracies and their individual occupancies are known. For a periodic cube with ground energy zero, write its excited energies as $\epsilon_{\mathbf{j}} = \epsilon_L(j_x^2 + j_y^2 + j_z^2)$, where $\epsilon_L = 2\pi^2\hbar^2/(mL^2)$. The first shell has six states, the next twelve and the third eight. Replacing each shell by one level without its vector degeneracy artificially drives particles into the ground state.

For a specified finite number N, solve $N = \sum_{\mathbf{j}} [e^{\beta(\epsilon_{\mathbf{j}} - \mu)} - 1]^{-1}$ with a chemical potential strictly below zero. This is a monotone root problem: as μ goes from minus infinity toward zero the sum grows from zero without bound owing to the ground state. Use an energy cutoff well above the available thermal energy and estimate the omitted tail, increasing the cutoff to check convergence. Near zero, evaluate $e^x - 1$ with a numerically stable small-argument function. A naive evaluation of $e^x - 1$ can erase the tiny denominator representing a macroscopically occupied orbital.

A useful artificial example has $\beta\epsilon_L = 2$ and a mean ground occupation of 99. This fixes $z = 99/100 = 0.99$. Each of the six first-shell levels contains on average $1/[e^2/0.99 - 1] = 0.1548$ particles, contributing 0.929 in total. The twelve next-shell levels at $4k_BT$ contribute another 0.222 together. Their combined occupation is already greater than one even though every individually excited orbital is weakly occupied. A code can compare an explicit vector enumeration with a shell-degeneracy table and verify exact agreement for the same cutoff.

The finite grand-canonical solution fixes mean N, not the integer count in each run. A canonical simulation with exactly N can yield similar mean populations in an appropriate large system but very different ground-state number fluctuations. In the grand ensemble one isolated orbital has a geometric number distribution with variance $\langle n\rangle(1 + \langle n\rangle)$. A mean ground population of 99 then corresponds to a standard deviation close to the mean. A sealed sample of exactly 100 atoms cannot physically realise this unbounded distribution. Comparing its observed condensate fluctuations to the independent grand-canonical orbital formula without enforcing number conservation would be misleading.

The heat capacity sees the excited cloud, not a hidden latent heat

Below the bulk transition in a uniform ideal gas the fugacity tends to one. The energy expression becomes

$$U = (3/2)Vk_BT\lambda_T^{-3}\zeta(5/2),$$

which grows as $VT^{5/2}$. At fixed box size and fixed total atom number, the excited fraction adjusts with temperature while the ground state supplies or accepts the remaining particles. Differentiating gives

$$C_V = (15/4)Vk_B\lambda_T^{-3}\zeta(5/2).$$

Dividing by the total particle count gives a heat capacity per particle proportional to $(T/T_c)^{3/2}$.

At the critical temperature just below the transition, the coefficient per particle is $15\zeta(5/2)/(4\zeta(3/2)) \simeq 1.93k_B$. At half that temperature it is about $0.681k_B$. The energy and entropy are continuous at the uniform ideal transition; saturation of the excited population is not the same as a first-order phase change with a latent heat between two bulk densities. In the harmonic trap the different density of states yields a different power and a different shape of the heat-capacity anomaly. Thermodynamic measurements can therefore help distinguish trap geometry when imaging alone is ambiguous.

A calorimetric estimate of temperature often assumes the gas can re-equilibrate after a controlled energy pulse. In a perfectly noninteracting isolated gas, occupation numbers of stationary orbitals do not redistribute by themselves. Weak collisions, a reservoir or an explicit thermalisation protocol must establish equilibrium. The ideal gas can be an excellent equation-of-state model while weak interactions remain necessary for reaching that state on laboratory timescales. If the waiting time after a pulse is shorter than the redistribution time, substituting the equilibrium heat capacity into $\Delta T = \Delta U/C$ may falsely imply anomalous heating.

One practical diagnostic is to vary the waiting time while monitoring both the broad-wing kinetic temperature and central fraction. Thermal equilibrium predicts a definite relation between those two quantities at fixed total number and geometry. A lagging central fraction together with apparently stable wings suggests a slow redistribution process or atom losses. Repeating with matched final number and with independently calibrated detection efficiency is more discriminating than fitting temperature from the central peak alone.

Detect bosonic bunching without assuming a condensate

Occupation can also be probed through intensity or number correlations. For a single thermal bosonic mode with geometric distribution, the normally ordered second-order correlation is

$$g^{(2)}(0) = \langle n(n-1)\rangle/\langle n\rangle^2 = 2.$$

A coherent-state Poisson number distribution gives one. A precisely fixed number n gives $1 - 1/n$. These statements concern the statistics of the selected mode and the preparation. They must not be

read as a universal test in which every condensed cloud has the same value under every ensemble or detection geometry.

In a detector pixel collecting M independent identically populated thermal modes, the mean total number is $M\bar{n}$, variance $M\bar{n}(1+\bar{n})$. Substituting into the factorial moment gives

$$g^{(2)}(0) = 1 + 1/M.$$

Ten unresolved thermal modes therefore give 1.10 rather than 2, even though each separately bunches fully. Finite temporal averaging, velocity-bin width and spatial point spread can increase the effective number of collected modes. A near-unity measured correlation does not by itself demonstrate that the gas has developed a macroscopically coherent orbital.

Why use a factorial moment? For independent shot noise the ordinary second moment contains a self-count term $\langle n \rangle$. Subtracting it isolates pair coincidences between distinct particles. A detector that cannot distinguish simultaneous events may preferentially drop exactly those pairs, creating artificial suppression. Two independent detection channels after a calibrated splitter can estimate coincidences without a single-pixel dead-time bias, provided cross talk and timing jitter are characterised.

A useful exercise varies both integration width and total atom number. A correlation excess falling with broader detection windows supports collection of multiple unresolved thermal modes. A central mode whose occupancy remains a fixed fraction as the whole gas grows supplies different information. Neither test replaces a correctly modelled spatial coherence measurement, yet together they constrain interpretations that can fit one mean momentum image.

How a number solver reveals its own near-critical sensitivity

At fixed temperature, derivative of total ideal boson occupation with respect to chemical potential is $\beta \sum_j n_j(1+n_j)$. A highly occupied lowest orbital can dominate the derivative even if the measured energy is primarily carried by the excited states. Hence a small error in chemical potential can produce a large error in number while barely changing calculated energy. A solver that stops when energy converges may report the wrong condensate fraction.

For a mean ground occupation 10^4, its chemical-potential distance below the lowest level is $k_B T \log(1 + 10^{-4})$, approximately $10^{-4} k_B T$. A shift in that tiny gap by ten percent changes the occupancy by roughly ten percent while leaving the apparent fugacity close to one to four significant digits. Track chemical potential relative to the exact ground energy and report particle-number residual directly. Conversely, a small absolute error in a reported fugacity near saturation is not by itself evidence for a large uncertainty in the empirically counted atom number.

Finally, compare continuum estimates on several finite boxes at the same density. The leading discrepancy often comes from low discrete modes rather than the high-energy cutoff. Refine the low-energy calculation and the tail independently: enumerate a set of low shells exactly and replace only the dense upper spectrum by an integral. Check for double counting at the matching cutoff. An apparently excellent bulk critical temperature at one size can otherwise result from two cancelling approximations with different size dependence.

Infer a box density from a harmonically confined thermal cloud

Far from degeneracy a harmonically confined gas has classical spatial density $n(\mathbf{r}) = n(0)e^{-V(\mathbf{r})/(k_B T)}$. For an isotropic trap $V = m\omega^2 r^2/2$, each coordinate is Gaussian with variance $s^2 = k_B T/(m\omega^2)$. Normalising its full three-dimensional density gives

$$n(0) = N/[(2\pi)^{3/2} s^3].$$

An absorption image generally measures a column integral, whose central value is $N/(2\pi s^2)$, with units per area. Substituting this column density into a volumetric condensation criterion is dimensionally incorrect.

For $N = 10^5$, mass 87 unified atomic units, ordinary trap frequency 100 Hz and temperature 400 nK, the classical one-axis size is about 9.85 micrometres. The inferred central volume density is 6.65×10^{18} per cubic metre. The phase-space density estimate $n(0)\lambda_T^3$ is about 0.17; the classical profile is therefore a useful leading reference rather than a precision model at all radii. A ten-percent systematic underestimate of width inflates the inferred central density by about 37 percent because width enters cubed.

Near the transition the local-density approximation replaces the constant chemical potential by $\mu - V(\mathbf{r})$ inside the uniform-gas density function. This preserves the global chemical potential while reducing local effective fugacity away from the centre. Integrating the resulting local density over position recovers the harmonic-trap density-of-states counting in the semiclassical regime. A central phase-space density approaching the uniform saturation value can therefore coincide with a trap-wide critical atom number governed by $\zeta(3)$, not $\zeta(3/2)$. These statements use different spatial integrations and are consistent.

A practical reconstruction fits widths along multiple axes and includes known imaging point spread. For a Gaussian image convolved with a Gaussian response, variances add. A measured width of 10.0 micrometres and known response width 3.0 micrometres imply true width 9.54 micrometres, changing a cubed-density estimate by roughly fifteen percent. A narrow condensate component often has a non-Gaussian profile or a size comparable with the resolution, so the simple Gaussian deconvolution should not be applied indiscriminately to both components.

Compare inferred atom counts over increasing fields of view. Truncating broad thermal wings undercounts the thermal population and makes the apparent condensed fraction too large, particularly near onset. Independent normalisation from the complete cloud and fits to held-out outer regions help separate this effect from an actual change in the fraction occupying the lowest orbital.

There is a useful contrast with a single-occupancy adsorption site. Its occupation variance at mean fraction f is $f(1-f)$, whereas one thermal Bose orbital at mean occupation $\bar{n}$ exhibits $\bar{n}(1+\bar{n})$. At mean 0.50 these are respectively 0.25 and 0.75. The sign difference expresses whether multiple occupancy is excluded or enhanced. Vary the chemical potential in the single-site experiment and compare the response-to-noise identity before evaluating the Bose derivative on paper. Both obey the same grand-ensemble fluctuation formula, yet their number variances differ at the same mean because the allowed microscopic configurations differ. An observed sub-Poisson count can alternatively arise from detector dead time, and multiple unresolved thermal modes reduce bunching, so the ideal comparison is strongest for calibrated single-mode or single-site measurements.

Compare the measured single-site fluctuation law with the Bose orbital variance

Count the excited atoms in a trapped gas

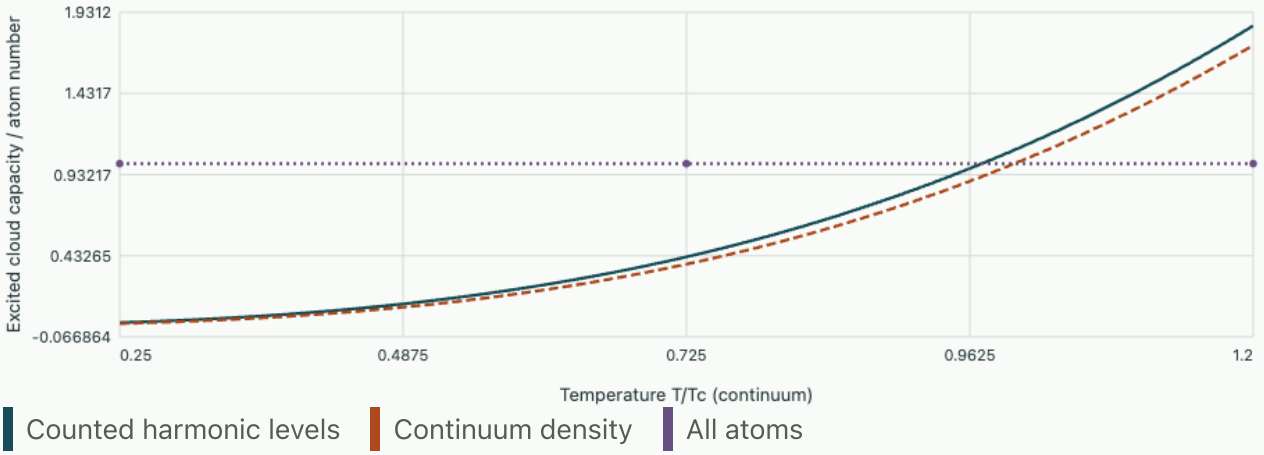
Change atom number and confinement frequency. Compare the discrete level sum with the continuum estimate and relate the difference to the density of states.

Open this experiment online

Try the calculation

Atom number **20000**

Isotropic trap frequency (Hz) **100**



Continuum Tc 122.52 nK. The discrete excited levels reach N at 119.17 nK.

Tutorial · Check what changes during an adiabatic box expansion

Below the transition a dilute ideal condensed box gas expands slowly and reversibly with fixed N. Determine how temperature and condensate fraction vary with volume. Would an isothermal expansion give the same outcome?

▼ **Use the entropy and critical-temperature scalings together**

Below transition the bulk chemical potential tends to zero. Since $U = 3pV/2$, $S = (U + pV - \mu N)/T$ is proportional to $VT^{3/2}$ for the saturated excited gas. At constant entropy $T \propto V^{-2/3}$. The box critical temperature at fixed N has exactly the same volume dependence, so T/T_c and the leading condensate fraction stay constant. In an isothermal expansion T_c falls while T stays fixed, decreasing condensed fraction and eventually exhausting it. The result rests on ideal-gas equilibrium and reversible expansion; a sudden opening into a second container is not that adiabatic path.

Tutorial · Determine whether a trap can be treated as three-dimensional

An anisotropic harmonic trap has ordinary frequencies 40, 40 and 8000 Hz at a measured $T = 100$ nK. Compute the spacings expressed as temperatures and assess using the three-dimensional semiclassical trap formula. Propose a consistent calculation near the dimensional crossover.

▼ **Identify which quantum numbers are actually sampled**

Using $h/k_B = 4.799 \times 10^{-11}$ K s, the level spacings are 1.92, 1.92 and 384 nK. Many transverse levels in the soft directions are populated, whereas a single quantum along the tight axis costs nearly four thermal energies. A three-dimensional integral over all three quantum numbers would overcount accessible low-energy states. Keep the tight-axis sum discrete while approximating the two soft directions semiclassically. Separate the exact global ground orbital from that hybrid excited-state count before solving for chemical potential at fixed total number.

At a few hundred nanokelvin the next tight-axis branches start to contribute appreciably. Observe widths or occupation across a temperature scan instead of announcing a fixed dimensionality based solely on trap shape. The model should recover a two-dimensional low-temperature limit and cross over smoothly toward three-dimensional counting when all spacings become small relative to thermal energy.

Tutorial · When is a ground-state peak extensive?

Three otherwise comparable measurements at fixed estimated areal density and temperature use planar boxes of areas in the ratio one, two and four. They yield approximately the same central-state population of 150 atoms but total numbers 3000, 6000 and 12000. Compare with the ideal planar fugacity estimate above and explain what conclusion follows.

▼ **Examine scaling and the resolution of a single mode**

The fractions are 0.050, 0.025 and 0.0125, falling inversely with area. This is consistent with a highly occupied but nonextensive single level at fixed fugacity near 0.993. A homogeneous three-dimensional bulk condensate at fixed reduced temperature would instead have approximately constant fraction under the appropriate volume scaling. Absolute occupation larger than one is insufficient to distinguish the two regimes.

Confirm that the same experimentally defined central window isolates a single momentum mode as the box size changes. Allowed wavevector spacing shrinks with system size, so a fixed physical momentum aperture can sum an increasing number of levels. Correct for finite resolution, compare the complete momentum distributions and control interactions before treating the simple size dependence as a proof about the thermodynamic limit.

Cambridge Statistical Physics · ideal quantum gases and Bose condensation

Degenerate fermions, heat and magnetic response

Statistical mechanics from microscopic counting to material response

An ordinary metal at room temperature contains electrons with characteristic kinetic energies of several electronvolts, yet its electronic heat capacity is small compared with the classical estimate. Identify which electrons respond to heat, compression and a weak magnetic field. Then compare three independent ways to determine a density of states and investigate what a disagreement would mean.

Count states and fill them with fixed particle number

A spin-half noninteracting fermion orbital of specified spin can be empty or occupied once. Its grand partition factor is $1 + e^{-\beta(\epsilon-\mu)}$, giving mean occupation $f = 1/[e^{\beta(\epsilon-\mu)} + 1]$. At zero temperature the occupations form a step. In a uniform three-dimensional box, fill both spin states at each wavevector up to radius k_F. Counting reciprocal-space cells gives

$$n = \frac{k_F^3}{3\pi^2},$$

$$E_F = \frac{\hbar^2 k_F^2}{2m},$$

$$v_F = \frac{\hbar k_F}{m},$$

$$D(E) = \frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{E}.$$

Here D counts both spin projections per unit energy for the whole sample. It obeys $D(E_F) = 3N/(2E_F)$. Distinguish this density from a density per volume or per spin, since factors of volume and two matter in susceptibility and specific heat. A lattice band may have a quite different dispersion and density of states; simply substituting the vacuum electron mass is a model to be tested for a given solid.

At $n = 8.5 \times 10^{28}$ electrons per cubic metre and the free electron mass, $E_F = 7.05$ eV, $T_F = 81{,}800$ K and $v_F = 1.57 \times 10^6$ m/s. Room temperature is only 0.00367 of this Fermi temperature. The gas is highly degenerate despite not being near absolute zero in everyday terms. These speeds are random occupied-state group velocities, with equal positive and negative contributions to net current in equilibrium. A metal with no applied field does not transport its entire electron density in one direction at that speed.

Integrating $ED(E)$ to the Fermi level gives

$$U_0 = 3NE_F/5.$$

Volume differentiation at fixed N, using $E_F \propto V^{-2/3}$, gives pressure $p_0 = 2nE_F/5$. Thus a nonzero kinetic pressure remains at zero temperature. It describes the electron subsystem and is balanced by other forces in a neutral solid. It is not the pressure one could read by attaching an ordinary fluid pressure gauge to a solid metal specimen.

Derive the low-temperature expansion without losing particle conservation

For a smooth function $H(E)$, compare $\int H(E)f(E)\,dE$ with its zero-temperature integral up to μ. Write the difference as contributions above and below μ, substitute $x = |E-\mu|/(k_BT)$ and use $1 - f(\mu - k_BTx) = f(\mu + k_BTx)$. The difference then samples $H(\mu + k_BTx) - H(\mu - k_BTx)$, so only odd derivatives of the smooth function contribute. Integrating the leading term gives

$$\int_0^\infty H(E)f(E)\,dE = \int_0^\mu H(E)\,dE + \frac{\pi^2}{6}(k_BT)^2H'(\mu) + O(T^4).$$

This expansion assumes the Fermi level is many thermal energies above the bottom and that H varies smoothly through the thermal shell. Near a band edge or a very narrow density-of-states peak the assumption can fail even when temperature looks modest relative to some larger energy. Direct numerical quadrature retaining the correct band limits and a particle-number solution for μ is then preferable.

For fixed N, use $H = D$ and demand that the result remain its zero-temperature value. The chemical potential shift is $\delta\mu = -(\pi^2/6)(k_BT)^2D'(E_F)/D(E_F)$. For $D \propto \sqrt{E}$ it becomes

$$\mu \simeq E_F[1 - (\pi^2/12)(T/T_F)^2].$$

Next use $H = ED(E)$ in the energy integral. Its T^2 contribution from the shifted upper limit cancels the part involving E_FD', leaving

$$U - U_0 = \frac{\pi^2}{6}D(E_F)(k_BT)^2,$$

$$C_V = \frac{\pi^2}{3}k_B^2TD(E_F).$$

The typical number of thermally active state slots scales as thermal energy times $D(E_F)$. This produces a heat capacity linear in T, not proportional to the entire electron count times a temperature-independent classical factor. At 300 K in the example the electronic heat capacity is $0.0181Nk_B$. Its volume coefficient $C_V/(VT)$ is about 70.8 J per cubic metre per kelvin squared. A much larger lattice contribution may hide it at room temperature. At low temperature fit C/T against T^2 to separate an electronic intercept from a Debye T^3 lattice term where these asymptotic laws are valid.

Compare the conserved-number chemical potential with the Sommerfeld approximation

Apply a weak field to the spin populations

For a weak uniform magnetic induction B neglect orbital coupling temporarily. Spin moments of magnitude μ_B acquire energy offsets $\mp\mu_BB$ for their two alignments in the simple electron model. At fixed total density and zero temperature, one spin population gains states at its Fermi surface and the other loses an equal number to first order. The density of states per spin is $D(E_F)/2$; the difference in population is thus $D(E_F)\mu_BB$. Magnetisation is this difference times μ_B/V, giving the weak-field SI susceptibility

$$M = \frac{\mu_B^2 D(E_F)}{V} B,$$

$$\chi_P = \frac{\partial M}{\partial H} \simeq \mu_0 \mu_B^2 D(E_F)/V.$$

We have used $B \simeq \mu_0 H$ in a weakly magnetic macroscopic state and ignored demagnetising corrections. The result is approximately temperature-independent for a strongly degenerate smooth band. Independent localised moments would give a Curie susceptibility proportional to inverse temperature in their high-temperature weak-field regime. A small low-temperature upturn can therefore come from a small concentration of local moments even when most carriers remain itinerant.

For the earlier density $\chi_P \simeq 1.22 \times 10^{-5}$. A field of 1.0 T splits spin energies by $2\mu_B B = 0.116$ meV, vastly smaller than 7.05 eV. The fractional spin imbalance is $3\mu_B B/(2E_F) \simeq 1.23 \times 10^{-5}$. Tiny spin imbalance can produce a reproducible measurable magnetisation because the sample contains a macroscopic number of electrons. Linear susceptibility should be tested by field reversal. A hysteretic or saturating impurity background cannot be removed reliably by fitting one positive-field point.

At finite temperature the response replaces sharp $D(E_F)$ by $\int D(E)[-\partial f/\partial E]dE$ at the corresponding chemical potential. The negative occupation derivative integrates to approximately one and has width several thermal energies. This Fermi-window average is another reason the susceptibility and heat capacity emphasise nearby states rather than the entire filled band. Their precise kernels differ, so a rapidly varying band structure or interactions can alter their ratio.

Infer the same states through a density response

At zero temperature density varies with chemical potential as $\partial n/\partial \mu = D(E_F)/V$. Since $dp = n\, d\mu$ at fixed temperature, the electronic inverse compressibility $1/\kappa_T = n^2/(\partial n/\partial \mu) = 2nE_F/3$. In the example this corresponds to about 64 GPa. A compressibility measurement of the full solid includes lattice elasticity and charge-neutrality effects, so it cannot generally be substituted directly for an isolated electron-fluid result.

A confined electronic layer contacted through a gate can instead provide a thermodynamic density response through its quantum capacitance. If an induced charge increment per area is $e\, dn$, part of the applied electrochemical voltage changes chemical potential by $d\mu/e$. The corresponding capacitance per area is $C_Q/A = e^2(\partial n_{2D}/\partial \mu)$. In a simple series combination the geometric dielectric capacitance and quantum capacitance both reduce the measured gate capacitance. Comparing a measured total directly to $e^2 D$ without removing the geometrical part can create an artificial density-of-states suppression.

Three independent measurements of low-temperature C/T, Pauli spin response and a properly interpreted compressibility all probe the density of states in this noninteracting model. Their ratio then removes $D(E_F)$ from some combinations. Correlations can renormalise spin response differently from specific heat. Orbital diamagnetism, closed-shell ions and temperature-dependent impurity moments must also be separated from a measured total susceptibility. A disagreement is an opportunity to distinguish these separate ingredients, not immediate evidence that state counting itself has failed.

A nonlinear zero-temperature spin response beyond Pauli's slope

At low magnetic fields the shift of either spin population is linear and nearly invisible in total number. To see where the susceptibility approximation comes from, retain the exact zero-temperature band populations. For a parabolic three-dimensional band with field-independent mass, absorb the two equal spin degeneracies into a coefficient A such that $n_{\pm} = A[\mu \pm \mu_B B]_+^{3/2}$. The positive part enforces an empty spin band if its shifted edge lies above the chemical potential. The total $n_+ + n_-$ must remain fixed in an isolated neutral sample, while $M = \mu_B(n_+ - n_-)$.

Expand both powers at small field. Linear terms cancel in the total density, so the chemical potential changes only at quadratic order. At fixed original Fermi energy the solution is

$$\mu = E_F - (\mu_B B)^2/(4E_F) + \cdots.$$

Population imbalance retains the linear term $3\mu_B B/(2E_F)$ as a fraction of total density. This derivation checks why simply leaving μ fixed gives the correct weak-field magnetisation slope yet fails in higher-order density calculations.

Complete spin polarisation in this model occurs when $\mu - \mu_B B$ reaches zero. All particles then occupy the lower spin band, with effective Fermi energy relative to that band's shifted bottom equal to $2^{2/3}E_F$. It follows that $\mu_B B_{\mathrm{pol}} = 2^{-1/3}E_F$. For a 7.05 eV metal this corresponds to roughly 9.7×10^4 tesla, far outside the regime of an ordinary static bench measurement and outside a naive description of unmodified material bands. The number is a scale comparison, not a proposed magnet design. A system with a much lower Fermi energy or an enhanced effective magnetic moment can show nonlinearity at far lower applied fields.

A practical diagnostic compares positive and negative fields after subtracting the holder signal. Pure weak Pauli response is odd and linear in field with little temperature dependence. Localised impurity spins may also be odd but exhibit curvature or saturation when their Zeeman energy competes with temperature. A carrier density changing with field can alter the inferred Pauli slope. Therefore pair a magnetic measurement with a density or transport measurement before attributing every nonlinearity to spin-band depletion.

Extract a two-dimensional density of states from a gate measurement

Take a uniform conducting layer gated by a planar metal through an insulating spacer. An increment in gate voltage partly changes electrostatic voltage across the insulator and partly shifts the electronic chemical potential. In a simple uniform single-sheet model the charge-density increment $dq = e\,dn_{2D}$ obeys $dV_g = dq/c_g + d\mu/e$, where lowercase c_g is geometric capacitance per area. Writing $c_q = e^2\,dn_{2D}/d\mu$ gives

$$1/c_{\mathrm{meas}} = 1/c_g + 1/c_q.$$

This derivation assumes a stable uniform equilibrium state and a calibrated voltage reference with negligible leakage.

Suppose independent dielectric calibration gives $c_g = 5.00\,\mu\mathrm{F\,cm^{-2}}$, while a low-frequency gate measurement yields

$$c_{\mathrm{meas}} = 3.00\,\mu\mathrm{F\,cm^{-2}}.$$

The inferred quantum capacitance is $c_q = 7.50\,\mu\mathrm{F\,cm^{-2}} = 0.0750$ F per square metre. Dividing by elementary charge squared gives 2.92×10^{36} states per joule per square metre, or 4.68×10^{17} states per eV per square metre. Confusing total sample capacitance with capacitance per area would produce a quantity that cannot be compared across different device sizes.

Sensitivity of the inverse step depends strongly on the ratio of the two measured values. If $c_q \gg c_g$, the measured value approaches the geometric one. Small calibration offsets then dominate their reciprocal difference. For example, a measured value only one percent below c_g implies $c_q \simeq 99 c_g$; an absolute error of one percentage point can remove any useful upper bound. In that limit the data support a lower bound on density of states rather than a precise central value.

Vary gate frequency to test whether traps or charge-relaxation delays contribute. A legitimate equilibrium compressibility response should be approached in the low-frequency limit under conditions where the electronic layer remains coupled to its reservoir. Interface states with slower charging can add a dispersive contribution that looks like extra density of states at one chosen frequency. Compare in-phase capacitance and loss, control for parasitic capacitance and measure several layer areas. Area-proportional capacitance and perimeter-dependent fringe capacitance can often be separated with devices of differing aspect ratio.

A differential heat pulse resolves the small electronic component

At low temperature a metal frequently has $C_{\mathrm{sample}} = \gamma T + \beta_D T^3$ over a limited range. A platform and thermometer add their own heat capacities. A measurement might apply a known energy pulse and fit a temperature relaxation through a calibrated weak thermal link. To infer instantaneous heat capacity from $Q/\Delta T$, the pulse must be small enough that C is nearly constant across the excursion and losses over the pulse negligible, or the energy balance must explicitly include them.

For a 10 mmol sample with $\gamma = 0.80$ mJ per mole per kelvin squared and $\beta_D = 0.10$ mJ per mole per kelvin to the fourth, molar heat capacity at 2 K is $1.60 + 0.80 = 2.40$ mJ per mole kelvin. The sample heat capacity is 24.0 microjoules per kelvin. A 1.2 microjoule pulse would raise it by about 0.050 K neglecting heat leakage and addenda. The electronic fraction here is two thirds. At 4 K the lattice contribution exceeds the electronic term, changing which temperatures most strongly constrain the intercept in a C/T fit.

An exact small-range energy calibration uses the integral $Q = \gamma(T_f^2 - T_i^2)/2 + \beta_D(T_f^4 - T_i^4)/4$ times the sample amount. This avoids confusing a finite pulse average with the capacity at the initial temperature. Measure the empty platform through the same range and apply the same thermal model. A mismatch in thermal contact may create a second relaxation time and imitate sample heat capacity. Two pulses of different sizes and a delayed thermometer provide an independent test of the assumed lumped-temperature approximation.

The density of states inferred from γ measures excitations that can equilibrate during the experiment. A superconducting transition, magnetic ordering, a low-lying impurity contribution or a gap around the Fermi level would change the temperature dependence qualitatively. Rather than forcing a linear plus cubic fit over such structure, identify a normal-state range or include a separately justified physical term. A visually straight line on a narrow interval can hide curvature large enough to alter the intercept materially.

How a Fermi window rounds a sharp density-of-states edge

A two-dimensional band with a constant density of states starting abruptly at energy zero provides an exactly soluble compressibility edge. Differentiating its number equation gives $dn/d\mu = D_0/[1 + e^{-\beta\mu}]$ if D_0 is per area. At zero temperature the response jumps from zero to D_0 as the chemical potential crosses the band edge; at finite temperature it rises as a logistic function. The width between ten and ninety percent equals $2k_B T \log 9$, approximately $4.39 k_B T$.

At 10 K that intrinsic energy width is about 3.79 meV. At 30 K it becomes 11.36 meV. A gate experiment converts voltage to energy using a lever arm that may depend on the same quantum capacitance being measured. Calibrate that conversion before inferring temperature from edge width. Static disorder, inhomogeneous local band edges or electrical noise broaden the edge even at low temperature. Plot width against temperature and compare the full convolution model rather than interpreting a nonzero intercept as an electronic thermal floor without independent evidence.

This edge example shows exactly where a low-temperature expansion in smooth density derivatives fails. Every derivative of the constant band interior is zero, yet if the chemical potential lies within a few thermal energies of its abrupt edge there is a strong measurable density response crossover. Numerical integration of the full Fermi function or the exact analytic expression automatically retains this boundary. A calculation based solely on the local derivative of a density of states sampled within the band cannot recover it.

Separate Pauli spin response from independent local moments with a concentration test

A common experimental model of weakly magnetic metallic matter combines an approximately constant itinerant-electron susceptibility with a dilute Curie term C/T from paramagnetic impurities. At modest fields the local moment of independent spin-half sites with magnetic moment magnitude μ has average $m = \mu \tanh(\mu B / k_B T)$. Its small-field derivative gives volume susceptibility $\mu_0 n_i \mu^2 / (k_B T)$. The conduction-electron spin response instead scales with the density of states at the Fermi level. Two subsystems exposed to the same bath can therefore have markedly different temperature dependences.

Suppose two volume-susceptibility readings in SI units are 2.20×10^{-5} at 10 K and 1.45×10^{-5} at 40 K. A fit to $\chi_0 + C/T$ gives $C = 1.00 \times 10^{-4}$ K and $\chi_0 = 1.20 \times 10^{-5}$. For independent moment magnitude μ_B, the implied impurity number density is $C k_B / (\mu_0 \mu_B^2) \simeq 1.28 \times 10^{25}$ per cubic metre. Relative to a carrier density of 8.5×10^{28} this is only about 150 parts per million. A small population can dominate the temperature-dependent change without representing most of the material.

Two points cannot verify the Curie form, so repeat across temperatures and fields. At sufficiently high B/T the independent moments approach saturation while the ideal degenerate electron subsystem remains approximately linear. At fixed B/T, plotting the temperature-dependent residual should approximately collapse for independent moments. Interacting or clustered impurities can display a shifted Curie–Weiss temperature, blocking or history dependence, and a distribution of moments can broaden the saturation shape. A successful two-point subtraction alone does not establish a unique background.

A related check uses calorimetry. A split two-level impurity contributes a field-dependent Schottky anomaly when its gap is several thermal energies. Changing field shifts this anomaly, whereas the leading electronic linear coefficient of a simple normal metal varies much less at modest fields. The integrated entropy of an isolated two-state population tends toward $n_i k_B \log 2$ when both orientations become equally populated. A count inferred from susceptibility can therefore be tested by a separately integrated excess heat capacity, once lattice and electronic backgrounds are constrained.

Take care with units and signs when comparing papers or instruments. Some report molar susceptibility, others volume susceptibility, and cgs and SI magnetic conventions differ by factors involving 4π and volume conversion. Divide a volume susceptibility by number density only after identifying whether the number refers to all host atoms, carriers or the inferred localised impurities. Agreement to a few significant figures after an unverified unit conversion has little diagnostic value.

A final test estimates when local spins should start to saturate at a given field. At 5 tesla and 2 K a magnetic moment equal to one Bohr magneton has $\mu_B B/(k_B T) \simeq 1.68$. The independent-spin alignment is $\tanh(1.68) = 0.933$, already close to saturation. In the same field the example metal's relative itinerant spin imbalance remains only around 6×10^{-5}. A common high-field measurement can therefore discriminate two very different microscopic populations even if both looked nearly linear at much smaller fields. Remanent magnetisation on returning to zero would be inconsistent with both ideal noninteracting equilibrium models. It would motivate a check for ferromagnetic contamination, anisotropic barriers or hysteresis in the measurement apparatus before refining a fit to either simple susceptibility law.

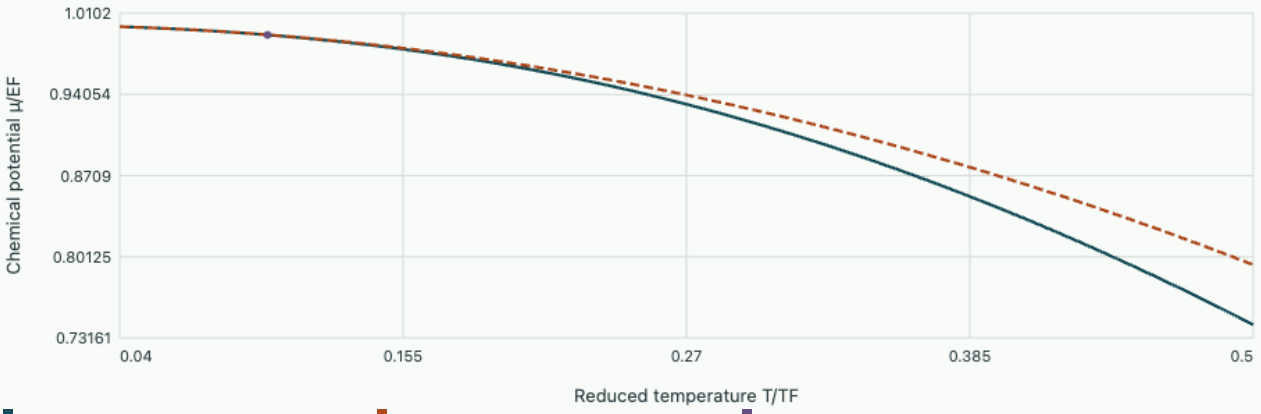
Decide which electronic states contribute to heat capacity

Change temperature relative to Fermi energy and compare the low-temperature approximation with the statistical occupation.

Open this experiment online

Try the calculation

Reduced temperature T/TF **0.1**



Fixed-density quadrature Quadratic approximation Selected solution

Fixed-number μ/EF = 0.9916, quadratic prediction 0.9918. Leading heat capacity $C_v/(Nk_\beta)$ = 0.4935.

Tutorial · Recover a density of states from a calorimeter

A sample exhibits a low-temperature molar heat capacity $C/T = 0.80\,\mathrm{mJ\,mol^{-1}K^{-2}} + bT^2$. Use the intercept to estimate the total density of electronic states per mole per electronvolt. Compare with the prediction for one conduction electron per atom and $E_F = 7.0$ eV.

▼ **Extract the electronic coefficient and convert energy units**

In joules the total molar density is $D = 3(0.00080)/(\pi^2 k_B^2) = 1.28 \times 10^{42}$ states per joule per mole. Multiplying by the number of joules per electronvolt gives 2.05×10^{23} states per eV per mole. For one electron per atom the ideal value is $3N_A/(2E_F) = 1.29 \times 10^{23}$ states per eV per mole. The fitted effective density is therefore about 1.59 times larger. A molar heat capacity must be paired with a molar density of states; using a sample volume here without a molar volume would create a dimensional mismatch.

To decide whether the enhancement reflects a nonfree band, interactions or a wrong temperature window, examine the linearity of C/T versus T^2 over several restricted fit ranges. A low-lying impurity or additional vibrational mode can produce curvature that changes the intercept. Compare with independent band information and spin response after relevant backgrounds are subtracted.

Tutorial · Determine a weak-field magnetisation and its background

At 1.0 T the measured volume magnetisation is 7.0 A/m, independent of temperature over a low-temperature range. The sample's predicted Pauli susceptibility is 1.22×10^{-5}. Estimate the residual susceptibility if the magnetisation is linear and explain why its sign alone does not identify an orbital model.

▼ **Compare quantities in the same magnetic convention**

One tesla corresponds to $H \simeq 7.96 \times 10^{5}$ A/m in this weakly magnetic setting. The predicted Pauli magnetisation is 9.71 A/m. The residual is minus 2.71 A/m or $\chi_{\mathrm{residual}} \simeq -3.4 \times 10^{-6}$. Orbital diamagnetism of an ideal bulk free-electron gas supplies a negative contribution, but core diamagnetism, band orbital effects and sample-holder background can also contribute. Measure an empty holder and both field signs, consider specimen shape and calibrate the sample volume. A nearly temperature-independent negative residual does not uniquely identify any one mechanism.

Tutorial · Why a constant two-dimensional density of states behaves differently

For an ideal two-dimensional parabolic band of constant total density D_0 above energy zero, derive the finite-temperature chemical potential at fixed $N = D_0 E_F$. Compare its leading low-temperature change with the power-law shift in three dimensions.

▼ **Evaluate the number integral before expanding**

The integral gives

$$N = D_0 k_B T \log[1 + e^{\mu/(k_B T)}].$$

Solving yields

$$\mu = k_B T \log[e^{E_F/(k_B T)} - 1] = E_F + k_B T \log[1 - e^{-E_F/(k_B T)}].$$

Its deviation from E_F at low temperature is exponentially small, consistent with $D'(E_F) = 0$. A program that always imposes the three-dimensional negative quadratic shift would violate particle conservation in this band. The low-temperature heat capacity remains linear because D_0 itself is nonzero.

At high temperature the same exact expression approaches the classical logarithmic chemical potential. Evaluate the low-temperature numerical form as $E_F + k_B T \log(1 - e^{-E_F/k_B T})$ to avoid overflowing the large exponential. Here numerical stability and correct dimensional density of states are independent requirements; solving one does not fix the other.

Cambridge Statistical Physics · Fermi gases, the Fermi surface and low-temperature response

Finite-size criticality and the evidence for a phase transition

Statistical mechanics from microscopic counting to material response

A finite magnet in zero field can reverse between two equally magnetised directions. Its exact mean magnetisation is then zero even below the temperature where a bulk material orders. Learn how to extract order, correlation length and critical exponents from finite measurements, and which apparent crossings or peaks would survive the thermodynamic limit.

What a finite magnet can and cannot spontaneously break

Consider a ferromagnetic Ising model with spins $s_i = \pm 1$, $N = L^d$ sites, short-range coupling $J > 0$, and uniform field h in energy units. Its Hamiltonian is $-J \sum_{\langle ij \rangle} s_i s_j - h \sum_i s_i$. A finite partition sum contains finitely many positive analytic terms at positive temperature. For finite N it cannot develop a true nonanalyticity on the real positive-temperature axis. Yet derivatives can become large and measured distributions strongly bimodal as size grows.

Define total magnetisation $M = \sum_i s_i$ and intensive order $m = M/N$. At zero field, spin inversion pairs configurations of M and $-M$, making the exact finite equilibrium mean zero. To define spontaneous order in the positive direction, take the thermodynamic limit at positive h, then let $h \to 0^+$. Reversing that order gives zero. A finite record that stays on one side because reversals are rare has not performed the full finite-system equilibrium average. It can approximate a symmetry-selected state only with that interpretation stated.

The absolute magnetisation or its root-mean-square value remains positive in finite zero-field calculations. Even independent disordered spins produce $\langle |m| \rangle$ of order $N^{-1/2}$. Therefore a nonzero magnitude in one simulation cannot establish bulk order. Repeat over sizes and inspect its scaling. A weak applied field, surfaces or slow field drift can also select one sign in a laboratory, modifying the finite-size comparison.

Derive a response from connected fluctuations

Differentiate the finite partition function at fixed temperature. The susceptibility per site to energy field h is

$$\chi_L = \frac{\partial \langle m \rangle}{\partial h} = \frac{\beta}{N} \left(\langle M^2 \rangle - \langle M \rangle^2 \right) = \beta \sum_{\mathbf{r}} G_c(\mathbf{r}),$$

where the last equality assumes a translationally invariant system and $G_c(\mathbf{r}) = \langle s_0 s_{\mathbf{r}} \rangle - \langle s_0 \rangle \langle s_{\mathbf{r}} \rangle$. Long correlations amplify the response because many sites contribute together when one responds to a field. For an experiment using magnetic induction instead of the energy field, include magnetic moments and unit-system factors.

In a symmetric finite zero-field sample below the ordering temperature, transitions between the two bulk-like peaks can give a large contribution proportional to N to the full fluctuation susceptibility. To estimate within-phase fluctuations one sometimes forms $\beta N (\langle m^2 \rangle - \langle |m| \rangle^2)$. This is a different observable and not exactly the derivative of the finite symmetric mean with respect to h. Report which definition is

plotted. Mixing those definitions in a size collapse can change both amplitudes and the apparent location of extrema.

A correlation length may be estimated from the large-distance connected correlation decay away from criticality, or from the ratio of small-wavevector structure factors. In lattice units, an often used periodic second-moment length is

$$\xi_L = \frac{1}{2\sin(\pi/L)}\sqrt{\frac{S(0)}{S(k_{\min})} - 1},$$

$$k_{\min} = 2\pi/L.$$

This relation uses the small-k expansion of the correlation structure factor and a lattice wavevector convention. Subtract means consistently, especially if estimating connected rather than symmetric correlations. If a small noisy $S(0)/S(k_{\min}) - 1$ is negative, replacing it by zero without describing the noise can bias a length. Larger samples or more suitable observables may be preferable.

Derive the finite-size variable from competing lengths

Write reduced temperature $t = (T - T_c)/T_c$. Suppose the bulk correlation length diverges as $\xi_\infty = \xi_0 |t|^{-\nu}$. A finite periodic sample cannot support correlations extending arbitrarily beyond L. The competition is governed by L/ξ_∞, equivalently $tL^{1/\nu}$ with its sign and a metric factor. If the bulk magnetisation, susceptibility and heat capacity have exponents β_m, γ, α, a standard finite-size ansatz below the appropriate upper critical dimension and without extra dangerous variables is

$$\langle |m| \rangle = L^{-\beta_m/\nu} F_m(tL^{1/\nu}),$$

$$\chi_L = L^{\gamma/\nu} F_\chi(tL^{1/\nu}) + \chi_{\mathrm{reg}}.$$

Here β_m is an order-parameter exponent, distinct from reciprocal thermal energy β. The functional forms match the bulk powers when the scaled argument becomes large at fixed nonzero t. A rounded finite-size peak of χ occurs at an approximately fixed scaled argument, so its temperature shift and width scale roughly as $L^{-1/\nu}$, allowing boundary and subleading corrections. Its height has an additional smooth background that can be significant for modest sizes.

For a two-dimensional short-range Ising universality class the known values $\nu = 1$, $\gamma = 7/4$, $\beta_m = 1/8$ give useful benchmarks. Doubling L at criticality should raise leading susceptibility by $2^{7/4} = 3.364$ and reduce order magnitude by only $2^{-1/8} = 0.917$. The very slow decrease of order is easy to mistake for nonzero bulk order at the critical temperature. The heat-capacity singularity in this example is logarithmic rather than a simple positive power; fitting a few peak heights to a power law can give a misleading small apparent exponent.

Compare a mean-field onset and susceptibility with the finite-size scaling argument

A dimensionless crossing and its systematic drift

A convenient shape diagnostic is $U_L = 1 - \langle m^4 \rangle / [3 \langle m^2 \rangle^2]$. A zero-mean Gaussian distribution has fourth moment three times variance squared and gives zero. Two infinitely narrow equal-weight peaks at $\pm m_0$ give $2/3$. At criticality, rescaling m by $L^{-\beta_m/\nu}$ cancels between numerator and denominator, so curves for different sizes can approach an approximately common crossing. The limiting value depends on geometry, boundary conditions and convention, despite the exponent universality.

Include a leading correction $U_L = U_* + a\,tL^{1/\nu} + bL^{-\omega} + \cdots$. Equating the expansions for L and sL shows that the pair crossing drifts as $t_\times \propto L^{-(1/\nu+\omega)}$ when the coefficients are nonzero. An exactly common crossing is therefore not guaranteed in finite data. Sample temperature resolution, uncertainty of higher moments and correlations between neighbouring reweighted temperatures can also shift the visual crossing.

For a data set with sizes 16, 32 and 64, suppose adjacent pair crossings occur at reduced-temperature-like coordinates 2.31 and 2.28 when plotted as $k_B T/J$. Without a calibrated correction exponent, taking their simple mean gives no justified bulk critical temperature. If a model independently predicts a crossing displacement decreasing by four per doubling, solve $T_{16,32} = T_c + A$, $T_{32,64} = T_c + A/4$, giving $T_c = 2.27 J/k_B$. Vary the correction assumption and use an additional larger size to test it. Extrapolated digits considerably beyond the scatter of these alternatives are not warranted.

Separate thermodynamic rounding from sampling error

Near criticality long spatial correlations often coincide with long relaxation times. A sequential Monte Carlo update can require many sweeps to produce independent configurations. If an observable has normalised autocorrelation $\rho(k)$, its effective independent sample count for a long record is approximately $N_{\mathrm{samp}} / [1 + 2 \sum_{k>0} \rho(k)]$. Blocking time records and recomputing Binder ratios from blocks is more reliable than treating individual sweeps as independent.

If a susceptibility peak is represented by ten nearby temperatures generated from the same energy histogram, their errors are correlated. Temperature reweighting uses factors $e^{-(\beta' - \beta)E}$ and is dependable only where sampled energy distributions overlap. A smooth reweighted curve extending outside sampled energies can look precise while depending on a few rare configurations. Use independent anchor simulations or directly measured temperatures to verify both tails.

Inspect distributions as well as means. A first-order transition can also show a large susceptibility peak and bimodality, with leading peak scaling proportional to volume in many finite systems. A continuous transition's order distribution may itself be broad or even possess two lobes. Distinguish using latent heat or its size dependence, interface free-energy barriers, scaling of peak width and appropriate thermal observables. One finite-size double peak in magnetisation is not a universal transition-order test.

Check how a long but finite correlation length can arise without a positive-temperature transition in one dimension

An exact finite chain tests what a rapidly growing length means

A one-dimensional nearest-neighbour Ising chain in zero field has no positive-temperature bulk ordering transition, but at low temperature its correlation length grows extremely rapidly. This makes it a valuable counterexample when interpreting a finite magnet. The transfer matrix has entries $e^{Kss'}$ with $K = J/(k_B T)$. Its two eigenvalues are

$$\lambda_+ = 2\cosh K,$$
$$\lambda_- = 2\sinh K.$$

For a periodic chain of L spins,

$$Z_L = \lambda_+^L + \lambda_-^L.$$

Keeping only the larger eigenvalue is permissible in the thermodynamic free-energy density at fixed positive temperature but need not accurately represent a short cold chain.

In the infinite chain the zero-field correlation is $\langle s_0 s_r \rangle = (\tanh K)^r$, yielding $\xi = -1/\log(\tanh K)$ in lattice spacings. At low temperature $\xi \simeq e^{2K}/2$. A chain of length 100 can therefore be strongly correlated across its entire length around $K \simeq 2.65$, while the infinite chain still has zero spontaneous order at every finite K. Setting $\xi \sim L$ gives an apparent ordering scale $k_B T_L/J \simeq 2/\log(2L)$, which moves slowly toward zero as the sample is enlarged.

For a periodic finite chain there are two paths connecting site zero and site r. The exact correlation is $[q^r + q^{L-r}]/[1 + q^L]$, $q = \tanh K$. Fitting it to a single exponential out to separations near half the box would overestimate the length because the second periodic image contributes. Use the periodic expression explicitly or restrict fits to a window that is both larger than microscopic distances and smaller than half the sample. A high-quality apparent fit over a narrow middle interval can still mask the wrong boundary geometry.

Numerically simulate chains of lengths 32, 64 and 128 at a sequence of K. Compare the measured absolute magnetisation, full connected correlation and transfer-matrix predictions. If a simulation remains stuck in one sign sector its short-distance correlation can agree with theory while the time average of signed magnetisation does not. Increasing the number of Monte Carlo sweeps and increasing the physical chain length are separate limits. Following only one of them cannot establish a nonzero thermodynamic order parameter.

A response derivative offers an exponent without partition-function normalisation

Let U_L be a dimensionless shape ratio such as the Binder cumulant. If $U_L = \mathcal{U}(tL^{1/\nu}) + \cdots$, its slope with respect to reduced temperature at the critical point scales as $L^{1/\nu}$. A logarithmic slope ratio between two sizes can estimate the correlation-length exponent independently of the magnetisation or susceptibility exponents. Corrections to scaling still matter, especially if the sampled point is not near a fixed critical argument for both sizes.

An equilibrium derivative can also be computed without a finite temperature difference. For any state observable A with no explicit temperature dependence, differentiating its canonical mean gives

$$\partial_\beta \langle A \rangle = -\mathrm{cov}(A, E).$$

Apply this to the second and fourth moments entering a Binder ratio, then use the chain rule. The resulting slope uses energy–magnetisation correlations from one run. Taking two numerically close independent temperatures instead can amplify noise by division by their small separation. A covariance estimator avoids that particular subtraction but requires accurate sampling of joint tails.

Suppose measured absolute crossing slopes on sizes 24, 48 and 96 are 5.7, 10.9 and 21.5 per unit reduced temperature. Adjacent estimates are $1/\nu = \log(10.9/5.7)/\log 2 = 0.935$ and 0.980. They approach the expected value one in this benchmark. A single fit through all sizes would give a smaller formal statistical uncertainty but might obscure a systematic drift. Report adjacent effective exponents and test sensitivity to omitting the smallest size.

The numerical derivative must use the same temperature variable for each size. A slope against inverse temperature differs by $d\beta/dT$, a known factor near a common critical temperature; plotting against independently fitted centre temperatures can introduce an additional calibration dependence. A nonlinear temperature sensor or a slowly drifting coupling also changes the physical conversion from laboratory temperature to the simulated dimensionless control. Use an independently determined parameter scale if the goal is to compare experiment with an absolute transition temperature rather than just a scaling exponent.

Estimate an interface cost from a probability trough

At a first-order transition, a finite periodic sample may switch between two phases separated by an interfacial configuration. A histogram of an appropriate extensive observable can show peaks corresponding to nearly pure phases and a trough representing intermediate states. At equal bulk phase weight, a probability ratio gives a free-energy difference $\Delta F_L = -k_B T \log(P_{\mathrm{trough}}/P_{\mathrm{peak}})$, provided the same bin measure or a properly defined density is used. An exponentially small trough amplifies sampling errors, so the measured minimum is often only an upper bound on its probability.

In a slab geometry under periodic boundaries, an intermediate configuration may contain two approximately flat interfaces. If each interface costs γL^{d-1}, the leading barrier is $2\gamma L^{d-1}$, with subleading capillary fluctuations, geometry corrections and possible droplet-to-slab crossovers. A cube with open boundaries can instead permit interfaces of different number or shape. Counting two interfaces automatically for every order-parameter histogram would misestimate the inferred tension.

For an illustrative two-dimensional periodic square, suppose the measured barrier grows from $8k_BT$ at $L = 40a$ to $15k_BT$ at $L = 80a$, where a is lattice spacing. From the difference of the two barriers and the increase in total interface length $2(40a)$, an effective line tension is $0.0875\,k_BT/a$. Using the absolute smaller-size barrier divided by $2L$ gives $0.100\,k_BT/a$, revealing a finite-size offset. More sizes, an observed slab geometry and adequate trough sampling are needed before assigning a bulk interface tension.

This analysis should not be transplanted unchanged to a continuous critical distribution with two shoulders in finite volume. A critical finite-size order distribution can be non-Gaussian without an interfacial barrier growing as L^{d-1} at its critical point. Plot the barrier's size dependence and inspect actual spatial configurations. A peak-height ratio alone contains no direct information about how many interfaces or what type of correlations generated the intermediate states.

Estimate a ratio of moments without losing its covariance

The second and fourth magnetisation moments in a Binder cumulant usually come from the same record. Treating their errors as independent can grossly misestimate the uncertainty because the fourth moment tends to be large precisely in configurations contributing a large squared moment. Divide the record into blocks longer than its integrated autocorrelation time, recompute the ratio after deleting each block and estimate uncertainty using a block jackknife or use independent replicate simulations. Recomputing the nonlinear statistic inside each resample retains numerator–denominator covariance.

Histogram reweighting across nearby temperatures should likewise be recomputed within each block when extracting a crossing. Fitting smooth curves to the full sample and then adding independent random errors to individual temperature points ignores strong correlations induced by the shared configuration weights. Repeat the root-finding step for each independent or block-resampled pair of systems to obtain an uncertainty on the crossing temperature itself.

Check block-size convergence. If estimated errors continue growing as block duration is increased, slow modes may not be resolved or the record may have drifted. A huge number of recorded spin configurations cannot compensate for fewer than a handful of independent reversals when the statistic depends on those reversals. This issue is measurable: compare direct covariance estimates, block estimates and independent starts, and extend the run where those diagnostics disagree.

A block-average variance exposes a correlation length without fitting a tail

Divide a large zero-field spin image into square blocks of side b and record the average spin in each block. Its variance is the sum of all connected correlations within the block divided by the squared number of sites. For independent spins of variance one, it is exactly $1/b^2$ in two dimensions. If neighbouring spins are correlated over a length ξ much smaller than the block, the variance still falls approximately with inverse block area, but with a larger prefactor proportional to the integrated correlation. Below block sizes comparable to ξ, this averaging becomes much less effective.

At a critical point the connected correlation decays algebraically rather than exponentially over the accessible range. If in dimension d it behaves as $G(r) \propto r^{-(d-2+\eta)}$, integrating pair separations to the block scale gives $\mathrm{var}(m_b) \propto b^{-(d-2+\eta)}$ in the scaling regime. For the two-dimensional short-range Ising benchmark with $\eta = 1/4$, doubling the block side reduces the variance by only $2^{-1/4} = 0.841$, compared with 0.25 for independent spins. This is an observable distinction even when a full bulk thermodynamic limit is unavailable.

Block analysis of one image does not generate independent thermodynamic systems. Neighbouring blocks share correlated surroundings, overlapping blocks are statistically dependent, and every block inherits global constraints or drift in the original preparation. Increasing block size also reduces the available count sharply. Repeat across independent configurations, compare several whole-image sizes and study block-size windows well above a lattice spacing and well below the full image. An apparently accurate critical exponent from the final two largest blocks of one image is particularly unreliable.

A real magnetic image may measure a blurred field projection rather than the local spin itself. Spatial point spread creates detector-induced correlations at small separations. Determine it with a reference pattern and omit or model blocks comparable to the blur. Conversely spatially varying illumination adds long-range covariance unrelated to the sample. Flat-field calibration and sample rotations can separate these artefacts from material correlations. A second-moment structure-factor estimate and a block-variance estimate respond differently to these measurement effects, providing a useful cross-check.

This method also links microscopic and macroscopic noise. The scaled block variance in the large-block regime approaches the same integrated connected correlation controlling uniform susceptibility, with boundary terms that shrink as block dimensions exceed ξ. Comparing a weak-field response with block statistics therefore tests equilibrium fluctuation–response and spatial calibration simultaneously. If the discrepancy grows only as blocks approach the full system size, a fixed total magnetisation constraint or incomplete sampling of global sign reversals deserves examination.

For a concrete noise budget, consider a camera averaging adjacent pairs of independent sites before reporting them as displayed pixels. Those displayed pixels can have a positive nearest-neighbour correlation if their averaging footprints overlap, even when the original spins are completely uncorrelated. A measurement of the apparent short-range correlation length would then reflect the instrument kernel. Blocking on sizes much larger than the overlap can recover the correct large-scale inverse-area trend, while direct comparison of single-pixel variance with independent-spin theory fails. Calibrating only the detector's global gain cannot fix this spatial response. A measured point-spread map, an uncorrelated reference preparation and tests of blocks at several image resolutions can separate the resolution scale from the material's growing correlation length as temperature is varied.

Compare a correlation length with a finite specimen

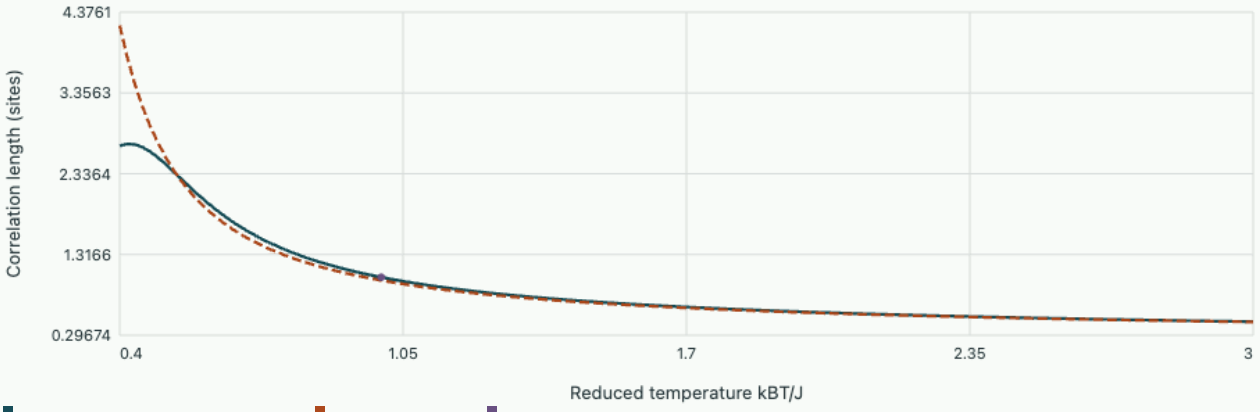
Determine how the subleading transfer-matrix eigenvalue controls the length seen in a colour measurement. A finite chain much shorter than this length can look ordered.

Open this experiment online

Try the calculation

Reduced temperature kBT/J **1**

Reduced colour field H/J **0.15**



Selected colour bias | Zero bias | Selected temperature

Leading eigenvalue 4.9881, second 1.8884, longest length 1.0295 sites.

Download the OneSimulation model

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

```
python spectra-and-orbits.py chain --temperature 1 --bias 0.15
```

Tutorial · Extract two exponent ratios from a size doubling

At a candidate critical temperature susceptibilities for sizes 32 and 64 are 110 and 360 in common units. Absolute magnetisations are 0.310 and 0.282. Estimate the corresponding exponent ratios and compare with the two-dimensional Ising benchmarks.

▼ **Use logarithmic ratios and inspect the small differences**

The susceptibility ratio gives

$$\gamma/\nu = \log(360/110)/\log 2 = 1.71.$$

For magnetisation $\beta_m/\nu = -\log(0.282/0.310)/\log 2 = 0.137$. These are near 1.75 and 0.125, but no error bars or corrections have been supplied. An additive susceptibility background or a slight critical-temperature offset can bias both effective exponents. Use several sizes and a dimensionless crossing diagnostic before assigning a universality class.

Tutorial · Determine whether a finite sample really switched phases

A zero-field record contains 20,000 sweeps of positive magnetisation near $+0.6$, followed by one reversal and 20,000 sweeps near -0.6. It yields a near-zero mean and large variance. Why does the nominal record length not establish a precise symmetric susceptibility? What extra records or controls would help?

▼ **Identify the slow event determining the uncertainty**

Within each sign sector there may be many locally decorrelated fluctuations, yet the relative weight of the two peaks is estimated from essentially one reversal. Treating 40,000 sweeps as independent underestimates the error of the peak weights and of statistics sensitive to global sign. Starting multiple independent runs with both signs can improve coverage, but deliberately balancing starting signs is not a measurement of true equilibrium weights if a weak residual field exists.

Record switching-time distributions over much longer runs or use an equilibrium-preserving algorithm that efficiently samples both sectors, while distinguishing computational updates from physical kinetics. Apply small positive and negative fields to test susceptibility and residual bias. For a within-phase susceptibility condition on a clearly defined phase and report the changed observable.

Tutorial · Explain a crossing that drifts with the boundary

Two groups study the same lattice Hamiltonian. One uses square periodic samples; the other uses twice-as-long rectangular samples with open ends. Their Binder curves cross near similar temperatures but at different cumulant values. Is this evidence against universal critical behaviour?

▼ **Separate universal exponents from finite-geometry functions**

Changing aspect ratio and boundary conditions changes the finite-size scaling function and hence the crossing value of a dimensionless ratio. Universal bulk exponent comparisons do not require identical Binder values across different finite geometries. Match boundary conditions, shape, operator and cumulant convention before comparing purported universal amplitude ratios.

Each group can still extrapolate its own crossing temperatures, examine moment slopes proportional to $L^{1/\nu}$ and test bulk exponent ratios with geometry fixed across its own size series. If one dimension is held fixed while the other grows, however, the family approaches a strip rather than a bulk two-dimensional limit. The thermodynamic sequence itself is an experimental variable, not just a plotting preference.

Cambridge Statistical Physics · phase transitions, criticality and scaling

How many lattice sites remember the colour of a neighbour?

Counting, correlations and thermal measurements

A chain of three-state spins gains energy when adjacent colours agree. A weak field favours one colour over the other two. Use a transfer matrix to calculate both free energy and the distance over which correlations survive. Recover the same numbers from OneSimulation eigenvalues without relying on symmetry reduction.

Find a local weight whose product counts each energy once

Each site has colour $a = 0, 1, 2$. A matching pair contributes energy $-J$, with $J > 0$, and a site of colour zero earns an additional $-H$. In a periodic ring each site belongs to two adjacent bonds. Give each of those bonds half the local field energy so the product of their weights counts each field term exactly once. Define $K = J/(k_B T)$ and $h = H/J$.

$$M_{ab} = \exp\left[K\left(\delta_{ab} + \frac{h}{2}(\delta_{a0} + \delta_{b0})\right)\right],$$

$$Z_N = \operatorname{tr}(M^N).$$

To see the trace, write the partition function as a sum over $a_1, \ldots, a_N$ of the product $M_{a_1 a_2} \cdots M_{a_N a_1}$. Summing an intermediate index is matrix multiplication; closing the initial and final index is the trace. This is exact even for a short periodic chain defined by that bond count. The transfer matrix contains no time step: multiplying it advances spatial position.

Compare the leading eigenvalue with the next one

Without bias the matrix has diagonal e^K and all other entries one. The uniform vector has eigenvalue $e^K + 2$. Every vector with components summing to zero has eigenvalue $e^K - 1$, twice degenerate. Thus $Z_N = (e^K + 2)^N + 2(e^K - 1)^N$. The thermodynamic free energy per site is $-k_B T \ln(e^K + 2)$. To measure a connected spatial correlation, insert a diagonal colour observable at sites separated by r transfer operations. The uniform contribution cancels from the connected part; each nonuniform mode contributes a factor $(\lambda_j/\lambda_1)^r$ when its overlap is nonzero. The longest correlation length in sites is

$$\xi^{-1} = \ln(\lambda_1/|\lambda_2|).$$

At $k_B T = J$ and zero bias, $\lambda_1 = 4.71828,\ \lambda_2 = 1.71828$ and $\xi = 0.9900$ sites. At low temperature e^K is large and $\xi \simeq e^K/3$. A finite chain can then appear ordered simply because it is shorter than its correlation length. There is no finite positive temperature at which these two one-dimensional eigenvalues cross.

Keep track of a mode that a chosen detector cannot see

At $h \neq 0$, colours one and two remain interchangeable. Their difference vector is an eigenmode of eigenvalue $e^{K} - 1$. In the orthonormal basis comprising colour zero and the sum of one and two, the remaining block is

$$\begin{pmatrix} e^{K(1+h)} & \sqrt{2}e^{Kh/2} \\ \sqrt{2}e^{Kh/2} & e^{K}+1 \end{pmatrix}.$$

Its eigenvalues follow from the trace and the square root of the squared diagonal difference plus four times the squared off-diagonal element. Sort all three eigenvalues. A detector sensitive only to whether the spin is zero is invariant under interchange of colours one and two. It has no overlap with their odd difference mode, even if that mode sets the mathematically longest length. A resolved measurement of colours one and two can see it. Quoting a correlation length from eigenvalues requires specifying which observation can actually couple to that eigenvector.

Change temperature and examine the spectral gap

Vary reduced temperature and reduced field H/J independently. The plot compares the longest colour-correlation length to the exact zero-field result. At a chosen point the downloadable model constructs the full three-by-three matrix and computes its modes using OneSimulation's contour eigensolver. It reports free energy per site and all three eigenvalues. In this small system the direct reduction above is available as an independent check. The same matrix construction generalises to more states even when no colour symmetry reduces the size.

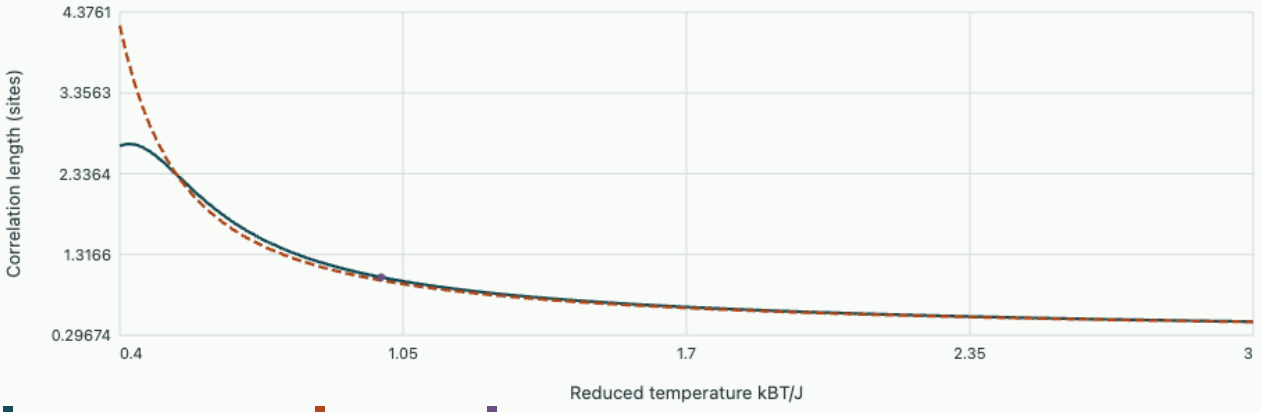
Compare two chains at the same temperature. Set the colour bias to zero and apply a positive bias of 0.30 J. Use the transfer-matrix eigenvalue ratio to explain the change in the longest correlation length.

Open this experiment online

Try the calculation

Reduced temperature kBT/J **1**

Reduced colour field H/J **0.15**



Selected colour bias | Zero bias | Selected temperature

Leading eigenvalue 4.9881, second 1.8884, longest length 1.0295 sites.

Download the OneSimulation model

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

```
python spectra-and-orbits.py chain --temperature 1 --bias 0.15
```

Tutorial

At $k_B T = J,\ h = 0$ what fraction of the asymptotic leading contribution to a connected correlation survives after three lattice sites?

▼ **Work through the solution**

The factor is $(1.71828/4.71828)^3 \simeq 0.0483$. The correlation observable sets its initial amplitude; the eigenvalue ratio sets this fraction.

Why does squaring a single bond weight to represent the effect of two bonds give the wrong two-step transition amplitude?

▼ **Work through the solution**

Two steps can pass through any of the three intermediate colours. Matrix multiplication sums all possibilities $M_{ab} M_{bc}$ over b. Squaring the single number M_{ac} loses those intermediate configurations.

Suppose only the leading eigenvalue is reliably resolved. Which thermodynamic-limit prediction is possible and which remains undetermined?

▼ **Work through the solution**

One can obtain free energy per site, and with its parameter dependence the derivatives giving energy or magnetisation. A connected spatial length requires a subleading eigenvalue and overlap of the chosen observable. A precise leading free energy alone cannot give that numerical length.

Compare this transfer calculation with the two-state Ising chain experiment.

Periodic potentials, interfaces and experimental probes

Follow a state across a boundary, a band gap or an avoided crossing. Use optical, tunnelling and calorimetric probes to distinguish an energy gap from a loss peak.

1. Scattering and tunnelling in one dimension
2. Transitions under a timed drive
3. Boundary conditions that carry a current around a ring
4. Adiabatic following and a finite sweep through an avoided crossing
5. When a periodic dielectric has no propagating Bloch wave
6. Measuring magnetic core loss by a hysteresis loop and by calorimetry
7. From a Coulomb orbit to the fine-structure multiplet
8. Two electrons, exchange splitting and the cost of localisation
9. Measure tunnelling from a four-site spectrum

Scattering and tunnelling in one dimension

Quantum mechanics

Match stationary waves across a potential discontinuity and use probability current to distinguish reflection, propagation and an evanescent field.

FROM FIRST PRINCIPLES

Obtain reflected and transmitted probabilities at a potential step

Suppose a constant-mass particle reaches a flat potential step from the left with energy above both sides. In a region of constant potential the stationary Schrödinger equation has waves $e^{\pm ikx}$ with $\hbar^2k^2/2m = E - V$. Set the incoming amplitude to one and use asymptotic waves $e^{ik_1x} + re^{-ik_1x}$ at left and te^{ik_2x} at right. Integration of the differential equation across the finite step makes the derivative continuous; the wave function itself is continuous as well:

$$1 + r = t,$$
$$k_1(1 - r) = k_2 t.$$

Solving the two linear conditions gives

$$r = (k_1 - k_2)/(k_1 + k_2),$$
$$t = 2k_1/(k_1 + k_2).$$

The current from the previous chapter for one travelling component is $\hbar k|A|^2/m$ in its propagation direction. Ratios to incoming current therefore yield

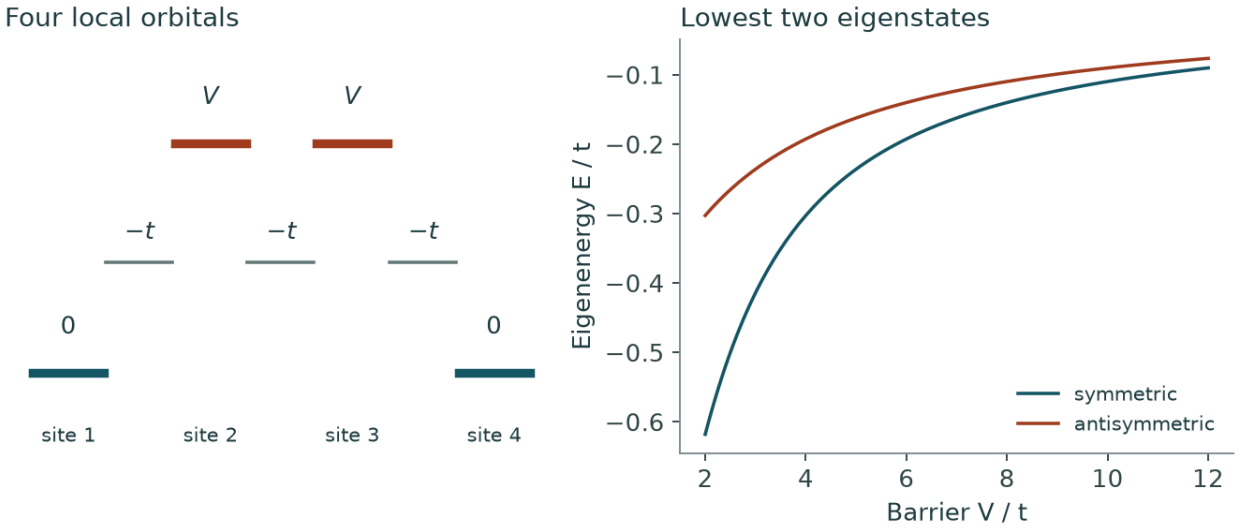
$$R = \frac{(k_1 - k_2)^2}{(k_1 + k_2)^2},$$
$$T = \frac{4k_1k_2}{(k_1 + k_2)^2},$$
$$R + T = 1.$$

Equal potentials give unity transmission. Near a right-hand threshold $k_2 \to 0$, a finite boundary amplitude can coexist with vanishing transmitted flux because the outgoing velocity also tends to zero.

Background · Probability current

$$j = (\hbar/m)\,\mathrm{Im}(\psi^*\partial_x\psi)$$

This expression assumes a scalar particle with no vector potential. Transmission probabilities require outgoing relative to incoming current, even when their wave numbers differ.



From a barrier drawing to a measurable splitting

Put the four on-site energies on the diagonal of a matrix and put –t between neighbours. An end-to-end transition visits both central sites. At large V it requires three hops and two off-resonant denominators, so its amplitude scales as t^3/V^2. The graph shows the two exact low energies whose difference tests that prediction.

Try it · Physics lab double well splitting

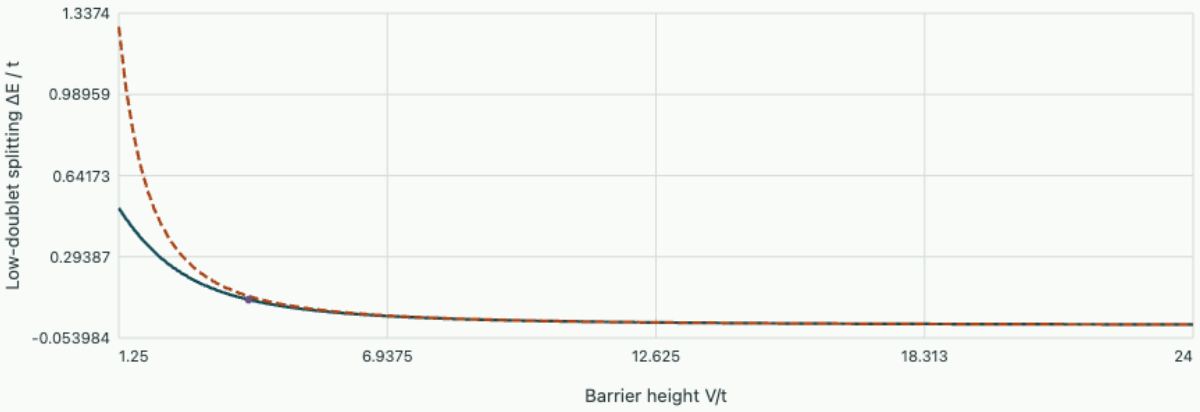
Explore the same physical relation with adjustable inputs. This independent parameter study has its own displayed reference settings; compare its assumptions and units before matching the static example.

Open this live calculation

Try the calculation

Middle site energy V (meV) 4

Hopping t (meV) 1



Exact parity sectors | Three-hop approximation | Selected barrier

Splitting 0.11019 meV. First transfer between dressed end states at 18.766 ps.

Download the OneSimulation model

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

```
python spectra-and-orbits.py wells --barrier 4 --hopping 1
```

DERIVATION AND APPLICATIONS

Match both faces of a finite barrier

A decaying wave inside a classically inaccessible interval does not alone determine its transmission. Boundary matching at the exit also contributes. Let $V(x) = V_0$ for $0 < x < L$ and zero outside, with a particle of constant mass and $0 < E < V_0$ incident from the left. Define $k = \sqrt{2mE}/\hbar$ and $\kappa = \sqrt{2m(V_0 - E)}/\hbar$. Set the phase reference at the exit so the outgoing wave is $te^{ik(x-L)}$. At $x = L$, continuity gives both the exit value t and its derivative ikt.

Inside, $\psi'' = \kappa^2\psi$. Propagate these exit data back to the entrance using the even and odd solutions:

$$\psi(0) = t[\cosh(\kappa L) - i(k/\kappa)\sinh(\kappa L)],$$
$$\psi'(0) = t[-\kappa\sinh(\kappa L) + ik\cosh(\kappa L)].$$

On the left write $e^{ikx} + re^{-ikx}$. Adding its value to its derivative divided by ik removes the reflected amplitude. Inserting the propagated data yields

$$2 = t\,[2\cosh(\kappa L) + i(\kappa/k - k/\kappa)\sinh(\kappa L)].$$

Both asymptotic waves propagate in the same material, so their velocity factors cancel in the current ratio. Take the squared magnitude and use $\cosh^2 u = 1 + \sinh^2 u$:

$$T = |t|^2 = \left[1 + \frac{(k^2+\kappa^2)^2}{4k^2\kappa^2}\sinh^2(\kappa L)\right]^{-1}.$$

A barrier of vanishing width transmits fully. An exponential estimate at large κL neglects the interface-dependent prefactor. Close to the top, the $\sinh$ factor and its apparently singular coefficient must instead be taken to the limit together.

WORKED EXAMPLE

Infer a width from a tunnelling measurement

Change the inputs and investigate this derivation.

At $E = V_0/2$, $k = \kappa$ and the result simplifies to $T = \operatorname{sech}^2(\kappa L)$. A device with $\kappa = 2.00$ nm⁻¹ has $T = 0.250$. Thus $\cosh(\kappa L) = 2$ and $L = \operatorname{arcosh}(2)/2.00 = 0.658$ nm. Doubling that width gives $\cosh(2\kappa L) = 2(2)^2 - 1 = 7$, hence

$$T = 1/49 = 0.0204.$$

SECOND APPLICATION

Approach the top from below

Change the inputs and investigate this derivation.

Let $E \to V_0$ at fixed L. With $\sinh(\kappa L) \simeq \kappa L$ the transmission becomes

$$T = [1 + (kL)^2/4]^{-1}.$$

For $k = 3.00$ nm⁻¹ and $L = 0.500$ nm this is 0.640. A vanishing inverse decay length therefore does not guarantee complete transmission. Exactly at the barrier top the internal solution becomes linear and still has to match the slopes on both sides.

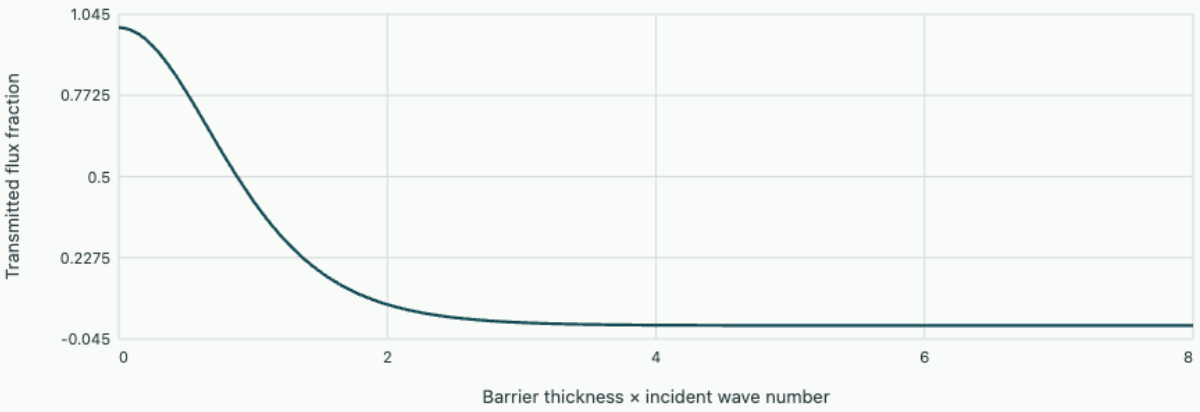
Try it · Barrier transmission

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

Barrier height / incident energy 2



Exact rectangular barrier

At threshold V0=E the limiting formula is 1/[1+(kL)²/4]. Below threshold repeated internal reflections give resonances; above it an evanescent region suppresses transmission.

Build a scattering state

Send a narrow band of particles from the left towards a region where the potential changes. Far from the obstacle a stationary energy component has incoming and outgoing travelling waves. Such plane waves are not normalisable on the whole line, but their flux ratios describe what happens to a sufficiently long incident packet whose momentum distribution is narrowly centred. Assume a particle of constant mass and a real, finite, time-independent potential.

The stationary Schrödinger equation is

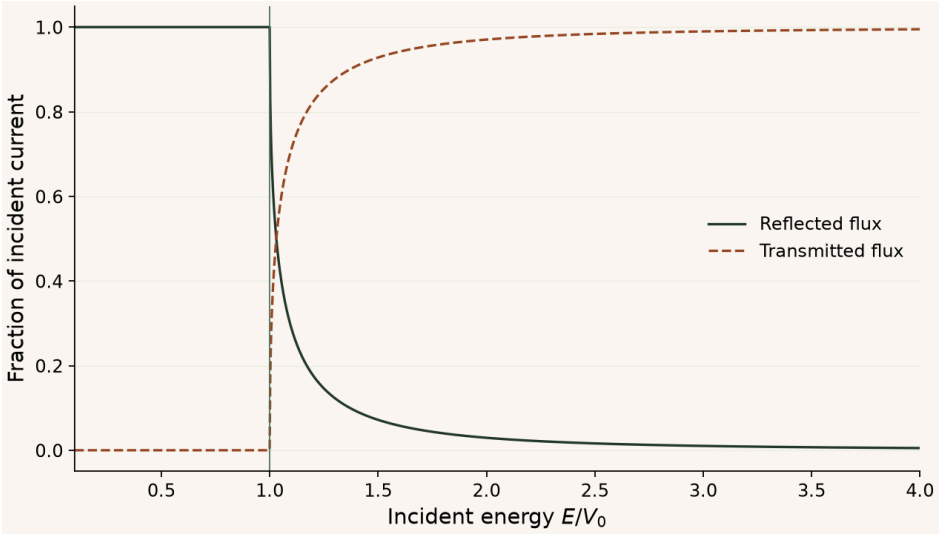
$$-\hbar^2\phi''/(2m) + V\phi = E\phi.$$

Integrating it over a shrinking interval around a finite step gives continuity of ϕ'; a finite jump in ϕ itself would introduce a more singular derivative, so ϕ is continuous as well. These matching laws change for an ideal delta potential, an infinite wall or an effective mass discontinuity. Stating the model at an interface matters as much as solving the constant-potential regions.

Partly reflect above a potential step

Take $V = 0$ for $x < 0$ and $V = V_0$ on the right. For $E > V_0$, write $\phi_L = e^{ikx} + re^{-ikx}$ and $\phi_R = te^{iqx}$, with $k = \sqrt{2mE}/\hbar$ and $q = \sqrt{2m(E - V_0)}/\hbar$. Matching gives $1 + r = t$ and $k(1 - r) = qt$. Thus $r = (k - q)/(k + q)$ and $t = 2k/(k + q)$. A wave can reflect from a change of wavelength even if a classical particle has enough energy to proceed.

A right-moving component Ae^{ikx} carries current $\hbar k|A|^2/m$. Consequently $R = |r|^2$ while $T = (q/k)|t|^2$, and direct substitution gives $R + T = 1$. For $E = 4V_0/3$, $q = k/2$, so $r = 1/3$, $t = 4/3$, $R = 1/9$ and $T = 8/9$. Although the transmitted wavefunction amplitude is larger than the incident one, its reduced group velocity keeps the transmitted probability flux below the incident flux.



For a semi-infinite potential step the flux fractions sum to unity at each energy. Partial reflection persists above the step; below threshold the decaying wave carries no net transmitted current.

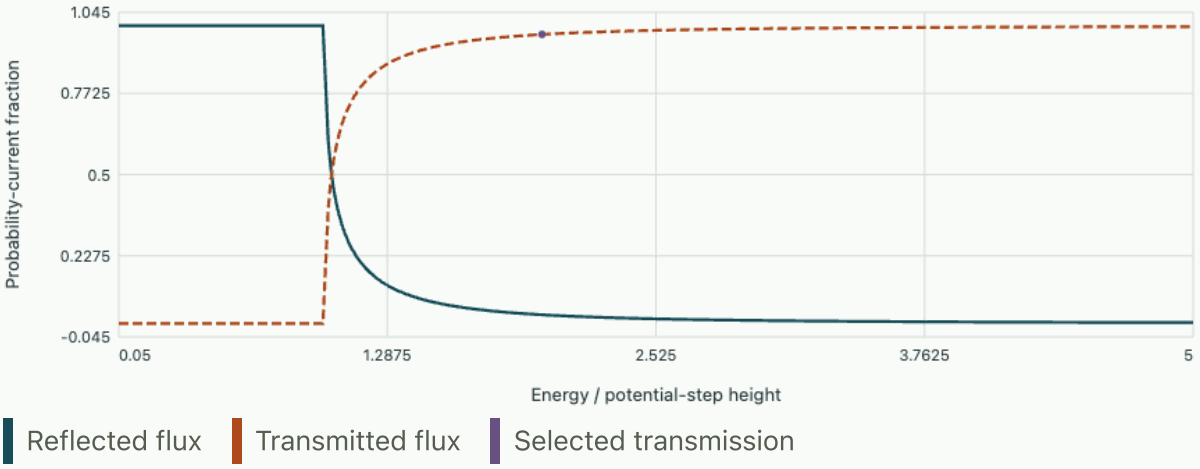
Try it · Step scattering

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 energy / step height 2



Reflected flux Transmitted flux Selected transmission

R + T = 1. At the selected energy T = 0.97056. The step extends to infinity; evanescent density below threshold carries no transmitted current.

Background · Why can a transmitted amplitude exceed one?

For incidence from the left, match the wavefunction and its derivative at a potential step with equal particle masses and positive real wavenumbers k_L, k_R. With unit incident amplitude the transmitted amplitude is $t = 2k_L/(k_L + k_R)$. A detector downstream counts probability crossing its position per unit time. The current in a right-travelling plane wave includes its group velocity.

$$T = \frac{k_R}{k_L}|t|^2,$$

$$R = \left|\frac{k_L - k_R}{k_L + k_R}\right|^2.$$

For $k_R = k_L/3$, the transmitted amplitude is 1.5, its intensity factor is 2.25, and its transmitted current fraction is 0.75. The reflected fraction supplies the other 0.25. The density increases because the transmitted particles propagate more slowly. Using flux-normalised channel states absorbs the velocity factor into the scattering amplitude.

Follow an evanescent field

If $0 < E < V_0$, the right-hand wave number becomes $i\kappa$ with $\kappa = \sqrt{2m(V_0 - E)}/\hbar$. The physical solution for a semi-infinite step decays as $e^{-\kappa x}$. This single exponential times a constant has zero current, since its spatial derivative has the same complex phase. It nevertheless has finite density a short distance into the forbidden region. Matching at the interface now gives a reflected wave of unit magnitude with an energy-dependent phase.

A finite barrier differs because an allowed region begins again at $x = L$. Both decaying and growing exponentials must be retained within the barrier, and their relative complex coefficient allows nonzero current. Match the wavefunction and derivative at both ends. For equal zero potentials on the outside the transmission probability reduces to

$$T = \left[1 + \frac{V_0^2 \sinh^2(\kappa L)}{4E(V_0 - E)}\right]^{-1}.$$

The large-width dependence comes from $\sinh(\kappa L) \sim e^{\kappa L}/2$. At fixed barrier height it is the exponential sensitivity to width that makes tunnelling useful as a distance probe. A single semi-infinite exponential fitted to density cannot by itself supply the finite-barrier prefactor.

Estimate the effect of a subnanometre change

For an electron with $E = 1.0$ eV and $V_0 = 2.0$ eV, $\kappa \simeq 5.12$ nm⁻¹ and the prefactor of $\sinh^2$ is exactly one. At width 0.50 nm,

$$T = 1/\cosh^2(2.56) \simeq 0.0236.$$

At 1.00 nm, $T \simeq 1.43 \times 10^{-4}$. Doubling this subnanometre gap suppresses transmitted current by roughly a factor of 165 for equal incident current.

The calculation assumes a sharp rectangular barrier, one occupied incident energy and elastic coherent scattering. A measured electric tunnel current generally combines many available occupied and empty states across an energy window. A current ratio is therefore most directly compared with the single-energy exponential when barrier shape and occupation stay sufficiently similar across the distance change.

A more detailed test uses scattering from opposite sides of a real, static barrier at the same energy. Time-reversal symmetry implies equal transmission probabilities for a single channel, even when the spatial barrier profile is asymmetric. The two reflection phases can differ. If the exterior potentials differ, the transmitted amplitude ratios from the two sides need not coincide because their travelling states carry different current per unit amplitude. Normalizing the prepared and detected states to unit flux removes that ambiguity. Checking incident current against the sum of outgoing currents separately tests whether the model conserves probability.

Solve for scattering and binding at one narrow defect

A point interaction illustrates how the same Hamiltonian predicts both scattering and bound states. On an infinite line take $V(x) = g\delta(x)$, with g having units of energy times length. This limiting model replaces a narrow feature when the wavelength of interest greatly exceeds its microscopic width. The wavefunction remains continuous, but integrating the stationary Schrödinger equation from $-\eta$ to $+\eta$ before taking $\eta \to 0$ gives the derivative jump

$$\psi'(0^+) - \psi'(0^-) = \frac{2mg}{\hbar^2}\psi(0).$$

For a particle incident from the left, write $\psi = e^{ikx} + re^{-ikx}$ on the left and te^{ikx} on the right. Continuity gives $t = 1 + r$. Inserting both derivatives into the jump condition and defining $\xi = mg/(\hbar^2 k)$ gives

$$t = 1/(1 + i\xi),$$
$$r = -i\xi/(1 + i\xi).$$

The external speeds are equal on both sides, so $T = 1/(1 + \xi^2)$ and $R = \xi^2/(1 + \xi^2)$. The sum is one for real g. Transmission probability depends only on the magnitude of the interaction, but the complex phase of the transmitted wave distinguishes its sign.

For negative energies the exterior solution must decay as $e^{-\kappa|x|}$. Its derivatives jump by $-2\kappa\psi(0)$, giving $\kappa = -mg/\hbar^2$. A normalizable state requires $\kappa > 0$, hence only an attractive defect binds. The bound energy is $-\hbar^2\kappa^2/(2m)$. A one-dimensional narrow attractive defect can therefore support binding even when a scattering-probability measurement alone cannot distinguish its sign.

Compare two defects with the same transmission

Use electron mass and take $|g| = 0.0762$ eV nm. Since $\hbar^2/(2m_e) = 0.0381$ eV nm², $m_e|g|/\hbar^2 = 1.00$ nm⁻¹. An incoming wave with $k = 2.00$ nm⁻¹ has energy 0.152 eV and $|\xi| = 0.500$. Both signs of g transmit 0.800 and reflect 0.200 of the incident probability current. The transmission phases, however, are $\mp \arctan(0.500) = \mp 26.6^\circ$, a difference observable against a reference path.

For the attractive sign the decay length of the bound state is $1/\kappa = 1.00$ nm and its energy is -0.0381 eV. The repulsive sign has no such state. At high incident k both become nearly transparent; at small k both produce nearly complete reflection for fixed nonzero strength. The order of these limits matters: taking g to zero first gives perfect transmission at any finite incident wavenumber.

An attractive defect with the inferred decay constant has normalized bound state $\psi = \sqrt{\kappa}\, e^{-\kappa|x|}$. The probability outside distance b from the defect is $e^{-2\kappa b}$ after integrating both tails. With $b = 1.00$ nm in the example this gives 0.135, despite each wave amplitude at that distance being $1/e$ of its central value. A numerical calculation in a box can use this estimate to choose outer boundaries. Placing artificial zero-value boundaries only one decay length away would remove substantial tail probability and shift the apparent bound energy upward.

Increasing the box symmetrically and tracking the bound energy is more informative than merely refining the spatial mesh in the original box. Mesh convergence at fixed outer boundaries can certify a well-resolved solution of the wrong finite-domain physical problem.

Transitions under a timed drive

Quantum mechanics

Derive the line shape of a weak finite-duration drive and distinguish Fourier width from a system's intrinsic lifetime.

FROM FIRST PRINCIPLES

How the time profile of a drive becomes a transition amplitude

Assume an initially occupied eigenstate $|i\rangle$ of a time-independent Hamiltonian H_0, perturbed by a weak time-dependent operator $V(t)$. Expand the state in its unperturbed energy eigenvectors and factor out their known phases. Inserting the expansion into the time-dependent Schrödinger equation cancels the unperturbed energy terms:

$$|\psi(t)\rangle = \sum_n c_n(t) e^{-iE_n t/\hbar}|n\rangle,$$

$$i\hbar \dot{c}_f = \sum_n V_{fn}(t) e^{i(E_f - E_n)t/\hbar} c_n.$$

At first order the right side uses the initial amplitudes $c_i = 1, c_{n\neq i} = 0$. Integrating from zero through pulse duration T gives

$$c_f^{(1)}(T) = -\frac{i}{\hbar}\int_0^T V_{fi}(t) e^{i\omega_{fi}t} dt,$$

$$\omega_{fi} = (E_f - E_i)/\hbar.$$

A square nearly resonant sinusoidal drive has a slowly varying integrand proportional to $e^{i\Delta t}$ plus a rapidly oscillating component. Keeping the slow term produces amplitude proportional to $T\,\mathrm{sinc}(\Delta T/2)$ and probability proportional to its square. The approximation requires small population transfer; a longer resonant drive eventually requires coupled state amplitudes.

DERIVATION AND APPLICATIONS

Why a strong resonant pulse stops increasing the transition probability

Perturbation theory predicts a probability growing as pulse duration squared on resonance. To follow the subsequent return of population, keep both amplitudes coupled. Consider an isolated pair of levels driven by a monochromatic field with a nearly resonant frequency. Use the rotating-wave approximation, justified when the detuning and coupling are both small compared with the carrier frequency. Let Δ be the angular detuning and Ω the resonant angular Rabi frequency. During a square pulse the rotating-frame Hamiltonian, up to an identity term, is

$$H_R = \tfrac{\hbar}{2}(\Omega\sigma_x + \Delta\sigma_z),$$
$$H_R^2 = \tfrac{\hbar^2}{4}(\Omega^2 + \Delta^2)I.$$

The square follows because the two Pauli matrices anticommute. In the exponential propagator the even powers of H_R therefore multiply the identity and the odd powers multiply H_R. Sum the two Taylor series with $\Omega_e = \sqrt{\Omega^2 + \Delta^2}$:

$$U(t) = I\cos(\Omega_e t/2) - i\,\frac{\Omega\sigma_x + \Delta\sigma_z}{\Omega_e}\sin(\Omega_e t/2).$$

Starting in the lower bare state, only σ_x contributes to the upper-state amplitude. Its squared magnitude is

$$P_e(t) = \frac{\Omega^2}{\Omega^2 + \Delta^2}\sin^2\left(\frac{\sqrt{\Omega^2 + \Delta^2}\,t}{2}\right).$$

On resonance a pulse area $\Omega t = \pi$ transfers the entire population; twice that duration returns it to the lower state. At short times the leading term is $\Omega^2 t^2/4$, recovering perturbation theory. Detuning both speeds up the oscillation and reduces its attainable maximum. These two signatures help distinguish mistuning from a change in drive strength.

WORKED EXAMPLE

Design a population inversion pulse

Change the inputs and investigate this derivation.

If $\Omega/(2\pi) = 2.00$ MHz, the resonant inversion time is $\pi/\Omega = 0.250$ microseconds. A 0.125 microsecond pulse gives $P_e = 0.500$. Holding the drive on for 0.500 microseconds instead gives zero in the ideal model. The source supplies and subsequently receives energy during a reversible oscillation.

SECOND APPLICATION

Infer the mistuning from an incomplete oscillation

Change the inputs and investigate this derivation.

Suppose the same drive produces maxima of 0.640 and full population oscillations repeat every 0.400 microseconds. The period gives

$$\Omega_e/(2\pi) = 2.50\,\text{MHz}.$$

The maximum gives $\Omega/\Omega_e = \sqrt{0.640} = 0.800$, recovering the calibrated 2.00 MHz drive. The detuning magnitude is $\sqrt{2.50^2 - 2.00^2} = 1.50$ MHz in ordinary-frequency units. A population trace alone leaves its sign undetermined.

How a time-dependent perturbation transfers population

For a time-independent Hamiltonian H_0 the energy components of a state accumulate phases without changing populations. Let a weak applied perturbation $V(t)$ act for a known interval. Prepare the unperturbed eigenstate $|i\rangle$ at $t = 0$ and write $|\psi(t)\rangle = \sum_n c_n(t) e^{-iE_n t/\hbar}|n\rangle$. The interaction-picture amplitudes isolate changes beyond the free phases. Substitution into Schrödinger's equation and projection onto $\langle f|$ give

$$i\hbar \dot{c}_f = \sum_n V_{fn}(t) e^{i\omega_{fn} t} c_n,$$

$$\omega_{fn} = (E_f - E_n)/\hbar.$$

At first order retain $c_i = 1$ on the right-hand side. For $f \neq i$ the resulting amplitude is the Fourier component of the perturbation at the transition frequency,

$$c_f(T) = -(i/\hbar) \int_0^T V_{fi}(t) e^{i\omega_{fi} t}\, dt.$$

An energetically resonant drive still fails to populate a state if the coupling matrix element vanishes by symmetry. The operator describing the drive is as important as its frequency.

A rectangular pulse

Let $V_{fi}(t) = V_{fi}\cos(\omega t)$ for $0 < t < T$, with an upward transition so $\omega_{fi} > 0$. Splitting the cosine into exponentials gives frequencies $\omega_{fi} - \omega$ and $\omega_{fi} + \omega$. Near resonance, for a pulse lasting many carrier cycles, the rapidly oscillating sum-frequency term contributes little. With detuning $\Delta = \omega_{fi} - \omega$ the remaining integral gives

$$P_{i\to f} \simeq \frac{|V_{fi}|^2 T^2}{4\hbar^2} \left[\frac{\sin(\Delta T/2)}{\Delta T/2}\right]^2.$$

At exact resonance all contributions build with the same phase, so for a weak short pulse probability grows quadratically in time. The first zeros occur at $|\Delta| = 2\pi/T$, or ordinary-frequency detuning $1/T$. The line shape measures coherent addition over the finite observation time. It can be broad even when the isolated energy levels have negligible natural width.

Calibrate a pulse

Take a 2.00 ms rectangular pulse with $|V_{fi}|/\hbar = 200$ rad/s. The resonant transition probability is

$$(200 \times 0.002/2)^2 = 0.0400.$$

At 250 Hz detuning the sinc argument is $\pi/2$, leaving probability 0.0162. The first zeros lie 500 Hz on either side. Doubling pulse duration at the same coupling predicts four times the peak probability and half

the first-zero spacing within the weak-drive approximation. A measurement can test both the area dependence and frequency scaling.

Once the predicted population is no longer small, depletion of the initial level removes the first-order approximation. For an isolated resonantly driven two-state system the rotating-wave solution is $\sin^2(|V_{fi}|T/(2\hbar))$, describing coherent population transfer and return. Lifetime decay or field fluctuations introduce still other envelopes. Varying pulse duration at fixed drive strength distinguishes finite-time broadening from linewidths that persist as the pulse is lengthened.

The pulse envelope provides an additional independent control. Replacing the sudden rectangular turn-on and turn-off with smooth edges reduces spectral sidelobes because the Fourier transform of the smoother envelope falls faster away from resonance. Peak transfer at resonance, in the weak-drive regime, depends on the time integral of the drive envelope. Two pulses of different shapes can therefore have the same low-power resonant transition probability yet excite a nearby unwanted level with different probabilities. Specifying duration alone would be insufficient to reproduce such a measurement.

From a finite pulse to Fermi’s golden rule

An isolated resonant final state initially collects amplitude proportional to time. A dense set of final states can instead produce a nearly constant total transition rate. To see the change, let a weak drive connect one initial state to levels around energy $E_\star = E_i + \hbar\omega$. Denote the matrix element of the resonant Fourier component by M. For a cosine drive it contains the factor of one half. First-order time integration over a square pulse gives a phase times $(Mt/\hbar)\,\mathrm{sinc}[(E-E_\star)t/(2\hbar)]$, where $\mathrm{sinc}(u)=\sin(u)/u$.

The different final states are orthogonal, so their probabilities add. If the density $\rho(E)$ counts accessible final states per unit energy, and ρ and M vary slowly over the line width,

$$P(t) \simeq \frac{|M|^2 t^2}{\hbar^2} \int \rho(E)\,\mathrm{sinc}^2\!\left(\frac{(E-E_\star)t}{2\hbar}\right) dE.$$

Set $u=(E-E_\star)t/(2\hbar)$. Away from a band edge one can extend the integral across the line and use $\int_{-\infty}^{\infty}\mathrm{sinc}^2 u\,du=\pi$. One power of time cancels against the narrowing energy interval, leaving $P(t)\simeq \Gamma t$ with $\Gamma = 2\pi|M|^2\rho(E_\star)/\hbar$. This intermediate-time result assumes many final levels within the line and little depletion of the initial state. For a small isolated system the long-time dynamics can resolve individual levels again.

As a numerical example, use $M=0.500\,\mu\mathrm{eV}=0.000500\,\mathrm{meV}$, $\rho=1000\,\mathrm{meV}^{-1}$ and $\hbar=0.6582\,\mathrm{meV\,ps}$. The rate is $0.00239\,\mathrm{ps}^{-1}$, or a transition probability of 0.0239 after $10.0\,\mathrm{ps}$. The first zero lies $2\pi\hbar/t=0.414\,\mathrm{meV}$ from resonance, spanning about 414 level spacings on each side for a locally uniform density. The small transition probability and dense spectrum are distinct checks on the approximation.

Separate phase accumulation from the drive pulses

A single spectroscopy pulse ties spectral resolution to its own duration. The Ramsey method separates two short coherent pulses by an interval of free evolution. Work with two isolated states $|g\rangle$, $|e\rangle$ in the rotating frame of a stable reference. Approximate each pulse as short compared with the free interval and strong enough to rotate the Bloch vector by $\pi/2$ about the same transverse axis. Ignore relaxation and pulse detuning during the rotations. Starting in $|g\rangle$, the first pulse prepares $(|g\rangle - i|e\rangle)/\sqrt{2}$ for the chosen pulse phase.

Over free interval T the excited component gains relative phase $e^{-i\Delta T}$, with $\Delta = \omega_{eg} - \omega_{\mathrm{ref}}$. The second pulse mixes both paths to $|e\rangle$. Using the rotation matrix $R_x(\pi/2) = (I - i\sigma_x)/\sqrt{2}$ gives

$$c_e = -\frac{i}{2}(1 + e^{-i\Delta T}),$$

$$P_e = \frac{1 + \cos(\Delta T)}{2}.$$

A controllable phase on the second pulse shifts these fringes; on exact resonance two aligned pulses complete a π rotation and transfer all population. For a statistical mixture with the same half populations after pulse one, the two contributions to the final probability have no mutual phase and give one half regardless of delay. Thus the separated-pulse signal measures phase retained between pulses, even when no energy transfer takes place during most of the sequence.

Measure a small reference-frequency error

Take a free interval of 2.00 ms. Adjacent maxima occur at ordinary detunings separated by $1/T = 500$ Hz. The first minimum on either side occurs at 250 Hz. The separation of the two half-height points immediately surrounding resonance is

$$1/(2T) = 250\,\mathrm{Hz}.$$

These three figures refer to different linewidth or spacing conventions and should be identified before comparing experimental reports.

Set the second pulse phase so that the calibrated signal near zero detuning is $P_e = [1 + \sin(2\pi\delta\nu T)]/2$. A measured $P_e = 0.600$ gives $\delta\nu = \arcsin(0.200)/(2\pi T) = 16.0$ Hz on the branch nearest the reference. This setting is more sensitive to small signed frequency errors than the top of a fringe, where the first derivative vanishes. A first scan or a shorter free interval is needed to identify which repeated fringe is being tracked.

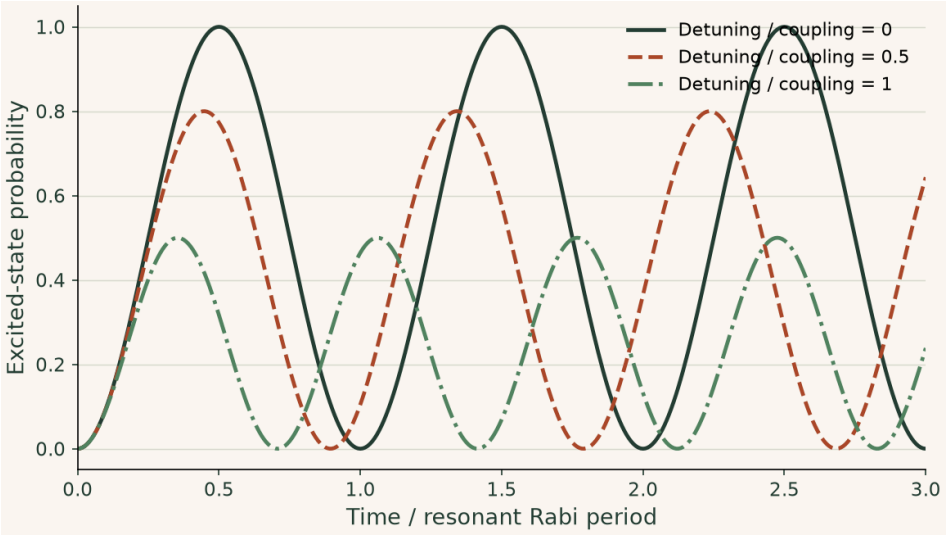
If the transverse coherence decays during free evolution, observed fringe amplitude is multiplied by the coherence factor, often approximated by e^{-T/T_2}. Extending the delay narrows fringe spacing but eventually reduces contrast. Measuring two phase settings in quadrature can separate an actual phase shift from a loss of visibility due to decoherence or pulse-area errors.

The frequency sensitivity per detected particle can also be estimated. At half population an ensemble of N independent measurements has standard error $1/(2\sqrt{N})$ in the estimated excited fraction. On the

central quadrature slope $dP_e/d\delta\nu = \pi T$, giving frequency uncertainty approximately $1/(2\pi T\sqrt{N})$ for ideal contrast. With the 2.00 ms delay and ten thousand independent detections this is about 0.80 Hz for a single ensemble. Imperfect contrast reduces the slope, and an unknown phase offset adds a systematic error that collecting more particles alone cannot remove.

Separate detuning from shortened observation time

Start a two-level system in its ground state, apply a monochromatic drive of resonant Rabi frequency Ω, and neglect relaxation. In the rotating-wave approximation detuning δ changes the oscillation frequency to $\sqrt{\Omega^2 + \delta^2}$ and reduces its maximum transfer.



At detuning equal to the resonant Rabi frequency the largest excited population is one half and its oscillation is √2 times faster. Sampling only at the nominal resonant π-pulse time confounds both effects.

Try it · Rabi

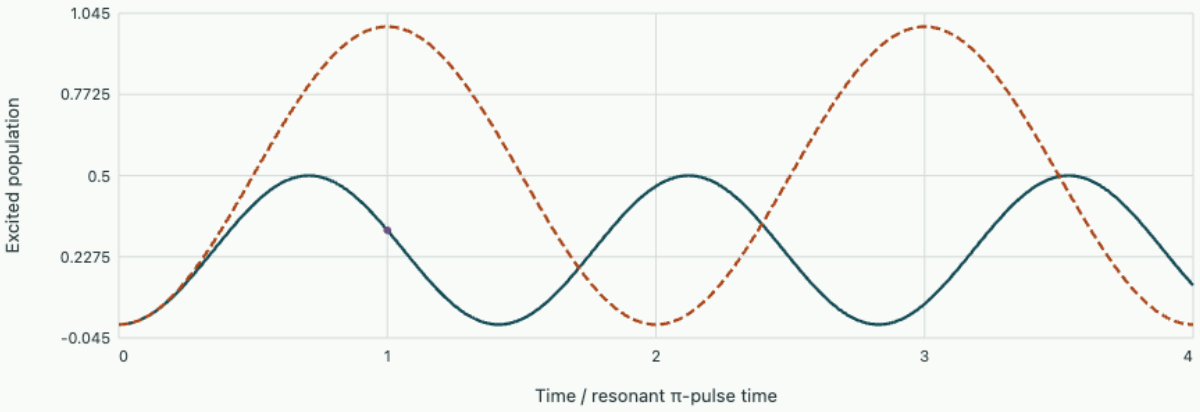
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Detuning / resonant Rabi frequency **1**

Pulse length in resonant π times **1**



Maximum excited population 0.5; pulse-end population 0.31656. Rotating-wave, coherent two-level model.

▶ **Read values from the plot**

Boundary conditions that carry a current around a ring

Quantum mechanics

An electron on an ideal narrow ring can sustain current in a stationary state. Work out how the topology of its configuration space and a magnetic flux through the hole select the allowed currents.

Recover the allowed state space from a boundary term

Wrap a coordinate s of length L around the ring. The points $s = 0$ and $s = L$ label the same physical place. For a uniform one-dimensional kinetic operator, integration by parts shows that the difference between $\langle \phi, H\psi \rangle$ and $\langle H\phi, \psi \rangle$ is proportional to

$$[\, \phi^* \partial_s \psi - (\partial_s \phi)^* \psi \,]_0^L .$$

Periodic values and periodic first derivatives make this expression vanish for any two admissible states. They also make the probability current continuous at the join. A plane wave $e^{iks}/\sqrt{L}$ then requires $kL = 2\pi n$ with integer n. The labels n and $-n$ have equal energy at zero flux but circulate oppositely. A real combination of the pair has zero net current. An energy measurement on its own does not tell us which combination was prepared.

Put flux through the hole

For a particle of charge q, the kinetic momentum along the ring is $-i\hbar\partial_s - qA_s$. Choose a constant tangential vector potential $A_s = \Phi/L$ giving the required circulation $\oint A_s ds = \Phi$. Define signed flux $\varphi = q\Phi/h$. Periodic plane waves remain eigenfunctions, while their kinetic momenta and energies now change:

$$p_{\text{kin}} = \frac{h}{L}(n - \varphi),$$

$$E_n = E_L (n - \varphi)^2,$$

$$E_L = \frac{h^2}{2mL^2}.$$

A gauge transformation can eliminate the constant A_s locally. It multiplies a wavefunction by a phase whose value after one circuit differs from its initial value by $e^{-2\pi i\varphi}$ in this gauge. The transformed wavefunction must carry that twist at the join. Imposing both zero vector potential and unchanged periodicity would discard the physical flux dependence. Increasing φ by one reorders the spectrum by a shift of the integer labels.

Work out what a current measurement sees

One particle of velocity $v = p_{\text{kin}}/m$ goes around the ring in a time $L/|v|$. The signed charge passing a fixed section per unit time is $I = qv/L$. Differentiating E_n at fixed n gives the same result:

$$I_n = \frac{qh}{mL^2}(n-\varphi) = -\frac{\partial E_n}{\partial \Phi}.$$

Take $\varphi = 0.30$. In units of E_L, the $n = 0$ energy is 0.09 and the $n = 1$ energy is 0.49. The currents in units of $qh/(mL^2)$ are −0.30 and +0.70 respectively. A probe that selects the lower-energy branch here sees a circulating state even though its probability density is uniform around the ring.

At $\varphi = 0.50$ the two energies agree at $0.25E_L$. Following a given eigenstate through this crossing and allowing a system to relax into its lowest energy at every flux describe different experiments. An ideal isolated perfectly uniform ring conserves its angular momentum label. Weak imperfections or a reservoir can change the populations and modify the behaviour near degeneracy.

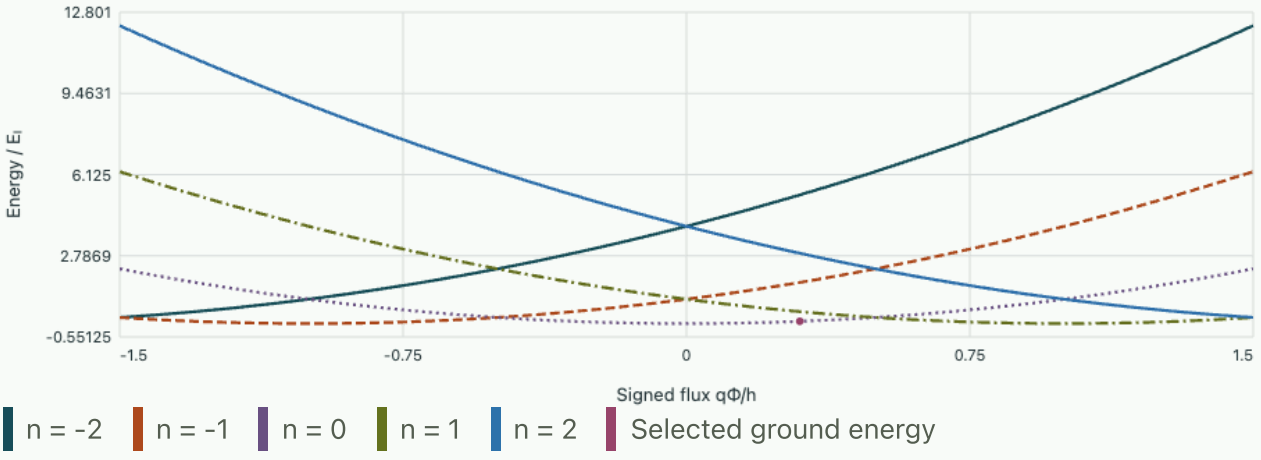
Inspect the neighbouring branches

Change the signed flux continuously. Compare the slope of the branch containing $n = 0$ with its current. Predict where the lowest pair changes ordering and check that the complete set repeats with unit flux.

Open this experiment online

Try the calculation

Signed flux qΦ/h **0.3**



Lowest n: 0. Ground energy 0.09 E_I. The n = 0 current in units of qh/(mL²) is -0.3.

Tutorial on a ring of another size

Double the circumference at fixed mass and flux while preparing the same integer state. Predict the changes of its wave number, energy and current.

▼ **Work through the solution**

The canonical and kinetic wave numbers each halve because the integer label and signed flux are fixed. E_L scales as L^{-2}, and so does the current unit $qh/(mL^2)$. Energy and current both fall to one quarter. The flux period $h/|q|$ stays unchanged.

A ring with fixed flux is enlarged to twice its original circumference without changing the particle mass. How do energy gaps and branch currents change? Does the location of a crossing in flux move?

▼ Work through the solution

$E_L = h^2/(2mL^2)$ falls by a factor of four. The current scale $qh/(mL^2)$ falls by the same factor. The signed flux variable contains no circumference when total flux is held fixed, so crossings stay at half-integer flux. Holding the magnetic field fixed as ring area grows would instead change the enclosed flux.

At $\varphi = 0.50$, prepare an incoherent mixture with equal populations on $n = 0$ and $n = 1$. Calculate its current and energy. Can observing zero average current distinguish this mixture from zero flux?

▼ Work through the solution

The currents are opposite, each of magnitude half the current unit, and the mean vanishes. Both energies are $0.25E_L$. A zero-flux ground state also has zero current, but its energy is zero in the same calibration. Spectroscopy or a change of population separates the two cases.

Further reading · Tong, Particles in Magnetic Fields, gauge transformations and Aharonov–Bohm effect · Fließbach, Quantenmechanik, Eigenwerte und Eigenfunktionen.

Adiabatic following and a finite sweep through an avoided crossing

Quantum mechanics

Two unperturbed energies cross as a control rises. A nonzero coupling opens a gap. Whether a prepared particle follows the lower energy through the crossing depends on how quickly the eigenvectors rotate compared with the dynamical phase they accumulate.

Find the gap and the rotation rate

Let $\hbar = 1$ temporarily, with Hamiltonian $H(t) = \begin{pmatrix} vt/2 & g \\ g & -vt/2 \end{pmatrix}$, $v > 0$, $g > 0$. The fixed basis is called diabatic because each basis vector keeps its identity as the detuning vt changes sign. Squaring the Hamiltonian gives $[(vt/2)^2 + g^2]I$; its energies are $\pm h(t)$ with $h = \sqrt{(vt/2)^2 + g^2}$. The minimum gap is $2g$. The corresponding Bloch axis points along $(g, 0, vt/2)$; write its polar angle as $\theta = \operatorname{atan2}(2g, vt)$.

Differentiating the angle gives

$$\dot{\theta} = -2gv/[(vt)^2 + 4g^2].$$

A spinor aligned with this axis rotates through half its geometric angle in Hilbert space. The magnitude of the off-diagonal derivative coupling is $|\dot{\theta}|/2$, largest at the crossing where it equals $v/(4g)$. Comparison with the minimum gap frequency $2g/\hbar$ gives a local adiabatic ratio $\hbar v/(8g^2)$. A slow sweep or large coupling suppresses transitions between instantaneous energy states. Following the lower adiabatic state exchanges the two fixed basis labels across the crossing.

Derive the fast-sweep scale independently

In the interaction picture of the diagonal part, the energy difference vt gives an accumulated relative phase $vt^2/(2\hbar)$. First-order perturbation theory gives a transition amplitude proportional to $(g/\hbar)\int_{-\infty}^{\infty} e^{ivt^2/(2\hbar)}dt$. The magnitude of this Fresnel integral is $\sqrt{2\pi\hbar/v}$, obtained by continuing the convergent Gaussian integral from a positive real damping parameter. The leading probability to change fixed basis states is therefore $2\pi g^2/(\hbar v)$, valid only when small. The exact infinite linear-sweep Landau–Zener solution resums this dependence: the probability of retaining the initial diabatic label is $\exp[-2\pi g^2/(\hbar v)]$. It refers to endpoints asymptotically far from the avoided crossing with constant coupling and no decoherence.

For $g = 0.5$ and $v = 1$ in units where $\hbar = 1$, the infinite-sweep retention probability is

$$e^{-\pi/2} = 0.2079.$$

Halving v squares it to 0.0432. The slow sweep therefore gives greater transfer between fixed basis states while preserving the lower-energy identity more faithfully. This distinction prevents reversing the interpretation of a measured probability.

Evolve a finite preparation numerically

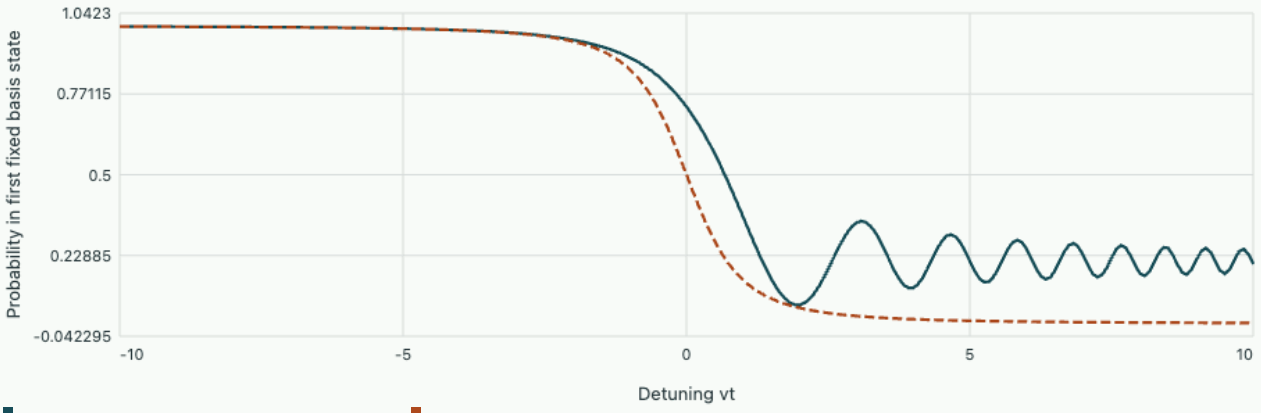
Prepare the instantaneous ground state at detuning $vt = -10$ and finish at $+10$. Over each short step freeze H at its midpoint. Since $H^2 = h^2 I$, the constant-step propagator is exactly $\cos(h\Delta t)I - i\sin(h\Delta t)H/h$ in the chosen units. The product is unitary up to roundoff; finite time resolution still changes accumulated phases. Compare the population of the original fixed basis state with its population in the instantaneous ground state. Their separation through the crossing directly shows where ideal adiabatic following fails. The calculated final population can differ from the infinite-sweep expression because of endpoint mixing and phases.

Open this experiment online

Try the calculation

Coupling g ($\hbar$ = 1) **0.5**

Sweep rate v ($\hbar$ = 1) **1**



Final fixed-basis probability 0.2005, asymptotic Landau–Zener value 0.2079. Propagated state norm squared 1.

Tutorial on measurement choices

The coupling doubles while all else is fixed. How must the sweep rate change to leave the ideal infinite-sweep probability unchanged? Does a sweep over fixed detuning endpoints remain exactly equivalent?

▼ **Work through the solution**

Keep g^2/v fixed by multiplying v by four. The infinite-sweep exponent is unchanged. At fixed endpoints the ratios $|vt|/g$ at preparation and readout are now smaller, so the finite endpoints are less asymptotic. Their state mixing can change the observed result.

At zero detuning an energy measurement finds the lower eigenvalue with certainty. What would an immediate measurement in the fixed basis return?

▼ **Work through the solution**

The lower eigenstate there is $(1, -1)/\sqrt{2}$ in the chosen convention. Each fixed-basis outcome therefore occurs with probability one half. Certainty in the instantaneous energy is compatible with maximal uncertainty concerning the diabatic label.

Further reading · Tokmakoff, time-dependent quantum mechanics and Landau–Zener transitions · Rubbmark and colleagues, experimental passage through an avoided crossing.

When a periodic dielectric has no propagating Bloch wave

Fields and the response of matter

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.

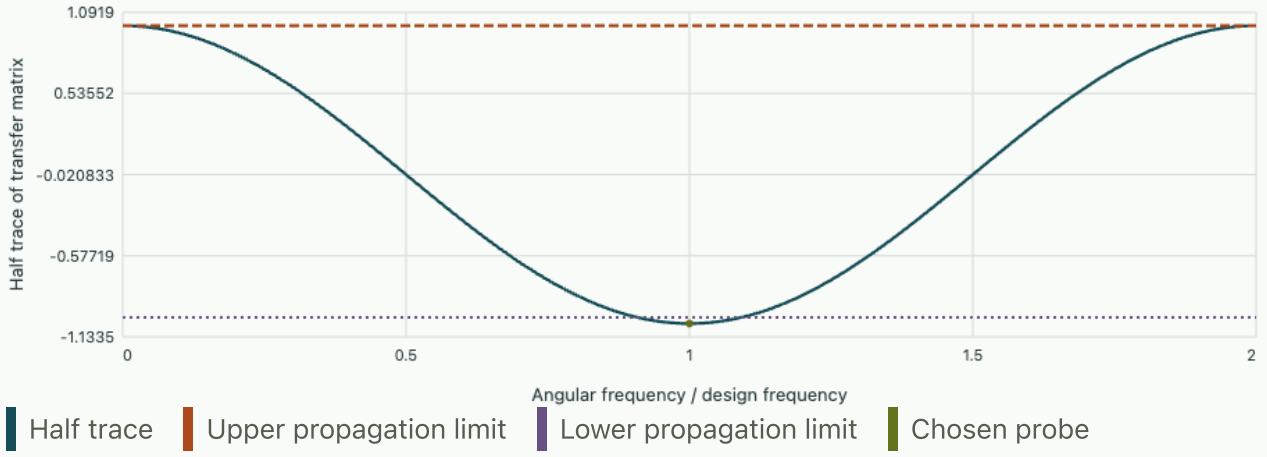
Compare a frequency scan with a finite mirror

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.

Measuring magnetic core loss by a hysteresis loop and by calorimetry

Electromechanics, optics and measurement

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.

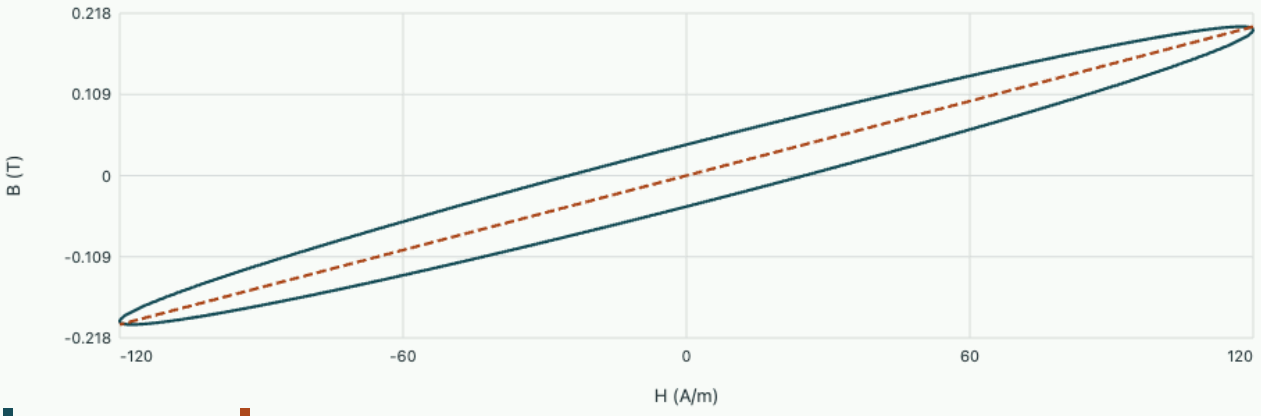
Sweep phase and read the oriented loop

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.

From a Coulomb orbit to the fine-structure multiplet

Quantum states, dynamics and measurement

Separate the hydrogenic Schrödinger equation, identify the origin of its large degeneracy and evaluate the three dominant relativistic corrections. The cancellation among them predicts an equality that spectroscopy can then test.

Separate centre-of-mass and relative motion

A light particle of mass m_e and charge $-e$ interacts with a pointlike nucleus $+Ze$ of mass M. Use centre of mass $\mathbf{R} = (m_e\mathbf{r}_e + M\mathbf{r}_N)/(m_e + M)$ and separation $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_N$. Expanding the sum of the two kinetic energies cancels all cross terms and gives a free centre-of-mass kinetic energy plus relative energy $p^2/(2\mu)$, with reduced mass $\mu = m_eM/(m_e + M)$. The relative potential is $-k/r,\ k = Ze^2/(4\pi\epsilon_0)$. Its rotational symmetry implies $[H, L_i] = 0$. Use a spherical harmonic Y_{lm} and write $\psi = u(r)Y_{lm}/r$, so the probability in a thin radial shell is $|u(r)|^2dr$ for angular harmonics of unit norm.

Substituting in the Laplacian separates the radial equation

$$-\frac{\hbar^2}{2\mu}u'' + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{k}{r}\right]u = Eu.$$

For the regular three-dimensional solution $u(0) = 0$ and a bound state also decays at infinity. The centrifugal term is an effective energy associated with angular variation, not a separate force added by hand. For $E < 0$ define $\kappa = \sqrt{-2\mu E}/\hbar,\ \rho = 2\kappa r$, and $a = \hbar^2/(\mu k)$. The small-radius behaviour of u is r^{l+1}; the large-radius tail contains $e^{-\kappa r}$. Pull both out by $u = \rho^{l+1}e^{-\rho/2}v(\rho)$. Straight differentiation leaves

$$\rho v'' + (2l + 2 - \rho)v' + (1/\kappa a - l - 1)v = 0.$$

For $v = \sum c_j\rho^j$, coefficient matching gives

$$c_{j+1}/c_j = (j + l + 1 - 1/\kappa a)/[(j + 1)(j + 2l + 2)].$$

The regular series must terminate to avoid an exponential asymptotic part that defeats the decaying prefactor. Thus $1/(\kappa a) = n_r + l + 1 = n$, and $E_n = -\mu k^2/(2\hbar^2 n^2)$. The normalised radial polynomials are associated Laguerre functions. This gives $l = 0, \ldots, n - 1$ and $\sum_l(2l + 1) = n^2$ states before including spin. Rotational symmetry guarantees the degeneracy in m, but a different central potential can split levels of different l. The additional equality is specific to the Coulomb potential.

Calculate the radial moments needed for a correction

To isolate the leading relativistic corrections now take an infinitely heavy nucleus and set $m = m_e$, $a = \hbar^2/(mk)$. Including reduced mass in the leading energy is straightforward; a complete recoil correction to the fine structure requires additional interactions. A useful identity is $\langle 1/r\rangle = 1/(an^2)$, obtained by differentiating E_n with respect to k. The derivative of an eigenvalue equals the expectation

of $\partial H/\partial k = -1/r$ for the normalised eigenstate, which follows by differentiating $H|\psi\rangle = E|\psi\rangle$ and cancelling terms using its adjoint. The same method applied to the radial centrifugal parameter yields $\langle r^{-2}\rangle = 1/[a^2 n^3 (l+1/2)]$ at fixed radial node count. Treat l as a continuously varied radial parameter during differentiation, then evaluate at its physical integer value.

For $l > 0$, the stationary radial force identity $\langle dV_{\mathrm{eff}}/dr\rangle = 0$ with vanishing endpoint term gives

$$\langle r^{-3}\rangle = 1/[a^3 n^3 l(l+1/2)(l+1)].$$

Here V_{eff} is the potential in the radial equation. The step uses sufficient behaviour of u, u' at the origin. For $l = 0$, the integral of r^{-3} diverges in an s state and the neglected endpoint term need not vanish. Substituting $l = 0$ formally into this expression would therefore be invalid. The contact density of an s state needed below is $|\psi_{n00}(0)|^2 = 1/(\pi a^3 n^3)$, while states with $l > 0$ vanish at the origin.

Track the kinetic, spin–orbit and contact terms separately

The small parameter is $Z\alpha = k/(\hbar c)$, with α the fine-structure constant for $Z = 1$. Expand the relativistic dispersion about small momentum after subtracting rest energy: $p^2/(2m) - p^4/(8m^3c^2) + \cdots$. For an energy eigenfunction $p^2\psi = 2m(E_n + k/r)\psi$, so $\langle p^4\rangle = 4m^2\langle (E_n + k/r)^2\rangle$ follows by taking the squared norm of $p^2\psi$. Squaring the two operators as if p^2 and $1/r$ commute is unnecessary and would not be an operator identity. Inserting the first two radial moments gives

$$\Delta E_{\mathrm{kin}} = -C_n\left[\frac{1}{l+1/2} - \frac{3}{4n}\right],$$

$$C_n = \frac{mc^2(Z\alpha)^4}{2n^3}.$$

The Pauli expansion of the Dirac Hamiltonian adds $H_{\mathrm{so}} = (1/[2m^2c^2r])V'(r)\mathbf{L}\cdot\mathbf{S}$ and the Darwin contact $H_D = \hbar^2\nabla^2 V/(8m^2c^2)$. These are effective interactions through order $(Z\alpha)^4$; a purely nonrelativistic scalar equation does not generate the spin or contact term. For $V = -k/r$, $V' = k/r^2$ and $\nabla^2 V = 4\pi k\delta^3(\mathbf{r})$. Thus the contact shifts only s states by $+C_n$. For $l > 0$ insert the third inverse radial moment and the angular identity $\mathbf{L}\cdot\mathbf{S} = \hbar^2[j(j+1) - l(l+1) - 3/4]/2$. For the branch $j = l + 1/2$ the bracket is l; for $j = l - 1/2$ it is $-(l+1)$. Reducing fractions then gives

$$\Delta E_{\mathrm{so}} = C_n[1/(l+1/2) - 1/(j+1/2)].$$

For $l = 0$ the spin–orbit operator vanishes and the Darwin term supplies the required difference. The sum in either case is

$$\Delta E_{nj} = -C_n\left[\frac{1}{j+1/2} - \frac{3}{4n}\right].$$

At fixed n the answer depends on j but no longer separately on l. The cancellation cannot be inferred by merely attaching a phenomenological $\mathbf{L}\cdot\mathbf{S}$ splitting to a nonrelativistic p state. In particular it predicts

identical energies for $2s_{1/2}$ and $2p_{1/2}$ in this approximation, while $2p_{3/2}$ lies higher. Radiative, recoil and finite-nuclear-size contributions can resolve some of these residual equalities. An experimentally resolved residual is evidence that the specified Hamiltonian omitted a physical term; it does not remove the cancellation inside that Hamiltonian.

Compare a small energy correction with a resolvable line

For hydrogen with $mc^2 \simeq 511\,000$ eV and $\alpha \simeq 1/137.04$, the $2p_{3/2}$–$2p_{1/2}$ interval is $C_2/2 \simeq 4.53 \times 10^{-5}$ eV or 10.95 GHz. Compare with the leading $n = 2$ binding energy 3.40 eV. A small correction to total binding can give a prominent splitting when a measurement subtracts nearly equal large energies. Resolving it in a Fourier-limited time-domain experiment requires a coherent observation time at least of order the inverse frequency difference, here around 0.09 ns. Actual linewidths, lifetimes and pulse shapes affect how two spectral components can be distinguished.

A useful numerical audit is to fit the radial spectrum with three successive physical models. First solve the nonrelativistic radial boundary problem with the correct near-origin power and check that 2s and 2p converge toward the same energy. Then evaluate expectation values of the three corrections on those numerical states. Finally combine terms. Converged binding energies alone do not guarantee a converged contact density, which samples one spatial endpoint. A coarse radial grid can imitate an anomalous splitting between states which the analytical correction predicts to be equal. Check the origin slope of u_{n0}, not only its norm or far tail.

Find where the electron is most likely to be detected

Confusion between spatial density and radial probability can lead to very different-looking descriptions of the same atom. For a hydrogenic ground state the normalised position-space wavefunction is $\psi_{100} = (\pi a^3)^{-1/2} e^{-r/a}$. Its density is largest at the centre, but a thin spherical shell at radius r contains probability $P(r)dr = 4r^2 e^{-2r/a} dr/a^3$. The shell volume tends to zero at the origin. Differentiating $\ln P$ gives $2/r - 2/a = 0$, so the most likely shell radius is a. The mean radius $\int rP(r)dr = 3a/2$ is different again because the shell distribution has a long tail. Neither number alone gives the probability that a finite spatial detector clicks.

Integrating the shell probability between zero and R gives

$$\Pr(r < R) = 1 - e^{-2R/a}\left(1 + \frac{2R}{a} + \frac{2R^2}{a^2}\right).$$

At the most likely shell radius $R = a$, the probability inside is $1 - 5e^{-2} = 0.3233$, substantially less than one half. At $R = 2a$ it is

$$1 - 13e^{-4} = 0.7619.$$

Taylor expansion at a small radius $R \ll a$ gives $4(R/a)^3/3$, equal to the ball volume times central density to leading order. This is the local scaling underlying finite-nucleus corrections to a contact-sensitive observable. Using the most likely radius as though the electron followed a thin circular orbit

would give none of these cumulative probabilities, even though the same Coulomb length scale appears in the Bohr model.

For $2s$ and $2p$, both states have principal index two and the same leading energy. The $2s$ radial amplitude is proportional to $(2 - r/a)e^{-r/(2a)}$; its node at $2a$ separates regions of opposite amplitude sign. The $2p$ radial amplitude is proportional to $re^{-r/(2a)}$, with no positive-radius radial node. At short distance, the probability inside R scales as R^3 for a nonzero s amplitude and as R^5 for the p wave. A sufficiently localised weak perturbation near the nucleus can therefore split the nonrelativistic degeneracy even if it is rotationally symmetric. Rotational symmetry still protects equal shifts within the m multiplet. A radial perturbation and a field selecting a direction leave different patterns of protected degeneracy.

Diagonalise a degeneracy before expanding a weak electric field

The $n = 2$ Coulomb degeneracy creates a striking response to a weak uniform electric field along z. In the purely nonrelativistic model with spin ignored the relevant spatial states are $2s,\ 2p_0,\ 2p_{\pm1}$. The perturbation for an electron in electric field $\mathcal{E}\hat{z}$ can be written $e\mathcal{E}z$ up to an additive constant. Each unperturbed parity eigenstate has zero diagonal expectation of this odd operator. Applying ordinary nondegenerate first-order perturbation theory separately to all four states would therefore predict zero for every shift. But first-order perturbation theory in a degenerate manifold requires diagonalising the projected perturbation, including off-diagonal amplitudes.

The $z = r\cos\theta$ operator has $q = 0$ and changes parity, so within $n = 2$ it couples only $2s$ and $2p_0$. Its angular matrix element between Y_{00} and Y_{10} is $1/\sqrt{3}$. In a conventional phase convention the radial functions are $R_{20} = (2 - r/a)e^{-r/(2a)}/(2\sqrt{2}\,a^{3/2})$ and $R_{21} = (r/a)e^{-r/(2a)}/(2\sqrt{6}\,a^{3/2})$. The remaining radial integral is $\int r^3 R_{20}R_{21}\,dr$. Setting $s = r/a$ gives

$$a(2 \cdot 4! - 5!)/(8\sqrt{3}) = -3\sqrt{3}a.$$

Multiplying the angular factor gives $-3a$. Rephasing one basis state would change this sign but leave all level shifts invariant.

The projected matrix has zero diagonal and off-diagonal $-3ea\mathcal{E}$. Its symmetric and antisymmetric combinations have linear shifts $\mp 3ea\mathcal{E}$, while $2p_{\pm1}$ remain unshifted to this order. The new states have definite first-order dipoles along opposite directions even though their zero-field basis states had zero dipoles. For $a = 5.29 \times 10^{-11}$ m and $\mathcal{E} = 1000$ V/m, an individual linear shift has magnitude 1.59×10^{-7} eV, corresponding to 38.4 MHz. This field is weak compared with the gross level separation, but whether the degenerate result is accurate in a real atom also depends on much smaller fine, hyperfine and radiative splittings. “Weak field” must specify the comparison energy.

A more precise two-state model retains a zero-field splitting Δ between the coupled states. After subtracting their mean the Hamiltonian is $(\Delta/2)\sigma_z + d\mathcal{E}\sigma_x$, with $d = 3ea$ in magnitude. The energies are $\pm\sqrt{(\Delta/2)^2 + d^2\mathcal{E}^2}$. At $d|\mathcal{E}| \ll |\Delta|$ the shift from each initial level is quadratic, $\pm d^2\mathcal{E}^2/|\Delta|$, while at $d|\mathcal{E}| \gg |\Delta|$ the slopes approach the linear-degenerate prediction. An observed crossover distinguishes a small unresolved internal splitting from a strictly zero projected dipole matrix element. Tracking energies over both positive and negative fields also tests for a stray dc field shifting the centre of the response.

Estimate a finite-size correction from the potential itself

The contact term derived from the Pauli expansion is not the same physical effect as giving the nucleus a finite charge radius. A simple electrostatic model separates the two scalings. Replace a point nucleus by a uniformly charged sphere of radius R_N. Inside it the electron potential becomes

$$V_{\mathrm{sph}}(r) = -k[3 - (r/R_N)^2]/(2R_N);$$

outside it is unchanged at $-k/r$. The perturbation $\delta V = k[1/r - 3/(2R_N) + r^2/(2R_N^3)]$ is positive for interior points and zero at the surface. The smoothed potential is less attractive than the point law near the centre.

For an s state and $R_N \ll a$, the unperturbed density varies little across the tiny sphere. First-order perturbation theory therefore gives

$$\delta E \simeq 4\pi k|\psi(0)|^2 \int_0^{R_N} \left(r - \frac{3r^2}{2R_N} + \frac{r^4}{2R_N^3} \right) dr$$

$$\frac{2\pi}{5}\, k|\psi(0)|^2 R_N^2.$$

The first two radial integrals cancel and leave the $R_N^2/10$ contribution before multiplication by 4π. Equivalently write $\delta E = (2\pi/3)k|\psi(0)|^2\langle r_N^2\rangle$, since the uniform sphere has mean-square charge radius $3R_N^2/5$. The result is a positive shift, decreasing binding. Corrections to the constant-density approximation involve additional powers of R_N/a. A different microscopic charge distribution changes higher moments, while the leading small-size shift depends on its mean-square charge radius.

For a model $R_N = 1.0$ fm and the ordinary electronic hydrogen length $a = 5.29 \times 10^{-11}$ m, $R_N/a = 1.89 \times 10^{-5}$. For the ground state compare with $|E_1| = k/(2a)$ and $|\psi(0)|^2 = 1/(\pi a^3)$. The relative shift is $(4/5)(R_N/a)^2 = 2.86 \times 10^{-10}$, equivalent to roughly 3.9×10^{-9} eV. This illustrative uniform charge radius is input data of the model; do not equate it directly to a measured rms radius. For a heavier orbiting lepton a shorter Bohr radius increases the fractional sensitivity strongly. A complete precision spectrum would additionally include radiative and recoil effects and a realistic nuclear response. The present derivation isolates how one local perturbation scales and why s levels are especially sensitive.

Use dilation to audit stationary Coulomb energies numerically

A variational scaling argument gives a powerful independent check on a computed wavefunction. For any normalised trial state ψ, define the normalised rescaling $\psi_\lambda(\mathbf{r}) = \lambda^{3/2}\psi(\lambda\mathbf{r})$. The Coulomb potential is homogeneous of degree minus one, and the gradient acquires a factor λ. Its energy expectation is therefore

$$E(\lambda) = \lambda^2\langle T\rangle + \lambda\langle V\rangle.$$

For an exact stationary bound state the energy functional is stationary with respect to norm-preserving variations, giving $2\langle T\rangle + \langle V\rangle = 0$ at $\lambda = 1$. Hence $\langle T\rangle = -E,\ \langle V\rangle = 2E$, consistent with the independently derived inverse-radius mean. The argument requires appropriate normalisation and boundary behaviour so the scaled trial function is allowed.

Suppose a numerical $1s$ reconstruction yields $\langle T\rangle = 13.2$ eV and $\langle V\rangle = -26.7$ eV. Its energy is -13.5 eV, yet its virial residual is -0.3 eV. Along the scale family the minimum occurs at $\lambda_* = -\langle V\rangle/(2\langle T\rangle) = 1.0114$, a slightly more concentrated state. The resulting energy reduction is only about 0.0017 eV, much smaller than the change in individual kinetic and potential terms. A small total-energy error can coexist with a noticeably incorrect radial scale because first-order energy errors cancel near a stationary solution. Monitoring separate expectation values, the virial residual and a contact-sensitive observable provides a more informative test of a numerical atomic calculation than reporting the total binding energy alone.

Distinguish orbital density from radial probability

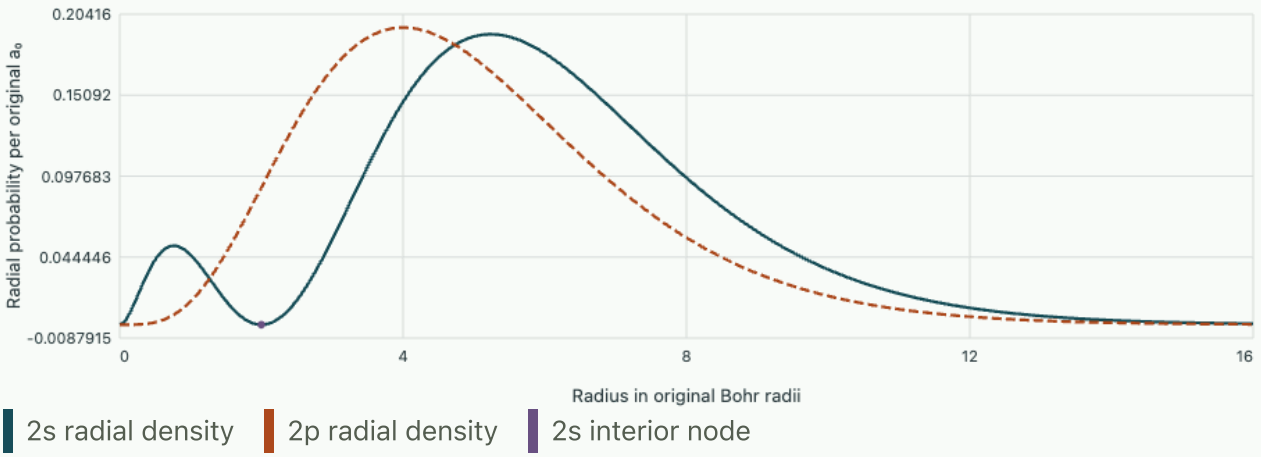
Compare radial states at the same principal quantum number. Inspect their nodes before considering which additional interactions can lift their degeneracy.

Open this experiment online

Try the calculation

Nuclear charge Z

1



The 2s node is at 2 original Bohr radii. The most probable 2p radius is 4 original Bohr radii.

Appraise a finite-grid Coulomb spectrum from convergence and degeneracy

To compute the radial Schrödinger eigenvalues directly, work in dimensionless radius $s = r/a$, energy unit $\hbar^2/(2ma^2)$, and write $[-d^2/ds^2 + l(l+1)/s^2 - 2/s]u = \varepsilon u$. On a uniform grid $s_j = jh$, use

$u_0 = 0$ and a far boundary $u_{N+1} = 0$. The diagonal approximation is $2/h^2 + l(l+1)/s_j^2 - 2/s_j$ and its neighbours are $-1/h^2$. All potential terms are evaluated at positive interior grid points. Inserting an arbitrary finite value for $1/s$ at the origin and treating it as an ordinary interior point changes the boundary approximation. A tridiagonal symmetric eigensolver returns orthogonal numerical vectors; physical radial norm is $h \sum_j |u_j|^2$, so normalise appropriately before reconstructing spatial densities.

For $l = 0$ on an infinite grid the lowest mode permits an unusually informative exact lattice check. Try $u_j = jq^j$ for $0 < q < 1$. The finite-difference kinetic row divided by this trial value is $[2 - q - q^{-1}]/h^2 - (q - q^{-1})/(jh^2)$. Adding potential $-2/(jh)$, the $1/j$ coefficient vanishes when $q^{-1} - q = 2h$, or $q = \sqrt{1+h^2} - h$. The eigenvalue is then

$$\varepsilon_h = [2 - 2\sqrt{1+h^2}]/h^2 = -2/[1 + \sqrt{1+h^2}].$$

It tends to the exact continuum ground value -1 as $-1 + h^2/4 + O(h^4)$. This analytic discrete answer separates algebraic solver correctness from discretisation error for the exact same potential and boundary treatment.

At $h = 0.20$ the infinite-grid ground value is approximately -0.99020; at $h = 0.10$ it is -0.99751. Their errors differ by nearly a factor of four as predicted. A finite far wall shifts both energies further; repeat first at fixed h and increasing boundary radius to isolate that effect. An excited $n = 2$ state has a more extended tail and requires a larger radius than the ground state for comparable truncation bias. Solving $l = 0$ and $l = 1$ on a small common box can yield a spurious nonrelativistic 2s–2p splitting. Check separately that both energies tend to $-1/4$ before interpreting a much smaller physical fine splitting. Subtracting independently computed energies amplifies absolute numerical biases even if each number appears visually close to its expected value.

The distinctive near-origin behaviour supplies a second check. Since $\psi_{100}(0)$ is proportional to the slope of the reduced radial function at zero, a first forward difference u_1/h typically incurs a first-order error even when the energy error is second order. For the lattice shape above $u_j \propto jq^j$, the first two positive grid values can be used to estimate the endpoint derivative by $(4u_1 - u_2)/(2h)$. Test its convergence under the physical radial normalisation. For a contact perturbation, squaring an endpoint derivative doubles its small relative error to leading order. This is why an atomic calculation should include an observable-specific refinement criterion in addition to an energy tolerance.

Use an eigenvalue benchmark to separate numerical and physical shifts

A report on the infinite-grid hydrogenic ground state quotes energy -0.99750 at $h = 0.10$ and attributes the difference from -1 to a short-range interaction. Predict the lattice-only change on refinement to $h = 0.05$, and propose a test that could reveal a real interaction independently.

▼ Derive and check the result

The lattice expression gives $-2/[1 + \sqrt{1.0025}] \simeq -0.999376$ at 0.05, so the discretisation alone accounts for the original difference at the stated accuracy. A local physical correction with fixed continuum strength should converge toward a fixed nonzero shift as $h \to 0$, while the lattice bias vanishes approximately as $h^2/4$. Compute control and perturbed energies on exactly the same grids and

compare their difference under refinement, with consistent treatment of the perturbation's physical range. An unresolved delta perturbation represented as a fixed diagonal value on one grid point changes its effective continuum strength upon refinement and would produce another artificial shift. Cross-check with the first-order expectation of a resolved physical potential or with the independently converged contact density.

Use an isotope pair to cancel a common transition-energy reference

High-precision spectroscopy often obtains its most reliable insight from combinations of frequencies rather than absolute energies. Consider two hydrogenic model isotopes with the same nuclear charge, different reduced masses μ_1, μ_2 and different mean-square charge radii R_1^2, R_2^2. For a specified transition the leading point-Coulomb frequency can be written $C\mu$, since its n-dependent dimensionless numbers are fixed. A local nuclear-size perturbation shifts each participating s level by a factor proportional to $|\psi_{ns}(0)|^2 R_i^2 \propto \mu_i^3 R_i^2/n^3$, while a p level receives a much smaller finite-size correction in the same limit. The different powers of reduced mass and radial quantum number permit differences of transitions to cancel selected terms.

For illustration compare a 1s–2s optical frequency with a 1s–2p frequency within one model isotope, assuming the fine, radiative and differing upper-state structures have been independently accounted for in the desired order. Subtracting these frequencies cancels the entire common lower 1s energy and the gross equal 2s/2p Coulomb energy, leaving the upper 2s–2p differences. If instead one sought the large 1s contact contribution, subtracting these same two lines would remove precisely the effect of interest. An apparently convenient common reference can therefore eliminate physical signal as well as noise. Write the coefficient of each contributing level explicitly before selecting a frequency combination.

Suppose two independently fitted line frequencies ν_a, ν_b each have a 10 MHz common reference error and independent residual standard errors of 0.30 and 0.40 MHz. The covariance from the shared reference is $100\,\mathrm{MHz}^2$. The difference has variance $0.30^2 + 0.40^2 = 0.25\,\mathrm{MHz}^2$, hence 0.50 MHz standard error. Treating both total errors as independent instead would suggest roughly 14 MHz. Conversely, an error in a calibration scale proportional to frequency leaves a residual proportional to the frequency difference; its cancellation is only approximate. Reporting the provenance of uncertainties and their covariance can matter more than recording an extra digit of each absolute transition.

Recover a contact shift from two excited-state probes

In a diagnostic model a local perturbation gives an unknown positive 1s shift D, an ns shift D/n^3, and zero shifts of p levels. Relative to a matched reference, how do 1s–2s, 1s–2p and 1s–3s frequencies change? Which experimental difference removes the unknown common lower-state shift yet recovers D?

▼ Derive and check the result

The changes are $(-7D/8)/h,\ -D/h,\ (-26D/27)/h$, respectively. Subtracting the 1s–2p shift from the 1s–2s shift gives $D/(8h)$, since the common lower-state contribution cancels. For a measured

difference of 0.90 ± 0.05 MHz the implied $D/h = 7.2 \pm 0.4$ MHz. The 1s–3s shift should then be -6.93 MHz in this stipulated model. If unavailable corrections also differ among the upper states, that value is not an absolute nuclear-radius extraction. The third line gives an independent test of the $1/n^3$ contact scaling before attributing all measured differences to the single parameter D.

Tutorial · measure degeneracy through independent calculations

Verify two inverse moments directly in the $2p$ state

Use $R_{21}(r) = re^{-r/(2a)}/(2\sqrt{6}\,a^{5/2})$ and normalised Y_{1m}. Evaluate $\langle r^{-2}\rangle$, $\langle r^{-3}\rangle$ without parameter differentiation.

▼ **Work through the solution**

The radial measure is r^2dr, so after inserting R the general moment is $(24a^5)^{-1}\int_0^\infty r^{4-q}e^{-r/a}dr$ for r^{-q}. For $q = 2$ the integral is $2a^3$, giving $1/(12a^2)$. For $q = 3$ it is a^2, giving $1/(24a^3)$. Both agree with the formulas for $n = 2, l = 1$. Dropping the radial Jacobian would change both powers and could falsely suggest singular behaviour. The result is independent of m because the operator is rotationally scalar.

Calculate the complete $n = 2$ shifts in units of C_2

Resolve the kinetic, spin–orbit and Darwin contributions for $2s_{1/2}$, $2p_{1/2}$ and $2p_{3/2}$. Which comparison isolates the effect of the spin–orbit term cleanly at this order?

▼ **Work through the solution**

For s, kinetic gives $-[2 - 3/8] = -13/8$ and Darwin $+1$, summing to $-5/8$. For both p states the kinetic value is

$$-2/3 + 3/8 = -7/24.$$

The spin–orbit corrections are $2/3 - 1 = -1/3$ for $j = 1/2$ and $2/3 - 1/2 = 1/6$ for $j = 3/2$, giving $-5/8$ and $-1/8$. Subtracting the two p levels removes their shared kinetic term and zero contact term. Their interval is $C_2/2$. Comparing 2s with either p instead involves two different radial wavefunctions and does not isolate the same correction.

Separate a reduced-mass shift from fine structure

Consider two otherwise identical hydrogenic nuclei with $M/m = 1836$ and 3672. Predict the ratio of gross transition frequencies and compare its fractional change with the order of a fine-structure correction. Which common interpretation would conflate them?

▼ **Work through the solution**

Gross transition frequencies scale as reduced mass $\mu = m/(1 + m/M)$. The heavier-to-lighter ratio is $(1 + 1/1836)/(1 + 1/3672) = 1.000272$ approximately. The order of a relative fine-structure correction for $Z = 1$ is $\alpha^2 \simeq 5.3 \times 10^{-5}$ with state-dependent coefficients. A change of isotope can shift a gross line by more than a particular resolved fine splitting, even with identical nuclear charge. Centroid measurements across isotopes and intervals within one multiplet are distinct observables. Recoil and nuclear magnetism require further inputs for precision predictions.

Further reading

- Fließbach · Dreidimensionale Probleme
- Cambridge · hydrogen, helium and fine structure
- Course investigation · perturbations and level mixing

Two electrons, exchange splitting and the cost of localisation

Quantum states, dynamics and measurement

Build a two-electron state that respects exchange, derive its charge and spin correlations, and then solve a two-site model exactly. The two examples reveal when a triplet is favoured and when an apparently weak hopping lowers the singlet instead.

Construct states before introducing a force

A two-electron wavefunction has two spatial arguments and two spin labels. The labels 1 and 2 identify argument slots; they do not attach persistent tags to individual identical electrons. Exchanging all position and spin variables must multiply an electronic state by minus one. This is the fermionic exchange postulate. It constrains which vectors describe electrons even when their Hamiltonian contains no spin-dependent term. A measurement with no access to the ordering of slots is represented by an observable symmetric under the same exchange.

Let $a(\mathbf{r}), b(\mathbf{r})$ be normalised orthogonal spatial orbitals. The symmetric and antisymmetric spatial combinations are $\Phi_\pm = [a(1)b(2) \pm b(1)a(2)]/\sqrt{2}$. Combine Φ_+ with the spin singlet $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$, and Φ_- with any of the three symmetric triplet spin states. For equal positions the antisymmetric spatial combination vanishes exactly. Its joint probability has an exchange hole at coincidence even for noninteracting same-spin fermions. When $a = b$, the antisymmetric combination vanishes identically and only the symmetric spatial state with opposite spins in a singlet is available. Occupation of the same spatial mode by two electrons is thus compatible with the exclusion principle if the full spin orbitals differ.

For two distinct orthogonal modes, both Φ_+ and Φ_- have the same one-particle number density $|a(\mathbf{r})|^2 + |b(\mathbf{r})|^2$, normalised to two electrons. To see this, integrate $|\Phi_\pm(\mathbf{r}_1, \mathbf{r}_2)|^2$ over $\mathbf{r}_2$ and multiply by two. Cross terms contain $\langle a|b\rangle = 0$. Coincidence and correlation measurements can therefore distinguish preparations that share their entire one-particle density. The spin-antisymmetric singlet can also give correlated opposite spin outcomes along any common measurement axis, while both individual spins are unpolarised. Position and spin comparisons probe different reduced descriptions of the same total state.

Separate the direct and exchange integrals

For an independent-orbital estimate take $H = h(1) + h(2) + v(1,2)$ with $ha = \epsilon_a a,\ hb = \epsilon_b b$, and a symmetric spin-independent repulsion. Expand the expectation of v in $\Phi_\pm$. Diagonal terms give the same direct integral and off-diagonal terms are conjugates:

$$J_{ab} = \iint |a(1)|^2 v(1,2) |b(2)|^2 \, d^3r_1 d^3r_2,$$

$$K_{ab} = \iint a^*(1) b(1) v(1,2) b^*(2) a(2) \, d^3r_1 d^3r_2.$$

For Coulomb repulsion K_{ab} is nonnegative: it is the positive Coulomb quadratic form of the transition density $a^{*}b$, seen for instance from its Fourier representation weighted by $4\pi/k^{2}$. Orthonormality only makes the spatial integral of the transition density vanish, not its weighted self-interaction. The expectation energies are $\epsilon_a + \epsilon_b + J_{ab} \pm K_{ab}$. In this fixed pair of spatial orbitals the triplet is lower by $2K_{ab}$. If the orbitals have disjoint support the exchange integral vanishes, although the direct repulsion need not. A splitting caused by spatial symmetry does not require an extra microscopic spin–spin repulsive force.

Set $J_{ab} = 0.80$ eV and $K_{ab} = 0.070$ eV in an illustrative excited configuration. Its Coulomb shifts are 0.87 eV and 0.73 eV. A weighted average over all four spin states is

$$J - K/2 = 0.765\,\mathrm{eV}.$$

Optical line strengths generally do not populate those four spin states equally. The energy difference is a statement about this restricted configuration; another spatial occupation can place a singlet below the entire triplet manifold. Predicting a ground-state spin therefore requires admitting the competing occupations and the one-body and correlation energies they carry.

Allow hopping and double occupation

Now take two orthonormal localised orbitals L, R of equal one-body energy. A fermion hops between them at amplitude $-t$, and placing opposite spins on the same site costs $U > 0$. Define $H = -t\sum_{\sigma}(c^{\dagger}_{L\sigma}c_{R\sigma} + c^{\dagger}_{R\sigma}c_{L\sigma}) + U\sum_i n_{i\uparrow}n_{i\downarrow}$. This is a specified two-site Hubbard model, not a complete microscopic Coulomb Hamiltonian. Its direct intersite exchange integral is set to zero here, so its spin ordering need not follow the fixed-orbital calculation above. Position, spin and field units enter through the experimentally or computationally determined t, U.

In the two-electron sector a state with two up spins must occupy one orbital on each site. Hopping would create two up spins in one orbital and therefore annihilates this state. Rotational symmetry implies the same energy zero for all three triplets. Define the singly occupied singlet $|S\rangle$ and symmetric doubly occupied state $|D_{+}\rangle = (|\uparrow\downarrow, 0\rangle + |0, \uparrow\downarrow\rangle)/\sqrt{2}$. A consistent fermionic ordering gives an off-diagonal $-2t$ between them; each direction and spin contributes and the normalisations leave the factor two. Changing the phase of D_{+} reverses that matrix element but leaves the eigenvalues unchanged. The orthogonal antisymmetric doublon remains at energy U. The relevant singlet block is

$$H_S = \begin{pmatrix} 0 & -2t \\ -2t & U \end{pmatrix}, \qquad E_{\pm} = \frac{U \pm \sqrt{U^2 + 16t^2}}{2}.$$

The ground energy $E_{-} < 0$ for $t \neq 0$. At $U \gg |t|$ expand the square root to obtain

$$E_{-} \simeq -4t^2/U + 16t^4/U^3 + \cdots.$$

The singlet–triplet splitting $J_{\mathrm{eff}} = E_T - E_{-} \simeq 4t^2/U$ can be represented in the singly occupied subspace by $J_{\mathrm{eff}}(\mathbf{S}_L \cdot \mathbf{S}_R/\hbar^2 - 1/4)$. The singlet can lower its energy by virtual double occupation

while spin symmetry blocks the triplets. The inferred spin coupling scales as the square of hopping and inversely with the energy cost. This mechanism is commonly called antiferromagnetic superexchange.

Determine a correlation without reading an eigenvector

The fraction of probability with double occupation in the ground state is $\langle \partial H/\partial U \rangle$. Differentiating its energy gives

$$D = \tfrac{1}{2}[1 - U/\sqrt{U^2 + 16t^2}].$$

For positive large U, $D \simeq 4t^2/U^2$. Comparing the derivative with the explicit eigenvector provides a useful independent test of signs and normalisations. From $Es = -2td$ one obtains

$$d/s = -E/(2t).$$

Normalising gives $D = E^2/(4t^2 + E^2)$, identical to the derivative expression for either lower-branch parameter value.

For $t = 10$ meV and $U = 100$ meV, $\sqrt{U^2 + 16t^2} = 107.70$ meV and the exact splitting is 3.85 meV. The leading 4.00 meV expression is about 3.9% high. Double occupation is only 3.58%, yet forbidding it at the outset would predict exactly zero splitting. A virtually occupied state can matter appreciably in a low-energy Hamiltonian even when a projective charge measurement rarely finds it occupied. At $U = 0$ the ground energy is $-2|t|$, corresponding to two opposite spins in the bonding one-electron orbital; the double-occupation fraction is one half. The large-U expression diverges as $U \to 0$ because its perturbative premise disappears. The exact two-state matrix remains finite.

Derive an effective screening charge for a helium-like ground state

A two-electron atom provides a continuous-space counterpart to the two-site model. In atomic units set $\hbar = m_e = e^2/(4\pi\epsilon_0) = 1$ and take an infinitely heavy nucleus of charge Z. The nonrelativistic Hamiltonian is

$$H = \sum_{i=1}^{2}[-\nabla_i^2/2 - Z/r_i] + 1/r_{12}.$$

The two one-electron ground orbitals may be identical in space because opposite spins can form an antisymmetric singlet. Use a trial product of two normalised hydrogenic functions $u_\zeta(\mathbf{r}) = (\zeta^3/\pi)^{1/2} e^{-\zeta r}$ with a variable positive exponent ζ. This single parameter lets the spatial cloud expand in response to mutual repulsion. Choosing the exponent equal to Z without testing it would impose the noninteracting size on an interacting problem.

A single orbital gives $\langle -\nabla^2/2 \rangle = \zeta^2/2$ and $\langle 1/r \rangle = \zeta$, either by direct integration or by hydrogenic scaling. Two electrons thus contribute $\zeta^2 - 2Z\zeta$ to the expectation. To derive the repulsion term,

average $1/|\mathbf{r}_1 - \mathbf{r}_2|$ over two independent uniform angles at fixed radii. Gauss's shell result, or a direct change of variable in the angular integral, gives $1/\max(r_1, r_2)$. The radial probability is

$$P(r) = 4\zeta^3 r^2 e^{-2\zeta r}.$$

Symmetry between the two electron labels permits restricting one integral to $r_2 < r_1$, then doubling:

$$J = 2\int_0^\infty \frac{P(r)}{r}\, dr \int_0^r P(s)\, ds,$$

$$\int_0^r P(s) ds = 1 - e^{-2\zeta r}(1 + 2\zeta r + 2\zeta^2 r^2).$$

The first part of the resulting r integral gives 2ζ. The three exponential subtraction terms sum to $11\zeta/8$, leaving $J = 5\zeta/8$. Their evaluation uses $\int_0^\infty r^n e^{-ar} dr = n!/a^{n+1}$. Both the dependence linear in ζ and the positive sign follow from scaling an inverse-distance repulsion; only the numerical coefficient needed the angular and radial integration. Replacing the expectation of $1/r_{12}$ by $1/\langle r_{12}\rangle$ would generally give a different coefficient because inverse distance is nonlinear.

Consequently $E(\zeta) = \zeta^2 - 2(Z - 5/16)\zeta$ is an upper bound on the ground energy. It minimises at $\zeta_* = Z - 5/16$, with $E_{\mathrm{var}} = -(Z - 5/16)^2$ hartree for positive ζ_*. For helium $Z = 2$, $\zeta_* = 1.6875$, giving -2.84766 hartree, compared with -2.75 hartree at $\zeta = 2$. The one-electron ion has energy -2 hartree, so this variational value lies below the threshold for an electron plus the ion and demonstrates binding within the model. The trial ionisation energy 0.84766 hartree is a lower bound on the exact ionisation energy when that threshold is known exactly. An upper bound on the neutral energy gives a lower bound on the energy required to remove an electron.

The interpretation of $5/16$ as a screening correction is limited. At short radii each electron feels nearly the full nuclear singularity, while far outside it sees the remaining net charge. One constant exponent cannot reproduce both limits and cannot describe how the two electron positions avoid one another at fixed one-particle radius. The product trial permits nonzero joint weight at electron coincidence. Multiplying by a factor involving r_{12} introduces genuine correlation beyond independent orbitals. Optimising a broader variational family containing this product can only lower or retain the variational energy, provided its integrals and normalisation are computed consistently. Monitoring energy together with the pair-separation distribution tests whether the enlargement captures the physics of repulsion rather than merely adjusting a one-body radius.

Infer a local two-electron boundary condition from the Coulomb singularity

Near coincidence of two unlike-spin electrons, the Coulomb repulsion diverges like $1/r_{12}$. The wavefunction itself can remain finite at coincidence, but its derivative must cancel this singularity in the Schrödinger equation. Consider a fixed smooth centre coordinate and average the wavefunction over the direction of the interelectron separation r, away from simultaneous nuclear coincidence. Assume an expansion $\bar{\Psi}(r) = A + Br + \cdots$. In atomic units the reduced mass is one half, so the relative spherical kinetic operator is $-[d^2/dr^2 + (2/r)d/dr]$ for the locally s-wave part. Its leading singular term is

$-2B/r$. Adding $\bar{\Psi}/r$ gives coefficient $A - 2B$ of $1/r$. A finite energy eigenvalue requires this coefficient to vanish:

$$\left.\frac{d\bar{\Psi}}{dr}\right|_{0} = \frac{1}{2}\,\bar{\Psi}(0).$$

The positive cusp slope means the spatial amplitude rises as the electrons move apart from contact, consistent with repulsion. The corresponding electron–nucleus cusp has opposite sign and coefficient set by the nuclear charge and reduced mass. In a same-spin state the scalar contact amplitude vanishes by antisymmetry; this s-wave formula becomes $0 = 0$ and a leading higher-partial-wave analysis is needed. Forgetting either angular averaging or exchange symmetry can turn an instructive local condition into a false statement about every directional derivative of every wavefunction.

A product of independent $1s$ orbitals is smooth in r_{12} at a generic fixed centre and has no required linear cusp in the spherical average. A trial factor $1 + r_{12}/2$ in atomic units can impose the first-order unlike-spin cusp while leaving additional variational freedom to model long-range behaviour. Merely enforcing the cusp does not solve the entire two-electron problem. It constrains a local ratio of derivative to value, while energy depends on a six-dimensional integral. In variational Monte Carlo one can diagnose this distinction through the local energy $H\Psi/\Psi$ near coincidence. A cusp-mismatched trial may have a singular local energy even though the volume element $r^2 dr$ keeps its integrated average finite. Imposing the cusp can reduce variance and improve short-distance pair observables more dramatically than it changes the energy expectation.

Use one-body occupation measurements to detect missing correlations

A single determinant of orthonormal spin orbitals χ_i has a one-particle reduced density matrix $\gamma = \sum_{\text{occupied}} |\chi_i\rangle\langle\chi_i|$, normalised to the electron number N_e. Thus $\gamma^2 = \gamma$ and its eigenvalues are zero or one. This is an algebraic signature of a single pure Slater determinant in the spin-orbital convention. For a correlated pure fermion state, natural occupations can become fractional while remaining between zero and one and summing to N_e. Ensemble mixing can produce fractional occupations too, so knowing whether the entire state is pure matters if one wants to interpret fractional one-particle populations as evidence of correlation.

Consider the normalised two-electron singlet $|\Psi\rangle = \sqrt{0.8}\,|a\uparrow, a\downarrow\rangle + \sqrt{0.2}\,|b\uparrow, b\downarrow\rangle$, with each ket a determinant. A one-body operator changes at most one occupied spin orbital, so its matrix element between these two doubly different determinants vanishes. The four natural spin-orbital occupations are $0.8, 0.8, 0.2, 0.2$. Their sum is two while $\operatorname{tr}\gamma^2 = 1.36$, smaller than the determinant value two. An incoherent mixture of the two determinants at the same weights has the same one-body matrix, although it no longer has their joint coherence. A two-body pair-transfer operator that converts both a electrons into b electrons, plus its adjoint, can reveal a relative phase between these branches. Thus a complete one-body spectrum still need not determine a two-body state.

The symmetric Hubbard dimer ground state has another direct relation between energy derivatives and its one-body coherence. Let $T = \sum_{\sigma}(c^{\dagger}_{L\sigma}c_{R\sigma} + c^{\dagger}_{R\sigma}c_{L\sigma})$. The energy derivative gives

$$\langle T \rangle = -\partial E_{-}/\partial t = 8t/\sqrt{U^2 + 16t^2}.$$

At $U = 0$ and $t > 0$, $\langle T \rangle = 2$ because both electrons occupy the bonding orbital. At large U/t this coherence approaches zero even though the exact finite-t ground state remains a total spin singlet. In the bonding/antibonding spatial basis, the sum of both populations is two and their difference is $\langle T \rangle$. For the example $U = 100,\ t = 10$ meV this gives spatial occupations roughly 1.371 and 0.629. Dividing by two gives the spin-orbital occupations for each spin in this singlet. Site populations alone are one on each site by symmetry at every U/t; their apparent constancy completely misses this changing form of correlation.

Break site symmetry and measure the onset of charge transfer

Add a static site-energy imbalance with energies $-\Delta/2$ on L and $+\Delta/2$ on R. The singly occupied configuration still has total site energy zero, while double occupations cost $U - \Delta$ and $U + \Delta$. For $|t| \ll U - |\Delta|$ the singly occupied singlet gains the second-order shift $-2t^2/(U - \Delta) - 2t^2/(U + \Delta)$, since its matrix element to either double occupation has magnitude $\sqrt{2}|t|$. The triplets remain blocked. The effective exchange splitting is $4t^2U/(U^2 - \Delta^2)$. At zero imbalance it reduces to the earlier $4t^2/U$; small imbalance changes it quadratically, consistent with exchanging left and right sites.

For $U = 100,\ t = 5,\ \Delta = 40$ meV, the estimate gives 1.19 meV compared with 1.00 meV at zero bias. Increasing imbalance can make one virtual double occupation especially inexpensive, changing the gap before the mean site charge substantially relocates. Near $\Delta = U$ the denominator vanishes and the perturbative model predicts a divergence. The true three-state singlet matrix instead has an avoided crossing of singly and doubly occupied configurations. A sweep through that region should retain both double occupations and compute site charge from $-2\,\partial E/\partial\Delta$, with the sign matched to the specified imbalance. Here an apparent runaway effective spin coupling marks the failure of eliminating charge degrees of freedom, rather than an arbitrarily large spin energy in the actual finite Hamiltonian.

Measure singlet–triplet exchange without a spin force

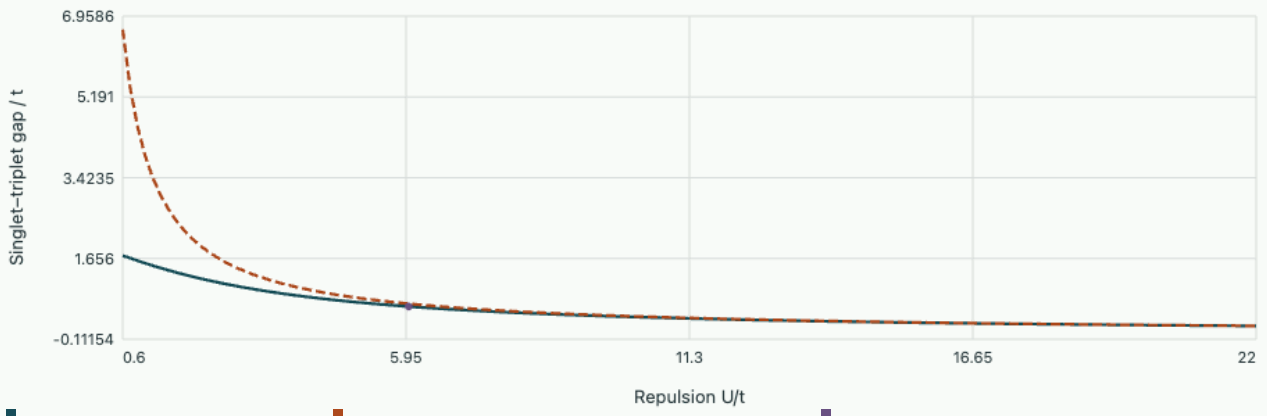
Change the on-site repulsion in a two-electron dimer. Check whether exchange inferred from the spectrum follows the virtual-hopping estimate.

Open this experiment online

Try the calculation

On-site energy U (meV) 6

Hopping t (meV) 1



Exact two-site singlet | Virtual hopping 4t/U in these units | Selected dimer

Triplet energy 0. Lower singlet energy -0.60555 meV. Estimated gap $4t^2/U$ = 0.66667 meV.

See exchange interference directly in a two-port experiment

A lossless balanced beam splitter for two orthonormal single-particle modes can be chosen to act on creation operators as $a^\dagger \to (c^\dagger + d^\dagger)/\sqrt{2},\ b^\dagger \to (c^\dagger - d^\dagger)/\sqrt{2}$. Bring two identical particles to the device, one per input, with wave-packet timing and internal states controlled. The input $a^\dagger b^\dagger|0\rangle$ contains no explicit interaction. Nevertheless the coincidence statistics after the device depend on exchange because two paths can lead to the same unordered output configuration. This provides a way to test exchange far from the Coulomb energy calculations of the previous sections.

For identical spinless bosons multiply the two sums and use commuting creation operators. The cross terms cancel, leaving $[(c^\dagger)^2 - (d^\dagger)^2]|0\rangle/2 = (|2,0\rangle - |0,2\rangle)/\sqrt{2}$. Both particles leave the same port and coincidence probability vanishes. The $\sqrt{2}$ in a normalised doubly occupied Fock state is essential. For identical fermions in the same internal state the repeated creation terms vanish and the cross terms add because $d^\dagger c^\dagger = -c^\dagger d^\dagger$. The outgoing state has one particle in each port with unit probability. Neither result requires an attractive or repulsive force at the splitter. Their interpretation turns on whether the alternatives remain indistinguishable and on the symmetry of the total input.

Now give the two particles internal wave packets $|u\rangle, |v\rangle$, which may encode time, frequency or spin, and suppose the detectors report output ports but do not resolve internal labels. One coincidence alternative puts u in c and v in d; the other interchanges them. The interference cross term contains the product $\langle u|v\rangle\langle v|u\rangle = |\langle u|v\rangle|^2$. At a balanced splitter the coincidence probability is $[1 - |\langle u|v\rangle|^2]/2$ for initially independently prepared bosons and $[1 + |\langle u|v\rangle|^2]/2$ for corresponding fermions. At zero overlap both reduce to one half, as expected for two distinguishable independent particles each taking a random output.

Take transform-limited Gaussian wavepacket intensities of common time variance σ_t^2 with delay τ between their centres and no carrier mismatch. Integrating their two normalised amplitude envelopes gives

$$|\langle u|v\rangle|^2 = \exp[-\tau^2/(4\sigma_t^2)].$$

For $\sigma_t = 0.50$ ps and delay 1.00 ps this is

$$e^{-1} = 0.368.$$

The predicted bosonic coincidence fraction is 0.316 and the fermionic value 0.684. A narrow time gate can itself identify the particles differently at different delays and must be included in the counting model. Measuring the single-output rates together with coincidences tests splitter balance and detection bias before converting dip or peak contrast into an overlap estimate.

Infer distinguishability with an unbalanced splitter

A calibrated splitter has intensity transmittance $T = 0.60$, reflectance $R = 0.40$ and negligible loss. Two photons from separate inputs give coincidence probability 0.184 at equal central arrival time. Infer their internal squared overlap. Would equal-time arrival alone have guaranteed perfect bunching?

▼ **Derive and check the result**

For indistinguishable paths the two coincidence amplitudes have opposite signs with magnitudes T, R. A nonresolving detector gives

$$P_c = T^2 + R^2 - 2TR|\langle u|v\rangle|^2.$$

Substitution yields squared overlap $(0.52 - 0.184)/0.48 = 0.70$. Even perfect internal overlap at this unbalanced splitter would leave $(T - R)^2 = 0.040$ coincidences. Equal arrival time controls only one possible distinguishing label; orthogonal polarisation or spectral mismatch can still reduce overlap. A far-delay reference should give $T^2 + R^2 = 0.52$ for distinguishable, otherwise comparable photons. This independent control helps separate exchange interference from lost counts or a miscalibrated splitter.

Tutorial · discriminate spatial exchange from virtual hopping

Normalise two nonorthogonal orbitals before predicting coincidences

Let a, b each have norm one but overlap $s = 0.30$. Find the normalisations of their symmetric and antisymmetric products. Does the antisymmetric state still vanish at common position?

▼ **Work through the solution**

The inner product $\langle ab|ba\rangle = |s|^2 = 0.09$. The normalisation denominators are

$$\sqrt{2(1+0.09)} = 1.476,$$

$$\sqrt{2(1-0.09)} = 1.349.$$

At identical positions the antisymmetric numerator cancels pointwise for every s. When the orbitals become identical its norm tends to zero and there is no independent antisymmetric state of that pair. For nonzero overlap the one-body cross terms in the earlier density calculation no longer integrate away, so those equal-density formulas must be rederived. Orthonormalising local orbitals changes the parameters of a corresponding Hubbard model as well.

Infer a hopping amplitude from a measured singlet gap

A calibrated two-site sample has $U = 120$ meV and a singlet–triplet gap 5.0 meV. In the model above find t exactly, predict the double-occupation probability and compare with a large-U inference.

▼ **Work through the solution**

Writing $J = (\sqrt{U^2 + 16t^2} - U)/2$ gives $4t^2 = J(U + J)$, hence $|t| = 12.50$ meV. The large-U estimate $\sqrt{JU}/2 = 12.25$ meV is close but measurably different. The double-occupation fraction is

$$\tfrac{1}{2}[1 - U/(U + 2J)] = J/(U + 2J) = 0.03846.$$

Gap alone cannot distinguish signs of t in a two-site model because rephasing one orbital reverses the sign. Measuring charge fluctuations independently tests whether the observed spin gap can be explained by this same virtual process.

Predict coherent double-occupation oscillations

Start suddenly in the singly occupied singlet of the symmetric two-site model. Find the probability of observing either doubly occupied site after time τ and its maximum at $U = 100,\ t = 10$ meV. Predict the triplet control.

▼ **Work through the solution**

Subtract $U/2$ times identity from the two-dimensional Hamiltonian. The traceless matrix squares to $(U^2 + 16t^2)I/4$. Exponentiating gives

$$P_D(\tau) = 16t^2 \sin^2[\sqrt{U^2 + 16t^2}\,\tau/(2\hbar)]/(U^2 + 16t^2).$$

Its maximum is

$$1600/11600 = 0.138.$$

This is larger than the ground-state admixture because the sudden preparation is a coherent superposition of both energy eigenstates. A triplet evolves with no doubly occupied component in this ideal Hamiltonian. Observing double occupation after triplet preparation would test spin mixing, state-preparation fidelity or additional orbitals.

Further reading

- Fließbach · Mehrteilchensysteme
- Cambridge · electron exchange, helium and determinants
- Course investigation · exchange of identical particles

Measure tunnelling from a four-site spectrum

Computational quantum experiments

Four localised orbitals form a minimal double well. Raise the inner two site energies and ask why two states concentrated in the outer wells still acquire different energies. Measure the splitting in the spectrum, predict an oscillation in time and reproduce the levels independently with OneSimulation.

Construct an operator before drawing an energy diagram

Label the sites $1, 2, 3, 4$ in order. The end orbitals have on-site energy zero, the middle two have energy $V > 0$ and every neighbouring pair has real hopping matrix element $-t$ with $t > 0$. The four occupation states form an orthonormal basis for one particle. Equal diagonal energies of the ends do not make their symmetric and antisymmetric combinations degenerate when there is a virtual path through the interior. There are no direct links between sites one and three or one and four. Keeping track of every allowed step will fix the power of t in the final answer.

$$H = \begin{pmatrix} 0 & -t & 0 & 0 \\ -t & V & -t & 0 \\ 0 & -t & V & -t \\ 0 & 0 & -t & 0 \end{pmatrix}.$$

Reflection maps $1 \leftrightarrow 4$ and $2 \leftrightarrow 3$, and commutes with H. In each parity sector use the normalised outer and inner basis vectors $(|1\rangle \pm |4\rangle)/\sqrt{2}$ and $(|2\rangle \pm |3\rangle)/\sqrt{2}$. The diagonal outer element vanishes in both sectors; the hopping inside the middle pair shifts its even and odd energies to $D_e = V - t$ and $D_o = V + t$. The outer–inner matrix element remains $-t$. This reduces the four-dimensional problem to two symmetric two-by-two matrices without neglecting any state.

Eliminate the barrier amplitudes and estimate the splitting

In a sector with interior energy D, the secular equation is

$$E(E - D) = t^2.$$

The lower root is $E_-(D) = [D - \sqrt{D^2 + 4t^2}]/2$, or the numerically better $-2t^2/(D + \sqrt{D^2 + 4t^2})$. For $D \gg t$ the shift is approximately $-t^2/D$. It is more negative for the even state because $D_e < D_o$. Subtract the two shifts before approximating their difference.

$$\Delta E = E_-(D_o) - E_-(D_e) \simeq t^2 \left(\frac{1}{V - t} - \frac{1}{V + t} \right) \simeq \frac{2t^3}{V^2}.$$

A single outward excursion produces an order-t^2/V shift in both end states. Only a three-hop path connects the different ends, so their splitting is smaller. Measuring an overall displacement of two spectral lines and measuring the distance between them test different physical mechanisms.

Predict a time trace and define what is measured

Denote the two low-energy eigenstates by $|e\rangle, |o\rangle$. Prepare $(|e\rangle + |o\rangle)/\sqrt{2}$. In the corresponding opposite superposition $(|e\rangle - |o\rangle)/\sqrt{2}$, the occupation at time τ equals $\sin^2(\Delta E\,\tau/2\hbar)$. Its first maximum is at $\pi\hbar/\Delta E$. These superpositions include interior weight, especially at low barrier. A detector on site four alone measures a different projector. Preparing bare site one excites both high and low doublets and can generate additional fast oscillations.

For $V/t = 4$, the two lower eigenvalues in units of t are $(3 - \sqrt{13})/2$ and $(5 - \sqrt{29})/2$, about −0.30278 and −0.19258. Their separation is about $0.11019t$. The high-barrier approximation gives $0.125t$. With $t = 1\,\mathrm{meV}$, $\hbar = 0.65821\,\mathrm{meV\,ps}$ and the first complete transfer in the dressed basis takes about 18.8 ps. If a measurement integrates over many such oscillations, it may return one half despite perfectly coherent evolution.

Compare spectra from three independent constructions

Use the parameter controls to vary the middle barrier and hopping. The plot shows the exact splitting and its leading virtual-hopping prediction over a range of V/t. The printed result gives an actual transfer time. The downloadable OneSimulation model builds the original four-site Hamiltonian and solves its eigenvalue problem by contour integration. Compare all four eigenvalues with the two parity blocks before interpreting their small difference. A wrong sign on one edge of this open chain can be removed by a local change of basis and may leave its spectrum unchanged, so also check any phase-sensitive observable used for preparation.

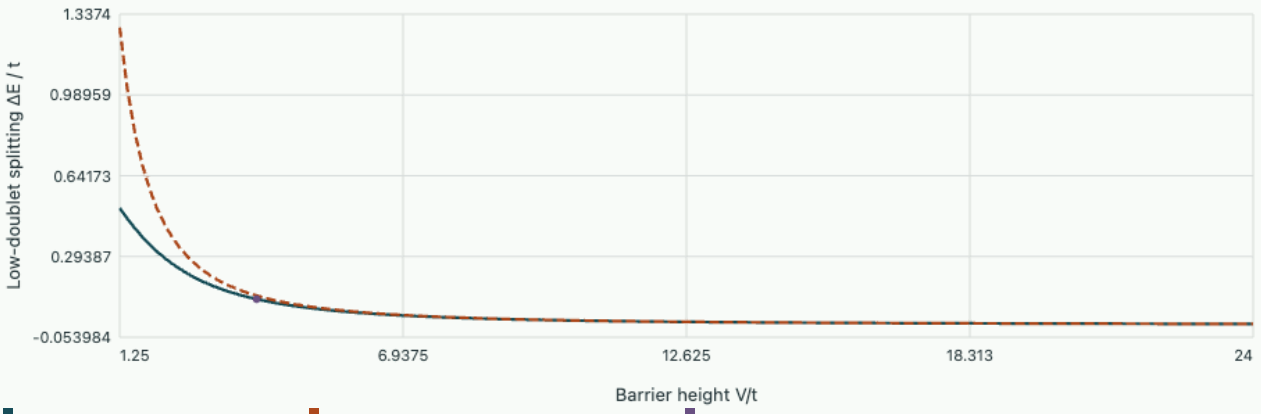
Keep the hopping fixed at 1 meV and compare two barriers. Start with a 4 meV barrier and raise it to 8 meV. Record both splittings and estimate the change in the first transfer time before reading off the prediction.

Open this experiment online

Try the calculation

Middle site energy V (meV) 4

Hopping t (meV) 1



Exact parity sectors | Three-hop approximation | Selected barrier

Splitting 0.11019 meV. First transfer between dressed end states at 18.766 ps.

Download the OneSimulation model

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

```
python spectra-and-orbits.py wells --barrier 4 --hopping 1
```

Tutorial

Hold V fixed and halve t in the high-barrier limit. How do the common shift, splitting and transfer time change?

▼ **Work through the solution**

The leading common shift scales as t^2/V and falls by four. The splitting scales as t^3/V^2 and falls by eight. The transfer time grows eightfold. Halving both V and t instead keeps their ratio fixed and halves every eigenvalue.

Explain why a small static asymmetry between the end wells can suppress almost complete transfer while increasing their measured frequency splitting.

▼ **Work through the solution**

In a two-state reduction, detuning δ and coupling J give separation $\sqrt{\delta^2 + 4J^2}$. Starting in one localised state, the maximum transfer probability is $4J^2/(\delta^2 + 4J^2)$. Both the frequency and the contrast are needed to infer the coupling.

How could one check that a contour solver has found both members of a closely spaced doublet rather than printing one twice?

▼ **Work through the solution**

Check the rank of the independent projected vectors and the residual norm for each vector. Then exploit reflection parity or solve in each symmetry sector independently. Repeating a low residual for two nearly parallel vectors does not establish a two-dimensional eigenspace.

The continuous barrier calculation provides a different model in which barrier width eventually produces an exponential dependence.

Applied study

Use lattice symmetry to connect vibrations, electronic bands and gaps in a periodic optical material.

From a row of atoms to a dispersion relation

Why can a crystal carry waves of some frequencies but exclude others? Start with the simplest solid. Identical masses m sit at sites $x_j = ja$, where a is the lattice spacing. Each atom is coupled to its nearest neighbours by springs of stiffness K. Small longitudinal displacements u_j obey

$$m\ddot{u}_j = K(u_{j+1} - 2u_j + u_{j-1}).$$

A displacement pattern can reproduce itself after a translation by one cell, apart from a phase. Insert $u_j = Ue^{i(kja-\omega t)}$ and cancel the common factor. This yields

$$\omega^2 = (4K/m)\sin^2(ka/2).$$

Two wave numbers that differ by $2\pi/a$ assign the same phase to every lattice site. One interval, the first Brillouin zone $-\pi/a \le k \le \pi/a$, therefore contains all distinct patterns.

For long waves the dispersion relation becomes

$$\omega \simeq a\sqrt{K/m}\,|k|.$$

A narrow wave packet travels at the group velocity $d\omega/dk$. This speed approaches the sound speed near the zone centre and falls to zero at the edge. There the neighbours move oppositely in a standing pattern.

How a periodic perturbation opens a gap

In a periodic potential a single electron has Bloch states $\psi_k(x) = e^{ikx}v_k(x)$, with $v_k(x+a) = v_k(x)$. For a free electron, the waves at $+\pi/a$ and $-\pi/a$ have the same energy E_0. Their wave numbers differ by a reciprocal lattice vector. A periodic potential can mix them.

Retain these two waves and absorb the mean potential into E_0. The diagonal entries of their Hamiltonian are both E_0. Its off-diagonal entries are V_G and its complex conjugate, the Fourier components of the potential that couple the waves. The characteristic equation $(E_0 - E)^2 - |V_G|^2 = 0$ gives

$$E_\pm = E_0 \pm |V_G|.$$

A coupling of 0.12 eV opens a 0.24 eV splitting at this crossing. The two standing waves place their probability density differently relative to the ions.

Band filling matters. At zero temperature a band with available neighbouring states at the Fermi energy can respond to a weak electric field. An ideal crystal with a full valence band separated from empty states by an energy gap needs excitations across that gap for electronic conduction. To infer such a gap in a real material, check the band extrema throughout the zone.

Read a gap in an optical lattice

Light in alternating transparent layers also obeys a Bloch condition. A transfer matrix maps the electric and magnetic fields across one complete cell. For a lossless cell at normal incidence its determinant is one, so the two eigenvalues are reciprocal. A propagating wave has eigenvalues $e^{\pm ika}$, hence half the matrix trace equals $\cos(ka)$. An absolute half-trace greater than one leaves no real wave number at that frequency. The field then decays into a thick stack.

Consider equal layers with the same relative permittivity 13 and relative permeability one. They behave as a uniform medium with $n = \sqrt{13}$. At the zone edge $2\pi n f/c = \pi/a$, so the reduced frequency $fa/c = 1/(2\sqrt{13}) = 0.1387$. The folded branches touch. Give the second layer a different refractive index and light reflects at each interface. Near the zone edge successive reflections add coherently, lifting this degeneracy. This gap is an interval of optical frequencies. The electronic example above measures an interval of energies.

Problem class

Infer a bond stiffness from a sound speed

A model crystal has spacing 0.30 nm, site mass 4.0×10^{-26} kg and measured long-wave speed 3000 m/s. Estimate the spring stiffness and the frequency in Hz at the zone edge.

▼ **Work through the lattice estimate**

Since $v_s = a\sqrt{K/m}$, the stiffness is

$$K = m(v_s/a)^2 = 4.0\,\mathrm{N/m}.$$

At the zone edge $\omega = 2\sqrt{K/m} = 2.0 \times 10^{13}$ rad/s, giving $f = 3.18$ THz. The same spring scale determines the sound speed and the upper end of this model branch.

Test which waves a stack can support

The half-trace of a lossless cell is 0.50 at one frequency and –1.10 at another. Which wave propagates? At the second frequency estimate how many cells the decaying field needs to lose a factor e in amplitude.

▼ **Use the Bloch multiplier**

The first frequency admits $ka = \pm\pi/3$. For the second, set $ka = \pi + i\gamma$. Then $\cos(ka) = -\cosh\gamma = -1.10$, giving $\gamma = \operatorname{arcosh}(1.10) = 0.444$. After N cells the decaying amplitude is multiplied by $e^{-N\gamma}$. The decay length is $1/\gamma = 2.25$ cells. Intensity decays twice as fast on a logarithmic scale.

ONESIMULATION LABORATORY

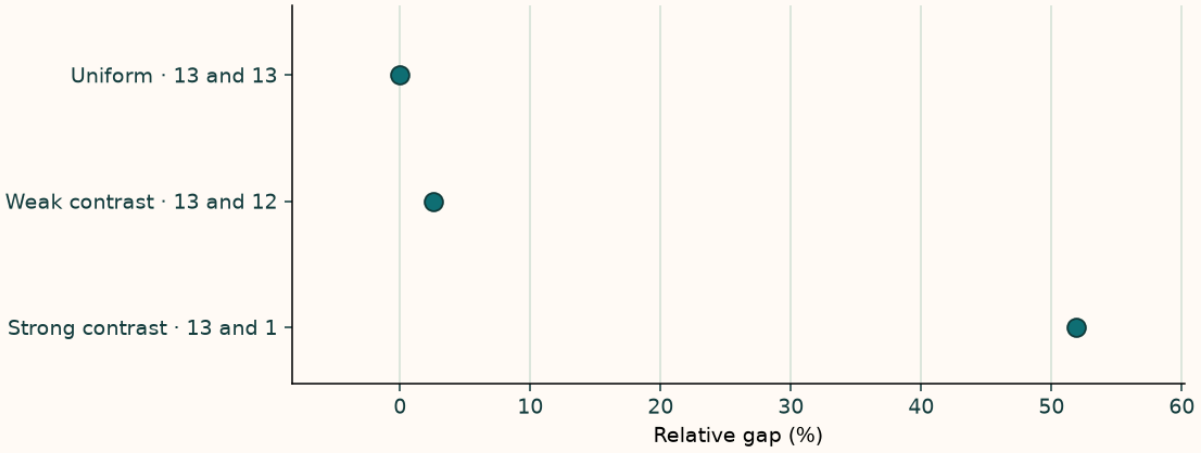
Open a band gap in a periodic dielectric

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.

Use one periodic cell and keep its layer widths fixed. The finite-element calculation solves the electromagnetic eigenproblem for several Bloch wave numbers. Change the second relative permittivity in `cases`. Predict how the two lowest zone-edge frequencies separate as the index contrast increases.

OneSimulation zone-edge frequencies, expressed as fa/c. Relative gap width is the gap divided by its midfrequency.

Layer permittivities	Lower edge	Upper edge	Relative gap
Uniform · 13 and 13	0.1387	0.1387	0.00%
Weak contrast · 13 and 12	0.1396	0.1433	2.55%
Strong contrast · 13 and 1	0.1509	0.2566	51.89%



The relative photonic gap rises from zero in the uniform medium to 2.55% at weak contrast and 51.89% at strong contrast.

Replace the existing `weak contrast` pair with `(13.0, 6.0)` and rerun the notebook. Compare the FEM edges with the transfer-matrix calculation included in the file. Then set `a = 2e-3` in the supplied geometry script, regenerate both STEP files and rerun with the same layer fractions. The reduced frequencies stay fixed, so the frequencies in Hz halve. Inspect the computed band diagrams.

Download experiment files (ZIP)

Extract the files together and open `one_dimensional_photonic_crystal_bands.ipynb` in a local OneSimulation environment.

Try it · Opening the Bloch gap in equal-width layers

Both layers occupy half the physical period, exactly as in the OneSimulation cell. Keep the first permittivity at 13 and change the second from 13 to 12 and then 1 to compare the uniform, weak and strong contrasts in the table. The transfer matrix gives the reference band edges; doubling the period halves their physical frequencies.

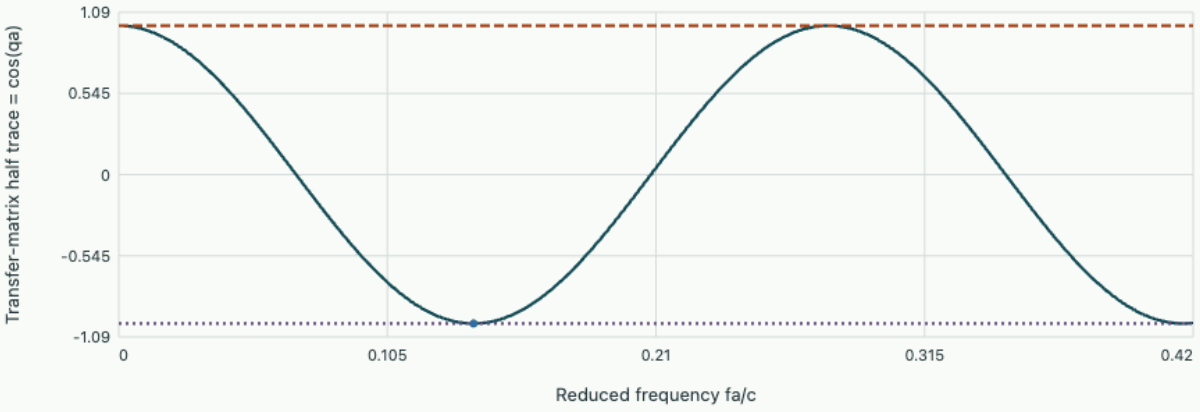
Open the corresponding OneSimulation notebook, code and output

Open this live calculation

Try the calculation

Second-layer relative permittivity **13**

Physical period a (mm) **1**



Zone-edge frequencies fa/c = 0.13868 and 0.13868; gap / midgap 0.000%. Physical edges 41.574 and 41.574 GHz. Exact transfer matrix for $\varepsilon_1 = 13$ and equal half-period layers, matching the OneSimulation Bloch cell; |half trace| > 1 is a gap. Scaling the period changes physical frequency but not these reduced curves.

Problem sheet 3 · Reconstruct the edges of a photonic gap

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.

Band velocity

$$\mathbf{v}_n(\mathbf{k}) = \hbar^{-1}\nabla_{\mathbf{k}} E_n$$

This is the group velocity of a band wave packet in the ordinary semiclassical approximation. The Fermi surface determines which states can respond at low temperature.

Derivation · Bloch waves, band edges and effective mass

Cyclotron mass

$$m_c = (\hbar^2/2\pi)\,\partial A/\partial E$$

The area is that of a closed constant-energy orbit in reciprocal space normal to the magnetic field. This is a measured orbital mass rather than a universal free-particle parameter.

Derivation · Counting carriers and weighing a cyclotron orbit

Density of states

$$g(E) = dN_{\text{states}}/dE$$

State density and state occupation are independent factors. Their product, integrated over a band, gives the mean particle number.

Derivation · Filling electronic states and measuring the Fermi surface

Fermi level

$$f(E) = 1/[1 + \exp((E-\mu)/k_B T)]$$

At equilibrium a one-electron level at the chemical potential has mean occupancy one half per available state. Spin multiplicity enters the number of states separately.

Derivation · Filling electronic states and measuring the Fermi surface

London length

$$B(x) = B(0)e^{-x/\lambda_L}$$

This is a field profile in a thick homogeneous superconductor in a linear screening regime. A thin slab requires both boundary conditions.

Derivation · Superconductivity, magnetic screening and flux quantisation

Reciprocal lattice

$$\mathbf{G} \cdot \mathbf{R} = 2\pi n$$

A reciprocal-lattice vector leaves the phase unchanged under every direct-lattice translation. A basis of different atoms can still suppress a reflection.

Derivation · Translational symmetry and diffraction

Further reading

- Compare the reciprocal-lattice construction with the nearly-free-electron picture. Return to the lattice-vibration chapters to contrast phonon and electronic densities of states. David Tong · Solid State Physics

Compare specific topics with the university sources