Engineering thermodynamics and heat transfer

Apply energy, entropy and exergy balances to steam cycles, heat exchangers and enclosures, and derive how heat transfer sets operating temperatures.

Read the subject in depth

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

- ### Thermal processes and transport in engineering

  Close energy and entropy balances around real processes. Use measured states and heat flows to separate component performance from modelling assumptions.

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. ### Properties, equilibrium and phase change

   Predict the sign of a response from a potential, then track what a measured fraction or coexistence curve says about the state of the material.

2. ### Transport, heat sources and diagnostics

   Separate sensible heating from transported or radiated energy. Test the temperature signal against a field loss measurement and a conservative heat discretisation.

How the course fits together

1. **Chapters 1–4**

   Choose first a closed system, then a flowing control volume. Energy balances determine work and heat, entropy restricts their conversion, and returning to the initial state creates a cycle whose performance can be measured.

2. **Chapters 5–8**

   A thermal resistance sets a steady heat rate. Adding storage makes temperatures time-dependent. Phase change and moist air then require tracking composition as well as sensible energy.

3. **Chapters 9–12**

   Use a reference environment to measure available work. Follow water around a steam cycle, transfer energy between two moving streams, and close the course with radiative exchange between surfaces.

Contents

1. Chapter 1 · State, process and the first law for a closed system
2. Chapter 2 · Open systems, enthalpy and steady-flow devices
3. Problem sheet 1 · Energy accounts for stored and flowing matter
4. Chapter 3 · Entropy, irreversibility and useful-work loss
5. Chapter 4 · Thermodynamic cycles, refrigeration and phase change
6. Problem sheet 2 · Entropy production and performance limits
7. Chapter 5 · Steady conduction, thermal resistance and insulation
8. Chapter 6 · Thermal diffusion and transient surface cooling
9. Problem sheet 3 · Heat paths and temperature histories
10. Chapter 7 · Phase equilibrium and steam quality
11. Chapter 8 · Gas mixtures, humidity and condensation
12. Problem sheet 4 · Saturation and moist air
13. Chapter 9 · Useful work and exergy destruction
14. Chapter 10 · Steam power cycles and turbine efficiency
15. Problem sheet 5 · Exergy and steam-cycle performance
16. Chapter 11 · Heat-exchanger effectiveness and finite capacity rates
17. Chapter 12 · Thermal radiation and enclosure exchange
18. Problem sheet 6 · Heat exchangers and radiation shields
19. Chapter 13 · Choosing a pressure ratio for a gas-turbine cycle
20. Chapter 14 · Why a real gas cools or warms across a throttle
21. Chapter 15 · When another centimetre of cooling fin stops helping
22. Chapter 16 · A moving boundary in a freezing layer
23. Chapter 17 · Diffusion with a stationary carrier gas
24. Chapter 18 · Reading a vapour-compression refrigerator from its four enthalpies
25. Chapter 19 · Finding the vapour fraction of an equilibrium flash
26. Chapter 20 · Where a thermoelectric cooler rejects its heat
27. Problem sheet 7 · Budgeting cooling power for a freezing experiment
28. Problem sheet 8 · Separate heat capacity, contact conductance and leakage in a calorimeter
29. Thermal processes and transport in engineering
30. Properties, equilibrium and phase change
31. Transport, heat sources and diagnostics
32. Applied study
33. Notation and assumptions
34. Further reading

Chapter 1 · State, process and the first law for a closed system

Distinguish properties from path-dependent transfers and determine heat and work for a piston-cylinder process.

FROM FIRST PRINCIPLES

Obtain closed-system work from the motion of a boundary

Choose a fixed mass of gas enclosed by a piston and cylinder. No mass crosses this boundary. The first law is the empirical balance of energy: with heat positive into the gas and work positive out, change of internal, kinetic and potential energy equals net heat minus net work. For a piston moving normally through dx against a uniform resisting pressure p_{ext}, force times displacement gives

$$\delta W_b = p_{\text{ext}} A\, dx = p_{\text{ext}}\, dV,$$
$$\Delta(U + KE + PE) = Q - W.$$

For slow internal equilibration with tiny pressure differences, one can use gas pressure in the integral. If bulk kinetic and potential changes are negligible and no shaft or electrical work crosses the boundary, the integrated heating requirement is

$$Q_{12} = U_2 - U_1 + \int_{V_1}^{V_2} p_{\text{ext}}\, dV.$$

The area under the specified pressure–volume path is therefore part of the heat calculation. For a rigid vessel it vanishes. For a constant-pressure piston heating an ideal gas, $p\Delta V = mR\Delta T$ adds to the internal-energy change. The end temperatures alone determine the ideal-gas internal-energy change, while the moving boundary determines the work.

DERIVATION AND APPLICATIONS

Determine the heat input when a spring loads the piston

A freely sliding piston does not necessarily hold the gas pressure constant. Attach a linear spring to a frictionless piston of area A and let its resisting force increase as the gas expands. Consider a closed mass m of ideal gas, heated slowly enough that piston force balance holds at each intermediate state. Neglect piston kinetic energy, heat leakage and changes of elevation except as already included in the constant external pressure p_1. At the initial position the gas is in equilibrium at absolute pressure p_1 and volume V_1.

The displacement x produces a volume increment Ax and an extra resisting force kx. Dividing by piston area gives the pressure as a function of the evolving volume:

$$V - V_1 = Ax,$$
$$p(V) = p_1 + \frac{k}{A^2}(V - V_1).$$

Boundary work is performed at this changing external pressure. Integrating the straight-line path over the swept volume $\Delta V = V_2 - V_1$ separates work against the initial pressure from work stored in the spring:

$$W_{12} = \int_{V_1}^{V_2} p(V)\,dV = p_1 \Delta V + \frac{k}{2A^2}(\Delta V)^2.$$

For constant specific heat c_v and specific gas constant R, the first law gives the heat required, positive into the gas. State temperatures follow from the ideal-gas relation, including the final spring pressure:

$$Q_{12} = mc_v(T_2 - T_1) + W_{12},$$
$$T_i = p_i V_i/(mR).$$

Only the temperature change determines the internal-energy change; the mechanical law selects the work term. Setting $k = 0$ recovers expansion against constant pressure. Abrupt heating that accelerates the piston would invalidate the assumed quasistatic pressure path.

WORKED EXAMPLE

Calculate the heater input for an air cylinder

Change the inputs and test this calculation.

Take $m = 0.0100$ kg, $V_1 = 0.0100$ m³, $p_1 = 100$ kPa, $A = 0.0100$ m² and $k = 1000$ N/m. Increase volume by 0.00200 m³. The final pressure is 120 kPa; the work is

$$(100 + 120)(0.00200)/2 = 0.220\,\mathrm{kJ}.$$

For $R = 0.287$ and $c_v = 0.718$ kJ/(kg K), $\Delta T = (1.440 - 1.000)/(0.0100 \times 0.287) = 153$ K. The gas receives $1.10 + 0.220 = 1.32$ kJ of heat.

SECOND APPLICATION

Separate actuator work from stored spring energy

Change the inputs and test this calculation.

On the same stroke the piston travels 0.200 m. The initially unloaded incremental spring stores $kx^2/2 = 20.0$ J, exactly the triangular part of the pressure–volume work. The constant load receives 200 J. If the spring were removed while retaining the same initial state and swept volume, the final pressure would stay at 100 kPa and the temperature rise would be only 69.7 K. The two processes end at different thermodynamic states despite having the same final volume.

Try it · Spring piston

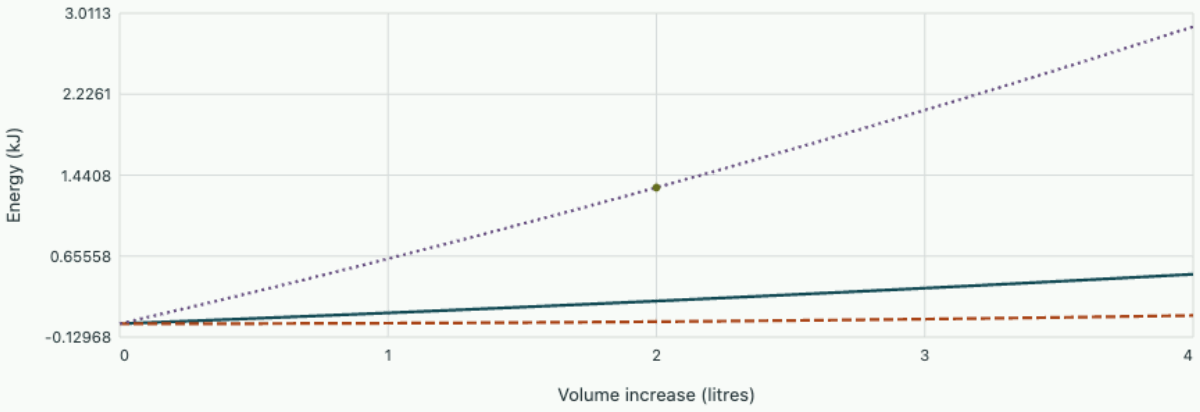
Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Spring stiffness (N/m) **1000**

Volume increase (litres) **2**



Air mass 0.010 kg, initial p=100 kPa, V=10 litres, piston area 0.010 m². Selected final pressure 120 kPa; temperature rise 153.31 K; work 0.22 kJ; heat 1.32077 kJ. Quasistatic ideal gas, cv=718 J/(kg K).

Decide what crosses the boundary

Draw a boundary around the gas in a piston cylinder. If no gas crosses its surface, the system is closed even though the piston moves. A rigid vessel is a special closed system with fixed volume. An insulated wall prevents heat transfer; a shaft or electrical lead can still transmit work through it. These descriptions impose different constraints and should be identified before choosing an equation.

At equilibrium a simple compressible substance can be described using independent state variables such as temperature and specific volume. Pressure p, volume V, temperature T and internal energy U are properties. An equation of state relates some of them. For a dilute ideal gas $pV = mRT$, where R is the specific gas constant. Always use absolute pressure and absolute temperature in this relation. Two

initial and final states determine changes in properties, but do not specify the heat and work transfers between them.

Construct the process energy balance

Take work done by the system as positive and heat entering it as positive. The first law is

$$\Delta U + \Delta KE + \Delta PE = Q - W.$$

For a stationary vessel ignore changes in kinetic and gravitational energy of the system as a whole. A moving piston of area A against external pressure p_{ext} performs incremental work $p_{\text{ext}} A\, dx = p_{\text{ext}} dV$. For a quasistatic mechanically reversible motion the boundary pressure can be represented by the gas pressure throughout the process.

If gas expands from V_1 to V_2 at constant external pressure,

$$W = p_{\text{ext}}(V_2 - V_1).$$

If it instead follows a reversible isothermal ideal-gas path, inserting $p = mRT/V$ and integrating gives

$$W = mRT \ln(V_2/V_1).$$

Both expressions have units of energy. Their different volume dependence makes a pressure-volume diagram useful: the area under each pressure path has a direct work interpretation when its boundary-force assumptions hold.

Background · Which pressure enters the boundary-work integral?

The traction at a moving boundary determines the mechanical work across it. For a piston use the pressure exerted against the load. Replacing it by a uniform equilibrium gas pressure additionally assumes slow mechanical evolution. A limiting comparison makes this distinction measurable.

$$W_{\text{by}} = \int_{V_1}^{V_2} p_{\text{ext}}\, dV.$$

Remove a partition in an insulated rigid vessel so that an ideal gas expands into a vacuum. The external pressure against the advancing gas is zero. There is no heat transfer or boundary work, so the gas returns to the same equilibrium temperature. A reversible isothermal expansion between those same gas volumes requires heat from a bath and yields positive work $nRT \ln(V_2/V_1)$. The two endpoint states alone do not define a work measurement.

Heat air below a freely moving piston

A weight-loaded, frictionless piston keeps 0.50 kg of ideal air at 150 kPa absolute as a heater raises its temperature from 300 to 360 K. Use constant properties $R = 0.287$ and $c_v = 0.718$ kJ/(kg K). The ideal-gas law gives the volumes at the two ends of the process.

$$V_1 = \frac{mRT_1}{p} = 0.2870\ \mathrm{m^3},$$
$$V_2 = \frac{mRT_2}{p} = 0.3444\ \mathrm{m^3}.$$

The pressure is constant, so the work is the area of a rectangle on a pressure–volume diagram. Internal energy depends on the gas temperature. These are separate contributions to the heat input.

$$\begin{aligned} W &= p(V_2 - V_1) = 150(0.0574)\ \mathrm{kJ} = 8.61\ \mathrm{kJ}, \\ \Delta U &= mc_v(T_2 - T_1) = 0.50(0.718)(60)\ \mathrm{kJ} \\ &= 21.54\ \mathrm{kJ}. \end{aligned}$$

The required heat is

$$Q = \Delta U + W = 30.15\,\mathrm{kJ}.$$

The same result follows from $mc_p\Delta T$ because for this constant-pressure ideal-gas process with only boundary work $c_p = c_v + R = 1.005$ kJ/(kg K). The extra heat above the internal-energy increase moved the piston weight. Using $Q = mc_p\Delta T$ without inspecting the constraint would give the wrong heat input to a rigid tank.

Compare a rigid-wall experiment

Lock the piston at its initial volume before heating an identical mass from 300 to 360 K. Now $W = 0$, so only 21.54 kJ of heat is needed. The final pressure is

$$150(360/300) = 180\,\mathrm{kPa}.$$

Equal temperature changes in an ideal gas imply equal changes in internal energy, even though final pressures and energy transfers differ. For a real fluid near a phase boundary, constant ideal-gas heat capacities may fail; use tabulated state properties or a suitable equation of state. In either case the first-law accounting is unchanged once the initial and final internal energies and the actual boundary work are known.

A different path between specified volumes

An ideal gas initially at 300 kPa and 0.020 m³ expands quasistatically and isothermally to twice its initial volume. Since $pV = p_1V_1$, work out is

$$W = p_1 V_1 \ln(V_2/V_1) = 6 \ln 2 = 4.16\,\mathrm{kJ}.$$

Ideal-gas internal energy depends only on temperature, so $\Delta U = 0$ and heat entering equals 4.16 kJ. A freely expanding ideal gas in an insulated rigid outer vessel can reach the same final temperature and occupied volume while doing zero boundary work. The state variables alone do not record which process occurred.

Determine heat and work along a measured compression path

A measured pressure–volume trace contains more information than its endpoints. For some slow piston processes an empirical polytropic law $pV^n = \text{constant}$ fits the path. Here n is a fitted process exponent and need not equal the gas heat-capacity ratio. Assume a closed cylinder with negligible kinetic and potential energy change, a quasistatic pressure load and a calorically perfect ideal gas. Work out of the gas is the area $W = \int_{V_1}^{V_2} p\,dV$, which is negative for compression. Substitute $p = p_1 V_1^n V^{-n}$ and integrate:

$$W = \frac{p_2 V_2 - p_1 V_1}{1 - n} \quad (n \neq 1),$$

$$W = p_1 V_1 \ln(V_2/V_1) \quad (n = 1).$$

The logarithmic case is the continuous limit as n approaches one. The ideal-gas equation also gives $T_2/T_1 = (V_1/V_2)^{n-1}$, so $\Delta U = m c_v (T_2 - T_1)$. With $Q - W = \Delta U$ and $pV = mRT$, the heat input is $Q = m[c_v + R/(1-n)](T_2 - T_1)$ for $n \neq 1$. Setting $n = \gamma = c_p/c_v$ eliminates Q. This inference relies on the ideal-gas, constant-property, quasistatic assumptions. An arbitrary irreversible compression with the same endpoints cannot be called isentropic from the fitted exponent alone.

Recover a cylinder heat loss from three measurements

A test traps air at $p_1 = 100\,\mathrm{kPa}$ absolute, $V_1 = 0.0100\,\mathrm{m^3}$, $T_1 = 300\,\mathrm{K}$, then compresses it to $V_2 = 0.0040\,\mathrm{m^3}$. The fitted exponent is 1.25. Using $R = 0.287,\ c_v = 0.718\,\mathrm{kJ/(kg\,K)}$, the trapped mass is

$$p_1 V_1/(RT_1) = 0.01161\,\mathrm{kg}.$$

The predicted pressure is

$$100(2.5)^{1.25} = 314.4\,\mathrm{kPa},$$

$$T_2 = 300(2.5)^{0.25} = 377.2\,\mathrm{K}.$$

Note that absolute pressure is required in both the fit and the gas law. A pressure transducer referenced to atmosphere would need its reference restored.

The gas does $W = [314.4(0.0040) - 100(0.0100)]/(1 - 1.25) = -1.03\,\mathrm{kJ}$. Its internal energy rises by $0.01161(0.718)(77.2) = 0.644\,\mathrm{kJ}$, giving $Q = \Delta U + W = -0.386\,\mathrm{kJ}$. Compression work enters the gas while part leaves as heat through the cylinder wall. If the process were isothermal at 300 K, the same volume change would require only $-W = 0.916\,\mathrm{kJ}$ and the full amount would leave as heat. A mid-stroke measurement gives a stronger check than the end pressure: the fitted law predicts $100(0.010/0.007)^{1.25} = 156.2\,\mathrm{kPa}$ at 7.0 litres. A substantial discrepancy would favour direct numerical integration of the measured trace over using a single exponent.

An apparent exponent can also be estimated directly from pressure ratios:

$$n = \ln(p_2/p_1)/\ln(V_1/V_2).$$

Compare estimates from the first and second halves of the stroke. Agreement means the same log–log slope describes both halves, while disagreement may indicate changing heat transfer, leakage or phase lag between the pressure and position sensors. It does not identify which of those mechanisms is responsible. As a separate check, suppose a water jacket gains the estimated 0.386 kJ each stroke and the compressor completes ten such strokes per minute. The mean heat removal from these strokes would be $0.386(10)/60 = 0.0643\,\mathrm{kW}$. A jacket measurement of the complete machine includes intake and exhaust processes and mechanical friction as well, so experimental comparison requires either an isolated compression interval or a balance around the rest of the cycle.

A change in operating speed can also change the fitted exponent because it changes the time available for heat exchange during a single compression stroke.

Chapter 2 · Open systems, enthalpy and steady-flow devices

Use an inlet-outlet balance to distinguish shaft work, kinetic-energy conversion and pressure reduction.

FROM FIRST PRINCIPLES

Derive the steady-flow energy balance including flow work

Surround a turbine, nozzle or mixing chamber with a stationary control surface. In addition to transporting its internal, kinetic and gravitational energies, an entering fluid is pushed across the inlet by upstream pressure. A mass increment dm occupies volume $v\,dm$ and requires pressure work $pv\,dm$ to enter. The outlet carries the corresponding work away. Combining internal energy u and flow work defines enthalpy:

$$h = u + pv,$$

$$e_{\text{stream}} = h + V^2/2 + gz.$$

Here V denotes speed, v specific volume and z elevation. In steady operation, sum these stream contributions and include heat into and shaft work out of the control volume. For one inlet and one outlet with equal mass rate, rearrangement gives

$$\dot{Q} - \dot{W}_s = \dot{m}\,[h_2 - h_1 + (V_2^2 - V_1^2)/2 + g(z_2 - z_1)].$$

All terms have units of power. A nozzle can accelerate flow without shaft work by lowering static enthalpy. A throttle with no heat, no shaft work and negligible speed and elevation changes is instead approximately isenthalpic. Omitting the pressure work while also using internal energy would miss both conclusions.

Background · Specific enthalpy

$$h = u + pv$$

Specific quantities are per unit mass. Enthalpy accounts for internal energy and flow work when mass passes through a control surface.

DERIVATION AND APPLICATIONS

Convert a nozzle energy balance into a mass-flow limit

A steady nozzle converts static enthalpy into bulk kinetic energy. A converging passage has a maximum possible mass flow at a specified inlet stagnation state. Derive this for a steady, one-dimensional,

adiabatic ideal gas with constant c_p and γ, reversible acceleration and negligible inlet-reservoir velocity. Write reservoir conditions as T_0, p_0 and the local static conditions as T, p. With no shaft work the energy balance gives

$$c_p T_0 = c_p T + u^2/2,$$
$$u = \sqrt{2 c_p T_0 (1 - T/T_0)}.$$

Isentropic acceleration links temperature and pressure. Let $r = T/T_0$, so $p/p_0 = r^{\gamma/(\gamma-1)}$. The mass flux is ρu and ideal-gas density is $p/(RT)$. Substitution gives a function of just one local variable:

$$G(r) = \frac{p_0}{R T_0} \sqrt{2 c_p T_0}\, r^{1/(\gamma-1)} \sqrt{1 - r}.$$

To find its maximum differentiate $\ln G = (\ln r)/(\gamma - 1) + \ln(1 - r)/2 + \text{constant}$. Equating the two derivative terms gives the sonic temperature and pressure ratios:

$$r_* = \frac{2}{\gamma + 1},$$
$$p_*/p_0 = \left(\frac{2}{\gamma + 1}\right)^{\gamma/(\gamma-1)},$$
$$u_*^2 = \gamma R T_*.$$

The flow through a converging exit is choked when downstream pressure is low enough for this state to be reached. Further lowering downstream pressure changes the flow after the exit; the ideal mass flux at that exit remains fixed. The model omits wall friction, heat transfer and discharge-coefficient corrections.

WORKED EXAMPLE

Size a sonic air metering aperture

Change the inputs and test this calculation.

For $p_0 = 200$ kPa absolute, $T_0 = 300$ K, $\gamma = 1.40$ and $R = 287$ J/(kg K), the critical state has $T_* = 250$ K and $p_* = 105.7$ kPa. Then $\rho_* = 1.47$ kg/m³ and $u_* = 317$ m/s, giving $G_* \simeq 467$ kg/(m² s). A 20.0 mm² effective aperture meters about 9.33 g/s if the downstream pressure allows choking.

SECOND APPLICATION

Decide whether a higher receiver pressure matters

Change the inputs and test this calculation.

Raise the receiver pressure to 160 kPa. Its ratio to reservoir pressure is now 0.800, above the critical 0.528. An ideal converging nozzle discharges at this pressure with $r = 0.800^{(0.4/1.4)} = 0.938$. The

energy equation gives $u \approx 193$ m/s, static density about 1.98 kg/m³, and mass flux about 382 kg/(m² s). Through the same aperture the predicted flow falls to 7.64 g/s. A model assuming choked flow at both settings would miss this change.

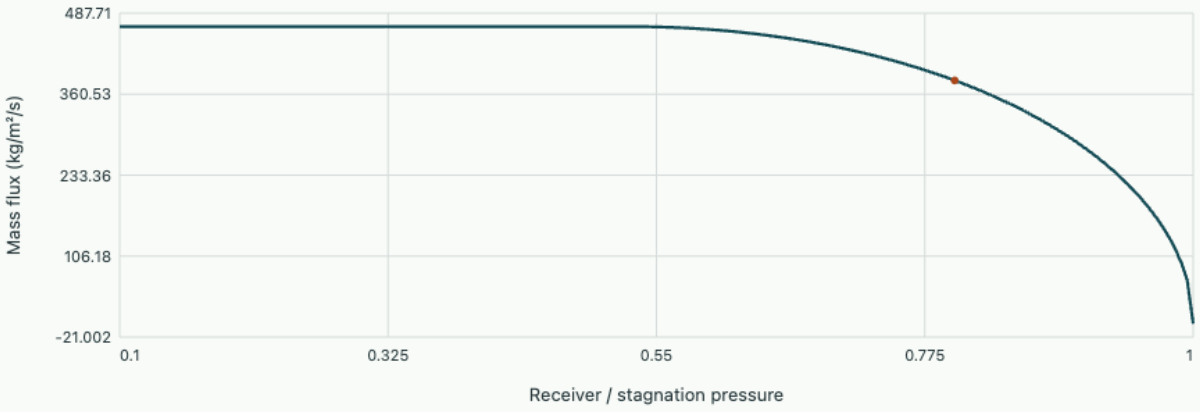
Try it · Nozzle

Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Stagnation pressure (kPa)	200
Stagnation temperature (K)	300
Receiver / stagnation pressure	0.8



Converging ideal nozzle | Selected receiver

Critical pressure ratio 0.528282; choked mass flux 466.712 kg/(m² s). A 20 mm² aperture gives 7.64291 g/s at the selected receiver pressure. The converging exit is unchoked.

Start from this chapter. Enclose each stream in a steady-flow control volume. Use its inlet and outlet enthalpy to infer heat rate. Check the hot- and cold-side estimates against one another in the numerical run.

OneSimulation laboratory · Explain why greater exchanger effectiveness can give less power

Account for fluid entering and leaving

A closed-system piston can be followed with its gas. A continuously operating turbine has fresh fluid crossing its boundary every second. Enclose the hardware and its internal flow in a stationary control volume. Flowing mass carries internal energy u, kinetic energy $v^2/2$ and gravitational energy gz per kilogram. Upstream pressure also does work pushing fluid in, while downstream pressure opposes its exit. Per unit mass this flow work is p/ρ, motivating the specific enthalpy $h = u + p/\rho$.

Adding all incoming and outgoing energy rates gives, for one inlet, one outlet and constant stored energy,

$$\dot{Q} - \dot{W}_s = \dot{m}\left[h_2 - h_1 + (v_2^2 - v_1^2)/2 + g(z_2 - z_1)\right].$$

Here $\dot{W}_s$ is shaft power out of the control volume. Pressure work at the flowing boundaries is already contained in enthalpy; adding it once more would count it twice. One inlet and outlet have the same steady mass rate even when their volumetric flow rates differ dramatically because gas expands.

Read the dominant energy conversion

In a short, insulated nozzle there is no shaft, the elevation difference is small and falling enthalpy raises velocity. The property $h + v^2/2$ is then constant, often called stagnation enthalpy. In an insulated turbine with small inlet and outlet speeds, the enthalpy drop instead appears mainly as positive shaft work. An insulated compressor takes work in, raising enthalpy. An insulated throttling valve has neither shaft power nor a large net change of bulk kinetic energy between distant ports, so its inlet and outlet enthalpies are approximately equal. These components can all drop pressure during expansion, yet their energy effects differ.

For ideal-gas air with nearly constant c_p, an enthalpy difference may be approximated by $c_p\Delta T$. Keeping h in kJ/kg while using $v^2/2$ in J/kg is a common factor-of-one-thousand error. At 100 m/s the kinetic energy is 5 kJ/kg. That estimate often decides whether an inlet velocity may be omitted from a turbine work calculation.

Evaluate heat loss and exhaust speed

A turbine passes 2.0 kg/s of gas from $h_1 = 820$ to $h_2 = 600$ kJ/kg. Inlet speed is 40 m/s, exhaust speed 160 m/s and the casing loses 10 kW of heat. Elevations are equal. The gas loses $2(820 - 600) = 440$ kW of enthalpy flow. Its kinetic-energy rate grows by $2(160^2 - 40^2)/2000 = 24.0$ kW. The unit delivers $440 - 24 - 10 = 406$ kW shaft power. Counting all 440 kW as shaft output would overstate the result by both the residual exhaust energy and the heat leak.

A separate nozzle using gas at $h_1 = 820$ kJ/kg and 40 m/s could reach $h_2 = 800$ kJ/kg without a shaft if its speed rose to $\sqrt{40^2 + 2(20\,000)} = 204$ m/s. This estimate is an energy condition. The required area ratio, pressure drop and possibility of choking also depend on compressible-flow relations.

Extend the same balance to a mixer

Two insulated water lines mix at negligible speed change. One supplies 0.20 kg/s at 360 K and the other 0.30 kg/s at 300 K. A mass balance gives 0.50 kg/s at exit. With equal, constant c_p and no shaft, summing the two inlet enthalpy rates yields an outlet temperature $[0.20(360) + 0.30(300)]/0.50 = 324$ K. The temperature lies between the inlet values, closer to the higher mass rate. The mixer conserves energy while creating entropy. The next lecture explains how both statements can hold simultaneously and why reversing the mixer without another effect is impossible.

Interpret a measured stagnation temperature

For an insulated diffuser with negligible shaft work, reducing speed from 200 to 40 m/s raises static enthalpy by $(200^2 - 40^2)/2 = 19.2\,\mathrm{kJ/kg}$. For air with $c_p = 1.005\,\mathrm{kJ/(kg\,K)}$ that corresponds to a 19.1 K static temperature rise. Stagnation enthalpy $h + V^2/2$ stays constant even if internal friction produces entropy. A stationary temperature probe that slows the arriving air may approach a stagnation reading, so comparison with a predicted static temperature requires knowing where and how the measurement was taken.

Minimise compression work by selecting an intermediate pressure

A continuous air supply can be compressed in two stages and cooled at intermediate pressure. For a steady adiabatic compressor with negligible kinetic and potential changes, shaft work input is the stream enthalpy rise. With an ideal gas, constant c_p, γ, and a reversible stage,

$$w = c_p T_{\mathrm{in}} [(p_{\mathrm{out}}/p_{\mathrm{in}})^{a} - 1],$$

where $a = (\gamma - 1)/\gamma$. This follows by integrating $ds = c_p\,dT/T - R\,dp/p = 0$, then applying $dh = c_p\,dT$. If both stages use the same isentropic efficiency, divide each ideal rise by that efficiency.

Assume complete intercooling to the original inlet temperature T_1, no pressure losses and prescribed first and last pressures. Put $r_1 = p_i/p_1,\ r_2 = p_3/p_i$, so $r_1 r_2 = p_3/p_1 = r$. Total ideal work is $c_p T_1 [r_1^a + (r/r_1)^a - 2]$. Differentiate with respect to $\ln r_1$; the derivative is $c_p T_1 a (r_1^a - r_2^a)$. The sum has a minimum at equal ratios, giving intermediate pressure $p_i = \sqrt{p_1 p_3}$. This is a geometric mean of absolute pressures. A gauge-pressure ratio would give a completely different design.

Compare two cooled stages with a single compression

Supply air at 100 kPa and 300 K and deliver it at 900 kPa absolute. For $\gamma = 1.4,\ c_p = 1.005\,\mathrm{kJ/(kg\,K)}$, an ideal single stage gives $T_3 = 300(9^{0.2857}) = 562.0\,\mathrm{K}$ and work about 263 kJ/kg. Two equal-ratio stages operate through 300 kPa. Each reaches $300(3^{0.2857}) = 410.6\,\mathrm{K}$ before its outlet or intercooler, requiring $1.005(110.6) = 111.2\,\mathrm{kJ/kg}$. Total work is about 222.4 kJ/kg, roughly 15.5% lower. The intercooler must remove 111.2 kJ/kg between the stages under the ideal complete-cooling assumption.

With equal stage efficiencies of 0.80, each real enthalpy rise becomes 139.0 kJ/kg and the intermediate discharge temperature is about 438 K. Cooling back to 300 K then removes 139.0 kJ per kilogram, with a corresponding load of 27.8 kW at 0.20 kg/s. The equal-ratio choice survives because both stage work expressions acquired the same multiplier. An intercooler that can cool only to 330 K increases the weight of second-stage compression work; the new optimum shifts towards a larger first-stage pressure ratio. A design comparing only compressor shaft power would also need to account for coolant pumping, fan power and pressure losses before choosing the number of stages.

For an incomplete intercooler the optimal shift can be calculated. Take stage two inlet temperature $T_m = 330$ K independently maintained and identical efficiencies. The variable portion of total work is proportional to $T_1 r_1^a + T_m (r/r_1)^a$. Differentiating with respect to log first-stage ratio gives $T_1 r_1^a = T_m r_2^a$, or $r_1/r_2 = (330/300)^{1/a} = 1.40$. At overall ratio nine,

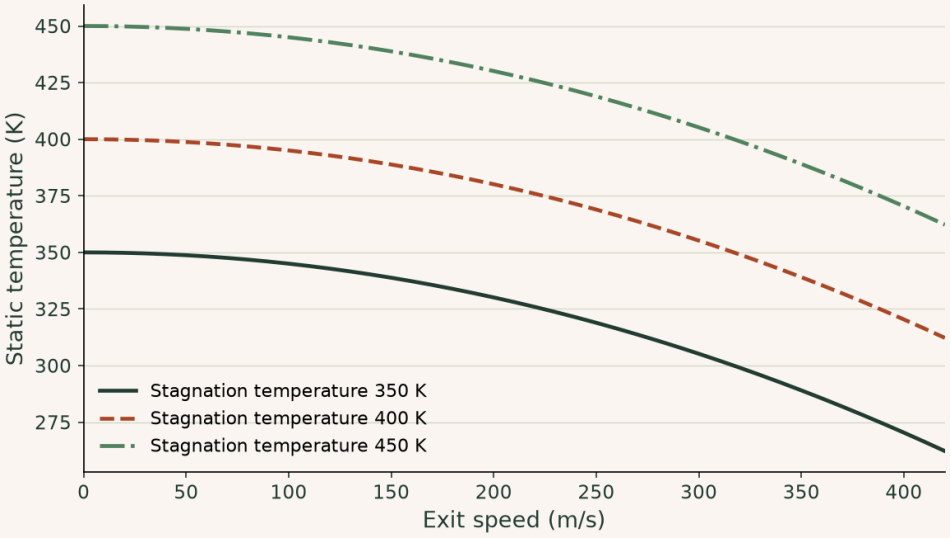
$$r_1 = \sqrt{9(1.40)} = 3.54,\; r_2 = 2.54.$$

Compressing farther while the air is colder is energetically preferable within this simplified comparison. These ratios presume the cooler can maintain 330 K over its revised duty and that both discharge temperatures stay within their allowable ranges. A compressor map may impose further flow and speed limits beyond the thermodynamic minimum.

At fixed pressure endpoints, the measured intermediate pressure alone would thus be insufficient for diagnosing whether the cooler or a compressor caused an observed power increase.

How much static temperature does a nozzle convert to speed?

Air flows steadily through an insulated nozzle with negligible shaft work and elevation change. Take $c_p = 1005$ J/(kg K). Divide the conserved stagnation enthalpy $c_p T + u^2/2$ by c_p to find the exit temperature.



At 300 m/s the static temperature is about 44.8 K below stagnation temperature, regardless of pressure in this enthalpy calculation. Predicting exit pressure as well would require an entropy or loss model.

▶ **Read values from the plot**

Interactive calculation · change the inputs of this example

Change stagnation temperature and heat capacity independently. The speed-related temperature drop follows the energy balance.

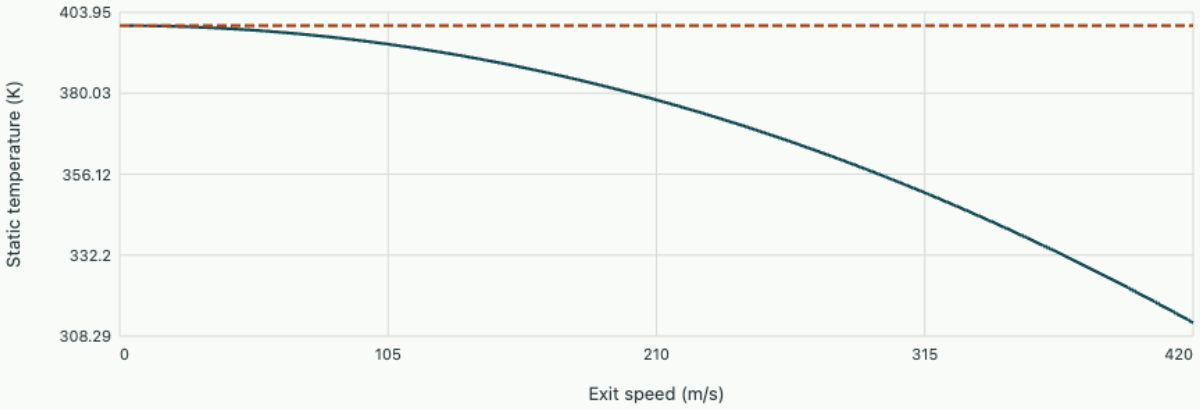
The curves below are recalculated from this section's equations in your browser. The fixed reference plot above remains available for comparison.

Open this live calculation

Try the calculation

Stagnation temperature (K) **400**

Specific heat (J/kg K) **1005**



Selected stagnation state Inlet stagnation temperature

At 300 m/s the static temperature drop is 44.776 K. Pressure needs a separate entropy/loss relation.

Study the full derivation

Staged compression with an imperfect intercooler · Thermal processes and transport in engineering

A laboratory compressed-air supply has two compressor stages and an intercooler. Equal pressure ratios minimise work only under particular assumptions. Derive the optimum, compare it with a single-stage machine and identify which temperature and power measurements reveal imperfect cooling or unequal efficiency.

Problem sheet 1 · Energy accounts for stored and flowing matter

Draw a boundary, identify work and heat with consistent signs, and test the effect of changing the boundary.

A gas under a spring-loaded piston

A closed gas expands from 0.010 to 0.016 m³. The resisting absolute pressure rises linearly from 160 to 280 kPa. Its internal energy increases by 0.900 kJ. Compute the boundary work and the heat entering the gas. If the ambient pressure is 100 kPa, how much of the boundary work is available after displacing the atmosphere?

▼ **Work, heat and the atmosphere**

The area under the straight pressure–volume line is the mean ordinate times the volume interval. With quasistatic motion,

$$W = (160 + 280)(0.006)/2 = 1.320\,\mathrm{kJ}.$$

The first law with work positive out gives

$$Q = \Delta U + W = 2.220\,\mathrm{kJ}.$$

Displacing the atmosphere consumes $100(0.006) = 0.600\,\mathrm{kJ}$. The remaining 0.720 kJ raises the piston or compresses the spring. Subtract atmospheric work when assessing useful work, but use absolute pressure for the closed gas energy balance.

Is the exit speed consistent?

Air enters a steady insulated nozzle at 360 K and 30 m/s, and leaves at 325 K. Take constant specific heat 1.005 kJ/(kg K). Find the exit speed and the exit area for 0.24 kg/s if the exit density is 1.08 kg/m³. Explain what changes if the nozzle loses 3 kJ/kg as heat at the same temperatures.

▼ **Balance and continuity**

Per unit mass,

$$V_2^2 = V_1^2 + 2c_p(T_1 - T_2) = 900 + 2(1005)(35) = 71250\,\mathrm{m^2/s^2}.$$

Thus $V_2 = 266.9\,\mathrm{m/s}$. The area is

$$\dot m/(\rho_2 V_2) = 8.33 \times 10^{-4}\,\mathrm{m^2}.$$

Losing 3000 J/kg reduces V_2^2 by 6000, giving 255.4 m/s. Energy that leaves the control volume as heat is unavailable to accelerate the stream.

Filling an evacuated rigid receiver

An insulated, initially evacuated tank is filled from a large air line at 300 K. The final mass is 0.80 kg. Assume perfect gas, $c_p = 1.005$, $c_v = 0.718$ kJ/(kg K), negligible kinetic energy in the final tank, no shaft work. Find the final energy and temperature. Why would a closed, constant-volume calculation yield a different interpretation?

▼ **Integrate the unsteady balance**

Each kilogram carries inlet enthalpy 301.5 kJ/kg. Integrating over filling gives $U_2 = m_2 h_{\mathrm{in}} = 241.2\,\mathrm{kJ}$ with a common zero-energy convention. Hence

$$T_2 = h_{\mathrm{in}}/c_v = 419.9\,\mathrm{K}.$$

The temperature rise accompanies the flow work carried in with the gas. A closed balance that follows the entering parcel must explicitly include the work that pushes the parcel into the receiver.

Explore the governing equations after solving the sheet: State, process and the first law for a closed system; Open systems, enthalpy and steady-flow devices. Compare the assumptions and units before transferring your chosen parameters; the controls may cover a different numerical range from the written exercise.

Chapter 3 · Entropy, irreversibility and useful-work loss

Check a proposed process with the second law and quantify the effect of a finite temperature difference.

FROM FIRST PRINCIPLES

Separate entropy transfer from entropy generated inside a boundary

For a reversible differential process the thermodynamic definition of entropy is

$$dS = \delta Q_{\mathrm{rev}}/T.$$

Irreversible changes reach the same state-function difference but can generate entropy internally. For a closed system exchanging heats Q_j across boundary portions at known local temperatures T_j, the second law gives an experimentally testable inequality:

$$S_2 - S_1 = \sum_j \int \frac{\delta Q_j}{T_j} + S_{\mathrm{gen}}, \qquad S_{\mathrm{gen}} \geq 0.$$

To evaluate the left side independently of the actual path, use a reversible path between the same equilibrium states. For an ideal gas the fundamental relation $Tds = du + p\,dv$, together with $du = c_v dT$ and $pv = RT$, yields

$$s_2 - s_1 = c_v \ln(T_2/T_1) + R \ln(v_2/v_1)$$

for constant specific heat. Inserting this property change into the balance isolates entropy generation of the actual process. A gas might cool and lose entropy while the combined gas plus surroundings generates entropy. For heat exchanged across a wall separating two reservoirs, the two reservoirs' boundary transfer terms have opposite signs and different temperature denominators.

Background · Entropy generation

$$\dot{S}_{\mathrm{gen}} = \dot{S}_{\mathrm{system}} - \sum \dot{Q}_j/T_j - \dot{S}_{\mathrm{mass,in}} + \dot{S}_{\mathrm{mass,out}}$$

Use boundary temperatures at which heat actually crosses. All flows, storage and units must refer to the same control volume.

DERIVATION AND APPLICATIONS

Count the entropy generated when two ideal gases mix

Two separated ideal gases can have the same temperature and pressure, yet their mixing increases entropy. Consider distinct nonreacting species in insulated rigid compartments. Remove the separator and let them equilibrate in the common volume. Initially both sides have temperature T and pressure p, and no shaft work is performed. The total internal energy of ideal gases depends only on temperature, so the final equilibrium temperature stays T. Denote mole numbers by n_i and total volume by V.

For each species, the entropy change can be calculated along a reversible isothermal expansion into its accessible final volume, even though the actual mixing is irreversible. Integrating $\delta Q_{\mathrm{rev}}/T = p_i\,dV/T = n_i R\,dV/V$ gives

$$\Delta S_i = n_i R \ln(V/V_i).$$

The initial volumes were proportional to mole numbers because p and T were common. Thus $V_i/V = n_i/\sum_j n_j = y_i$, the final mole fraction. Add the species changes and use zero boundary entropy transfer:

$$\Delta S_{\mathrm{mix}} = S_{\mathrm{gen}} = -R\sum_i n_i \ln y_i.$$

For two gases with y and $1-y$, the generation per mole of mixture is maximal when both are equally abundant. The derivative and its zero are

$$\frac{d}{dy}\frac{\Delta S_{\mathrm{mix}}}{n} = R\ln[(1-y)/y] = 0 \quad\Longrightarrow\quad y = \tfrac{1}{2}.$$

The entropy change tends to zero when either fraction tends to one. If the compartments contain the same gas and the initial temperatures and pressures already agree, removing the partition does not create distinguishable species; assigning artificial labels would predict a spurious mixing term.

WORKED EXAMPLE

Estimate the lost work of mixing two samples

Change the inputs and test this calculation.

Mix 1.00 mol of one ideal gas and 3.00 mol of another at 300 K. The mole fractions are 0.250 and 0.750. The generated entropy is

$$-8.314[\ln 0.25 + 3\ln 0.75] = 18.7\,\mathrm{J/K}.$$

Relative to an environment at 300 K, the destroyed work potential is $300(18.7) = 5.61$ kJ. No energy has disappeared from the insulated vessel.

SECOND APPLICATION

Compare equal and strongly unequal amounts

Change the inputs and test this calculation.

For a fixed total of 4.00 mol, an equal mixture generates $4R\ln 2 = 23.1$ J/K. With only 0.0400 mol of one species the fractions become 0.0100 and 0.990. The generation falls to

$$4R[-0.01\ln 0.01 - 0.99\ln 0.99] = 1.86\,\mathrm{J/K}.$$

Each rare molecule gains access to much more volume, but there are few such molecules; both contributions must be counted before comparing samples.

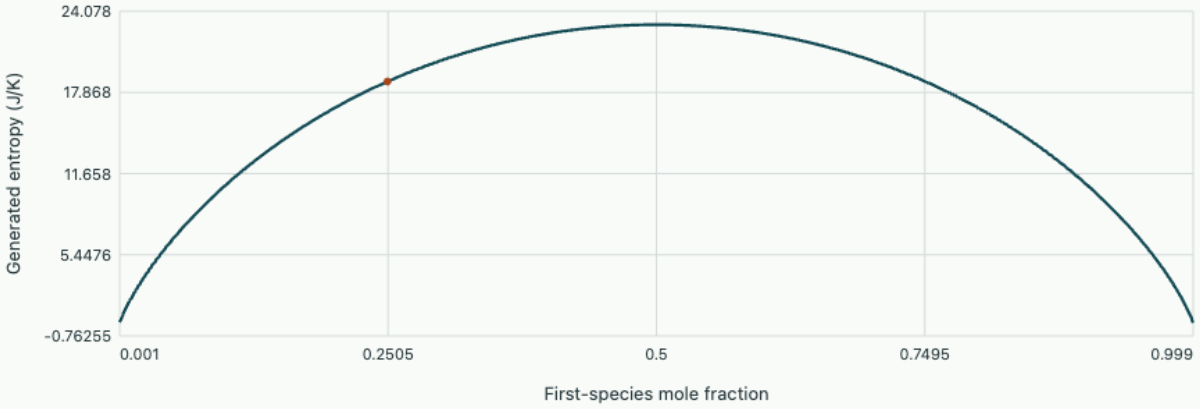
Try it · Mixing

Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Total amount (mol)	4
First-species mole fraction	0.25
Environment temperature (K)	300



Two distinct ideal species | Selected mixture

Entropy generation 18.7021 J/K; destroyed work potential T0 Sgen = 5.61062 kJ. Equal initial temperature and pressure; the same-species case has no mixing term.

Energy bookkeeping is not sufficient

In an insulated two-body experiment, heat lost by one body is gained by the other. That energy balance alone would allow heat to move spontaneously from a colder body to a hotter one. Entropy supplies the missing direction constraint. It is a property of a state. For a simple equilibrium system a reversible differential heat transfer defines $dS = \delta Q_{\mathrm{rev}}/T$. An irreversible process between the same endpoints has the same entropy difference; the actual heat transfer divided by the body temperature need not equal that difference.

For a closed system, the integrated entropy balance reads $\Delta S = \int \delta Q/T_b + S_{\mathrm{gen}}$, where T_b is the absolute temperature at the boundary where the heat crosses. Entropy generation satisfies $S_{\mathrm{gen}} \geq 0$. A

system can lose entropy by exporting it in heat or in a mass stream. A negative change in the entropy of a single hot body therefore need not violate the second law.

Background · Why does a small temperature difference produce a second-order loss at fixed conductance?

Let a link of constant conductance G join reservoirs held at $T + \Delta T/2$ and $T - \Delta T/2$. In steady operation the heat current is $G\Delta T$. Subtract the entropy leaving the hot reservoir from that entering the cold reservoir.

$$\dot{S}_{\mathrm{gen}} = G\Delta T\left(\frac{1}{T - \Delta T/2} - \frac{1}{T + \Delta T/2}\right) = \frac{G(\Delta T)^2}{T^2 - (\Delta T/2)^2}.$$

At fixed conductance, halving a small difference approximately quarters the entropy production per second, while it also halves the heat current. If a design must keep the heat rate fixed, halving the difference requires nearly twice the conductance and only halves the production. This distinction matters when comparing a bigger exchanger with the original one.

Calculate a property change with a convenient path

For a calorically perfect ideal gas, combine $T\,ds = du + p\,dv$ with $du = c_v\,dT$ and $p = RT/v$. This gives

$$ds = c_v\,dT/T + R\,dv/v.$$

Replacing v using the ideal-gas law yields the alternative $ds = c_p\,dT/T - R\,dp/p$. Integrate between the initial and final equilibrium states. These are property relationships, even when the actual path between states includes friction, throttling or free expansion.

An ideal gas passing through an insulated valve with equal inlet and outlet speeds has $h_2 = h_1$, so its temperatures are equal in the constant-c_p approximation. Dropping air from 400 to 100 kPa gives $s_2 - s_1 = -0.287\ln(100/400) = 0.398$ kJ/(kg K). With no heat transfer or other streams, the positive entropy change per kilogram is entirely generated. A reversible insulated expansion turbine would generally produce a lower outlet temperature and shaft work.

Trace entropy across two reservoirs

An insulated conductor transfers 600 W from a reservoir at 400 K to another at 300 K. Its stored energy and entropy are steady. The hot reservoir loses entropy at $600/400 = 1.50$ W/K, while the cold one gains $600/300 = 2.00$ W/K. The total generation rate is 0.50 W/K. Heat rate is the same at both ends; entropy transfer in that heat differs because the temperatures differ.

$$\dot{S}_{\mathrm{gen}} = \dot{Q}\left(\frac{1}{T_c} - \frac{1}{T_h}\right).$$

If the surroundings defining the available-work reference are at $T_0 = 300$ K, the corresponding destruction of available work is $T_0 \dot{S}_{\mathrm{gen}} = 150$ W. For these particular reservoirs a reversible engine supplied with 600 W at 400 K could have produced $600(1 - 300/400) = 150$ W before rejecting the remainder. The conductor instead rejects all 600 W to the cold sink. Energy did not disappear; an opportunity for work was lost.

Locate sources and choose a boundary

A heat exchanger can have little external heat leak yet generate appreciable entropy internally through finite temperature differences and pressure losses. Write $\dot{S}_{\mathrm{gen}} = \sum_{\mathrm{out}} \dot{m} s - \sum_{\mathrm{in}} \dot{m} s - \sum_j \dot{Q}_j / T_{b,j}$ for steady operation. Both streams belong inside the same control volume when testing internal heat exchange. If the boundary encloses only one stream, heat crossing the wall carries entropy with it and must appear explicitly. The largest reduction in entropy generation per extra unit of area or pumping effort is a design question. The thermodynamic balance identifies losses without alone determining the economically preferred geometry.

Compare the conditions on a valve and a turbine

Air passes through a steady insulated throttling valve from 400 to 100 kPa with negligible changes of speed. With no shaft work the first law gives equal inlet and outlet enthalpy. For ideal air this implies equal temperature. Its entropy change follows from $\Delta s = c_p \ln(T_2/T_1) - R \ln(p_2/p_1) = 0.287 \ln 4 = 0.398\,\mathrm{kJ/(kg\,K)}$. At 0.10 kg/s the entropy generation is 0.0398 kW/K. An ideal adiabatic turbine undergoing the same pressure drop has constant entropy and delivers work as its temperature falls. Pressure ratio alone therefore cannot determine the temperature change across an unfamiliar device.

Obtain a work bound for two finite bodies

The reversible efficiency $1 - T_c/T_h$ describes operation between reservoirs held at fixed temperatures. If the two bodies are finite, each increment of transferred heat changes the temperatures and the instantaneous efficiency. Consider two incompressible blocks in a common insulated enclosure. Their heat capacities C_h, C_c are taken as constant. A reversible engine operates between them and exports mechanical work until no temperature difference remains. The working fluid returns to its starting state after each incremental cycle, so conservation of entropy for this reversible composite system determines the final temperature.

$$C_h \ln(T_f/T_h) + C_c \ln(T_f/T_c) = 0,$$
$$T_f = T_h^{C_h/(C_h+C_c)} T_c^{C_c/(C_h+C_c)}.$$

Energy conservation gives maximum work $W_{\max} = C_h T_h + C_c T_c - (C_h + C_c) T_f$ for consistent internal-energy references. The reversible final temperature is a weighted geometric mean. If the two blocks are simply put in thermal contact without work output, first-law conservation instead gives the weighted arithmetic mean. The arithmetic mean is larger for unequal initial temperatures; additional energy remains as sensible heat after irreversible contact. These different final temperatures matter when comparing two ways of using the same finite resources.

In a real engine running at finite heat-transfer rates there are temperature differences between each block and the working fluid. Total entropy then increases. For a specified common final temperature the entropy change of the blocks is fixed by state properties; measuring it indicates the departure from the ideal reversible construction without needing the working-fluid cycle.

Compare direct contact and a reversible engine quantitatively

Take a hot assembly of heat capacity 2.0 kJ/K at 450 K and a cooler assembly of 1.0 kJ/K at 300 K. Solving the log balance gives

$$T_f = (450^2\, 300)^{1/3} = 393.1\,\mathrm{K}.$$

The hot assembly releases about $2(450 - 393.1) = 113.8\,\mathrm{kJ}$ and the cold one receives 93.1 kJ. Their difference gives approximately 20.7 kJ of maximum work. Acting as though the initial temperatures stayed at 450 and 300 K throughout would assign an incorrectly high efficiency to later parts of the transfer.

By direct contact the common temperature would be $(2 \times 450 + 300)/3 = 400\,\mathrm{K}$. The entropy generated is

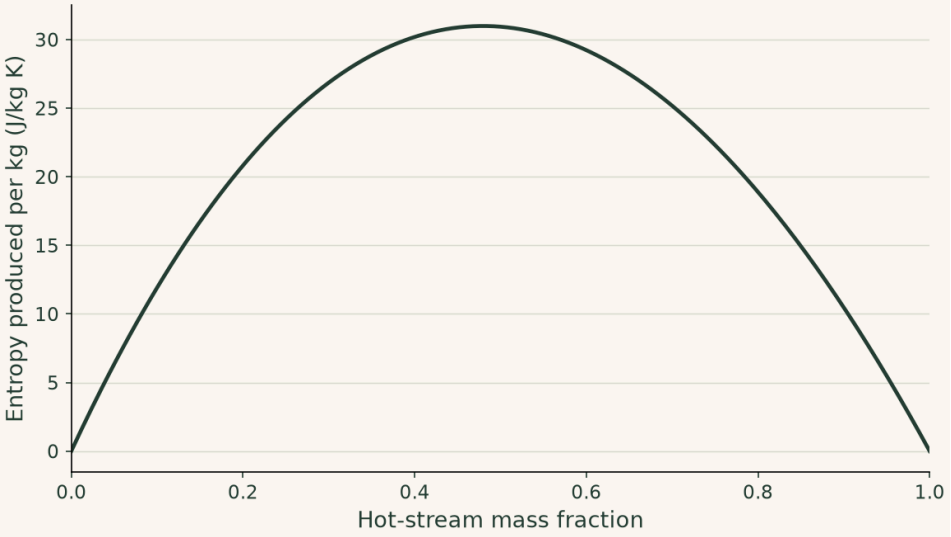
$$2 \ln(400/450) + \ln(400/300) = 0.0521\,\mathrm{kJ/K}.$$

If instead a device delivers a measured 12 kJ of work before the blocks reach a common temperature, energy fixes that temperature as $(1200 - 12)/3 = 396\,\mathrm{K}$. The measured process then generates $2 \ln(396/450) + \ln(396/300) = 0.0220\,\mathrm{kJ/K}$, a positive value between the two limiting cases. Entropy generation multiplied by an environmental temperature gives destroyed exergy relative to that environment; comparing the finite-body work outputs also requires acknowledging their different final stored energies.

The geometric-mean result has an informative limiting case. If the cold assembly is so large that C_c/C_h tends to infinity, its temperature remains effectively fixed, and the engine work reduces to $C_h[(T_h - T_c) - T_c \ln(T_h/T_c)]$. This is the integral of the reversible work recovered as the finite hot block approaches a constant sink. It is smaller than applying the initial Carnot efficiency to all of the hot block’s released heat.

Which mixture generates the most entropy?

Mix streams of the same incompressible liquid at 370 K and 290 K, at equal pressure, with constant $c_p = 4180$ J/(kg K). The energy balance fixes the outlet temperature. Evaluate the outlet entropy minus the mass-weighted inlet entropy using $c_p \ln T$.



Either pure stream gives zero entropy production. An equal-mass mixture settles at 330 K and produces about 30.9 J/(kg K). The maximum is near, but not exactly at, equal fractions because temperature appears inside a logarithm.

▶ **Read values from the plot**

Interactive calculation · change the inputs of this example

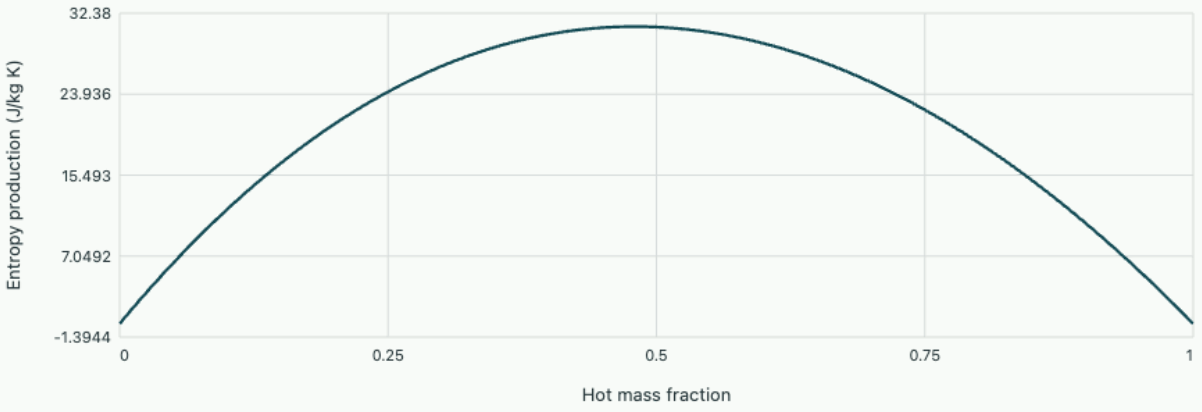
Move the two inlet temperatures toward each other and watch entropy production vanish throughout the mixture range.

The curves below are recalculated from this section's equations in your browser. The fixed reference plot above remains available for comparison.

Open this live calculation

Try the calculation

Hot stream (K)	**370**
Cold stream (K)	**290**
Liquid heat capacity (J/kg K)	**4180**



Adiabatic mixture

Equal mass fractions reach 330.000 K and generate 30.935 J/(kg K). Endpoints are unmixed and generate zero.

Chapter 4 · Thermodynamic cycles, refrigeration and phase change

Analyse a vapour-compression cycle from supplied state properties and compare its performance with a reversible bound.

FROM FIRST PRINCIPLES

Obtain cycle efficiency and refrigerator performance from two balances

A working fluid completing a cycle returns to its original thermodynamic state, so its internal energy and entropy have zero net changes. Denote positive magnitudes of heat received from the hot reservoir and given to the cold reservoir by Q_h, Q_c, with absolute temperatures $T_h > T_c$. For a heat engine the first law then gives work delivered as their difference, and the cycle entropy balance relates both heat amounts:

$$W_{\text{out}} = Q_h - Q_c,$$
$$0 = Q_h/T_h - Q_c/T_c + S_{\text{gen}}.$$

Since generated entropy is nonnegative, $Q_c/Q_h \geq T_c/T_h$. Thus the maximum fraction of the absorbed heat available as work occurs for a reversible cycle:

$$\eta = W_{\text{out}}/Q_h = 1 - Q_c/Q_h \leq 1 - T_c/T_h.$$

For a refrigerator run in the opposite direction the useful output is extracted cold heat and the input is compressor work. Applying the same balances with heat directions reversed gives

$$\text{COP} = Q_c/(Q_h - Q_c) \leq T_c/(T_h - T_c).$$

These limits assume exactly two thermal reservoirs. The actual cycle depends on how its individual compression, heat exchange and expansion processes are carried out.

DERIVATION AND APPLICATIONS

Combine two refrigeration stages without counting intermediate heat twice

A cascade refrigerator can move heat across a large temperature interval using two separate loops. For an ideal limit, assume each stage is reversible, works steadily, and exchanges heat with constant-temperature reservoirs. The lower stage extracts Q_c per cycle at T_c and rejects Q_i at intermediate temperature T_i. The upper stage takes precisely this Q_i and rejects Q_h at T_h. All temperatures are absolute; ignore other heat leaks, pressure drops and energy storage at the intermediate exchanger.

For a reversible cycle the working fluid returns to its initial entropy, so its incoming and outgoing heat-to-temperature ratios match. Apply this statement independently to both stages:

$$Q_i/Q_c = T_i/T_c,$$
$$Q_h/Q_i = T_h/T_i.$$

Energy conservation gives the work delivered to each compressor. In the sum the intermediate heat cancels because it leaves the first loop and enters the second:

$$W_1 = Q_i - Q_c,$$
$$W_2 = Q_h - Q_i,$$
$$W_{\text{tot}} = Q_h - Q_c.$$

Combine the two ratios to remove T_i. The minimum work and highest possible overall cooling coefficient of performance are

$$W_{\min} = Q_c(T_h/T_c - 1),$$
$$\text{COP}_{\text{cascade}} = Q_c/W_{\min} = T_c/(T_h - T_c).$$

The ideal intermediate temperature controls how work is split but cannot change the reversible overall limit. A real cascade can benefit from smaller pressure ratios in its components. Heat exchanged across a finite temperature difference at the interstage exchanger produces entropy and raises work above the reversible value.

WORKED EXAMPLE

Allocate heat and work in a laboratory freezer

Change the inputs and test this calculation.

Take $T_c = 240$ K, $T_i = 270$ K, $T_h = 300$ K and cooling duty 2.00 kW. The lower stage rejects 2.25 kW, consuming 0.250 kW. The upper rejects 2.50 kW and also consumes 0.250 kW. The overall COP is

$$2.00/0.500 = 4.00.$$

The upper loop has to process all the heat extracted from the cold space plus the lower compressor work.

SECOND APPLICATION

Quantify the cost of a finite interstage temperature difference

Change the inputs and test this calculation.

Keep the end temperatures and cooling duty, but let the lower stage reject heat at 275 K while the upper stage takes it at 265 K. Treat each separate stage as internally reversible. The shared heat rate is now

$2(275/240) = 2.292$ kW. Upper rejection is

$$2.292(300/265) = 2.594\,\mathrm{kW}.$$

The total work rises to 0.594 kW and COP falls to 3.37. The exchanger generates $2.292(1/265 - 1/275) = 0.000315$ kW/K; multiplied by 300 K this accounts for the extra compressor work.

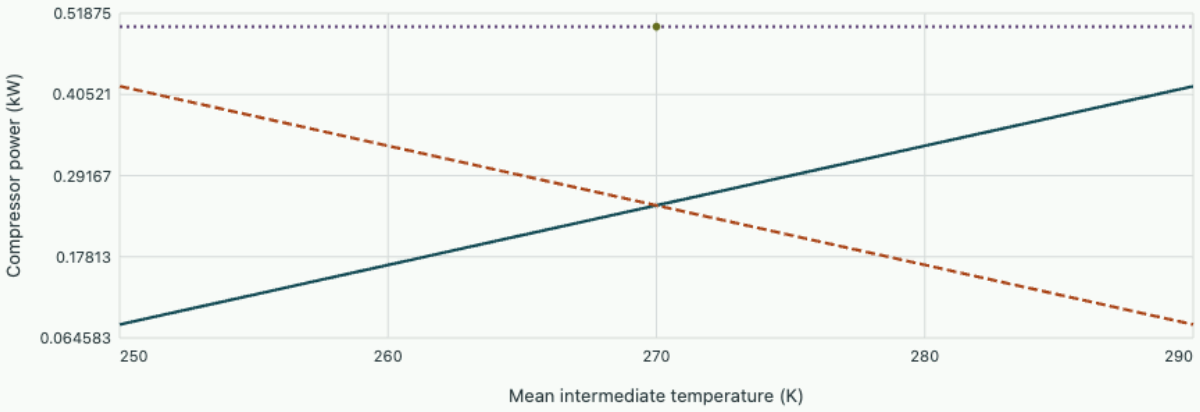
Try it · Cascade

Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Mean intermediate temperature (K)	270
Interstage temperature gap (K)	0
Cooling duty (kW)	2



Cold/hot reservoirs 240/300 K. Total work 0.5 kW; COP 4. Interstage entropy generation 0 W/K; 300 K × Sgen = 0 kW equals the extra work above the reversible limit.

Account for transfers over a complete cycle

A cyclic machine returns a working fluid to its initial state. Its internal energy and entropy changes over one complete cycle are zero. Heat and shaft work can nevertheless have nonzero net values. For an engine receiving Q_h, rejecting positive magnitude Q_c and delivering work W_{out}, the first law gives

$$W_{\text{out}} = Q_h - Q_c.$$

Thermal efficiency is W_{out}/Q_h. A refrigerator instead consumes work to remove heat from a cold space. Its useful effect is cooling, so its performance measure is $\text{COP}_R = Q_c/W_{\text{in}}$.

Between reservoirs at absolute temperatures T_h, T_c a reversible cycle transports equal entropy $Q_h/T_h = Q_c/T_c$. Combining this condition with the first law gives $\eta_{\text{rev}} = 1 - T_c/T_h$ for an engine and $\text{COP}_{R,\text{rev}} = T_c/(T_h - T_c)$ for a refrigerator. The latter can exceed one because cooling duty includes transported heat. A heat pump uses Q_h as useful output and has $\text{COP}_{HP} = \text{COP}_R + 1$ for the same cycle.

Follow the refrigerant once around the loop

Label the compressor inlet 1, discharge 2, condenser outlet 3 and expansion-valve outlet 4. Assume steady operation, negligible kinetic and elevation changes and insulated compressor and valve. The evaporator absorbs $q_c = h_1 - h_4$ per kilogram, the compressor consumes $w = h_2 - h_1$, the condenser rejects $q_h = h_2 - h_3$ and throttling preserves enthalpy $h_4 = h_3$. Enter enthalpies from a consistent refrigerant table or supplied state data. A constant ideal-gas heat capacity cannot represent latent heat across a two-phase region.

At saturation a mixture at quality x has $h = (1-x)h_f + x h_g$, where x is vapour mass fraction and h_f, h_g are saturated liquid and vapour enthalpies at the same pressure. A throttle can create vapour as the pressure falls even though no external heat enters the valve. The mixture redistributes its enthalpy between phases at the new saturation temperature.

Determine cooling, power and flash fraction

For an illustrative refrigerant take enthalpies $h_1 = 390$, $h_2 = 430$, $h_3 = h_4 = 250$ kJ/kg. At evaporator pressure the saturated values are $h_f = 200,\ h_g = 390$ kJ/kg. With mass flow 0.060 kg/s, the evaporator removes $0.060(390 - 250) = 8.40$ kW, the compressor needs 2.40 kW and the condenser rejects 10.80 kW. The powers satisfy $8.40 + 2.40 = 10.80$ kW, and cooling COP is 3.50. The valve outlet quality is $(250 - 200)/(390 - 200) = 0.263$: just over a quarter of the mass is already vapour when it enters the evaporator.

Suppose that at the actual discharge pressure an isentropic compression would give $h_{2s} = 420$ kJ/kg. The compressor isentropic efficiency is

$$(h_{2s} - h_1)/(h_2 - h_1) = 30/40 = 0.75.$$

Losses raise the required shaft power and increase condenser duty at unchanged evaporator inlet and outlet conditions.

Distinguish a property assumption from a bound

If heat is removed from a 268 K reservoir and rejected to a 308 K reservoir, the reversible refrigeration COP is

$$268/(308-268) = 6.70.$$

A measured COP of 3.50 lies below that limit. Real heat exchangers need finite temperature differences, compressors generate entropy and valves destroy expansion work. Reducing the required temperature lift improves the reversible limit. The working fluid’s actual pressures and enthalpies must be redetermined to estimate a real COP change. A cycle sketch and a numbered state table prevent accidentally using a compressor outlet temperature as the external hot-reservoir temperature in the bound.

Follow the energy after the compressor stops

Consider a cold cabinet held at constant temperature by a refrigerator. If it receives a 100 W heat leak and a 25 W electric fan dissipates power inside it, the evaporator must remove 125 W in steady operation. At cooling COP 2.5 the compressor draws 50 W, and the condenser releases $125 + 50 = 175$ W into the room. If both the condenser and the cabinet exterior exchange heat with that room, its net thermal gain from the refrigerator system is

$$175 - 100 = 75\,\mathrm{W}.$$

This equals the total electrical input to fan and compressor. Boundary selection explains why the condenser duty and the net heating of the entire room have different values.

Derive a gas-turbine efficiency from pressure ratio

A gas turbine running on a closed idealised Brayton cycle provides a useful counterpart to a refrigerating cycle. Model compression and expansion as isentropic ideal-gas processes with constant heat capacities. Heat is supplied and rejected at constant pressure. Let T_1 be compressor inlet, T_2 compressor outlet, T_3 turbine inlet and T_4 exhaust, and set $r_p = p_2/p_1 = p_3/p_4$. For $a = (\gamma - 1)/\gamma$, the isentropic relations give

$$T_2/T_1 = T_3/T_4 = r_p^{a}.$$

The heat quantities per kilogram are $q_{\mathrm{in}} = c_p(T_3 - T_2)$ and $q_{\mathrm{out}} = c_p(T_4 - T_1)$.

Rewriting the latter difference as $(T_3 - T_2)/r_p^{a}$ yields the ideal thermal efficiency $1 - r_p^{-a}$. This simple rise with pressure ratio does not mean that available shaft work rises indefinitely for a fixed permissible turbine inlet temperature. Net specific work divided by c_p is $T_3(1 - 1/y) - T_1(y - 1)$ with $y = r_p^{a}$. Differentiating gives $T_3/y^2 - T_1 = 0$, so maximum work at fixed end temperatures occurs at $y = \sqrt{T_3/T_1}$. When compression heats the gas all the way to T_3, external heat input and net output both vanish in the ideal model.

Calculate the fraction of turbine output consumed by compression

Let $T_1 = 300\,\mathrm{K}$, $T_3 = 1200\,\mathrm{K}$, $r_p = 8$, $\gamma = 1.4$, and $c_p = 1.005\,\mathrm{kJ/(kg\,K)}$. Then $y = 1.811$,

$$T_2 = 543.4\,\mathrm{K},$$

$$T_4 = 662.5\,\mathrm{K}.$$

Compressor work is 244.7 kJ/kg, turbine work is 540.2 kJ/kg and net output is about 295.6 kJ/kg. The back-work ratio is $244.7/540.2 = 0.453$, and the thermal efficiency is 44.8%. The ideal maximum-work pressure ratio would be $2^{1/a} = 11.31$ for these temperature limits.

Now specify compressor isentropic efficiency 0.82 and turbine efficiency 0.87 at the same pressure ratio and inlet temperatures. Actual compressor outlet is $300 + (543.4 - 300)/0.82 = 596.9\,\mathrm{K}$ and actual turbine outlet is

$$1200 - 0.87(1200 - 662.5) = 732.3\,\mathrm{K}.$$

The net work becomes

$$1.005[(1200 - 732.3) - (596.9 - 300)] = 171.7\,\mathrm{kJ/kg}.$$

Heat supplied is $1.005(1200 - 596.9) = 606.1\,\mathrm{kJ/kg}$, giving efficiency 28.3%. A relatively small fractional degradation in each machine causes a larger fractional drop in net power because compression consumes a large portion of gross turbine work. If a recuperator is considered, the actual exhaust being hotter than the compressor discharge offers a possible heat-recovery driving temperature difference.

The actual exhaust and compressor temperatures also quantify the limit of regeneration. If both streams have the same heat capacity rate and an idealised recuperator has effectiveness 0.70, the compressed stream can gain $0.70(732.3 - 596.9) = 94.8$ K before entering the external heater. Its required heat input falls by $1.005(94.8) = 95.3\,\mathrm{kJ/kg}$, to about 511 kJ/kg. If turbine and compressor pressure ratios, inlet temperatures and work remain unchanged, thermal efficiency rises to

$$171.7/511 = 33.6\%.$$

Check both exchanger terminal temperatures to ensure heat still flows in the intended direction. Pressure drop in an actual recuperator can reduce net work, while the lower fuel requirement might still justify the component. A recuperator cannot improve the ideal cycle at an operating condition where turbine exhaust is already colder than compressor delivery.

Counterflow heat exchange here transfers energy internally, so the cycle energy account includes the reduced external heating requirement and leaves the recuperator transfer out of its net external balance.

How does piston compression ratio set the ideal cycle efficiency?

Consider a fixed mass of air in a piston cylinder. Represent heat exchange by two constant-volume processes and represent rapid compression and expansion by reversible adiabatic paths. This air-standard Otto cycle is a closed reference cycle. With constant specific heats, c_v determines the change in internal energy and $\gamma = c_p/c_v$ determines the adiabatic temperature change. Let $r = V_1/V_2$, with state 1 just before compression and state 2 just after. Heat is supplied from 2 to 3 at the smaller volume and rejected from 4 to 1 at the larger.

The isentropic law $TV^{\gamma-1} = \text{constant}$ gives $T_2 = r^{\gamma-1}T_1$ for compression and $T_3 = r^{\gamma-1}T_4$ for expansion. The constant-volume heat transfers have magnitudes $q_{\text{in}} = c_v(T_3 - T_2)$ and $q_{\text{out}} = c_v(T_4 - T_1)$, since the boundary work vanishes on each of those legs. Subtract the two temperature relations to obtain

$$T_3 - T_2 = r^{\gamma-1}(T_4 - T_1).$$

The complete cycle returns to the same internal energy, giving

$$\eta = \frac{w_{\text{net}}}{q_{\text{in}}} = 1 - \frac{q_{\text{out}}}{q_{\text{in}}} = 1 - r^{1-\gamma}.$$

For $T_1 = 300\,\text{K}$, $T_3 = 1500\,\text{K}$, $r = 8$, $\gamma = 1.40$, and $c_v = 0.718\,\text{kJ}\,\text{kg}^{-1}\text{K}^{-1}$, compression ends at $689.2\,\text{K}$ and expansion at $652.9\,\text{K}$. The supplied heat is $582.1\,\text{kJ}\,\text{kg}^{-1}$ and rejected heat $253.4\,\text{kJ}\,\text{kg}^{-1}$. The difference gives $328.7\,\text{kJ}\,\text{kg}^{-1}$ of net work and $\eta = 56.5\%$. Raising r to 10 while retaining both temperature limits yields $\eta = 60.2\%$, but only $535.9\,\text{kJ}\,\text{kg}^{-1}$ can now be supplied. Net work falls to $322.6\,\text{kJ}\,\text{kg}^{-1}$. Thus a larger fraction of the available heat becomes work, although the specified cycle does less work per kilogram. The temperature constraint matters when comparing design alternatives. Varying specific heats would change these numerical estimates.

Practise on sheet 2 · Test a heat input against a peak-temperature limit. Further reading · MIT · 16.Unified Thermodynamics and Propulsion · §3.5.1 Efficiency of an ideal Otto cycle.

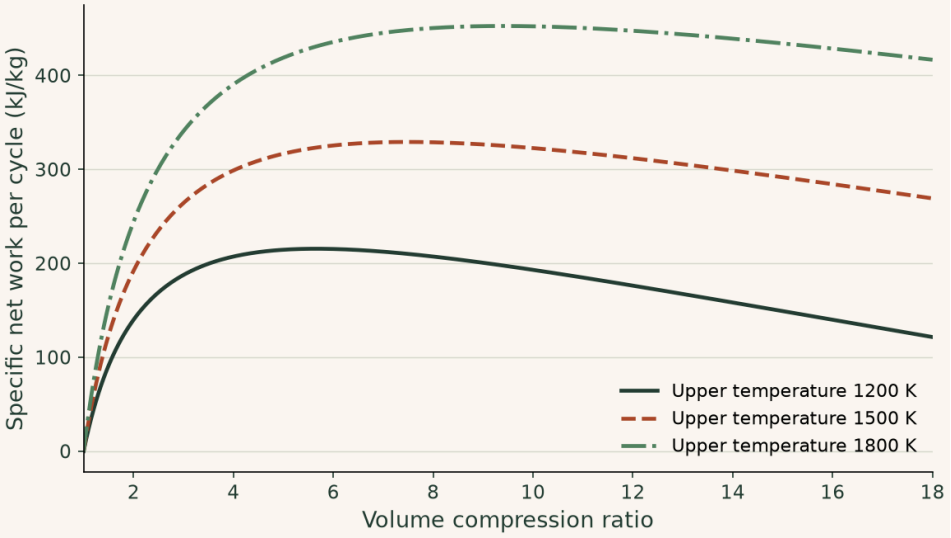
Where does raising compression stop increasing work?

Set the cold state to 300 K and keep each upper temperature fixed. With constant $\gamma = 1.4$ and $c_v = 0.718$ kJ/(kg K), the isentropic legs give

$$T_2 = T_1 r^{\gamma-1},$$

$$T_4 = T_3 / r^{\gamma-1}.$$

Subtract the rejected heat from the supplied heat.



At 1500 K the work peaks near a compression ratio of 7.48. The ideal thermal efficiency continues to rise beyond it while the available heat input decreases. The maximum-work ratio moves upward when the admissible upper temperature increases. These curves hold the temperature limits fixed.

▶ **Read values from the plot**

Interactive calculation · change the inputs of this example

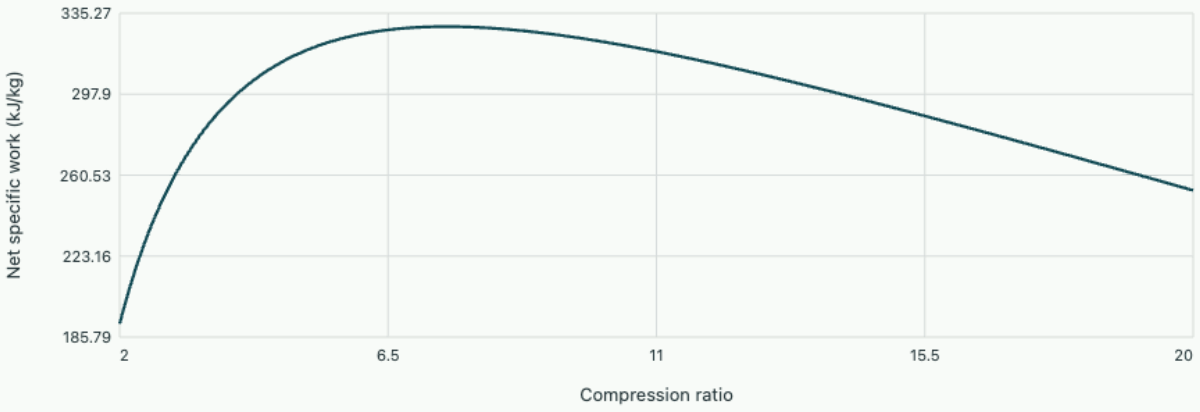
Vary the upper temperature at fixed lower temperature and locate the maximum-work compression ratio.

The curves below are recalculated from this section's equations in your browser. The fixed reference plot above remains available for comparison.

Open this live calculation

Try the calculation

Upper temperature (K)	**1500**
Lower temperature (K)	**300**
Heat-capacity ratio	**1.4**



Fixed temperature limits

Maximum-work compression ratio 7.477. Positive heat addition requires r below 55.902. The fixed cv is 0.718 kJ/(kg K).

Problem sheet 2 · Entropy production and performance limits

Calculate an irreversibility and use a reversible benchmark to assess engines and refrigerators.

Two insulated copper blocks

Blocks of masses 2 and 1 kg at 400 and 280 K touch inside an insulated enclosure. Use $c = 0.390\,\mathrm{kJ/(kg\,K)}$. Find the final temperature and total entropy generation. Why may one block lose entropy?

▼ **Combine energy and entropy balances**

Energy conservation gives

$$T_f = (2 \times 400 + 280)/3 = 360\,\mathrm{K}.$$

The warmer block has $\Delta S_h = 0.780 \ln(360/400) = -0.0822\,\mathrm{kJ/K}$. The other gains $0.390 \ln(360/280) = 0.0980\,\mathrm{kJ/K}$. The net 0.0158 kJ/K is generated internally. Entropy crosses the interface with heat, so only the total for the insulated pair has to be nondecreasing.

Interpret a heat-engine test

An engine absorbs 15 kW from a reservoir at 750 K and rejects 10 kW to a second reservoir at 300 K. Find its power, thermal efficiency, entropy generation and reversible maximum power at the same heat input.

▼ **Interpret the missing output**

The cyclic energy balance gives $15 - 10 = 5\,\mathrm{kW}$ of power and efficiency $5/15 = 0.333$. The reservoirs generate a combined entropy change rate $10/300 - 15/750 = 0.01333\,\mathrm{kW/K}$; the working substance returns to its initial state. A reversible engine reaches $1 - 300/750 = 0.600$, or 9 kW. The difference of 4 kW equals $300\dot{S}_{\mathrm{gen}}$ at this sink temperature.

Measure cooling at the evaporator

Refrigerant enthalpies before and after compression are 405 and 451 kJ/kg. Liquid leaves the condenser at 257 kJ/kg and throttles into the evaporator. For 0.045 kg/s, calculate cooling duty, compressor input and heating COP. If the evaporator and condenser reservoirs are at 263 and 308 K, check the heating COP against its reversible bound.

▼ **Close the cycle**

The valve gives

$$h_4 = h_3 = 257\,\mathrm{kJ/kg}.$$

Evaporation absorbs $0.045(405 - 257) = 6.66\,\mathrm{kW}$; compression requires

$$0.045(451 - 405) = 2.07\,\mathrm{kW}.$$

The condenser duty is 8.73 kW, hence the heating COP is 4.22. The reversible heating bound is

$$308/(308 - 263) = 6.84.$$

Both COPs refer to heat delivered to the warm side per unit work, which permits values above unity.

Test a heat input against a peak-temperature limit

Start an ideal air-standard Otto cycle at 290 K with $r = 7.00$, $\gamma = 1.40$ and $c_v = 0.718\,\mathrm{kJ\,kg^{-1}K^{-1}}$. An electrical heater supplies $500\,\mathrm{kJ\,kg^{-1}}$ at constant volume. Determine the peak temperature and the net work. If the permitted peak is 1250 K, find the allowable input and the resulting net work.

▼ **Show the worked solution**

Compression gives

$$T_2 = 290(7^{0.4}) = 631.6\,\mathrm{K}.$$

The nominal heat input raises this to $T_3 = T_2 + 500/0.718 = 1328.0\,\mathrm{K}$, exceeding the limit by 78 K. Expansion would give

$$T_4 = 609.7\,\mathrm{K}.$$

At this compression ratio $\eta = 1 - 7^{-0.4} = 0.54084$, so the nominal work is $270.4\,\mathrm{kJ\,kg^{-1}}$. Limiting the supplied heat to $0.718(1250 - 631.6) = 444.0\,\mathrm{kJ\,kg^{-1}}$ respects the peak. Net work becomes $240.1\,\mathrm{kJ\,kg^{-1}}$. The compression ratio and ideal efficiency are unchanged.

Revisit chapter 4 · How does piston compression ratio set the ideal cycle efficiency?.

Explore the governing equations after solving the sheet: Entropy, irreversibility and useful-work loss; Thermodynamic cycles, refrigeration and phase change. Compare the assumptions and units before transferring your chosen parameters; the controls may cover a different numerical range from the written exercise.

Chapter 5 · Steady conduction, thermal resistance and insulation

Derive the heat leak through a cylindrical wall and identify when increasing insulation thickness can increase heat transfer.

FROM FIRST PRINCIPLES

Derive series thermal resistance from conservation of heat rate

Consider conduction along a prescribed path through a stationary material, with local area $A(x)$ normal to the heat flow and conductivity $k(x) > 0$. Assume a one-dimensional steady state, no side losses and no heat generation. A balance on an elementary slice gives a single heat rate $\dot{Q}$ at every position. Fourier's measured constitutive relation links that heat rate to the temperature gradient:

$$\frac{d\dot{Q}}{dx} = 0,$$

$$\dot{Q} = -k(x)A(x)\frac{dT}{dx}.$$

Separate variables and integrate from the hot to the cold surface. An exterior fluid film described by $\dot{Q} = hA_s(T_s - T_\infty)$ introduces an additional drop proportional to the same heat rate:

$$T_{\text{hot}} - T_\infty = \dot{Q}\left[\int_{\text{solid}} \frac{dx}{k(x)A(x)} + \frac{1}{hA_s}\right].$$

The bracket is a thermal resistance in K/W. For a plane homogeneous layer it gives $L/(kA)$; for radial conduction in a cylinder the local area $2\pi rL$ varies, giving a logarithm. Adding resistances is justified by a common heat rate. Heat-generating layers or side branches require a separate balance before using a series formula.

DERIVATION AND APPLICATIONS

How the changing area sets resistance in a spherical shell

Heat passing through the insulation of a spherical vessel spreads over a larger surface as it travels outward. Treat an inner surface at radius r_i as isothermal, a concentric outer radius r_o, constant thermal conductivity k, and uniform outside convection coefficient h. Assume steady radial conduction with no internal heat generation, perfect interfaces and a uniform ambient temperature T_∞. With no accumulation between concentric spheres, the heat rate is constant even though local heat flux decreases outward.

Fourier’s empirical law specifies flux from temperature gradient. Multiply it by the local area and rearrange so that each radial slice contributes its own temperature drop:

$$\dot{Q} = -4\pi r^2 k\, dT/dr,$$

$$T_i - T_o = \frac{\dot{Q}}{4\pi k}\int_{r_i}^{r_o} \frac{dr}{r^2}.$$

The outer convective drop equals the same heat rate divided by the outer film conductance. Adding both drops gives a series relation:

$$\frac{T_i - T_\infty}{\dot{Q}} = R = \frac{1}{4\pi k}\left(\frac{1}{r_i} - \frac{1}{r_o}\right) + \frac{1}{4\pi h r_o^2}.$$

Increasing r_o adds conduction resistance but also raises film area. The sign of the derivative decides which effect dominates:

$$\frac{dR}{dr_o} = \frac{1}{4\pi k r_o^2} - \frac{2}{4\pi h r_o^3},$$

$$r_{\text{critical}} = 2k/h.$$

For an initially smaller radius, adding a thin layer can reduce total resistance; after this minimum the resistance grows. The result presumes fixed h and surface temperature T_i. If changes in geometry substantially change natural convection or radiation, the assumed film law must be revisited.

WORKED EXAMPLE

Evaluate an insulated sensor bead

Change the inputs and test this calculation.

Take $r_i = 3.00$ mm, $r_o = 6.00$ mm, $k = 0.0400$ W/(m K) and $h = 10.0$ W/(m² K). The shell contributes 331.6 K/W and the film 221.0 K/W. A 40.0 K inner-to-ambient difference gives 0.0724 W. The uncoated bead had outer resistance 884 K/W and lost only 0.0452 W. At 6 mm the coating is still below the 8.00 mm critical radius.

SECOND APPLICATION

Determine whether a change in airflow reverses the trend

Change the inputs and test this calculation.

With forced airflow raising h to 40.0 W/(m² K), the critical radius falls to 2.00 mm, inside the bare bead. The bare and coated resistances become 221 and $331.6 + 55.3 = 386.9$ K/W. The coating now reduces heat loss from 0.181 to 0.103 W. The same layer can change the sign of its effect because the surrounding film is part of the experiment.

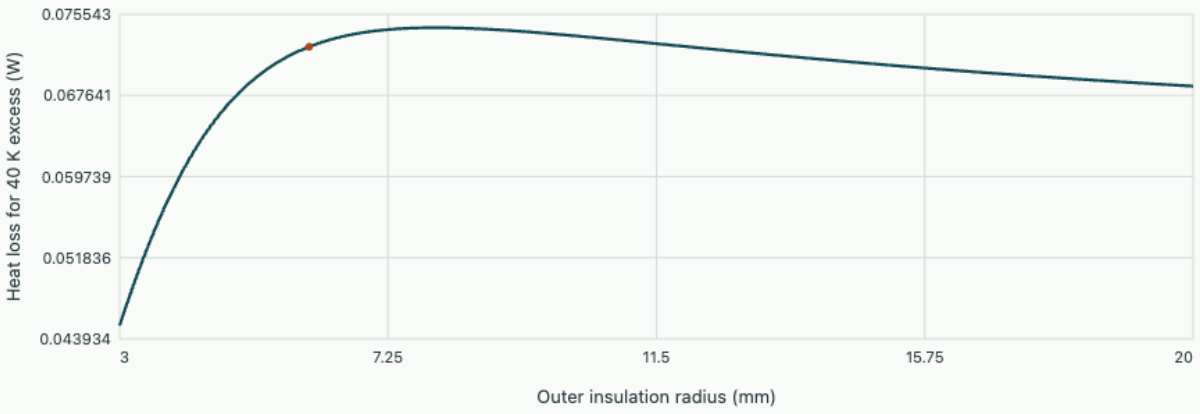
Try it · Spherical insulation

Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Insulation conductivity (W/m/K)	**0.04**
Outer convection coefficient (W/m²/K)	**10**
Selected outer radius (mm)	**6**



Spherical shell + surface film | Selected coating

Inner radius 3 mm. Critical radius 2k/h = 8 mm. Bare loss 0.0452389 W; selected coated loss 0.0723823 W. The film coefficient and inner temperature are held fixed.

Turn heat flux into a temperature field

Energy moves through a stationary homogeneous solid down its temperature gradient. For isotropic conduction Fourier's law is

$$\mathbf{q} = -k\nabla T,$$

where $\mathbf{q}$ is heat flux in W/m² and k is thermal conductivity. A steady one-dimensional wall with no internal heat generation passes the same heat rate at every cut. Heat flux is uniform only when the area of the cuts is uniform. That distinction matters for pipes and spherical shells.

Let an inner cylindrical surface at radius r_i be warmer than the outer one at r_o. Assume a long concentric wall of length L, constant k, negligible end loss and axisymmetry. Fourier's law gives

$$\dot{Q} = -2\pi r L k \, dT/dr,$$

with outward rate taken positive. The area grows with radius, so the magnitude of dT/dr decreases outward even though the heat rate is unchanged.

Derive and combine thermal resistances

Separate and integrate over the cylindrical wall. This yields

$$T_i - T_o = \dot{Q} \ln(r_o/r_i)/(2\pi k L).$$

The cylindrical conduction resistance is the factor multiplying $\dot{Q}$. At an outer convection boundary $\dot{Q} = h \, 2\pi r_o L (T_o - T_\infty)$, giving another resistance between the surface and the bulk surroundings. For a prescribed inner surface temperature,

$$\dot{Q} = \frac{T_i - T_\infty}{\ln(r_o/r_i)/(2\pi k L) + 1/(2\pi r_o L h)}.$$

For multiple concentric layers, add the conduction resistance of each. The drops across series resistances sum because the same rate passes through them all. Poor contact between layers can add another temperature jump. A parallel heat leak bypassing the insulation carries part of the total heat under a common temperature difference, so its conductance adds instead.

Estimate the heat leak from a warm pipe

Take $L = 1.0$ m, $r_i = 0.020$ m and $r_o = 0.040$ m. The insulation has $k = 0.050$ W/(m K) and the surroundings have $h = 10$ W/(m² K), $T_\infty = 293$ K. The inner insulation surface is held at 353 K. The conduction resistance is $\ln 2/[2\pi(0.050)(1)] = 2.206$ K/W, and the outer film resistance is

$$1/[10(2\pi)(0.040)(1)] = 0.398 \, \mathrm{K/W}.$$

Total heat loss is $60/(2.206 + 0.398) = 23.0$ W per metre of pipe.

The outer surface is

$$293 + (23.0)(0.398) = 302.2 \, \mathrm{K}.$$

The rest of the 60 K difference is within the insulation. The heat flux at the outer surface is half that at the inner surface, because the heat-transfer area is twice as large. Assigning the outer circumference to the entire wall in a flat-wall formula would miss this variation. When thickness is very small relative to radius, expanding the logarithm recovers the plane-wall approximation.

Explain the critical radius

Increasing r_o makes the conduction resistance grow but makes the convection resistance shrink. At fixed T_i, T_∞, h, k, differentiating the total resistance gives

$$dR/dr_o = 1/(2\pi L k r_o) - 1/(2\pi L h r_o^2).$$

Its minimum occurs at $r_c = k/h$; that radius maximises heat loss. Here $r_c = 0.0050$ m, already below the bare pipe radius, so adding this insulation reduces loss from the outset. A much thinner wire could initially lose more heat because its exposed area increases faster than the wall resistance. This conclusion assumes the same external coefficient as the diameter changes. If natural convection or radiation changes appreciably, the boundary model and its temperature dependence must be reconsidered.

Recover an unknown contact resistance

A 0.010 m² test sandwich passes 8 W at steady state between measured outer face temperatures of 50 and 40 °C. Its measured total resistance is

$$10/8 = 1.25\,\mathrm{K/W}.$$

The two known solid layers together account for 0.75 K/W, leaving 0.50 K/W attributable to the interface if other parasitic paths are small. The area-normalised contact resistance is $R'' = AR = 0.0050\,\mathrm{m^2K/W}$, corresponding to conductance $1/R'' = 200\,\mathrm{W/(m^2K)}$. A pair of thermocouples on opposite sides of the interface should then show a 4 K temperature difference. That independent measurement helps distinguish contact resistance from an unaccounted external heat loss.

Derive the temperature field along a cooling fin

A fin adds area for convection but must first conduct heat from its base to that area. Treat a straight fin with constant cross-sectional area A_c, exposed perimeter P, conductivity k and uniform film coefficient h. Its base is held at T_b and surrounding air at T_∞. Assume the temperature is nearly uniform across each fin section; a small transverse Biot number supports this approximation. Take the tip as adiabatic, appropriate when tip area is small or insulated. Define excess temperature $\theta = T - T_\infty$.

A slice dx loses heat $hP\theta\,dx$ to convection. Its decrease in axial heat rate must supply that loss. Fourier's law $Q_x = -kA_c\,d\theta/dx$ then yields $\theta'' - m^2\theta = 0$, where $m = \sqrt{hP/(kA_c)}$. The boundary conditions are

$$\theta(0) = \theta_b,\ \theta'(L) = 0.$$

Using a hyperbolic cosine referenced to the insulated end gives

$$\theta(x) = \theta_b \frac{\cosh[m(L-x)]}{\cosh(mL)},$$

$$Q_b = \sqrt{hPkA_c}\,\theta_b \tanh(mL).$$

If the whole fin stayed at base temperature, heat removal would be $hPL\theta_b$. The ratio $\eta_f = \tanh(mL)/(mL)$ is the fin efficiency under the stated tip model. A very long fin can have low efficiency although it removes more heat than a short one; efficiency measures use of installed area, while total heat removal measures cooling duty.

Decide whether doubling fin length is worthwhile

An aluminium fin has rectangular section 20 mm by 2 mm, length 50 mm and conductivity 180 W/(m K). Air at 25 °C flows over it with $h = 35\,\mathrm{W/(m^2K)}$, and the base is at 75 °C. Thus

$$A_c = 4.0 \times 10^{-5}\,\mathrm{m^2},\ P = 0.044\,\mathrm{m},\ \theta_b = 50\,\mathrm{K}.$$

The transverse length $A_c/P = 0.909\,\mathrm{mm}$ gives $hA_c/(kP) = 0.000177$, consistent with small cross-sectional gradients. The fin parameter is $m = \sqrt{35(0.044)/[180(4\times10^{-5})]} = 14.6\,\mathrm{m^{-1}}$; hence $mL = 0.731$ and $\tanh(mL) = 0.624$.

The base heat rate is

$$\sqrt{35(0.044)(180)(4\times10^{-5})}(50)(0.624) = 3.29\,\mathrm{W}.$$

Fin efficiency is

$$0.624/0.731 = 0.854.$$

Doubling length makes $mL = 1.46$ and raises heat removal to about $5.27\tanh(1.46) = 4.73\,\mathrm{W}$. Twice the material therefore gives only 44% more cooling. The original covered base area by itself would convect $35(4\times10^{-5})(50) = 0.070\,\mathrm{W}$, so the fin still offers a large improvement. Adding several fins can change the film coefficient through crowding, and finite contact resistance can lower the effective base temperature. Those effects must be checked before multiplying a single-fin heat rate by the number installed.

A surface thermometer offers a second way to test the heat-transfer model. The insulated-tip solution gives

$$T(L) = 25 + 50/\cosh(0.731) = 64.1^\circ\mathrm{C}.$$

If a thin thermocouple instead measures nearly 75 °C all along a fin while the base still sheds only the predicted few watts, the effective film coefficient may be smaller than assumed. A much colder measured tip could indicate strong local tip convection or poor base contact. Interpret a profile together

with the base heat rate: several combinations of conductance, base temperature and film conditions can reproduce one measured temperature, but they need not reproduce both the profile and the total cooling duty.

The root of the fin offers a particularly direct simultaneous measurement of axial temperature gradient and supplied heat rate.

Infer wall cooling from a measured skin-friction force

The film coefficient in a thermal-resistance calculation carries information about transport within the flow. For a smooth isothermal plate, take a steady parallel free stream of speed U and temperature T_e. Define $\theta = (T - T_w)/(T_e - T_w)$. Both u/U and θ rise from zero at the wall to one in the outer stream. If momentum and heat diffuse at equal rates, their transport equations and boundary conditions match for a constant-property, zero-pressure-gradient boundary layer starting at the same leading edge. Under this approximation their profiles and their normalized wall gradients coincide. In turbulent flow the analogy additionally relies on similar turbulent diffusivities.

For a wall hotter than the stream the heat flux into the fluid and the magnitude of skin friction are $q'' = k(T_w - T_e)\partial_y\theta|_w$ and $\tau_w = \mu U \partial_y(u/U)|_w$. Divide the equations. At $Pr = \mu c_p/k = 1$, the unknown gradient cancels. With $q'' = h(T_w - T_e)$ the Reynolds analogy becomes

$$h \simeq \frac{c_p \tau_w}{U},$$

$$St = \frac{h}{\rho U c_p} \simeq \frac{C_f}{2},$$

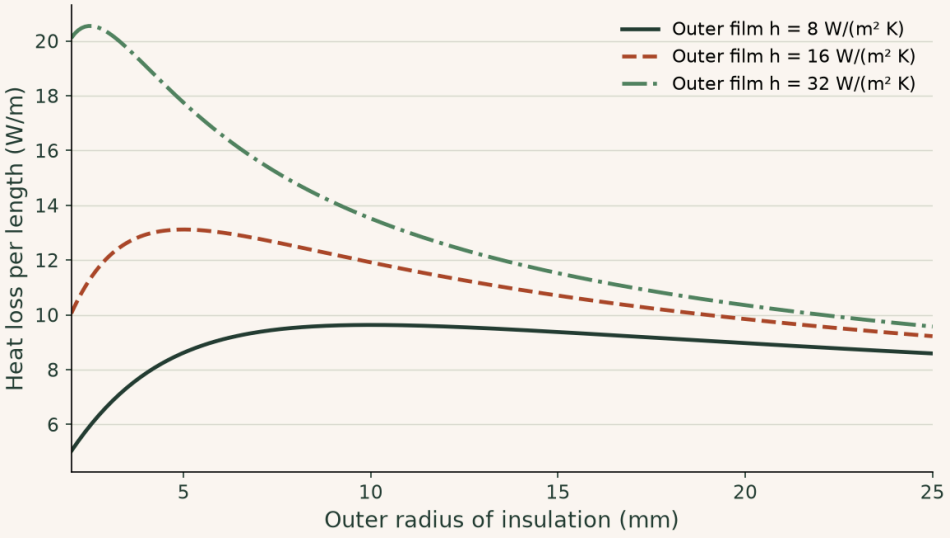
$$C_f = \frac{\tau_w}{\frac{1}{2}\rho U^2}.$$

A small heat-transfer tunnel has effective Pr close to one. At $U = 10.0\,\mathrm{m\,s^{-1}}$, the stream has $\rho = 1.20\,\mathrm{kg\,m^{-3}}$ and $c_p = 1000\,\mathrm{J\,kg^{-1}K^{-1}}$. A force balance isolates the skin friction on one flush plate face of area $0.50\,\mathrm{m^2}$, and measures $0.180\,\mathrm{N}$. The mean shear is $0.360\,\mathrm{Pa}$, hence $\bar{C}_f = 0.00600$ and $\bar{h} \simeq 36.0\,\mathrm{W\,m^{-2}K^{-1}}$. A uniform wall 20 K above the stream loses approximately $36(0.50)(20) = 360\,\mathrm{W}$ by convection. Average drag produces an average coefficient because both quantities are integrated over the same face. Pressure drag from a protruding instrument, an unheated entry length or a fluid with a very different Prandtl number would invalidate that identification.

Practise on sheet 3 · How much skin friction is consistent with a cooling requirement?. Further reading · MIT · 16.Unified Thermodynamics and Propulsion · §17.1 The Reynolds Analogy.

Can adding insulation initially increase heat loss?

A long wire of radius 2 mm is held 50 K above ambient. Insulation of conductivity 0.08 W/(m K) adds radial resistance but increases exterior convection area. Add cylindrical conduction and convection resistances per unit length before converting back to heat rate.



At h = 8 W/(m² K) heat loss peaks at the critical radius k/h = 10 mm; at h = 32 it peaks at 2.5 mm. This effect assumes constant surface film coefficient and prescribed inner temperature. A fixed-power heater poses a different question.

▶ **Read values from the plot**

Interactive calculation · change the inputs of this example

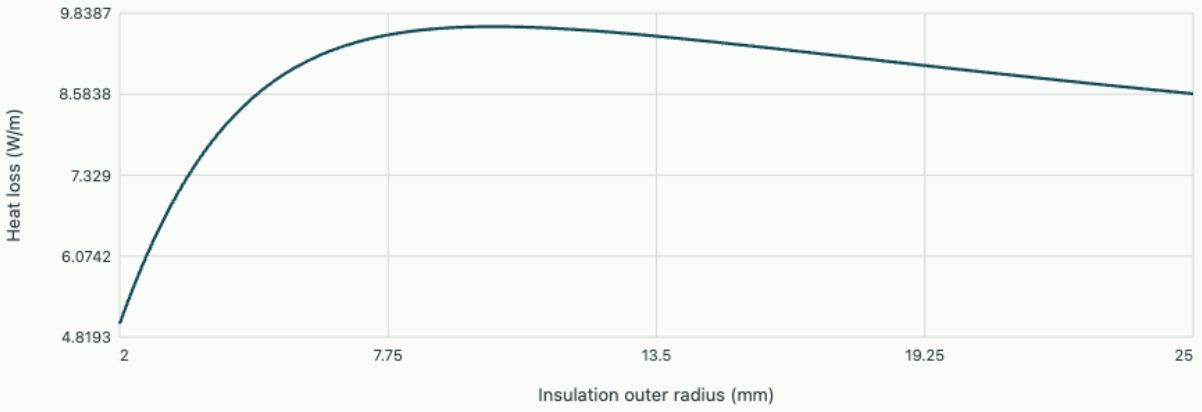
Raise the film coefficient until the critical radius falls below the wire radius. Check whether adding insulation can still increase heat loss.

The curves below are recalculated from this section's equations in your browser. The fixed reference plot above remains available for comparison.

Open this live calculation

Try the calculation

Conductivity (W/m K)	**0.08**
Film coefficient (W/m² K)	**8**
Wire overtemperature (K)	**50**



Conduction and convection

Critical radius k/h = 10.000 mm. The core radius is fixed at 2 mm and its temperature is prescribed.

Study the full derivation

How a gravity-drained condensate film limits heat transfer · Thermal processes and transport in engineering

Saturated vapour condenses on a cool vertical plate. A liquid film thickens as it drains, adding thermal resistance toward the bottom. Derive the local heat-transfer coefficient from coupled momentum and energy balances, then determine which observations would invalidate the smooth-film model.

Chapter 6 · Thermal diffusion and transient surface cooling

Use diffusion time, Biot number and penetration depth to choose between a uniform-body model and a spatial temperature field.

FROM FIRST PRINCIPLES

Derive thermal diffusion from stored energy and conductive flux

Let a stationary homogeneous solid have density ρ, specific heat c and thermal conductivity k. Assume no heat generation and consider one-dimensional temperature variation. A slice of area A and thickness dx stores energy at rate $\rho c A\, dx\, \partial_t T$. The difference between the conductive flux entering and leaving gives

$$\rho c A\, dx\, \partial_t T = A[q''(x) - q''(x+dx)] \quad \Longrightarrow \quad \rho c\, \partial_t T = -\partial_x q''.$$

Fourier's law $q'' = -k\partial_x T$ closes this balance. For constant properties the diffusion equation and its characteristic time over distance L follow:

$$\partial_t T = \alpha\, \partial_{xx} T,$$
$$\alpha = \frac{k}{\rho c}, \qquad t_{\mathrm{diff}} \sim L^2/\alpha.$$

Initial temperatures and boundary heat-flux or temperature conditions are needed to select a specific solution. If conduction inside the solid is sufficiently rapid relative to transfer at its surface, one can integrate the balance using an almost uniform interior temperature. Comparing the internal and external resistance scales yields the Biot number hL_c/k used for that simplification.

Background · Biot number

$$\mathrm{Bi} = hL_c/k$$

Here h is a surface heat-transfer coefficient and k is conductivity. A small Biot number supports a nearly uniform specimen temperature for the specified boundary condition.

DERIVATION AND APPLICATIONS

Predict how a thermal probe follows a changing fluid temperature

A sensor calibrated by an abrupt temperature step can lag behind a fluid that is itself heating steadily. Take a body of mass m, heat capacity c, exposed area A and film coefficient h with negligible internal temperature gradients. This lumped description requires internal conduction to be fast compared with

surface heat transfer. Treat properties and h as constant and neglect self-heating, radiation and lead conduction. Denote the body temperature by T, fluid temperature by T_f, and set $\tau = mc/(hA)$.

An instantaneous energy balance gives a first-order differential equation. For a linearly increasing fluid temperature $T_f = T_0 + \beta t$, define the lag $D = T_f - T$:

$$mc\dot{T} = hA(T_f - T),$$
$$\dot{D} + D/\tau = \beta.$$

Suppose the probe and fluid start in equilibrium at T_0. Multiply the lag equation by $e^{t/\tau}$ and integrate from zero. The integrating factor converts the left side to a single derivative:

$$e^{t/\tau} D(t) = \beta \int_0^t e^{s/\tau}\, ds = \beta\tau(e^{t/\tau} - 1).$$

Subtract the resulting lag from the prescribed fluid history to obtain the reading:

$$T(t) = T_0 + \beta[t - \tau(1 - e^{-t/\tau})].$$

For $t \ll \tau$ the rise begins quadratically, since the initial film temperature difference was zero. For $t \gg \tau$ the probe warms at the same rate as the fluid but reads low by $\beta\tau$. Cooling ramps reverse the sign. A measured constant lag is therefore evidence for this model only after the startup transient has decayed.

WORKED EXAMPLE

Evaluate the reading during a ramp calibration

Change the inputs and test this calculation.

A probe has $mc = 0.800$ J/K and $hA = 0.100$ W/K, hence $\tau = 8.00$ s. Start at 20.0°C and heat the fluid by 0.500 K/s. At $t = 16.0$ s the fluid is at 28.0°C and the lag is $4.00(1 - e^{-2}) = 3.46$ K, so the probe reads 24.5°C. Simply subtracting the eventual 4.00 K would overcorrect during startup.

SECOND APPLICATION

Estimate a time constant without an abrupt temperature step

Change the inputs and test this calculation.

During a long calibration ramp of 0.200 K/s, a second probe remains 1.20 K below an independent fluid reference. Its inferred time constant is

$$\tau = 1.20/0.200 = 6.00\,\mathrm{s}.$$

If the fluid is then held fixed, the remaining lag decays exponentially. Waiting $6.00\ln(1.20/0.100) = 14.9$ s brings the error below 0.100 K. The hold and the ramp test two different

responses using the same fitted parameter.

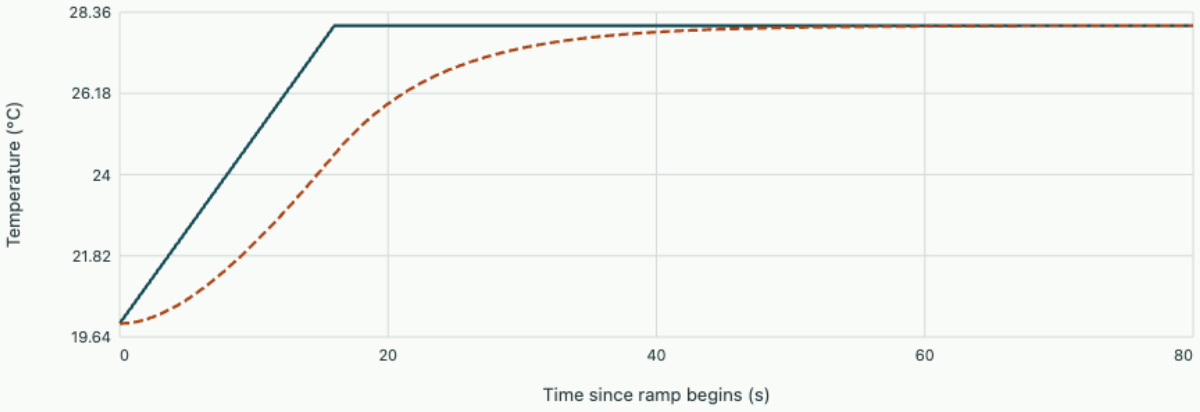
Try it · Probe ramp

Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Probe time constant (s)	8
Fluid ramp rate (K/s)	0.5
Ramp duration before hold (s)	16



At end of ramp fluid 28 °C, probe 24.5413 °C, signed lag 3.45866 K. Long-ramp lag βτ=4 K; the hold tests the same time constant independently. Internal temperature gradients are neglected.

Allow energy to accumulate in the material

A surface thermocouple responds rapidly to a cold wash, while the interior of a thick part can remain hot. To predict the time needed to cool the centre, steady thermal resistance alone is insufficient. Enclose a differential slice of material. Incoming conductive heat minus outgoing heat plus local generation equals the rate of change of its stored energy. Substituting Fourier’s law in a homogeneous stationary solid gives

$$\rho c_p \frac{\partial T}{\partial t} = k\nabla^2 T + \dot{q}_v,$$

$$\alpha = \frac{k}{\rho c_p}.$$

Thermal diffusivity α in m²/s compares a material's ability to spread heat with its ability to store it. A large conductivity does not alone guarantee fast adjustment if volumetric heat capacity is also large. Without heat generation, a length L gives a characteristic diffusion time $t_d = L^2/\alpha$. The dimensionless Fourier number $Fo = \alpha t/L^2$ measures how long diffusion has acted relative to that scale.

Decide whether the interior stays nearly uniform

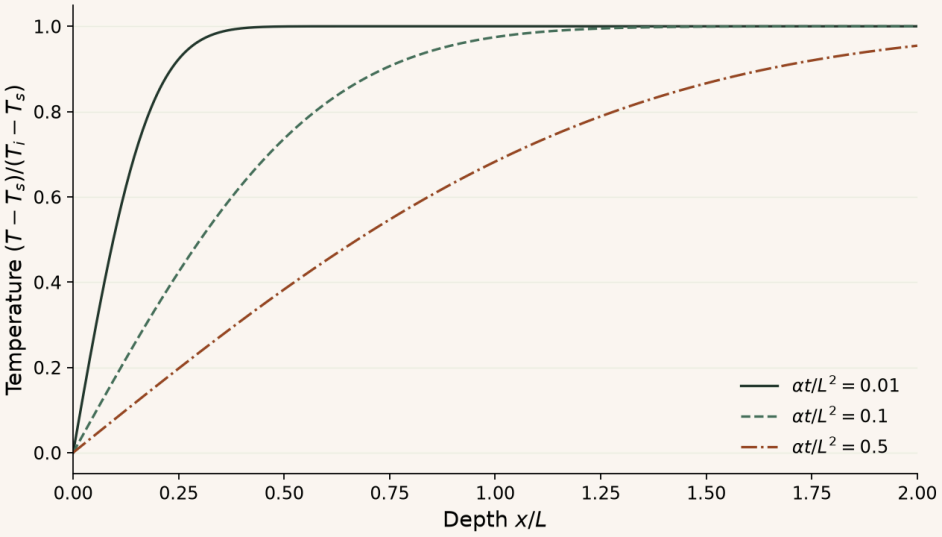
At a convecting boundary $-k\partial T/\partial n = h(T_s - T_\infty)$. Comparing the resistance of the interior with a fluid film leads to $Bi = hL_c/k$, often using $L_c = V/A_s$ for a lumped model. If this ratio is small, commonly below about 0.1 as a first engineering check, the internal temperature difference is small compared with that across the film. Integrating energy over a uniform body then gives

$$\rho c_p V\, dT/dt = -hA_s(T - T_\infty),$$

$$(T - T_\infty)/(T_0 - T_\infty) = e^{-t/\tau},$$

with $\tau = \rho c_p V/(hA_s)$.

The Biot and Fourier numbers answer different questions. A small Fourier number says a disturbance has penetrated only a short distance compared with L. A small Biot number says the boundary transfers heat slowly enough for internal gradients to remain small. Choose the length for each comparison explicitly, especially for a slab cooled on one side rather than both.



Semi-infinite cooling profiles plotted against a fixed reference length L. Increasing time by a factor of ten increases the depth of a given temperature ratio by the square root of ten.

Solve a sudden prescribed-temperature change

Consider initially uniform material occupying $x > 0$ whose surface is suddenly held at T_s, with its far interior still at T_i. Neglect internal generation and use constant properties. The initial and boundary conditions have no geometric length scale. Try the similarity coordinate $\eta = x/(2\sqrt{\alpha t})$ and set $\Theta = (T - T_s)/(T_i - T_s)$. Differentiating and substituting into the diffusion equation yields $\Theta'' + 2\eta\Theta' = 0$, with derivatives in η. Integrating once gives

$$\Theta' = Ce^{-\eta^2}.$$

The limits $\Theta(0) = 0,\ \Theta(\infty) = 1$ give

$$\Theta = \operatorname{erf}(\eta).$$

At $x = 0$ the imposed surface temperature holds exactly. At a fixed depth and very early time η is large, so the temperature tends to its initial value. The distance of any fixed fractional temperature change grows as $\sqrt{t}$. Quadrupling exposure time doubles that distance, a scaling quite different from a body travelling at a constant speed.

Predict depth and validate the time window

A thick solid initially at 373 K is held at 293 K at its surface. With $\alpha = 1.0 \times 10^{-5}$ m²/s and $t = 10$ s, $2\sqrt{\alpha t} = 0.020$ m. At depth 10 mm, $\eta = 0.50$ and $\operatorname{erf}(0.50) = 0.5205$, giving $T = 293 + 80(0.5205) = 334.6$ K. At 20 mm, $\operatorname{erf}(1) = 0.8427$ gives 360.4 K. The surface has cooled by 80 K while the deeper sensor has cooled by only 12.6 K.

At 40 s the 20 mm sensor reaches the same fraction previously measured at 10 mm. This prediction assumes the remote boundary has not begun to influence the measurement and that the surface temperature is actually prescribed. A finite slab or a convective wash may require a finite-domain diffusion solution instead. Record surface and interior histories to distinguish uncertain boundary conditions from uncertain diffusivity.

Interpret a penetration depth

If two embedded sensors are at 2 and 6 mm in the same semi-infinite specimen, the deeper sensor reaches the same normalised temperature at nine times the shallow sensor’s time because $x/\sqrt{\alpha t}$ must match. An observed factor much smaller than nine can indicate a heat leak along a probe or a far boundary that has begun to influence the temperature field.

Resolve a slab with a finite surface film

A prescribed surface temperature and a uniformly cooling lump are two useful limits. A finite slab exchanging heat through a finite film generally satisfies neither. Let L be half-thickness, $x = 0$ the

midplane, and T_i the initial uniform temperature. Both sides suddenly see fluid at T_∞ with constant h. Using $X = x/L,\ Fo = \alpha t/L^2$ and $\Theta = (T - T_\infty)/(T_i - T_\infty)$, the diffusion equation becomes

$$\Theta_{Fo} = \Theta_{XX},$$

with symmetry $\Theta_X(0) = 0$ and film condition $-\Theta_X(1) = Bi\,\Theta(1)$, where $Bi = hL/k$.

Seek a separated mode $\Theta = \cos(\lambda X) e^{-\lambda^2 Fo}$. Its derivative vanishes at the midplane. Imposing the convective surface condition gives

$$\lambda \tan \lambda = Bi.$$

The uniform initial condition fixes modal amplitudes by cosine orthogonality on $0 < X < 1$:

$$\Theta = \sum_n A_n \cos(\lambda_n X) e^{-\lambda_n^2 Fo},$$

$$A_n = \frac{4 \sin \lambda_n}{2\lambda_n + \sin(2\lambda_n)}.$$

For $Bi \ll 1$, the smallest root has $\lambda_1^2 \simeq Bi$, its shape approaches a constant, and $A_1 \simeq 1$. The decay becomes $e^{-Bi\,Fo}$, exactly the lumped time scale because $Bi\,Fo = ht/(\rho c L)$. For a large film coefficient the first root tends towards $\pi/2$, corresponding to a prescribed surface temperature. The same boundary-value model thus recovers both limiting behaviours.

Predict when the core and surface meet different targets

A 40 mm thick flat slab with $k = 2.0\,\mathrm{W/(m\,K)}$ and $\alpha = 1.0 \times 10^{-6}\,\mathrm{m^2/s}$ is cooled on both sides by $h = 50\,\mathrm{W/(m^2 K)}$. It starts at 100 °C in fluid at 20 °C. Here $L = 0.020$ m and $Bi = 0.50$. The first root of $\lambda \tan \lambda = 0.5$ is 0.6533 and $A_1 = 1.070$. After 600 s, $Fo = 1.50$. The first term gives centre excess ratio $1.070 \exp[-(0.6533)^2 (1.50)] = 0.564$, hence $T_c = 65.1^\circ\mathrm{C}$. At the surface multiply by $\cos(0.6533) = 0.794$, giving about 55.9 °C. Higher roots are strongly attenuated at this Fourier number.

A process requiring the centre to fall below 50 °C needs centre ratio $(50 - 20)/(100 - 20) = 0.375$. Solve $Fo = \ln(1.070/0.375)/0.6533^2 = 2.46$, or approximately 984 s. By contrast a lumped estimate gives $Fo = -\ln(0.375)/0.5 = 1.96$, about 785 s. It underestimates the core cooling time by over three minutes because the film and internal gradients both matter. An early-time surface prediction requires more modes than this late-time centre criterion. In particular the first term alone gives a centre ratio above unity at time zero, signalling that the remaining terms are necessary to reconstruct the original uniform state.

If product quality depends on average energy content rather than the hottest spot, average the first-mode shape over half the thickness. Its factor $\int_0^1 \cos(\lambda_1 X) dX = \sin \lambda_1 / \lambda_1 = 0.930$ gives mean excess ratio $0.564(0.930) = 0.525$ after 600 s. The predicted bulk temperature is therefore about 62.0 °C, between centre and surface as expected. A calorimetric measurement made after removing and mixing the material is closer to the mean than to either point thermometer. A decision to stop cooling

based on such a reading can therefore satisfy an average enthalpy requirement while leaving the centre several kelvins above a local temperature limit.

The surface flux can be checked against the rate of thermal energy loss. At 600 seconds each face transfers $50(55.9-20) = 1795\,\mathrm{W/m^2}$ to the air. Per square metre of plan area the slab therefore loses approximately 3590 W. The volumetric heat capacity obtained independently from k/α is $2.0\times10^{6}\,\mathrm{J/(m^3K)}$. Multiplying by total 0.040 m thickness gives an areal capacity of 80000 J/(m² K), and a predicted instantaneous mean cooling rate $3590/80000 = 0.0449\,\mathrm{K/s}$. Differentiating the one-term mean temperature gives the same decay rate to within rounding.

This independent balance is especially useful if measurements suggest two different film coefficients on opposite sides. An unequal pair of measured fluxes cannot be fitted by a strictly symmetric half-slab model even if it reproduces one centre temperature. The full-thickness conduction problem then requires both surface boundary conditions; simply replacing them by their arithmetic mean may match total heat over one interval while shifting the hottest location away from the midplane.

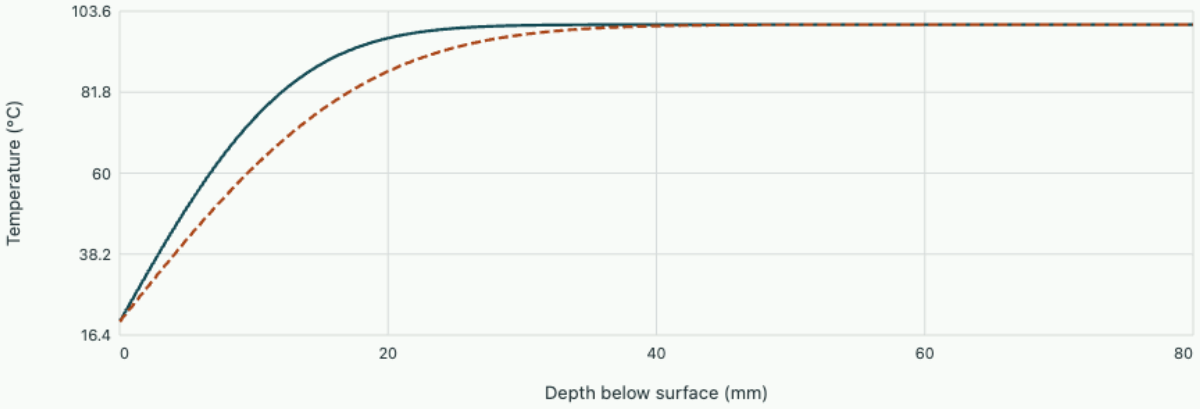
Interactive calculation · Change the inputs of this example

Change the parameters of this reference plot and compare its limiting cases. The fixed illustration remains beside the derivation; the adjustable calculation exposes the same dependence on geometry, material properties or loading.

Open this live calculation

Try the calculation

Thermal diffusivity (mm²/s)	5
Elapsed time (s)	10
Initial temperature (°C)	100



The diffusion scale2√(αt) is 14.142 mm. Semi-infinite solid, initially uniform, surface suddenly held at20°C.

Study the full derivation

Periodic heating as a measurement of thermal diffusivity · Thermal processes and transport in engineering

Two buried thermometers see the same applied surface temperature oscillation with different amplitudes and phases. Derive how a solid filters that oscillation, recover diffusivity independently from both observations and identify when a rear boundary or a poor sensor contact changes the inference.

Problem sheet 3 · Heat paths and temperature histories

Identify which geometry sets a thermal resistance, and decide whether a transient can be represented by one temperature.

A wall with two films

A flat panel of area 2.0 m² and thickness 40 mm has conductivity 0.080 W/(m K). Inside and outside air are at 22 and 2 °C; film coefficients are 8 and 20 W/(m² K). Find the heat rate and both surface temperatures. Which design change would halve the rate at fixed film coefficients?

▼ **Combine the three resistances**

The films give 0.0625 and 0.0250 K/W. The solid contributes $0.040/(0.080 \times 2) = 0.250\,\mathrm{K/W}$. Hence

$$Q = 20/0.3375 = 59.26\,\mathrm{W}.$$

The inside face is $22 - 59.26(0.0625) = 18.30$ °C, the outside face $2 + 59.26(0.025) = 3.48$ °C. Halving the rate requires 0.3375 K/W of extra resistance, or 54 mm more of this material over the same area. Doubling just the existing insulation gives a smaller reduction.

Time to reach a handling temperature

A metal sphere of radius 12 mm cools in air at 20 °C from 140 to 50 °C. Its conductivity is 45 W/(m K), volumetric heat capacity 3.7 MJ/(m³ K), and $h = 35\,\mathrm{W/(m^2K)}$. Assess the uniform-temperature approximation, estimate cooling time and determine the new time if the initial temperature becomes 80 °C.

▼ **Select a length and integrate**

For a sphere $V/A = R/3 = 0.004\,\mathrm{m}$, giving $\mathrm{Bi} = 35(0.004)/45 = 0.00311$. The time constant is

$$3.7 \times 10^{6}(0.004)/35 = 422.9\,\mathrm{s}.$$

The original excess falls from 120 to 30 K, taking $\tau \ln 4 = 586.2\,\mathrm{s}$. With a 60 K initial excess it takes $\tau \ln 2 = 293.1\,\mathrm{s}$. The time constant is unchanged under the given constant-property conditions.

Reproduce a centreline temperature

Two thick slabs have half-thicknesses 10 and 25 mm and the same diffusivity $8.0 \times 10^{-6}\,\mathrm{m^2/s}$. Each starts at the same uniform temperature and suddenly acquires the same prescribed surface temperatures on both faces. If the thinner slab reaches a given centreline temperature after 18 s, find the corresponding Fourier number and the time for the thicker slab. What if surface temperatures are replaced by convective boundary conditions?

▼ **Match dimensionless time and boundary data**

The thinner slab gives

$$\mathrm{Fo} = \alpha t / L^2 = 1.44.$$

Matching it gives

$$t_2 = 18(25/10)^2 = 112.5\,\mathrm{s}.$$

The surfaces and initial profile have the same dimensionless temperatures. For convective cooling the dimensionless temperature history also depends on hL/k, so equal Fourier number alone is insufficient. With the same material, the thicker slab needs a film coefficient $10/25$ times that of the thinner specimen to match Biot number.

How much skin friction is consistent with a cooling requirement?

An isothermal face of area $0.20\,\mathrm{m}^2$ must lose 150 W to a $Pr \simeq 1$ stream. The allowable wall excess is 15 K,

$$c_p = 1000\,\mathrm{J\,kg^{-1}K^{-1}},$$

$$U = 8.0\,\mathrm{m\,s^{-1}}.$$

Estimate the required area-averaged coefficient and mean skin-friction force. The balance reports 0.110 N, of which 0.040 N is support pressure drag. Estimate the achievable heat rate.

▼ **Show the worked solution**

The target coefficient is

$$150/(0.20 \times 15) = 50\,\mathrm{W\,m^{-2}K^{-1}}.$$

The corresponding wall stress is $hU/c_p = 0.400\,\mathrm{Pa}$, or $0.0800\,\mathrm{N}$ on this face. After subtracting support drag the measured skin force is $0.0700\,\mathrm{N}$, giving $\bar{\tau} = 0.350\,\mathrm{Pa}$ and $\bar{h} \simeq 43.75\,\mathrm{W\,m^{-2}K^{-1}}$. At 15 K excess the predicted heat rate is $131.3\,\mathrm{W}$, roughly 19 W below target. Total balance force would overestimate cooling because pressure drag lacks the wall-gradient relation used in the analogy.

Revisit chapter 5 · Infer wall cooling from a measured skin-friction force.

Explore the governing equations after solving the sheet: Steady conduction, thermal resistance and insulation; Thermal diffusion and transient surface cooling. Compare the assumptions and units before transferring your chosen parameters; the controls may cover a different numerical range from the written exercise.

Chapter 7 · Phase equilibrium and steam quality

Locate a two-phase state from two independent measurements and use mass fractions to find its volume and energy.

FROM FIRST PRINCIPLES

Derive the slope of a pure-substance coexistence line

At liquid–vapour equilibrium, transferring a small amount of matter between phases cannot lower the total Gibbs energy. For a pure substance this requires equality of the specific Gibbs energies $g_f = g_g$ at common temperature and pressure. To find how pressure must change when temperature changes while maintaining coexistence, differentiate this equality along the saturation curve. The fundamental differential $dg = -s\,dT + v\,dp$ gives

$$-s_f\,dT + v_f\,dp = -s_g\,dT + v_g\,dp.$$

Rearranging expresses the measured slope as the entropy change of vaporization divided by its specific-volume change. Reversible phase change at constant temperature absorbs specific enthalpy $h_{fg} = T(s_g - s_f)$, so

$$\left(\frac{dp}{dT}\right)_{\text{sat}} = \frac{s_g - s_f}{v_g - v_f} = \frac{h_{fg}}{T(v_g - v_f)}.$$

The units are Pa/K. No ideal-gas approximation has been made. If the vapour is dilute and the liquid volume negligible, further substitution can simplify the slope, but this fails near the critical point. The slope describes neighbouring coexistence states; determining a system's vapour mass fraction additionally requires an extensive property such as specific enthalpy or volume.

Background · Vapour quality

$$x = m_v/(m_v + m_l)$$

This mass fraction applies in a two-phase equilibrium region. Mole fraction, mass fraction and fraction of volume occupied by vapour can be very different.

DERIVATION AND APPLICATIONS

Trace an isentropic pressure change through the saturation dome

A vapour fraction inferred from one property must reproduce the others at the same pressure. Consider a pure substance in local liquid–vapour equilibrium. Let x be the vapour mass fraction and use subscripts f, g for saturated liquid and vapour at the current pressure. Neglect interfacial contributions

and assume each phase has the tabulated saturation state. Specific extensive properties then obey the lever rule because their totals add over the masses of both phases:

$$s = (1 - x)s_f + xs_g = s_f + xs_{fg},$$
$$h = h_f + xh_{fg}.$$

In a steady adiabatic reversible expansion the entropy balance gives

$$s_2 = s_1.$$

At the specified exit pressure solve the entropy relation before using the enthalpy relation:

$$x_{2s} = (s_1 - s_{f,2})/s_{fg,2},$$
$$h_{2s} = h_{f,2} + x_{2s}h_{fg,2}.$$

This interpolation is physical only when $0 \le x_{2s} \le 1$. For an actual adiabatic turbine with negligible kinetic-energy change, use isentropic efficiency $\eta_t = (h_1 - h_2)/(h_1 - h_{2s})$ and check the actual phase state separately:

$$h_2 = h_1 - \eta_t(h_1 - h_{2s}),$$
$$x_2 = (h_2 - h_{f,2})/h_{fg,2}$$

if the actual outlet is still a mixture. Irreversibility reduces extracted work, which can leave the exhaust drier at the same pressure. A drier exit therefore does not by itself demonstrate higher efficiency.

WORKED EXAMPLE

Infer equilibrium quality for two expansion histories

Change the inputs and test this calculation.

At a low-pressure outlet take rounded steam properties $h_f = 192$, $h_{fg} = 2393$ kJ/kg, $s_f = 0.649$ and $s_{fg} = 7.50$ kJ/(kg K). Steam enters with $h_1 = 3200$ kJ/kg and $s_1 = 6.65$ kJ/(kg K). The reversible quality is $(6.65 - 0.649)/7.50 = 0.800$, giving $h_{2s} = 2107$ kJ/kg. For $\eta_t = 0.800$, the actual exit enthalpy is about 2326 kJ/kg and the quality 0.892. The actual specific work is about 874 kJ/kg.

SECOND APPLICATION

Detect when a reported quality has no physical meaning

Change the inputs and test this calculation.

At the same outlet pressure an independent calculation gives $h_2 = 2630$ kJ/kg. Interpolating would give

$$x = (2630 - 192)/2393 = 1.019.$$

Since the saturated vapour enthalpy is only 2585 kJ/kg, the enthalpy lies above the saturation range. The outlet is superheated; finding its temperature requires superheated-vapour properties. Reporting a vapour fraction above unity would hide a change of thermodynamic region.

Try it · Turbine quality

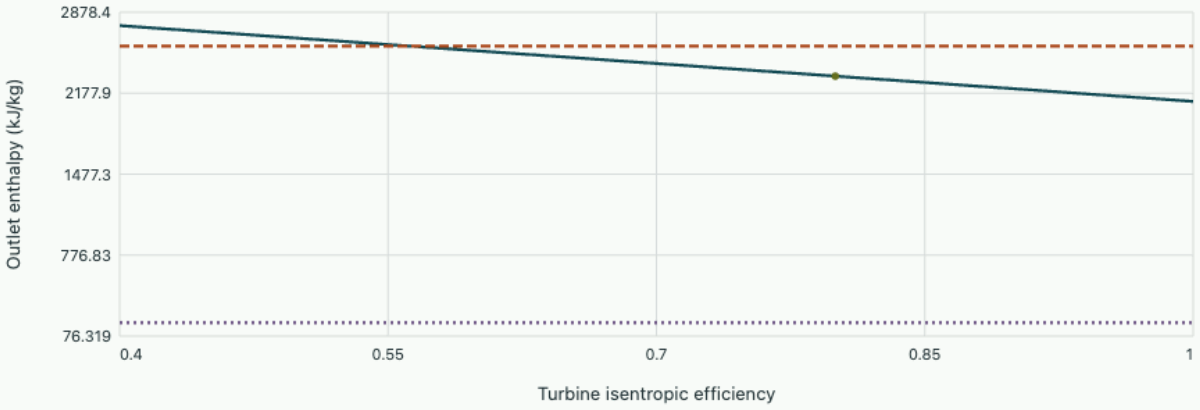
Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Turbine isentropic efficiency **0.8**

Inlet entropy (kJ/kg/K) **6.65**



Isentropic outlet quality 0.800133; selected work 874.625 kJ/kg. Actual mixture quality 0.891507. Rounded outlet saturation properties are fixed as in the derivation.

Temperature can stay fixed while heat enters

In an open pan at a fixed ambient pressure, liquid water begins to boil at a particular temperature. Additional heat can then convert liquid to vapour while the temperature remains close to that value. Specifying both saturation temperature and saturation pressure of a pure substance does not locate a unique mixture state. Along the coexistence curve they are linked, and the fraction of material in each

phase supplies the missing information. Outside that curve pressure and temperature can usually be used as independent coordinates for a simple compressible substance.

Define quality $x = m_g/(m_f + m_g)$ as the vapour mass fraction, using subscripts f for saturated liquid and g for saturated vapour. A property such as volume or enthalpy is additive over the two masses when interfacial contributions are negligible. Writing $V = m_f v_f + m_g v_g$ and dividing by total mass gives the mixing relation. The same derivation applies to specific internal energy, enthalpy and entropy.

$$v = v_f + x(v_g - v_f),$$

$$h = h_f + x h_{fg},$$

$$h_{fg} = h_g - h_f.$$

Quality lies between zero and one in a saturated mixture. A calculated value beyond either endpoint signals a single-phase state or inconsistent measurements. Temperature does not interpolate: both phases in equilibrium have the same temperature.

Distinguish vapour mass from space occupied

A closed vessel contains 2.00 kg of saturated water and steam at approximately 100 kPa in 0.500 m³. Use the supplied rounded values $v_f = 0.00104,\ v_g = 1.694\,\mathrm{m^3/kg}$,

$$h_f = 417.5,\ h_{fg} = 2258\,\mathrm{kJ/kg}.$$

The measured specific volume is 0.250 m³/kg. Quality follows as $(0.250 - 0.00104)/(1.694 - 0.00104) = 0.1471$. Thus vapour mass is 0.294 kg while 1.706 kg remains liquid.

The vapour occupies $0.2941(1.694) = 0.4982\,\mathrm{m^3}$, almost the entire vessel. The liquid occupies roughly 0.0018 m³. Specific enthalpy is

$$417.5 + 0.1471(2258) = 749.6\,\mathrm{kJ/kg}.$$

Specific internal energy follows from $u = h - pv = 749.6 - 100(0.250) = 724.6\,\mathrm{kJ/kg}$ when using kPa and m³/kg. A level indicator measures occupied volume; using its reading directly as a mass fraction would give a large error.

Connect latent heat to the coexistence curve

For two pure phases in equilibrium their specific Gibbs free energies are equal. Along the coexistence line this equality persists, so $dg_f = dg_g$. From $dg = v\,dp - s\,dT$ one obtains

$$dp_{\mathrm{sat}}/dT = (s_g - s_f)/(v_g - v_f) = h_{fg}/[T(v_g - v_f)].$$

At about 373 K using the same values, the slope is

$$2258/[373(1.693)] = 3.58\,\mathrm{kPa/K}.$$

A small increase of boiling pressure by 3.6 kPa therefore raises saturation temperature by about one kelvin locally. Larger changes require integrating the changing properties or reading a saturation table rather than treating this local slope as a constant.

Predict flash after throttling

A saturated liquid with inlet enthalpy 500 kJ/kg passes through an insulated valve to the pressure of the drum above. With negligible work and velocity change, enthalpy remains 500 kJ/kg. The downstream quality is therefore

$$(500 - 417.5)/2258 = 0.0365.$$

Only about 3.7% of the water flashes by mass, but the large specific volume of the vapour can make its occupied volume dominate the downstream pipe.

Trace a saturated mixture along a constant-volume path

The drum example determines one state. If a sealed rigid vessel is heated, its pressure and temperature change together along saturation while its specific volume remains fixed. Provided equilibrium is maintained and no other phase appears, the quality at every pressure is

$$x = (v - v_f)/(v_g - v_f).$$

The energy equation for the vessel is $Q = \Delta U$ when it does no boundary or shaft work. Replacing internal-energy change by enthalpy change would add the spurious $V\Delta p$ term of a rigid container.

Quality does not change solely because latent heat has entered. Differentiate $v = (1 - x)v_f + x v_g$ along the saturation curve at fixed v. This gives

$$\frac{dx}{dT} = -\frac{(1 - x)\,dv_f/dT + x\,dv_g/dT}{v_g - v_f}.$$

Saturated liquid usually expands as temperature rises while saturated vapour specific volume decreases over the ordinary water range. The sign therefore depends on the initial mixture and local property slopes. A vessel containing almost all liquid by volume can, under some conditions, follow the saturated-liquid boundary on heating rather than drying out. Determine the actual phase region from the property data after each change. When v lies outside the interval between v_f and v_g, a saturated quality formula is no longer applicable.

Find how much energy changes phase in a sealed vessel

Take 1.00 kg of water in a rigid 0.300 m³ vessel. Use rounded saturation data at 100 kPa:

$$v_f = 0.00104,\ v_g = 1.694\,\mathrm{m^3/kg},\ u_f = 417.4,\ u_g = 2506\,\mathrm{kJ/kg}.$$

At 200 kPa use $v_f = 0.00106,\ v_g = 0.8857,\ u_f = 504.5,\ u_g = 2529\,\mathrm{kJ/kg}$. The volume bounds contain 0.300 m³/kg at both pressures. Initial quality is $(0.300 - 0.00104)/(1.694 - 0.00104) = 0.1766$; final quality is

$$(0.300 - 0.00106)/(0.8857 - 0.00106) = 0.3379.$$

Thus about 0.161 kg additionally evaporates.

Mixture internal energies are $417.4 + 0.1766(2506 - 417.4) = 786.3\,\mathrm{kJ/kg}$ and $504.5 + 0.3379(2529 - 504.5) = 1188.6\,\mathrm{kJ/kg}$, giving heat input about 402 kJ. Merely multiplying evaporated mass by one latent heat would overlook the sensible energy changes in each phase and the variation of saturation properties. As an independent check, $\Delta H = \Delta U + V\Delta p$ exceeds the required heat by $0.300(200 - 100) = 30.0\,\mathrm{kJ}$. If additional heating brings saturated vapour specific volume down to 0.300 m³/kg, the last liquid disappears. Beyond that state the fixed-volume path enters the superheated region and one must determine temperature using a single-phase equation of state or superheated-vapour data.

The vapour volume fraction in the second state is $m x v_g / V = 0.3379(0.8857)/0.300 = 0.998$ approximately. Although about one third of the mass is vapour, almost every part of the vessel volume is vapour-filled. A level probe that detects a very small liquid volume is consistent with the presence of substantial residual liquid mass. This has a direct consequence for rounding: replacing v_f by zero only gives a modest error in calculated quality here, but using that same approximation to predict the remaining liquid depth gives a nonsensical zero. Retain the finite liquid volume whenever the question concerns an interface height or the mass recovered through a bottom drain.

The units of the property table matter just as much as the phase labels when predicting a small liquid inventory.

Can an equation of state predict energy changes during expansion?

A caloric model $du = c_v\,dT$ works for an ideal gas. To discover the extra term for a real fluid, let v, s and u denote molar volume, entropy and energy for a single homogeneous phase. At a fixed composition the fundamental relation is

$$du = T\,ds - p\,dv.$$

Introduce the Helmholtz energy $f = u - Ts$. Then $df = -s\,dT - p\,dv$. Equality of the two mixed partial derivatives of the smooth function $f(T, v)$ gives the Maxwell relation $(\partial s/\partial v)_T = (\partial p/\partial T)_v$.

Now expand ds in T, v. The term $T(\partial s/\partial T)_v$ equals c_v, obtained directly by holding v fixed in the fundamental relation. Substituting both coefficients gives

$$du = c_v\,dT + \left[T\left(\frac{\partial p}{\partial T}\right)_v - p\right] dv.$$

The bracket has units of pressure. For $p = RT/v$ it vanishes. For a van der Waals model $p = RT/(v-b) - a/v^2$, the repulsive volume correction cancels between the two terms and the attraction term leaves a/v^2. Assuming constant c_v, integration yields

$$\Delta u = c_v(T_2 - T_1) + a(1/v_1 - 1/v_2).$$

An insulated, rigid container contains a gas on one side of a removable internal partition and vacuum on the other. Removing the partition causes no external boundary work and exchanges no heat. States before and after settling therefore have equal internal energy. Let $T_1 = 300\,\mathrm{K}$, $v_1 = 0.00200\,\mathrm{m^3\,mol^{-1}}$, $v_2 = 0.00400\,\mathrm{m^3\,mol^{-1}}$, $a = 0.150\,\mathrm{Pa\,m^6\,mol^{-2}}$, $b = 4.0\times10^{-5}\,\mathrm{m^3\,mol^{-1}}$ and $c_v = 24.0\,\mathrm{J\,mol^{-1}K^{-1}}$. The volume contribution is

$$0.150(500 - 250) = 37.5\,\mathrm{J\,mol^{-1}}.$$

To keep energy constant the temperature falls by $37.5/24.0 = 1.56\,\mathrm{K}$, to $298.44\,\mathrm{K}$. The differential identity links the equilibrium endpoints even though the actual expansion was irreversible. The value of b matters for their pressures and cancels from this energy change.

Practise on sheet 4 · Compare free and isothermal expansion of the same model gas. Further reading · University of Cincinnati · C. Vaz, Thermodynamics and Statistical Mechanics · §4.2 Maxwell Relations.

Chapter 8 · Gas mixtures, humidity and condensation

Translate relative humidity into a water mass balance and determine how much condensate a cooling coil removes.

FROM FIRST PRINCIPLES

Obtain humidity ratio from partial pressures and dry-air mass

Treat moist air as an ideal mixture of dry air and water vapour occupying the same volume V at temperature T. Dalton's law gives total pressure $p = p_{da} + p_v$. For each component $m_i = p_i V/(R_i T)$, so dividing vapour mass by dry-air mass cancels the shared volume and temperature:

$$\omega = \frac{m_v}{m_{da}} = \frac{R_{da}}{R_v}\,\frac{p_v}{p - p_v} \simeq 0.622\,\frac{p_v}{p - p_v}.$$

Relative humidity is a different observable. It compares actual vapour partial pressure with the saturation pressure at the same temperature. Cooling without water removal keeps ω constant, and approximately constant total pressure then keeps p_v constant until the saturation curve reaches it:

$$\varphi = p_v / p_{\mathrm{sat}}(T),$$
$$p_{\mathrm{sat}}(T_{\mathrm{dew}}) = p_v.$$

Below the dew point some water can leave the gas as liquid, while dry-air mass is conserved. A steady coil receiving humidity ratio ω_i and delivering ω_o therefore removes condensate at rate $\dot{m}_{da}(\omega_i - \omega_o)$. Inferring the same removal from relative humidity alone would ignore its dependence on temperature.

DERIVATION AND APPLICATIONS

Measure the bypass fraction of a cooling and dehumidifying coil

Humid air crossing a coil may leave warmer and wetter than air in equilibrium with the coldest effective surface. A common reduced model treats a fraction b of the dry-air stream as bypassing the surface while the remainder reaches a saturated state at the apparatus dew-point temperature. Consider a steady coil at nearly constant total pressure, neglect fan work and heat gains, and use dry-air mass as the conserved basis. Denote inlet, saturated-surface and mixed outlet states by i, s, o, with humidity ratio ω and moist-air enthalpy h per kilogram of dry air.

Water-vapour and mixed-stream enthalpy balances give two independent ways to estimate the same bypass fraction:

$$\omega_o = b\omega_i + (1 - b)\omega_s,$$
$$h_o = bh_i + (1 - b)h_s.$$

Solving either for b exposes a consistency check on measured data and the assumed single surface state:

$$b = \frac{\omega_o - \omega_s}{\omega_i - \omega_s} = \frac{h_o - h_s}{h_i - h_s}.$$

The condensed water remains a separate outlet from the coil. With condensate leaving at liquid enthalpy h_ℓ in the same reference convention, an overall balance gives positive heat removal

$$\dot{m}_\ell = \dot{m}_{da}(\omega_i - \omega_o),$$

$$\dot{Q}_{\text{out}} = \dot{m}_{da}(h_i - h_o) - \dot{m}_\ell h_\ell.$$

The mixed-state model requires

$$0 \le b \le 1.$$

It also requires incoming air wet enough for condensation on the assumed surface. A dry surface with solely sensible cooling cannot be represented by forcing a saturated leaving branch.

WORKED EXAMPLE

Evaluate air and water leaving a coil

Change the inputs and test this calculation.

Use $\omega_i = 0.0160$, $\omega_s = 0.00760$ kg/kg dry air, $h_i = 71.0$, $h_s = 29.3$ kJ/kg dry air and $b = 0.250$. Then $\omega_o = 0.00970$ and $h_o = 39.7$ kJ/kg dry air. With 1.20 kg/s dry air, the condensate rate is 0.00756 kg/s. For $h_\ell = 42.0$ kJ/kg, heat removal is

$$1.20(71.0 - 39.725) - 0.00756(42.0) = 37.2\,\text{kW}.$$

SECOND APPLICATION

Test whether three measurements share one apparatus dew point

Change the inputs and test this calculation.

On another run with the same inlet and assumed surface, measured outlet values are $\omega_o = 0.01096$ and $h_o = 41.8$ kJ/kg dry air. The humidity ratio implies $b = 0.400$, while enthalpy gives $b = 0.300$. At $b = 0.400$ the model would predict $h_o = 46.0$ kJ/kg. The 4.2 kJ/kg discrepancy indicates that measurement uncertainty, a more distributed surface temperature or another heat transfer must be considered before inferring a unique bypass fraction.

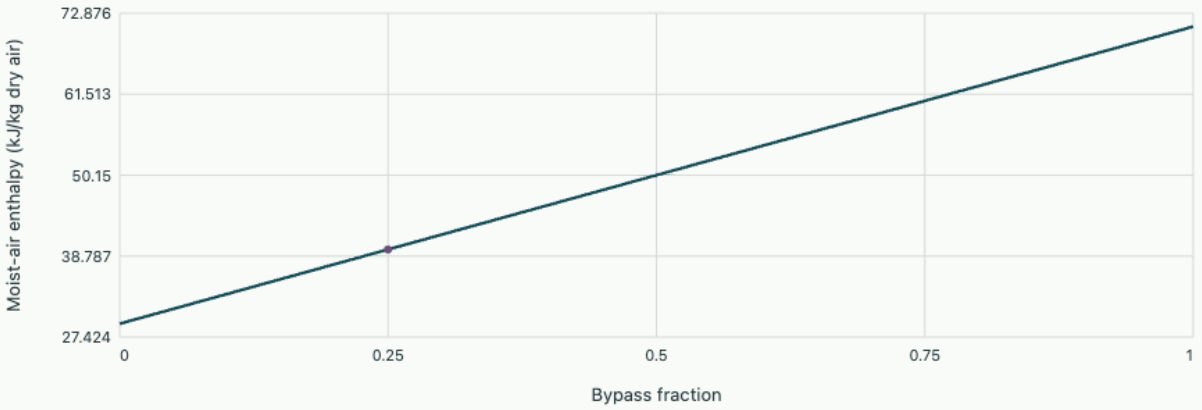
Try it · Coil bypass

Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Bypass fraction inferred from humidity	**0.25**
Measured outlet enthalpy (kJ/kg dry air)	**39.725**



Humidity predicts ωout=0.0097 kg/kg and hout=39.725 kJ/kg; enthalpy measurement implies b=0.25. Discrepancy 0 kJ/kg. At the model state and 1.2 kg/s dry air: condensate 0.00756 kg/s; heat removal 37.2125 kW including liquid enthalpy 42 kJ/kg.

Choose a quantity that survives condensation

A ventilation duct can carry an unchanged mass flow of dry air while water evaporates into it or condenses out. It is therefore convenient to express water content and mixture energy per kilogram of dry air. Model the dry air and water vapour as ideal gases sharing a common temperature and total volume. The partial pressure of each gas is the pressure it would exert if it occupied this volume alone. Dalton’s law gives

$$p = p_a + p_v.$$

Relative humidity $\phi = p_v / p_{\mathrm{sat}}(T)$ describes proximity to saturation at the current temperature. It is not the mass fraction of water. For humidity ratio $w = m_v / m_a$, divide the two ideal-gas mass equations. The common temperature and volume cancel, leaving a ratio of molar masses multiplied by the partial-pressure ratio.

$$w = \frac{M_v}{M_a}\frac{p_v}{p - p_v} \simeq 0.622\frac{p_v}{p - p_v},$$

$$w_{\mathrm{sat}}(T) = 0.622\frac{p_{\mathrm{sat}}(T)}{p - p_{\mathrm{sat}}(T)}.$$

Before any water condenses during cooling at constant total pressure, the humidity ratio and vapour partial pressure remain fixed. Relative humidity rises because saturation pressure falls. The dew point is the temperature at which saturation pressure equals that fixed vapour pressure. Once that point is crossed, further cooling can remove water.

Find water removal and the associated cooling duty

Air enters a coil at 30 °C, 100 kPa and 60% relative humidity, and leaves saturated at 12 °C. Use $p_{\mathrm{sat}}(30) = 4.24$ kPa and $p_{\mathrm{sat}}(12) = 1.40$ kPa. The inlet vapour partial pressure is 2.544 kPa, giving $w_1 = 0.01624$ kg water per kg dry air. The outlet ratio is $w_2 = 0.00883$. For dry-air flow 0.50 kg/s the condensate flow is $0.50(0.01624 - 0.00883) = 0.00371\,\mathrm{kg/s}$, about 13.3 kg/h.

Using a 0 °C liquid-water reference, approximate moist-air enthalpy per kilogram dry air by $h = 1.005T + w(2501 + 1.86T)$ kJ/kg dry air with T in °C. The duct streams carry about 71.7 and 34.3 kJ/kg dry air. Condensate leaving at 12 °C carries $h_l \simeq 4.18(12) = 50.2\,\mathrm{kJ/kg}$. Energy conservation gives heat removed per kilogram dry air as $h_1 - h_2 - (w_1 - w_2)h_l$, approximately 37.0 kJ/kg dry air, or 18.5 kW.

Separate humidity ratio from relative humidity

Suppose the outlet stream is reheated to 20 °C without adding water. Its vapour pressure remains 1.40 kPa at the same total pressure. Since $p_{\mathrm{sat}}(20) = 2.34\,\mathrm{kPa}$, the new relative humidity is 59.8%. Neither evaporating water nor drying the air was required during this reheat. A sensor reporting 60% at both the inlet and reheated outlet therefore describes two air streams containing rather different masses of water. The coil water balance, written on the dry-air basis, explains the apparent coincidence.

Bracket an unknown dew point

At the original inlet the vapour pressure was 2.544 kPa. Suppose a saturation table gives 2.487 kPa at 21 °C and 2.644 kPa at 22 °C. Linear interpolation over this short interval puts the dew point at $21 + (2.544 - 2.487)/(2.644 - 2.487) = 21.36^{\circ}\mathrm{C}$. A coil surface warmer than this value cannot start condensing the entering air locally under the stated conditions.

Balance an adiabatic spray chamber

A direct evaporative cooler behaves differently from a refrigerated cooling coil. Water is added to an airstream, the dry-bulb temperature falls, and the humidity ratio increases. Assume steady operation, negligible shaft work and heat leakage, unchanged total pressure, and no liquid droplets carried out. Express enthalpy per kilogram dry air. The dry-air flow is conserved; added liquid flow per kilogram dry air equals $w_2 - w_1$. The resulting energy balance is

$$h_1 + (w_2 - w_1) h_{\ell,\mathrm{in}} = h_2.$$

Use the same enthalpy reference for liquid and moist air. If the added water has low liquid enthalpy on that reference, the moist-air enthalpy is approximately constant. Often the recirculated liquid is near the leaving-air temperature; the more accurate energy equation above can be used without that approximation. An ideal infinitely long saturated contactor has $\phi_2 = 1$, but a finite cooler generally leaves unsaturated air. An outlet temperature or a contact effectiveness is needed in addition to the balances unless saturation is specified.

With constant heat capacities one may write $h_2 = c_{pa} T_2 + w_2 (h_{g0} + c_{pv} T_2)$. Substituting this form into the balance and grouping the unknown w_2 shows explicitly which measured quantities are needed. The process cannot cool below the adiabatic saturation limit for the given inlet state and water supply merely by increasing the spray rate.

Check outlet humidity and water consumption

A cooler receives air at 35 °C and 100 kPa with inlet humidity ratio 0.0080 kg/kg dry air and dry-air flow 0.60 kg/s. Its measured outlet temperature is 23 °C and liquid makeup enters at 20 °C. Use $h = 1.005T + w(2501 + 1.86T)$ kJ/kg dry air and $h_\ell = 4.18T$ kJ/kg. The inlet carries $1.005(35) + 0.008[2501 + 1.86(35)] = 55.70\,\mathrm{kJ/kg}$ dry air; liquid enters with $h_\ell = 83.6\,\mathrm{kJ/kg}$.

The outlet equation is $55.70 + (w_2 - 0.0080)83.6 = 1.005(23) + w_2[2501 + 1.86(23)]$, giving $w_2 = 0.01297\,\mathrm{kg/kg}$ dry air. Thus liquid consumption is $0.60(0.01297 - 0.0080) = 0.00298\,\mathrm{kg/s}$, or about 10.7 kg/h. The outlet vapour pressure is

$$100(0.01297)/(0.622 + 0.01297) = 2.04\,\mathrm{kPa}.$$

At a supplied saturation pressure of 2.81 kPa for 23 °C, this corresponds to about 73% relative humidity, consistent with a finite contact process.

Treating air as a dry sensible stream would suggest a cooling duty of roughly $0.60(1.005)(35 - 23) = 7.24\,\mathrm{kW}$ removed from the cabinet. In this adiabatic device most of the sensible decrease supplies evaporation inside the boundary; no corresponding cooling-coil duty exists. A downstream process sensitive to water content may therefore react very differently to evaporative cooling and cooling plus dehumidification, even when both report the same outlet dry-bulb temperature.

An independent check can be made by weighing the makeup reservoir over a stable ten-minute test. At the predicted consumption it should lose about $10.7/6 = 1.78$ kg, provided no liquid is stored in the

spray circuit during that interval. A greater measured loss together with the same outlet humidity suggests entrained droplets, drainage or leakage rather than necessarily more evaporation. A humidity probe downstream of an effective droplet separator and a separate liquid-drain measurement distinguish those paths and make the water balance close.

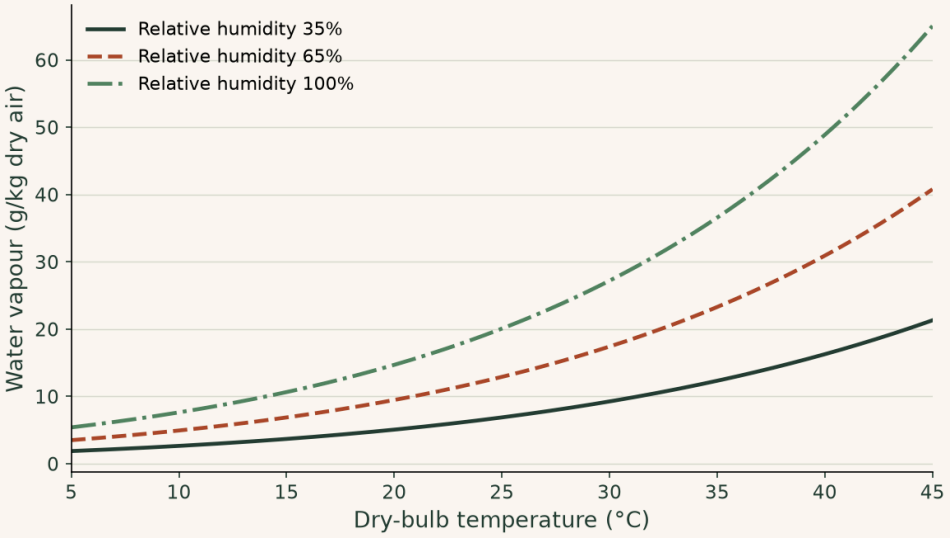
Identify condensation from a humidity-ratio chart

At a total pressure of 101.325 kPa,

$$w = 0.62198\, p_v/(p - p_v),$$

$$p_v = \phi\, p_{\mathrm{sat}}(T).$$

The plot uses the Buck saturation-pressure correlation over liquid water. Follow a sealed parcel cooled at constant pressure with no condensation until it reaches saturation.



Air at 30°C and 65% relative humidity contains about 17.4 g/kg dry air. Following that horizontal humidity ratio to the saturation curve gives a dew point near 22.7°C. Continuing cooling at constant total water content then produces liquid condensate.

▶ **Read values from the plot**

Interactive calculation · change the inputs of this example

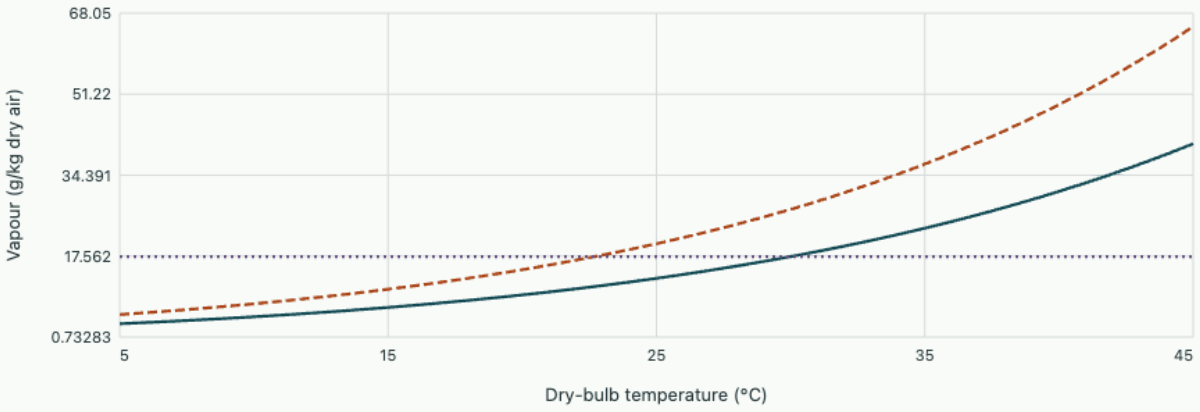
Choose an initial temperature and relative humidity, then follow the horizontal parcel-water line to saturation.

The curves below are recalculated from this section's equations in your browser. The fixed reference plot above remains available for comparison.

Open this live calculation

Try the calculation

Relative humidity	**0.65**
Total pressure (kPa)	**101**
Initial dry bulb (°C)	**30**



Initial humidity 17.470 g/kg dry air; dew point 22.701 °C. Below that point cooling at fixed total water produces liquid.

Problem sheet 4 · Saturation and moist air

Infer phase mass fractions from measured properties and use dry-air balances through cooling and mixing processes.

A vessel containing liquid and vapour

At a specified saturation pressure use $v_f = 0.0011,\ v_g = 0.900\,\mathrm{m^3/kg}$,

$$h_f = 505,\ h_{fg} = 2200\,\mathrm{kJ/kg}.$$

A 0.180 m³ vessel contains 1.0 kg in equilibrium. Determine quality, vapour volume fraction and specific enthalpy. Does its liquid mass or vapour mass dominate?

▼ **Reveal the worked solution**

Quality is

$$(0.180 - 0.0011)/(0.900 - 0.0011) = 0.1990.$$

The vapour occupies $0.1990(0.900) = 0.1791\,\mathrm{m^3}$, approximately 99.5% of the volume. Enthalpy is

$$505 + 0.1990(2200) = 942.8\,\mathrm{kJ/kg}.$$

The liquid still supplies about 80.1% of the mass.

Establish a condensate flow

An air stream at 100 kPa total pressure has vapour partial pressure 2.80 kPa. Cooling reduces this to 1.20 kPa through condensation. For a dry-air flow of 0.80 kg/s determine both humidity ratios, hourly condensate flow and relative humidity after reheating to a state with saturation pressure 3.00 kPa.

▼ **Reveal the worked solution**

The ratios are $0.622(2.8)/(100 - 2.8) = 0.01792$ and $0.622(1.2)/(100 - 1.2) = 0.00755$ kg/kg dry air. Condensate flow is

$$0.80(0.01792 - 0.00755)3600 = 29.9\,\mathrm{kg/h}.$$

Reheating at fixed composition and total pressure preserves vapour partial pressure, giving a final relative humidity of $1.20/3.00 = 40\%$.

Predict an outlet humidity

Two streams at the same pressure carry 0.30 and 0.70 kg/s of dry air with humidity ratios 0.006 and 0.016 kg/kg dry air. Mixed air leaves at 100 kPa and a temperature with saturation pressure 2.60 kPa. Find outlet humidity ratio, vapour pressure and relative humidity. Verify that the assumption of no condensation is consistent.

▼ **Reveal the worked solution**

Dry-air flow is 1.00 kg/s and water flow $0.30(0.006)+0.70(0.016)=0.0130\,\mathrm{kg/s}$. The humidity ratio is 0.0130. Solving $w=0.622p_v/(p-p_v)$ gives

$$p_v=100(0.0130)/(0.622+0.0130)=2.047\,\mathrm{kPa}.$$

Relative humidity is 78.7%, below saturation at this outlet temperature.

Compare free and isothermal expansion of the same model gas

One mole of gas obeys the van der Waals law with $a=0.180\,\mathrm{Pa\,m^6\,mol^{-2}}$,

$$b=5.0\times10^{-5}\,\mathrm{m^3\,mol^{-1}},$$

$$c_v=25.0\,\mathrm{J\,mol^{-1}K^{-1}}.$$

It expands from 0.0030 to $0.0060\,\mathrm{m^3\,mol^{-1}}$, initially at 320 K. Find its final temperature for insulated expansion into vacuum. Then calculate the heat needed for reversible isothermal expansion between the same volumes. Use $R=8.314\,\mathrm{J\,mol^{-1}K^{-1}}$.

▼ **Show the worked solution**

The potential-energy contribution is $0.180(1/0.0030-1/0.0060)=30.0\,\mathrm{J}$ for the mole. Free expansion has $\Delta U=0$, hence

$$T_2=320-30/25=318.8\,\mathrm{K}.$$

Along the isotherm,

$$\Delta U=30.0\,\mathrm{J}.$$

The reversible work done by the gas is $\int p\,dv=RT\ln[(v_2-b)/(v_1-b)]-a(1/v_1-1/v_2)$. The first law gives

$$Q=\Delta U+W=8.314(320)\ln(0.00595/0.00295)\simeq1.87\,\mathrm{kJ}.$$

Here the reservoir supplies heat while the gas performs boundary work. Equal volume endpoints alone do not determine the heat transfer.

Revisit chapter 7 · Can an equation of state predict energy changes during expansion?.

Explore the governing equations after solving the sheet: Phase equilibrium and steam quality; Gas mixtures, humidity and condensation. Compare the assumptions and units before transferring your chosen parameters; the controls may cover a different numerical range from the written exercise.

Chapter 9 · Useful work and exergy destruction

Evaluate the useful-work potential of a steady stream and distinguish energy conservation from the destruction of useful-work potential.

FROM FIRST PRINCIPLES

Derive maximum useful work relative to an environment

For a closed body changing toward equilibrium with surroundings at T_0, p_0, consider the reversible path that gives the largest possible useful work. Neglect bulk kinetic and gravitational energy. With heat Q_0 positive into the body and total boundary plus shaft work W positive out, the first law gives

$$W = U_i - U_f + Q_0.$$

Reversible heat exchange with this environment requires

$$Q_0 = T_0(S_f - S_i).$$

Part of total work simply displaces ambient pressure, so

$$W_{\text{useful,rev}} = U_i - U_f + T_0(S_f - S_i) - p_0(V_f - V_i).$$

Choose the final state as the dead state, denoted by subscript zero. Grouping the available initial differences defines nonflow physical exergy:

$$B = (U - U_0) + p_0(V - V_0) - T_0(S - S_0).$$

If the process generates entropy, the same first and second law combination lowers the useful work by $T_0 S_{\text{gen}}$. The expression depends on the specified surrounding state and assumes no other heat reservoirs or chemical changes. A stream crossing a control volume carries its pressure flow work as well, giving a distinct enthalpy-based expression for flow exergy.

Background · Exergy destruction

$$\dot{X}_{\text{dest}} = T_0 \dot{S}_{\text{gen}}$$

The environmental reference temperature enters in kelvin. Irreversibility destroys useful-work potential even when the total energy balance closes.

DERIVATION AND APPLICATIONS

Derive the work potential of a compressed flowing gas

A stream can deliver useful work by cooling or by expanding, even when no fuel or chemical reaction is involved. Choose a steady reference environment at T_0, p_0 with the same gas composition. Consider one kilogram taken reversibly from its inlet state to the environmental state. Neglect kinetic and potential energy. Let q_0 be heat supplied from the environment, positive into the device, and w shaft work delivered. The steady energy balance includes flow work in enthalpy:

$$w = h - h_0 + q_0.$$

The only external heat exchange occurs with the environment. Reversibility then gives $s_0 - s = q_0/T_0$, eliminating the unmeasured heat transfer. The maximum work is the specific physical flow exergy

$$b = (h - h_0) - T_0(s - s_0).$$

For an ideal gas with constant c_p and R, insert $dh = c_p\,dT$ and $ds = c_p\,dT/T - R\,dp/p$. Integrate relative to T_0, p_0:

$$b = c_p[(T - T_0) - T_0 \ln(T/T_0)] + RT_0 \ln(p/p_0).$$

The pressure term survives at ambient temperature. Near $T = T_0$ the thermal term is quadratic in the temperature difference because energy and the ambient heat contribution cancel to first order. The formula describes work relative to a specified reference and assumes the inlet and outlet composition is the same. An actual dissipative path destroys part of the reversible work potential.

WORKED EXAMPLE

Estimate the work available from warm compressed air

Change the inputs and test this calculation.

Take air at $T = 360$ K, $p = 300$ kPa entering an environment at $T_0 = 300$ K and $p_0 = 100$ kPa. With $c_p = 1.005$ and $R = 0.287$ kJ/(kg K), the thermal contribution is $1.005[60 - 300 \ln 1.2] = 5.33$ kJ/kg. The pressure contribution is $0.287(300) \ln 3 = 94.6$ kJ/kg. Total available work is approximately 99.9 kJ/kg, mostly associated with pressure.

SECOND APPLICATION

Account for loss in an ambient-temperature throttle

Change the inputs and test this calculation.

Let 0.100 kg/s of ideal air throttle adiabatically from 300 to 100 kPa, both at 300 K. Enthalpy stays constant, no shaft work is extracted, and the exergy destruction is $0.100(94.6) = 9.46$ kW. Independent

entropy accounting gives $\dot{S}_{\mathrm{gen}} = \dot{m} R \ln 3 = 0.0315$ kW/K. Its product with the ambient 300 K recovers the same 9.46 kW. A device balancing energy perfectly can therefore expend substantial work potential.

Try it · Flow exergy

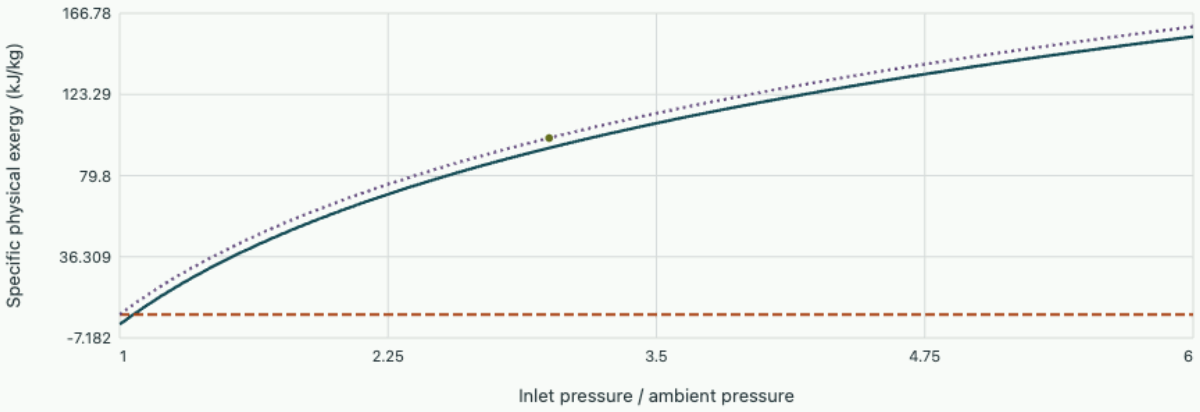
Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Inlet temperature (K) **360**

Selected pressure / ambient pressure **3**



Pressure contribution · Thermal contribution · Total physical flow exergy · Selected inlet

Environment 300 K, 100 kPa. Thermal 5.33005, pressure 94.5905, total 99.9206 kJ/kg. At ambient temperature an ideal-gas throttle from this pressure to ambient destroys 9.45905 kW at 0.1 kg/s without changing enthalpy.

How much work is lost when two finite bodies touch?

Two insulated metal blocks start at different temperatures. A small cyclic heat engine could transfer heat from one to the other and deliver shaft work until the temperatures equalise. Simply bringing the blocks into thermal contact also gives a common final temperature, but it conserves a different quantity. Calculate the two end states before estimating how much work the initial temperature difference made available.

Assume constant positive heat capacities C_h, C_c, no phase change, no external heat exchange and no mechanical work by either block. In direct passive contact no outside work is produced. The total energy change is therefore zero, giving $T_{\mathrm{mix}} = (C_h T_h + C_c T_c)/(C_h + C_c)$. If the intervening engine instead operates reversibly, each small entropy transfer leaving the hot block equals the entropy entering the cold one. At the final common temperature T_f,

$$C_h \ln(T_f/T_h) + C_c \ln(T_f/T_c) = 0,$$

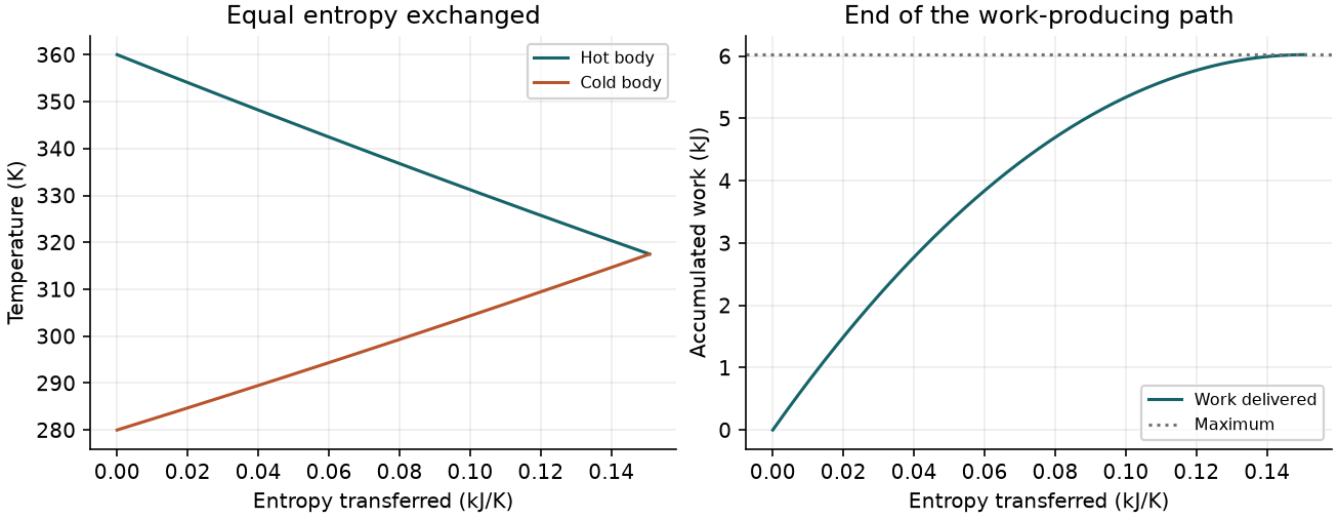
$$T_f = \exp\left[\frac{C_h \ln T_h + C_c \ln T_c}{C_h + C_c}\right].$$

Here all temperatures are absolute. More formally the logarithms can be taken of ratios to the same reference temperature; that reference cancels. The weighted geometric mean lies below the weighted arithmetic mean. The reversible final state thus contains less internal energy than the passive final state. The difference has emerged as work,

$$W_{\max} = (C_h + C_c)(T_{\mathrm{mix}} - T_f).$$

An irreversible engine generates entropy and ends at a higher common temperature with smaller available work.

Take two blocks with capacities of 1.2 kJ/K, initially at 360 K and 280 K. Passive contact ends at 320.0 K; the reversible engine ends at $\sqrt{360 \times 280} = 317.49$ K and delivers 6.02 kJ. Passive contact generates 0.01890 kJ/K of entropy. To visualise the reversible path, let s be the entropy transferred from one block into the other. The instantaneous temperatures are $T_h(s) = T_h(0)e^{-s/C_h}$ and $T_c(s) = T_c(0)e^{s/C_c}$. The next increment of work is $(T_h - T_c)ds$, so the signed area between the two temperature curves reproduces the total work.



The slope of the work curve is the temperature gap. It vanishes at the reversible end state even though the engine delivered finite work earlier.

The scaling becomes particularly clear for equal capacities C when the starting temperatures are $T_m + \Delta$ and $T_m - \Delta$. The mixed temperature is T_m, while the reversible final temperature is $T_m\sqrt{1 - (\Delta/T_m)^2}$. Expanding the square root for small temperature gaps gives

$$W_{\max} \simeq C\Delta^2/T_m.$$

For $T_m = 320$ K and $\Delta = 40$ K this is 6.0 kJ, close to the exact 6.02 kJ. Halving the temperature gap at the same mean temperature gives roughly a quarter of the work. The energy initially contained in the warmer block above the mean temperature scales only linearly with the gap, so that energy alone is a poor measure of usable work.

If the surroundings never exchange heat with either block, should their temperature appear in this maximum-work calculation?

▼ **Work through the solution**

The reversible engine can use only the two specified bodies. Their initial states and conservation of their total entropy determine how far the work extraction goes. A different problem allows ultimate equilibration with an external reservoir at T_0. In that enlarged problem the standard exergy destruction associated with the irreversible contact is $T_0 S_{\mathrm{gen}}$. These work comparisons have different permitted end states. The notebook keeps the enclosure insulated and checks the work using the energy balance on both bodies throughout the exchange.

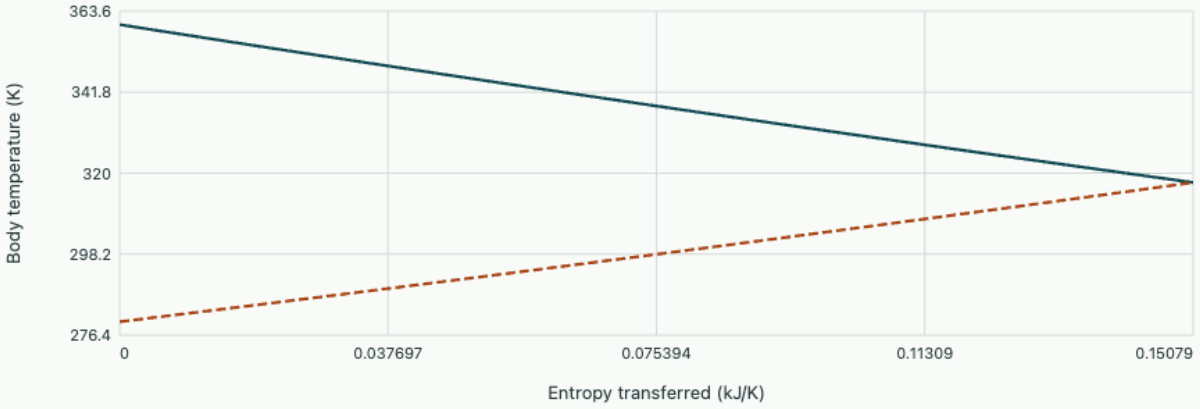
Interactive calculation · Finite source temperatures

Vary the notebook inputs and compare both plotted observables. The same equations generate the browser curves and the notebook reference; fixed material, geometry and model assumptions are stated in the readout.

Open this live calculation

Try the calculation

Hot initial temperature (K)	**360**
Cold initial temperature (K)	**280**
Hot / cold heat capacity	**1**



Hot finite source Cold finite source

Common reversible temperature 317.490 K; maximum work 6.024 kJ. Cold heat capacity1.2 kJ/K. Direct contact reaches a warmer arithmetic-weighted mean and generates entropy.

Download the editable notebook for this example

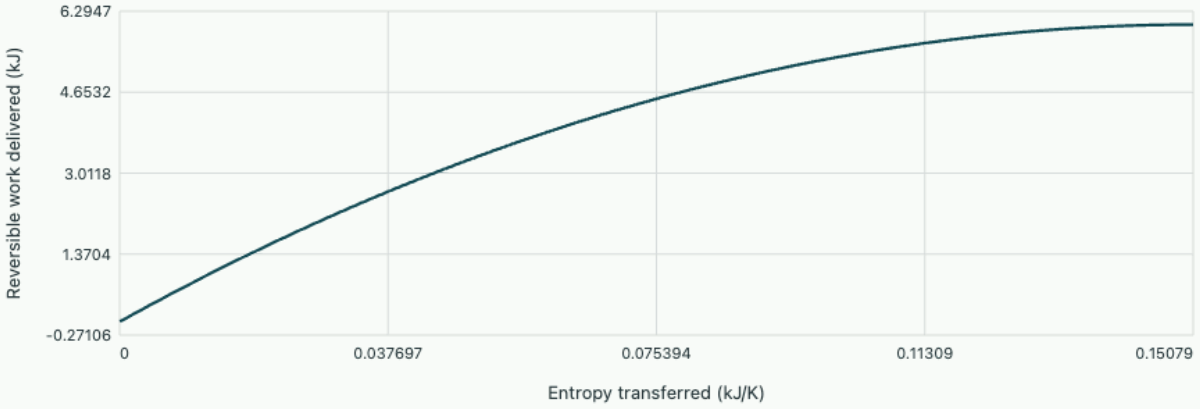
Interactive calculation · Finite source work

Vary the notebook inputs and compare both plotted observables. The same equations generate the browser curves and the notebook reference; fixed material, geometry and model assumptions are stated in the readout.

Open this live calculation

Try the calculation

Hot initial temperature (K)	360
Cold initial temperature (K)	280
Hot / cold heat capacity	1



Accumulated reversible work

Common reversible temperature 317.490 K; maximum work 6.024 kJ. Cold heat capacity1.2 kJ/K. Direct contact reaches a warmer arithmetic-weighted mean and generates entropy.

Download the editable notebook for this example

Open editable notebook Download .ipynb

Specify the surroundings before rating a resource

A warm stream contains energy relative to a chosen reference. The amount that could become useful work depends on the surroundings available to receive heat. If stream and environment reach mutual thermal and mechanical equilibrium, there is no further physical work potential associated with their difference. This equilibrium is the physical dead state at T_0, p_0. Exclude chemical reactions here and consider a steady flow with constant composition.

Imagine a reversible device that takes a stream from its current state to the dead state while interacting thermally only with the environment. The steady first law relates delivered work to changes in stream

enthalpy, kinetic and potential energy, and heat. The reversible entropy balance gives heat received from the environment $q_0 = T_0(s_0 - s)$. Eliminating that heat gives the specific flow exergy

$$\psi = (h - h_0) - T_0(s - s_0) + V^2/2 + g(z - z_0).$$

Use consistent units for the kinetic and potential terms. The entropy contribution reduces the fraction of enthalpy that a reversible device can turn into useful work. At the dead state, with the same velocity and elevation as the reference, that work potential vanishes.

Cool a liquid to the environment

Take a stream of liquid water at approximately ambient pressure, 360 K, with negligible motion and elevation change. It flows at 0.20 kg/s towards a large environment at 300 K. Assume $c_p = 4.18\,\mathrm{kJ/(kg\,K)}$ constant. Integration gives

$$h - h_0 = c_p(T - T_0),$$
$$s - s_0 = c_p \ln(T/T_0).$$

The stream's specific physical exergy is therefore

$$4.18[60 - 300\ln(360/300)] = 22.2\,\mathrm{kJ/kg}.$$

Its enthalpy difference is much larger, 250.8 kJ/kg.

Cooling all the way to 300 K makes at most $0.20(22.2) = 4.44\,\mathrm{kW}$ of useful work available, even though the stream releases 50.16 kW of heat. Direct heat exchange with the environment delivers no shaft work. The environment entropy gain is $50.16/300 = 0.1672\,\mathrm{kW/K}$, while the stream loses $0.20(4.18)\ln(360/300) = 0.1524\,\mathrm{kW/K}$. The difference, about 0.0148 kW/K, is the entropy production rate. Multiplication by T_0 gives the lost 4.44 kW of work potential.

Include heat crossing another temperature boundary

Heat Q supplied reversibly by a reservoir at a uniform boundary temperature T_b carries exergy $Q(1 - T_0/T_b)$. This follows by comparing its energy with the environmental heat required to discharge the associated entropy Q/T_b. A heat input of 10 kW from a 600 K reservoir in a 300 K environment carries 5 kW of exergy. A transfer from a finite cooling stream has changing source temperature and requires integration, as in the water example. Irreversibility in the device destroys exergy at rate $T_0\dot{S}_{\mathrm{gen}}$; directing heat to a useful secondary process can retain part of the exergy that direct cooling would destroy.

Include the work potential of pressure

For an ideal gas already at T_0, an elevated pressure still gives physical flow exergy $RT_0 \ln(p/p_0)$. Air at 300 K and twice environmental pressure has about $0.287(300)\ln 2 = 59.7\,\mathrm{kJ/kg}$ of pressure-related work potential although its ideal-gas enthalpy equals that at the dead state.

Derive the least work needed to heat a finite stream

Heat pumps are commonly compared by heat delivered per unit of electrical work. A second-law comparison additionally accounts for the temperature at which heat is delivered. Consider a steady stream heated from T_i to T_o at nearly environmental pressure, with constant heat capacity rate $C = \dot{m}c_p$. A reversible heat pump extracts heat from surroundings at T_0 and delivers it through a temperature-matched sequence of infinitesimal condensers. For one increment $dQ = C\,dT$ at stream temperature T, the reversed Carnot relation gives

$$dW_{\min} = dQ(1 - T_0/T).$$

Integrating over the stream yields

$$W_{\min} = C[(T_o - T_i) - T_0 \ln(T_o/T_i)].$$

This is the increase of its physical flow exergy, and is smaller than evaluating the Carnot work factor at the final high temperature for the whole heat duty. A device that condenses at a single temperature above the warmest outlet generates entropy during delivery to the colder portions. The result assumes that the source of heat is the environment at T_0; a colder source would introduce an additional resource requirement.

For steady equipment receiving only electricity and environmental heat, with negligible parasitic flows, a meaningful second-law efficiency is the stream’s exergy increase divided by actual electrical input. Electric power is entirely work potential, whereas a kilowatt of heat near room temperature carries only a small fraction of a kilowatt of exergy.

Compare a heat pump with a resistance heater

Water flows at 0.10 kg/s from 293 K to 333 K in a 293 K environment. With $c_p = 4.18\,\mathrm{kJ/(kg\,K)}$, heat duty is $0.418(40) = 16.72\,\mathrm{kW}$. The exergy increase is $0.418[40 - 293\ln(333/293)] = 1.05\,\mathrm{kW}$ approximately. This relatively small figure reflects the modest average temperature lift. A well-insulated heat pump drawing 4.20 kW from the electrical supply has heating COP $16.72/4.20 = 3.98$ and exergy efficiency about 25%. Its surroundings provide 12.52 kW of heat under the stated energy balance.

A resistance heater supplying the identical water service draws 16.72 kW when its losses to the room are negligible. It is nearly 100% efficient on the simple electrical-energy-to-heat basis, yet its exergy efficiency is only about 6.3%. For the heat pump the implied destruction of work potential is $4.20 - 1.05 = 3.15\,\mathrm{kW}$; for resistance heating it is about 15.67 kW. Electricity used by a circulating pump, external heat leaks or a source below environmental temperature belongs in the appropriate

device boundary before drawing a practical comparison. The heat-duty equality ensures that the machines are judged against the same delivered temperature service.

For the resistance case the entropy balance can be checked without using the exergy formula. Electrical work carries no entropy across the control surface, and assume no heat leakage. The stream increases its entropy at $0.418 \ln(333/293) = 0.0535\,\mathrm{kW/K}$. Multiplying by the 293 K environment gives 15.7 kW destroyed, consistent with the difference between input work and stream exergy gain. This agreement depends on keeping the heater and heated stream in the same chosen boundary.

Chapter 10 · Steam power cycles and turbine efficiency

Close a Rankine cycle energy balance and predict how turbine irreversibility changes power, condenser duty and exhaust quality.

FROM FIRST PRINCIPLES

Obtain Rankine-cycle efficiency by following one kilogram

Consider a simple steady steam loop with a pump, boiler, turbine and condenser. Number states 1 at pump inlet, 2 at pump outlet, 3 at turbine inlet and 4 at turbine exhaust. Neglect kinetic and potential changes and external heat leaks. The steady-flow energy balance for one kilogram through the adiabatic pump and turbine converts their enthalpy changes directly into shaft work; in the boiler and condenser there is no shaft work:

$$w_p = h_2 - h_1,$$
$$w_t = h_3 - h_4,$$
$$q_{\text{in}} = h_3 - h_2,$$
$$q_{\text{out}} = h_4 - h_1.$$

Subtract the pump input from the turbine output and divide by heat supplied. Rearrangement gives a second expression that checks cycle energy closure:

$$\eta_{\text{th}} = \frac{(h_3 - h_4) - (h_2 - h_1)}{h_3 - h_2} = 1 - \frac{h_4 - h_1}{h_3 - h_2}.$$

These identities use actual enthalpies. An idealized isentropic turbine supplies h_{4s} from its outlet pressure and inlet entropy; isentropic efficiency then estimates actual h_4. Liquid pump work often admits the incompressible estimate $v(p_2 - p_1)$, explaining why its enthalpy rise can be much smaller than the turbine enthalpy drop.

DERIVATION AND APPLICATIONS

Close a steam-cycle balance using the condenser cooling-water measurement

The condenser often transfers more heat than the plant delivers in shaft work. Its cooling-water measurement supplies an independent check on assumed steam flow and moisture fraction. Choose a steady control volume enclosing both sides of a surface condenser, with no mixing, negligible heat leaks or pressure losses, and negligible kinetic and potential energy changes. Steam leaves as saturated liquid at the condenser pressure. Cooling water remains a single-phase liquid with constant c_w. Denote steam and cooling-water mass rates by $\dot{m}_s, \dot{m}_w$.

Separate mass balances keep both flow rates constant across their respective sides. The steady energy balance couples the enthalpy lost by steam to the energy gained by water:

$$\dot{m}_s(h_{\mathrm{ex}} - h_f) = \dot{m}_w c_w (T_{w,o} - T_{w,i}).$$

If the turbine exhaust is a saturated mixture of quality x, then $h_{\mathrm{ex}} = h_f + x h_{fg}$. Solving for quality gives a quantity inferred independently of a turbine efficiency assumption:

$$x = \frac{\dot{m}_w c_w (T_{w,o} - T_{w,i})}{\dot{m}_s h_{fg}}.$$

To connect this measurement to the complete cycle, write the boiler heat per kilogram as $q_{\mathrm{in}} = h_3 - h_2$ and the condenser rejection as $q_{\mathrm{out}} = x h_{fg}$. Over a full ordinary Rankine cycle the working fluid stores no net energy:

$$w_{\mathrm{net}} = q_{\mathrm{in}} - q_{\mathrm{out}},$$

$$\eta_{\mathrm{th}} = 1 - \frac{x h_{fg}}{h_3 - h_2}.$$

These relations assume no extraction, reheating or additional heat exchange elsewhere. If the exhaust is superheated or the condensate is significantly subcooled, the first energy balance remains valid but the latent-heat-only expression must be replaced by actual endpoint enthalpies.

WORKED EXAMPLE

Infer exhaust quality and plant output

Change the inputs and test this calculation.

Suppose $\dot{m}_s = 10.0$ kg/s, $\dot{m}_w = 500$ kg/s, $c_w = 4.18$ kJ/(kg K), water temperature rise 10.0 K and $h_{fg} = 2393$ kJ/kg. The water carries off 20.9 MW, giving $x = 20\,900/(10.0 \times 2393) = 0.873$. If boiler heat is 3000 kJ/kg, the steam receives 30.0 MW and the cycle delivers 9.10 MW net, corresponding to efficiency 30.3%.

SECOND APPLICATION

Detect a cooling-water instrument inconsistency

Change the inputs and test this calculation.

With unchanged steam flow and latent heat, a reported rise of 12.0 K at the same water rate would imply 25.08 MW and $x = 1.048$. For a saturated exhaust the greatest possible rise in this setup is

$$10.0(2393)/(500 \times 4.18) = 11.45\,\mathrm{K}.$$

The larger value suggests checking water flow, sensor bias or an assumption such as superheated inlet steam before accepting the inferred shaft output.

Try it · Condenser calorimetry

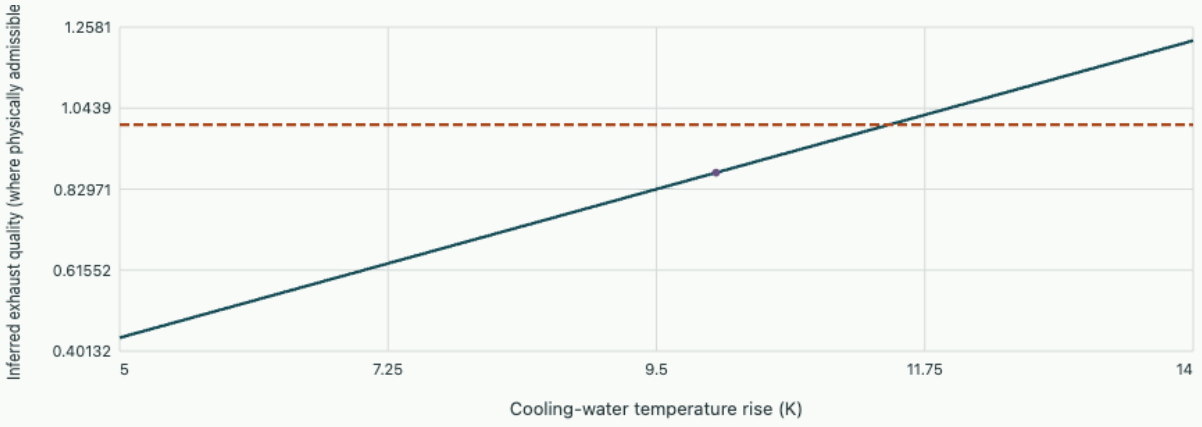
Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Cooling-water mass rate (kg/s) **500**

Measured water temperature rise (K) **10**



10 kg/s steam; latent heat 2393 kJ/kg. Condenser duty 20.9 MW. Quality 0.873381; ordinary-cycle net power 9.1 MW for 30 MW boiler heat. Greatest rise compatible with saturated exhaust 11.4498 K.

Follow a kilogram through the loop

A simple steam plant pumps condensed water to boiler pressure, heats it to steam, expands it through a turbine and condenses it at low pressure. Water is nearly incompressible in the pump, while vapour occupies a much greater specific volume during expansion. A reversible steady adiabatic device satisfies $dh = v\,dp$ because $T ds = dh - v\,dp$ and $ds = 0$. Pressure increase therefore costs little work per kilogram when the fluid is liquid. This is one reason for condensing the working fluid before pumping it.

Number the saturated liquid leaving the condenser as state 1, the pump outlet as 2, the turbine inlet as 3 and the exhaust as 4. Neglect changes in kinetic and potential energy. Using positive magnitudes for pump work input and turbine work output gives

$$w_p \simeq v_1 (p_2 - p_1),$$
$$w_t = h_3 - h_4,$$
$$q_{\text{in}} = h_3 - h_2,$$
$$\eta_{\text{th}} = (w_t - w_p)/q_{\text{in}}.$$

The exhaust pressure is set by the low-pressure side. An ideal turbine exit at that pressure has $s_{4s} = s_3$; a real adiabatic turbine delivers a smaller enthalpy drop. Its isentropic efficiency is $\eta_t = (h_3 - h_4)/(h_3 - h_{4s})$.

Calculate an ideal and an actual expansion

Use illustrative rounded state data. Condenser pressure is 10 kPa, with $h_1 = 191.8\,\text{kJ/kg}$, $v_1 = 0.00101\,\text{m}^3/\text{kg}$, $s_f = 0.649$, $s_{fg} = 7.50\,\text{kJ/(kg K)}$ and $h_{fg} = 2393\,\text{kJ/kg}$. Boiler pressure is 4000 kPa and superheated inlet properties are

$$h_3 = 3214\,\text{kJ/kg}, s_3 = 6.77\,\text{kJ/(kg K)}.$$

The ideal pump requires $0.00101(4000 - 10) = 4.03\,\text{kJ/kg}$, so $h_2 = 195.8\,\text{kJ/kg}$.

At the low pressure, the ideal exhaust quality is $x_{4s} = (6.77 - 0.649)/7.50 = 0.816$, giving $h_{4s} = 191.8 + 0.816(2393) \simeq 2145\,\text{kJ/kg}$. With turbine efficiency 0.85 the actual exit enthalpy is

$$3214 - 0.85(3214 - 2145) = 2305\,\text{kJ/kg}.$$

The corresponding quality is

$$(2305 - 191.8)/2393 = 0.883.$$

At fixed inlet and exhaust pressure this less efficient turbine leaves more enthalpy in the exhaust, and therefore a higher quality inside the two-phase region.

Interpret the condenser and generator duties

The actual turbine delivers about 909 kJ/kg. Subtracting pump work leaves approximately 905 kJ/kg net shaft work. Boiler heat input is $3214 - 195.8 = 3018\,\text{kJ/kg}$, giving thermal efficiency about 30.0%. The condenser rejects $2305 - 191.8 = 2113\,\text{kJ/kg}$, which closes the cycle energy balance within the rounding of the property data. At 2.0 kg/s, the idealised net mechanical output is about 1.81 MW before generator and auxiliary losses.

A lower condenser pressure generally increases the available expansion work for a fixed turbine inlet. It can also lower exhaust quality in the ideal expansion, increasing liquid droplet loading on the later blades. Evaluating such a change requires new saturation properties and a fresh state calculation. Treating the exhaust as an ideal gas would omit the phase fraction precisely where it affects both the energy balance and turbine operating requirements.

Estimate the consequence of a pump loss

If the pump's isentropic efficiency falls to 0.70, actual input rises from 4.03 to $4.03/0.70 = 5.76\,\mathrm{kJ/kg}$. The turbine work remains the same under the specified inlet and outlet assumptions. Boiler duty falls by 1.73 kJ/kg because its feed arrives with more enthalpy. Both effects should enter the revised efficiency, even though pump work is small compared with the turbine output.

Derive the extraction fraction of an open feedwater heater

In a regenerative steam cycle, part of the flow is withdrawn from the turbine to preheat the boiler feedwater. This raises the average temperature of external heat input but reduces the mass continuing through the low-pressure turbine. Draw the control volumes per kilogram entering the boiler. Let fraction y be extracted at intermediate pressure and enthalpy h_b, and let the remaining $1 - y$ pass to the condenser. A first pump returns the condensed fraction to the extraction pressure at enthalpy h_c. In an insulated open heater the streams mix at the same pressure to a single liquid outlet of enthalpy h_f.

Mass and energy balances reduce to $y h_b + (1 - y) h_c = h_f$, hence $y = (h_f - h_c)/(h_b - h_c)$. This fraction must lie between zero and one for passive mixing at the stated states. A second pump then raises the full boiler flow to boiler pressure. If turbine inlet and final exhaust enthalpies are h_3, h_4, work per kilogram admitted to the boiler is

$$w_t = (h_3 - h_b) + (1 - y)(h_b - h_4),$$

$$w_p = (1 - y) w_{p1} + w_{p2}.$$

The first stage always carries the whole flow, as does the second pump. Only the low-pressure turbine, condenser and first pump carry the reduced mass. Applying the factor $1 - y$ to the entire turbine drop is a common bookkeeping mistake. Actual component efficiencies can be included through the enthalpies used in the same balances.

Quantify the exchange between specific work and thermal efficiency

Use an illustrative cycle with $h_3 = 3300,\ h_b = 2780,\ h_4 = 2300\,\mathrm{kJ/kg}$. Condensate leaves at 192.0 kJ/kg and requires $w_{p1} = 0.5\,\mathrm{kJ/kg}$ to reach the heater pressure, so $h_c = 192.5$. Mixed feedwater leaves at $h_f = 640\,\mathrm{kJ/kg}$ and the second pump requires 5.0 kJ/kg. The extraction fraction is

$$(640 - 192.5)/(2780 - 192.5) = 0.1730.$$

About 83% of the boiler flow therefore reaches the low-pressure exhaust.

Turbine work is

$$520 + 0.8270(480) = 917.0\,\mathrm{kJ/kg}.$$

Pump input is $0.8270(0.5) + 5 = 5.41\,\mathrm{kJ/kg}$, giving net 911.5 kJ/kg. Boiler input is $3300 - (640 + 5) = 2655\,\mathrm{kJ/kg}$, or 34.3% thermal efficiency. Condenser duty, calculated independently from the remaining stream, is $0.8270(2300 - 192) = 1743.3\,\mathrm{kJ/kg}$, matching the difference between heat input and net work after rounding.

As a simplified comparison, let all steam expand to the same exhaust state and let a single pump require 5.5 kJ/kg with no regeneration. Net work rises to 994.5 kJ/kg but heat input rises to $3300 - 197.5 = 3102.5\,\mathrm{kJ/kg}$; efficiency falls to 32.1%. A boiler with a fixed heat supply favours the more efficient cycle, while a turbine limited by inlet mass flow may deliver more shaft power without extraction. The appropriate selection depends on which physical resource sets the constraint.

For a boiler limited to 10 MW of heat supply, the regenerative cycle can process $10000/2655 = 3.77\,\mathrm{kg/s}$ of boiler flow and supply about 3.43 MW net. Without extraction the same heat limit permits only $10000/3102.5 = 3.22\,\mathrm{kg/s}$, yielding approximately 3.21 MW net. These power outputs reverse the ranking of work per kilogram because the limiting quantity has changed from mass flow to heat input.

Problem sheet 5 · Exergy and steam-cycle performance

Compare useful-work bounds with actual work and estimate separate turbine and pump contributions to a power cycle.

A finite source cools partway

A liquid with heat capacity 2.0 kJ/(kg K) flows at 0.50 kg/s from 400 to 340 K at ambient pressure. The environment is at 300 K. Find released heat rate, maximum useful work rate from this cooling interval and entropy generated if heat goes directly into the environment.

▼ **Reveal the worked solution**

Released heat is

$$0.50(2)(400-340)=60\,\mathrm{kW}.$$

The drop in flow exergy is

$$0.50(2)[60-300\ln(400/340)]=11.24\,\mathrm{kW}.$$

With direct exchange, the reservoirs and stream generate $60/300-0.50(2)\ln(400/340)=0.03748\,\mathrm{kW/K}$. Multiplication by 300 K recovers the unused 11.24 kW of work potential.

Compare two exhaust states

At a fixed exhaust pressure $h_f=190,\ h_{fg}=2400$ kJ/kg. A turbine has $h_3=3200$ kJ/kg, ideal outlet enthalpy 2200 kJ/kg and isentropic efficiency 0.80. Find shaft work per kilogram, actual exhaust quality and output power at 1.5 kg/s. Which outlet contains the larger liquid fraction?

▼ **Reveal the worked solution**

The actual work is $0.8(3200-2200)=800\,\mathrm{kJ/kg}$, so $h_4=2400\,\mathrm{kJ/kg}$. Actual quality is $(2400-190)/2400=0.921$, and shaft output 1.20 MW. Ideal quality is $(2200-190)/2400=0.838$; the ideal exit has the larger liquid fraction because it has lost more enthalpy in the expansion.

Include pump efficiency

A cycle takes saturated liquid with $v=0.0010\,\mathrm{m^3/kg}$ and $h_1=200\,\mathrm{kJ/kg}$ through a 3.5 MPa pressure rise. Pump efficiency is 0.70; turbine inlet and outlet enthalpies are 3000 and 2250 kJ/kg. Find pump input, net work, heat input and thermal efficiency.

▼ **Reveal the worked solution**

The ideal pump needs $0.0010(3500) = 3.50\,\mathrm{kJ/kg}$; actual input is

$$3.50/0.70 = 5.00.$$

Thus $h_2 = 205\,\mathrm{kJ/kg}$, net work $750 - 5 = 745\,\mathrm{kJ/kg}$ and boiler heat $3000 - 205 = 2795\,\mathrm{kJ/kg}$. Efficiency is 26.65%. Rejection $2250 - 200 = 2050\,\mathrm{kJ/kg}$ closes the cycle energy balance.

Explore the governing equations after solving the sheet: Useful work and exergy destruction; Steam power cycles and turbine efficiency. Compare the assumptions and units before transferring your chosen parameters; the controls may cover a different numerical range from the written exercise.

Notebook practice · How much work is lost when two finite bodies touch?

Chapter 11 · Heat-exchanger effectiveness and finite capacity rates

Derive the counterflow effectiveness relation and determine outlet temperatures when neither stream temperature is fixed.

FROM FIRST PRINCIPLES

Derive exchanger effectiveness from the limiting heat balance

Consider a steady two-stream heat exchanger with no external heat gain, constant specific heats and no phase change. Each stream has a capacity rate $C = \dot{m}c_p$, the power required for one kelvin of temperature change. Applying the first law independently to each stream gives a common transferred heat rate

$$\dot{Q} = C_h(T_{h,i} - T_{h,o}) = C_c(T_{c,o} - T_{c,i}).$$

In passive exchange neither exit can cross the opposite stream's inlet temperature. Therefore the maximum available temperature change on either stream is bounded by the same inlet difference $T_{h,i} - T_{c,i}$. The smaller capacity rate limits the largest conceivable common duty. Define effectiveness from this physically accessible upper scale:

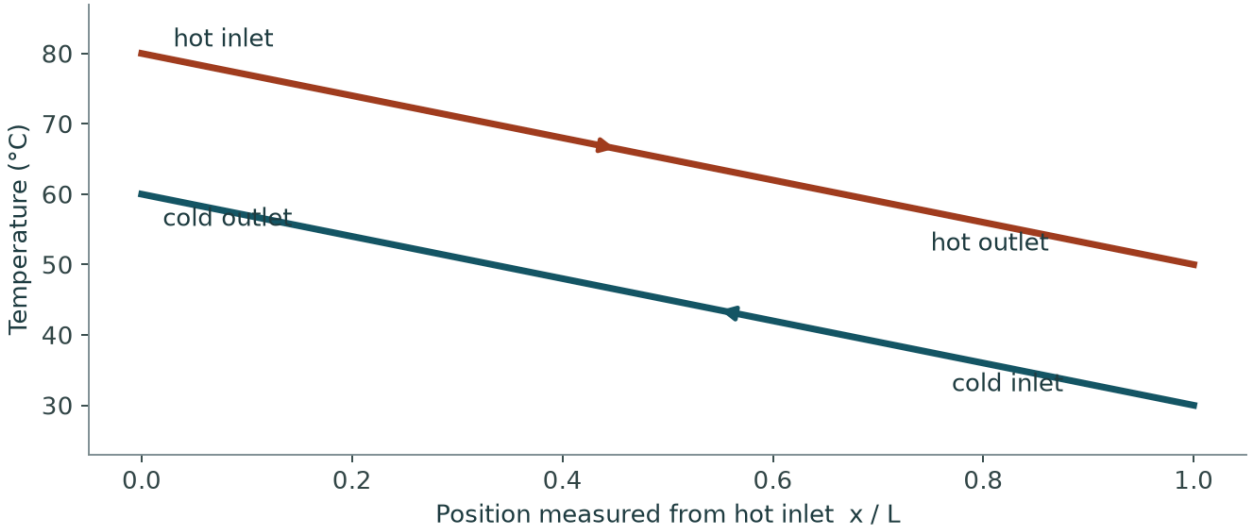
$$\dot{Q}_{\max} = C_{\min}(T_{h,i} - T_{c,i}),$$
$$\varepsilon = \dot{Q}/\dot{Q}_{\max},$$
$$\mathrm{NTU} = UA/C_{\min}.$$

The bound need not be reached by a finite exchanger or by every flow arrangement. Local conductance $U\,dA$ and the changing local temperature difference determine the actual effectiveness. Knowing the two capacity rates and ε suffices to reconstruct both outlet temperatures using the separate energy balances.

Background · Heat-exchanger effectiveness

$$\epsilon = \dot{Q}/[C_{\min}(T_{h,i} - T_{c,i})]$$

Heat-capacity rate is mass flow times specific heat for a single-phase stream. The smaller rate limits the maximum possible sensible heat transfer.



Read exchanger terminal temperatures at the same physical end

The hot stream enters at 80 °C and leaves at 50 °C. The cold stream enters at 30 °C on the right and leaves at 60 °C on the left. Equal heat-capacity rates and uniform conductance per length give parallel lines. The local driving difference is 20 K throughout, even though the cold outlet is hotter than the hot outlet.

DERIVATION AND APPLICATIONS

Where the logarithmic mean temperature difference comes from

A local heat-transfer coefficient multiplies the temperature difference at that location. Integrating one exchanger therefore requires the changing driving difference across its whole area. Take a steady counterflow exchanger with no phase change, constant positive capacity rates $C_h = \dot{m}_h c_h$, $C_c = \dot{m}_c c_c$, constant overall coefficient U and no external heat loss. Measure area coordinate A' along the hot flow. Let $d\dot{Q}$ be the positive heat transfer from hot to cold over the increment. Both fluid temperatures decrease along this chosen coordinate because the cold stream flows the other way.

$$d\dot{Q} = U(T_h - T_c)\, dA',$$

$$dT_h = -d\dot{Q}/C_h,$$

$$dT_c = -d\dot{Q}/C_c.$$

Thus $\Delta = T_h - T_c$ changes linearly with cumulative heat transfer. Write its two physical terminal values as $\Delta_1 = T_{h,i} - T_{c,o}$ and $\Delta_2 = T_{h,o} - T_{c,i}$:

$$d\Delta = (1/C_c - 1/C_h) d\dot{Q},$$

$$\Delta_2 - \Delta_1 = (1/C_c - 1/C_h)\dot{Q}.$$

To recover total area, integrate $dA' = d\dot{Q}/(U\Delta)$ using the preceding linear relation. Eliminating the reciprocal capacity difference leaves

$$\dot{Q} = UA\,\frac{\Delta_1 - \Delta_2}{\ln(\Delta_1/\Delta_2)}.$$

For equal terminal differences the apparent zero divided by zero has limit Δ_1. When one positive approach tends toward zero at a fixed heat duty, the necessary area diverges. Terminal differences must pair the temperatures at the same physical end, not two inlet temperatures chosen from opposite ends.

WORKED EXAMPLE

Size an exchanger from measured terminal temperatures

Change the inputs and test this calculation.

A hot stream cools from 90 to $55\,^{\circ}$C while the counterflow cold stream warms from 20 to $60\,^{\circ}$C. The terminal differences are 30 and 35 K, so their logarithmic mean is

$$5/\ln(35/30) = 32.4\,\mathrm{K}.$$

For a duty of 70.0 kW and $U = 500$ W/(m² K), the required area is 4.32 m². The capacities 2.00 and 1.75 kW/K independently deliver the same 70.0 kW.

SECOND APPLICATION

Compare equal duty at a tighter terminal approach

Change the inputs and test this calculation.

A redesign has measured terminal differences of 5.00 and 35.0 K at the same duty and coefficient. The logarithmic mean drops to $30/\ln 7 = 15.4$ K and the required area rises to 9.08 m². Using the arithmetic mean 20.0 K would give only 7.00 m². The region with low driving difference consumes disproportionately more area per kilowatt transferred.

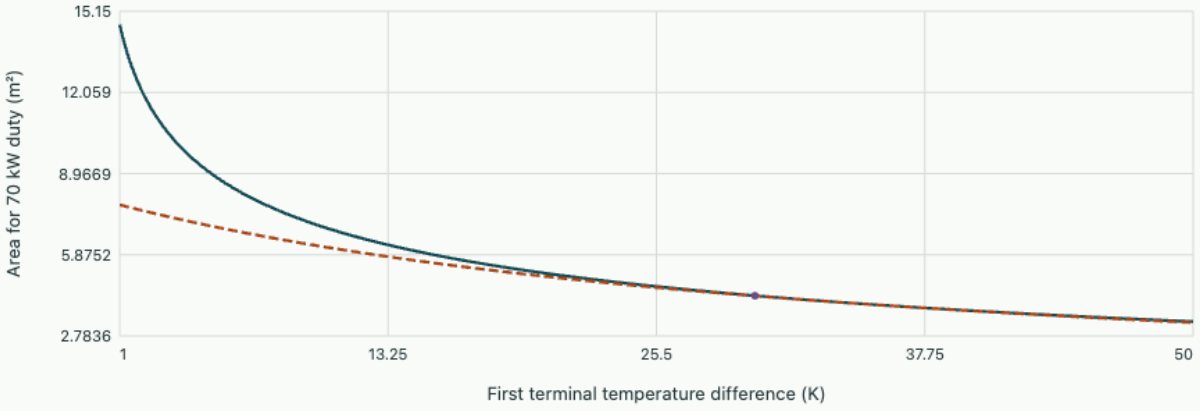
Try it · Lmtd area

Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Selected first terminal difference (K)	30
Second terminal difference (K)	35
Overall coefficient (W/m²/K)	500



LMTD 32.4358 K; required area 4.31622 m². Equal terminal differences use the continuous limit. Positive paired approaches at each physical end are required.

Start from this chapter. Predict outlet temperatures using the counterflow effectiveness relation. At fixed conductance, increasing NTU here reduces the smaller heat-capacity rate. Determine whether effectiveness and total power move together.

OneSimulation laboratory · Explain why greater exchanger effectiveness can give less power

Set the maximum possible heat rate

In a two-stream heat exchanger, the hot temperature falls as the cold temperature rises. Treat each stream as single phase with constant heat capacity rate $C = \dot{m}c_p$ in watts per kelvin. Neglect heat leakage, axial conduction and changes in kinetic energy. A steady overall balance gives

$$Q = C_h(T_{hi} - T_{ho}) = C_c(T_{co} - T_{ci}).$$

Performing this balance alone leaves both outlet temperatures and the transferred heat rate unknown.

The stream with smaller capacity rate would undergo the larger temperature change for a given heat rate. Even in an ideal counterflow exchanger it cannot pass the inlet temperature of the other stream at its nearest end. Thus $Q_{\max} = C_{\min}(T_{hi} - T_{ci})$. Effectiveness is $\varepsilon = Q/Q_{\max}$, which immediately gives both outlet temperatures once the heat transfer law fixes ε.

Apply a balance to an area element

Let A increase along the hot stream and write the local temperature difference $\Delta = T_h - T_c$. An element transfers $dQ = U\Delta\, dA$. In counterflow both temperature profiles decrease when followed in the hot-stream direction, giving $dT_h = -dQ/C_h$, $dT_c = -dQ/C_c$. Hence $d\ln\Delta/dA = U(1/C_c - 1/C_h)$. Integrate over area and express each end difference in terms of inlet temperatures and total heat rate using the stream balances.

$$NTU = UA/C_{\min},$$

$$C_r = C_{\min}/C_{\max},$$

$$\varepsilon = \frac{1 - \exp[-NTU(1 - C_r)]}{1 - C_r \exp[-NTU(1 - C_r)]}.$$

The sign in the differential balance follows from the chosen spatial direction. In parallel flow the cold stream warms along that coordinate and the local temperature difference decays more rapidly. For equal capacity rates the counterflow formula has a removable zero over zero. Its limit is $NTU/(1 + NTU)$, which can also be obtained from the constant temperature difference in that special case.

OneSimulation · Can the cold outlet leave hotter than the hot outlet?

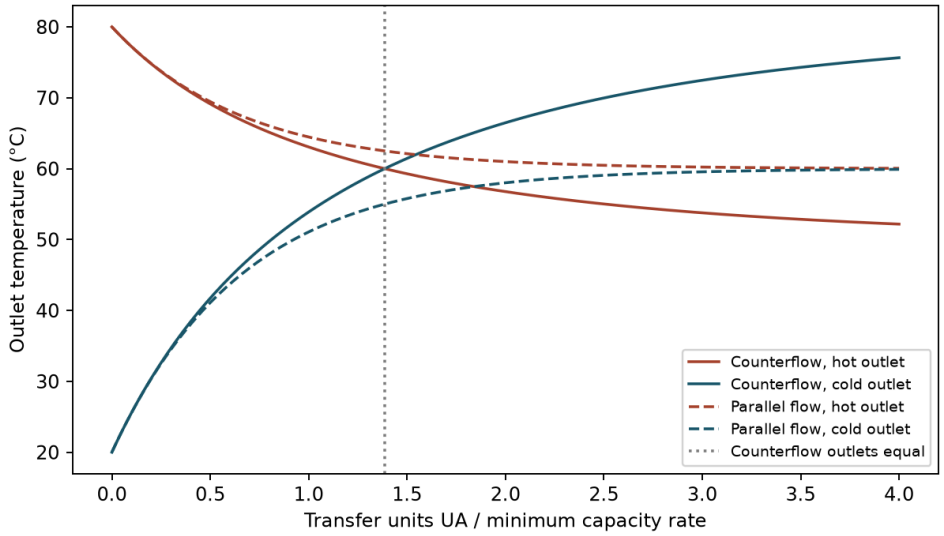
The hot stream enters at 80 °C with a heat-capacity rate of 2 kW/K. The cold stream enters at 20 °C with a capacity rate of 1 kW/K. Compare the same heat exchanger conductance in parallel flow and counterflow. Obtain the two discharges from the same transferred heat rate $\dot{Q}$, using $\dot{C} = \dot{m}c_p$.

$$T_{h,o} = T_{h,i} - \dot{Q}/\dot{C}_h,$$

$$T_{c,o} = T_{c,i} + \dot{Q}/\dot{C}_c.$$

Equal outlets occur at 60 °C and require 40 kW, or effectiveness $2/3$. Parallel flow approaches this common temperature asymptotically. Counterflow reaches it at $\mathrm{NTU} = 2\ln 2$. At NTU 2 the two OneSimulation models predict cold and hot outlets of 66.5 °C and 56.8 °C in counterflow, against 58.0 °C and 61.0 °C in parallel flow. The counterflow outlets sit at opposite ends; heat still travels locally from the hotter fluid to the colder one.

```
"""Check for an outlet temperature cross in a two-stream heat exchanger."""
import matplotlib.pyplot as plt
import numpy as np
from onesim.physics.systems.process import counterflow_effectiveness, parallel_flow_effectiveness

hot_in, cold_in = 80., 20.       # °C
hot_rate, cold_rate = 2., 1.     # kW/K
capacity_min = min(hot_rate, cold_rate)
ratio = capacity_min/max(hot_rate, cold_rate)
ntu = np.linspace(0, 4, 201)

fig, ax = plt.subplots(figsize=(7, 4), constrained_layout=True)
for name, model, style in [("Counterflow", counterflow_effectiveness, "-"),
                           ("Parallel flow", parallel_flow_effectiveness, "--")]:
    heat = model(ntu, ratio)*capacity_min*(hot_in-cold_in)  # kW
    hot_out = hot_in - heat/hot_rate
    cold_out = cold_in + heat/cold_rate
    assert np.allclose(hot_rate*(hot_in-hot_out), cold_rate*(cold_out-cold_in))
    ax.plot(ntu, hot_out, style, color="#a5442e", label=f"{name}, hot outlet")
    ax.plot(ntu, cold_out, style, color="#19596a", label=f"{name}, cold outlet")

ax.axvline(2*np.log(2), color="0.5", linestyle=":", label="Counterflow outlets equal")
ax.set(xlabel="Transfer units UA / minimum capacity rate", ylabel="Outlet temperature (°C)")
ax.legend(fontsize=8)
fig.savefig("engineering-thermodynamics.png", dpi=170)
```



Compare outlets at the same horizontal coordinate. A temperature cross in counterflow describes two different positions along the exchanger.

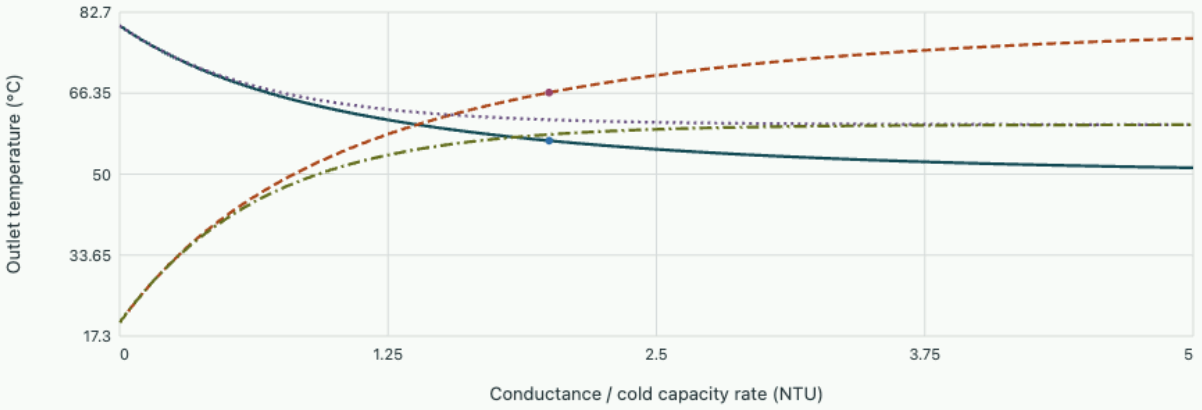
Download the Python example

Change the inputs of this exact example and compare the recalculated curves with the recorded OneSimulation output above. The browser evaluates the stated analytical or modal reference.

Open this live calculation

Try the calculation

Hot inlet (°C)	80
Cold inlet (°C)	20
Cold / hot capacity rate	0.5
Selected NTU	2



Hot, counterflow · Cold, counterflow · Hot, parallel flow · Cold, parallel flow · Chosen hot counterflow · Chosen cold counterflow

At selected NTU, counterflow cold/hot=66.5/56.8°C; parallel=58.0/61.0°C. Counterflow outlets occur at opposite physical ends. Constant properties, no ambient loss.

Predict both outlet temperatures

A counterflow cooler receives a hot stream at 90 °C with $C_h = 2.0\,\mathrm{kW/K}$ and a cold stream at 20 °C with $C_c = 1.0\,\mathrm{kW/K}$. Its conductance UA is 1.50 kW/K. Thus $C_r = 0.5$, $NTU = 1.5$, and the exponential is

$$e^{-0.75} = 0.4724.$$

Effectiveness is 0.6908. Consequently $Q = 0.6908(1.0)(90 - 20) = 48.4\,\mathrm{kW}$,

$$T_{ho} = 90 - 48.4/2 = 65.8^{\circ}\mathrm{C},$$

$$T_{co} = 20 + 48.4 = 68.4^{\circ}\mathrm{C}.$$

The cold outlet is warmer than the hot outlet. They are at opposite physical ends, so this does not imply a local reversal of heat flow. The end differences are $90 - 68.4 = 21.6$ K and $65.8 - 20 = 45.8$ K, both positive. Their logarithmic mean is approximately $(45.8 - 21.6)/\ln(45.8/21.6) = 32.2\,\mathrm{K}$, independently recovering $Q = UA\Delta T_{\mathrm{lm}} \simeq 48.3\,\mathrm{kW}$. This local check catches errors caused by swapping the terminal temperatures of a counterflow exchanger.

Predict the counterflow experiment

A smaller stream capacity rate raises NTU when the exchanger conductance is held fixed, yet it can reduce heat transfer in watts because less thermal capacity passes each second. For inlet temperatures 350 and 290 K and $C_r = 0.60$, use the counterflow expression to predict the effectiveness at NTU 1.5 and 3.0. Then halve both capacity rates, as the model does to obtain the second NTU, and predict the ratio of heat rates. Check both outlet temperatures and the hot-stream heat rate, paying attention to which stream has the smaller capacity. Open the heat-exchanger experiment.

Identify which side controls an exchanger’s conductance

The effectiveness relation uses a single conductance UA, but a designer may need to determine whether the hot film, cold film or fouling dominates it. Write the total thermal resistance in K/W as $R_{\mathrm{tot}} = 1/(h_h A_h) + R_w + R_f + 1/(h_c A_c)$. Areas must be kept explicit if the two sides use different reference areas. Suppose a test can vary hot-side flow while holding cold-side flow, wall resistance and deposit condition fixed. Let a supported correlation in the tested region imply $h_h \propto \dot{m}_h^{\,n}$. With $m_r = \dot{m}_h / \dot{m}_{\mathrm{ref}}$, the relation takes the linear form

$$(UA)^{-1} = R_{\mathrm{fixed}} + R_{h,\mathrm{ref}}\, m_r^{-n}.$$

This is the basis of a Wilson-plot style separation. Each measured UA can be obtained from balanced heat duty and the correct logarithmic mean temperature difference. The intercept estimates the combined fixed resistance; the slope estimates the hot-film resistance at reference flow. One set of measurements alone does not separate cold film from wall and deposits individually. Temperature-dependent material properties and changes in the opposite stream can spoil the assumed linear relation.

Treating UA as proportional to flow whenever convection rises would ignore the series resistances left in its path. As m_r becomes large the hot-film term tends to zero, leaving an asymptote $1/R_{\mathrm{fixed}}$ rather than unbounded conductance. Pumping power typically rises over this same change of flow.

Use a flow sweep to distinguish fouling from film resistance

A clean counterflow rig at reference hot-side mass flow has $UA = 500\,\mathrm{W/K}$, giving total resistance $0.00200\,\mathrm{K/W}$. At twice the flow it has $UA = 600\,\mathrm{W/K}$, or $0.001667\,\mathrm{K/W}$. Over this interval use measured exponent $n = 0.80$, so

$$2^{-0.8} = 0.5743.$$

Subtracting the two equations gives

$$R_{h,\mathrm{ref}} = (0.00200 - 0.001667)/(1 - 0.5743) = 0.000783\,\mathrm{K/W}.$$

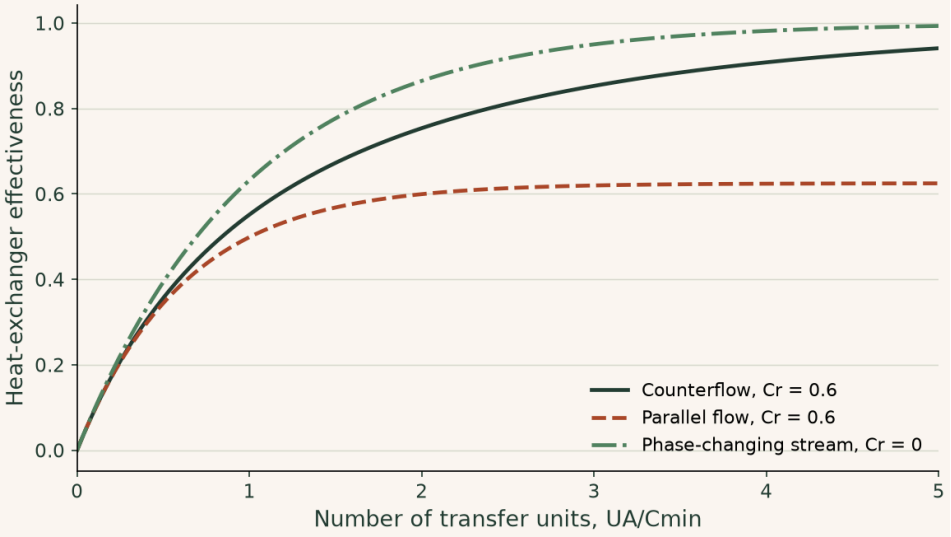
The combined fixed resistance is $0.001217\,\mathrm{K/W}$. Even an ideal hot film could raise conductance only to about 822 W/K under the unchanged fixed-side assumptions.

A later test at reference flow gives $UA = 400\,\mathrm{W/K}$, so total resistance has grown to 0.00250 K/W. If a second flow sweep confirms the same hot-film slope, the added fixed resistance is 0.00050 K/W. For a heat-transfer area of 2.0 m² on the chosen basis this corresponds to fouling resistance $R''_f = 0.0010\,\mathrm{m^2K/W}$. If instead the slope has changed, assigning the entire difference to deposits is unwarranted; changed flow distribution or fluid properties may contribute. Temperature measurement also matters strongly when the two terminal differences are close to sensor uncertainty. Compare the hot and cold enthalpy duties before fitting the conductance and retain the residual across all measured flow settings.

A third clean test can challenge the fitted exponent. At four times reference hot flow, the predicted resistance is $0.001217 + 0.000783(4^{-0.8}) = 0.001475\,\mathrm{K/W}$, or $UA \simeq 678\,\mathrm{W/K}$. A systematic departure that grows at high flow could reflect changing cold-side properties, thermocouple bias or failure of the hot-film correlation. Two calibration points always fit a two-parameter straight line; the independent third point actually tests the assumed dependence.

Determine whether a specified exchanger effectiveness is attainable

Fix both inlet temperatures and $C_{\min}$. Changing area changes NTU. For constant properties and no losses the effectiveness divides actual heat rate by $C_{\min}(T_{h,i} - T_{c,i})$. Compare two flow arrangements and the isothermal-stream limit.



With capacity ratio 0.6, parallel flow approaches only $1/1.6 = 0.625$ however large the area. Counterflow can exceed 0.8. An area calculation that ignores the arrangement can return a target that the selected exchanger cannot reach.

▶ **Read values from the plot**

Interactive calculation · change the inputs of this example

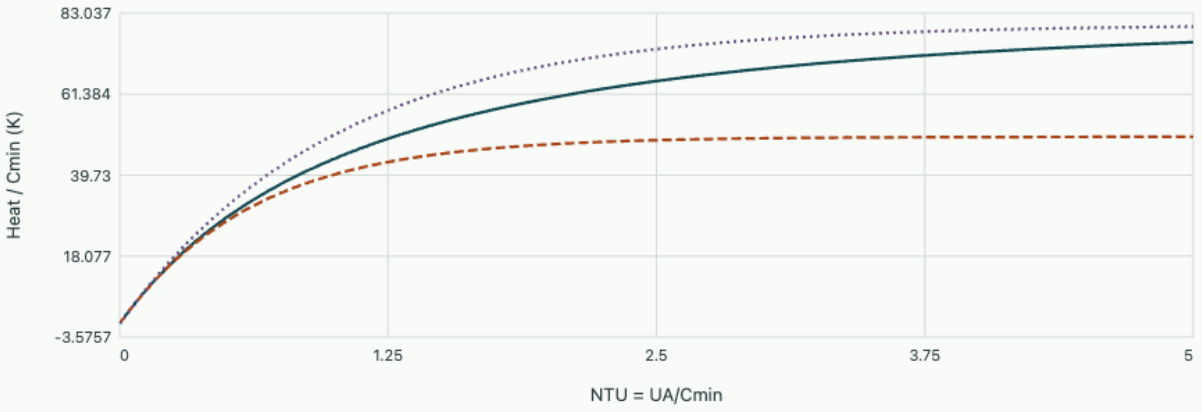
Move the capacity-rate ratio toward one and compare the finite-area behaviour with the phase-change limit.

The curves below are recalculated from this section's equations in your browser. The fixed reference plot above remains available for comparison.

Open this live calculation

Try the calculation

Capacity-rate ratio	0.6
Hot inlet (°C)	100
Cold inlet (°C)	20



At NTU 2, counterflow effectiveness 0.754. Parallel flow cannot exceed 0.625.

Study the full derivation

How a gravity-drained condensate film limits heat transfer · Thermal processes and transport in engineering

Saturated vapour condenses on a cool vertical plate. A liquid film thickens as it drains, adding thermal resistance toward the bottom. Derive the local heat-transfer coefficient from coupled momentum and energy balances, then determine which observations would invalidate the smooth-film model.

Chapter 12 · Thermal radiation and enclosure exchange

Use surface emission and reflected irradiation to derive the net heat exchange between parallel grey surfaces.

FROM FIRST PRINCIPLES

Derive enclosure heat exchange from emission, reflection and absorption

Consider an opaque diffuse grey surface at uniform absolute temperature T. Let G denote incoming radiant power per area and J total outgoing radiant power per area. The Stefan–Boltzmann law gives blackbody emission $E_b = \sigma T^4$. Emissivity ϵ scales actual emission; under the grey thermal model absorptivity equals ϵ, so the reflected fraction is $1 - \epsilon$. Add emitted and reflected contributions to obtain

$$J = \epsilon E_b + (1 - \epsilon) G,$$
$$q''_{\text{out}} = J - G = \epsilon (E_b - G).$$

In an enclosure each surface receives irradiation from the radiosities it sees. For two large parallel facing plates, $G_1 = J_2$ and $G_2 = J_1$. Solving the two linear radiosity equations for their net outward flux difference yields

$$q''_{1 \to 2} = \frac{\sigma (T_1^4 - T_2^4)}{1/\epsilon_1 + 1/\epsilon_2 - 1}.$$

The denominator incorporates repeated reflections. Equal temperatures give zero net exchange even though each surface radiates. Black plates reduce to the difference of their blackbody emissions. A small surface that sees itself or other walls requires the corresponding geometric view factors in its irradiation balance.

DERIVATION AND APPLICATIONS

Determine a small surface temperature in a large radiating enclosure

A small opaque specimen suspended in a large enclosure receives radiation over its entire visible hemisphere. Under diffuse grey conditions Kirchhoff's relation gives equal absorptivity and emissivity ϵ; a large isothermal black enclosure at T_w supplies irradiation $G = \sigma T_w^4$. Assume the specimen is isothermal at T, sees only the enclosure, and also exchanges heat with a gas at T_g through a known constant film coefficient h. Neglect conduction through supports. Let p'' be a uniform internal heat supply per unit radiating area.

Emission and absorption occur separately, so the net outward radiative heat flux is the emitted flux minus the absorbed irradiation. The steady balance becomes

$$q''_{\text{rad}} = \epsilon\sigma(T^4 - T_w^4),$$
$$p'' = h(T - T_g) + \epsilon\sigma(T^4 - T_w^4).$$

To compare film and radiation effects without approximating a fourth power, factor the difference. Define a secant radiation coefficient between the two actual temperatures:

$$h_r(T, T_w) = \epsilon\sigma(T + T_w)(T^2 + T_w^2),$$
$$q''_{\text{rad}} = h_r(T - T_w).$$

For perturbations about an operating point, hold T_w, T_g fixed and differentiate the balance. The slope governing a small additional input is different from the secant coefficient:

$$\frac{dT}{dp''} = \frac{1}{h + 4\epsilon\sigma T^3}.$$

Both coefficients have units of W/(m² K); their numerical values coincide as $T \to T_w$. Fourth powers require absolute temperatures. The balance is monotone for positive T, giving one physically relevant operating temperature for a compatible specified input.

WORKED EXAMPLE

Size a heater for a test coupon

Change the inputs and test this calculation.

Let $T = 350$ K, $T_w = 300$ K, $T_g = 310$ K, $h = 8.00$ W/(m² K) and $\epsilon = 0.800$. With $\sigma = 5.6704 \times 10^{-8}$ W/(m² K⁴), radiation loss is 313 W/m² and convection loss 320 W/m². A coupon with 0.0100 m² total exposed area requires 6.33 W. The corresponding secant radiation coefficient is $313/50 = 6.26$ W/(m² K).

SECOND APPLICATION

Predict a small calibration shift

Change the inputs and test this calculation.

At the same operating point the tangent radiation coefficient is $4(0.800)\sigma(350)^3 = 7.78$ W/(m² K). Adding 0.100 W over the same area increases p'' by 10.0 W/m². Linearization predicts a temperature rise of $10/(8.00 + 7.78) = 0.634$ K. The differential coefficient applies to this incremental measurement; the secant coefficient gave the original heat exchanged with the enclosure.

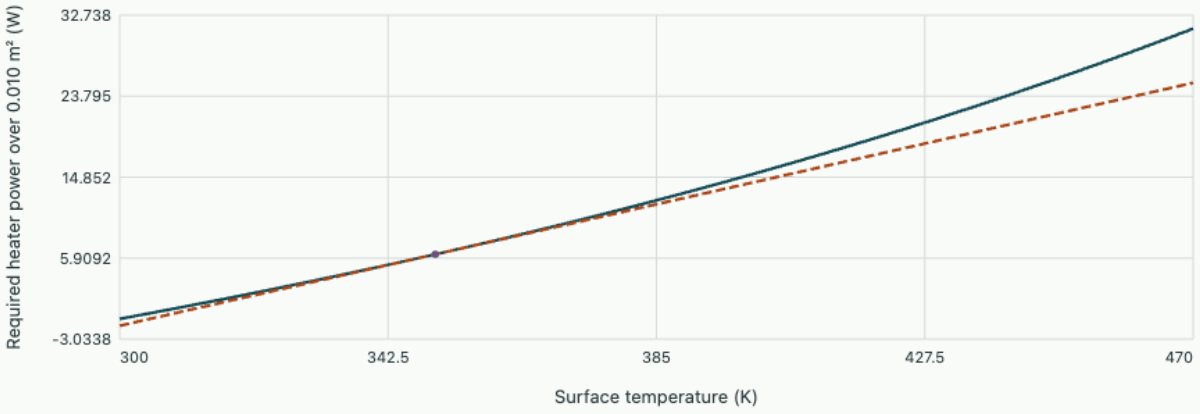
Try it · Radiation heater

Set the inputs to the values in the worked example, then vary one quantity while holding the others fixed. Compare the curve and numerical balance with the second application. The calculation uses the assumptions derived above; its readout identifies bounds such as choking, phase-region changes and inconsistent measurements.

Open this live calculation

Try the calculation

Surface emissivity	0.8
Convection coefficient (W/m²/K)	8
Selected surface temperature (K)	350



Enclosure 300 K, gas 310 K. Heater 6.33288 W; secant radiation coefficient 6.26576 W/(m² K); tangent 7.77975 W/(m² K). Additional 0.100 W gives linearised ΔT=0.633723 K.

Separate what a surface emits from what leaves it

A hot surface transfers energy across a vacuum through electromagnetic radiation. A black surface emits total hemispherical power $E_b = \sigma T^4$ per unit area, using absolute temperature. A diffuse grey opaque surface emits a fraction εE_b and reflects a fraction of incident irradiation G. Under the grey equilibrium-property assumptions, absorptivity equals emissivity and reflectivity is $1 - \varepsilon$. The total radiation leaving the surface, called radiosity, is then

$$J = \varepsilon E_b + (1 - \varepsilon) G.$$

Net radiative heat flux is

$$q'' = J - G.$$

Eliminate the unknown irradiation using the radiosity relation. For a surface of area A and $0 < \varepsilon < 1$, this gives

$$Q = (E_b - J)/[(1-\varepsilon)/(A\varepsilon)].$$

The denominator behaves like a resistance between the blackbody emission potential and the surface radiosity. It vanishes for an ideal black surface, whose outgoing radiation is set solely by its temperature.

Couple two large facing plates

Consider two large parallel diffuse grey plates of equal area with negligible edge losses. Each plate sees only the other, so $G_1 = J_2$ and $G_2 = J_1$. The net rate from 1 to 2 is $A(J_1 - J_2)$. Combine that space resistance $1/A$ with the surface resistances of the two plates in series. After cancellation the net flux is

$$q'' = \frac{\sigma(T_1^4 - T_2^4)}{1/\varepsilon_1 + 1/\varepsilon_2 - 1}.$$

The fourth powers cannot generally be evaluated using Celsius temperatures, even though a simple temperature difference has the same magnitude in kelvins and degrees Celsius. For finite or angled surfaces, a view factor describes how much of one surface’s outgoing radiation reaches the other. Equal area alone does not imply a view factor of one.

Add a freely floating radiation shield

Take $T_1 = 500\,\mathrm{K}$, $T_2 = 300\,\mathrm{K}$, both plate emissivities 0.80 and $\sigma = 5.670 \times 10^{-8}\,\mathrm{W/(m^2K^4)}$. The blackbody difference is $3084\,\mathrm{W/m^2}$. Division by $1/0.8 + 1/0.8 - 1 = 1.5$ gives $2056\,\mathrm{W/m^2}$. Insert an opaque thin shield with emissivity 0.10 on both faces and negligible temperature drop through its thickness. There are now two radiation gaps, each with denominator $1/0.8 + 1/0.1 - 1 = 10.25$.

At steady state the shield exchanges equal heat flux with both plates and stores no energy. With equal gap denominators it follows that $T_s^4 = (T_1^4 + T_2^4)/2$, so $T_s \simeq 433\,\mathrm{K}$. The total resistance denominator is 20.5 and heat flux drops to about 150 W/m². The floating shield temperature is higher than the arithmetic average of the plate temperatures. Solid supports connecting the shield to either plate would add conductive paths and change its temperature and the total heat leak. This explains why a low measured radiative load alone may not determine the performance of an insulated assembly.

Compare radiation with a convection coefficient

For black exchange, factor $T_1^4 - T_2^4 = (T_1 - T_2)(T_1 + T_2)(T_1^2 + T_2^2)$. For the original grey plates the corresponding effective coefficient is

$$2056/(500 - 300) = 10.28\,\mathrm{W/(m^2K)}.$$

It can be compared directly with a convective coefficient for the same temperature difference, but its value changes when either absolute temperature changes.

Combine view-factor reciprocity with grey-surface resistances

Two large parallel plates have equal area and see one another completely. For opposite-sized diffuse grey surfaces in a two-surface enclosure, exchange through the intervening space depends on view factor F_{12}, the fraction of energy leaving surface 1 that directly reaches surface 2. Reciprocity requires

$$A_1 F_{12} = A_2 F_{21}.$$

A convex inner body entirely surrounded by a shell has $F_{12} = 1$; radiation leaving the larger concave shell need not all reach the inner body because some hits the shell again.

Net transfer through the space is $Q = A_1 F_{12}(J_1 - J_2)$, with radiosities J_1, J_2. Combine this relation with both surface resistance expressions and eliminate the radiosities. For uniform isothermal surfaces that can each be represented by one radiosity, the result is

$$Q = \frac{\sigma(T_1^4 - T_2^4)}{(1 - \varepsilon_1)/(A_1\varepsilon_1) + 1/(A_1 F_{12}) + (1 - \varepsilon_2)/(A_2\varepsilon_2)}.$$

For concentric long cylinders with negligible ends $F_{12} = 1$ and $A_1/A_2 = r_1/r_2$. The rate per inner area simplifies to $\sigma(T_1^4 - T_2^4)/[1/\varepsilon_1 + (A_1/A_2)(1/\varepsilon_2 - 1)]$. As the outer area becomes very large its surface resistance has little effect on transfer per unit inner area. When the two areas approach equality the familiar parallel-surface denominator is recovered. This area weighting matters when a shiny outer enclosure surrounds a relatively small part.

Separate radiation and support conduction in a vacuum vessel

A long heated cylinder has outer radius 10 mm and active length 0.50 m, concentric with a cylindrical vacuum jacket of inner radius 40 mm. Take $T_1 = 400\,\mathrm{K}$, $T_2 = 300\,\mathrm{K}$, $\varepsilon_1 = 0.70$, $\varepsilon_2 = 0.20$, and ignore end exchange. Areas satisfy $A_1/A_2 = 0.25$, and $A_1 = 2\pi(0.010)(0.50) = 0.03142\,\mathrm{m^2}$. The effective denominator is

$$1/0.70 + 0.25(1/0.20 - 1) = 2.429.$$

The blackbody flux difference is $5.670 \times 10^{-8}(400^4 - 300^4) = 992.3\,\mathrm{W/m^2}$, giving radiative loss $992.3(0.03142)/2.429 = 12.8\,\mathrm{W}$.

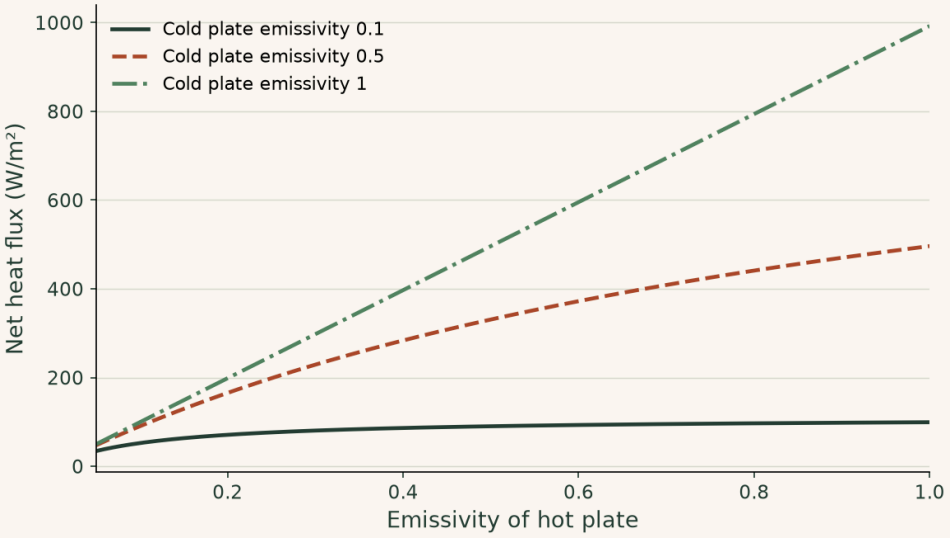
Two solid supports each have area 4.0 mm², length 20 mm and mean conductivity 15 W/(m K). Conduction through both is $2(15)(4.0 \times 10^{-6})(100)/0.020 = 0.60\,\mathrm{W}$, so holding the inner cylinder at its stated temperature requires approximately 13.4 W. A measured heater power of 15 W would leave about 1.6 W unaccounted for, potentially from end radiation, electrical leads, residual gas or emissivity uncertainty. If the jacket is blackened to $\varepsilon_2 = 1$, the radiation rises to $992.3(0.03142)(0.70) = 21.8\,\mathrm{W}$. Improving absorption of the outer wall increases the heat loss from the inner component even though neither temperature changes.

An intercept test can distinguish temperature-dependent pathways more effectively than a single operating point. At a small excess ΔT around 300 K, radiative loss varies approximately as $4\sigma T^3 A_1 \Delta T/[1/\varepsilon_1 + (A_1/A_2)(1/\varepsilon_2 - 1)]$. The original cylinder has linearised radiative conductance about $4(5.67 \times 10^{-8})(300)^3(0.03142)/2.429 = 0.0792\,\mathrm{W/K}$, compared with 0.0060 W/K for its two supports. At large temperature differences the effective radiation conductance rises with absolute temperature while ideal support conduction stays linear. Recording heater power at several balanced temperatures lets those trends be tested, provided the surface emissivities and support conductivity do not themselves vary significantly over the range.

A measurement with electrical power off also establishes any background sensor offset before estimating these comparatively small near-ambient conductances from a temperature difference.

Which surface treatment changes net radiation most?

Two large diffuse-grey opaque plates at 400 K and 300 K face each other across a nonparticipating gap. For unit view factor the net radiation resistance is proportional to $1/\epsilon_h + 1/\epsilon_c - 1$. Vary the hot coating and keep the cold one fixed.



Two black plates exchange approximately 992 W/m². If the cold plate has emissivity 0.1, even a perfectly black hot plate exchanges only about 99 W/m². Treating the temperature difference as a difference of fourth powers is essential here.

▶ **Read values from the plot**

Interactive calculation · change the inputs of this example

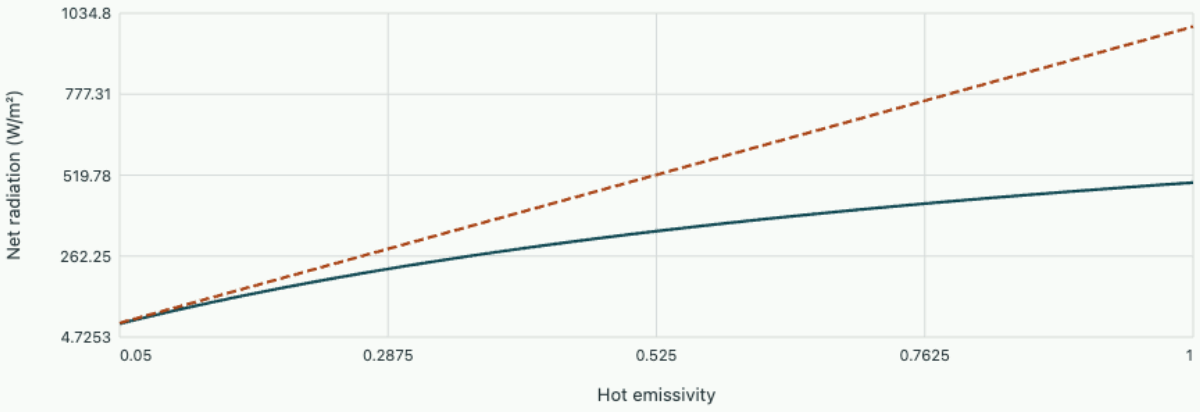
Change the cold coating and compare its effect with changing the hot coating. Temperature enters as an absolute fourth power.

The curves below are recalculated from this section's equations in your browser. The fixed reference plot above remains available for comparison.

Open this live calculation

Try the calculation

Hot temperature (K)	400
Cold temperature (K)	300
Cold emissivity	0.5



Black/black exchange 992.3 W/m². View factor one; diffuse-grey opaque plates.

Study the full derivation

Radiative exchange in a grey enclosure · Thermal processes and transport in engineering

A thin radiation shield does not need to absorb all incoming radiation to reduce heat transfer. Multiple reflections connect every surface to every other visible surface. Build the radiosity equations, recover the two-plate result and infer what can be learned from a measured shield temperature.

Problem sheet 6 · Heat exchangers and radiation shields

Relate thermal conductance and surface properties to local temperature differences and net energy transport.

A balanced counterflow exchanger

Both stream capacity rates are 1.5 kW/K and conductance $UA = 3.0$ kW/K. Inlets are at 80 and 20 °C. Determine effectiveness, heat rate, both outlets and the local temperature difference. Explain why both end differences agree.

▼ **Reveal the worked solution**

At $C_r = 1$, $NTU = 2$ and $\varepsilon = 2/3$. The exchanger transfers $(2/3)(1.5)(60) = 60\,\mathrm{kW}$. Hot outlet is 40 °C and cold outlet 60 °C. At each end the hot stream exceeds the adjacent cold stream by 20 K. Equal capacity rates produce equal profile slopes when followed along the same physical direction.

Infer a required exchanger size

A hot stream of capacity rate 4 kW/K cools from 100 to 75 °C. A counterflow cold stream enters at 25 °C with capacity rate 2 kW/K. Calculate its outlet, the terminal differences and required UA. What is the effectiveness?

▼ **Reveal the worked solution**

The duty is $4(100 - 75) = 100\,\mathrm{kW}$, giving cold outlet $25 + 100/2 = 75^\circ\mathrm{C}$. Terminal differences are 25 and 50 K. Their logarithmic mean is $(50 - 25)/\ln 2 = 36.07\,\mathrm{K}$, so $UA = 2.773\,\mathrm{kW/K}$. The maximum duty is $2(100 - 25) = 150\,\mathrm{kW}$; effectiveness is 0.667.

A shield in a black enclosure gap

Large black parallel walls face one another at 400 and 300 K. A thin black shield is inserted midway using negligible conducting supports. Find unshielded flux, shielded flux and equilibrium shield temperature. Does equal spacing enter the result?

▼ **Reveal the worked solution**

The unshielded flux is

$$\sigma(400^4 - 300^4) = 992\,\mathrm{W/m^2}.$$

Each black gap contributes the same unit denominator, halving the flux to 496 W/m². The shield satisfies $T_s^4 = (400^4 + 300^4)/2$, giving $T_s = 360.3\,\mathrm{K}$. Equal spacing is unnecessary for two effectively infinite parallel plates whose view factors are unity.

Close the energy and exergy balances of a heat exchanger

An insulated counterflow exchanger cools water from 90°C to 60°C at 1.0 kg/s. Another water stream enters at 20°C and leaves at 50°C. Both have c_p = 4.18 kJ/(kg K). Find the cold mass flow, thermal conductance UA, effectiveness and rate of exergy destruction at an ambient temperature of 293.15 K. Ignore pressure losses.

▼ **Worked solution**

The hot side releases 125.4 kW. Conservation of energy gives a cold mass flow of 1.0 kg/s because the two temperature changes have equal magnitude. For counterflow, the end temperature differences are 90−50 = 40 K and 60−20 = 40 K. Their logarithmic mean is 40 K by continuity, giving UA = 3.135 kW/K.

The minimum heat-capacity rate is 4.18 kW/K, and the maximum possible duty is 4.18 × 70 = 292.6 kW. The effectiveness is 0.4286. As a check, NTU = UA/Cmin = 0.75 and for balanced counterflow ε = NTU/(1+NTU) = 0.4286.

With absolute temperatures, the entropy creation rate is 4.18[ln(333.15/363.15)+ln(323.15/293.15)] kW/K, about 0.0469 kW/K. Multiplying by ambient temperature gives about 13.7 kW of exergy destruction. Insulation eliminates the external heat flux, but finite temperature differences inside the exchanger still produce entropy.

Explore the governing equations after solving the sheet: Heat-exchanger effectiveness and finite capacity rates; Thermal radiation and enclosure exchange. Compare the assumptions and units before transferring your chosen parameters; the controls may cover a different numerical range from the written exercise.

Recover an exchanger capacity rate from outlet data

For a steady counterflow exchanger, the baseline hot stream cools from 350 to 309.84 K and the cold stream warms from 290 to 314.20 K. Hot-side power loss is 0.2168 W. Infer the hot heat-capacity rate, estimate the cold rate assuming equal powers, and infer UA from the stated NTU of 1.5. Why is the inferred cold rate approximate?

▼ **Work through the solution**

The hot balance gives $C_h = \dot{Q}/(350 - 309.84) = 0.00540$ W/K. Setting the cold gain equal gives $C_c = 0.2168/(314.20 - 290) = 0.00896$ W/K. Their ratio is 0.603, close to the input 0.60. The smaller capacity rate is the hot one, so $UA = \mathrm{NTU}\, C_{\min} = 0.00810$ W/K. The finite mesh yields a small mismatch between the two streamwise estimates of heat rate. Using the measured cold-side power instead would directly test that balance.

Use the OneSimulation laboratory to check your answer

Vary the corresponding laboratory reference interactively. Compare the recalculated observable with the recorded OneSimulation runs and the approximation used in this exercise.

Chapter 13 · Choosing a pressure ratio for a gas-turbine cycle

Why does a bigger pressure ratio eventually reduce the work from a gas turbine? Follow one kilogram of gas round four devices, keep separate accounts for heat and work, and see what changes when the compressor and turbine are imperfect.

Label four states before comparing designs

Use an air-standard Brayton cycle. A compressor takes gas from state 1 to 2. A constant-pressure heater raises it to the maximum allowed temperature at 3. A turbine expands the gas to the original pressure at 4. A cooler returns it to 1. Neglect kinetic and potential energy changes and pressure losses in the heat exchangers. The heater and cooler replace combustion and exhaust in this closed model. Use a calorically perfect gas with constant c_p and $\gamma = c_p/c_v$. For an ideal adiabatic reversible device the entropy change vanishes. Integrating the ideal-gas entropy differential gives

$$0 = c_p \ln(T_{2s}/T_1) - R \ln r,$$

$$x = r^{(\gamma-1)/\gamma},$$

$$T_{2s} = x T_1,$$

$$T_{4s} = T_3/x.$$

The pressure ratio $r = p_2/p_1 = p_3/p_4$ is dimensionless. The subscript s identifies a state at the specified exit pressure reached along an isentrope. The turbine and compressor compare different inlet states. A common bookkeeping mistake is to use the compressor inlet temperature to calculate turbine work.

Subtract the compressor demand before optimizing output

Write positive turbine output and positive compressor consumption separately. The first law then gives the net delivered work w and the heat input q per kilogram. Maximizing w at fixed T_1, T_3 amounts to differentiating with respect to x.

$$\frac{w}{c_p} = T_3(1 - 1/x) - T_1(x - 1),$$

$$\frac{q}{c_p} = T_3 - x T_1,$$

$$\frac{1}{c_p}\frac{dw}{dx} = T_3/x^2 - T_1.$$

The maximum occurs at $x = \sqrt{T_3/T_1}$. The second derivative is negative. There the compressor exit and turbine exit have the same temperature. Dividing work by heat also gives the ideal efficiency $1 - 1/x$, which increases throughout the range that admits positive heat input. A design chosen for maximum shaft work and a design chosen for maximum thermal efficiency can therefore have different pressure ratios. At the extreme $x = T_3/T_1$, both heat input and net work vanish.

A design point and its sensitivity to losses

At 300 K and 1200 K with $\gamma = 1.4,\ c_p = 1.005$ kJ/(kg K), maximum ideal work occurs at $x = 2$, hence

$$r = 2^{3.5} = 11.31.$$

Both intermediate temperatures are 600 K. The compressor consumes 301.5 kJ/kg and the turbine produces 603.0 kJ/kg. The delivered 301.5 kJ/kg for 603.0 kJ/kg heating gives 50% efficiency.

Now set the compressor isentropic efficiency to 0.80 and the turbine efficiency to 0.90. The actual compressor rise is divided by its efficiency, while the ideal turbine drop is multiplied by its efficiency. At the same pressure ratio the actual temperatures are 675 K and 660 K, the net work is 165.8 kJ/kg, and heat input is 527.6 kJ/kg. The corresponding work derivative vanishes at $x = \sqrt{\eta_t \eta_c T_3 / T_1}$. This moves the optimum to a smaller pressure ratio. Use the curves to locate both maxima before calculating their positions.

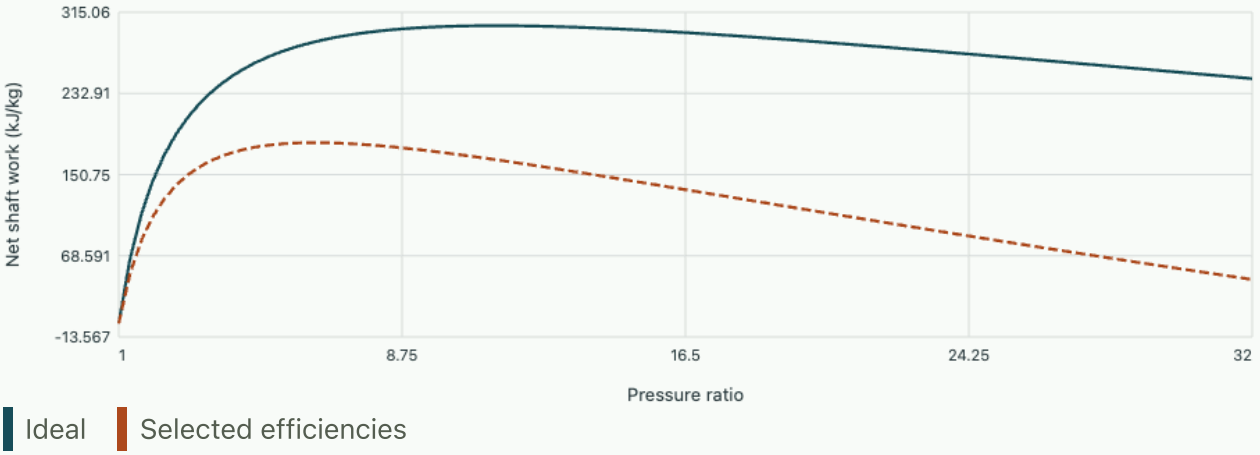
Start with both device efficiencies set to one, record the pressure ratio at maximum work, then reduce only compressor efficiency. Predict which side of the old optimum becomes more costly and compare your reasoning to the shifted peak.

Open this experiment online

Try the calculation

Compressor efficiency **0.8**

Turbine efficiency **0.9**



The best pressure ratio for net work with these component efficiencies is 6.37, giving 183.1 kJ/kg. At the ideal optimum 11.31 the real cycle gives 165.8 kJ/kg.

A related OneSimulation calculation

A regenerator adds hot and cold stream constraints to the shaft-work calculation. Use the course counterflow-exchanger study to examine how effectiveness and actual transferred power respond when heat-capacity rates change.

Compare heat-exchanger cases in OneSimulation

Tutorial · can the exhaust preheat the compressed gas?

Take the same imperfect components at $r = 4$. Calculate temperatures 2 and 4. An insulated counterflow regenerator has equal capacity rates and effectiveness 0.70. Determine the saving in heater duty and explain why its benefit vanishes at certain pressure ratios.

▼ **Show the state and heat calculations**

Here $x = 1.4860$, so $T_2 = 482.3$ K and $T_4 = 846.8$ K. The compressed stream gains $0.70(846.8 - 482.3) = 255.2$ K, saving about 256.5 kJ/kg of external heat. Both shaft devices retain the same inlet and exit states. The model predicts no change in their work. When the exhaust arrives colder than the compressed stream it cannot serve as the hot source for this duty. A practical design would also account for exchanger pressure loss.

Further reading · MIT Unified Engineering · the ideal Brayton cycle

Study the full derivation

Staged compression with an imperfect intercooler · Thermal processes and transport in engineering

A laboratory compressed-air supply has two compressor stages and an intercooler. Equal pressure ratios minimise work only under particular assumptions. Derive the optimum, compare it with a single-stage machine and identify which temperature and power measurements reveal imperfect cooling or unequal efficiency.

Chapter 14 · Why a real gas cools or warms across a throttle

A gas passes through an insulated porous plug. Its pressure drops and its measured temperature changes. Start with the energy flowing through the plug and obtain a prediction from the equation of state.

Establish which property travels through the plug

A steady control volume with one inlet, one outlet and no shaft work has $h_1 + u_1^2/2 + g z_1 = h_2 + u_2^2/2 + g z_2$ when the casing is insulated. Choose the measurement stations in large, level pipes where changes in mean kinetic energy are negligible. Then $h_1 = h_2$. The pressure gradient is sustained by irreversible friction in the plug. Along the family of equilibrium outlet states, a small pressure change produces $dT = \mu_{JT} dp$, with $\mu_{JT} = (\partial T/\partial p)_h$. This family lets us compute endpoints without resolving the interior pore flow.

Recover the pressure dependence of enthalpy

Use molar properties. For a simple compressible pure substance, $dh = T\,ds + v\,dp$. View entropy as a function of T, p. At fixed pressure its temperature derivative is c_p/T. To obtain its pressure derivative, differentiate $g = h - Ts$. Its differential is $dg = -s\,dT + v\,dp$. Equality of mixed derivatives gives

$$(\partial s/\partial p)_T = -(\partial v/\partial T)_p.$$

Substitution yields

$$dh = c_p\,dT + \left[v - T\left(\frac{\partial v}{\partial T}\right)_p\right] dp,$$

$$\mu_{JT} = \frac{T(\partial v/\partial T)_p - v}{c_p}.$$

For an ideal gas $v = RT/p$ the two numerator terms cancel at all temperatures. Real gas attractions and finite molecular volume affect this near cancellation. Since $dp < 0$ for the pressure drop, positive μ_{JT} indicates cooling. Check units by dividing cubic metres per mole by joules per mole per kelvin. The result is kelvins per pascal.

An instructive gas with an inversion temperature

Represent a dilute model gas by $v = RT/p + B(T)$, with $B = b - a/(RT)$. Here a, b are constants and the omitted equation-of-state terms become important at higher density. The ideal part cancels identically. Differentiating the correction gives

$$\mu_{JT} \simeq \frac{TB'(T) - B(T)}{c_p} = \frac{2a/(RT) - b}{c_p},$$

$$T_{\mathrm{inv}} = \frac{2a}{Rb}.$$

With $b = 4.0 \times 10^{-5}$ m³/mol, $a = 0.080$ Pa m⁶/mol², $R = 8.314$ J/(mol K) and $c_p = 30$ J/(mol K), the inversion lies near 481 K. At 300 K, $\mu_{JT} = 0.0805$ K/bar. A modest drop from 3 bar to 1 bar cools this gas by about 0.161 K. At 600 K the derivative is about −0.0264 K/bar. The same pressure drop then warms the gas by 0.0528 K. Treat this as a local calculation. Over a large temperature range one should integrate the property derivative using a validated equation of state.

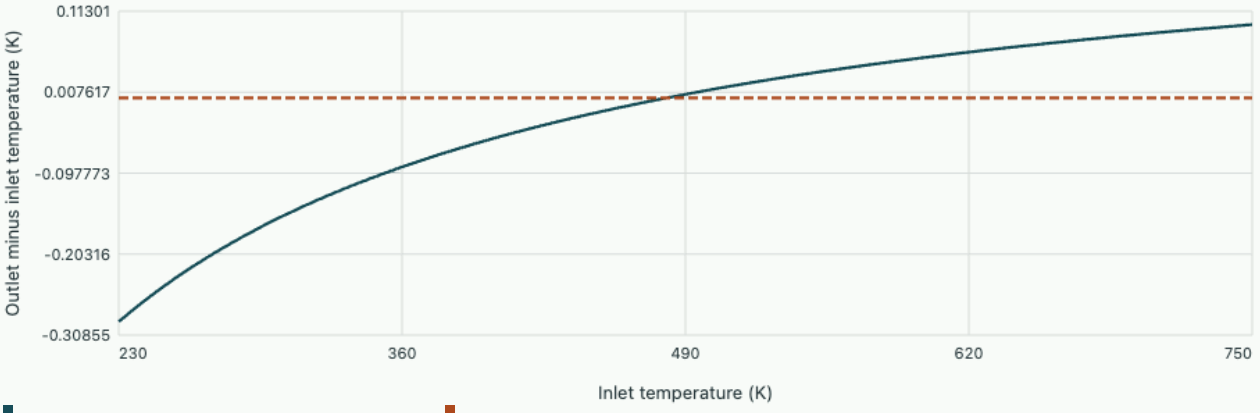
Alter the pressure drop while holding the attractive constant fixed. Locate the zero crossing at two pressure drops. Explain why it remains at the same temperature in this first-order calculation and why the cooling magnitudes change.

Open this experiment online

Try the calculation

Attractive constant a (Pa m⁶/mol²) **0.08**

Pressure drop (bar) **2**



The inversion temperature is 481.1 K. At 300 K the local pressure-drop estimate gives an outlet temperature change of -0.161 K.

Tutorial · infer the attractive correction

A low-density gas at 320 K gives a positive slope of 0.060 K/bar in a throttle experiment. Assume the same b, c_p as above. Estimate a and the inversion temperature. How would an unaccounted heat leak from a warm room bias the inference if the gas cools?

▼ **Show the inference and sign check**

Convert the slope to 6.0×10^{-7} K/Pa. Rearranging gives $a = (RT/2)(c_p \mu_{JT} + b) = 0.0772$ Pa m⁶/mol², whence $T_{\mathrm{inv}} = 464$ K. A warm leak partially cancels cooling, making the observed outlet temperature drop too small. Because the pressure drop is also negative, the apparent positive slope is smaller. In this parameter fit the estimated a is therefore too low.

Further reading · MIT 2.43 · property relations and the Joule–Thomson coefficient

Chapter 15 · When another centimetre of cooling fin stops helping

A fin adds surface area, but heat must travel down its length before it can reach that area. Derive the temperature profile from an energy balance and quantify how much of the nominal area contributes to cooling.

Balance a short slice with distributed convection

Consider a straight prismatic fin of conductivity k, cross-section A, exposed perimeter P and length L. Hold the root at T_b, let a fluid at T_∞ surround the fin and use a uniform convection coefficient h. Temperature is assumed nearly uniform across each cross-section. A small transverse Biot number $h(A/P)/k$ supports that assumption. Ignore radiation and heat loss from the tip. Define $\theta = T - T_\infty$. The axial heat entering a slice, less the heat leaving its far face, must equal $hP\theta\,dx$.

$$q_x = -kA\,\theta',$$

$$-\frac{dq_x}{dx} = hP\theta,$$

$$\theta'' - m^2\theta = 0,$$

$$m = \sqrt{hP/(kA)}.$$

A positive temperature excess makes the heat flow diminish downstream. For an insulated tip $q_x(L) = 0$, so a hyperbolic cosine centered at the tip satisfies the zero-slope condition. Its coefficient follows from the root temperature.

$$\frac{\theta(x)}{\theta_b} = \frac{\cosh[m(L-x)]}{\cosh(mL)},$$

$$Q_b = kAm\theta_b \tanh(mL).$$

Efficiency follows by comparing with an isothermal fin

If every part of the lateral area stayed at T_b, the heat transfer would be $hPL\theta_b$. The actual-to-ideal ratio is therefore

$$\eta_f = \tanh(mL)/(mL).$$

For small mL the temperature falls little along the wall and doubling the area nearly doubles the heat flow. For large mL the root heat flow approaches $kAm\theta_b$. Near the long-fin limit increasing conductivity or the root cross-section can buy more cooling than adding length. Efficiency alone does not measure total heat removal because it contains the nominal area in its denominator.

Two fins made from the same stock

Use a rectangular aluminium strip, 20 mm wide, 2 mm thick, $k = 200$ W/(m K), $h = 25$ W/(m² K) and $\theta_b = 50$ K. All four long faces are exposed. With $A = 4.0 \times 10^{-5}$ m² and $P = 0.044$ m, $m = 11.73$ m⁻¹. At 50 mm length, $mL = 0.586$, $Q_b = 2.47$ W and $\eta_f = 0.899$. The tip remains about 42.5 K above ambient.

At 150 mm, the heat rate is 4.43 W and the efficiency drops to 0.536. The material and exposed area tripled, while the heat transfer increased by about 79%. A fin root bonded through a finite contact resistance R_c would run cooler than the heated base. Replacing the temperature excess by $T_{\text{base}} - T_\infty - Q_b R_c$ in the root heat equation gives a direct way to include the bond in a thermal circuit.

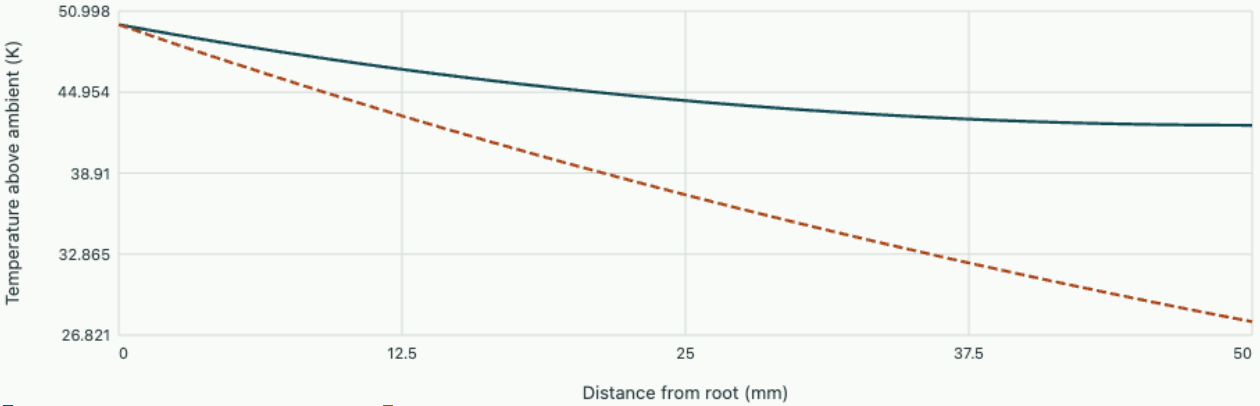
Choose a short fin and inspect how far the infinite-fin approximation is from the insulated-tip temperature. Increase length until the root heat rate is almost saturated. Estimate the added cooling from another centimetre at that point.

Open this experiment online

Try the calculation

Length (mm) **50**

Convection coefficient (W/m² K) **25**



The finite tip is 42.5 K above ambient. Root heat flow is 2.47 W and the fin efficiency is 89.9%.

Tutorial · select a length from a marginal-benefit criterion

For the same strip, require 95% of the infinitely long fin's heat rate. Find the required length. Then estimate the heat rate if the base is held 50 K above ambient through a contact resistance of 1 K/W.

▼ **Show the inverse calculation**

Set $\tanh(mL) = 0.95$, giving $mL = \frac{1}{2}\ln(1.95/0.05) = 1.832$ and $L = 0.156$ m. The root conductance there is $G = kAm(0.95) = 0.0891$ W/K. Adding the contact resistance gives

$$Q = 50 / (1/G + 1) = 4.09\,\mathrm{W}.$$

The temperature drop across the contact is about 4.09 K. This check distinguishes the fin-root temperature from the base temperature reported by a sensor under the bond.

Further reading · Lienhard and Lienhard · A Heat Transfer Textbook, chapter 4 on conduction and fins

Chapter 16 · A moving boundary in a freezing layer

A flat cold surface grows a layer of ice from water initially at its melting temperature. Estimate when it will reach a specified thickness, then discover how a finite heat-transfer resistance changes its early growth.

Write a balance at a boundary whose position is unknown

Let the solid occupy $0 < x < s(t)$. The boundary at x=0 is held at $T_w < T_m$, the liquid remains at T_m, and the solid-liquid interface is at its equilibrium melting temperature. Treat density, latent heat of fusion L_f and solid conductivity k as constants. Ignore volume change and fluid convection. When the front moves by ds, each unit area releases latent energy $\rho L_f \, ds$. In this one-phase problem conduction into the cold side removes that energy. Taking the magnitude of heat extracted, the interface condition reads

$$\rho L_f \frac{ds}{dt} = k \left. \frac{\partial T}{\partial x} \right|_{s^-}.$$

The derivative is positive because temperature increases towards the liquid. Solving the full problem requires time-dependent conduction in a domain with a moving outer end. A useful approximation emerges when the sensible heat needed to cool the new solid is small beside its latent heat. Measure that ratio with the Stefan number $\mathrm{Ste} = c_p(T_m - T_w)/L_f$.

Derive the square-root law and check the approximation

For small Ste the temperature profile in the solid adjusts quickly compared with front growth. Approximate the gradient by $(T_m - T_w)/s$. Multiplying the interface balance by s separates the variables. Integrating from an initial layer s_0 gives

$$s \, ds = \frac{k(T_m - T_w)}{\rho L_f} \, dt,$$

$$s^2 = s_0^2 + \frac{2k(T_m - T_w)}{\rho L_f} t.$$

With zero initial thickness the predicted speed starts unbounded because the model inserts a finite temperature difference across zero resistance. A finite surface heat-transfer coefficient removes this singularity. If coolant at T_c reaches the wall through a film $1/h$, the heat flux is $(T_m - T_c)/(1/h + s/k)$. For a layer growing from zero, integration gives

$$t = \frac{\rho L_f}{T_m - T_c} \left(\frac{s}{h} + \frac{s^2}{2k} \right).$$

The linear term dominates while the film presents the larger resistance. Once the solid is thick, the quadratic term recovers diffusion-controlled growth. These two regimes can be distinguished by recording thickness at equal time intervals.

An ideal wall and a finite film

For water freezing at 0 °C take $k = 2.2$ W/(m K), $\rho = 917$ kg/m³, $L_f = 334$ kJ/kg and $c_p = 2.1$ kJ/(kg K). With a wall at −5 °C the Stefan number is 0.0314. The ideal-wall time to form 10 mm is $917(334000)(0.01)^2/[2(2.2)(5)] = 1392$ s, about 23 minutes. Reaching 20 mm requires four times that duration. The increasing conduction distance accounts for the slowing front.

With coolant at −5 °C and $h = 200$ W/(m² K), the added time for 10 mm is 3063 s. The resulting total is about 74.2 minutes. At that point the film resistance is 0.005 m² K/W and the ice contributes 0.00455 m² K/W. Both terms are of the same order. Vary h and find when a second thickness measurement would provide more information about ice conductivity than a first one.

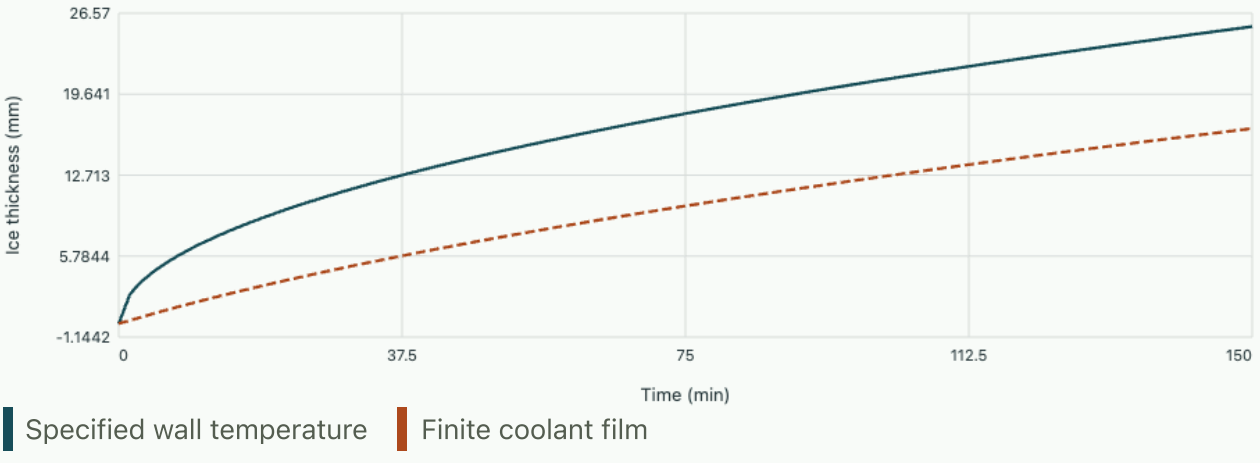
At a fixed coolant temperature compare the early and late parts of the thickness histories. Increase the film coefficient strongly. Predict where the finite-film curve should approach the fixed-wall result first in relative terms, then read the completion time.

Open this experiment online

Try the calculation

Coolant depression below freezing (K) **5**

Film coefficient (W/m² K) **200**



At the selected film coefficient 10 mm requires 74.2 minutes. The fixed-wall prediction is 23.2 minutes.

Tutorial · recover the coolant coefficient from one observation

A 15 mm layer is observed after 8000 s with the same properties and coolant temperature. Estimate h from the integrated equation. Would a time below the ideal-wall value be compatible with a positive h in this model?

▼ **Show the resistance fit**

The ideal-wall contribution is about 3132 s. The remaining 4868 s equals $\rho L_f s/[h(T_m - T_c)]$, giving $h \simeq 189$ W/(m² K). Both resistances are positive. A shorter time than the ideal-wall prediction would require checking the temperature, initial ice thickness, geometry or material data before assigning a heat-transfer coefficient.

Further reading · RWTH Aachen · Computational Modeling of Multiphysics, chapter 12 on the Stefan problem

Study the full derivation

How a gravity-drained condensate film limits heat transfer · Thermal processes and transport in engineering

Saturated vapour condenses on a cool vertical plate. A liquid film thickens as it drains, adding thermal resistance toward the bottom. Derive the local heat-transfer coefficient from coupled momentum and energy balances, then determine which observations would invalidate the smooth-film model.

Chapter 17 · Diffusion with a stationary carrier gas

A solvent evaporates into a gas layer. Its molecules diffuse relative to the mixture, while evaporation also induces a bulk molar flow. Work through the species balances to see why the flux ceases to scale linearly with surface concentration.

Choose the reference frame for a flux

Let A be the evaporating species and B an insoluble carrier. A planar film of thickness L lies between the evaporating surface, at mole fraction y_s, and a well-mixed reservoir at $y_\infty < y_s$. Keep total pressure, temperature, total molar concentration $c = p/(RT)$ and binary diffusivity D uniform. In a stationary laboratory frame the total species flux equals diffusion relative to the molar-average velocity plus transport by that velocity. If positive z points towards the reservoir, write

$$N_A = -cD\frac{dy_A}{dz} + y_A(N_A + N_B).$$

Steady conditions in one dimension make N_A, N_B independent of z. Carrier B cannot enter the liquid and has zero net flux there. This fixes $N_B = 0$ throughout the film. B still diffuses towards the interface relative to the moving mixture. Its diffusive flux balances its outward convective flux. The constraint $N_A + N_B = 0$ would instead describe equimolar counterdiffusion between two exchanging species.

Separate variables and enforce both boundary concentrations

Collect the terms in N_A, divide by $1 - y_A$ and integrate from the surface to the reservoir. The remaining integral is a logarithm of the carrier mole fraction.

$$N_A(1 - y_A) = -cD\frac{dy_A}{dz},$$

$$N_A = \frac{cD}{L}\ln\frac{1 - y_\infty}{1 - y_s}.$$

$$1 - y_A(z) = (1 - y_s)\left(\frac{1 - y_\infty}{1 - y_s}\right)^{z/L}.$$

The concentration profile curves because a given net flux requires a different concentration gradient at different carrier fractions. A dilute expansion of the logarithm gives

$$N_A \simeq (cD/L)(y_s - y_\infty).$$

The exact-to-dilute flux ratio exceeds unity when the surface emits A. The discrepancy becomes measurable long before the carrier disappears. At fractions very close to unity, the assumptions of a stagnant, isothermal film also deserve new scrutiny.

Small and large amounts of vapour at the surface

At 300 K and 1 bar, $c = 40.09$ mol/m³. Take $D = 2.0 \times 10^{-5}$ m²/s and $L = 5$ mm. The coefficient $cD/L = 0.1604$ mol/(m² s). With a vapour-free reservoir and $y_s = 0.05$, the logarithm is 0.05129 and the predicted flux is 0.00823 mol/(m² s). A linear gradient would give 0.00802, about 2.5% lower.

For $y_s = 0.50$, the flux rises to 0.111 mol/(m² s), versus 0.0802 for the dilute estimate. At mid-depth the exact fraction is $1 - \sqrt{0.5} = 0.293$, while linear interpolation gives 0.25. An experimental check can therefore use either a mass-loss measurement or a local concentration measurement. Evaporative cooling changes the saturation mole fraction at the surface, so a measured surface temperature matters even if room temperature seems stable.

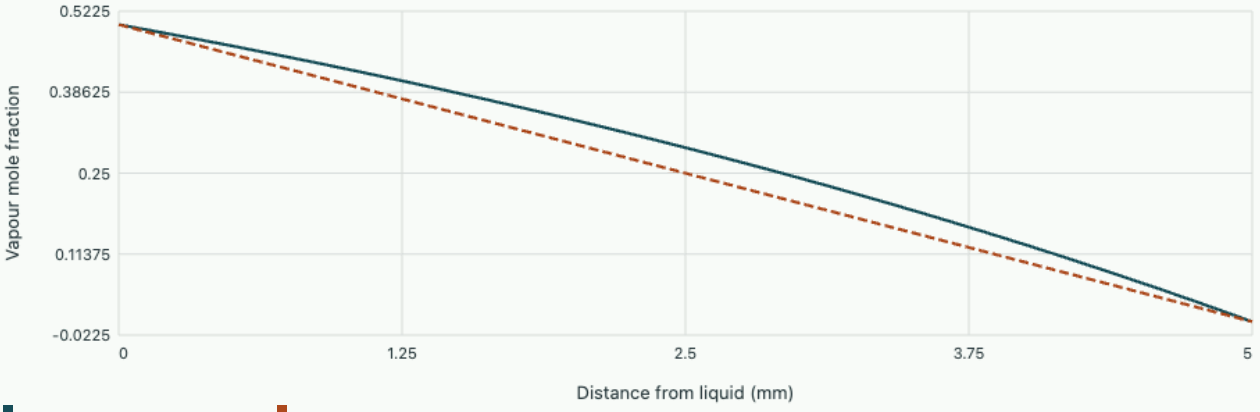
Use a vapour-free reservoir and increase the surface fraction from 0.10 to 0.80. Estimate which midpoint concentration lies farther from linear interpolation. Compare the resulting flux errors and identify the physical transport term responsible.

Open this experiment online

Try the calculation

Surface vapour mole fraction **0.5**

Reservoir vapour mole fraction **0**



The evaporation flux is 0.1112 mol/(m² s). The dilute estimate is 0.0802 mol/(m² s).

Tutorial · recover the effective film thickness

At the same temperature and diffusivity the surface mole fraction is 0.30, the reservoir fraction is 0.05 and a balance reports $N_A = 0.050$ mol/(m² s). Estimate L. What film thickness would be inferred if the outward molar flow were ignored?

▼ **Show the two fits**

Solving the exact expression gives

$$L = cD \ln(0.95/0.70)/N_A = 4.90\,\mathrm{mm}.$$

The linear fit yields

$$L = cD(0.25)/N_A = 4.01\,\mathrm{mm}.$$

The latter fit makes the film thinner to compensate for omitted advection. In an experiment that effective thickness would predict the wrong flux when the surface composition changes appreciably.

Further reading · Lienhard and Lienhard · A Heat Transfer Textbook, chapter 11 on mass transfer by diffusion

Chapter 18 · Reading a vapour-compression refrigerator from its four enthalpies

A cooling machine circulates refrigerant through a compressor, condenser, valve and evaporator. Follow one kilogram around the circuit and see why compressor efficiency and liquid subcooling change the coefficient of performance by different mechanisms.

Write separate steady-flow balances on each component

State 1 is low-pressure vapour at compressor inlet. State 2 is high-pressure vapour at the discharge. At 3 the condenser delivers high-pressure liquid, which throttles to the evaporator pressure at 4. Neglect velocity and height changes and assume insulated compressor and valve. For the valve the flow energy balance gives

$$h_4 = h_3.$$

For the other devices write positive magnitudes for useful cooling, input work and rejected heat:

$$q_L = h_1 - h_4,$$
$$w_c = h_2 - h_1,$$
$$q_H = h_2 - h_3,$$
$$\mathrm{COP}_R = \frac{q_L}{w_c}.$$

Summing the component balances gives

$$q_H = q_L + w_c.$$

The COP may exceed one because a refrigerator transports heat using work. The maximum for two specified external reservoir temperatures is $T_L/(T_H - T_L)$ in kelvins. No particular refrigerant's enthalpy table is implied in the following illustrative property data. In a design calculation determine the properties at consistent pressure, temperature or quality from one validated refrigerant source.

Get the compressor exit and valve exit by different paths

At specified discharge pressure define h_{2s} by the same entropy as state 1. The isentropic efficiency of the compressor is $\eta_c = (h_{2s} - h_1)/(h_2 - h_1)$. Rearranging gives

$$h_2 = h_1 + (h_{2s} - h_1)/\eta_c.$$

At the valve exit, if $h_f < h_4 < h_g$ at evaporator pressure, the vapour mass fraction follows from $h_4 = (1 - x_4)h_f + x_4 h_g$. Entropy increases through the valve even though enthalpy is constant. Assigning a constant entropy there would alter both flash quality and refrigeration effect.

$$x_4 = \frac{h_3 - h_f}{h_g - h_f},$$

$$\dot{m} = \frac{\dot{Q}_L}{h_1 - h_3},$$

$$\dot{W}_c = \dot{m}\,\frac{h_{2s} - h_1}{\eta_c}.$$

Keeping evaporation and condensation pressures fixed isolates the role of each state change. In an installed system the pressures, compressor map, flow and heat exchangers interact. A subcooler with a separate electrical load must also be included when evaluating a whole-system COP.

Evaluate a baseline and a subcooled variant

Take $h_f = 200,\ h_g = h_1 = 400,\ h_{2s} = 430,\ h_3 = 250$ kJ/kg with $\eta_c = 0.75$. The actual outlet has $h_2 = 440$ kJ/kg, cooling per kilogram is 150 kJ/kg and compressor input is 40 kJ/kg. The evaporator inlet is 25% vapour by mass, and COP is 3.75. At 6.0 kW cooling demand the flow is 0.0400 kg/s, compressor input 1.60 kW and rejected heat 7.60 kW.

At the same pressure levels and compressor inlet state, lower h_3 by 10 kJ/kg before throttling. Valve-exit quality becomes $40/200 = 0.20$, and 160 kJ/kg can be absorbed in the evaporator. COP becomes 4.00. For the specified cooling duty the flow falls to 0.0375 kg/s and compressor power to 1.50 kW. By comparison raising compressor efficiency from 0.75 to 0.85 with the original condenser exit reduces the work per kilogram to 35.29 kJ/kg and gives COP 4.25 at the original mass flow. One change acts on refrigeration effect and the other on compression work.

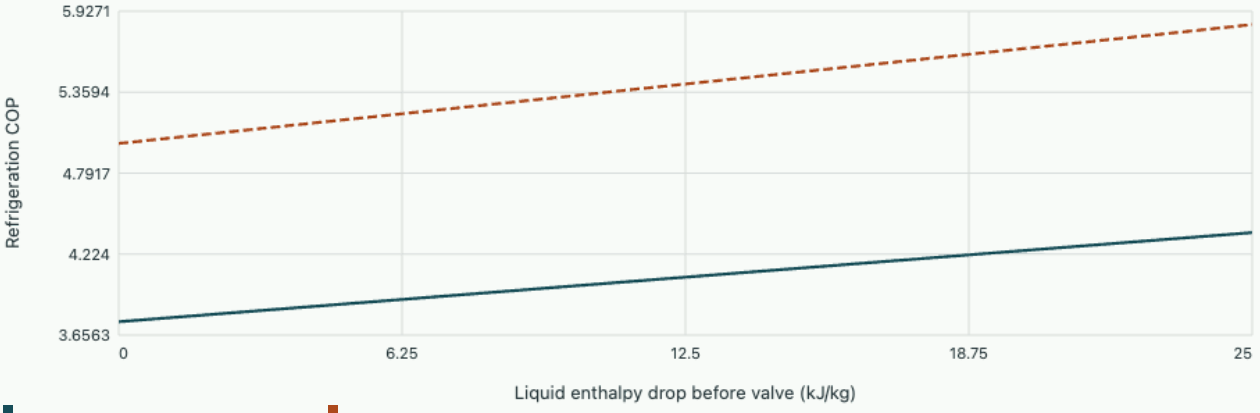
At unchanged compressor efficiency reduce the liquid enthalpy by 20 kJ/kg. Predict how much flash vapour is avoided for each kilogram passing the valve. Then reduce compressor efficiency and determine which of the two changes affects inlet quality.

Open this experiment online

Try the calculation

Compressor isentropic efficiency **0.75**

Selected liquid enthalpy drop (kJ/kg) **10**



At the selected 10 kJ/kg subcooling enthalpy drop, COP is 4.00, valve-exit vapour quality is 0.200 and a 6 kW duty takes 1.500 kW shaft power.

Tutorial · diagnose a measured power excess

A rig with the original inlet and condenser enthalpies delivers 5.4 kW cooling while the measured refrigerant-side shaft input is 1.8 kW. Infer mass flow, actual outlet enthalpy and compressor isentropic efficiency. An electrical measurement gives 2.0 kW. What additional efficiency does that measurement quantify?

▼ **Close the component boundaries**

The evaporator gives $\dot{m} = 5.4/150 = 0.0360$ kg/s. The shaft input gives $w_c = 1.8/0.0360 = 50.0$ kJ/kg and $h_2 = 450$ kJ/kg. Hence

$$\eta_c = 30/50 = 0.60.$$

With 2.0 kW electrical input, the upstream motor/drive delivers $1.8/2.0 = 0.90$ of that power to the shaft. Shaft-based COP is 3.0, electrical COP is 2.7. Comparing measured COP values requires identifying which boundary each power meter encloses.

Further reading · University of Waterloo, ME 354 · vapour-compression refrigeration cycles

Study the full derivation

Staged compression with an imperfect intercooler · Thermal processes and transport in engineering

A laboratory compressed-air supply has two compressor stages and an intercooler. Equal pressure ratios minimise work only under particular assumptions. Derive the optimum, compare it with a single-stage machine and identify which temperature and power measurements reveal imperfect cooling or unequal efficiency.

Chapter 19 · Finding the vapour fraction of an equilibrium flash

A mixture enters a separator at given overall composition, pressure and temperature. Some components prefer the vapour more strongly than others. Derive a scalar equation for the fraction that evaporates and test for a two-phase solution before iterating.

Connect composition to the equilibrium ratios

Let $F = L + V$ be the total molar feed and liquid and vapour flow rates. Overall component fractions are z_i, liquid fractions x_i and vapour fractions y_i. At equilibrium the chemical potential of each species agrees across phases. In an ideal gas over an ideal liquid this reduces to Raoult's law $y_i p = x_i p_i^{\mathrm{sat}}(T)$, so $K_i = y_i/x_i = p_i^{\mathrm{sat}}(T)/p$. A larger K enriches that species in the vapour. For a real mixture K can depend on phase composition and require an outer property iteration. Here use fixed positive K values to isolate the material balance.

Eliminate the component flows

Define $\beta = V/F$. Component conservation gives

$$z_i = (1-\beta)x_i + \beta y_i = x_i[1 + \beta(K_i - 1)].$$

Thus $x_i = z_i/[1 + \beta(K_i - 1)]$. For both compositions to sum to one, enforce $\sum_i (y_i - x_i) = 0$, giving the Rachford–Rice equation:

$$R(\beta) = \sum_i \frac{z_i(K_i - 1)}{1 + \beta(K_i - 1)} = 0,$$

$$R'(\beta) = -\sum_i \frac{z_i(K_i - 1)^2}{[1 + \beta(K_i - 1)]^2}.$$

The denominators are positive for $0 \le \beta \le 1$. With at least one non-unity K of positive feed fraction, R decreases strictly, so at most one physical root exists. A positive R(0) and a negative R(1) bracket two phases. If $R(0) \le 0$ the stable fixed-K split is all liquid; if $R(1) \ge 0$ it is all vapour, with equality corresponding to incipient phase appearance. When every K is one this balance cannot specify β. A bracketed numerical method preserves the meaningful interval even near a phase boundary.

Perform a binary flash and change the feed composition

Choose $K_A = 2,\ K_B = 0.5,\ z_A = z_B = 0.5$. Then

$$R(0) = 0.25,\ R(1) = -0.25.$$

Multiplying by the two denominators gives $0.25 - 0.50\beta = 0$, hence $\beta = 0.50$. Liquid fractions are $x_A = 1/3,\ x_B = 2/3$, and vapour fractions are

$$y_A = 2/3,\ y_B = 1/3.$$

For 10 mol/s feed each outlet carries 5 mol/s. The A balance reads $5(1/3) + 5(2/3) = 5$ mol/s.

Increase z_A to 0.60 at the same T, p and K. The physical root becomes $\beta = 0.80$. The phase compositions remain at the same two endpoints; their proportions change to accommodate the new feed. Indeed the overall A fraction is

$$0.20(1/3) + 0.80(2/3) = 0.60.$$

At $z_A = 0.75$ the feed lies beyond the vapour endpoint. Then R(1) is positive, and searching outside the interval would produce a vapour fraction above unity. Report a single vapour phase instead. An adiabatic flash at unknown temperature would require a further enthalpy balance.

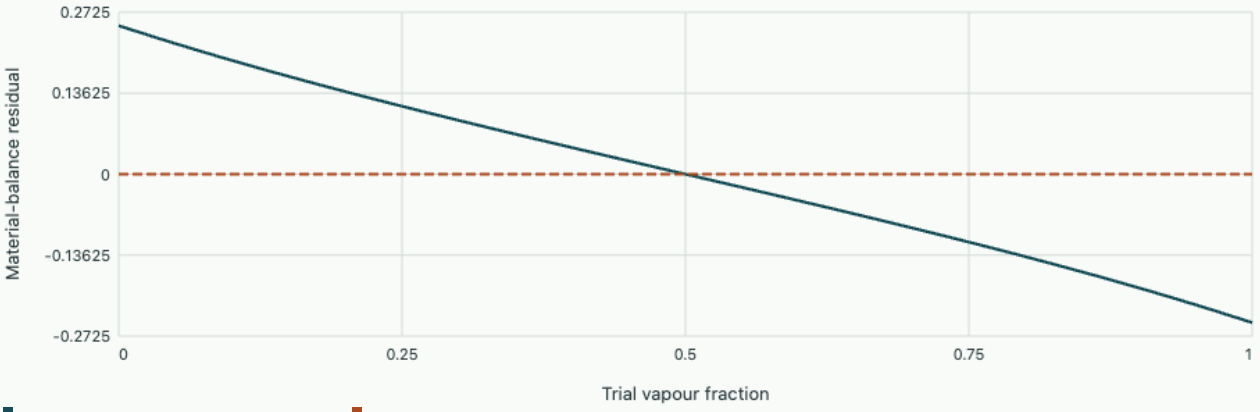
Start from equal feed fractions and increase the fraction of A until the root leaves the physical interval. Predict which pure-phase endpoint is reached. Compare residual signs at both endpoints before interpreting the reported phase.

Open this experiment online

Try the calculation

Feed fraction of A **0.5**

Equilibrium ratio K of A **2**



Bracketed balance gives vapour fraction 0.5000. A fractions in the liquid and vapour are 0.3333 and 0.6667.

Tutorial · determine whether an iteration is warranted

At an operating point take $K_A = 3,\ K_B = 0.6$. Compute the liquid and vapour A fractions when both phases coexist. For $z_A = 0.30$, calculate β. Repeat the phase test for $z_A = 0.10$ and explain why normalized hypothetical phase compositions can hide a failed balance.

▼ **Solve phase closure and the lever rule**

Vapour closure gives $3x_A + 0.6(1 - x_A) = 1$, hence $x_A = 1/6$, $y_A = 1/2$. For overall fraction 0.30,

$$\beta = (0.30 - 1/6)/(1/2 - 1/6) = 0.40.$$

At $z_A = 0.10$,

$$R(0) = 0.10(2) + 0.90(-0.4) = -0.16.$$

The feed remains liquid. Renormalizing an unphysical trial vapour composition could make it sum to unity while violating the original component and equilibrium equations. Check the physical endpoint signs first.

Further reading · Penn State, PNG 520 · elementary vapour–liquid equilibrium, flash objective function and numerical solution

Chapter 20 · Where a thermoelectric cooler rejects its heat

A detector has to operate ten kelvin below the laboratory air temperature. A thermoelectric module beneath it can transport heat against that temperature difference when an electric current flows. Increasing the current also generates heat inside the module and warms its heat sink. Build an energy balance that predicts when additional current helps, when it costs more than it delivers and why a nominal cooling-capacity figure alone cannot specify an operating point.

Identify the two faces and the power supply

Treat the detector plate as an isothermal cold face at T_c and the face in contact with the heat sink as an isothermal hot face at T_h. The laboratory air is at temperature T_a. All three temperatures enter the calculation in kelvin. Temperature differences in kelvin and degrees Celsius have the same numerical value, but the Peltier term contains an *absolute* temperature. A calculation using a cold-side value of 15 instead of 288 would suppress its principal heat-transport mechanism.

Define Q_c as the steady heat absorbed from the cold plate and Q_h as the steady heat given to the sink, both in watts. The electrical power entering the module is IV. The complete module has no net energy storage in steady state, hence $Q_h = Q_c + IV$. For a detector dissipating 15 W, its refrigeration requirement may be larger than this dissipation if wiring or insulation admits additional environmental heat. The load balance belongs to the detector, while the module characteristic supplies the deterministic prediction of what it can remove at specified temperatures.

Use effective parameters for the assembled device: Seebeck coefficient α in V/K, electrical resistance R in ohms and thermal conductance K in W/K. Current flows through the series electrical elements, while the thermoelectric legs supply largely parallel paths for heat conduction. An effective module coefficient cannot be silently replaced by the coefficient of a single leg. Manufacturer data can vary with temperature, mechanical mounting and electrical contact. The constant-property approximation used here permits an exact algebraic check before introducing fitted temperature-dependent coefficients.

Separate transported heat, back conduction and ohmic heating

A reversible thermoelectric contact transports heat ΠI where its Peltier coefficient is

$$\Pi = \alpha T.$$

This relation follows from the Kelvin connection between the reversible Peltier and Seebeck effects. With the sign of current chosen for cooling, the cold junction removes $\alpha T_c I$ from the plate. Heat conducts down the temperature gradient from the hot face toward the cold one at a rate $K(T_h - T_c)$, reducing the net amount the module can absorb from the detector.

Resistance generates $I^2 R$ of heat throughout the legs. The familiar one-half allocation can be derived rather than treated as bookkeeping by convention. In a homogeneous leg of constant thermal conductivity and equal cross section, a uniform volumetric heat source produces a parabolic correction to the imposed linear temperature profile. The source-driven correction vanishes at both isothermal ends and its slopes there have equal and opposite signs. Each face therefore receives half of the internally

generated heat. A different material arrangement or strongly temperature-dependent coefficients can change that split.

$$Q_c = \alpha T_c I - \tfrac{1}{2} I^2 R - K(T_h - T_c),$$
$$Q_h = \alpha T_h I + \tfrac{1}{2} I^2 R - K(T_h - T_c).$$

Subtracting the two equations gives

$$Q_h - Q_c = I[\alpha(T_h - T_c) + IR].$$

Thus the electrical terminal voltage compatible with these heat flows is $V = IR + \alpha(T_h - T_c)$. The second term is the back electromotive force associated with the imposed temperature difference. Omitting it underestimates the input power and inflates the reported coefficient of performance. As a sign check, set $I = 0$. Heat still leaks through a module with $T_h > T_c$; Q_c is then negative.

Open editable notebook Download .ipynb

Close the hot-side balance and solve for temperature

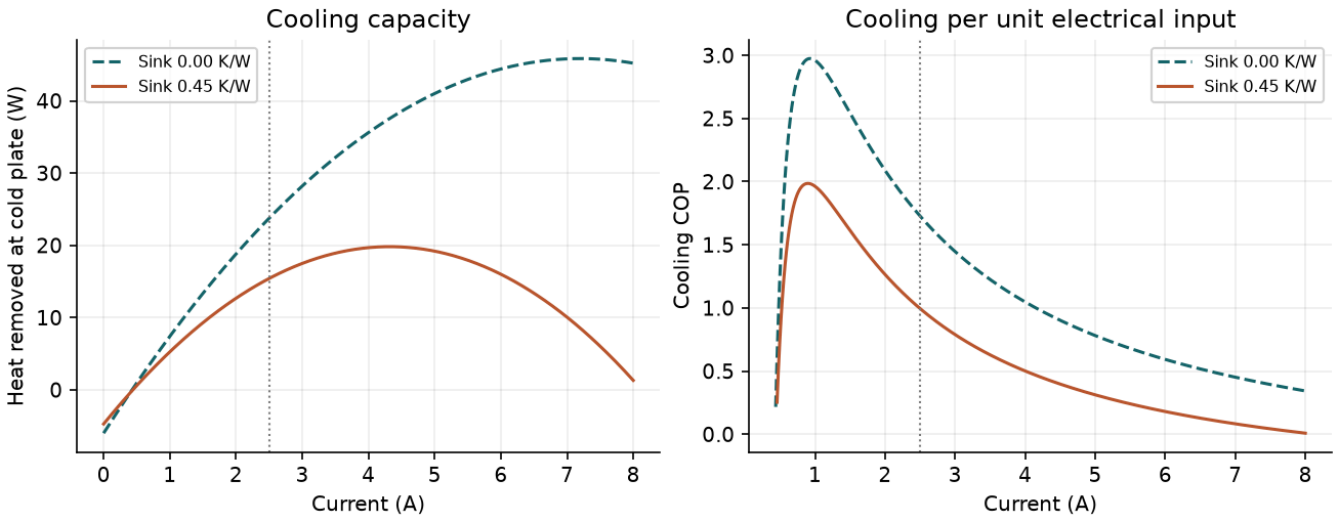
A heat sink with thermal resistance R_s must satisfy

$$T_h = T_a + R_s Q_h.$$

Insert the hot-face expression above, collect every occurrence of T_h on the left, and solve before evaluating the cold-face capacity. The result is

$$T_h = \frac{T_a + R_s(K T_c + I^2 R/2)}{1 + R_s(K - \alpha I)}.$$

When $R_s = 0$, the hot junction stays at ambient temperature. For finite R_s it responds to the rejected heat, which includes the electrical input. A steady-state candidate also requires a positive stabilising slope in the hot-face balance. With a finite hot-side heat capacity the linear coefficient of a temperature perturbation is proportional to $\alpha I - K - 1/R_s$. The denominator in the formula is positive precisely when this perturbation decays, for $R_s > 0$. Extremely large currents can invalidate constant material properties well before a formal singularity is reached.



A finite sink moves the largest cooling capacity to a lower current. The operating point that gives the greatest capacity differs from the one that delivers the most cooling per watt consumed.

Interactive calculation · Cooler capacity

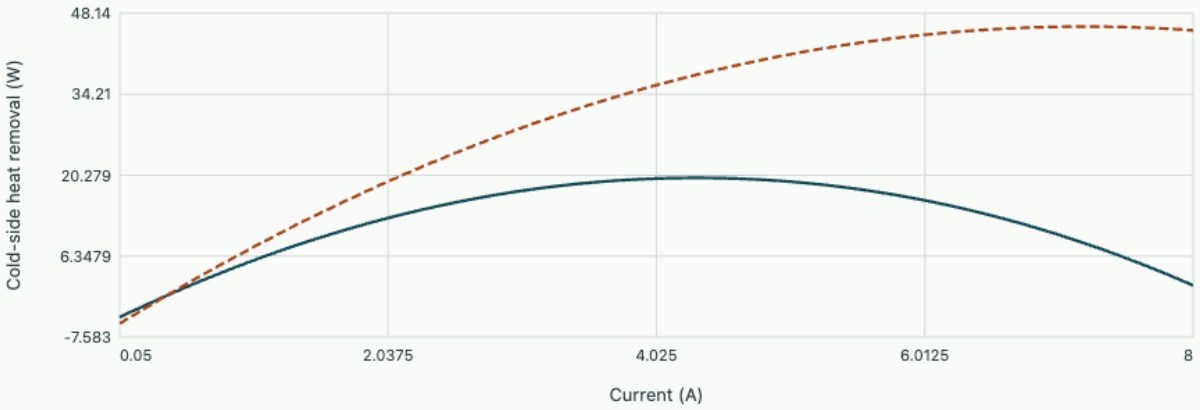
Change the physical and numerical inputs of the notebook model. Compare the updated curves and balance checks with the recorded notebook figures. Each plot uses the same equations and assumptions documented in that notebook; its readout names any analytical or reduced-model reference.

Open this live calculation

Try the calculation

Sink resistance (K/W) **0.45**

Cold-face temperature (K) **288**



Selected heat sink Ideal ambient-temperature sink

At2.5 A, hot face 311.90 K and capacity 15.41 W. COP is shown only where cooling and electrical input are positive. Constant module properties as in the notebook.

Download this example's editable notebook

Interactive calculation · Cooler cop

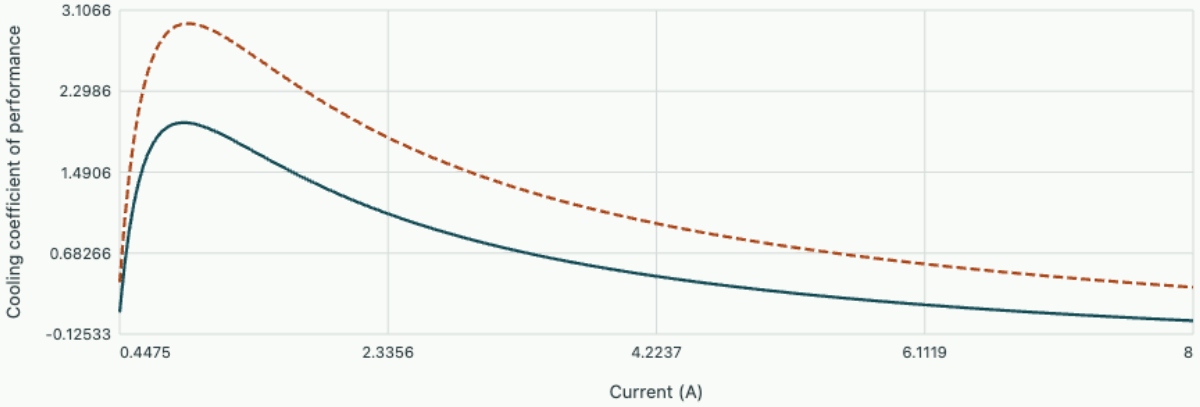
Change the physical and numerical inputs of the notebook model. Compare the updated curves and balance checks with the recorded notebook figures. Each plot uses the same equations and assumptions documented in that notebook; its readout names any analytical or reduced-model reference.

Open this live calculation

Try the calculation

Sink resistance (K/W) **0.45**

Cold-face temperature (K) **288**



Selected heat sink Ideal ambient-temperature sink

At2.5 A, hot face 311.90 K and capacity 15.41 W. COP is shown only where cooling and electrical input are positive. Constant module properties as in the notebook.

Download this example's editable notebook

For $T_c = 288$ K, $T_a = 298$ K, $\alpha = 0.050$ V/K, $R = 2.0$ ohms, $K = 0.60$ W/K, $R_s = 0.45$ K/W and $I = 2.5$ A, the hot face settles at 311.90 K. The cold-side terms are 36.0 W of transported heat, 6.25 W of returning ohmic heat and 14.34 W of thermal leakage. Their net is 15.41 W. The required voltage is approximately 6.195 V. The sink has to dissipate about 30.90 W, consistent with its 13.90 K temperature rise above ambient. Check both balances independently before using a single net wattage in a larger thermal model.

Find the feasible current interval for a prescribed load

With a perfect sink, hold T_c and T_h fixed. The capacity is a concave quadratic in I. Differentiation gives the peak at $I = \alpha T_c / R$, or 7.2 A for the example. At this point the cold-junction transported heat still grows with current, but its marginal gain is exactly cancelled by the marginal ohmic loss. If a positive

target is below the maximum available capacity, the same target may be reached at a low-current and a high-current operating point. At fixed face temperatures the low-current point uses less electrical power.

For a finite sink there is an additional derivative $-K\,dT_h/dI$ in dQ_c/dI. The hot junction becomes warmer as the supply current rises in the operating band. As a result the capacity curve in the notebook reaches its maximum close to 4.32 A. The first current on its finite sweep that meets a 15 W load is about 2.43 A. Refining the current grid or solving $Q_c(I) - 15 = 0$ gives a sharper estimate. A point chosen just at the load threshold has essentially no margin for an unexpectedly warm laboratory.

The refrigeration coefficient of performance is $Q_c/(IV)$ only where $Q_c > 0$ and $IV > 0$. At 2.5 A it is about 0.995. The ideal reversible upper bound for heat pumping between the actual two face temperatures is $T_c/(T_h - T_c) \simeq 12.1$. Using room air in place of T_h yields another bound for a differently defined system whose losses include the heat exchanger. Neither comparison supplies a substitute for the heat-flow balances. They independently warn about an implausibly optimistic calculated performance.

Estimate the sink resistance from an operating test

Imagine that the laboratory supply measures 2.50 A and 6.20 V while two thermometers indicate 311.9 K at the hot face and 298.0 K in the entering air. A calibrated electrical heater on the cold plate supplies 15.4 W. With negligible other heat leaks the module lifts the heater power. The sink receives $15.4 + 2.50(6.20) = 30.9$ W. Dividing the measured temperature difference by this load gives an effective sink resistance of 0.450 K/W. A measurement of the sink temperature alone would be insufficient for this estimate. The electrical input must be measured at the device terminals, and the two thermal measurements must refer to the endpoints of the resistance being inferred.

Consider instead that an unmeasured heat leak of 2 W also enters the cold plate. The true hot-side load rises to 32.9 W at the stated electrical readings. Assigning the full temperature rise to only 30.9 W overestimates the thermal resistance by a factor $32.9/30.9$, or about 6.5%. Repeating the measurement with a second heater power at the same cold temperature can help separate an approximately constant parasitic load from a change in resistance. Plot hot-face rise against heater power plus measured electrical power for both points. The slope estimates the resistance and an intercept can reveal the omitted heat load. Changes in fan speed or ambient temperature during the test would add another explanation for that intercept.

The first feasible current in a tabulated sweep depends on the grid spacing. The notebook refines the lower crossing with OneSimulation's bracketed inverse in a local installation. It chooses two adjacent currents on the increasing branch. A bracket spanning the entire current interval could include a second solution on the decreasing branch.

Exercises on operating point and experimental interpretation

1. Repeat the 2.5 A calculation with the hot face clamped to room temperature. How much cooling does an ideal heat sink recover?

▼ **Work through the solution**

With $T_h - T_c = 10$ K the cold capacity is

$$36.0 - 6.25 - 6.0 = 23.75\,\mathrm{W}.$$

Compared with 15.41 W, the ideal sink recovers about 8.34 W by reducing back conduction. Its terminal voltage is also smaller, $5 + 0.05(10) = 5.50$ V, since the Seebeck contribution falls with the temperature difference.

---

2. A technician sizes the heat sink for 15.4 W because that is the measured cooling capacity. What temperature prediction results and what balance does it violate?

▼ **Work through the solution**

The assumed temperature rise is $0.45(15.4) = 6.93$ K, about half the actual rise. The electrical supply adds approximately $2.5(6.195) = 15.49$ W. The correct hot-face load includes both contributions. A cooled detector can warm its surrounding air by more heat than the detector generates because the power supply is part of the complete energy balance.

---

3. Which measurement can disentangle electrical resistance from the Seebeck coefficient when both are uncertain?

▼ **Work through the solution**

At a fixed measured difference between face temperatures, sweep small currents and measure terminal voltage. The slope gives R and the zero-current voltage gives $\alpha(T_h - T_c)$, provided the face temperatures remain controlled and the leads contribute negligible voltage offsets. Reverse the temperature difference or calibrate the leads to test that offset. A current sweep without measuring both face temperatures mixes the slope of the ohmic term with the change in hot-face temperature.

Check the module definitions against an engineering reference

The Ferrotec reference on mathematical modelling of modules gives performance equations and temperature-dependent parameter descriptions. When moving from this example to a particular module, check the property definitions and the temperatures at which the supplier evaluated them.

Problem sheet 7 · Budgeting cooling power for a freezing experiment

A cold plate grows a planar ice layer while a metal support fin conducts heat into the cold enclosure. Construct an energy budget that distinguishes the rate at which the front moves, the total latent heat extracted, and the electrical input of the refrigerator.

1 · Predict the completion time and distinguish two resistances

Water starts at its melting point. Use $\rho = 917$ kg/m³, latent heat 334 kJ/kg and solid conductivity 2.2 W/(m K). Coolant is 5 K below the melting point with film coefficient 200 W/(m² K). Find the time to grow 15 mm on 0.0100 m². Use the small-Stefan-number model and explain why doubling area does not change this time.

▼ **Integrate the local flux balance**

The heat flux magnitude is $5/(1/h + s/k)$. Equating it to $\rho L_f \dot{s}$ gives

$$t = (\rho L_f/5)[s/h + s^2/(2k)].$$

Film and solid contributions are about 4594 and 3132 s, respectively. Total time is 7727 s, or 2.15 h. Both extracted heat and interface mass scale with area, so area cancels from the local balance for a planar front. The plate and coolant must deliver the corresponding total power over the chosen footprint.

Check the finite-film and fixed-wall growth curves at 15 mm

2 · Convert front position into a heat-load history

Determine latent heat extracted and mean latent heat rate over the run. Calculate initial and final heat rates and explain why the simple average of the endpoint rates differs from the integral mean.

▼ **Close the integral and instantaneous budgets**

Ice mass is $917(0.0100)(0.015) = 0.13755$ kg, hence $Q = 45.94$ kJ and mean rate 5.95 W. Initially the finite film limits flux to $h\Delta T = 1000$ W/m² or 10.0 W. At 15 mm it falls to $5/[0.005 + 0.015/2.2] = 423$ W/m² or 4.23 W. The average of these endpoints is 7.12 W. Heat rate varies nonlinearly with time and spends longer near its lower values. Integrating through the latent energy gives the correct time average.

---

3 · Include a parasitic fin connected to a warm mount

An aluminium support strip 20 mm wide, 2 mm thick and 50 mm long has a base 25 K warmer than the cold enclosure. Take $k = 200$ W/(m K), $h = 25$ W/(m² K), all four long faces exposed, insulated tip. Calculate the extra steady heat deposited in the enclosure and the mean electrical power at a measured refrigerator COP of 2.2.

▼ **Evaluate the distributed heat leak**

With $A = 4.0 \times 10^{-5}$ m² and $P = 0.044$ m, $m = \sqrt{hP/(kA)} = 11.73$ m⁻¹. Fin heat is $kAm(25)\tanh(mL) = 1.24$ W. During freezing the enclosure therefore requires mean heat removal $5.95 + 1.24 = 7.19$ W. Electrical power is $7.19/2.2 = 3.26$ W using unrounded numbers. The chapter's fin study uses a 50 K root excess, so its heat rate for the same geometry is twice this value by linearity. The peak electrical requirement under the constant-COP assumption is about 5.11 W at startup.

Check the support-strip heat rate and use its linear temperature scaling

4 · Test a thermodynamic bound

The refrigerator rejects heat to a room at 20 °C while its cold reservoir is at −5 °C. Compute the maximum reversible COP and its minimum possible mean work input at the preceding load. Would a claimed 0.50 W mean input be compatible with these two reservoirs?

▼ **Use absolute temperatures and both energy streams**

The reversible COP is $T_c/(T_h - T_c) = 268.15/25 = 10.73$, so even a reversible refrigerator needs about $7.18/10.73 = 0.670$ W. A reported 0.50 W would imply COP 14.4. Check whether other cooling paths, stored cold or an incorrect averaging interval entered that measurement. The warm side must reject the sum of extracted heat and electrical work, giving about 10.45 W for the measured COP.

Problem sheet 8 · Separate heat capacity, contact conductance and leakage in a calorimeter

Two small metal assemblies, each internally close to uniform temperature, are joined through a removable thermal contact. Their thermometer traces approach one another. Use the common final temperature, the transient temperature difference and the subsequent slow drift to decide which combinations of thermal properties the experiment can identify.

Separate the conserved and relaxing coordinates

Let the capacities of initially hot and cold bodies be $C_h = 120$ and $C_c = 300$ J/K. Initial temperatures are 70 and 20 degrees Celsius, respectively. Their contact conducts $H = 2.5$ W/K. Assume capacities and conductance are constant and heat losses initially negligible. Derive two independent combinations of temperature and identify which one a thermometer difference measures.

▼ **Add the balances and subtract the temperature rates**

The contact heat rate is $H(T_h - T_c)$, entering the cold body and leaving the hot body. Their balances are $C_h\dot{T}_h = -H(T_h - T_c)$ and $C_c\dot{T}_c = H(T_h - T_c)$. Add them to obtain conservation of $C_hT_h + C_cT_c$. An arbitrary Celsius reference cancels from changes in this energy; absolute temperature will be necessary if entropy is required. Subtract the equations after dividing by the respective capacities to obtain

$$\bar{T} = \frac{C_hT_h + C_cT_c}{C_h + C_c},$$

$$T_h - T_c = (T_{h0} - T_{c0})e^{-t/\tau},$$

$$\tau = \frac{C_hC_c}{H(C_h + C_c)}.$$

The equilibrium temperature is 34.286 degrees Celsius and the relaxation time 34.286 seconds. Their matching numerical values are coincidental and their dimensions differ. Changing the contact conductance changes only the relaxation time, while changing a body's capacity changes both. The arithmetic temperature mean is not conserved for unequal capacities.

Predict an intermediate reading and irreversible entropy production

Compute both thermometer readings at 40 seconds. Does energy conservation imply zero total entropy change? Use the complete final equilibration to obtain a numerical entropy check and explain why using temperatures in degrees Celsius would give a spurious result.

▼ **Reconstruct both temperatures and integrate the entropy changes**

The difference after 40 seconds is

$$50e^{-40/34.286} = 15.57\,\mathrm{K}.$$

Solving the weighted-mean and difference equations gives $T_h = \bar{T} + C_c(T_h - T_c)/(C_h + C_c) = 45.41$ degrees Celsius and $T_c = 29.84$ degrees Celsius. Substitution into the original energy balance gives equal and opposite heat transfer. The instantaneous contact heat current is about 38.9 W, below its initial value of 125 W.

At any instant, $dS_{\mathrm{total}}/dt = H(T_h - T_c)(1/T_c - 1/T_h) \geq 0$ using kelvin. Integrating each body's reversible state-function differential $C\,dT/T$ between its initial and final equilibrium temperatures gives $C_h \ln(T_f/T_{h0}) + C_c \ln(T_f/T_{c0})$, about 1.09 J/K. Temperature differences in conduction may use either Celsius or kelvin, but a reciprocal temperature or ratio in an entropy calculation requires an absolute scale. Internal irreversibility produces entropy while the isolated system conserves energy.

Recover a contact from data with one independently known capacity

Suppose the hot capacity is known and the cold capacity is to be calibrated. The equilibrium and temperature-difference measurements are as predicted. Show how to recover C_c and H, then demonstrate a family of parameter sets producing identical temperatures if neither capacity is known.

▼ **Use one temperature ratio and one time constant**

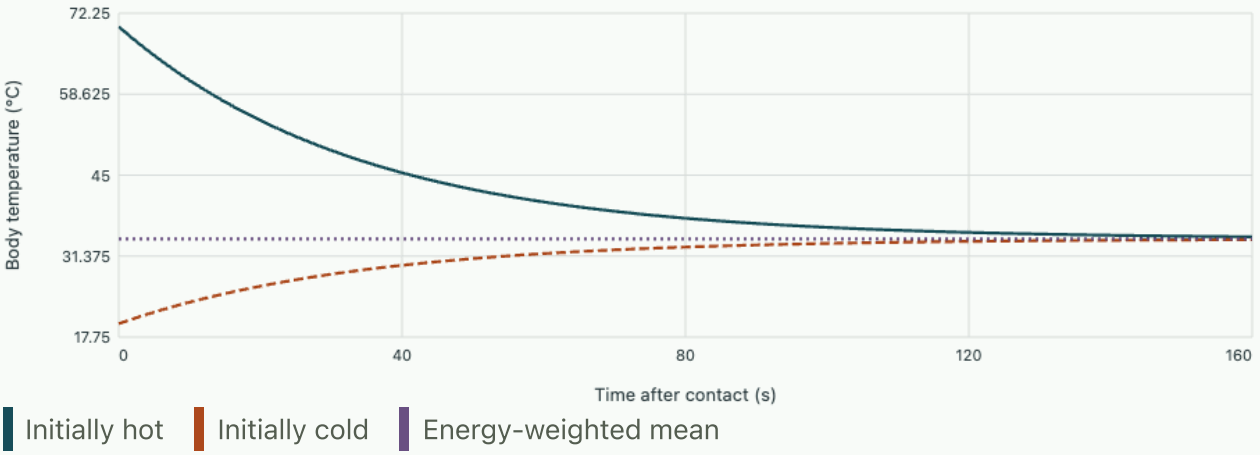
The final energy balance gives $C_c/C_h = (T_{h0} - T_f)/(T_f - T_{c0}) = 2.50$, so $C_c = 300$ J/K. The slope of a log difference against time equals $-1/\tau$; substitution gives $H = 2.50$ W/K. Fit the signed temperature difference before taking a logarithm near its noise floor. Two-channel gain or lag mismatch can make the measured difference turn negative late in the record even while actual contact heat flow has not reversed.

Multiplying both capacities and the conductance by the same positive number leaves both differential equations for temperature unchanged. No temperature trace from this isolated preparation can fix their absolute scale without additional information. A known electrical heat pulse delivered to one isolated body can determine its capacity if heat loss and thermometer lag are controlled. Changing the initial temperatures alone supplies more data but cannot remove this exact scaling degeneracy.

Open this experiment online

Try the calculation

Hot body capacity (J/K)	120
Cold body capacity (J/K)	300
Contact conductance (W/K)	2.5



The isolated common temperature is 34.29 °C. The difference decays with time constant 34.29 s; after 40 s the readings are 45.41 and 29.84 °C.

Diagnose a warm enclosure and a delayed thermometer

After the fast transient, the two nearly equal temperatures cool slowly toward an enclosure at 20 degrees Celsius, with common slope initially minus 0.010 K/s near 34.3 degrees Celsius. Estimate total external thermal conductance. Can the same trace determine which body leaks the heat?

▼ **Form the whole-system balance and design a separate perturbation**

When internal equilibration is fast, a common-temperature approximation gives

$$G_{\mathrm{loss}} \simeq -(C_h + C_c)\dot{T}/(\bar{T} - T_{\mathrm{room}}).$$

The measured slope implies about 0.294 W/K. Its external cooling time is roughly 1430 s, much longer than the contact relaxation. A slow external loss may still bias a final-temperature estimate taken very late. Fit the full two-temperature model or extrapolate the common cooling trajectory back toward the end of internal equilibration.

The common slope identifies the sum of conductances to the enclosure, not their allocation when both temperatures are equal. Heat one body separately with the link removed or swap external insulation while keeping contact unchanged. A sensor's own heat capacity and time lag can introduce another apparent time constant. Calibrating both thermometers on one body during a known step tests that

alternative. Those independent checks decide whether a second observed decay belongs to contact physics, environmental loss or the measurement itself.

Lienhard and Lienhard · A Heat Transfer Textbook, transient conduction and lumped systems

Notebook practice · Where a thermoelectric cooler rejects its heat

Thermal processes and transport in engineering

Close energy and entropy balances around real processes. Use measured states and heat flows to separate component performance from modelling assumptions.

1. ## Staged compression with an imperfect intercooler

   A laboratory compressed-air supply has two compressor stages and an intercooler. Equal pressure ratios minimise work only under particular assumptions. Derive the optimum, compare it with a single-stage machine and identify which temperature and power measurements reveal imperfect cooling or unequal efficiency.

2. ## How a gravity-drained condensate film limits heat transfer

   Saturated vapour condenses on a cool vertical plate. A liquid film thickens as it drains, adding thermal resistance toward the bottom. Derive the local heat-transfer coefficient from coupled momentum and energy balances, then determine which observations would invalidate the smooth-film model.

3. ## Radiative exchange in a grey enclosure

   A thin radiation shield does not need to absorb all incoming radiation to reduce heat transfer. Multiple reflections connect every surface to every other visible surface. Build the radiosity equations, recover the two-plate result and infer what can be learned from a measured shield temperature.

4. ## Periodic heating as a measurement of thermal diffusivity

   Two buried thermometers see the same applied surface temperature oscillation with different amplitudes and phases. Derive how a solid filters that oscillation, recover diffusivity independently from both observations and identify when a rear boundary or a poor sensor contact changes the inference.

5. ## What temperature does a probe in a hot gas measure?

   A small bead in a duct exchanges energy with both the gas and the duct walls. Calculate the indicated temperature from an energy balance, infer the gas temperature from the reading, and test how emissivity changes the measurement.

6. ## Where does the temperature of a reacting gas come from?

   A heater in an ideal cycle can supply a prescribed amount of heat. A steady combustor converts chemical enthalpy into sensible enthalpy. Compute a state from atomic balances and show where the complete-conversion model begins to overpredict its temperature.

7. ## How can a flow measurement determine a wall temperature?

   Move from a heat-transfer coefficient supplied as data to one inferred from geometry and flow. Compare a turbulent pipe, a laminar external boundary layer and a plate cooled by buoyancy.

8. Why can a heated surface suddenly become much hotter?

Trace the pool-boiling curve of a saturated liquid. Compare changing a heater's temperature with changing its electrical power and estimate the hydrodynamic scale at which nucleate boiling becomes unstable.

Thermal processes and transport in engineering · 1. Staged compression with an imperfect intercooler

A laboratory compressed-air supply has two compressor stages and an intercooler. Equal pressure ratios minimise work only under particular assumptions. Derive the optimum, compare it with a single-stage machine and identify which temperature and power measurements reveal imperfect cooling or unequal efficiency.

Work follows the actual enthalpy rise

Model dry air as an ideal gas with constant $c_p = 1.005$ kJ per kilogram kelvin and $\gamma = 1.40$. Assume steady adiabatic compressor stages, negligible changes in kinetic and potential energy, and negligible intercooler pressure drop initially. For a fixed inlet temperature T_i and pressure ratio r_i, an isentropic reference ends at $T_{os} = T_i r_i^a$, $a = (\gamma - 1)/\gamma$. Isentropic compressor efficiency compares reference enthalpy rise with actual rise, so actual shaft work per unit mass is $w_i = c_p T_i (r_i^a - 1)/\eta_i$.

An intercooler reduces the second-stage inlet temperature by rejecting heat to a colder stream. It does not undo the pressure work already delivered by the first stage. For overall ratio R and first-stage ratio r, assume the measured second-stage inlet T_2 can be held fixed across the comparison. Adding both work contributions gives

$$w(r) = c_p \left[\frac{T_1}{\eta_1} (r^a - 1) + \frac{T_2}{\eta_2} ((R/r)^a - 1) \right].$$

Direct differentiation with respect to $\log r$ balances the incremental work costs of transferring pressure ratio between stages. The stationary split is

$$r_{\mathrm{opt}} = \sqrt{R} \left(\frac{T_2 \eta_1}{T_1 \eta_2} \right)^{1/(2a)}.$$

Both exponential-power terms have positive second derivatives in $\log r$, giving a unique interior minimum when the root lies between one and R. Otherwise the constrained optimum is a boundary under this simplified model. A hot or inefficient second stage favours assigning more ratio to the first. Perfect intercooling with identical efficiencies gives $r = \sqrt{R}$, the geometric mean rule for intermediate absolute pressure.

Background · What if the two compressor stages start at different temperatures?

Consider an ideal gas with constant specific heat. Use equal isentropic efficiencies and a fixed total pressure ratio R. Prescribe inlet temperatures T_1 and T_3 independently of the intermediate pressure. With $a = (\gamma - 1)/\gamma$, the speed-independent part of shaft work is proportional to $T_1(r_1^a - 1) + T_3[(R/r_1)^a - 1]$. Differentiate with respect to $\ln r_1$ and set the derivative to zero.

$$T_1 r_1^a = T_3 r_2^a,$$

$$r_1 = \sqrt{R}\left(\frac{T_3}{T_1}\right)^{1/(2a)},$$

$$r_1 r_2 = R.$$

At the minimum the two ideal discharge temperatures coincide. Incomplete intercooling with $T_3 > T_1$ shifts more compression to the cooler first stage. For very different temperatures the stationary point may leave the allowable interval $1 \le r_1 \le R$; check the boundary then. A cooler with specified effectiveness can make T_3 vary with r_1, which requires an additional derivative.

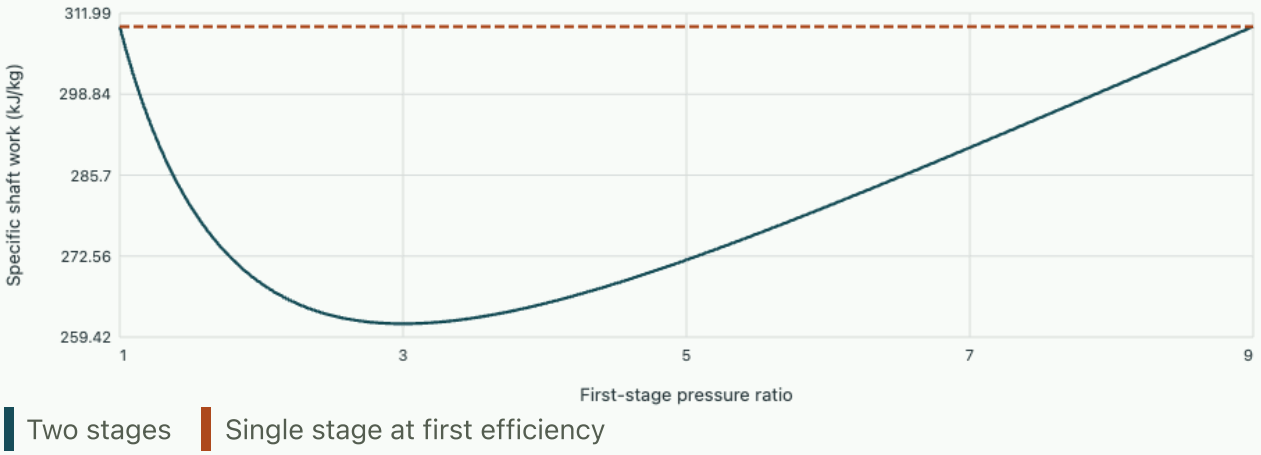
Use pressure, temperature and electrical power together

For $R = 9, T_1 = T_2 = 300$ K and efficiencies 0.85, each stage ideally uses ratio three. Each consumes 130.79 kJ/kg and the combined work is 261.59 kJ/kg. The otherwise identical single stage uses 309.82 kJ/kg. At 0.020 kg/s, two-stage shaft power is 5.23 kW before motor losses. The first-stage discharge temperature is about 430.1 K; cooling it back to 300 K rejects 130.8 kJ/kg, numerically equal to that first adiabatic stage's work under these property assumptions.

Open the live experiment

Try the calculation

Overall pressure ratio	9
Second inlet temperature (K)	300
Second efficiency	0.85



Minimum work is 261.59 kJ/kg at first ratio 3.000. The equal-ratio split requires 261.59 kJ/kg.

Raise the second inlet temperature or reduce its efficiency. The minimum now moves toward a higher first pressure ratio. The graph compares total specific work, keeping final absolute pressure fixed. It does not guarantee that a real compressor can operate at every displayed split: stage flow maps, surge, choke, motor limits and heat exchanger capacity may restrict the admissible point. Measure actual mass flow and pressure drop at each tested condition. A wattmeter upstream of a motor reports more than shaft power; use measured or justified motor efficiency when comparing with a gas enthalpy rise.

Close the work and rejected-heat balances separately

The two-stage minimum is a statement about shaft work per kilogram of ideal gas. An experimenter can test it by weighing delivered gas over a timed interval and measuring electrical energy or shaft torque. Converting an electrical reading to specific shaft work requires the motor efficiency at the measured operating point. A small motor often operates at a different efficiency when the intermediate pressure changes. Dividing the same electrical input by two different mass-flow readings can obscure a compressor change behind a motor change.

For an illustrative mass flow of 0.10 kg/s, a computed 262 kJ/kg corresponds to 26.2 kW at the shaft. At 90 percent motor efficiency the electrical input is about 29.1 kW. If the first compressor delivers at 430 K and the intercooler returns the gas to 300 K, its separate steady heat duty at the stated constant heat capacity is approximately 13.1 kW. A flow of cooling water warming by 10 K at 4.18 kJ per kilogram kelvin would need about 0.31 kg/s. These two measurements test different assumptions. Work compares both compressor stages; the cooling-water enthalpy rise compares only the intermediate exchanger, including any loss through its casing.

When measuring intake and discharge temperatures, distinguish stagnation from static temperature if velocity changes appreciably in connecting ducts. A thermocouple close to a wall may also respond to conduction along its sheath. Repeat each setting after allowing the intercooler and casing to approach steady temperatures. Heat retained in a warming machine changes the short-time coolant balance while leaving the algebraic optimum calculated for fully steady operation unchanged.

Tutorial · Find the new optimum for incomplete cooling

Suppose $R = 9, T_1 = 300$ K, $T_2 = 330$ K and both efficiencies remain 0.85. Predict the optimal first ratio and decide whether it is physically sensible. Estimate its work advantage over equal ratios using the graph or direct substitution.

▼ **Balance the two marginal temperature-weighted costs**

Since $1/(2a) = 1.75$, $r_{\mathrm{opt}} = 3(1.10)^{1.75} = 3.545$, so the second ratio is about 2.54. The hotter second stage now receives less compression work. Direct substitution gives approximately 273.5 kJ/kg versus 274.7 kJ/kg at equal ratios, an improvement only around one kJ/kg. Imperfect cooling itself costs much more relative to the original 261.6 kJ/kg. A noisy power sweep can therefore locate a rather broad minimum poorly even when cooling quality clearly matters.

Tutorial · Include an intercooler pressure loss

The cooler outlet pressure is a fraction $1 - \delta$ of first-stage discharge pressure. For fixed compressor inlet and final discharge pressures, what happens to the product of the actual compressor-stage ratios? Compute the equal-efficiency, fully cooled optimum at $R = 9$, $\delta = 0.05$.

▼ **Use absolute pressures on both sides of the exchanger**

With $p_2 = (1 - \delta) p_1 r_1$, final pressure $p_f = p_2 r_2$, so

$$r_1 r_2 = R/(1 - \delta) = 9.474.$$

Under identical stage inlet temperatures and efficiencies the optimum actual compressor ratios are equal to $\sqrt{9.474} = 3.078$. Thus the pressures around the cooler straddle the ideal loss-free geometric-mean intermediate pressure. A pressure tap only at first discharge cannot determine the second-stage ratio when exchanger loss is appreciable.

Tutorial · A temperature measurement conflicts with the power estimate

A trial records 0.020 kg/s, first-stage inlet 300 K and discharge 430 K. The drive wattmeter reports 3.25 kW for that stage. Estimate gas enthalpy power and explain two possible sources of the difference that have distinguishable signatures.

▼ **Draw separate drive and gas control volumes**

The gas enthalpy gain is $0.020(1.005)(130) = 2.61$ kW. Relative to electrical input this gives an apparent conversion fraction of 0.804. Electrical and mechanical motor losses are candidates, but heat rejected by a nonadiabatic compressor, temperature probe recovery errors, flow calibration and auxiliary equipment can also enter. Compare shaft torque times angular speed with electrical power to isolate drive losses. Measure casing heat rejection or obtain an energy balance around compressor plus coolant if shaft power remains above the gas gain. A single inferred efficiency cannot identify which physical boundary carries the missing 0.64 kW.

Ohio State thermal sciences · multistage compression with intercooling

Thermal processes and transport in engineering · 2. How a gravity-drained condensate film limits heat transfer

Saturated vapour condenses on a cool vertical plate. A liquid film thickens as it drains, adding thermal resistance toward the bottom. Derive the local heat-transfer coefficient from coupled momentum and energy balances, then determine which observations would invalidate the smooth-film model.

Find the liquid flow for a specified film thickness

Let x increase downward along a vertical isothermal wall, y normal to it, and $\delta(x)$ be the liquid thickness. Assume a steady laminar film with negligible inertia, constant liquid properties, negligible interfacial shear from slow vapour, saturation temperature at the interface, and a wall uniformly below saturation. Neglect sensible cooling of the condensate relative to latent heat and heat conduction along the plate. The common vapour pressure gradient is approximately hydrostatic, so net downward liquid driving per unit volume is $(\rho_l - \rho_v)g$.

Solve $\mu_l u_{yy} = -(\rho_l - \rho_v)g$ with no slip at $y = 0$ and zero shear at $y = \delta$. The velocity is a downward parabola $u = (\rho_l - \rho_v)g(\delta y - y^2/2)/\mu_l$. Integrating gives mass flow per unit width $\Gamma = \rho_l(\rho_l - \rho_v)g\delta^3/(3\mu_l)$. The velocity increases from the wall to its interface maximum; assigning the interface a second no-slip condition would substantially underpredict gravity drainage.

A growing mass flow changes the conductive resistance

If transverse heat conduction dominates, approximately linear temperature through the film gives wall heat flux $q'' = k_l(T_{\mathrm{sat}} - T_w)/\delta$. New condensate produced per unit wall length adds $d\Gamma/dx = q''/h_{fg}$. Substitute the cubic film-flow relation and integrate from an ideal leading edge of zero film thickness:

$$\delta(x) = \left[\frac{4\mu_l k_l \Delta T\, x}{\rho_l(\rho_l - \rho_v)g h_{fg}}\right]^{1/4},$$

$$h_x = \frac{k_l}{\delta(x)},$$

$$\bar{h}_L = \frac{4}{3}h_L.$$

The leading coefficient diverges at $x = 0$ in this idealisation, but its integral is finite. A real leading edge, substrate conduction and contact-angle effects regularise the immediate entrance. The factor $4/3$ comes from integrating $x^{-1/4}$, not averaging endpoint values. A plate made twice as long gains heat transfer area but increases film resistance downstream; total heat per unit width rises only as $L^{3/4}$.

As an illustrative water-like fluid take $\rho_l = 960$, $\rho_v = 0.6$ kg/m³, $k_l = 0.68$ W/(m K), $\mu_l = 2.9 \times 10^{-4}$ Pa s, $h_{fg} = 2.26$ MJ/kg, $\Delta T = 5.0$ K and $L = 0.20$ m. The outlet film is roughly 79 micrometres and the outlet coefficient about 8.6 kW per square metre kelvin. The average coefficient is about 11.5 kW per square metre kelvin; a 0.10 m wide plate removes about 1.15 kW and yields roughly 0.51 g/s condensate. Calculate film Reynolds number with the specified convention before accepting the assumed smooth regime.

Return to the compressor and intercooler calculation to compare this thermal coefficient with an actual heat duty

Test the quarter-power prediction at several stations

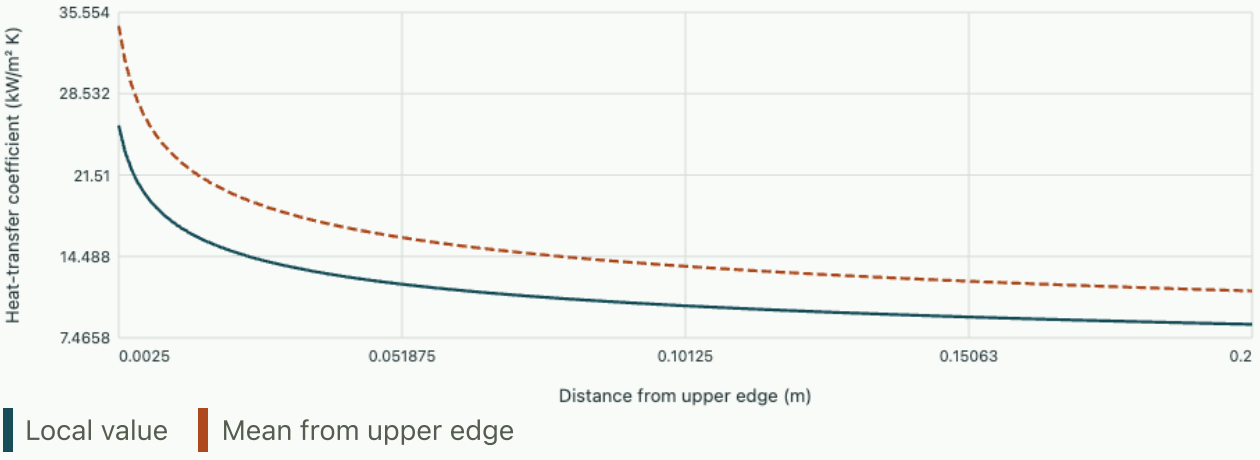
A film-thickness measurement is easiest away from the leading edge. There the ideal formula predicts a thin film and formally unbounded heat transfer coefficient. Real entrance geometry, surface condition, contact-line physics and measurement resolution take over. In the study the curves begin at a small positive fraction of plate length. This preserves the finite integral of the singular ideal heat flux yet avoids assigning a physical measured value at the mathematical edge.

Open the live experiment

Try the calculation

Wall subcooling (K) **5**

Plate length (m) **0.2**



The outlet film is 78.8 µm thick. A 0.10 m wide plate rejects 1.150 kW.

For an independent test place two thermal sensors along the downstream wall and estimate local heat flux by conduction across a thin backing of known thickness and conductivity. At two stations whose distances differ by a factor sixteen, the ideal prediction halves the local coefficient. A common wall heat loss can distort this ratio much more at the downstream station, where the useful heat flux is smaller. Repeat with vapour absent to estimate the backing loss and use an unheated reference section to check temperature gradients in the support.

The liquid mass collected at the bottom integrates all upstream condensation. Doubling plate length increases total condensation by a factor of two raised to three quarters under fixed conditions, even as the mean coefficient falls. If the collector instead finds exactly twice the condensate and local sensors show small axial variation, test for liquid dripping onto the plate from above, a nonuniform vapour

temperature or turbulent mixing. Keep track of condensate retained on the surface when using short timed collection intervals; intermittent release into the collector can dominate the apparent uncertainty.

Tutorial · Reconstruct a heat balance from collected mass

A plate of the stated size produces 0.48 g/s of condensate in a steady timed collection. Estimate latent heat removal and inferred mean coefficient. Does disagreement with the wall-coolant calorimeter necessarily imply an incorrect latent heat?

▼ **Convert the collected mass into a rate of thermal energy**

Latent duty is

$$0.00048(2.26 \times 10^{6}) = 1085\,\mathrm{W}.$$

The effective mean coefficient is $1085/(0.020 \cdot 5) = 10.85$ kW per square metre kelvin. Disagreement with a coolant energy balance can arise from heat leaks, condensate subcooling, liquid holdup or drainage that the collection misses. Weigh over a sufficiently long stationary interval and measure condensate outlet temperature. A transient wetting front stores liquid and invalidates equating instantaneous formation and collection.

Tutorial · Compare two uses of the same plate area

Two plates can be placed vertically in series so liquid feeds from the upper into the lower, or arranged side by side with independently drained leading edges. Each plate has length L and width b. Estimate the ideal heat-duty ratio at fixed temperatures.

▼ **Compare fresh versus pre-existing films**

Side-by-side total duty is $2Q(L)$. A continuous vertically stacked film behaves like one length $2L$, giving $Q(2L) = 2^{3/4}Q(L)$. Independently drained plates therefore give a factor $2^{1/4} = 1.189$ more ideal duty for the same area. Additional vapour routing, wall conduction and edge drainage can change real performance. The advantage arises from repeatedly beginning with a thin film, not from more total geometric area.

Tutorial · Diagnose a vapour-side resistance

A nominally identical plate shows dramatically smaller heat flux after a small amount of noncondensable gas enters the vapour volume, while liquid film thickness measurements barely change. Why can changing only liquid conductivity in a fit give a misleading result?

▼ Identify a missing interface boundary condition

The model assumed saturated pure vapour at an interface of known temperature. A noncondensable component can accumulate at the interface and require vapour diffusion through a gas layer. Local vapour partial pressure and interface saturation temperature can differ from estimates based on total chamber pressure. Heat and species transport on the vapour side then add resistance. Measure composition and partial pressure, sweep the vapour flow or purge the noncondensable fraction and compare flux with an independent film profile. Fitting a different liquid conductivity can mimic the overall flux but would generally predict the wrong response to those controls.

Lienhard and Lienhard · A Heat Transfer Textbook, film condensation and phase change

Thermal processes and transport in engineering · 3. Radiative exchange in a grey enclosure

A thin radiation shield does not need to absorb all incoming radiation to reduce heat transfer. Multiple reflections connect every surface to every other visible surface. Build the radiosity equations, recover the two-plate result and infer what can be learned from a measured shield temperature.

Separate emission, reflection and net heat flux

For opaque diffuse-grey isothermal surface i, let G_i be irradiation arriving from the enclosure and J_i be all radiative power leaving per unit area. The outgoing consists of thermal emission plus reflected irradiation. With emissivity ϵ_i, reflectivity $1-\epsilon_i$ and blackbody emissive power $E_{bi}=\sigma T_i^4$, one obtains

$$J_i=\epsilon_i E_{bi}+(1-\epsilon_i)G_i.$$

Net outward heat is $Q_i=A_i(J_i-G_i)$.

Geometry enters through view factors: $G_i=\sum_j F_{ij}J_j$ for a closed nonparticipating enclosure of diffuse surfaces, provided reciprocity $A_iF_{ij}=A_jF_{ji}$ has been used consistently. Each closed-enclosure row sums to one, and summing Q_i over all surfaces gives zero. A concave surface can see itself, so setting every diagonal view factor to zero would break some real enclosure models. For an opening to surroundings include an explicit environmental surface or known irradiance source.

Eliminate the unknown reflected fields

Two large parallel plates of equal area see only one another. Thus $G_1=J_2, G_2=J_1$. Eliminate both radiosities, or add a surface resistance for each plate and a unit space resistance in the diffuse network. Net flux from hot to cold becomes

$$q''=\frac{\sigma(T_1^4-T_2^4)}{1/\epsilon_1+1/\epsilon_2-1}.$$

The denominator is not simply the reciprocal of the product of emissivities because photons can reflect repeatedly before absorption. In the black limit both emissivities are one and the cold plate's own emission remains in the numerator. At equal temperature flux vanishes even when emissivities differ. Measuring only outgoing hot-side radiance without reflected incoming light cannot determine heat loss.

For $T_1=500$ K, $T_2=300$ K and emissivities 0.80 and 0.60, the blackbody power difference is 3085 W/m² and the denominator 1.9167, giving about 1609 W/m². Place an opaque thermally floating shield between the plates, with both faces of emissivity 0.10 and negligible conduction to its supports. The same steady flux crosses both gaps. The two gap resistances add in series, with total dimensionless denominator $1/\epsilon_1+2/\epsilon_s+1/\epsilon_2-2=20.9167$, leaving about 147.5 W/m².

Check the underlying blackbody emission and view-factor balances in the radiation chapter

Determine its temperature from the steady flux

For the above example the first gap denominator is $1/0.80 + 1/0.10 - 1 = 10.25$ and the second 10.6667. Write $E_{bs} = E_{b1} - q'' D_1$ or $E_{bs} = E_{b2} + q'' D_2$. The shield temperature is about 435 K. It is a weighted average of fourth powers, not the arithmetic average of temperatures. A measured shield temperature without net heat flux generally constrains resistance ratios, while the flux supplies their sum.

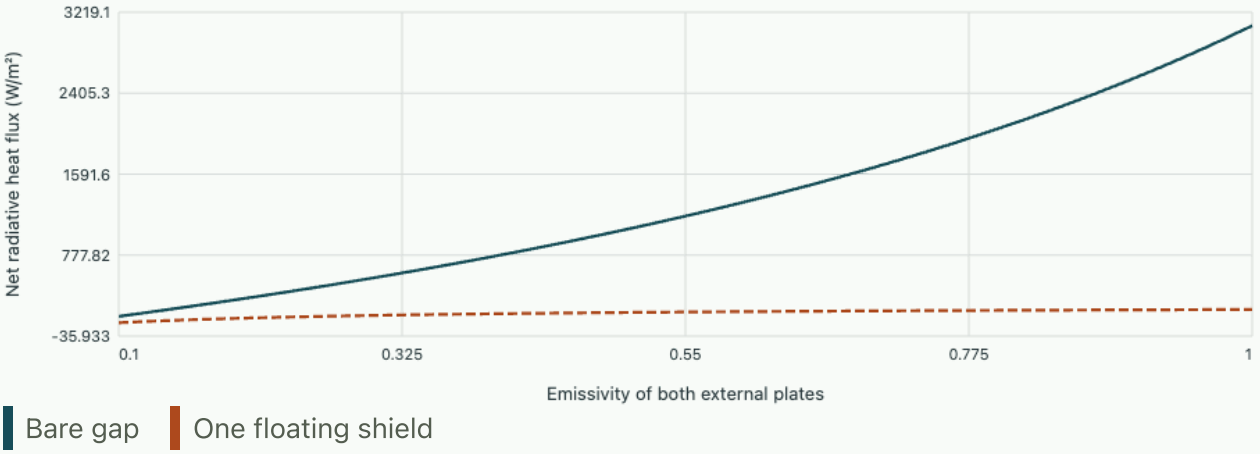
Use a floating shield as a two-sided calorimeter

Insert an opaque isothermal sheet between the two parallel walls. At steady state its net accumulation vanishes, so energy transferred across the hot gap equals energy leaving the cold gap. For identical external surface emissivities and identical shield-face emissivities, the two radiative resistances of the gaps are the same. Equality of the fourth-power temperature drops then places the shield temperature at the fourth-power mean of the two plate temperatures. It is generally above their arithmetic mean. Lower shield emissivity changes the size of the common heat flow while leaving the intermediate temperature fixed in this symmetric problem.

Open the live experiment

Try the calculation

Hot surface temperature (K)	500
Cold surface temperature (K)	300
Shield emissivity, both faces	0.15



At plate emissivity 0.8 the bare and shielded fluxes are 2056.5 and 223.0 W/m². Both faces of the isothermal shield share temperature 433.5 K.

Make a paired test using a matte shield and a polished shield of the same dimensions. A heat-flux transducer should respond strongly, while an ideal shield-temperature measurement would indicate the same equilibrium value. That contrast provides a cleaner test than relying on a single fitted emissivity. A

thermocouple lead can conduct heat away from the shield and change its temperature, particularly for the polished sheet when radiative conductance is small. A radiation thermometer aimed at that sheet can receive reflected hot-wall radiation. Use a conducting contact sensor with estimated lead loss and a separately characterised radiance measurement to investigate discrepancies.

If the shield faces have different coatings, reverse it without changing either plate. The series resistance and net heat flow stay the same under the simple diffuse-grey model because the two face resistances enter as a sum. The equilibrium shield temperature changes, however, because the allocation of resistance to the hot and cold gaps changes. This distinguishes an orientation-dependent temperature from an orientation-dependent heat flux.

Tutorial · Can two identical black plates benefit from a black shield?

Keep the same temperatures but make both plates and both shield faces black. Calculate flux with and without one floating shield. Would replacing one shield by two double the benefit again?

▼ **Count separate spaces and solve their fluxes**

Without a shield the net flux is about 3085 W/m². With one black shield two identical gap resistances divide it in half, about 1542 W/m². With two black shields there are three equal resistances and flux is one third of the unshielded value, about 1028 W/m². Each plate emits and absorbs on its own faces even when black. Equal changes in the count of shields add equal resistances in this ideal case rather than multiplying the previous flux by one half each time.

Tutorial · Test a view-factor proposal by reciprocity

Two surfaces have areas 0.020 and 0.050 square metres. A ray calculation gives $F_{12} = 0.70$. What should F_{21} be and why cannot the uncovered row sum automatically be assigned to the same neighbouring surface?

▼ **Use the area-weighted symmetry**

Reciprocity gives

$$F_{21} = 0.020(0.70)/0.050 = 0.28.$$

Row remainders are respectively 0.30 and 0.72. They can describe additional walls, an opening or a surface's self-view if geometry permits. Two convex finite unequal plates alone do not form a closed enclosure with unit rows. Give their surrounding environment a temperature and surface properties or prescribe incoming irradiation before computing the net powers.

Tutorial · Infer an omitted conductive heat leak

An instrument with the example shield and plate area 0.10 square metres reports net heat input of 19 W at the hot wall after other losses are removed. Assume the ideal radiation estimate holds. Estimate the parallel conductance directly connecting hot and cold walls and propose a second test.

▼ Compare absolute powers and their different temperature scaling

Predicted radiation is about 14.75 W, leaving 4.25 W not explained. At a 200 K temperature difference a direct linear bypass would require 0.0213 W/K. This could represent support conduction but is not unique: emissivity errors, finite plate-edge view to surroundings or residual gas conduction can also contribute. Repeat at a different mean absolute temperature while holding the same wall-temperature difference. Linear solid conduction with weak property variation remains similar, whereas radiative transfer grows strongly with mean temperature. Monitor the floating shield temperature too; direct bypass need not heat it, whereas a support attached to the shield changes its equilibrium.

Lienhard and Lienhard · A Heat Transfer Textbook, radiation exchange in enclosures

Thermal processes and transport in engineering · 4. Periodic heating as a measurement of thermal diffusivity

Two buried thermometers see the same applied surface temperature oscillation with different amplitudes and phases. Derive how a solid filters that oscillation, recover diffusivity independently from both observations and identify when a rear boundary or a poor sensor contact changes the inference.

Choose the decaying branch of a complex thermal wave

A homogeneous solid fills $x > 0$ and has constant conductivity k, density ρ and heat capacity c_p. For small one-dimensional disturbances its temperature excess obeys $\partial_t \theta = \alpha \partial_x^2 \theta$, $\alpha = k/(\rho c_p)$. Impose sinusoidal surface temperature with amplitude A and angular frequency ω, and wait until the initial transient has decayed. With a positive-exponential time convention look for $\theta = \mathrm{Re}[A e^{-sx + i\omega t}]$. Then $s^2 = i\omega/\alpha$, and boundedness far from the wall selects

$$s = (1 + i)/\delta,$$

$$\delta = \sqrt{2\alpha/\omega},$$

$$\theta(x, t) = A e^{-x/\delta} \cos(\omega t - x/\delta).$$

One penetration depth produces both an amplitude factor e^{-1} and a phase lag of one radian. This is diffusive propagation: it is not an undamped travelling wave and its phase velocity depends on frequency. Temperature is continuous at the surface condition, while required input heat flux has complex amplitude ksA. Its magnitude is $kA\sqrt{\omega/\alpha}$ and its phase leads surface temperature by 45 degrees. A known heat flux and surface amplitude can therefore supply information about thermal effusivity $\sqrt{k\rho c_p}$, distinct from diffusivity.

Extract a complex slope from two thermometer depths

For two sensors separated by known Δx, the amplitude ratio gives $\log(A_1/A_2) = \Delta x/\delta$ and unwrapped phase difference gives

$$\Delta\phi = \Delta x/\delta.$$

Either yields $\alpha = \omega(\Delta x)^2/(2\Delta\phi^2)$ when amplitude log is substituted for phase if desired. A common unknown drive amplitude and common timing phase cancel. Different sensor gains or unequal sensor time constants do not cancel automatically.

For a solid with $\alpha = 1.0 \times 10^{-5}$ square metres per second heated at 0.10 Hz, $\delta = 5.64$ mm. Sensors 3.0 mm apart should display amplitude ratio $A_2/A_1 = 0.588$ and lag 0.532 rad, about 30.5 degrees and 0.847 s. Increasing frequency by a factor four halves penetration depth and doubles the phase lag at the same separation. The signal at depth falls more strongly and may eventually become noise-dominated although the theoretical phase difference grows.

Compare time-dependent wall-temperature inference with a steady-stage energy balance

A finite slab creates a returning thermal disturbance

The half-space solution is valid only if the far boundary is several penetration depths beyond the region of interest and if lateral heat losses or heat spreading remain small. In a slab of thickness L with its back face held at zero temperature excess, an exact periodic solution satisfying both faces is proportional to $\sinh[s(L-x)]/\sinh(sL)$. An insulating back instead gives $\cosh[s(L-x)]/\cosh(sL)$. Both reduce to the decaying exponential far from the back at sufficiently high frequency. Fitting the half-space law at very low frequency in a thin specimen can bias diffusivity even with perfect temperature readings.

Use the time traces to separate a material delay from a sensor delay

In the experiment the two displayed signals share a clock and a known drive amplitude. Read off the time difference between successive maxima, multiply it by angular frequency, and compare with the logarithm of the two amplitudes. Both dimensionless readings estimate depth divided by penetration depth. Increase drive frequency while keeping all geometry fixed. Thermal propagation predicts the same square-root frequency scaling for both quantities. An electronic timing offset instead contributes a phase proportional to frequency and leaves amplitude unchanged.

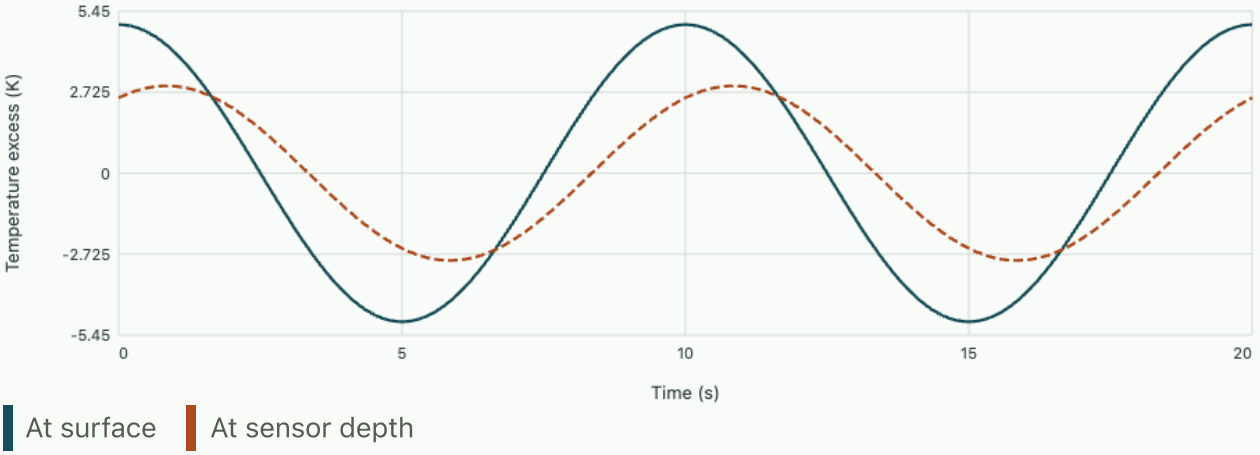
Open the live experiment

Try the calculation

Drive frequency (Hz) **0.1**

Thermal diffusivity (mm²/s) **10**

Sensor separation (mm) **3**



Penetration depth is 5.64 mm. At the selected depth the amplitude ratio is 0.588 and the phase lag is 30.5°.

A thermistor embedded in a drilled hole may exchange heat with adhesive and its own leads. A simple first-order sensor with time constant one second has gain magnitude about 0.847 and phase lag about

32 degrees at one tenth hertz. That is already comparable with the nominal three-millimetre material phase in the example. Exchange the two sensors at the same holes after a baseline run. A phase shift which follows the sensor points to unequal time response; one that stays attached to the location is more consistent with diffusion or a location-dependent contact layer.

To compare a finite solid with the semi-infinite model, estimate the reflected thermal contribution from the far boundary. A disturbance travelling a distance equal to five penetration depths decays in amplitude by a factor near 0.0067. A boundary much nearer than that may noticeably alter phase and attenuation, particularly for a low drive frequency. The relevant distance for a reflected disturbance includes both the outward and return journey. Repeating the measurement at two thicknesses can reveal this effect while preserving material properties and sensor spacing.

Tutorial · Check two independent diffusivity estimates

A run at 0.20 Hz with 2.0 mm separation gives amplitude ratio 0.600 and lag 0.500 rad. Infer diffusivity separately from the two measurements and decide whether a simple phase-wrap error is likely.

▼ **Invert the attenuation and lag laws**

The amplitude logarithm is 0.5108. Its estimate is $\alpha = 2\pi(0.20)(0.002)^2/[2(0.5108)^2] = 9.63 \times 10^{-6}$ square metres per second. The phase estimate is 1.01×10^{-5} in the same units. The small difference is plausible with roughly percent-scale amplitude or phase uncertainty. A full missed wrap would add 2π and change the inferred diffusivity dramatically. Repeat at adjacent frequencies to establish a continuous unwrapped phase rather than selecting a branch from a single fit.

Tutorial · Separate conductivity and volumetric heat capacity

A second measurement of imposed surface heat-flux amplitude and temperature identifies effusivity $e = 1.50 \times 10^4$ in SI units. Diffusivity is 1.0×10^{-5} square metres per second. Recover conductivity and volumetric heat capacity.

▼ **Combine the product and ratio of material properties**

Write $e = \sqrt{kC}$, $\alpha = k/C$, where $C = \rho c_p$. Then $k = e\sqrt{\alpha} = 47.4$ W/(m K) and $C = e/\sqrt{\alpha} = 4.74 \times 10^6$ J/(m³ K). Neither a temperature attenuation nor a phase slope alone identifies both quantities. Surface convection, heater thermal mass and contact resistance can distort the inferred heat actually entering the specimen, so an electrical heater input must be corrected before assigning the effusivity.

Tutorial · Design a frequency window for a finite coupon

A 20 mm specimen has its second sensor at depth 5 mm. A designer wants at least three penetration depths from that sensor to the back wall and a sensor-depth amplitude exceeding one tenth of the surface amplitude. Find an approximate interval of allowable penetration depths, then convert it to frequencies for the example diffusivity.

▼ **Satisfy attenuation and boundary constraints simultaneously**

The back-wall criterion gives

$$\delta \leq (20 - 5)/3 = 5.0\,\mathrm{mm}.$$

The signal criterion gives $e^{-5\mathrm{mm}/\delta} > 0.10$, or $\delta > 2.17$ mm. Since $f = \alpha/(\pi\delta^2)$, the interval is approximately 0.127 to 0.675 Hz. These are design heuristics for the stated geometry and signal requirement, not universal boundaries of diffusion theory. A finite-slab calculation at both endpoints can quantify actual back-boundary error before committing to the measurement.

Lienhard and Lienhard · A Heat Transfer Textbook, transient conduction and penetration depth

Thermal processes and transport in engineering · 5. What temperature does a probe in a hot gas measure?

A small bead in a duct exchanges energy with both the gas and the duct walls. Calculate the indicated temperature from an energy balance, infer the gas temperature from the reading, and test how emissivity changes the measurement.

Draw the control surface around the bead

Let the gas have temperature T_g and the bead a uniform temperature T_s. A large, isothermal enclosure at T_w fills the bead's field of view. Use kelvins throughout. Convection supplies the bead with heat at $hA(T_g - T_s)$. A diffuse grey bead emits $\epsilon\sigma A T_s^4$ and absorbs $\epsilon\sigma A T_w^4$ from the enclosure. The absorptivity and emissivity are equal for the grey surface. The common area A cancels from the steady energy balance when axial conduction through the support is negligible.

$$F(T_s) = h(T_g - T_s) - \epsilon\sigma(T_s^4 - T_w^4) = 0.$$

For a hotter gas, $F(T_w)$ is positive, $F(T_g)$ is nonpositive, and $F'(T_s) = -h - 4\epsilon\sigma T_s^3$ is strictly negative. There is exactly one equilibrium between the two reservoir temperatures. A colder gas reverses their ordering and keeps the same argument. This physical bracket gives a bisection solver two trustworthy endpoints. With positive heat capacity, a small rise in bead temperature makes the net heat input negative, which restores equilibrium.

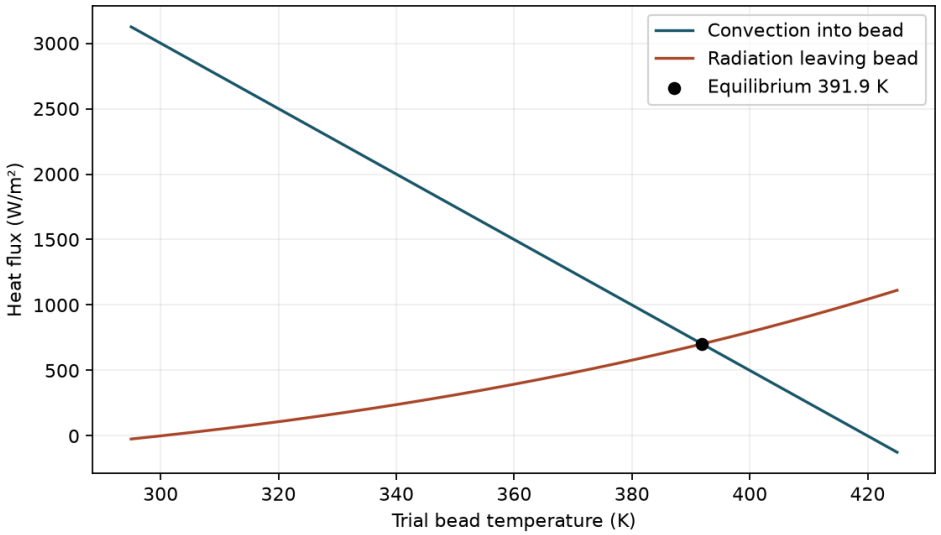
The factorization $T_s^4 - T_w^4 = (T_s - T_w)(T_s + T_w)(T_s^2 + T_w^2)$ gives an exact coefficient $h_r = \epsilon\sigma(T_s + T_w)(T_s^2 + T_w^2)$. At the root,

$$T_g - T_s = h_r(T_s - T_w)/h.$$

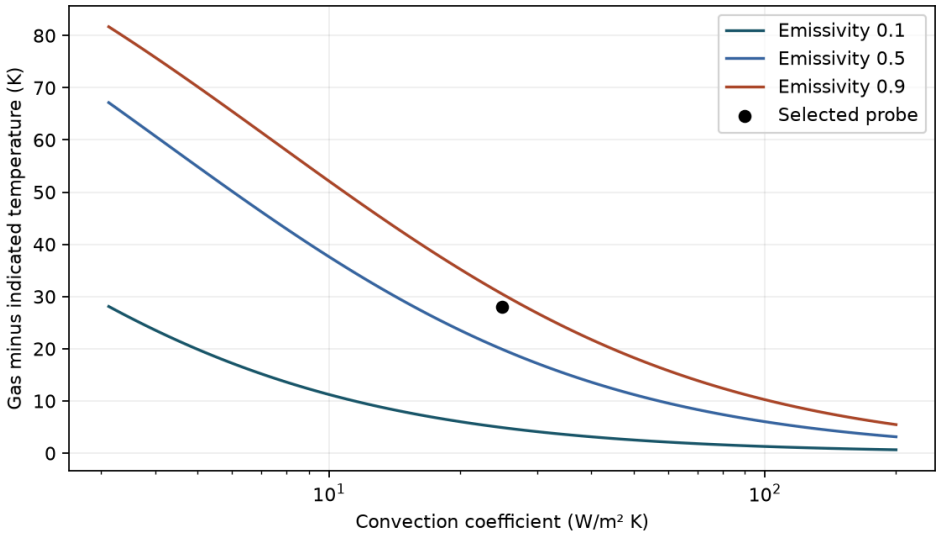
A shiny surface and a fast gas stream therefore act on different terms of the same equation. A larger bead changes the transient response through its mass-to-area ratio, and may also change h. Area alone cannot correct the steady reading at fixed h.

Predict the indicated temperature

Take $T_g = 420$ K, $T_w = 300$ K, $h = 25$ W m^{-2} K^{-1} and $\epsilon = 0.8$. At 390 K, convection supplies 750 W m^{-2} and the radiative loss is approximately 682 W m^{-2}. The bead thus warms. At 400 K the convective input has fallen to 500 W m^{-2} while radiative loss reaches 794 W m^{-2}. The intersection gives $T_s = 391.90$ K, an under-reading of 28.10 K. The two independently evaluated fluxes at equilibrium agree to round-off. Doubling h gives 403.34 K for the same gas and wall. Metal lead wires connected to a cold mount could supply a third heat path, which the two-flux residual cannot independently diagnose.



Follow the curves to the same bead temperature. Their vertical separation is the energy available to warm or cool the bead.



Track one surface along its curve to change the flow, or hold the heat-transfer coefficient fixed and compare surface finishes.

Problem 1 · Correct a recorded temperature

The probe reads 400 K. The walls remain at 300 K, $h = 25$ W m^{-2} K^{-1} and $\epsilon = 0.8$. Recover the gas temperature without iteration. How would a hotter enclosure affect the sign of this correction?

▼ Work through the solution

Rearranging the balance gives

$$T_g = T_s + \epsilon\sigma(T_s^4 - T_w^4)/h.$$

The correction is $793.85/25 = 31.75$ K, giving 431.75 K. If the walls were hotter than the bead, absorbed radiation would exceed emitted radiation and the correction to the reading would be negative. A Celsius temperature can be reported at the end after the fourth powers have been evaluated in kelvins.

Problem 2 · Specify a minimum convection coefficient

For the original gas at 420 K, find h required for the under-reading to be at most 5 K. Then reduce the emissivity to 0.2 and repeat. Explain the ratio.

▼ **Work through the solution**

At the limiting reading $T_s = 415$ K the required coefficient is $\epsilon\sigma(415^4 - 300^4)/5$. For $\epsilon = 0.8$ this is approximately 196 W m^{-2} K^{-1}. The required coefficient scales directly with emissivity at this prescribed reading, so the lower-emissivity bead needs about 49 W m^{-2} K^{-1}. Above those coefficients the gas transfers more heat at the limiting temperature and drives the equilibrium closer to 420 K.

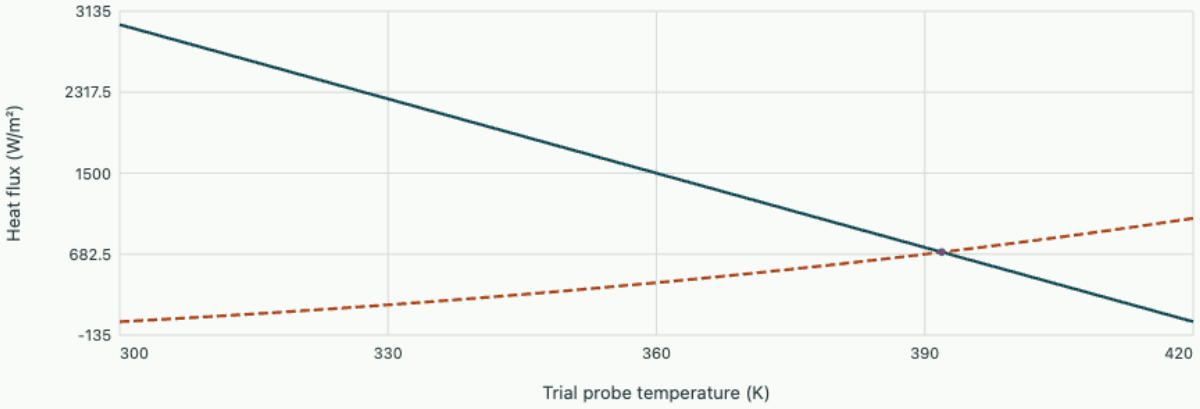
Interactive calculation · Probe balance

Change the physical and numerical inputs of the notebook model. Compare the updated curves and balance checks with the recorded notebook figures. Each plot uses the same equations and assumptions documented in that notebook; its readout names any analytical or reduced-model reference.

Open this live calculation

Try the calculation

Gas temperature (K)	420
Wall temperature (K)	300
Film coefficient (W/m² K)	25
Probe emissivity	0.8



Indicated temperature 391.90 K; bias 28.10 K. The equilibrium balances convection and grey-surface radiation.

Download this example's editable notebook

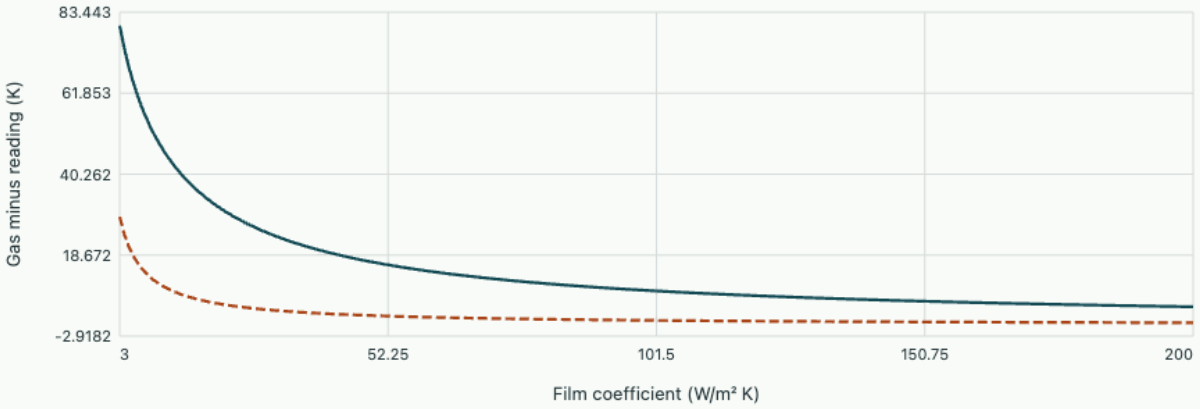
Interactive calculation · Probe bias

Change the physical and numerical inputs of the notebook model. Compare the updated curves and balance checks with the recorded notebook figures. Each plot uses the same equations and assumptions documented in that notebook; its readout names any analytical or reduced-model reference.

Open this live calculation

Try the calculation

Gas temperature (K)	420
Wall temperature (K)	300
Emissivity	0.8



At25 W/(m² K), bias 28.10 K. More convective coupling or less radiative coupling moves the reading toward the gas.

Download this example's editable notebook

Open editable notebook Download .ipynb

Thermal processes and transport in engineering · 6. Where does the temperature of a reacting gas come from?

A heater in an ideal cycle can supply a prescribed amount of heat. A steady combustor converts chemical enthalpy into sensible enthalpy. Compute a state from atomic balances and show where the complete-conversion model begins to overpredict its temperature.

Element balances determine the complete-conversion products

Choose one mole of methane as the feed basis, ideal dry air represented by oxygen plus 3.76 nitrogen molecules per oxygen molecule, and an air ratio λ relative to stoichiometric. For $\lambda \geq 1$ and complete conversion to carbon dioxide and water vapour, balance carbon, hydrogen, oxygen and nitrogen separately:

$$\mathrm{CH_4} + 2\lambda(\mathrm{O_2} + 3.76\,\mathrm{N_2}) \longrightarrow \mathrm{CO_2} + 2\mathrm{H_2O} + 2(\lambda - 1)\mathrm{O_2} + 7.52\lambda\mathrm{N_2}.$$

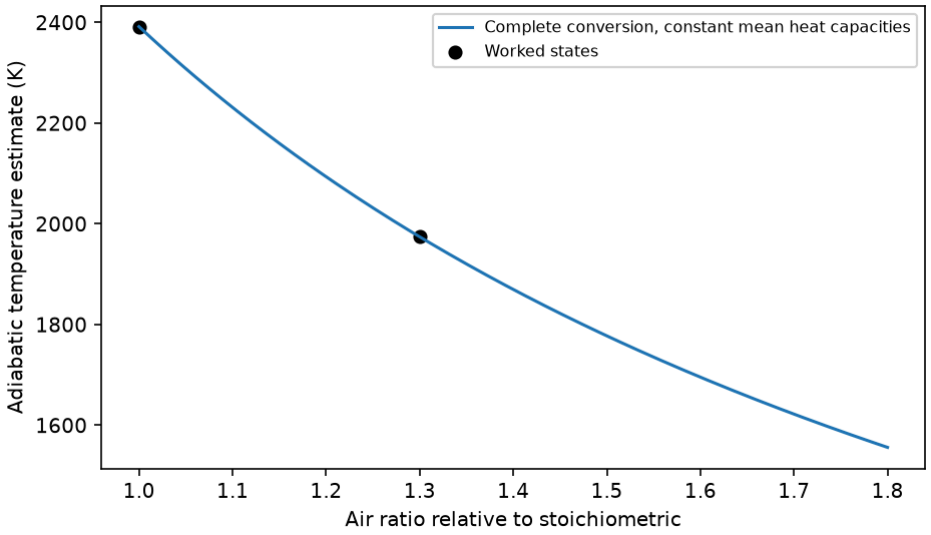
This balances atoms even though individual chemical species are created and consumed. At $\lambda = 1.3$ the products include 0.6 mol oxygen and 9.776 mol nitrogen per mole of fuel. The equivalence ratio is $1/\lambda$ for this feed. If oxygen is deficient, the assumed set of products cannot satisfy these balances with nonnegative amounts.

Use a common reference for chemical and sensible energy

In steady adiabatic operation without shaft work or appreciable kinetic-energy change the inlet and outlet enthalpy flows are equal. For an ideal gas write $h_i(T) = \Delta h^\circ_{f,i}(T_r) + \int_{T_r}^{T} c_{p,i}(u)du$. Elements in their reference states have zero standard formation enthalpy. Water vapour and liquid water have different formation enthalpies. Use the product phase appearing in the balance.

$$\sum_{p} n_i \int_{T_r}^{T_{ad}} c_{p,i}(u)\,du = -\Delta H^\circ_{rxn}(T_r) + \sum_{r} n_i \int_{T_r}^{T_{in,i}} c_{p,i}(u)\,du.$$

Use feeds at 298 K and an illustrative heat release of 802.3 kJ/mol of methane for vapour water. Take constant mean product heat capacities 51, 42, 35 and 33 J/(mol K) for carbon dioxide, water, oxygen and nitrogen, respectively. For stoichiometric air the total is $51 + 2(42) + 7.52(33) = 383.16$ J/K per mole of fuel. The resulting estimate is 2392 K. With 30% excess air the sum rises to 478.61 J/K and the estimate falls to 1974 K. Both are deliberately conditional on constant averaged heat capacities and the assumed product set.



Additional air raises the heat capacity of the products. Both plotted states assume complete conversion with water vapour and the mean heat capacities given above.

Close energy and composition together

At high temperature allow the product reaction $\mathrm{CO_2} \rightleftharpoons \mathrm{CO} + \frac{1}{2}\mathrm{O_2}$ to proceed by extent ξ. Replace carbon dioxide by $1 - \xi$, add ξ carbon monoxide, and add $\xi/2$ oxygen. For ideal gases define $a_i = y_i p / p^\circ$. Chemical equilibrium and the energy balance are now two equations for T and ξ:

$$\frac{a_{CO} a_{O_2}^{1/2}}{a_{CO_2}} = K_p(T),$$

$$\sum_i n_i(\xi) h_i(T) = H_{in}.$$

Dissociation stores energy in chemical form, altering temperature and oxygen content. With one illustrative reaction the calculation isolates that effect. A quantitative high-temperature property model requires all important species, pressure and temperature-dependent thermochemical data. The physical root must have nonnegative species amounts.

> **Is there heat crossing the wall?**
>
> In this adiabatic model wall heat transfer is zero. Formation enthalpies carry the chemical energy with the flowing matter. Adding a separate fuel heating value to a balance that already includes formation enthalpies counts the same conversion twice.

Solve two independent reactions and the energy equation numerically

Use the preceding methane basis with stoichiometric air, at 298.15 K and initially 1 bar. Permit the carbon reaction above with extent x and the water dissociation $\mathrm{H_2O} \rightleftharpoons \mathrm{H_2} + \tfrac{1}{2}\mathrm{O_2}$ with extent y. The product amounts in the order carbon dioxide, carbon monoxide, steam, hydrogen, oxygen and nitrogen are

$$n = (1 - x,\; x,\; 2 - y,\; y,\; (x + y)/2,\; 7.52).$$

These amounts preserve all four elemental inventories for any admissible pair. At an air ratio above one, add $2(\lambda - 1)$ mol oxygen and use 7.52λ mol nitrogen. Both reactions share the same oxygen activity. Solving them as independent single-reaction mixtures would give two incompatible oxygen partial pressures.

In the stated species order use standard formation enthalpies $(-393.52, -110.53, -241.83, 0, 0, 0)$ kJ/mol and standard entropies $(213.79, 197.66, 188.84, 130.68, 205.15, 191.61)$ J/(mol K) at $T_r = 298.15$ K. For this pedagogical approximation use constant mean heat capacities $(51, 33, 42, 31, 35, 33)$ J/(mol K) over the whole interval. The inlet formation enthalpy is −74.87 kJ per mole of methane. The common pressure reference is 1 bar. For each species integrate c_p for enthalpy and c_p/T for entropy. With reaction stoichiometric vectors ν_j, compute

$$h_i = h^\circ_{f,i} + c_{p,i}(T - T_r),$$

$$s^\circ_i = s^\circ_{i,r} + c_{p,i}\ln(T/T_r),$$

$$\ln K_j = -\frac{\sum_i \nu_{ji}(h_i - T s^\circ_i)}{RT}.$$

Enforce $\sum_i \nu_{ji}\ln[(n_i/\sum n)(p/p^\circ)] - \ln K_j = 0$ for both reactions and $\sum n_i h_i - H_{in} = 0$ for energy. The notebook bounds x between zero and one and y between zero and two. It scales the energy residual by 100 kJ and verifies both equilibrium residuals and the atomic balances independently. At 1 bar the solution has $x = 0.11117$, $y = 0.03649$ mol and $T = 2286.01$ K, about 106 K below the corresponding complete-conversion estimate. Products contain 0.07383 mol oxygen per mole of methane even though the feed was stoichiometric.

Interactive experiment · Change the pressure and air ratio of an adiabatic mixture

Move the pressure slider and compare the recomputed equilibrium temperature with the complete-conversion curve. The air-ratio slider places an operating point on that curve and reports the two reaction extents. Every update solves the same two chemical-equilibrium equations and enthalpy balance specified above, using damped Newton steps in the browser. A bounded composition parametrisation keeps all species amounts positive.

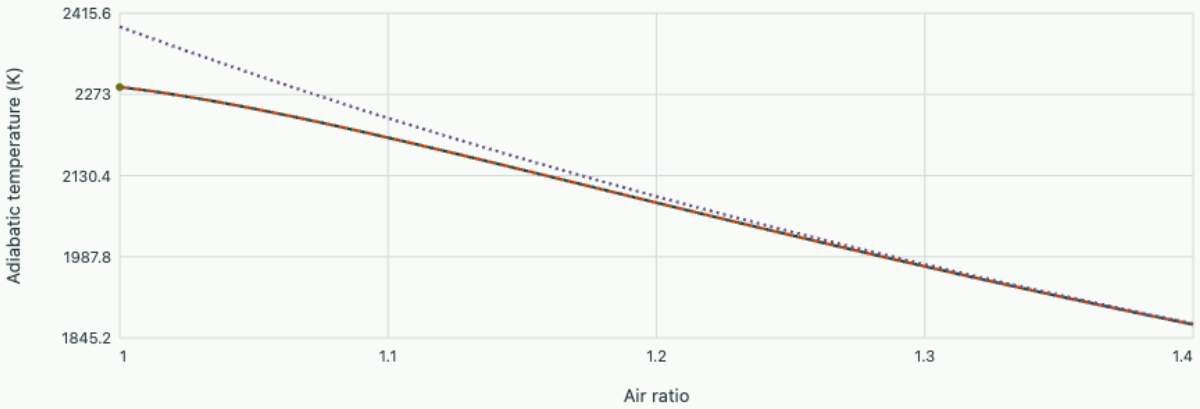
At stoichiometric feed, raise pressure from 1 to 5 bar and explain why both dissociation reactions are suppressed: each increases the total number of gas molecules. Now add excess air and identify where dilution dominates the temperature change. The readout includes a residual check on the solved equations. The fixed mean heat capacities and six-species product set are deliberately limited thermochemical assumptions, shared with the downloadable OneSimulation model.

Open this experiment online

Try the calculation

Air ratio **1**

Total pressure (bar) **1**



Selected state: 2286.01 K; CO extent 0.11117 mol; H_2 extent 0.03649 mol per mole of methane. Dissociation lowers the complete-conversion estimate by 106.07 K. Scaled equilibrium/energy residual 1.8e-13. Constant heat capacities and six ideal-gas species, as in the notebook.

Continue in the OneSimulation notebook · Download its Python model

OneSimulation experiment · Air ratio and total pressure change the temperature of the six-species adiabatic equilibrium mixture

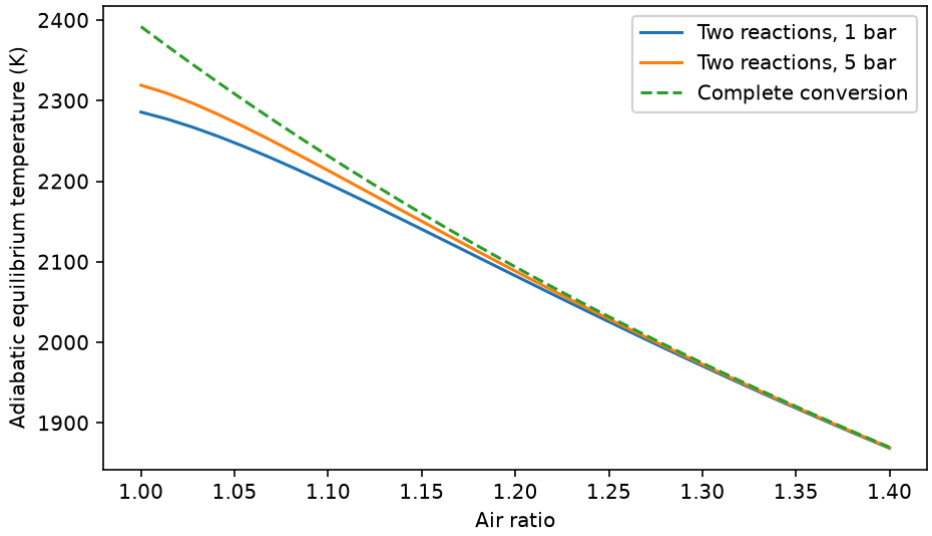
Run the notebook with OneSimulation 0.2.0, NumPy, SciPy and Matplotlib. Change the input data and compare your run with the recorded result.

```
"""Solve a two-reaction, six-product adiabatic equilibrium model for a methane-air stream."""
import numpy as np
import matplotlib.pyplot as plt
from onesim.numerics.nonlinear import LeastSquares

# Species order: CO2, CO, H2O(g), H2, O2, N2. Reference state 298.15 K and 1 bar.
reference = 298.15
gas_constant = 8.314462618  # J/(mol K)
hf = 1000*np.array([-393.52, -110.53, -241.83, 0.0, 0.0, 0.0])  # J/mol
entropy_ref = np.array([213.79,197.66,188.84,130.68,205.15,191.61])  # J/(mol K)
cp = np.array([51.,33.,42.,31.,35.,33.])  # constant teaching-model mean heat capacities
nu = np.array([[-1.,1.,0.,0.,0.5,0.], [0.,0.,-1.,1.,0.5,0.]])
element_counts = np.array([[1,1,0,0,0,0], [0,0,2,2,0,0],
                           [2,1,1,0,2,0], [0,0,0,0,0,2]])
inlet_enthalpy = -74870.0  # J per mole methane, with all reactants at 298.15 K

def equilibrate(air_ratio, pressure_bar):
    complete = np.array([1.,0.,2.,0.,2*(air_ratio-1),7.52*air_ratio])
    def residual(state):
        extents, temperature = state[:2],state[2]
        moles = complete + extents @ nu
        enthalpies = hf + cp*(temperature-reference)
        entropies = entropy_ref + cp*np.log(temperature/reference)
        log_k = -(nu @ (enthalpies-temperature*entropies))/(gas_constant*temperature)
        log_q = nu @ np.log(pressure_bar*moles/moles.sum())
        return np.r_[log_q-log_k, (moles @ enthalpies-inlet_enthalpy)/1e5]
    optimum = LeastSquares(options={"gtol":1e-13,"xtol":1e-13,"ftol":1e-13}).solve(
        residual,[0.06,0.04,2200.0], bounds=([1e-12,1e-12,500.],[1-1e-12,2-1e-12,3500.]))
    assert np.max(np.abs(optimum.fun)) < 1e-8
    moles = complete + optimum.x[:2] @ nu
    assert np.allclose(element_counts @ moles, element_counts @ complete)
    return optimum.x, moles

for pressure in [1.,5.]:
    result, moles = equilibrate(1.0, pressure)
    print(f"{pressure:.0f} bar: T = {result[2]:.2f} K, extents = {result[0]:.5f}, {result[1]:.5f}
mol")
    print("product moles",np.round(moles,5))
ratios = np.linspace(1.0, 1.4, 30)
fig,ax = plt.subplots(figsize=(6.4,3.7),layout="constrained")
for pressure in [1.,5.]:
    temperature = [equilibrate(r,pressure)[0][2] for r in ratios]
    ax.plot(ratios,temperature,label=f"Two reactions, {pressure:g} bar")
complete_temp = reference + 802310/(51+2*42+2*(ratios-1)*35+7.52*ratios*33)
ax.plot(ratios,complete_temp,"--",label="Complete conversion")
ax.set(xlabel="Air ratio",ylabel="Adiabatic equilibrium temperature (K)")
ax.legend()
plt.show()
```

Reference run

```
1 bar: T = 2286.01 K, extents = 0.11117, 0.03649 mol
product moles [0.88883 0.11117 1.96351 0.03649 0.07383 7.52   ]
```

```
5 bar: T = 2319.34 K, extents = 0.07711, 0.02403 mol
product moles [0.92289 0.07711 1.97597 0.02403 0.05057 7.52   ]
```



Air ratio and total pressure change the temperature of the six-species adiabatic equilibrium mixture

Download editable notebook · Download Python script

At 5 bar the computed extents are 0.07711 and 0.02403 mol. Reconstruct the oxygen amount and explain why the new temperature, 2319.34 K, is higher.

▼ **Worked solution**

The released oxygen is $(0.07711 + 0.02403)/2 = 0.05057$ mol. Both reactions increase the mole count. At higher pressure their equilibrium extents shrink, leaving more energy as sensible enthalpy. The difference between 2286 and 2319 K follows from the model, not a direct pressure dependence of ideal-gas species enthalpy.

How could this model approach a thermochemical-table calculation?

Replace the constant heat capacities by temperature-dependent species functions with a common reference, then enlarge the allowed product set to include radical species when required. The elemental constraints and energy residual still apply. The present six-species answer is reproducible for its named property law.

Reference gas data · NIST carbon dioxide · carbon monoxide · water · hydrogen · oxygen · nitrogen · methane.

Exercises on model closure

At $\lambda = 1.3$ preheat all feed gases so their combined sensible inlet enthalpy rises by 20 kJ per mole of methane. Estimate the temperature change.

▼ **Worked solution**

Holding the product composition and heat capacities fixed gives $20000 / 478.61 = 41.8$ K of additional temperature rise. At equilibrium the reaction extent changes and this simple ratio is only an initial estimate.

In an ideal-gas equilibrium calculation, increase total pressure while holding T and composition momentarily fixed. Which way does the dissociation quotient move?

▼ **Worked solution**

The net mole change is +1/2, so the quotient scales as $p^{1/2}$. It rises above its original equilibrium value, favouring recombination. The coupled adiabatic temperature must then be solved again.

Further reading · Çengel, Boles and Kanoğlu, Thermodynamics: An Engineering Approach, 10th ed., Chapters 16–17.

Thermal processes and transport in engineering · 7. How can a flow measurement determine a wall temperature?

Move from a heat-transfer coefficient supplied as data to one inferred from geometry and flow. Compare a turbulent pipe, a laminar external boundary layer and a plate cooled by buoyancy.

Distinguish a defining relation from a predictive law

At an impermeable wall heat crosses the fluid immediately adjacent to the solid by molecular conduction. Define h by equating that flux to the observable temperature difference. In a tube use the bulk mean temperature from the enthalpy flux. On an external plate use the free-stream temperature.

$$q'' = -k\,\partial_n T|_w = h(T_w - T_b),$$

$$Nu_L = hL/k,$$

$$Re_L = UL/\nu,$$

$$Pr = \nu/\alpha.$$

Nondimensionalising the advective and diffusive terms of the thermal energy equation produces $Re\,Pr$. Geometry, boundary conditions, momentum transport and entrance development decide the actual dependence of Nusselt number. The exponents fitted to turbulent data do not follow from dimensional analysis alone.

A heated smooth tube with developed turbulent flow

For a smooth, sufficiently long, fully developed turbulent tube with moderate property changes, a common first estimate for heating the fluid is $Nu_D = 0.023 Re_D^{0.8} Pr^{0.4}$. Use it here within its usual range, roughly $Re_D \gtrsim 10000$ and $0.7 \lesssim Pr \lesssim 160$. Check that the physical setting and temperature convention match before comparing correlations. For laminar fully developed circular flow the exact constant-wall-temperature and constant-wall-flux limits instead give 3.66 and 4.36, respectively.

At $Re_D = 30000$, $Pr = 5$, $D = 0.020$ m, $k = 0.60$ W/(m K) and imposed $q'' = 9000$ W/m², the turbulent estimate gives $Nu = 167.1$, $h = 5013$ W/(m² K) and $T_w - T_b = 1.80$ K. That temperature rise is small enough that sensor offsets matter. The following notebook fits one prefactor to deliberately perturbed synthetic readings and tests whether the residual size is compatible with thermometer precision.

OneSimulation experiment · Measured wall superheat decreases with flow rate in a heated tube

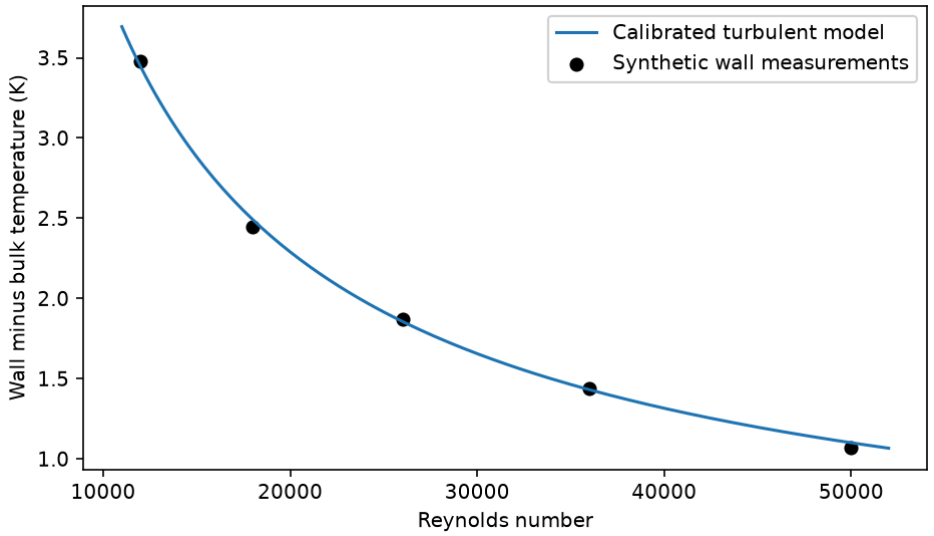
Run the notebook with OneSimulation 0.2.0, NumPy, SciPy and Matplotlib. Change the input data and compare your run with the recorded result.

```
"""Estimate a turbulent tube coefficient using synthetic wall temperature readings."""
import numpy as np
import matplotlib.pyplot as plt
from onesim.numerics.nonlinear import LeastSquares
```

```
reynolds = np.array([12000., 18000., 26000., 36000., 50000.])
q, diameter, conductivity, prandtl = 9000.0, 0.02, 0.60, 5.0
def superheat(coefficient, re):
    nusselt = coefficient*re**0.8*prandtl**0.4
    return q*diameter/(conductivity*nusselt)

readings = superheat(0.025, reynolds)+np.array([0.04,-0.04,0.02,0.01,-0.03])
fit = LeastSquares().solve(lambda c: superheat(c[0], reynolds)-readings,
                           [0.023], bounds=([0.005],[0.08]))
assert fit.success
coefficient = fit.x[0]
print(f"fitted prefactor = {coefficient:.5f}")
print(f"RMS temperature residual = {np.sqrt(np.mean(fit.fun**2)):.4f} K")
print(f"predicted rise at Re=30000 = {superheat(coefficient, 30000):.3f} K")
```

Reference run

```
fitted prefactor = 0.02494
RMS temperature residual = 0.0299 K
predicted rise at Re=30000 = 1.655 K
```



Measured wall superheat decreases with flow rate in a heated tube

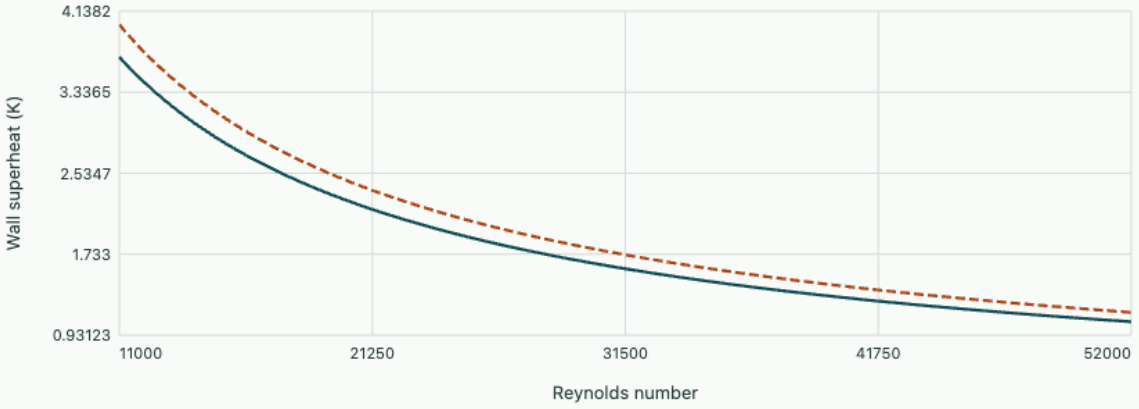
Interactive calculation · Change a turbulent heat-transfer closure

Vary heat flux, Prandtl number and correlation prefactor. The fixed synthetic calibration measurements in the notebook determine one prefactor; here you can inspect its predictive consequences.

Open this live calculation

Try the calculation

Heat flux (W/m²)	9000
Correlation prefactor	0.025
Prandtl number	5



At Re30000, predicted wall superheat 1.652 K. Tube diameter20 mm and conductivity0.60 W/(m K) are fixed; use the notebook to fit the prefactor from temperature data.

Download editable notebook · Download Python script

Changing the flow geometry changes the correlation

For a smooth isothermal plate with a laminar external boundary layer starting at the leading edge and moderate Prandtl number, use the area-mean value $\overline{Nu}_L = 0.664 Re_L^{1/2} Pr^{1/3}$. A local correlation at the trailing edge has half the prefactor: local and average heat transfer answer different measurement questions. For air with $\nu = 1.6 \times 10^{-5}$ m²/s, $Pr = 0.71$, $k = 0.026$ W/(m K), $L = 0.20$ m and $U = 2$ m/s, the mean result is $h = 12.18$ W/(m² K). At 4 m/s it becomes 17.22 W/(m² K).

For a still fluid by a vertical plate the motion instead follows from thermal expansion. Buoyant acceleration is of order $g\beta\Delta T$. Balancing it against thermal and momentum diffusion supplies $Ra_L = g\beta|T_w - T_\infty|L^3/(\nu\alpha)$. A vertical isothermal plate correlation across laminar and turbulent regimes is

$$\overline{Nu}_L = \left[0.825 + \frac{0.387 Ra_L^{1/6}}{[1 + (0.492/Pr)^{9/16}]^{8/27}}\right]^2.$$

With the same length, $\Delta T = 20$ K, $\beta = 1/300$ K⁻¹ and $\alpha = 2.25 \times 10^{-5}$ m²/s, $Ra_L = 1.45 \times 10^{7}$ and $h = 4.51$ W/(m² K). Evaluate properties consistently, usually at a representative film temperature. If both buoyancy and forced motion matter, their orientation can aid or oppose the flow.

> **What changes when the wall temperature is unknown?**
>
> For imposed heat flux in free convection, the temperature difference appears both in h and in $q'' = h\Delta T$. Solve the scalar nonlinear balance with updated properties as warranted. Treating h as independent of wall temperature suppresses the coupling driving the motion.

Exercises on correlation selection

Double the pipe velocity with constant properties inside the turbulent range and unchanged imposed heat flux. How far does the wall rise above the fluid?

▼ **Worked solution**

The coefficient rises by $2^{0.8} = 1.741$. The 1.80 K rise falls to about 1.03 K. A fixed wall-temperature experiment would observe the opposite dependent variable, an increased heat rate.

A stationary vertical plate in air has a known 20 K excess temperature. Its height is doubled. May the original coefficient simply be reused?

▼ **Worked solution**

Rayleigh number grows eightfold. Reevaluate the stated correlation at that Rayleigh number and divide the new Nu by the doubled length when recovering h. The heat transfer area also changes if width stays fixed.

Further reading · Bergman and Lavine, Fundamentals of Heat and Mass Transfer, 8th ed., Chapters 6–9.

Thermal processes and transport in engineering · 8. Why can a heated surface suddenly become much hotter?

Trace the pool-boiling curve of a saturated liquid. Compare changing a heater's temperature with changing its electrical power and estimate the hydrodynamic scale at which nucleate boiling becomes unstable.

Read a boiling curve as a family of states

Plot heat flux q'' against excess wall temperature $\Delta T_e = T_w - T_{sat}(p)$ for a saturated liquid at fixed pressure. Natural convection carries small heating rates. When active cavities support vapour embryos, rising bubbles renew liquid at the wall. Nucleate boiling heats the liquid much more effectively and the slope is usually steep. At the critical heat flux vapour structures begin to impede liquid return. The transition branch can have negative slope. Farther along, a continuous vapour film separates liquid from the surface and heat crosses a poorly conducting layer.

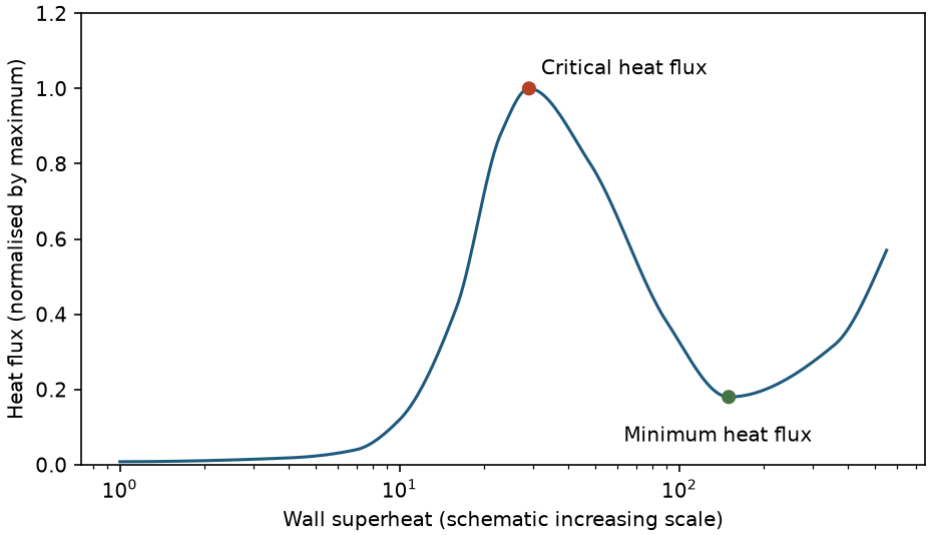
A boiling curve depends on wetting, surface finish, pressure, orientation and heating history. Saturation temperature gives the equilibrium temperature at a given pressure. It is insufficient to determine the local wall superheat or boiling branch. A condensing film and a boiling boundary involve the same latent heat but different interfacial dynamics.

Balance capillary and buoyant lengths

At a gravity-disturbed interface a pressure difference of order σ/ℓ competes with a hydrostatic difference $(\rho_l - \rho_v)g\ell$. Their equality gives capillary length $\ell_c = [\sigma/(g(\rho_l - \rho_v))]^{1/2}$. A vapour inertial stress of order $\rho_v u_v^2$ matching the same pressure sets a velocity scale. Multiplication by $\rho_v h_{fg}$ converts that speed to a latent-heat flux scale. A hydrodynamic model for saturated pool boiling on a large horizontal upward-facing plate gives the empirical/geometric coefficient $C \simeq 0.131$:

$$q''_{CHF} = C h_{fg} \sqrt{\rho_v}\, [\sigma g(\rho_l - \rho_v)]^{1/4}.$$

For saturated water near atmospheric pressure take $h_{fg} = 2.257$ MJ/kg, $\rho_l = 958$, $\rho_v = 0.598$ kg/m³ and $\sigma = 0.0589$ N/m. The scale predicts about 1.11 MW/m². This order-of-magnitude estimate pertains to a specific large-surface model, not a universal limit for small heaters or forced flow.



Schematic saturated pool-boiling curve. The marked extrema bound the negative-slope branch. Its coordinates illustrate the topology and are not experimental data.

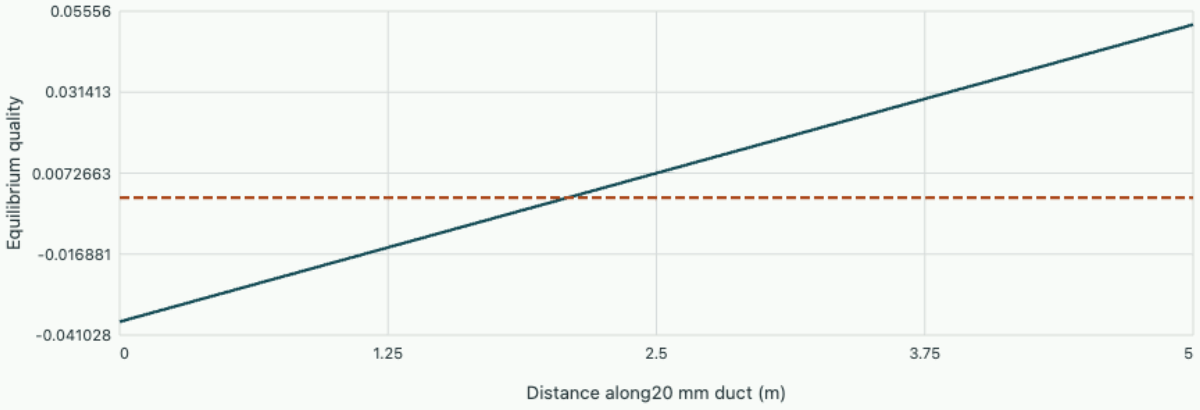
Interactive calculation · Locate bulk saturation in a heated tube

Increase heating or decrease mass flux and find where equilibrium quality first reaches zero. This energy balance does not predict the wall-boiling onset or dryout threshold.

Open this live calculation

Try the calculation

Heat flux (kW/m²)	100
Mass flux (kg/m² s)	500
Inlet subcooling (K)	20



Equilibrium quality Bulk saturation

Bulk saturation begins at 2.090 m. Negative quality denotes sensible subcooling, not a negative vapour fraction. This energy balance alone predicts neither flow pattern nor dryout.

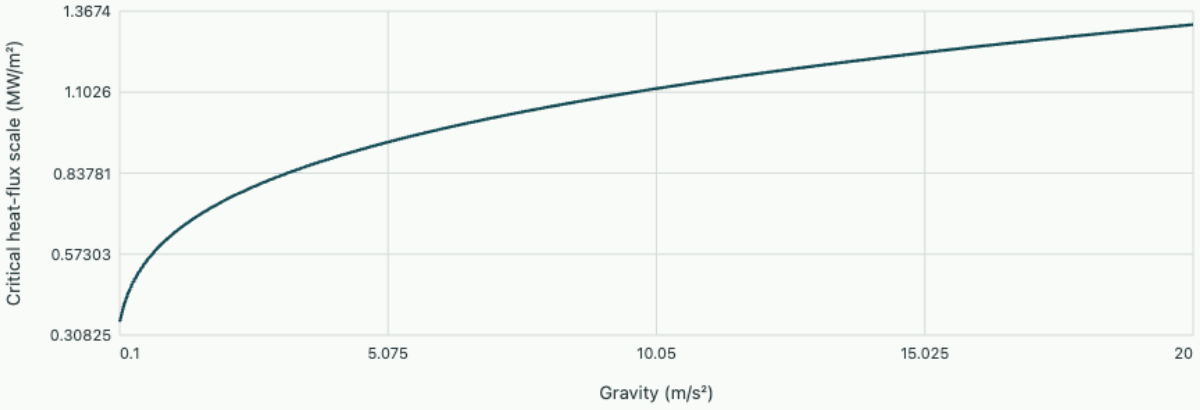
Interactive calculation · Explore the critical-flux scaling

Vary one property at a time and compare the gravity dependence. The property inputs are independent scaling experiments; they are not a thermodynamically coupled fluid-property table.

Open this live calculation

Try the calculation

Latent heat (MJ/kg)	2.25
Surface tension (N/m)	0.059
Vapour density (kg/m³)	0.6



Large horizontal surface model

At terrestrial gravity the scale is 1.108 MW/m². Liquid density958 kg/m³ is held fixed; property sliders explore scaling, not a consistent pressure-dependent fluid table.

Explain the jump using the heater energy balance

Write $C_A dT_w/dt = q''_{in} - q''_{out}(T_w)$ for heat capacity per area C_A. Perturb a steady point by δT at fixed input. Then $C_A\, d\delta T/dt = -q'_{out}\delta T$. A positive local slope suppresses a perturbation and a negative slope amplifies it. Consequently an imposed-power sweep can jump past much of the transition branch. A temperature-controlled experiment can in principle visit states that constant-power control cannot stabilize.

Does a target heat flux determine a unique wall temperature?

A horizontal line on a nonmonotone boiling curve can intersect more than one branch. Surface state, control mode and the direction of the preceding sweep decide which intersections can be reached.

Follow a flowing liquid from subcooling into a two-phase tube

For a uniformly heated circular duct, wall area per axial length is πD and the mass flow is $G\pi D^2/4$ with mass flux G. Neglect pressure change, kinetic-energy change, loss to surroundings and inlet vapour. A differential steady energy balance yields $dh/dz = 4q''/(GD)$. Define the equilibrium quality $x_{eq} = (h - h_f)/h_{fg}$ at the local pressure. Negative values measure subcooling. A positive value describes the mass vapour fraction when both phases are in local thermal equilibrium.

$$z_{sat} = \frac{GDc_{p,l}(T_{sat} - T_{in})}{4q''},$$

$$x_{eq}(z) = \frac{4q''(z - z_{sat})}{GDh_{fg}}.$$

Wall bubbles may form before the bulk reaches saturation and condense in the colder core. Downstream, bubbly or slug flow can give way to an annular liquid film with droplets in a fast vapour core. Dryout of that film and departure from nucleate boiling in a subcooled region are distinct routes to a sharp rise in wall temperature. Orientation, mass flux, pressure, quality and heating history influence the transitions.

Use boiling and momentum scales before choosing a closure

The boiling number $Bo_q = q''/(Gh_{fg})$ measures vapour generation relative to mass transport. The liquid-only Froude number $Fr_{lo} = G^2/(\rho_l^2 gD)$ compares inertia and gravity. The liquid-only Weber number $We_{lo} = G^2 D/(\rho_l \sigma)$ compares inertia and capillarity. Confinement number $Co = D^{-1}\sqrt{\sigma/[g(\rho_l - \rho_v)]}$ measures capillary length relative to diameter. These definitions specify their reference density. A regime map may instead use superficial velocities $j_g = Gx/\rho_v$ and $j_l = G(1-x)/\rho_l$. Enter that map with its own orientation, geometry and definitions.

For a vertical water tube take $D = 16$ mm, $G = 300$ kg/(m² s), $q'' = 42$ kW/m², 8 K inlet subcooling and representative 1 MPa saturation data $\rho_l = 887$, $\rho_v = 5.15$ kg/m³, $h_{fg} = 2.015$ MJ/kg, $c_{p,l} = 4.4$ kJ/(kg K), $\sigma = 0.042$ N/m. Bulk saturation occurs at 1.006 m. At 3 m the predicted quality is only 0.03464, yet $j_g = 2.018$ m/s and $j_l = 0.327$ m/s. The much smaller vapour density explains why a small mass fraction can dominate volumetric transport. The reference groups are $Bo_q = 6.95 \times 10^{-5}$, $Fr_{lo} = 0.729$, $We_{lo} = 38.7$ and $Co = 0.138$.

For a local illustrative closure suppose a validated superposition fit at this station supplies $h_{lo} = 4500$ W/(m² K), convective enhancement $F = 1.35$, pool reference $h_{nb} = 6500$ W/(m² K) and nucleate suppression $S = 0.60$. Then $h_{tp} = Fh_{lo} + Sh_{nb} = 9975$ W/(m² K) and the required superheat is $42000/9975 = 4.21$ K. Those coefficients are supplied station data. Extending them through flow-regime change would require a validated local closure. A single-phase fit cannot predict film dryout.

At what distance does the simple uniform-pressure model first predict quality 0.10? What additional evidence would justify calling that station a dryout location?

▼ **Worked solution**

The quality slope is $4q''/(GDh_{fg}) = 0.01737$ per metre. Adding $0.10/0.01737$ m to the saturation length gives 6.76 m. Dryout needs evidence of film depletion, measured wall-temperature excursion or a correlation validated for this pressure, mass flux, tube and heated length. A quality of 0.10 alone provides none of these.

Exercises on energy and stability

A saturated pool removes 200 kW/m² solely by evaporation. What is the associated mass flux of water vapour?

▼ **Worked solution**

The evaporation rate is $q''/h_{fg} = 200000/2257000 = 0.0886$ kg/(m² s). The fraction of area covered by vapour cannot be inferred from energy alone.

Near a transition point $q'_{out} = -4000$ W/(m² K) and $C_A = 2000$ J/(m² K). Predict the initial growth time of a small temperature error at fixed power.

▼ **Worked solution**

The linear growth exponent is $-q'_{out}/C_A = 2$ s⁻¹. An initial deviation grows by a factor e in 0.5 s until nonlinear changes of the heat-removal law become significant.

Further reading · Bergman and Lavine, Fundamentals of Heat and Mass Transfer, 8th ed., Chapter 10, especially sections 10.2–10.5.

Properties, equilibrium and phase change

Predict the sign of a response from a potential, then track what a measured fraction or coexistence curve says about the state of the material.

1. Equations of state and response functions
2. Heat, work and the limits of a cycle
3. Thermodynamic potentials and phase boundaries
4. Indistinguishable particles and quantum gases
5. Chemical equilibrium and reaction extent
6. How a pair potential changes the gas pressure
7. Why two coexisting phases share a tangent
8. When does a polar crystal change its preferred state?
9. Counting, absolute entropy and the entropy of mixing
10. Ensembles, thermodynamic potentials and measurable curvature
11. Pair correlations, the virial gas and a pressure measurement
12. Rotational and vibrational populations in a molecular gas
13. Phase nucleation, metastability and waiting-time statistics

Equations of state and response functions

Thermodynamics and statistical physics

Use compressibility and thermal expansion to predict what a pressure gauge will record when a material is heated in a closed vessel.

FROM FIRST PRINCIPLES

Derive ideal-gas pressure from impacts on a wall

Pressure is the normal force on a wall per unit area. Consider N classical molecules of mass m in a cubic box of side L. Neglect intermolecular forces and molecular volume. For an elastic collision at a wall perpendicular to the x axis, a molecule transfers normal momentum $2m|v_x|$. It visits the same wall once in a round trip of duration $2L/|v_x|$. Its time-averaged force there is

$$F_{x,1} = \frac{2m|v_x|}{2L/|v_x|} = \frac{mv_x^2}{L}.$$

Sum all contributions and divide by the wall area L^2. With $V = L^3$, equilibrium isotropy gives

$$\langle v_x^2 \rangle = \langle v^2 \rangle / 3.$$

Classical thermal equilibrium assigns mean energy $k_B T/2$ to each quadratic velocity component. These inputs give

$$p = \frac{Nm}{V}\langle v_x^2 \rangle = \frac{Nm}{3V}\langle v^2 \rangle,$$
$$pV = Nk_B T.$$

This equation of state connects three directly measurable quantities at fixed particle number. The kinetic calculation also identifies where a real gas can differ: intermolecular forces transfer momentum, and dense molecules occupy an appreciable fraction of the volume.

DERIVATION AND APPLICATIONS

Connect a sound-speed measurement to two compressibilities

A slowly compressed sample has time to exchange heat with its surroundings. A small sound pulse usually compresses a fluid element faster than heat can diffuse across a wavelength. The elastic response relevant to sound is then nearly adiabatic. Consider a homogeneous single-phase fluid at equilibrium and a small-amplitude pulse. Assume negligible viscosity and thermal diffusion over one oscillation, so the specific entropy of a moving element remains constant.

Let v be volume per unit mass, c_p the specific heat at constant pressure, $\alpha = v^{-1}(\partial v/\partial T)_p$, and $\kappa_T = -v^{-1}(\partial v/\partial p)_T$. From the specific Gibbs differential $dg = -s\,dT + v\,dp$, equality of mixed derivatives gives

$$(\partial s/\partial p)_T = -(\partial v/\partial T)_p.$$

The entropy differential determines a fluid element's temperature change under isentropic compression,

$$ds = \frac{c_p}{T}dT - v\alpha\,dp = 0, \qquad \left(\frac{\partial T}{\partial p}\right)_s = \frac{Tv\alpha}{c_p}.$$

Insert this derivative into the volume response $dv/v = \alpha\,dT - \kappa_T\,dp$. The resulting isentropic compressibility is

$$\kappa_s = -\frac{1}{v}\left(\frac{\partial v}{\partial p}\right)_s = \kappa_T - \frac{Tv\alpha^2}{c_p} = \kappa_T\frac{c_v}{c_p}.$$

The last step uses the heat-capacity difference for a simple compressible material. For $\rho = 1/v$, a density perturbation gives

$$p' = (\partial p/\partial \rho)_s \rho'.$$

Linear momentum balance and continuity then combine into a wave equation with speed c,

$$\partial_t^2 \rho' = c^2 \nabla^2 \rho',$$
$$c^2 = \left(\frac{\partial p}{\partial \rho}\right)_s = \frac{1}{\rho\kappa_s}.$$

Adiabatic heating adds a pressure response on compression for an ordinary expanding fluid. The two compressibilities coincide where thermal expansion vanishes.

WORKED EXAMPLE

Predict the speed in an ideal gas

Change the inputs and investigate this derivation.

For dry model air at $p = 1.00 \times 10^5\,\mathrm{Pa}$, take $\rho = 1.20\,\mathrm{kg/m^3}$ and $c_p/c_v = 1.40$. Since $\kappa_T = 1/p$, sound propagates at $\sqrt{1.40p/\rho} = 342\,\mathrm{m/s}$. An isothermal model gives $289\,\mathrm{m/s}$. Comparing the two speeds measures the square root of the heat-capacity ratio under these gas assumptions.

SECOND APPLICATION

Infer a heat-capacity ratio for a liquid

Change the inputs and investigate this derivation.

An ultrasonic pulse travels at $1480\,\mathrm{m/s}$ through a liquid of density $998\,\mathrm{kg/m^3}$. It gives

$$\kappa_s = 1/(998 \times 1480^2) = 4.57 \times 10^{-10}\,\mathrm{Pa^{-1}}.$$

A slow isothermal experiment gives $4.65 \times 10^{-10}\,\mathrm{Pa^{-1}}$. Their ratio yields $c_p/c_v = 1.017$, which explains why the choice of compressibility produces a much smaller acoustic correction for this liquid.

Try it · Sound compressibility

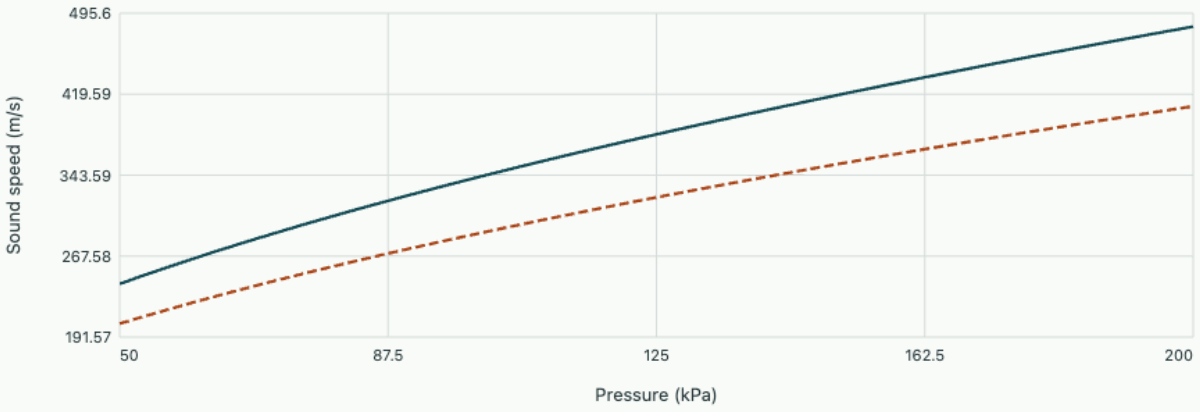
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Density (kg/m³) **1.2**

Heat-capacity ratio cp/cv **1.4**



For ideal gas κT=1/p, κs=κT/γ. At 100 kPa the speed is 341.57 m/s. A liquid instead requires measured compressibility.

Decide which variables specify a state

An equilibrium sample of one simple phase and fixed composition can be described by pressure, volume and temperature related by an equation of state. A piston experiment can prescribe pressure while changing volume; a sealed rigid vessel instead fixes volume. The distinction matters because the same temperature change can lead to expansion in one apparatus and a large pressure increase in the other. Assume quasistatic measurements so that a single pressure and temperature describe the sample.

An equation of state is a constitutive statement. The ideal-gas relation $pV = nRT$ treats particles as dilute enough for their mutual interactions to have negligible effect on the measured pressure. An incompressible-liquid approximation often gives an adequate volume balance at a chosen pressure but can make a constrained heating problem ill posed if applied literally. Both small thermal expansion and small but finite compressibility may be needed to compute the pressure response.

Describe a neighbourhood of equilibrium

Treat volume as a differentiable function $V(T, p)$ near a reference state. Its total differential contains the two independent slopes $(\partial V/\partial T)_p$ and $(\partial V/\partial p)_T$. Define the two responses at their respective fixed controls,

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p,$$
$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T.$$

Ordinary mechanically stable material contracts when pressure increases, which gives positive isothermal compressibility. Inserting both definitions into the volume differential gives

$$\frac{dV}{V} = \alpha\, dT - \kappa_T\, dp.$$

Both terms are dimensionless. Expansion has units of inverse kelvin, while compressibility has units of inverse pascal. For fixed volume the pressure slope is

$$(\partial p/\partial T)_V = \alpha/\kappa_T.$$

This relation is a consequence of using the same volume differential. It does not assume a particular molecular model and it remains informative for a liquid for which the ideal-gas equation would give a poor answer.

For an ideal gas at fixed amount, differentiating $V = nRT/p$ gives

$$\alpha = 1/T,$$
$$\kappa_T = 1/p.$$

Their ratio is p/T, as follows independently from $p = nRT/V$. At 300 K and 100 kPa, warming a sealed ideal gas by 3.0 K raises the pressure by 1.0 kPa. At fixed pressure instead the relative volume increase is one percent.

Background · Why do slow compression tests and sound measure different moduli?

For a sealed ideal gas in thermal contact with a bath, a sufficiently slow compression allows heat to escape and fixes T. Then pV is constant. During a short, nearly reversible acoustic oscillation in the bulk there can be too little time for heat to leave the compressed fluid element. For this adiabatic limit pV^{γ} is constant. Differentiate at the same initial pressure.

$$K_T = -V\left(\frac{\partial p}{\partial V}\right)_T = p,$$

$$K_S = -V\left(\frac{\partial p}{\partial V}\right)_S = \gamma p.$$

At one bar and $\gamma = 1.4$ the moduli are 100 and 140 kPa. Each measures the response at the same equilibrium state, under a different thermal constraint. In a narrow tube heat can diffuse to the wall during a cycle. The appropriate limit depends on whether the thermal diffusion length is small or large compared with the tube radius.

Separate intensive, extensive and measured responses

Doubling the amount of material at the same temperature and pressure doubles volume but leaves the two normalised coefficients unchanged. Such intensive coefficients can be compared across samples. A vessel’s own volume, however, can respond to both pressure and temperature. If its expansion coefficients are appreciable, the measured pressure slope describes the coupled sample and vessel. One cannot simply substitute the unconstrained expansion of the sample into a rigid-vessel formula.

To see this, let the vessel’s fractional volume response be $\alpha_v\,dT + C_v^{\mathrm{vessel}}dp$, where its mechanical compliance is positive for an internal pressure increase. Equal fractional changes of vessel and sample give

$$(\alpha - \alpha_v)dT = (\kappa_T + C_v^{\mathrm{vessel}})dp.$$

A compliant wall reduces the pressure rise, and a vessel that expands just as much as the sample can nearly cancel the temperature signal. Here the symbol for vessel compliance denotes a mechanical quantity and has different units from heat capacity.

Compare liquid and container data

Take a representative liquid with $\alpha = 2.5 \times 10^{-4}$ K⁻¹ and $\kappa_T = 5.0 \times 10^{-10}$ Pa⁻¹ near its operating point. In an ideally rigid container, a 2.0 K increase gives $\Delta p \simeq 1.0$ MPa. The unconstrained volume

change at fixed pressure would be only 0.050%. This combination explains how a small fractional expansion can produce a significant pressure response when the fluid is only weakly compressible.

If the container has $\alpha_v = 5.0 \times 10^{-5}$ K⁻¹ and $C_v^{\text{vessel}} = 3.0 \times 10^{-10}$ Pa⁻¹, the same warming gives about 0.50 MPa. Inverting a measured slope without the vessel data could thus give a misleading estimate of compressibility. These are local linear predictions. Over a broad pressure or temperature range the coefficients must be evaluated along the path, and crossing a phase boundary can invalidate the single-phase differential altogether.

Convert an isobaric calorimeter reading into an isochoric one

A sample can absorb different amounts of heat per kelvin depending on which variable is held fixed. For a fixed amount of a simple, single-phase material in equilibrium,

$$C_V = T(\partial S/\partial T)_V,$$

$$C_p = T(\partial S/\partial T)_p.$$

Apply the chain rule to $S(T, V)$ along the constant-pressure path. Subtracting the two derivatives leaves the effect of expansion. From $dF = -S\,dT - p\,dV$ equality of mixed derivatives gives

$$(\partial S/\partial V)_T = (\partial p/\partial T)_V.$$

Therefore

$$C_p - C_V = T\left(\frac{\partial p}{\partial T}\right)_V \left(\frac{\partial V}{\partial T}\right)_p = \frac{TV\alpha^2}{\kappa_T}.$$

Here α is the volume expansion coefficient and κ_T the isothermal compressibility. The last equality uses $dV/V = \alpha dT - \kappa_T dp$. For a mechanically stable material $\kappa_T > 0$, so the difference is nonnegative even for a material that contracts when warmed. For one mole of ideal gas, inserting $\alpha = 1/T$, $\kappa_T = 1/p$ and $pV = RT$ gives

$$C_p - C_V = R.$$

At 300 K a solid has molar volume 1.20×10^{-5} m³/mol, volumetric expansion coefficient 6.00×10^{-5} K⁻¹ and bulk modulus $B_T = 1/\kappa_T = 100$ GPa. The molar difference is $TV_mB_T\alpha^2 = 1.30$ J/(mol K). A constant-pressure measurement of 25.0 J/(mol K) thus implies $C_{V,m} = 23.7$ J/(mol K). For a cubic solid the linear expansion coefficient is approximately one third the volume coefficient. Using it directly in this formula would underestimate the correction ninefold. This conversion matters when comparing a tabulated calorimetric result with a vibrational model calculated at fixed volume.

Predict the temperature change of a throttled gas

Some pressure changes take place in flowing rather than closed matter. Consider a well-insulated porous plug that sustains a pressure difference between two steady streams of the same single-phase gas. There is no shaft work and inlet and outlet speeds and heights have negligible differences in kinetic and gravitational energy. Flow work at each end is already included in the molar enthalpy $h = u + pv$. The steady energy balance then gives

$$h_{\mathrm{out}} = h_{\mathrm{in}}.$$

A pressure drop through the restriction can nevertheless change temperature if the gas is nonideal.

From $dh = T\,ds + v\,dp$ and the Maxwell relation $(\partial s/\partial p)_T = -(\partial v/\partial T)_p$, write the enthalpy differential in terms of T, p. The coefficient of dT at fixed pressure is the molar heat capacity c_p. Setting $dh = 0$ gives the Joule–Thomson coefficient

$$\mu_{\mathrm{JT}} = \left(\frac{\partial T}{\partial p}\right)_h = \frac{T(\partial v/\partial T)_p - v}{c_p} = \frac{v(T\alpha - 1)}{c_p}.$$

For an ideal gas $\alpha = 1/T$, so an isenthalpic throttle produces no temperature change. Isentropic expansion in a turbine is a different process. A throttle's pressure drop is irreversible; constant enthalpy follows here from the steady energy balance and does not imply constant entropy.

Estimate a small temperature drop from gas data

Near a chosen operating point let the low-pressure molar volume be $v = RT/p + B(T)$, retaining a second virial coefficient $B = b - a/(RT)$. Here a and b characterize an approximate real-gas equation of state. The ideal-volume term cancels in $T(\partial v/\partial T)_p - v$. The remaining numerator is $TB' - B = 2a/(RT) - b$. At fixed local heat capacity, attractions and excluded volume therefore compete in determining the sign of the temperature response.

Take $a = 0.140$ Pa m⁶/mol², $b = 4.00 \times 10^{-5}$ m³/mol, $T = 300$ K and $c_p = 29.0$ J/(mol K). Then $2a/(RT) = 1.12 \times 10^{-4}$ m³/mol and $\mu_{\mathrm{JT}} = 2.49 \times 10^{-6}$ K/Pa, or 0.249 K/bar. A pressure change of -2.0 bar predicts temperature change -0.50 K if the properties remain close to these values. The leading model changes sign at $T = 2a/(Rb) = 842$ K. This inversion estimate should not be extrapolated into a dense or multiphase region. A proper finite pressure drop integrates the measured local coefficient along an isenthalp. Recording both enthalpy balance and inlet-to-outlet entropy increase offers an independent test of whether unwanted heat transfer through the apparatus affected the result.

The entropy change of the throttled stream can be checked without a calorimeter. Along an isenthalp the Gibbs relation gives

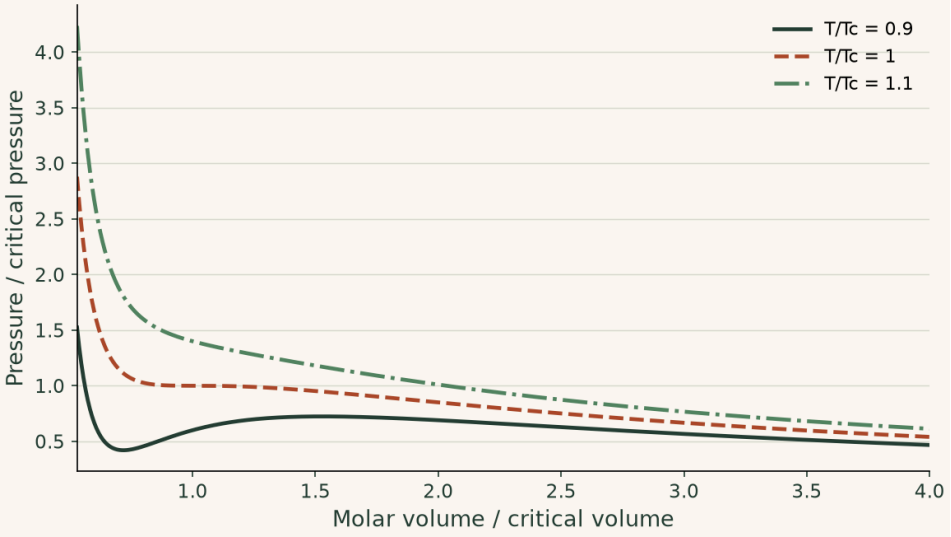
$$ds = -v\,dp/T.$$

Since molar volume and temperature are positive, any pressure decrease raises the molar entropy. Even when an ideal gas shows zero temperature change, its pressure decrease leaves a positive entropy difference. This reveals precisely why absence of observed cooling is insufficient evidence for thermodynamic reversibility. A pressure reading, inlet temperature and appropriate volume data together supply a sign test independently of the small temperature slope used to estimate the Joule–Thomson coefficient.

An independent heat-leak estimate is particularly important near inversion, where a small parasitic heat flow can dominate the intrinsic temperature change.

Find the mechanically unstable part of an isotherm

Use the reduced van der Waals equation $p = 8t/(3v - 1) - 3/v^2$. At fixed temperature a stable homogeneous phase must have positive isothermal compressibility, which requires a negative slope on this graph. Locate any interval in which the model reverses that slope.



Below the critical temperature, larger pressures along the rising segment of the formal loop correspond to larger volumes. This implies negative isothermal compressibility. The homogeneous state is mechanically unstable there, and equilibrium coexistence replaces part of the curve with a constant-pressure tie line.

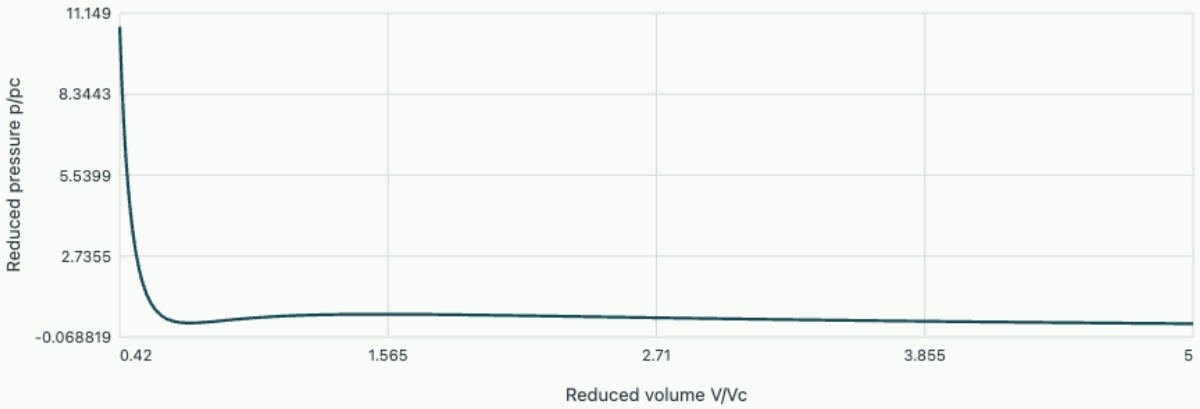
Try it · Vdw

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 **0.9**



Homogeneous van der Waals isotherm

The derivative is $-24Tr/(3Vr-1)^2 + 6/Vr^3$. A positive derivative means negative compressibility. Below Tc equilibrium requires a coexistence construction; the displayed loop is the homogeneous equation.

▶ **Read values from the plot**

Heat, work and the limits of a cycle

Thermodynamics and statistical physics

Track changes of internal energy around a gas cycle and derive the temperature bound on the efficiency of a reversible engine.

FROM FIRST PRINCIPLES

Derive the reversible engine efficiency from two balances

A working substance in a cyclic engine returns to its initial state. Its internal energy is a state function, so the energy gained from a hot reservoir must equal the useful work plus the heat delivered to a cold reservoir. Let Q_h, Q_c be positive heat magnitudes and W positive work delivered by the engine. Energy conservation gives

$$W = Q_h - Q_c,$$
$$\eta = \frac{W}{Q_h} = 1 - \frac{Q_c}{Q_h}.$$

To determine the smallest possible heat rejection, apply the second law. Reservoirs at constant absolute temperatures $T_h > T_c$ lose entropy Q_h/T_h and gain entropy Q_c/T_c respectively. The engine itself returns to its original entropy. The total entropy production is therefore

$$S_{\mathrm{gen}} = \frac{Q_c}{T_c} - \frac{Q_h}{T_h} \geq 0, \qquad \eta \leq 1 - \frac{T_c}{T_h}.$$

A reversible cycle attains equality. Friction or heat transfer across finite temperature differences produces entropy and reduces efficiency at fixed reservoir temperatures. The ratio uses absolute temperatures because reservoir entropy changes are defined by reversible heat divided by thermodynamic temperature.

DERIVATION AND APPLICATIONS

Derive the efficiency of an idealized four-stroke gas cycle

A compression ratio predicts an efficiency for a model cycle only after the heat-transfer paths are specified. Consider a fixed amount of ideal gas with constant molar heat capacities and $\gamma = C_p/C_V > 1$. In the Otto model the gas undergoes reversible adiabatic compression from state 1 to 2, heat addition at fixed volume from 2 to 3, reversible adiabatic expansion from 3 to 4, then heat rejection at fixed volume back to 1. Use positive magnitudes Q_h, Q_c for the absorbed and rejected heat.

For either adiabat the first law gives

$$nC_V dT = -p\,dV = -nRT\,dV/V.$$

Separation and integration give

$$TV^{\gamma-1} = \text{constant}.$$

Introducing $r = V_1/V_2 = V_4/V_3 > 1$ sets both adiabatic temperature ratios,

$$\frac{T_2}{T_1} = \frac{T_3}{T_4} = r^{\gamma-1}.$$

The two fixed-volume steps do no boundary work. Their respective heats are internal energy changes, and cyclic energy conservation converts the heat ratio to efficiency,

$$Q_h = nC_V(T_3 - T_2),$$
$$Q_c = nC_V(T_4 - T_1),$$
$$\eta = 1 - \frac{Q_c}{Q_h}.$$

Substitute $T_3 = r^{\gamma-1}T_4$ and $T_2 = r^{\gamma-1}T_1$. Both temperature differences contain the same factor $T_4 - T_1$, which cancels and leaves

$$\eta = 1 - r^{1-\gamma}.$$

The gas absorbs and rejects heat over intervals of temperature. Its efficiency therefore depends on the chosen paths even with reversible adiabatic strokes. At $r = 1$ the pressure-volume loop collapses and produces no net work. Increasing the compression ratio also raises the temperature reached before heat addition, which can invalidate a constant-γ model.

WORKED EXAMPLE

Close a cycle with measured endpoint temperatures

Change the inputs and investigate this derivation.

For one mole with $r = 8.00$, $\gamma = 1.40$, $T_1 = 300\,\mathrm{K}$ and $T_3 = 1600\,\mathrm{K}$, the adiabatic factor is 2.297. Hence $T_2 = 689\,\mathrm{K}$, $T_4 = 696\,\mathrm{K}$, and $C_V = R/0.40 = 20.8\,\mathrm{J/(mol\,K)}$. The gas absorbs $18.9\,\mathrm{kJ}$ and rejects $8.24\,\mathrm{kJ}$, yielding $10.7\,\mathrm{kJ}$ of work and efficiency 0.565.

SECOND APPLICATION

Infer a compression ratio for a different working gas

Change the inputs and investigate this derivation.

A model monatomic gas has $\gamma = 5/3$. Achieving efficiency 0.600 requires

$$r = (1 - 0.600)^{-1/(\gamma - 1)} = 3.95.$$

Starting at $280\,\mathrm{K}$, its temperature at the end of compression is

$$280/(1 - 0.600) = 700\,\mathrm{K}.$$

A cycle whose hottest temperature were $650\,\mathrm{K}$ could not realize the specified heat-addition stroke with this initial state.

Try it · Otto

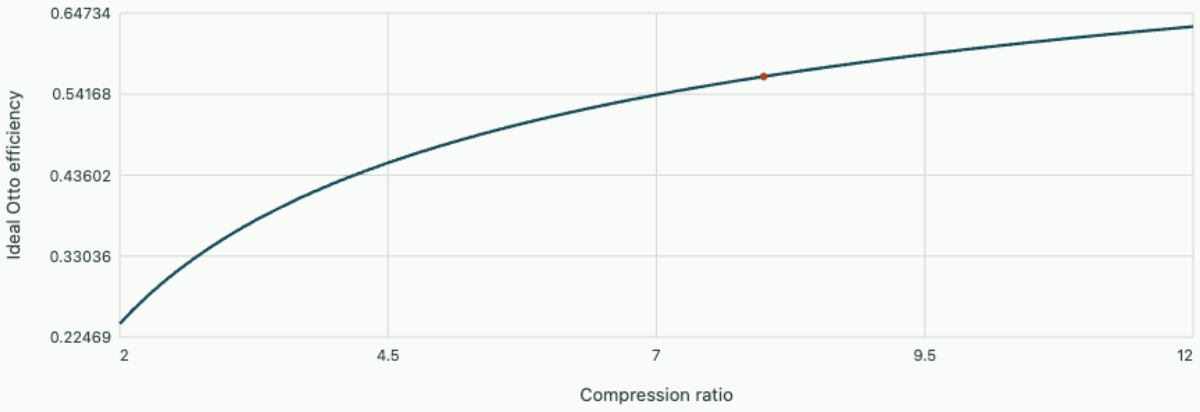
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Heat-capacity ratio γ **1.4**

Compression ratio **8**



Air-standard Otto cycle Selected ratio

Efficiency 1−r^(1−γ) = 0.56472. At initial 300 K, adiabatic compression reaches 689.22 K. Constant heat capacities and reversible adiabats are assumed.

Put signs on a physical boundary

Enclose a gas inside a movable cylinder. When the gas expands, it can raise an external weight. When heated at fixed volume, it can increase its temperature while doing no boundary work. To fix signs, take heat Q into the gas as positive and work W done by the gas as positive. The first law is

$$\Delta U = Q - W.$$

Internal energy is a state property; the other two quantities describe transfer during a process. A cyclic device returns to its initial state, so its net work equals its net heat input.

For a quasistatic expansion, a piston moving distance dx under external pressure p_{ext} receives work $p_{\mathrm{ext}} A\,dx = p_{\mathrm{ext}} dV$. In mechanical equilibrium p_{ext} approaches the gas pressure, giving the usual area under a pressure–volume curve. A suddenly expanding gas is not generally uniform or in mechanical equilibrium. Extending the equilibrium pressure formula to an unconstrained free expansion would predict work when no piston or other boundary force actually receives it.

Follow an ideal gas along two paths

For a calorically ideal gas, $dU = nC_{V,m}dT$ with constant molar heat capacity. Reversible isothermal expansion from V_1 to V_2 therefore gives zero internal energy change and $Q = W = nRT\ln(V_2/V_1)$. Revise the process so that the gas is thermally insulated but still expands reversibly. Then $nC_{V,m}dT = -p\,dV = -nRT\,dV/V$. Integrating and introducing $\gamma = C_{p,m}/C_{V,m} = 1 + R/C_{V,m}$ gives

$$TV^{\gamma-1} = \text{constant}.$$

Temperature now falls because the gas supplies work from its energy store.

The two paths can pass through a common starting state but they generally end at different pressures and temperatures for a given final volume. If one wants to compare work between identical endpoints one must specify a connecting step as well. This prevents a common misreading of a pressure–volume plot: an isotherm and an adiabat through the same point have different slopes and cannot serve as alternate paths to the same later point without additional processes.

Construct a reversible two-reservoir cycle

Let the gas expand isothermally at T_h, then expand adiabatically down to T_c, compress isothermally there and return adiabatically to the original state. If the first isotherm changes volume by a factor r, the two adiabatic relations force the cold compression to change volume by the reciprocal factor. The hot heat input is $Q_h = nRT_h\ln r$. The magnitude of heat expelled during cold compression is $Q_c = nRT_c\ln r$. Thus

$$W = Q_h - Q_c,$$
$$\eta = \frac{W}{Q_h} = 1 - \frac{T_c}{T_h}.$$

The working gas provides one explicit construction. More generally a reversible cycle exchanges entropy Q_h/T_h with the hot reservoir and returns exactly that entropy to the cold one. For an irreversible cyclic engine between the same two reservoirs, the reservoir entropy change $Q_c/T_c - Q_h/T_h$ is nonnegative. This gives the same expression as an upper bound on efficiency. The temperatures must be absolute and must refer to the reservoirs actually exchanging heat.

Test a proposed engine

An engine receives 1.50 kJ per cycle from a 450 K reservoir and rejects heat to a 300 K sink. Reversibility permits at most 0.500 kJ of useful work. A measured cycle delivering 0.360 kJ must reject 1.140 kJ by the first law. The working fluid's entropy returns to its original value, while the two reservoirs together gain $1140/300 - 1500/450 = 0.467$ J/K per cycle.

The energy and entropy balances answer different design questions. Energy accounting gives the amount of rejected heat. Entropy generation explains why some of the available work has been lost relative to the reversible limit. Raising the running frequency can increase power even at lower efficiency, since power depends on work per unit time. The reversible limit describes a thermodynamic bound, and achieving a useful power output additionally depends on finite rates of heat transfer and mechanical motion.

Extract work from two finite heat reservoirs

The usual Carnot limit assumes two fixed reservoir temperatures. Two finite bodies instead change temperature as a reversible engine transfers energy between them. Consider insulated hot and cold blocks with constant heat capacities C_h, C_c. A cyclic engine operating between them can deliver work to an external mechanical store. The block temperatures vary slowly relative to internal equilibration, and the reversible engine draws and rejects each small amount of heat at its corresponding instantaneous block temperature.

If an infinitesimal amount of entropy passes through the engine reversibly, the hot block loses exactly the amount gained by the cold block. Integrating until both reach a common final temperature T_f yields

$$C_h \ln(T_f/T_{h,0}) + C_c \ln(T_f/T_{c,0}) = 0,$$
$$T_f = T_{h,0}^{C_h/(C_h+C_c)} T_{c,0}^{C_c/(C_h+C_c)}.$$

The logarithms contain temperature ratios, so the product expression is dimensionally a temperature. Once temperatures coincide a cyclic engine has no thermal difference left from which to extract work. Conservation of energy gives

$$W_{\max} = C_h(T_{h,0} - T_f) + C_c(T_{c,0} - T_f).$$

Any irreversible entropy production raises the attainable common temperature relative to the reversible result at the same initial conditions, leaving more energy in the blocks and less exported as work.

Compare a reversible engine with passive contact

Let $C_h = C_c = 1.00$ kJ/K, $T_{h,0} = 400$ K and $T_{c,0} = 300$ K. The reversible common temperature is the geometric mean $\sqrt{400 \times 300} = 346.41$ K. The engine receives 53.59 kJ from the hot block and transfers 46.41 kJ to the cold block; their difference is 7.18 kJ of work. The overall work fraction of extracted hot-side heat is 0.134. At the beginning the instantaneous reversible efficiency is $1 - 300/400 = 0.250$, falling to zero at the end. Applying that initial efficiency throughout would overestimate the work.

Now put the same blocks directly in thermal contact without an engine. Energy conservation gives the arithmetic mean of 350 K and zero work. The positive entropy produced is

$$1000 \ln[350^2/(400 \times 300)] = 20.6 \,\mathrm{J/K}.$$

The final temperature exceeds the reversible result by 3.59 K for each block, precisely accounting for 7.18 kJ retained as thermal energy. In the limit of a very small initial temperature difference the geometric and arithmetic means coincide to first order. Recoverable work thus starts at second order in the initial difference even though the heat passed between the blocks is first order, explaining why small temperature contrasts are poor resources for mechanical work.

The arithmetic and geometric means also reveal how sensitive the possible work is to initial temperature calibration. For equal heat capacities, write initial temperatures as $T_m + \delta$ and $T_m - \delta$. The reversible common value is

$$\sqrt{T_m^2 - \delta^2} \simeq T_m - \delta^2/(2T_m).$$

Multiplying the difference from the arithmetic mean by the combined heat capacity gives

$$W_{\max} \simeq C\delta^2/T_m.$$

Thus a small absolute error in a temperature difference can have a doubled relative effect on the predicted maximum work. Measuring both blocks with matched instruments matters more here than obtaining many digits for the mean temperature alone.

Thermodynamic potentials and phase boundaries

Thermodynamics and statistical physics

Select the appropriate potential for a controlled experiment and infer the slope of a coexistence line from latent heat and volume change.

FROM FIRST PRINCIPLES

Derive a coexistence slope from equality of chemical potentials

At fixed temperature and pressure a closed single-component system lowers its Gibbs energy when matter moves into a phase with lower molar chemical potential. Transfer an infinitesimal amount dn from phase 1 to phase 2. Along the coexistence line, either direction is possible in equilibrium, so

$$dG = (\mu_2 - \mu_1)\, dn = 0,$$
$$\mu_1(T, p) = \mu_2(T, p).$$

Moving along the line must preserve that equality. The Gibbs differential for one mole gives

$$d\mu = -s_m\, dT + v_m\, dp,$$

with molar entropy s_m and volume v_m. Equating the differentials of the two phases leaves

$$(v_{m,2} - v_{m,1})\, dp = (s_{m,2} - s_{m,1})\, dT.$$

The latent heat per mole for reversible conversion $1 \to 2$ is $L_m = T(s_{m,2} - s_{m,1})$. Therefore the slope is $dp/dT = L_m/[T(v_{m,2} - v_{m,1})]$ for a nonzero volume jump. Its sign depends on both latent heat and volume change. This relation requires measured or modeled properties of the two phases at the same coexistence point.

Background · Helmholtz free energy

$$F = -k_B T \ln Z$$

At fixed particle number and temperature, differentiating with respect to volume yields minus the pressure. The ensemble specifies which variables to hold fixed.

DERIVATION AND APPLICATIONS

Calculate how a curved interface changes the equilibrium vapour

pressure

A small liquid droplet in its own vapour reaches equilibrium at a different vapour pressure from a flat liquid surface. Consider a spherical droplet of radius r at uniform temperature T, with surface tension γ independent of size. Let p_v and p_l be the exterior vapour and interior liquid pressures. A virtual change of radius changes area and volume by $dA = 8\pi r\, dr$ and $dV = 4\pi r^2 dr$. Balancing the pressure work against the surface free energy gives

$$(p_l - p_v) dV = \gamma\, dA,$$

$$p_l = p_v + \frac{2\gamma}{r}.$$

Use a flat interface at the same temperature as a chemical-potential reference. Both phases there have pressure p_∞. The difference of their chemical potentials is zero for the flat reference and for the equilibrium droplet. Integrating $d\mu = v_m\, dp$ along an isotherm for each phase then gives, for ideal vapour and nearly incompressible liquid of molar volume v_l,

$$RT \ln(p_v/p_\infty) = v_l(p_l - p_\infty) = v_l(p_v - p_\infty) + \frac{2\gamma v_l}{r}.$$

For a dilute vapour the molar volume greatly exceeds the liquid value. When $p_v - p_\infty$ is also small compared with the curvature-induced pressure jump, its contribution on the right is negligible. The Kelvin relation follows,

$$\ln(p_v/p_\infty) \simeq \frac{2\gamma v_l}{rRT}.$$

A convex liquid surface raises the required equilibrium vapour pressure. The result tends continuously to flat coexistence at large radius. At molecular radii the continuum description, including the assumption of a bulk surface tension, requires independent justification.

WORKED EXAMPLE

Estimate the supersaturation over a droplet

Change the inputs and investigate this derivation.

For water-like material at $298\,\mathrm{K}$ take $\gamma = 0.0720\,\mathrm{N/m}$, $v_l = 18.0 \times 10^{-6}\,\mathrm{m^3/mol}$, and $r = 20.0\,\mathrm{nm}$. The exponent is 0.0523, giving $p_v/p_\infty = 1.054$. The mechanical pressure jump is $7.20\,\mathrm{MPa}$. Even with $p_\infty = 3.17\,\mathrm{kPa}$, the exterior pressure increase is only about $170\,\mathrm{Pa}$, verifying the term dropped in obtaining the Kelvin relation.

SECOND APPLICATION

Infer the radius that can coexist with a given vapour

Change the inputs and investigate this derivation.

A different liquid has $\gamma = 0.0250\,\mathrm{N/m}$ and $v_l = 100 \times 10^{-6}\,\mathrm{m^3/mol}$ at $300\,\mathrm{K}$. A chamber stabilized at $p_v/p_\infty = 1.020$ matches radius $r = 2\gamma v_l/[RT\ln(1.020)] = 101\,\mathrm{nm}$. A smaller droplet has a higher required vapour pressure and tends to evaporate in that chamber. A larger one tends to grow, so matching the pressure selects an unstable balance if the vapour pressure is held fixed.

Try it · Kelvin droplet

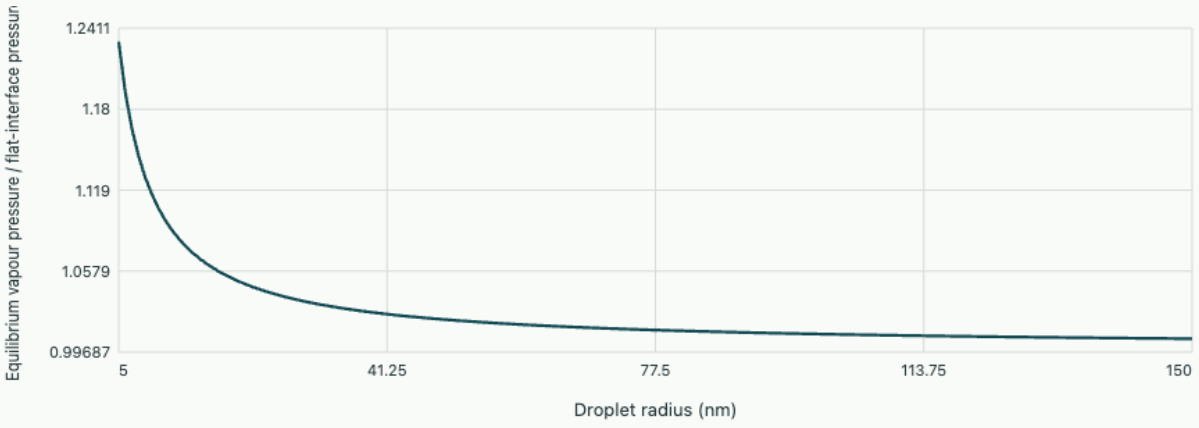
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Surface tension (N/m) **0.072**

Temperature (K) **300**



Kelvin curvature shift

Liquid molar volume fixed at 18 cm³/mol. Curvature raises the liquid chemical potential; molecular-scale droplets require corrections beyond this continuum model.

Change which variables are controlled

Consider a closed simple compressible system that can exchange heat with a reservoir and work by moving a boundary. In equilibrium its fundamental energy differential for fixed amount is $dU = T\,dS - p\,dV$. This expression makes entropy and volume natural independent variables, yet laboratory experiments more frequently control temperature and pressure. A change of thermodynamic potential exchanges one independent variable for its conjugate without losing any equilibrium information.

For the Helmholtz free energy $F = U - TS$, differentiation gives

$$dF = -S\,dT - p\,dV.$$

Its slope with respect to volume at fixed temperature is minus the pressure. For the Gibbs free energy $G = U + pV - TS$, the differential is $dG = -S\,dT + V\,dp$. At fixed temperature and pressure a spontaneous change in a closed sample with only pressure–volume work tends to lower its total G. The external reservoir can exchange heat during that change, so a fall in sample energy by itself would not be the correct selection criterion.

Extract a measurable cross derivative

The entropy of a gas is harder to measure directly than its thermal expansion. Since $G(T, p)$ is a state function with commuting mixed derivatives inside a smooth single phase,

$$-(\partial S/\partial p)_T = (\partial V/\partial T)_p.$$

This is one Maxwell relation. It converts an entropy response to a derivative accessible from volume measurements. The equality applies to the same material and the same choice of fixed variables on both sides.

For one mole of ideal gas $(\partial V/\partial T)_p = R/p$. Compressing isothermally therefore changes its entropy by $\Delta S = -R\ln(p_2/p_1)$. Doubling pressure gives -5.76 J/(mol·K), in agreement with an independent reversible-heat calculation. The Maxwell relation obtains the state difference without requiring the actual compression process to be reversible. Only an equilibrium path used to evaluate it has to connect well-defined states.

Derive a coexistence-line slope

For one pure substance the molar Gibbs energy is the chemical potential. At two-phase equilibrium both phases must have the same chemical potential; otherwise transferring matter to the phase of lower value decreases total G. Label the phases 1 and 2. Along their coexistence curve a small change in T and p changes their chemical potentials equally. Substituting $d\mu = -s_m dT + v_m dp$ for each yields

$$\frac{dp}{dT} = \frac{s_{m,2} - s_{m,1}}{v_{m,2} - v_{m,1}} = \frac{L_m}{T(v_{m,2} - v_{m,1})}.$$

Here L_m is the molar latent heat for the transition from phase 1 to 2. The slope can be negative for melting when the liquid has smaller molar volume than the solid. For vaporisation the volume difference is usually positive and large at modest pressure. A pressure reading and a thermometer together can therefore constrain a latent heat through measurements of the slope and molar volumes.

Background · How can pressure favour melting?

At coexistence the molar Gibbs free energies of solid and liquid agree. Increasing pressure at fixed temperature increases each free energy by its molar volume times the pressure increment. The phase with the smaller molar volume gains the thermodynamic advantage. For ordinary ice near its melting point, melting reduces the molar volume by about 1.6 cm³/mol and requires about 6.0 kJ/mol of heat.

$$\frac{dT_m}{dp} = \frac{T_m(v_l - v_s)}{\Delta h_{\mathrm{fus}}} \simeq -7.3 \times 10^{-8}\ \mathrm{K/Pa}.$$

A 1 MPa increase lowers the equilibrium melting temperature by about 0.073 K near 273 K. This is a local prediction along the coexistence curve. Extrapolating it to hundreds of megapascals would require the pressure dependence of the volumes and latent heat and the possible appearance of another solid phase.

Estimate a boiling temperature near atmospheric pressure

For vaporising water near 373 K use $L_m = 40.7$ kJ/mol, a vapour volume of 0.0306 m³/mol and a much smaller liquid volume. The coexistence slope is approximately $40700/[373(0.0306)] = 3.57$ kPa/K. An increase of pressure by 10 kPa therefore raises the boiling temperature by roughly 2.8 K over this limited interval. The dimensional check is J/(m³·K), equivalent to Pa/K.

Over a larger interval, approximate the vapour as ideal, neglect the liquid volume and treat L_m as constant. The slope becomes $d\ln p/dT = L_m/(RT^2)$, giving $\ln(p_2/p_1) = -(L_m/R)(1/T_2 - 1/T_1)$. This form clarifies why a straight line may appear on a logarithmic-pressure versus reciprocal-temperature plot. Curvature in measured data can carry information about temperature-dependent latent heat and departures from ideal-gas behaviour. Close to a critical point the assumptions behind the simple integrated expression cease to be reliable.

For a substance whose solid is less dense than its liquid, melting can have a negative coexistence slope even though the latent heat of melting is positive. In the Clapeyron relation the sign then comes from the volume difference. Increasing pressure at fixed temperature favours the phase taking less volume. This sign check connects the derivative calculation to an independently measurable density contrast.

Choose the equilibrium pressure inside a two-phase region

A nonmonotone model isotherm can predict several volumes at one pressure below the critical temperature. The middle segment with $(\partial p/\partial v)_T > 0$ is mechanically unstable. Removing that segment alone does not identify the two coexisting phases. At the coexistence pressure p_* both endpoints have equal molar chemical potential. Write molar Gibbs energy as $\mu = f + p_* v$, where f is molar Helmholtz energy on this isotherm. Since $df = -p(v)\,dv$ at fixed temperature, equality of endpoint values implies

$$0 = \mu_g - \mu_l = -\int_{v_l}^{v_g} p(v)\,dv + p_*(v_g - v_l).$$

Geometrically, the areas above and below the horizontal line p_* cancel. The same condition makes that line's negative slope the common tangent of $f(v)$ at v_l and v_g. Between these endpoint volumes the equilibrium material changes the relative amounts of liquid and gas at fixed pressure.

As a local model take $p = p_0 + A(x - x^3)$ with $x = (v - v_0)/b$, $A > 0$, $v_0 = 2.00 \times 10^{-4}$ and $b = 3.00 \times 10^{-5}$ m³/mol. At $p = p_0$, the outer roots $x = \pm 1$ have negative slopes, and the excess pressure is odd in x, verifying equal areas exactly. The stable molar volumes are 1.70×10^{-4} and 2.30×10^{-4} m³/mol. At average volume 2.12×10^{-4} m³/mol, conservation of volume gives a gas mole fraction $(2.12 - 1.70)/(2.30 - 1.70) = 0.700$. The spinodals, where the original slope vanishes, instead occur at $x = \pm 1/\sqrt{3}$. They delimit local stability. They are not the coexistence volumes. This cubic is a local representation of the loop and should not be extrapolated to very large or negative volumes.

Why a phase can persist beyond the coexistence line

Equality of bulk chemical potentials determines coexistence, yet a sample can remain liquid below its equilibrium freezing temperature. Forming a small solid nucleus introduces an interface whose free-energy cost can initially outweigh the bulk driving force. Consider a spherical nucleus of radius r in a pure, uniformly undercooled liquid. Assume an isotropic solid–liquid interfacial energy $\gamma > 0$ and a positive bulk free-energy advantage per volume Δg_v for the solid. Ignore elastic accommodation, walls, impurities and curvature dependence of γ.

The gain in stable bulk material is proportional to the volume while the cost is proportional to area. Relative to uniform liquid the work is

$$\Delta G(r) = 4\pi r^2 \gamma - \frac{4\pi}{3} r^3 \Delta g_v.$$

Differentiating gives the nonzero stationary size $r_* = 2\gamma/\Delta g_v$. The second derivative there is negative, so the critical nucleus lies at a barrier rather than at a stable droplet size. Smaller nuclei reduce their free energy by shrinking, and larger ones reduce it by growing. Substituting the radius into the original expression gives

$$\Delta G_* = 16\pi\gamma^3/(3\Delta g_v^2).$$

Near coexistence, the difference of the two bulk free energies changes as the entropy-of-fusion difference times undercooling. With latent heat per unit volume L_v, use $\Delta g_v \simeq L_v(T_m - T)/T_m$ for a small fractional undercooling.

Determine how an undercooling changes the barrier

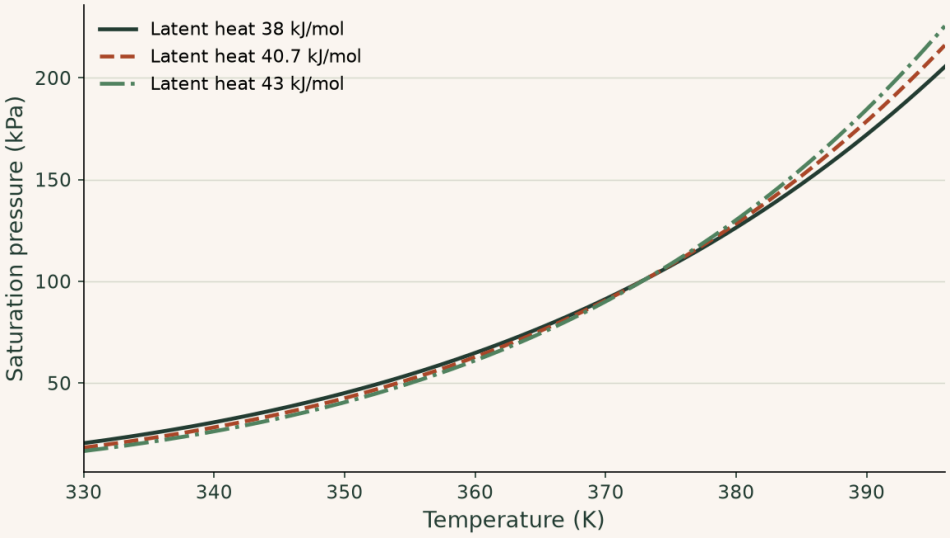
For an illustrative melt take $T_m = 400$ K, $T = 380$ K, $L_v = 3.00 \times 10^{8}$ J/m³ and $\gamma = 0.0300$ J/m². The bulk drive is 1.50×10^{7} J/m³. Critical radius is 4.00 nm and the barrier is approximately 2.01×10^{-18} J. Since $k_B T = 5.25 \times 10^{-21}$ J at 380 K, this is roughly 383 thermal energy units, a formidable homogeneous nucleation barrier. This comparison alone is insufficient for an exact waiting time, which also requires a kinetic prefactor and the sample volume.

Doubling undercooling to 40 K doubles the approximate bulk drive, halves the critical radius to 2.00 nm and reduces the energy barrier by four. The nucleation probability per attempt can thus vary by many orders of magnitude while the equilibrium transition temperature remains fixed. As undercooling tends to zero, both critical radius and barrier diverge in this ideal capillarity model. A container wall or an impurity can reduce the interface area needed and initiate crystallization much earlier. This distinction between the coexistence curve and observed onset explains why reducing the cooling rate or changing surface preparation can change a freezing record.

The spherical calculation exposes a distinction between a critical nucleus and a nucleus for which the total formation work has already become negative. Set the formation work itself, rather than its derivative, to zero. Apart from zero radius this gives $r = 3\gamma/\Delta g_v$, larger than the critical radius by a factor of one and a half. A cluster just above the critical radius can still have positive total formation work relative to the uniform liquid, yet growth reduces its free energy. Using the zero-work radius as a kinetic threshold would therefore overestimate the size needed for downhill growth.

Estimate the boiling temperature under reduced pressure

Anchor a liquid’s saturation curve at 373.15 K and 101.325 kPa. For a dilute ideal vapour, negligible liquid volume and approximately constant molar heat of vaporisation, integrate $d\ln p/dT = L/(RT^2)$. The three curves test sensitivity to the assumed latent heat.



For 40.7 kJ/mol the predicted pressure at 350 K is about 42.5 kPa. Uncertainty in latent heat matters increasingly far from the anchor point. This local model needs new property data before extrapolation toward a critical point.

Try it · Clausius

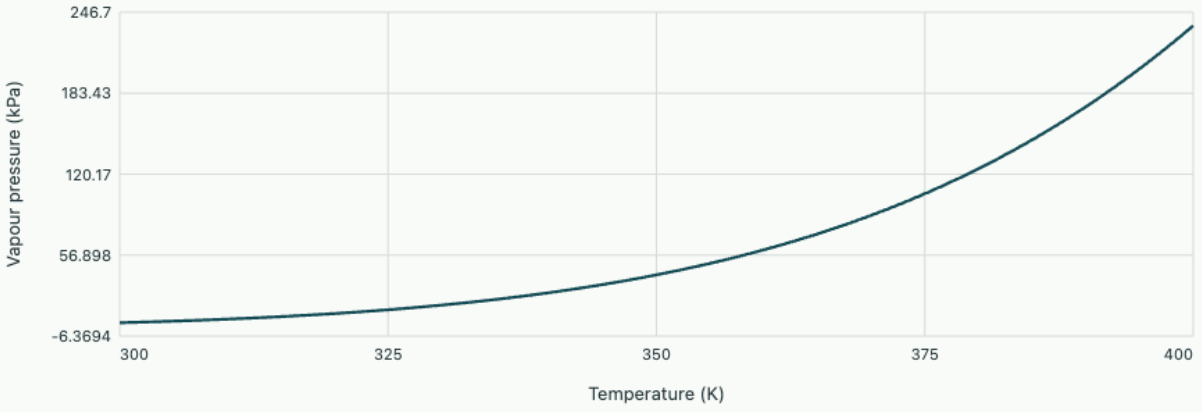
Change the inputs and compare the recalculated curves with the preceding derivation and baseline figure. The axes give the units; the numerical readout states the model assumptions and any limiting cases. Reset the inputs to return to the reference setting.

Open this live calculation

Try the calculation

Molar latent heat (kJ/mol) **40.5**

Anchor boiling temperature (K) **374**



Constant-latent-heat approximation

Pressure is anchored to 101.325 kPa at 374 K. Ideal vapour, negligible liquid molar volume and constant latent heat are assumed.

▶ **Read values from the plot**

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.

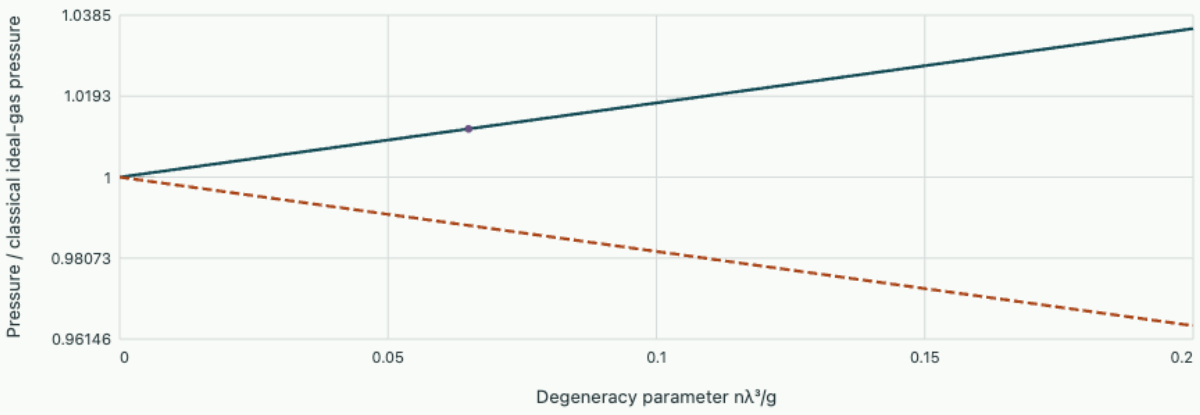
Try it · Quantum virial

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

$$T_c = (2\pi\hbar^2/mk_B)[n/\zeta(3/2)]^{2/3},$$

$$N_0/N = 1 - (T/T_c)^{3/2}.$$

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 $1/(e^{E/k_BT}-1)$. 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.

Chemical equilibrium and reaction extent

Thermodynamics and statistical physics

Minimize Gibbs energy along a permitted reaction and predict a composition change when the system is compressed.

FROM FIRST PRINCIPLES

Build the reaction quotient from chemical potentials

At fixed temperature and pressure, a closed reacting mixture equilibrates by minimizing Gibbs energy subject to conservation of its constituent elements. Let ν_i be the balanced stoichiometric coefficient of species i, positive for a product and negative for a reactant. Advancing the reaction by a small extent $d\xi$ changes each amount by $dn_i = \nu_i d\xi$. The free-energy change per mole of reaction is then

$$\frac{dG}{d\xi} = \sum_i \nu_i \mu_i = \Delta_r G.$$

For an interior equilibrium this derivative vanishes. Chemical potentials can be written $\mu_i = \mu_i^\circ + RT \ln a_i$ using dimensionless activities relative to specified standard states. Collecting all the logarithms gives

$$\Delta_r G = \Delta_r G^\circ + RT \ln Q,$$

$$Q = \prod_i a_i^{\nu_i},$$

$$K = e^{-\Delta_r G^\circ/(RT)}.$$

The equilibrium composition satisfies $Q = K$. For an ideal gas, $a_i = p_i/p^\circ$; nonideal mixtures require activities measured or modeled for the mixture. If $Q < K$, a small forward reaction lowers G. This thermodynamic preference determines a direction near the present state while leaving the reaction timescale to kinetics.

Background · Chemical potential

$$\mu = (\partial F/\partial N)_{T,V}$$

This change in free energy per added particle controls exchange between subsystems. In a finite small system an actual finite difference may matter.

DERIVATION AND APPLICATIONS

Propagate an equilibrium constant when reaction heat capacity changes the enthalpy

An equilibrium constant quoted at one temperature does not determine the equilibrium composition at another unless the energy change of the reaction is also known. Consider a reversible reaction with a dimensionless standard equilibrium constant $K(T)$ and unchanged standard-state choices. All reaction quantities below are per mole of reaction as written. At fixed standard pressure, the Gibbs differential $dG = -S\,dT$, together with $G = H - TS$, gives

$$d(G/T)/dT = -H/T^2.$$

Apply this identity to the stoichiometric differences of standard molar quantities,

$$\ln K = -\frac{\Delta_r G^\circ}{RT},$$
$$\frac{d\ln K}{dT} = \frac{\Delta_r H^\circ(T)}{RT^2}.$$

If heat capacities of products and reactants differ, the enthalpy changes with temperature. Over an interval where their stoichiometric difference $\Delta_r C_p^\circ$ can be treated as constant, integration of $dH^\circ = C_p^\circ dT$ gives

$$\Delta_r H^\circ(T) = \Delta_r H^\circ(T_1) + \Delta_r C_p^\circ (T - T_1).$$

Insert this linear enthalpy into the derivative for $\ln K$. The constant part integrates as $1/T^2$; the term proportional to temperature also produces a logarithm. The two-temperature relation is

$$\ln\frac{K_2}{K_1} = \frac{\Delta_r H^\circ(T_1)}{R}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) + \frac{\Delta_r C_p^\circ}{R}\left[\ln\frac{T_2}{T_1} + \frac{T_1}{T_2} - 1\right].$$

Both terms vanish when the two temperatures coincide. For a small temperature interval the heat-capacity correction begins at second order, consistent with the local slope depending only on the current reaction enthalpy. Phase changes would interrupt the assumed smooth enthalpy law.

WORKED EXAMPLE

Estimate a change in an association equilibrium

Change the inputs and investigate this derivation.

Take $K_1 = 4.00$ at $300\,\mathrm{K}$, $\Delta_r H^\circ(T_1) = -20.0\,\mathrm{kJ/mol}$, and $\Delta_r C_p^\circ = 100\,\mathrm{J/(mol\,K)}$. At $360\,\mathrm{K}$ the enthalpy term in $\ln(K_2/K_1)$ is -1.336 and the heat-capacity term is $+0.188$. Thus

$$K_2 = 4.00 e^{-1.148} = 1.27.$$

The association remains exothermic throughout the interval, although its magnitude declines on heating.

SECOND APPLICATION

Infer enthalpy from equilibrium measurements

Change the inputs and investigate this derivation.

For a different reaction K rises from 0.500 at $290\,\mathrm{K}$ to 0.900 at $310\,\mathrm{K}$. If the reaction heat-capacity difference is negligible here, inversion gives

$$\Delta_r H^\circ = R \ln(0.900/0.500)/(1/290 - 1/310) = 22.0\,\mathrm{kJ/mol}.$$

The positive sign is consistent with the increased equilibrium constant on heating. At $300\,\mathrm{K}$ the same enthalpy predicts a local logarithmic slope of $0.0293\,\mathrm{K}^{-1}$, testable with a more closely spaced measurement.

Try it · Vanthoff

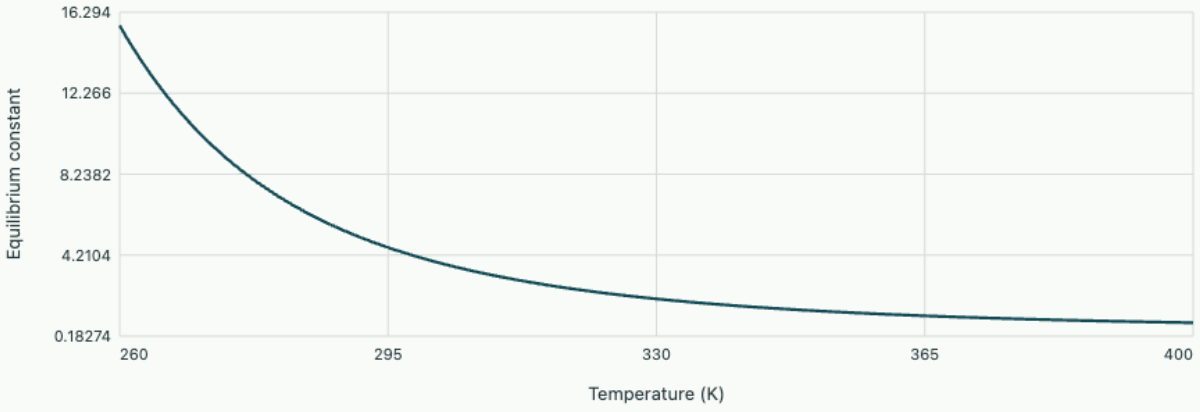
Vary the physical parameters in the derived relation. Use the units and assumptions in the readout to connect the curves to the two worked examples. A parameter study can test the forward prediction; the same relation can then be inverted to infer a parameter from a measurement.

Open this live calculation

Try the calculation

Reaction enthalpy at 300 K (kJ/mol) **-20**

Reaction heat-capacity difference (J/mol/K) **100**



Integrated temperature-dependent enthalpy

K=4 at 300 K, ΔH(T)=ΔH(300)+ΔCp(T−300). Activities and standard states are held consistent.

A reaction has one constrained direction

A closed vessel can change particle species while conserving the total number of atoms of each element. At fixed temperature and pressure the equilibrium state minimizes Gibbs energy. Introduce a reaction extent ξ in moles and stoichiometric coefficients ν_i, negative for reactants and positive for products. A small reaction advance changes amounts by $dn_i = \nu_i \, d\xi$. The Gibbs differential at fixed T, p therefore gives

$$dG = \sum_i \mu_i \, dn_i = \Delta_r G \, d\xi,$$

$$\Delta_r G = \sum_i \nu_i \mu_i.$$

An interior minimum satisfies $\Delta_r G = 0$. A negative value means that a small positive advance lowers G. Reversibility of the microscopic chemistry is assumed for equilibrium to be accessible. Thermodynamic preference alone does not provide the time needed to reach it.

Activities and the equilibrium constant

For ideal gases $\mu_i = \mu_i^\circ(T) + RT\ln(p_i/p^\circ)$. The activity is the dimensionless ratio $a_i = p_i/p^\circ$. Inserting this expression into the reaction Gibbs energy gives

$$\Delta_r G = \Delta_r G^\circ + RT\ln Q,$$

with $Q = \prod_i a_i^{\nu_i}$. The equilibrium value of Q is $K = \exp[-\Delta_r G^\circ/(RT)]$. Standard pressure supplies the reference that makes the logarithm and the equilibrium constant dimensionless even when the total molecular count changes.

For the dimerization $2A \rightleftharpoons A_2$, begin with one mole of monomer units and define α as the fraction of units bound into dimers. Then $n_A = 1-\alpha$, $n_{A_2} = \alpha/2$ and total gas amount is $1-\alpha/2$ moles on this basis. Partial pressures follow from mole fractions times total pressure. Substitution into the reaction quotient gives

$$Q = \frac{\alpha(2-\alpha)}{4(1-\alpha)^2}\,\frac{p^\circ}{p}.$$

The ratio of activities is determined by both composition and total pressure. Compression at the original equilibrium composition immediately lowers Q, making the forward reaction favourable. The subsequent association reduces the number of gas particles while conserving monomer units.

A pressure change at fixed temperature

Suppose $K = 0.750$ at the chosen temperature. Let $y = 1-\alpha$. Because $\alpha(2-\alpha) = 1-y^2$, solving $Q = K$ gives

$$y = [1 + 4Kp/p^\circ]^{-1/2}.$$

At $p = p^\circ$ this gives $\alpha = 0.500$, with one half of all units bound and only one third of gas particles being dimers. At $p = 4p^\circ$ the bound-unit fraction is

$$1 - 1/\sqrt{13} = 0.723.$$

The distinction between fractions of conserved units and fractions of gas molecules matters for interpreting a spectroscopic composition measurement.

A temperature change can also move equilibrium. Differentiating $\ln K = -\Delta_r G^\circ/(RT)$ and using the Gibbs–Helmholtz relation gives

$$d\ln K/dT = \Delta_r H^\circ/(RT^2).$$

For an exothermic association increasing temperature decreases K locally. Pressure changes at fixed temperature leave K itself fixed under the same standard-state convention. It is the composition-dependent quotient that adjusts to the imposed mechanical conditions.

The distinction between a chemical potential per mole and one per particle is essential in this derivation. Here partial pressures were combined with R because the extent was measured in moles. Molecular statistical mechanics instead uses k_B with energies per molecule, but both conventions give the same dimensionless ratio of equilibrium activities. A useful numerical check is to scale every initial amount and the container volume by the same factor at fixed temperature. Mole fractions and partial pressures stay fixed, so the equilibrium fraction is unchanged even though twice as many molecules may react.

Use a membrane to turn chemical equilibrium into pressure

Equality of chemical potential can constrain a species even when other species cannot move. Separate pure solvent at pressure p from a dilute ideal solution at pressure $p + \Pi$ by a membrane permeable only to solvent. Hold temperature fixed and assume a nearly incompressible solvent with approximately constant partial molar volume $\bar{v}_w$. The solvent can exchange across the membrane; solute cannot. Its equilibrium condition is therefore equality of solvent chemical potentials, even though total pressures and solute concentrations differ.

For an ideal solution the solvent chemical potential relative to pure solvent at the same pressure changes by $RT\ln x_w$, where x_w is the solvent mole fraction. Increasing pressure by Π raises the solvent chemical potential by $\int_p^{p+\Pi} \bar{v}_w\,dp' \simeq \bar{v}_w \Pi$. Equating the two sides gives

$$\Pi = -\frac{RT}{\bar{v}_w}\ln x_w \simeq \frac{RT}{\bar{v}_w} x_s \simeq RT c_s.$$

The last approximation assumes a small solute mole fraction x_s and solution volume close to the solvent volume, with c_s in moles of independently dissolved particles per cubic metre. The pressure difference is positive because mixing lowers the chemical potential of solvent. If pressure is released, solvent flows into the solution, modifying its concentration until a new balance is attained.

Estimate a hydrostatic head and a molar mass

At 298 K dissolve 2.00 g/L of a nonvolatile solute with molar mass 0.200 kg/mol. The mass concentration is 2.00 kg/m³, so $c_s = 10.0$ mol/m³. Predicted osmotic pressure is $8.314 \times 298 \times 10.0 = 24.8$ kPa. With solution density approximated by 1000 kg/m³, an open-column experiment would need height difference $\Pi/(\rho g) = 2.53$ m to balance it. The final concentration must be measured after the column rises, since transferred solvent can appreciably dilute a small sample.

Conversely, for the same 2.00 g/L preparation a pressure reading of 12.4 kPa would imply an apparent molar mass $RT(2.00)/12400 = 0.400$ kg/mol. Association of two formula units into one particle is one

possible explanation if the chemical analysis gives 0.200 kg/mol. Nonideal activities, leakage and concentration calibration can also change the inferred particle count. Repeating at lower concentrations tests whether Π / c_s approaches the ideal-gas constant times temperature. In the pure-solvent limit $x_w \to 1$ the chemical-potential difference vanishes so no osmotic pressure is required, despite the membrane still being present.

Concentration units can be a major source of error in this example. One mole per litre is a thousand moles per cubic metre, so the gas constant in joules per mole kelvin gives a pressure in pascals only after this conversion. For fully dissociated dilute electrolytes there may also be more than one independently mobile particle per formula unit. Conversely strong ion pairing or aggregation can reduce the count. Measuring a nonvolatile neutral reference solute of known molar mass in the same apparatus provides a way to test both the pressure scale and the effectiveness of the membrane.

How a pair potential changes the gas pressure

Thermodynamics and statistical physics

A low-density gas can measure its own intermolecular forces. At a prescribed temperature the first deviation from ideal behaviour is a volume. Its sign tells us whether overlap exclusion or attraction makes the larger contribution.

Separate independent pairs from correlated pairs

Let N classical particles occupy volume V, with a total potential energy equal to a sum of short-range pair potentials $u(r_{ij})$. Momentum integrals give the ideal-gas factors. In the remaining positional integral write each pair's Boltzmann weight as $e^{-\beta u} = 1 + f(r)$. The auxiliary function $f = e^{-\beta u} - 1$ vanishes where particles do not interact. In a large vessel the positional integral with one f is $V^{N-1} \int f(\boldsymbol{r})\, d^3 r$. One particle of the pair may be placed anywhere and its neighbour's relative coordinate is sampled against f. There are $N(N-1)/2$ possible unordered pairs.

In the low-density expansion of $\ln Z$, independent repetitions of separate pairs cancel against products from the logarithm. The first connected correction is $N(N-1)\int f/(2V)$. This is a density expansion of the free energy per particle. The total correction need not be small for a macroscopic sample. Differentiating with $p = k_B T (\partial \ln Z / \partial V)_{T,N}$ supplies a minus sign from the reciprocal volume. In the thermodynamic limit with number density $\rho = N/V$,

$$\frac{p}{\rho k_B T} = 1 + B_2(T)\rho + O(\rho^2),$$

$$B_2 = -\tfrac{1}{2}\int f\, d^3 r = -2\pi \int_0^\infty f(r) r^2\, dr.$$

An impenetrable core gives $f = -1$. An attractive annulus gives $f > 0$. The sign of B_2 follows from their integrated volumes, including the spherical shell factor r^2.

Integrate a square well

Set $u = \infty$ for $r < d$, $u = -\epsilon$ for $d < r < \lambda d$, and $u = 0$ outside. The two nonzero contributions are constants integrated over spherical volumes. With $B_{\mathrm{core}} = 2\pi d^3/3$, obtain

$$B_2 = B_{\mathrm{core}}\{1 - (\lambda^3 - 1)[e^{\epsilon/(k_B T)} - 1]\}.$$

A sphere of diameter d has actual volume $\pi d^3/6$. The hard-core coefficient is four times that volume. The forbidden positions concern the centre of a second particle, and the factor of one half avoids counting an interacting pair twice. For $\lambda = 1.5$ and $k_B T/\epsilon = 2$, the ratio $B_2/B_{\mathrm{core}} = 1 - 2.375(e^{0.5} - 1) \simeq -0.541$. A dilute sample with $\rho B_{\mathrm{core}} = 0.02$ consequently has $p/(\rho k_B T) \simeq 0.9892$ at this order.

Locate the temperature at which the slope vanishes

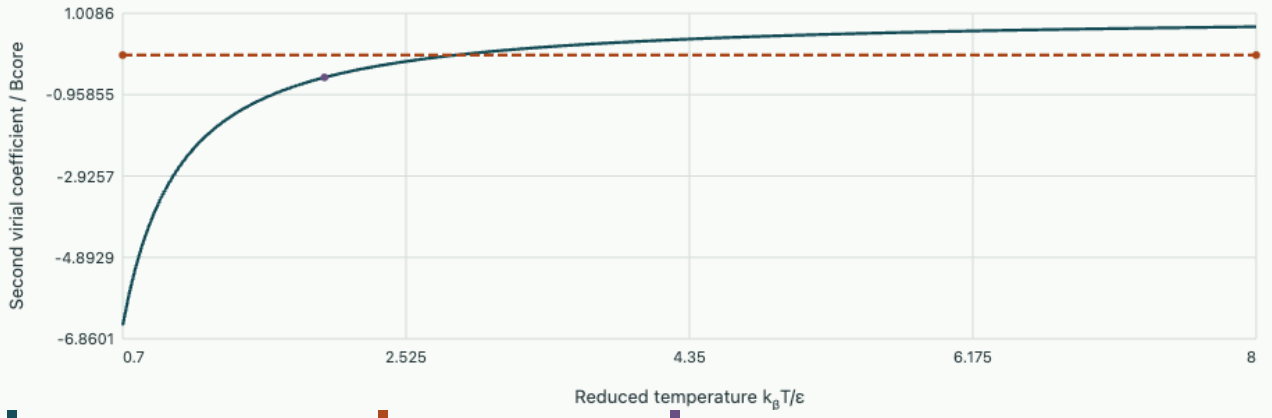
Set $B_2 = 0$ and solve for the Boyle temperature. The result is $k_B T_B/\epsilon = 1/\ln[1 + 1/(\lambda^3 - 1)]$, or 2.846 for the preceding range. Vary the width of the well. A larger attraction volume can offset the hard core even when the temperature is high. At T_B the pressure curve has zero first density correction. Three-particle contributions can still produce curvature. At very low temperatures pair binding can become strong enough that even a superficially low density fails to justify truncation.

Open this experiment online

Try the calculation

Outer radius / core diameter **1.5**

Reduced temperature $k_\beta T/\varepsilon$ **2**



B_2/Bcore = -0.5407. The Boyle temperature is 2.846 ε/k_β.

Tutorial on inferring the range

In a square-well model the measured Boyle temperature is

$$k_B T_B/\epsilon = 2.$$

Determine λ, then predict the sign of the pressure correction when the temperature is doubled.

▼ **Work through the solution**

Rearrangement gives $\lambda^3 = 1 + 1/(e^{1/2} - 1) = 2.5415$, hence $\lambda \simeq 1.365$. Doubling the temperature decreases the attractive Boltzmann factor and leaves the hard-core term fixed. Substitution at reduced temperature 4 gives

$$B_2/B_{\text{core}} \simeq 1 - 1.5415(0.2840) = 0.562.$$

At sufficiently low density the pressure exceeds the ideal prediction.

Can a negative B_2 at one temperature fix how much of the molecular attraction comes from its depth and how much from its range?

▼ **Work through the solution**

The coefficient in the square-well model depends on the product $(\lambda^3 - 1)[e^{\epsilon/k_B T} - 1]$. Many combinations of width and depth give the same product. Measurements at two temperatures can constrain both parameters if the core diameter is independently known and the model is appropriate. Additional temperatures expose inadequacies in the assumed well shape.

Take two measured compressibility factors $Z = 0.9950, 0.9900$ at $\rho B_{\text{core}} = 0.010, 0.020$ at the same temperature. Estimate B_2/B_{core} and check consistency of the truncation with these two measurements.

▼ **Work through the solution**

The slopes from the ideal limit are $(Z - 1)/(\rho B_{\text{core}}) = -0.50$ for both points. Their common value agrees with a leading pair correction and shows no detectable curvature at the stated precision. It provides no bound on higher-order terms outside that measured density interval.

Further reading · Cambridge statistical physics, classical gases and the cluster expansion.

Why two coexisting phases share a tangent

Thermodynamics and statistical physics

A material prepared with an intermediate composition can separate into two regions of different compositions. Its total composition is conserved. How can a free-energy graph predict both final compositions and the amount of each phase?

Minimise the free energy of a mixture

Work at fixed temperature and volume with a constant density of equivalent sites. Let $f(c)$ be the Helmholtz free energy per site of a spatially uniform state with fraction c of species B. If the two regions have compositions c_1, c_2 and the second contains fraction x of all sites, conservation requires

$$\bar{c} = (1-x)c_1 + xc_2.$$

Neglect interface energy per site in the macroscopic limit. The mixture free energy per site is then the weighted sum of the two f's, a point on the chord joining them.

Introduce a multiplier μ and minimise $(1-x)f(c_1) + xf(c_2) - \mu[(1-x)c_1 + xc_2 - \bar{c}]$. Varying c_1, c_2 when both regions are present gives

$$f'(c_1) = f'(c_2) = \mu.$$

This derivative measures the chemical work to replace an A particle by B. Varying x supplies the missing condition,

$$f(c_2) - f(c_1) = \mu(c_2 - c_1).$$

The common slope therefore equals the slope of the connecting chord. A straight line must touch the free energy at both endpoints and lie below competing homogeneous states. Equivalently $\int_{c_1}^{c_2} [f'(c) - \mu]\, dc = 0$. This equal-area condition is the composition version of Maxwell's construction. For liquids plotted as pressure versus specific volume, the conjugate slope acquires the sign from $p = -\partial F/\partial V$.

Find the coexistence and stability boundaries separately

Take $f(c) = A[(c-m)^2 - d^2]^2 + \mu_0 c$, with $A > 0$, $m = 0.5$ and $0 < d < 0.5$. Subtracting the linear term tilts all trial chords by the same amount. At $c = m \pm d$ the remaining term vanishes with zero derivative. Hence the common tangent is $\mu_0 c$, and the coexistence compositions are $m - d, m + d$. At $d = 0.30$ a sample of average composition 0.45 separates into 0.20 and 0.80 regions. Conservation gives $x = (0.45 - 0.20)/(0.80 - 0.20) = 5/12$ for the B-rich region.

Now perturb a uniform state into two equal parts at $c \pm \eta$. Their average energy differs from the uniform energy by $f''(c)\eta^2/2 + O(\eta^4)$. Here $f'' = 4A[3(c-m)^2 - d^2]$. Between $m - d/\sqrt{3}$ and $m + d/\sqrt{3}$ infinitesimal fluctuations reduce the energy. Between either of these limits and the nearest coexistence endpoint a uniform state resists small fluctuations but a sufficiently large rearrangement can still lower

the energy. This accounts for metastability and for the distinction between spinodal and coexistence boundaries. The uniform model does not predict a nucleation time because interfaces carry no energy in this construction.

Get a phase fraction from a graph

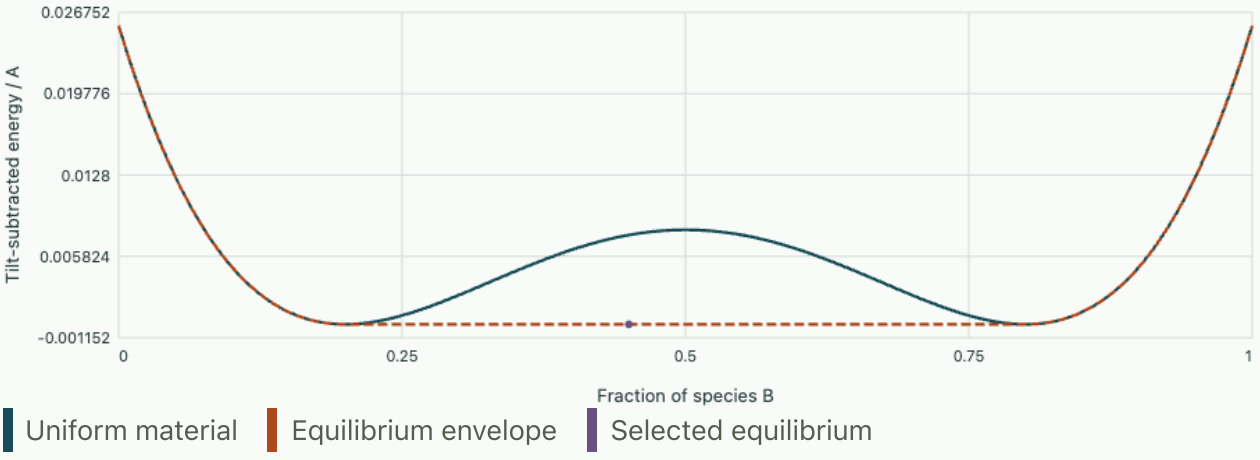
The display subtracts $\mu_0 c$ and measures energies in A. Move the mean composition across the whole interval. Outside the coexistence window the stable composition is homogeneous. Inside, the endpoints stay put while the rich-region fraction changes linearly. Compare $c = 0.45$ to $c = 0.55$ for the same gap. Reflection symmetry exchanges the two phase fractions and leaves the free-energy saving unchanged.

Open this experiment online

Try the calculation

Composition half-gap d **0.3**

Mean B fraction **0.45**



The B-rich phase occupies 41.67% of sites. Energy saving is 0.007656 A per site.

Tutorial on a composition quench

A homogeneous sample is prepared at $c = 0.26$ with $d = 0.30$. Does an infinitesimal composition fluctuation lower its energy? What fraction of the final mixture occupies the B-rich phase in equilibrium?

▼ **Work through the solution**

The lower spinodal lies at $0.5 - 0.3/\sqrt{3} = 0.327$, so the prepared state lies below it and has positive curvature. Nevertheless 0.26 lies between 0.20 and 0.80. Its equilibrium rich fraction is

$$0.06/0.60 = 0.10.$$

Local stability and the global free-energy minimum answer different experimental questions.

For $m = 0.50$, $d = 0.30$, compare free-energy gains of demixing at mean compositions 0.20 and 0.50 after subtracting the tilt. Report the gain in units of A.

▼ **Work through the solution**

At a coexistence endpoint the pure homogeneous material already lies on the common tangent and has zero gain. At the midpoint the homogeneous energy exceeds the tangent by $d^4A = 0.0081A$ per site. The equilibrium mixture uses equal amounts of the two compositions. The gain tends continuously to zero at either endpoint.

Would adding $b + mc$ to the free energy change the two selected compositions or spinodals? What would change in the tangent construction?

▼ **Work through the solution**

Conservation fixes the mean composition, so the new term contributes $b + m\bar{c}$ to every candidate mixture at that mean. The minimiser is unchanged. Second derivatives are unchanged, preserving spinodals. Both endpoint tangent slopes, and thus the chemical-potential difference assigned to coexistence, shift by m.

Further reading · Cambridge statistical physics, phase equilibrium and Maxwell construction.

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.

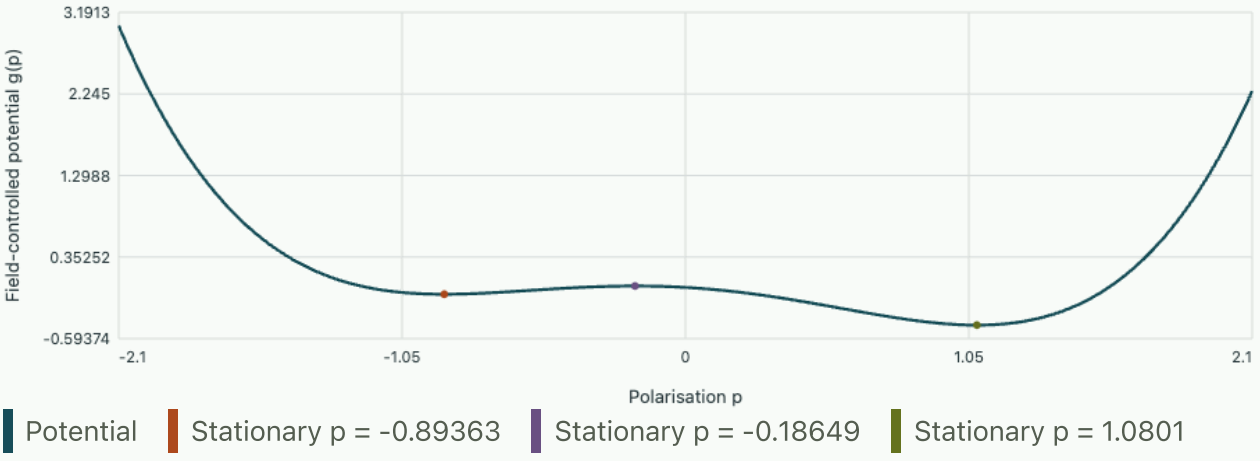
Read a landscape rather than tracing an unlabelled loop

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.

Counting, absolute entropy and the entropy of mixing

Statistical mechanics from microscopic counting to material response

An ideal gas expands into an evacuated bulb and its temperature can remain essentially unchanged. What changed if a thermometer reports the same value? Begin by specifying a microscopic count, recover the pressure and temperature of the gas, and test the meaning of distinguishability by removing a partition between two samples. The calculation also explains why the classical formula for absolute entropy needs a microscopic action scale.

Describe a count before taking its logarithm

A macrostate specifies only outcomes resolved by a chosen preparation and measurement, for example particle number, energy within a narrow interval and container volume. A microstate specifies detailed positions and momenta consistent with those outcomes, or a quantum state in the corresponding energy shell. The multiplicity is the number of allowed states under the stated resolution and constraints. The logarithm makes multiplicities of statistically independent systems additive, because the number of joint possibilities is their product. Boltzmann entropy is

$$S = k_B \log \Omega.$$

Before computing a gas, inspect an easier count. If eight distinguishable sites each have a variable taking values zero or one and exactly three sites are occupied, there are $\binom{8}{3} = 56$ configurations. Revealing which sites are occupied reduces the compatible set to one. This change of conditional information need not involve an actual heat flow in the original system. Thermodynamic entropy requires a physical specification of allowed states and operations; it cannot be inferred solely from how surprised a particular observer feels.

In equilibrium, equal a priori weighting in a narrow isolated energy shell is a postulate about the appropriate ensemble. It does not state that a finite trajectory visits every conceivable state in any chosen observation time. Long-lived constraints, such as a closed partition or a conserved component of angular momentum, change the accessible shell. A gas that has not exchanged particles between remote compartments can be assigned a separate particle number in each. Removing a constraint can enlarge the equilibrium count even while the volume of an exact fine-grained Hamiltonian distribution is preserved by microscopic dynamics. The thermodynamic description refers to a macroscopic resolved equilibrium state.

Count the spatial and momentum volumes of a monatomic gas

Take N indistinguishable noninteracting particles of mass m in volume V, with total kinetic energy at most E. Integrating over positions produces V^N. Allowed momenta lie inside a sphere in $3N$ dimensions of radius $\sqrt{2mE}$. Its volume follows from computing a multidimensional Gaussian integral both in Cartesian and radial coordinates. Divide by a microscopic cell h^{3N} to make the state count dimensionless and by $N!$ to prevent a permutation of indistinguishable particles from creating another physical state. The cumulative count is

$$\Phi(E,V,N)=\frac{V^N}{N!h^{3N}}\frac{(2\pi mE)^{3N/2}}{\Gamma(3N/2+1)}.$$

Using the shell count $\partial_E\Phi\,\delta E$ instead changes subextensive terms for a conventional narrow shell that still contains many states. The cumulative formula is especially convenient for this gas. Equalities at finite small N depend on which entropy is chosen, so state that choice before claiming an exact finite-size temperature. The usual bulk relations follow from either in the thermodynamic limit for this unbounded, regular system.

For large N, use $\log N!\simeq N\log N-N$ and the analogous gamma-function asymptotics. After collecting the terms,

$$S\simeq Nk_B\left[\log\left(\frac{V}{N}\left(\frac{4\pi mE}{3Nh^2}\right)^{3/2}\right)+\frac{5}{2}\right].$$

Doubling E,V,N at fixed ratios doubles the leading expression. Omitting $N!$ would leave a term growing like $N\log N$, violating this extensivity. The phase-space cell influences the entropy per particle and the chemical potential but drops out of classical fixed-particle derivatives for pressure and temperature. Quantum mechanics fixes the natural cell scale through state counting. It does not require physical particles to carry serial numbers while remaining otherwise identical.

Read temperature, pressure and chemical potential from one function

Write the entropy differential as $dS=dE/T+(p/T)dV-(\mu/T)dN$. Differentiating the count at fixed remaining variables gives

$$1/T=3Nk_B/(2E),$$
$$p/T=Nk_B/V.$$

Thus an energy exchange measurement and a force on a moving wall follow from derivatives of the same count. The thermal wavelength $\lambda_T=h/\sqrt{2\pi mk_BT}$ lets us rewrite the entropy and chemical potential as

$$\frac{S}{Nk_B}=\frac{5}{2}-\log(n\lambda_T^3),$$
$$\mu=k_BT\log(n\lambda_T^3),$$
$$n=N/V.$$

The chemical potential uses the same single-particle energy zero as the energy expression. Shifting every particle energy by a constant shifts μ by that constant, leaving equilibrium criteria involving consistent references unchanged. An unresolved internal ground-state degeneracy g would multiply the state count by g^N in the classical dilute limit and replace $n\lambda_T^3$ by $n\lambda_T^3/g$. Real molecules and atoms can

have internal states and additional degrees of freedom; never assign monatomic entropy to a molecular gas without justifying its internal contribution.

For ideal monatomic argon at 300 K and 1.00 bar, using $m = 39.948$ unified atomic mass units, $\lambda_T = 15.95$ pm and $n\lambda_T^3 = 9.79 \times 10^{-8}$. The result is approximately 154.97 J per mole kelvin. Its small degeneracy parameter supports the dilute classical approximation. Compressing isothermally by a factor of ten increases chemical potential by $k_B T \log 10$ and decreases entropy by $N k_B \log 10$. Extending the classical formula down to zero temperature at fixed density would eventually give a negative unbounded entropy, signalling the breakdown of the classical state occupation assumptions.

Why a partition between identical samples is different

Take two ideal gases initially separated at equal temperature and pressure. Their volumes are proportional to their particle numbers. For two distinguishable species each component expands to the full volume after the partition is removed. The internal kinetic energies are unchanged in the ideal model. Summing the two entropy increases and writing mole fractions $x_i = N_i/N$ gives

$$\Delta S_{\mathrm{mix}} = -N k_B \sum_i x_i \log x_i .$$

One mole containing fractions 0.30 and 0.70 gains 5.08 J/K upon ideal mixing from the pure components. At 300 K a reversible isothermal separation needs minimum work $T\Delta S_{\mathrm{mix}} = 1.52$ kJ for this mole if the ideal components and reference states are restored. The free expansion itself need not deliver that work. It is available in a reversible mixing path using selective barriers or equivalent controlled separation operations.

If both compartments initially contain identical gas at equal density, the final macrostate has the same density and same single species throughout. Combining two extensive entropies gives the final entropy to leading bulk order. There is no extensive mixing increment $N k_B \log 2$. Introducing artificial labels for the former left and right particles changes the physical question only if a real degree of freedom can be prepared, distinguished and selectively processed. A dye, isotope or internal spin state can serve as such a label under suitable experimental operations. Merely drawing colours on a computer representation of identical classical particles does not establish a membrane that could sort those labels.

A tempting argument says that making two different gases more and more similar must continuously reduce their ideal mixing entropy. In the ideal model the species occupation labels are either distinct or identical; the description does not specify a continuously degrading separation instrument. A model of partially distinguishable internal states or a noisy finite-resolution sorter is a different experiment. Its accessible work depends on the physical state preparation and measurement operations. Distinguish the mathematical idealisation from an actual membrane having imperfect chemical selectivity.

Check an entropy change using a reversible reference route

In an isolated free expansion of a dilute gas into a vacuum the external pressure vanishes and boundary work is zero. With negligible heat flow energy remains constant and for an ideal gas so does

temperature. To compute the state-function entropy change choose a hypothetical reversible isothermal expansion between the same endpoints. Reversible heat input is $\int p\,dV = Nk_BT\log(V_f/V_i)$, giving $Nk_B\log(V_f/V_i)$. This reference heat is not a claim that heat actually crossed the boundary during free expansion.

For a more exact real-gas experiment, first check that the initial and final temperatures are indeed the same. Internal energy may depend on volume. From the fundamental differential and a Maxwell relation one obtains

$$(\partial U/\partial V)_T = T(\partial p/\partial T)_V - p.$$

A temperature change during an isolated free expansion thus carries information about intermolecular interactions. Measuring only final pressure and assuming constant temperature could misattribute an interaction-induced energy change to the ideal state count.

Use the existing energy-sharing experiment to see how multiplicity changes with a constraint

Recover the entropy cost of a number fluctuation

A gas can lose an extensive mixing entropy when a physically selective partition is imposed, yet the finite-size cost of constraining only a particle count is much smaller. To see this, divide a box into two equal geometric halves with an imaginary plane and let all particles travel freely. At fixed total particle number each classical ideal particle has probability one half to be on either side. The probability of exactly k particles on the left is $\binom{N}{k}/2^N$, independently of whether the particles have names in a bookkeeping algorithm. The binomial coefficient counts the spatial arrangements compatible with that number. The resulting variance of the left count is $N/4$.

The chance of exactly the mean count for even N is approximately $\sqrt{2/(\pi N)}$, from Stirling's formula. If a protocol constrains the left count to precisely $N/2$, the entropy difference between allowing all counts and that constrained ensemble is therefore about $(k_B/2)\log(\pi N/2)$. It grows only logarithmically with total size. In contrast, if the constraint identifies which half contains every member of two distinguishable species, the mixing entropy is extensive. The physical variables that the partition controls decide which count is appropriate.

For $N = 100$ the probability of exactly fifty on the left is about 0.0796. The associated entropy cost of conditioning is $2.53k_B$. A macroscopic observer resolving the fraction only to a few percent can count a whole interval of nearby values as the same macrostate. As N grows the relative width shrinks like $1/\sqrt{N}$, placing almost all of the probability inside any fixed nonzero fractional interval about one half. This is how an apparently precise macroscopic density can coexist with an enormous number of fluctuating microscopic arrangements.

The far tails reveal a stronger distinction. For a fixed fraction $x = k/N$ away from one half, applying Stirling to the binomial gives probability proportional to $\exp[-NI(x)]/\sqrt{2\pi Nx(1-x)}$, with

$$I(x) = x\log(2x) + (1-x)\log[2(1-x)].$$

This rate function is nonnegative, vanishes at one half and starts quadratically there. At $x = 0.7$ it is about 0.0823, so the leading exponential cost for a thousand particles is $e^{-82.3}$. A Gaussian near the peak gives a similar rough scale here but fails to capture the exact tail or the hard limits $0 \le k \le N$. A useful numerical exercise calculates binomial coefficients using log factorials, compares the central Gaussian approximation with the full rate function and checks that the probabilities sum to one. Evaluating the three factorials directly can overflow while the final probability is perfectly representable.

A reversible separation engine and a deliberately imperfect measurement

Imagine two ideal distinguishable gases separated initially at equal pressure and temperature. To obtain work while mixing, use species-selective membranes and allow reversible isothermal expansion of each partial gas against an appropriate pressure. The membrane traversed by one component must not transmit that component to the wrong side, and every infinitesimal motion must occur at negligible pressure imbalance for the component doing the work. The sum of the recoverable works is the sum of each component's ideal expansion work. Reversing the device separates the mixture at the same thermodynamic cost when dissipation can be made negligible.

For a one-mole mixture containing 0.20 mole of one species and 0.80 of another at 300 K, the work contribution of the minority component is $-0.20 RT \log 0.20 = 0.803$ kJ and that of the majority is $-0.80 RT \log 0.80 = 0.445$ kJ. Although the minority has fewer particles, each expands across a larger volume ratio and its work contribution is larger. The total is 1.25 kJ. A bulk piston doing no selective separation sees equal total initial pressures and no obvious macroscopic pressure difference. The selective mechanical operation matters to the interpretation of free energy.

A useful idealised sorter determines a binary label with symmetric error probability e, with equal prior fractions. The fraction of correctly sorted particles is $1 - e$, yet that single number by itself says little about its thermodynamic performance unless both output compositions are measured. For the specified symmetric device the posterior composition in each equal outlet is $(1 - e, e)$ in opposite order. The increase in mixing free energy relative to one unseparated equimolar reservoir at the same total temperature, density and amounts is

$$\Delta F = N k_B T [\log 2 + (1 - e)\log(1 - e) + e \log e].$$

This is a state-function comparison of outlet compositions, not an assertion that the actual apparatus consumes this amount or that measuring a label is free. At $e = 0.10$ the bracket is 0.368, about 53 percent of the complete separation value. At $e = 0.5$ it vanishes even though a nominal label was printed for every particle. If one outlet loses particles or the two error probabilities differ, compute actual outlet fractions and chemical potentials from a material balance instead of using this symmetric formula.

An experiment can close that balance without seeing individual molecular trajectories. Measure both outlet amounts, composition, pressure and temperature, and account separately for work or heat supplied to the sorter. Compare the Gibbs free-energy change at controlled pressure with the actual electrical work. A device consuming several times the reversible minimum can still separate effectively. Conversely, apparent work below the calculated minimum suggests an omitted material stream, a

different final pressure, an energy supplied by an uncounted reservoir or uncertainty in composition. The mixing formula is only as meaningful as the thermodynamic endpoints used.

Follow isotope and quantum effects to the edge of the classical formula

The appearance of h in a classical-looking entropy sometimes suggests that the Sackur–Tetrode expression already describes any temperature so long as gas interactions are weak. These are two different approximations. Diluteness relative to the interaction range controls the importance of forces. Diluteness in phase space, measured by $n\lambda_T^3$, controls whether quantum exchange changes occupation weights. Atoms can interact weakly while exhibiting strong exchange statistics. Lowering temperature makes the thermal wavelength grow as inverse square root of temperature, quickly invalidating the classical count at a fixed density.

Consider two isotopes with one-particle masses in the ratio four to three, neglecting internal degeneracy differences and comparing at the same density and temperature. Their wavelengths have ratio $\sqrt{3/4}$. The heavier isotope has a translational molar entropy greater by $(3R/2)\log(4/3) = 3.59$ J per mole kelvin. This does not contradict the mass independence of the ideal mixing increment. The former compares two absolute entropy values of different substances; the latter is a change in each substance's accessible translational volume while its mass remains fixed.

The same mass scaling influences onset of quantum degeneracy. At fixed n, the temperature for a chosen value of $n\lambda_T^3$ varies inversely with mass. A heavier isotope reaches that value at a lower temperature. If the nuclei also have different spin degeneracies, distinguish the dimensionless density per internal component from the total density. Two isotopes can even have different Bose or Fermi exchange statistics, so a high-temperature entropy comparison need not extrapolate into similar low-temperature heat capacities.

A calorimetric consistency test illustrates the limits. Suppose an idealised solid has a measured low-temperature molar $C_p = AT^3$ up to 5 K, with $A = 0.20$ mJ per mole per kelvin to the fourth. Its entropy increment there is $AT^3/3 = 0.00833$ J per mole kelvin. Missing that contribution hardly changes a room-temperature entropy tabulated to a few tenths of a joule per mole kelvin. In contrast, omitting a latent heat of 1.0 kJ/mol at a transition near 100 K misses 10 J per mole kelvin. The low-temperature extrapolation and the phase transition enter the same entropy integral but can carry very different experimental leverage.

A finite thermometer measures a ratio of multiplicities

Temperature from an entropy derivative presupposes sufficiently smooth variation with energy. In a small discrete environment a two-level probe of gap Δ can instead measure an exact finite difference. If the whole isolated system has total energy E and the two probe states have equal degeneracy, their equilibrium odds are $\Omega_R(E-\Delta)/\Omega_R(E)$. Write the logarithm as $[S_R(E-\Delta) - S_R(E)]/k_B$. Interpreting it as $-\Delta/(k_B T_{\mathrm{probe}})$ defines an operational inverse temperature averaged over an interval of width Δ.

For a reservoir whose cumulative multiplicity is proportional to E^{α}, this gives $\beta_{\mathrm{probe}} = \alpha \log[E/(E-\Delta)]/\Delta$, while the local slope at E gives α/E. With $\Delta/E = 0.10$, the finite-gap value exceeds the endpoint slope by 5.36 percent. A second probe with one fifth that fractional gap reports only about a one-percent excess. Two well-calibrated probes can therefore disagree in a finite reservoir while each correctly reports its specified probability ratio.

The distinction vanishes for a macroscopically large heat bath whose entropy is locally almost linear over both gaps. It is experimentally testable in numerical energy-sharing models without calibrating a mechanical pressure. Count both occupancy ratios, establish whether the probe appreciably changes the reservoir's allowed states and compare against the exact discrete multiplicities before invoking a smooth temperature. If repeated measurements perturb the environment, the measured distribution is a sequence of new preparations and the equilibrium interpretation needs an explicit readout model.

Construct a thermodynamic cycle that isolates mixing entropy

The chemical-potential formula suggests a pressure measurement for testing mixing free energy. Let one side of a movable membrane transmit a solvent species freely and block a second dilute gas species, with common temperature maintained. In the ideal-gas case the permeable species equilibrates its partial pressure across the membrane. If on one side it is pure at pressure p_A, while on the other side it has mole fraction x at total pressure p_{mix}, equality gives

$$p_A = x p_{\mathrm{mix}}.$$

The remaining pressure load on the membrane is the partial pressure of the blocked species. This is a mechanical consequence of equality of the permeable component's chemical potential.

For an ideal mixture with $x = 0.80$ and total pressure 1.00 bar, the pure compartment must reach 0.80 bar for zero net flow of the permeable species. A square centimetre membrane then supports a net pressure difference of 0.20 bar, corresponding to 2.0 N, before any additional elastic support forces. A membrane sensitive only to the total pressure might move and alter the volumes until mechanical and chemical constraints are both fulfilled. The quoted force is a held-geometry measurement, not a prediction that a freely moving unsupported membrane can remain at that position.

A particularly important control switches the membrane for a nonselective porous partition. Both species can then flow and equilibrate their own chemical potentials; at final full equilibrium there is no maintained composition difference across these gas reservoirs. A pressure reading alone cannot establish whether selectivity, partial blockage, a small temperature difference or an external pump generated the observed imbalance. Measure both compositions and isolate slow leaks over the relevant time. An apparently stable jump on a short trace can be a long-lived transient.

The same thought experiment gives a check on how gas-phase mixing connects to osmotic pressure in dilute solutions. The permeable solvent's chemical potential is lowered by mixing at fixed total pressure; applying enough pressure to the solution can restore equality across a solvent-selective membrane. The ideal dilute result has a form proportional to solute number density times thermal energy, although a liquid solution requires its own solvent partial molar volume model. Equal-looking formulas arise from a

chemical-potential balance, not because dissolved solute molecules necessarily act as an ideal gas physically bouncing against every membrane.

Finally, a cyclic device must restore membrane positions, reservoir compositions and temperatures as well as a piston reading. Ignoring the work to regenerate selectivity or restore separated gas streams can turn a one-use release of stored mixing free energy into an apparent heat engine operating on one bath. Write the mole count of each component in every reservoir at start and finish and integrate mechanical and chemical work consistently. The criterion for a true thermodynamic cycle is return of the complete controlled system to its initial state.

A final microscopic check concerns the dimensions of the count. If one changed the formal classical cell size from h^3 to ah^3 per particle, the gas entropy would shift by $-Nk_B \log a$ at every fixed state. Pressure, temperature and the entropy change of a fixed-N expansion would be unaffected. Chemical potential would shift by $k_B T \log a$, and absolute comparisons between different particle species or chemical reactions would require a consistently calibrated microscopic count. This is why a piston experiment alone could never determine the natural action cell. Combining calorimetry with quantum level counting or fixed-reference chemical equilibria supplies information absent from fixed-N mechanical derivatives. The same bookkeeping distinction applies to any unobserved constant internal degeneracy: it can contribute entropy without producing a heat-capacity anomaly.

Connect ideal mixing to a separation-work measurement

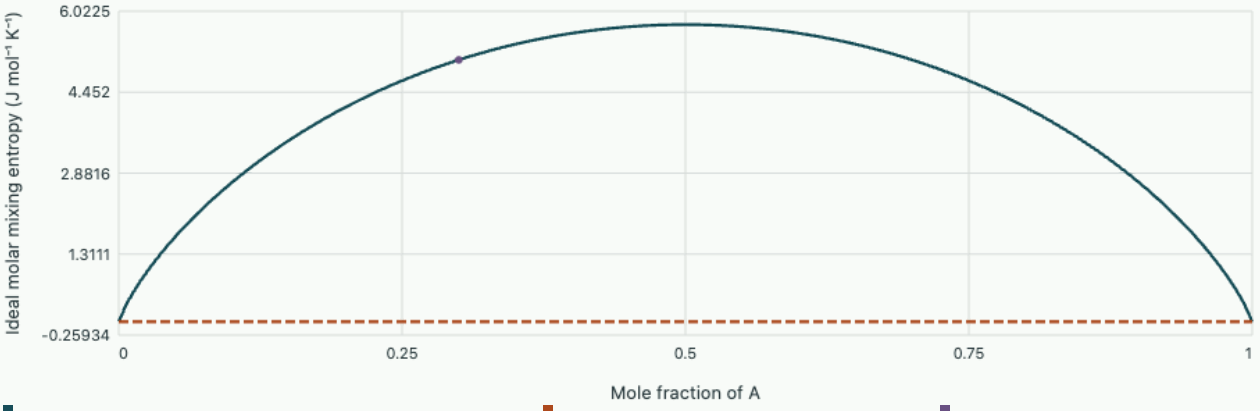
Vary the mixture composition and the reservoir temperature. Compare the entropy with reversible work for separately identifiable ideal components.

Open this experiment online

Try the calculation

Mole fraction of component A **0.3**

Separation reservoir temperature (K) **300**



Two distinguishable ideal components | Identical gas on both sides | Selected composition

Ideal entropy increase 5.079 J mol⁻¹ K⁻¹. Reversible isothermal separation of the distinguishable mixture requires 1.5237 kJ/mol.

Tutorial · Measure an entropy difference by two routes

A 0.50 mol sample of ideal monatomic gas expands from 2.0 to 5.0 litres in a thermally insulated free expansion. Calculate entropy change and change of chemical potential at 290 K. Now identify what changes if the gas instead expands reversibly and adiabatically between the same volumes.

▼ **Hold the state variables appropriate to each route**

For free expansion temperature stays at 290 K and entropy increases by $0.50R\log(2.5) = 3.81$ J/K. At fixed temperature $\mu_f - \mu_i = k_B T \log(n_f/n_i) = -k_B T \log 2.5$, or minus 2.21 kJ/mol. The chemical-potential difference is a single-particle or molar quantity, while the entropy answer refers to the entire sample. Dividing or multiplying by the sample amount twice produces an inconsistent unit.

In a reversible adiabatic expansion $\Delta S = 0$ and the gas does boundary work. The final temperature becomes $290(2/5)^{2/3} \simeq 157$ K. These are not the same endpoints as in the free expansion, despite identical start and end volumes. Substituting the free-expansion entropy into an adiabatic isentropic relation would falsely equate insulation with reversibility.

Tutorial · An isotope mixture and a finite-size count

Two equal amounts of isotopically distinct ideal monatomic gases at the same temperature and density are mixed. Explain why their masses enter absolute chemical potentials but cancel from each component's ideal isothermal mixing increment. For just ten sites individually occupied at random by one

of two distinguishable isotope labels, compare the exact number of arrangements with five of each label to the leading extensive entropy.

▼ **Cancel fixed internal terms and keep the finite combinatorics**

Each component's thermal wavelength is unchanged between its own initial and final states. Taking the difference of its entropy therefore removes the mass and internal factors. Both species expand by a factor of two and the large-sample increase is $Nk_B \log 2$. This conclusion assumes independently identifiable species and an ideal mixture, and does not state that their absolute entropies are identical.

For ten sites with five labels of each kind there are $\binom{10}{5} = 252$ arrangements. Their logarithm is 5.529 in units of k_B, versus $10 \log 2 = 6.931$ from an extensive approximation neglecting the composition constraint's subextensive term. Summing over all possible label counts gives 2^{10}, recovering the latter count for an ensemble in which the composition itself fluctuates. A small-system discrepancy is an opportunity to ask which constraint the count imposes.

Tutorial · Infer what a calorimetric entropy determination misses

A crystal's heat capacity is measured from a low reference temperature up to a temperature above a reversible melting transition. How could this measurement be used to estimate entropy, and which missing physical contributions could spoil an apparent comparison with a dilute gas entropy?

▼ **Include the low-temperature reference and all phase changes**

Along a constant-pressure reversible route add $\int C_p(T) dT/T$ in each phase and add latent heat divided by transition temperature at every first-order change. To compare with gas one must also traverse vaporisation or another controlled path to the desired gas state and pressure. Missing a transition silently omits its finite entropy even if the heat capacities in both adjacent phases were measured perfectly.

Below the lowest measured temperature extrapolate using a physically justified low-temperature model and add any residual ground-state entropy consistent with the sample preparation. A unique ordered ground state has zero thermodynamic entropy in the ideal third-law limit. A crystal retaining configurational disorder on laboratory timescales can carry residual entropy associated with its accessible arrangements. Unknown isotopic composition or low-lying internal levels likewise change the absolute comparison. A discrepancy between a calorimetric integral and a monatomic ideal expression should first be traced to these explicit reference states and model assumptions.

Cambridge Statistical Physics · foundational postulates, entropy and statistical ensembles

Fließbach, Statistische Physik · statistical foundations, the ideal gas and entropy

Ensembles, thermodynamic potentials and measurable curvature

Statistical mechanics from microscopic counting to material response

A sample sealed inside an insulated container, a sample weakly coupled to a temperature bath and a surface exchanging particles with a reservoir need different probability laws. Derive the three laws from one composite-system count. Then use the curvature of their generating functions to predict noise, response and parameter uncertainty without assuming that microscopic quantities take their mean values in every run.

Expand the reservoir count rather than postulating a probability

Let a small system occupy a microstate i of energy E_i while a weakly coupled reservoir holds the remainder of a fixed total energy. The composite system is isolated. Conditional on i, the number of compatible reservoir states is $\Omega_R(E_{\text{total}} - E_i)$. This weight is the same for degenerate system microstates of equal energy. For a macroscopic reservoir, expand its entropy about an appropriate reference energy. The first derivative is $1/T$, and its second derivative at fixed other extensive parameters is $-1/(T^2 C_R)$.

$$S_R(E_R - \epsilon) = S_R(E_R) - \frac{\epsilon}{T} - \frac{\epsilon^2}{2T^2 C_R} + \cdots.$$

Exponentiating and taking the limit where the quadratic and further corrections are negligible yields $p_i = e^{-\beta E_i}/Z$, $\beta = 1/(k_B T)$, and $Z = \sum_i e^{-\beta E_i}$. A constant shift in all system energies multiplies both numerator and partition sum by the same factor. Relative probabilities are therefore independent of energy-zero convention. If a calculation sums over energy values rather than microstates, insert their degeneracies explicitly. These two equivalent ways of counting should never be applied simultaneously.

The derivation assumes weak coupling so that assigning energies separately makes sense, and a reservoir whose temperature does not respond significantly to typical exchanges. A small heat capacity in the bath can suppress sample fluctuations even at the same mean temperature. The relevant test concerns excursions explored by the sample, not simply whether the bath contains more particles. Near a transition or in an engineered small reservoir a canonical distribution may need a measurable correction.

Compare sample-energy fluctuations for finite and effectively infinite heat reservoirs

Differentiate the normalisation to obtain energy and its variance

Differentiate $\log Z$ with respect to β at fixed energy levels. The first derivative yields minus mean energy. A second derivative gives the mean square minus the squared mean. Changing differentiation from β to temperature gives the equilibrium heat capacity:

$$\langle E \rangle = -\partial_\beta \log Z,$$

$$\operatorname{var}(E) = \partial_\beta^2 \log Z = k_B T^2 C_V.$$

This proves nonnegative canonical heat capacity for a temperature-independent Hamiltonian. If the model Hamiltonian contains parameters that themselves vary with temperature after other degrees of freedom have been eliminated, the derivative also acts on those parameters. Inserting measured effective levels into the simple identity without that dependence need not reproduce the calorimetric capacity.

A two-level system has energies zero and Δ, with upper probability $p = 1/(1 + e^{\beta\Delta})$ for equal degeneracies. Its variance is $\Delta^2 p(1-p)$ and $C/k_B = (\beta\Delta)^2 p(1-p)$. At $\Delta/k_B T = 2$, $p = 0.1192$ and $C = 0.420 k_B$. In one thousand independent copies the expected upper count is 119 and its standard deviation about 10.2. An actual count of 130 is unremarkable. An experiment resolving only one sample at one time cannot identify the variance from its energy alone.

Higher derivatives of $\log Z$ generate cumulants. A Gaussian approximation discards cumulants beyond the second. For a small two-level population near low occupancy the histogram is skewed and the approximation can yield artificial negative occupancies. An exactly binomial count or a direct discrete partition sum handles this elementary case without assuming that the central-limit regime has arrived. Sample-to-sample correlations can also prevent variances from scaling as the sum of independent variances.

Why the Helmholtz free energy knows the same equilibrium as entropy

Group the canonical sum by energy. A shell with entropy $S(E)$ has weight proportional to $e^{S(E)/k_B - \beta E}$. For a macroscopic system with a dominant smooth interior maximum this exponent is stationary at $(\partial S/\partial E)_{V,N} = 1/T$. The leading free energy $-k_B T \log Z$ is then the minimum of $E - TS(E)$; the width of the peak gives the energy fluctuations. The exact finite-system integral is a Laplace transform. Replacing it by its stationary contribution is an approximation, often called the Legendre-transform limit.

An equivalent variational argument works for finite discrete states without a saddle point. Choose arbitrary trial probabilities q_i, their mean energy $U[q]$ and Shannon entropy $S[q] = -k_B \sum_i q_i \log q_i$. Subtract the canonical free energy and insert the canonical weights. Algebra gives

$$U[q] - TS[q] - F = k_B T \sum_i q_i \log(q_i/p_i) \geq 0.$$

The final inequality follows from convexity of the logarithmic divergence and becomes an equality only for the canonical distribution on the permitted support. For an isolated system one instead maximises entropy subject to its prepared energy constraint. Helmholtz free-energy minimisation at fixed temperature is thus a consequence of reservoir exchange and normalisation, not an additional force sending every microstate to the lowest energy. At finite temperature high-multiplicity excited states can dominate the equilibrium probability.

Now include volume work in the fundamental differential $dU = TdS - pdV + \mu dN$. Product differentiation gives $dF = -SdT - pdV + \mu dN$, $dG = -SdT + Vdp + \mu dN$, and $d\Omega = -SdT - pdV - Nd\mu$ for $\Omega = F - \mu N$. A potential's natural variables specify which gradient directly returns an observable. From $F(T, V, N)$ one obtains

$$p = -(\partial F/\partial V)_{T,N},$$
$$S = -(\partial F/\partial T)_{V,N};$$

their cross derivatives agree if the potential is smooth. At phase boundaries use phase branches and appropriate coexistence conditions instead of differentiating through a kink as if it were smooth.

Let the reservoir exchange particles as well as energy

With both exchanges allowed, expand reservoir entropy in E and N. Its differential includes $-\mu_R dN_R/T$. A system microstate with E, N therefore acquires weight $e^{-\beta(E-\mu N)}$. Summing normalises the grand canonical distribution. At fixed temperature and volume, derivatives with respect to $\beta\mu$ give mean particle number and number variance. In a homogeneous equilibrium region with suitable thermodynamic limit, $\Omega = -pV$.

$$\Xi = \sum_N e^{\beta\mu N} Z_N,$$
$$\langle N\rangle = \partial_{\beta\mu} \log \Xi,$$
$$\mathrm{var}(N) = k_B T \left(\frac{\partial \langle N\rangle}{\partial \mu}\right)_{T,V}.$$

A surface of M independent adsorption sites, each empty or holding one particle with binding energy $-\epsilon$, has $\Xi = [1 + e^{\beta(\mu+\epsilon)}]^M$. Its occupancy fraction is logistic and its variance $Mf(1-f)$. At $M = 400$ and half coverage, the mean occupancy is 200 with standard deviation 10. The slope with respect to chemical potential is largest there. Near saturation, a small number variance and an apparently quiet detector indicate poor sensitivity to further increases of μ, not necessarily high precision in inferring μ.

Compare the adsorption response slope with the independently predicted number variance

The grand formula links number fluctuations to isothermal compressibility in a bulk one-component fluid. Write $n = \langle N\rangle/V$ and use $dp = n\,d\mu$ at constant T. Then $(\partial n/\partial\mu)_T = n^2\kappa_T$ and $\mathrm{var}\,N/\langle N\rangle = nk_BT\kappa_T$. A dilute classical ideal gas gives unity. These fluctuations refer to a region exchanging particles. If an observer counts all particles of a strictly sealed closed box, the number variance is zero even though its compressibility is not. The boundary condition of the measurement is part of the formula.

Infer a temperature and explain when it is identifiable

The logarithm of a measured population ratio is linear in β for nondegenerate known levels. With enough independent samples one can estimate temperature either from this ratio or by matching measured energy to its canonical expectation. If all observed states have nearly the same energy, changing β barely changes their probabilities. The available temperature information then becomes small, regardless of detector precision.

For independent canonical draws, differentiating a single log probability gives

$$\partial_\beta \log p_i = -(E_i - \langle E \rangle).$$

Its squared expectation is energy variance, the Fisher information per observation about inverse temperature. At n_s independent draws information adds. Equality between response and noise reappears as a statistical bound on temperature estimation. A long record with correlation time comparable to its length may contain few independent draws despite a large sample count. Recording more frequently without extending the observation window need not improve a temperature estimate.

A pressure reservoir selects enthalpy-weighted volume fluctuations

An isothermal piston in contact with a pressure reservoir supplies a fourth ensemble closely related to the first three. A configuration at volume V costs external work $p_{\mathrm{ext}} V$ in addition to its microscopic system energy. At fixed T, N, p_{ext}, integrate the canonical configurational partition over possible piston positions with weighting $e^{-\beta p_{\mathrm{ext}} V}$. For an ideal gas in a cylinder of fixed cross-sectional area the piston coordinate is proportional to V, so its geometric measure is a constant times dV. Mechanical kinetic factors independent of V cancel from its normalised distribution.

For $Z_N(V) \propto V^N$, the probability density for volume is proportional to $V^N e^{-\beta p_{\mathrm{ext}} V}$ on the positive real axis. Normalisation uses $\int_0^\infty V^N e^{-aV} dV = N!/a^{N+1}$. The maximum, mean and variance are

$$V_{\mathrm{mode}} = \frac{N}{\beta p_{\mathrm{ext}}},$$

$$\langle V \rangle = \frac{N+1}{\beta p_{\mathrm{ext}}},$$

$$\mathrm{var}(V) = \frac{N+1}{(\beta p_{\mathrm{ext}})^2}.$$

The finite-N offset between mode and mean comes from the skewed piston-volume distribution. The most likely volume obeys the usual ideal gas law evaluated using N, while the exact mean differs in this simple piston model. At $N = 99$ the difference is about one percent and relative volume fluctuations are ten percent. At thermodynamic sizes the fractional difference vanishes. When calibrating nanoscale pressure from a histogram, a mean, a peak and a mean of reciprocal volume should therefore not be treated as interchangeable.

Check the force equilibrium directly. The instantaneous gas pressure conditional on V is $N k_B T / V$. The gamma distribution has $\langle 1/V \rangle = \beta p_{\mathrm{ext}} / N$, so its mean gas pressure equals the imposed external pressure exactly. Substituting the mean volume into the nonlinear pressure formula instead gives $N p_{\mathrm{ext}} / (N+1)$. This calculation illustrates a general statistical rule: mechanical balance concerns the mean actual force, not necessarily that force evaluated at the mean position.

If the piston lies horizontally and dry static friction is appreciable, configurations can remain stuck with nonzero force imbalance. A distribution fitted to the equilibrium piston formula would then return

apparent biases in pressure or particle number. Reverse the pressure change and vary waiting times to expose hysteresis, or remove the mechanical sticking before interpreting volume fluctuations as an equilibrium susceptibility.

A matrix of fluctuations controls coupled responses

Many experiments permit exchange of more than one extensive quantity. Write a general ensemble with natural control parameters θ_a and weights proportional to $e^{\sum_a \theta_a X_a}$, with a fixed set of microscopic state values X_a. The first derivatives of the log normalisation give their means and the second derivatives give the full covariance matrix. Off-diagonal entries describe coupled responses. For the grand canonical ensemble choose $\theta_1 = -\beta$, $\theta_2 = \beta\mu$, $X_1 = E$, $X_2 = N$. One must keep the other natural parameter fixed while taking a partial derivative in this form.

A covariance matrix is positive semidefinite because the variance of any linear combination of the X_a is nonnegative. Consequently $\operatorname{var}(E)\operatorname{var}(N) \geq \operatorname{cov}(E, N)^2$. A measured matrix violating this inequality cannot describe these two observables on the same equilibrium sample space. Possible causes include mixing variances and covariances from different sample preparations, subtracting a poorly estimated instrumental background or an inconsistent energy reference when counting particles.

Suppose a volume exchanges particles each carrying a fixed energy e_0 plus other thermal energy. If the system energy reference shifts by a per particle, new energy is

$$E' = E + aN.$$

Its covariance with N changes by $a \operatorname{var}(N)$, although no physical heat changed. A useful combination for isolating fluctuations uncorrelated with particle transport is $E - bN$, with $b = \operatorname{cov}(E, N)/\operatorname{var}(N)$. Minimising the variance with respect to b gives residual variance $\operatorname{var}(E) - \operatorname{cov}(E, N)^2/\operatorname{var}(N)$, invariant under this reference shift. In a general nonlinear conditional distribution that residual is a best linear prediction variance, not automatically the exact variance conditional on a fixed integer N.

For a unit-normalised example, suppose energy variance is $25\epsilon^2$, number variance is 4 and their covariance is 8ϵ. The best linear energy increment per exchanged particle is 2ϵ, and the residual energy variance is $9\epsilon^2$. Replacing energy by $E - 3\epsilon N$ changes the covariance to -4ϵ, but recalculating the residual returns the same value. A data-analysis code can use this reference-invariance as a check on both numerical units and susceptibility derivatives.

The same covariance structure controls simultaneous inference. If two state observables are nearly proportional over all sampled configurations, the log-likelihood curvature has a nearly zero eigenvalue. The individual fields conjugate to them are then hard to determine even if one linear combination is tightly constrained. Preparing the system in a region with different accessible states or adding an independent observable can resolve the degeneracy. Increasing the sampling rate on exactly the same narrow manifold cannot create a missing independent direction.

When the maximum-probability replacement loses a phase

The passage from a Laplace transform to a Legendre transform assumes one dominant smooth contribution or handles multiple branches explicitly. Near a first-order transition two energy ranges can have comparable canonical weights. Replacing the energy distribution by a single Gaussian around its mean then produces a state that may be relatively improbable, lying in the trough between two peaks. The finite partition sum remains smooth, but its derivatives can vary rapidly over a narrowing range of temperature.

In a simple two-phase approximation take branch partition contributions proportional to $e^{-\beta F_1(T)}$ and $e^{-\beta F_2(T)}$. Their ratio is $e^{-\beta(F_2-F_1)}$. At coexistence the leading bulk free-energy densities agree. If their entropy densities differ, then a small temperature change produces an extensive branch free-energy difference proportional to volume times that temperature shift. The equal-weight finite-size window therefore narrows inversely with volume in this simplest picture. Interfacial states, boundaries and phase degeneracy can shift detailed crossing locations.

Consider an energy histogram with equal phase probabilities and sharply peaked branch energies differing by total ΔE. The mixture variance includes $p(1-p)(\Delta E)^2 = (\Delta E)^2/4$ in addition to within-phase fluctuations. As ΔE is extensive this gives an extensive heat-capacity peak per particle in a finite coexistence window, much larger than the smooth single-phase contribution. This does not imply an infinite heat capacity throughout an interval of bulk temperature. The peak becomes higher and narrower and encodes latent heat in the thermodynamic limit.

An experimental or simulated time series can undersample one branch when interfacial barriers make switching rare. Estimating the variance from one branch then yields a completely different apparent heat capacity from a record containing a few switches. Compare heating and cooling preparations, identify branch-specific energy distributions and establish a reproducible equilibrium weight ratio before treating the union as a canonical histogram. In a macroscopic sample, calorimetric latent heat often provides a more reliable bulk quantity than a noisy attempt to observe whole-sample switching.

Use one equilibrium histogram to predict a nearby temperature

If a canonical run at inverse temperature β_0 samples energies E_i, a nearby expectation at β_1 can be expressed as a ratio of weighted averages with factors $w_i = e^{-(\beta_1-\beta_0)E_i}$. Multiplying numerator and denominator by the same constant allows one to subtract a reference energy to maintain numerical precision. Mathematically the ratio is exact when the original distribution covers every important configuration. Finite data can render an exact identity practically useless outside its overlap window.

For an approximately Gaussian energy histogram of standard deviation σ_E, reweighting shifts its influential energy region by roughly $-(\beta_1-\beta_0)\sigma_E^2$. Expressed in original standard deviations the shift is $|\beta_1-\beta_0|\sigma_E$. A value of 0.3 stays near the observed centre; a value of five asks for configurations five original widths away. At the latter value a smooth computed curve can be governed almost entirely by one observation at the extreme edge of the histogram.

Inspect the effective weight count $(\sum_i w_i)^2/\sum_i w_i^2$ alongside the raw record length. This statistic does not account for time correlation or prove that missing configurations are irrelevant, but a collapse to a

handful of weights is a clear failure of overlap. Simulate or measure directly at a second temperature displaced toward the target and compare its actual histogram with the reweighted prediction. This connects the formal equivalence of ensemble descriptions to a concrete question of measurement coverage and uncertainty.

A finite-spin thermodynamic cycle tests mixed derivatives

Consider independent two-state magnetic moments with energies $\mp\mu B$ in a prescribed induction B. Ignore interactions and orbital response. Per moment the canonical sum is $2\cosh x$, $x = \mu B/(k_B T)$. Its Helmholtz free energy is $-k_B T \log(2\cosh x)$ and the average magnetic moment is

$$m = -\partial F/\partial B = \mu \tanh x.$$

The entropy follows from $-\partial F/\partial T$, yielding $S/k_B = \log(2\cosh x) - x\tanh x$. These component equations define an exactly soluble controlled system in which thermal and magnetic responses are coupled.

Equality of the mixed free-energy derivatives predicts

$$(\partial S/\partial B)_T = (\partial m/\partial T)_B.$$

Both are negative for positive field in this model. Increasing field at fixed temperature aligns spins and reduces their entropy. If the spin system is instead isolated reversibly, its entropy is constant, which requires x constant away from singular limiting cases. Reducing field then reduces temperature in proportion to field. This is the ideal paramagnetic demagnetisation argument, derived here from the canonical weights rather than an empirical cooling rule.

For an idealised system prepared at 3.0 K and 1.5 T, reducing the field quasistatically to 0.50 T under thermally isolated conditions gives 1.0 K. Its mean alignment fraction stays constant even as the Zeeman gap shrinks. A real specimen includes lattice heat capacity, residual interactions and inevitable heat leakage. If the lattice remains equilibrated with the moments, the total entropy includes both subsystems and the simple proportional temperature drop is modified. If spin–lattice exchange is too slow, a single temperature may fail to describe both during the field ramp.

A numerical test can use a rectangular loop in the T, B plane. Evaluate the change in F along one path using the integral of $-S\,dT - m\,dB$, then return along the other path. At successively finer quadrature the closed integral should approach zero. By contrast heat and magnetic work evaluated separately generally have nonzero loop integrals and cancel only in their appropriate state-function combination. A nonzero free-energy residual that does not shrink under refinement points toward a sign error, an inconsistent field-energy convention or different energy references on the two paths.

Finally use the weak-field limit to estimate experimental sensitivity. There $m \simeq \mu^2 B/(k_B T)$, so a thermometer inferred from magnetisation is sensitive to both the magnetic-field scale and the independent number of active moments. Saturation measurements can calibrate their total number, while a field reversal removes a constant magnetometer offset. A mixed-derivative test based on

independently measured calorimetry and magnetisation is more discriminating than fitting a temperature parameter to a single magnetic trace that already assumed the predicted population law.

The spin example also gives a direct thermal-noise measurement of susceptibility. For one moment the variance of its projection is $\mu^2[1-\tanh^2 x]$, and differentiating mean moment with respect to induction gives exactly this variance divided by thermal energy. At $x=1$, magnetisation is 0.762 of saturation and variance is $0.420\mu^2$. For a thousand independent moments the total rms fluctuation is about 20.5μ, even though their mean total moment is about 762μ. An instrument averaging over many reversals may see the mean accurately and suppress the variance; replaying the same averaging operation on a simulated equilibrium time series makes the comparison meaningful. A direct susceptibility inferred from a very large field modulation would instead sample the curved magnetisation over a finite interval. Use a small symmetric modulation and verify its amplitude scaling to compare with the local fluctuation slope.

Count the configurations of two finite reservoirs

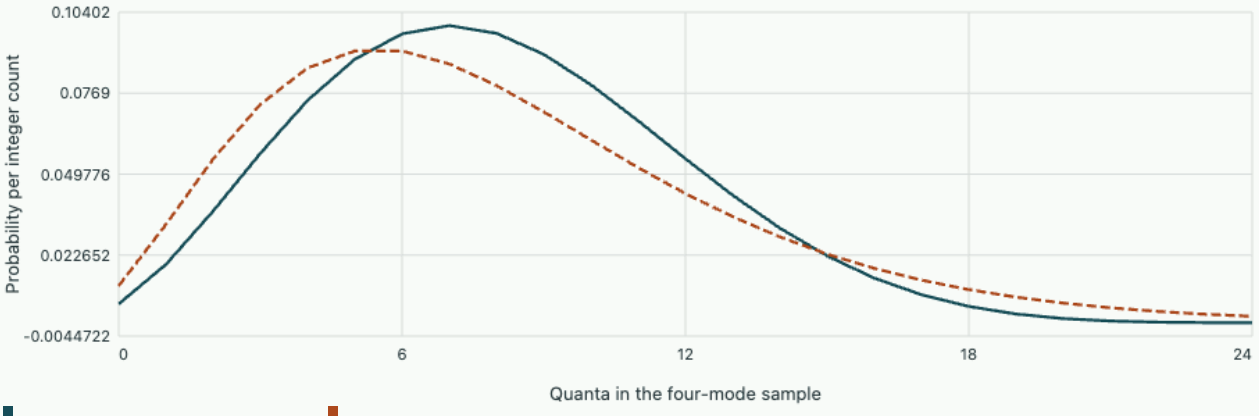
Change the reservoir size and compare the observed energy histogram with a canonical ensemble at the same mean.

Open this experiment online

Try the calculation

Reservoir modes **8**

Excitation quanta **24**



Mean 8 quanta; closed-pair variance 14.769. Canonical variance 24.

Tutorial · Can a finite bath explain unusually quiet energy data?

A sample has canonical heat capacity $60k_B$ at the mean temperature. Contact with a small reservoir reduces the measured equilibrium energy variance to three quarters of its canonical value. Approximate

both entropies quadratically near their common equilibrium and infer the bath heat capacity. What other measurement could mimic the reduction?

▼ **Add the two inverse susceptibilities and check the observation**

The joint log probability has curvature $-[1/C_s + 1/C_R]/(k_B T^2)$. Hence the effective variance is $k_B T^2 C_s C_R/(C_s + C_R)$. Dividing by $k_B T^2 C_s$ and setting the ratio to $3/4$ gives

$$C_R = 3C_s = 180 k_B.$$

The temperature at maximum probability still agrees for bath and sample at leading order. The suppressed fluctuations reflect their shared total-energy constraint.

Finite temporal averaging of a fast energy signal can also suppress measured variance. Vary detector bandwidth and check convergence of the measured area, then repeat with a different independently estimated reservoir capacity. A corrected variance scaling with the harmonic combination supports the finite-bath interpretation. A single reduced number cannot distinguish thermodynamic curvature from the detector response.

Tutorial · Determine a volume response from temperature data

For a fixed amount of a smooth single-phase material near one operating point, the thermal expansion coefficient is α, the isothermal compressibility is κ_T and the volume is V. Derive $C_p - C_V$ from a Maxwell relation. Evaluate it for $T = 300$ K, molar volume 1.0×10^{-5} cubic metres, $\alpha = 6.0 \times 10^{-5}$ per kelvin and $\kappa_T = 1.0 \times 10^{-11}$ per pascal.

▼ **Eliminate the unmeasured mixed entropy derivative**

Write $dS = (C_V/T)dT + (\partial p/\partial T)_V dV$, using the Helmholtz mixed derivative. On the constant-pressure path $dV = \alpha V dT$. Holding volume fixed in $dV/V = \alpha dT - \kappa_T dp$ gives

$$(\partial p/\partial T)_V = \alpha/\kappa_T.$$

Therefore $C_p - C_V = TV\alpha^2/\kappa_T = 1.08$ J per mole kelvin. Positive isothermal compressibility and positive temperature imply the difference is nonnegative even for negative expansion coefficient.

Tutorial · Infer two unknowns from occupancy measurements

At two distinct temperatures a set of dilute independent adsorption sites gives coverages 0.20 and 0.60. Why can these two readings fail to identify binding energy if the reservoir chemical potentials are not independently known? Express the unknown combination and suggest an experiment that would determine the binding energy.

▼ Write an equation for each observation before counting parameters

The log odds satisfy

$$\log[f/(1-f)] = (\mu + \epsilon)/(k_B T).$$

The measured log odds are minus 1.386 and plus 0.405. They determine the value of $\mu + \epsilon$ in each run multiplied by its known inverse thermal energy. With two unrelated reservoir potentials there are three unknown dimensional quantities and only two equations. Increasing coverage precision cannot resolve that degeneracy.

A separately calibrated reservoir supplies $\mu(T, p)$ relative to a consistent energy reference. Then several pressures at one temperature verify the model and several temperatures can test a constant binding energy against an additional binding entropy or site degeneracy. A reported equilibrium constant alone can always absorb a reference-state shift; the reference reservoir and site standard state are part of a physically interpretable adsorption energy.

Cambridge Statistical Physics · canonical and grand canonical ensembles

Pair correlations, the virial gas and a pressure measurement

Statistical mechanics from microscopic counting to material response

A dilute nonideal gas can display almost ideal pressure while its particles strongly avoid one another at short distance. Connect a two-particle interaction to pressure, internal energy and a measurable pair correlation. Then use the resulting integral equations to ask which features of a potential can actually be inferred from a limited set of pressure readings.

Separate kinetic motion from the configurational count

For a classical monatomic gas at fixed temperature take pair potential $u(r)$ and total interaction $\sum_{i<j} u(r_{ij})$. The canonical partition function factors into a Gaussian momentum part and an integral over all allowed particle positions. Divide by $N!\lambda_T^{3N}$ for indistinguishable particles. In the absence of interactions the configurational integral is V^N. A hard core removes configurations with overlapping particles, while an attractive well increases their equilibrium weights instead of adding geometric space.

Use the Mayer function $f(r) = e^{-\beta u(r)} - 1$. The position-dependent Boltzmann weight is exactly $\prod_{i<j}[1 + f(r_{ij})]$. Expanding the product separates independent particles, terms with one pair bond, and terms involving several bonds. Translation invariance allows each isolated pair integral to be written as $V \int f(\mathbf{r}) d^3r$ when boundary corrections are negligible. The logarithm of the configurational partition function retains connected contributions, a crucial step before taking the large-volume limit.

$$\frac{F}{N k_B T} = \log(n\lambda_T^3) - 1 + n B_2(T) + O(n^2),$$

$$B_2 = -\frac{1}{2}\int f(\mathbf{r})\, d^3r.$$

A naive truncation of the full partition function after its first pair term can fail even for small density because there are of order N^2 pairs. The connected expansion organises the intensive free energy, not the sum of all events in a macroscopic gas. Differentiating F with respect to volume at fixed T, N gives

$$p/(n k_B T) = 1 + n B_2 + \cdots.$$

It is the smallness of successive density terms that controls the virial prediction. A small accidental value of B_2 does not bound the next coefficient.

Compute excluded volume and attraction as separate contributions

For a spherical hard core of diameter σ, $f = -1$ for $r < \sigma$, zero otherwise. The second virial coefficient is $2\pi\sigma^3/3$, four times the physical volume of a sphere of diameter σ. The forbidden region for the centre of a second particle has radius σ, twice a particle's radius, and each unordered pair is counted once. Treating the excluded volume per particle as merely its physical volume loses both geometric factors.

A square well of depth $\epsilon > 0$ extends from σ to $\lambda\sigma$. Outside the well the interaction vanishes. Evaluate the shell integrals directly:

$$B_2 = b[1 - (\lambda^3 - 1)(e^{\epsilon/(k_B T)} - 1)],$$

$$b = 2\pi\sigma^3/3.$$

The Boyle temperature is the solution $B_2 = 0$, namely $k_B T_B/\epsilon = 1/\log[1 + 1/(\lambda^3 - 1)]$. At T_B hard-core and attractive pair contributions cancel in pressure through first order in density. Particles still have nontrivial correlations and excess internal energy. The Boyle point is a pressure cancellation, not a noninteracting gas state. At fixed T and very low density the absolute pressure correction is small even when $|B_2|/b$ is large.

For $\lambda = 1.50$ and $\epsilon/k_B = 100$ K, $T_B \simeq 285$ K. At 200 K the ratio $B_2/b = 1 - 2.375(e^{1/2} - 1) = -0.541$. If $nb = 0.030$, the compressibility factor $Z_p = p/(nk_BT)$ is approximately 0.9838. At $T = 400$ K the ratio becomes 0.325 and the pressure exceeds ideal by about 0.98 percent for the same density. A precise comparison should also estimate $n^2 B_3$, especially when the objective is to infer model parameters from percent-level deviations.

Sweep well range and temperature through the pressure cancellation

An interaction can exchange heat even at the Boyle point

The canonical identity $U = \partial(\beta F)/\partial\beta$ at fixed volume and particle number gives the leading excess internal energy per particle as $-k_B T^2 n\, dB_2/dT$. For the square well,

$$\frac{U_{\mathrm{ex}}}{N} = -nb(\lambda^3 - 1)\epsilon e^{\epsilon/(k_B T)} + O(n^2).$$

This result is negative for an attractive well. Repulsive hard-sphere exclusion instead changes pressure but has no potential energy on any permitted configuration. These contrasting examples show that a pressure deviation alone does not determine an energy deviation without its temperature dependence. Different virial models may pass through the same pressure at one temperature while predicting very different heat capacities.

In the numerical well at 200 K the excess energy per particle equals $-0.1175\,\epsilon$, or approximately minus 1.62×10^{-22} J. Compare with the positive translational mean $3k_BT/2$; the correction is about minus four percent of that kinetic energy. A configuration-based simulation offers a direct independent check: count pairs in the well, assign minus ϵ per unordered pair, and average. Counting both neighbours of every pair without halving would double the inferred mean energy.

Thermodynamic consistency supplies another useful cross-check. The identity $(\partial U/\partial V)_T = T(\partial p/\partial T)_V - p$ requires the pressure and energy virial calculations to share the same derivative of B_2. For a hard-sphere gas with temperature-independent core the two sides vanish through this order. For an attraction they do not vanish. If a fitted empirical repulsion diameter depends on temperature, differentiate that dependence as part of the fit and interpret the resulting effective energy with care.

What a pair histogram measures

Choose one particle as origin and count the expected number of other centres in a spherical shell $4\pi r^2 dr$. Define $g(r)$ by setting this count equal to $n g(r) 4\pi r^2 dr$. In the dilute limit a single neighbour at separation r sees the relative Boltzmann weight $e^{-\beta u(r)}$, so $g(r) \to e^{-\beta u(r)}$. A hard core produces an empty inner region, an attractive well a finite elevated shell and large separations approach one. Histograms per unit radius without division by the shell volume grow as r^2 even for an ideal gas.

The leading configurational energy follows by integrating the pair energy weighted by the expected number of neighbours and dividing by two for pair sharing. Similarly, expanding a system uniformly by changing scaled coordinates in the configurational integral gives the virial pressure for a differentiable finite-range pair potential:

$$\frac{U_{\mathrm{ex}}}{N} = \frac{n}{2} \int u(r) g(r)\, d^3 r,$$

$$p = n k_B T - \frac{2\pi n^2}{3} \int_0^\infty r^3 u'(r) g(r)\, dr.$$

The pressure formula includes the radial force, whereas the energy integral includes the potential. Taking the dilute Boltzmann form and integrating by parts returns the second-virial result. For a hard discontinuity ordinary pointwise differentiation of $u(r)$ is inappropriate. Treat it through contact jump conditions or a limit of a smooth steep interaction. A code that samples a hard-core derivative as zero everywhere it is finite will miss exactly the repulsive momentum transfer it seeks to measure.

An optical or simulated correlation measurement with a finite point spread blurs the exclusion edge and can populate apparent distances below the true diameter. Calibrate spatial resolution and two-particle tracking efficiency, particularly where particle images overlap. In a periodic cube, shells whose radius approaches half the box edge need an appropriate geometric normalisation. Removing the outer shell deficit by an arbitrary normalisation may preserve a pretty asymptote while biasing an integral used for compressibility.

Connect density fluctuations to the integral of the pair structure

For a large observation window in an open homogeneous fluid, total particle-number variance can be calculated in two ways. The diagonal single-particle terms give the mean number. Distinct pairs add the integral of the connected density correlation. Provided correlations decay rapidly enough compared with the window size, one obtains a zero-wavevector structure factor

$$S(0) = \frac{\operatorname{var} N}{\langle N \rangle} = 1 + n \int [g(r) - 1]\, d^3 r = n k_B T \kappa_T.$$

To leading order $g - 1 = f$, so $S(0) = 1 - 2 n B_2$. The compressibility derived by differentiating $p = n k_B T (1 + n B_2)$ is consistent to this order. The factor two relative to the pressure correction arises because pressure contains $n^2 B_2$ before differentiation. A mean pressure near the ideal value therefore need not imply that the full spatial correlation function is nearly flat.

One must specify the counting window. The exactly fixed total number in a sealed simulation box gives zero variance of that total. Measuring number fluctuations in subvolumes or the small-nonzero-wavenumber structure factor can estimate the thermodynamic value, with finite-size corrections. If the correlation length approaches the box size near a transition, extrapolation to zero wavevector using only one or two available modes is poorly conditioned.

An exactly soluble interacting gas as a test of truncation

A one-dimensional gas of hard rods provides a rare opportunity to compare an entire exact equation of state with its virial truncation. Take N rods of length a in an interval of length L with hard end walls. Order them by their left endpoints $x_1 < \cdots < x_N$. The no-overlap constraints become $x_{j+1} \geq x_j + a$, $0 \leq x_1$ and $x_N + a \leq L$. Introduce free coordinates $y_j = x_j - (j-1)a$. They obey $0 \leq y_1 \leq \cdots \leq y_N \leq L - Na$. Their configuration volume is consequently $(L - Na)^N / N!$. We have already chosen one ordering sector for indistinguishable rods, so another division by $N!$ would double-count the symmetry correction.

The momentum integral contributes one thermal wavelength per rod. Differentiating log partition function with respect to length gives the one-dimensional pressure, which has units of force:

$$p = \frac{N k_B T}{L - Na},$$

$$\frac{p}{n k_B T} = \frac{1}{1 - na} = 1 + na + (na)^2 + \cdots,$$

$$n = N/L.$$

In this geometry the second coefficient is a and successive virial terms are successive powers of the occupied length fraction. At $na = 0.10$ the first two terms give 1.10 rather than the exact 1.111, a one-percent relative pressure error. At $na = 0.40$ they give 1.40 instead of 1.667, a sixteen-percent error. The pressure becomes singular as free length vanishes even though the potential energy of every permitted configuration is zero.

A measurement of the mean internal energy alone would therefore find the same kinetic value as a one-dimensional ideal gas at each temperature and miss the entire excluded-volume pressure. Conversely a pressure fit over a narrow high-density interval can return an effective second coefficient quite unlike its true low-density limit. The exact hard-rod solution makes these two warnings quantitative without relying on a simulation or on a vaguely defined packing parameter.

There is also a contact-force interpretation. An arbitrary ordering gap between neighbouring rods has a broad distribution even if the total free length is strictly fixed. In the large-system limit the probability density for a free gap is exponential with mean $(1 - na)/n$. The pressure fixes its inverse scale $p/(k_B T)$, a one-dimensional illustration of how contact probabilities encode force on a wall. For a finite chain the fixed sum of all gaps correlates them, and their exact marginal distribution has a finite upper bound. A direct histogram of gap sizes is thus an independent test of whether a pressure-inferred excluded length really represents the particle geometry.

Infer a well range and depth from a two-temperature experiment

Suppose the core diameter has been determined from a high-temperature data set, and low-density pressure intercepts determine $B_2/b = -0.541$ at 200 K and $B_2/b = 0.325$ at 400 K. In the square-well model define $A = \lambda^3 - 1$ and $x = e^{\epsilon/(400 k_B)}$. The two observations imply $A(x^2 - 1) = 1.541$ and $A(x - 1) = 0.675$. Divide to obtain

$$x + 1 = 1.541 / 0.675 = 2.283.$$

Therefore $\epsilon/k_B = 400 \log 1.283 \simeq 99.8$ K, $A \simeq 2.38$, and $\lambda \simeq 1.50$. The algebra retrieves the earlier microscopic model using only two temperatures and a separately known core size.

The ratio calculation also reveals an ill-conditioned regime. At temperatures both high relative to the well depth, x differs from one by a small amount. A modest relative uncertainty in the measured excess over the hard-core baseline can then dominate $x - 1$, and range and depth will move in opposite directions in a fit. Put the estimated model back into both original pressure values, then predict a third intermediate temperature without refitting. A correlated misestimate of the core diameter shifts both ratios and can survive the first two consistency tests.

Consider an alternative attraction with a smooth tail but the same two measured virial coefficients. Since each coefficient is a weighted integral over all separations, agreement at two temperatures imposes only two scalar constraints on an entire function of distance. Measuring the dilute $g(r)$ supplies local information and can distinguish a broad shallow well from a narrow deep one. A third highly precise pressure reading cannot guarantee a unique functional form in the absence of a justified parametrisation, though it can falsify a particular two-parameter guess.

When estimating the pressure slope from experiment, do not fit Z_p at a single density as if the ideal intercept were measured there. Convert density units consistently, include the uncertainty in n entering both axes, fit several low-density points and test sensitivity to a quadratic term. If the inferred low-density intercept differs from one, check absolute pressure and temperature calibration before interpreting an empirical nonzero intercept as an interaction. No finite pair potential with a conventional dilute classical limit should create a constant deviation from ideality at zero density.

Use reciprocal-space scattering to test density fluctuations

A weak coherent probe scattered elastically by identical particles measures a momentum-transfer dependent structure factor. For positions $\mathbf{r}_j$, define $\rho_{\mathbf{q}} = \sum_j e^{-i \mathbf{q}\cdot\mathbf{r}_j}$. For nonzero $\mathbf{q}$ in a uniform sample the dimensionless static factor is $S(q) = \langle |\rho_{\mathbf{q}}|^2 \rangle / N$, after subtracting coherent mean components where required. Expanding the absolute square gives N self terms and pair terms. Isotropy then connects the measured angular spectrum to the radial correlation via

$$S(q) = 1 + 4\pi n \int_0^\infty r^2 [g(r) - 1] \frac{\sin(qr)}{qr} \, dr.$$

At large q the oscillatory factor tends to cancel smooth pair contributions, approaching the self term for suitable fluids. At very small q the factor tends to one, giving the compressibility sum rule stated earlier.

An experiment rarely measures exactly $q = 0$ because the incident beam itself occupies that direction. Extrapolation from small but finite angles is required. A correlation extending farther in real space produces structure at smaller q, so insufficient low-angle coverage is equivalent to limited knowledge of long separations.

For hard spheres at leading density order, $g - 1 = -1$ inside the core and zero outside. Integrate the sine kernel over that sphere to obtain a pair suppression proportional to $[\sin(q\sigma) - q\sigma\cos(q\sigma)]/(q\sigma)^3$. Its limit is one third before multiplication by the geometric coefficient. At $q = 0$ the correction is $-4\pi n\sigma^3/3 = -2nB_2$, exactly the pressure derivative prediction. At finite angle the correction oscillates even though the pair potential contains no oscillatory part. Interpreting every peak in an angular spectrum as a separate preferred particle spacing would overinterpret a Fourier transform.

Actual measured intensity also contains a one-particle form factor, detector background, finite beam shape and instrumental angular response. Determine the form factor in an independently diluted reference, then divide only in bands where it is sufficiently far from zero. Corrected $S(q)$ should approach one at sufficiently large accessible q for a simple fluid model. A pressure-based low-q estimate and a independently estimated real-space core radius give two anchors on the scattering normalisation. Agreement across all three measurements is more discriminating than an arbitrary smooth normalisation of a single spectrum.

Separating exchange from interaction in a dilute virial measurement

An observed second virial correction is not always purely the configuration-space Mayer integral. Even an ideal quantum gas has exchange corrections when the thermal wavelength is appreciable. Expand the fugacity series of an ideal three-dimensional gas at small fugacity. With internal degeneracy one,

$$n\lambda_T^3 = z \pm z^2/2^{3/2} + \cdots,$$

$$p\lambda_T^3/(k_B T) = z \pm z^2/2^{5/2} + \cdots,$$

where the upper signs are for bosons. Eliminate z to second order to find

$$B_2^{\mathrm{exchange}} = \mp\lambda_T^3/2^{5/2}.$$

A Bose gas has a negative pressure correction and a Fermi gas positive at fixed density even without a material pair potential.

The relative importance scales as $n\lambda_T^3$, growing on cooling and for lighter particles. This can be distinguished from a hard-sphere core whose classical second coefficient is temperature independent. A real weakly interacting quantum gas can have both effects, and a full quantum virial coefficient may require scattering phase shifts or bound-state information. A classical configuration integral with an added exchange term is at best an explicitly justified limit, not a universally accurate combination.

Predict a Joule expansion from the measured virial curve

An independently calibrated expansion gives an energy-based test of the low-density pressure law. Suppose an insulated vessel of gas expands freely into a second evacuated chamber. No external work or heat is exchanged, so the total internal energy is constant. For a general single-component simple fluid write $dU = C_V dT + [T(\partial p/\partial T)_V - p]dV$. This identity follows from the Helmholtz mixed derivative and is testable even if detailed pair potentials are inaccessible. Insert the leading virial pressure at fixed particle number to connect temperature change directly with $B_2'(T)$.

The excess energy per particle in this approximation is $-k_B T^2 n B_2'(T)$. For weak nonideality use the ideal monatomic per-particle heat capacity $3k_B/2$ in the leading temperature correction and evaluate B_2' at the initial temperature. At fixed energy this gives

$$\Delta T \simeq (2/3) T^2 B_2'(T)(n_f - n_i).$$

An attractive well with $B_2' > 0$ cools on expansion: fewer attractive neighbours raise potential energy, reducing the kinetic contribution. Pure hard-core exclusion has $B_2' = 0$ and does not change internal energy at fixed temperature in this model.

In the earlier square-well example the initial excess energy at 200 K and $n_i b = 0.030$ is approximately minus $11.75 k_B$ kelvin per particle. Doubling the available volume at nearly unchanged temperature halves that excess magnitude. The approximate increase in configurational energy is $5.88 k_B$ kelvin per particle, implying a translational temperature drop near 3.92 K. A fuller calculation solves the constant-energy equation with the final temperature inside the well Boltzmann factor; the difference between that and the linear estimate is a measure of the size of the correction itself.

Distinguish this experiment from throttling through a steady porous restriction, where enthalpy rather than internal energy is approximately conserved under idealised insulated flow conditions. A real-gas Joule–Thomson coefficient depends also on thermal expansion and molar volume. Pressure and density units can look almost identical in the two calculations, but the device boundary and the work term differ. Logging the direction of mass transfer and whether the total confined particle number remains fixed prevents a common misidentification.

Repeat the free expansion at several initial densities while controlling chamber heat capacity. A temperature change proportional to the density difference supports the leading pair correction; a constant offset suggests a thermometer or wall heat effect. An accidentally zero pressure deviation at the Boyle temperature need not make the expansion temperature change vanish, because the temperature derivative of the virial coefficient remains finite there. Jointly observing near-ideal pressure and measurable cooling is exactly the distinction between a function crossing zero and its derivative crossing zero.

A useful uncertainty estimate comes directly from the linear expansion thermometer. In the illustrated twofold expansion the roughly four-kelvin temperature shift scales with the initial density and with the temperature derivative of the virial coefficient. If both wall thermometers share a common zero offset, an initial-minus-final temperature difference largely cancels it, but a drifting temperature scale can remain. Use a control expansion of a gas well represented as ideal under the same laboratory conditions; an apparent cooling there bounds residual wall and valve effects. Independently reversing the gas transfer

between two matched bulbs changes wall-specific artefacts while retaining the expected sign for true free expansion. Quantitative agreement should involve the gas energy change plus heat transferred to thermometer mounts and vessel walls, not merely two indicated peak temperatures.

Recover a virial coefficient from a pair interaction

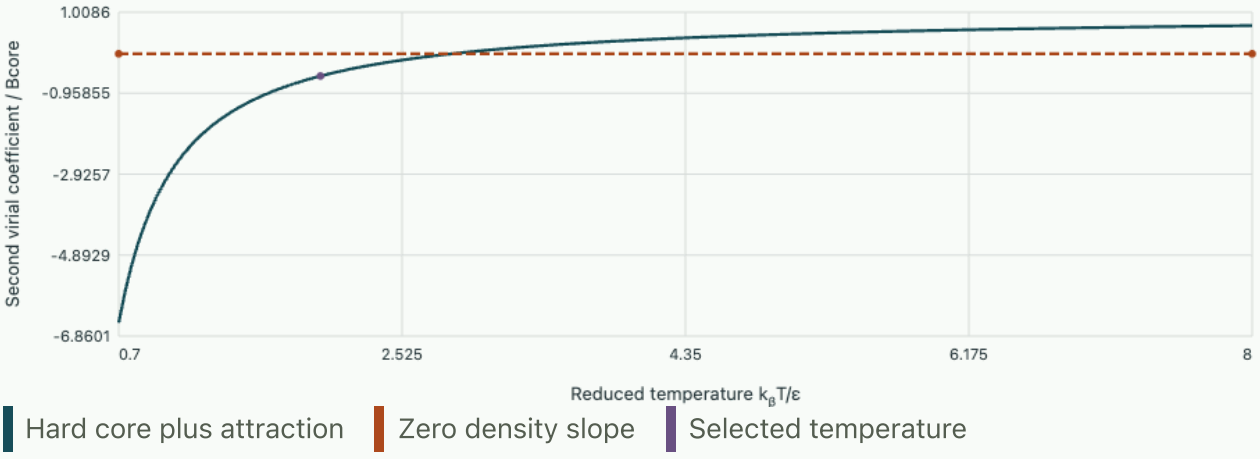
Vary the range and temperature of an attractive interaction. Identify where its attractive and excluded-volume contributions cancel.

Open this experiment online

Try the calculation

Outer radius / core diameter **1.5**

Reduced temperature $k_\beta T/\varepsilon$ **2**



B_2/B_{core} = -0.5407. The Boyle temperature is 2.846 ε/k_β.

Tutorial · Determine an effective diameter from high-temperature pressure

A gas at high enough temperature for its attraction to be negligible is dilute. A fit of $Z_p - 1$ versus particle density gives a slope 1.00×10^{-28} cubic metres. Infer the hard-sphere diameter. A student instead interprets that slope as the volume of one sphere. Quantify the error.

▼ **Apply the pair geometry and inspect the fit range**

Set the slope equal to $2\pi\sigma^3/3$, giving $\sigma = 3.63 \times 10^{-10}$ m. Using the physical sphere volume $\pi\sigma^3/6$ instead makes the diameter too large by $4^{1/3} = 1.587$, giving 0.576 nm rather than 0.363 nm. Check the fit extrapolated toward zero density and examine curvature as the maximum included density is reduced.

A small slope uncertainty based on a linear regression does not include a systematic neglected third virial term.

Tutorial · Use energy to break an attractive-potential degeneracy

At one temperature two square wells with the same core yield the same negative B_2, one deep and narrow, one shallow and broad. Show how an independent calorimetric energy measurement can distinguish them and specify a limiting regime in which the additional measurement becomes weak.

▼ **Identify two different combinations of range and depth**

At fixed temperature the pressure supplies $(\lambda^3-1)(e^x-1)$, with $x=\beta\epsilon$. The excess energy divided by $-nbk_BT$ instead supplies $(\lambda^3-1)xe^x$. Their ratio is $x/(1-e^{-x})$, a monotone function of positive x. With a known core and sufficiently accurate data it determines x, then the original pressure combination determines range. At $x\ll 1$ the ratio approaches $1+x/2$; its excess above one becomes small compared with calibration noise. Several temperatures can then be more informative than one highly uncertain difference.

Tutorial · Design an independent histogram check of pressure and energy

A dilute well fluid has measured $g=0$ in the core, $g=1.65$ across its attractive shell and $g=1$ outside. Take $\lambda=1.5$, $nb=0.03$ and $T=200$ K. Infer approximate well depth and mean neighbours in the shell. Compare the resulting potential energy with the thermodynamic prediction.

▼ **Integrate the actual shell volume and count unordered pairs**

The dilute relation gives

$$\epsilon/k_B = T\log 1.65 = 100.2\,\mathrm{K}.$$

The shell volume is $2b(\lambda^3-1)$, giving mean shell neighbours per particle $0.03(2)(2.375)(1.65)=0.235$. Dividing by two gives mean potential energy -0.1176ϵ per particle, matching the derivative calculation. Using an unnormalised radial pair histogram or counting the same bond twice would spoil this agreement.

Now repeat at higher density. A measured effective potential $-k_BT\log g$ is then a potential of mean force containing the effect of other neighbours. Inserting it into the original low-density two-body virial integral as if it were the bare pair potential double-counts many-body correlations. The direct energy integral using a known actual pair potential and the measured full g remains appropriate for a pairwise-additive fluid.

Cambridge Statistical Physics · interacting classical gases and cluster expansion

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 $\theta_r = \hbar^2/(2 I k_B)$. 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,

$$dS = C_V(T)dT/T + Nk_B\,dV/V.$$

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 $C_V = 2.5Nk_B$ 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 $0.00050k_B$, while at 1200 K it is about $0.309k_B$. 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

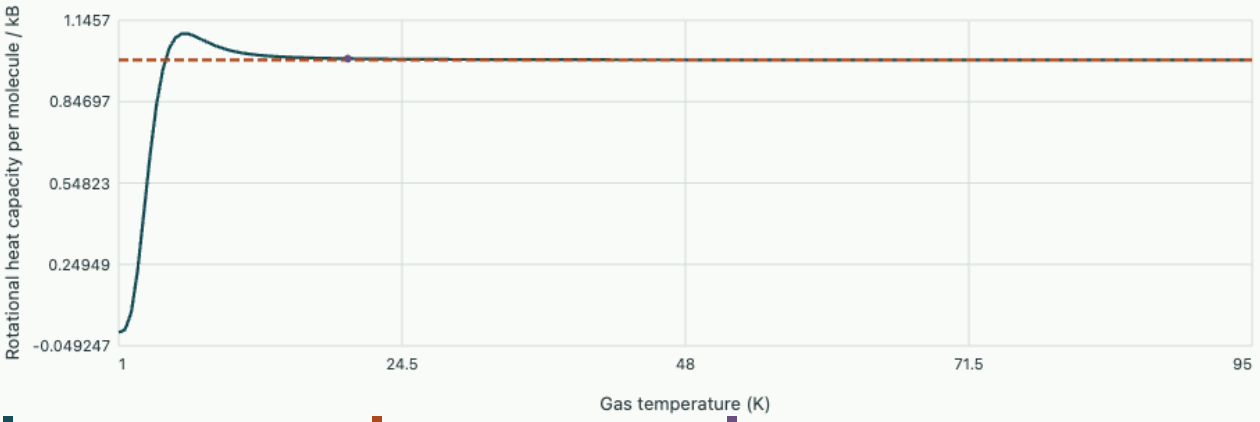
Sum the discrete levels of a linear rotor and compare the heat capacity with equipartition as temperature rises.

Open this experiment online

Try the calculation

Rotational temperature θrot (K) **8**

Chosen gas temperature (K) **20**



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 $x^2 e^{-x}/(1-e^{-x})^2 = 0.41$ 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

Phase nucleation, metastability and waiting-time statistics

Statistical mechanics from microscopic counting to material response

A cooled liquid can remain liquid below its equilibrium freezing temperature and then transform abruptly. How can a bulk phase be thermodynamically favoured while a small embryo of it tends to shrink? Balance interfacial cost against bulk gain, convert a barrier into a probabilistic waiting-time prediction and test which measurements can distinguish surface nucleation from a bulk mechanism.

Choose a driving free-energy density at fixed controls

At specified temperature and pressure, phases in equilibrium have equal molar Gibbs free energies. Below a melting temperature T_m the stable solid has lower chemical potential than the liquid. For modest undercooling, use $d\mu = -s\,dT + v\,dp$. At fixed pressure the chemical-potential difference per molecule is approximately $\mu_l - \mu_s \simeq \Delta h_f (T_m - T)/T_m$ if latent enthalpy of fusion per molecule varies little. Multiplying by the number density of the new phase gives a positive driving free-energy density Δg. State whether each quoted latent heat is per particle, per mole or per unit volume; converting it twice gives enormous radius errors.

The underlying free-energy landscape has stable and metastable minima separated by a barrier. In a phase-field picture a local bulk density penalises intermediate order within the metastable region, and a gradient penalty charges a finite energy for an interface. A finite amount of stable phase can therefore cost free energy before its bulk advantage wins. A coexistence construction determines which bulk phases minimise free energy; it does not provide the rate at which an existing metastable preparation can reach them.

Compare a bulk coexistence construction with the cost of an interface

Derive the critical radius and barrier

In the simplest capillarity model an embryo is a sphere of radius r, with isotropic positive interface energy γ per area and a bulk gain Δg per volume. Treat its interface as thin and its bulk properties as those of the macroscopic phases. Neglect elastic strain, long-range fields and depletion of the surrounding reservoir, then the reversible work to create it is

$$\Delta G(r) = 4\pi r^2 \gamma - \frac{4\pi}{3} r^3 \Delta g,$$

$$r_* = \frac{2\gamma}{\Delta g},$$

$$\Delta G_* = \frac{16\pi\gamma^3}{3\Delta g^2}.$$

The derivative vanishes at r_* and the second derivative there is $-8\pi\gamma$, so this is a maximum along the size coordinate. A slightly smaller embryo tends to shrink and a slightly larger one tends to grow, subject to noise and kinetic constraints. The radius where the total formation free energy first drops below its initial zero is

$$3\gamma/\Delta g = 1.5 r_*.$$

A supercritical nucleus can grow downhill even while its cumulative creation free energy is still positive. Confusing these two radii yields a fifty percent size error.

Take illustrative values $\gamma = 0.010$ J per square metre and $\Delta g = 1.0 \times 10^{7}$ J per cubic metre. Then $r_* = 2.0$ nm and $\Delta G_* = 1.68 \times 10^{-19}$ J, about 40.5 thermal energies at 300 K. Doubling driving density halves critical radius and quarters the barrier if γ is unchanged. At a two-nanometre radius the embryo may be just a few molecular spacings across. This calculated barrier is a sensitivity estimate; the assumed bulk interface tension, sharp interface and spherical shape become increasingly questionable there.

An alternative derivation uses the pressure jump $2\gamma/r$ across a curved interface together with a bulk chemical-potential difference. Near coexistence the condition of zero growth recovers the same critical curvature. In a vapour close to ideal behaviour the driving chemical-potential difference involves $k_B T \log S$ with supersaturation $S = p/p_{\mathrm{sat}}$, not simply $S - 1$ at arbitrary supersaturation. Near $S = 1$ the two are close. A quantitative cloud-chamber inference must also account for the evolving supersaturation and evaporation or growth after nucleation.

Background · Why does a faster temperature ramp move the apparent freezing point?

Assume independent homogeneous nucleation with rate density $J(T)$ in volume V, and prompt detection of the first nucleus. The chance that the untransformed sample survives the next interval dt is $1 - V J(T)\, dt$. Multiply these conditional probabilities along a cooling ramp with constant positive speed $v = -dT/dt$. Starting at T_0 gives

$$S(T) = \exp\!\left[-\frac{V}{v}\int_{T}^{T_0} J(T')\, dT'\right].$$

At fixed undercooling the faster ramp has accumulated less nucleation probability. The median observed freezing temperature satisfies an integrated hazard of $\ln 2$. Doubling the sample volume and the ramp speed together leaves the predicted distribution unchanged. Failure of that scaling can signal surface-dominated nucleation, thermal lag or a growth time long enough to affect detection.

Derive the barrier reduction for a spherical cap

A nucleus forming on a flat inert wall can replace one existing interface with another. Let it be a spherical cap of curvature radius r and contact angle θ through the new phase. The liquid-facing area is $2\pi r^2(1 - \cos\theta)$, contact area $\pi r^2 \sin^2\theta$ and cap volume $\pi r^3(2 - 3\cos\theta + \cos^3\theta)/3$. Young's contact-angle condition relates the wall energy difference to minus $\gamma\cos\theta$ per replaced contact area. Combining all terms gives the homogeneous spherical result multiplied by

$$f(\theta) = \frac{(2+\cos\theta)(1-\cos\theta)^2}{4}.$$

Both surface and bulk coefficients receive the same factor. Critical curvature radius therefore remains $2\gamma/\Delta g$ in this simple model, while the barrier becomes $f(\theta)\Delta G_*$. At 180°, $f = 1$ as the full sphere barely touches the wall. At 90°, $f = 1/2$ for a hemisphere. At 60°, $f = 0.15625$, lowering the illustrative 40.5-thermal-energy barrier to 6.33 thermal energies. Even a relatively small number of active wall sites can dominate observed events.

Complete wetting, a chemically patterned surface, a very small defect or appreciable energy per contact-line length can invalidate the simple contact geometry. Contact angle measured on a macroscopic equilibrium droplet need not equal the local microscopic value for a tiny forming embryo. A fitted effective angle that varies systematically with undercooling can signal a missing mechanism rather than a temperature-dependent property of a clean ideal wall.

Distinguish an attempt rate from an equilibrium barrier

A classical rate expression is $J = J_0 e^{-\Delta G_*/(k_B T)}$ per unit volume per unit time for homogeneous events. The prefactor contains the supply of available configurations and the kinetics for traversing the critical region. A barrier alone is insufficient to predict a waiting time. Lowering temperature can raise thermodynamic driving yet also reduce mobility and attachment, giving a nonmonotone total rate as the two effects compete.

If a well-mixed sample of volume V has constant independent hazard JV, the survival probability of the metastable state follows from $P_0(t+dt) = P_0(t)(1 - JV dt)$: $P_0 = e^{-JVt}$. Mean wait is $1/(JV)$, median $\log 2/(JV)$. A 1.0 microlitre sample with $J = 2.0 \times 10^8$ per cubic metre per second has hazard 0.20 per second and median 3.47 seconds. After five seconds 63.2 percent have transformed at least once, if the detector registers nucleation immediately.

If surfaces contribute independent rate per area J_s, the combined first-event hazard is $J_v V + J_s A$. Geometrically similar droplets have $A \propto V^{2/3}$, so changing sample size can distinguish these contributions. Vary active surface chemistry at fixed geometry as a second probe. If most samples contain a single rare defect not scaling smoothly with area, survival can be a mixture of exponentials rather than one exponential. A straight-line fit to mean waits alone obscures such heterogeneity.

A cooling ramp integrates many different hazards

Under a deterministic temperature history $T(t)$, survival is $P_0(t) = \exp[-\int_0^t J(T(s))V ds]$ for homogeneous, conditionally independent events. In a constant cooling ramp $q = -dT/dt > 0$, the exponent becomes $-(V/q)\int_T^{T_0} J(T')dT'$. The observed freezing-temperature histogram is weighted by survival as well as the instantaneous rate. The most probable freezing temperature need not minimise any free energy, and different ramp speeds can shift it even when all material properties are unchanged.

For a ramp calculation on gridded temperatures, update the remaining fraction by $e^{-\Delta H}$ where ΔH is the integrated hazard in that interval. Using a transition probability $JV\Delta t$ even when it exceeds one violates probability bounds. Subdivide the interval and check convergence of median and distribution width. At steep temperature dependence, a midpoint estimate over a coarse interval can move a reconstructed barrier enough to dominate the apparent contact-angle uncertainty.

The detected event may be a threshold in light scattering, latent heat or visual crystal size. A nucleus must grow before producing that signal. At fixed growth speed u_g, reaching detection radius r_d adds an approximate lag $(r_d - r_*)/u_g$; under cooling, both growth and hazard vary. Compare multiple detection thresholds or resolve the early growth trajectory and extrapolate a starting time. A shift in waiting time due entirely to the imaging threshold could otherwise be misinterpreted as altered microscopic nucleation.

Test how changing undercooling alters a barrier in the course's phase-equilibrium example

A finite reservoir can stabilise a droplet that would otherwise grow

The standard critical-radius derivation supposes that the driving free-energy density remains fixed while the nucleus grows. In a small sealed droplet or isolated simulation box, formation of the new phase can consume a significant fraction of the supersaturation. The surrounding phase then approaches coexistence and the bulk driving force falls. A single critical radius computed from initial supersaturation may describe early stability but cannot describe the entire late-time growth or the final phase fraction.

As a transparent model let conversion of new-phase volume v reduce the driving density linearly to $\Delta g(v) = \Delta g_0 - a v$, with $a > 0$ incorporating finite-reservoir compressibility or depletion. The bulk free-energy gain is the integral of the driving density, not its final value times volume. For a spherical interface $A(v) = (36\pi)^{1/3} v^{2/3}$, the total model is

$$\Delta G(v) = \gamma (36\pi)^{1/3} v^{2/3} - \Delta g_0 v + \frac{a}{2} v^2 .$$

The initial derivative is positive because a very small sphere has large area per volume. At large v the positive depletion derivative can make further growth unfavourable. If the derivative turns negative at intermediate volume there can be both a barrier top and a larger stable or metastable finite-droplet minimum. If it never turns negative the finite reservoir is too small or too weakly driven for a stable droplet in this model. This explains why a finite system can show a rare nucleus followed by arrested growth without contradicting the initial local barrier idea.

For a dimensionless illustration choose $\Delta G/(k_B T) = 4 v^{2/3} - 4 v + 0.25 v^2$, with a specified arbitrary reference volume absorbed into v. Its derivative is $(8/3) v^{-1/3} - 4 + 0.5 v$. At $v = 0.10$ it is about 1.80, at $v = 1$ about minus 0.83 and at $v = 10$ about 2.24. At least two stationary sizes therefore lie along the growth path. A sign test on neighbouring sizes distinguishes the early maximum from the later minimum without assuming every stationary root is a critical nucleus. Numerical root finding should bracket these branches separately instead of passing a single initial guess to a solver.

An experiment can test depletion by varying reservoir size at the same initial supersaturation and surface treatment. If early appearance times are comparable after correcting for trial volume but final crystal sizes grow with available material, finite feeding is plausible. Limited thermal transport, impurities blocking the interface or elastic stress can also arrest growth, so monitor the remaining parent-phase composition and temperature. A small system is precisely where the thermodynamic ensemble and the available amount of material need explicit definition.

A detection time can be dominated by diffusion after nucleation

After crossing the barrier a supercritical spherical domain may grow by diffusion of a dilute species. In a quasi-steady infinite exterior with surface concentration c_s and far concentration $c_\infty > c_s$, the concentration solves the radial Laplace equation. Its solution is

$$c(r) = c_\infty - (c_\infty - c_s) R / r.$$

The total inward solute flux is $4\pi D R (c_\infty - c_s)$. Equate this to solute incorporated into an advancing sphere, $4\pi R^2 (c_{\mathrm{new}} - c_s)\dot{R}$, to obtain

$$R^2(t) - R^2(0) = 2D \frac{c_\infty - c_s}{c_{\mathrm{new}} - c_s} t.$$

This square-root growth law assumes quasi-steady diffusion, a spherical shape, nearly constant far conditions and a surface near local equilibrium. If interface attachment instead limits growth, a roughly constant $\dot{R}$ may be a better first model over a short range. Measuring time to several optical thresholds distinguishes R_d^2 scaling from R_d scaling. Threshold variation therefore separates growth and nucleation more effectively than one image definition held fixed.

Take $D = 1.0 \times 10^{-9}$ square metres per second and concentration ratio 0.020. From a negligibly small initial radius, reaching 5.0 micrometres takes $25 \times 10^{-12} / (4.0 \times 10^{-11}) = 0.625$ s; reaching ten micrometres takes 2.50 s. If true median nucleation time is 1.0 s, the detected median at the larger threshold can be 3.5 s even when the microscopic barrier and initial supersaturation are unchanged. Calling its inverse a nucleation rate would seriously underestimate the true first-event hazard.

At very small radius a Gibbs–Thomson curvature shift increases the equilibrium surface concentration. The same curvature effect that opposes a small embryo can therefore slow its initial excess growth even after it has crossed the most probable barrier. At late times the bath may not be infinite and growing domains compete for material. These effects change the simple square-root law in different regions. Image a radius trajectory for each event instead of assigning the same unverified growth lag to all samples.

A thermal ramp with a tractable hazard shape

Over a restricted undercooling interval an experimentally measured hazard can rise approximately exponentially with undercooling $\Delta T = T_0 - T$: take $h(\Delta T) = h_0 e^{a \Delta T}$. This is a local fit to the full

barrier and prefactor, not a universal replacement for their pronounced temperature dependence. At constant cooling rate $q > 0$, direct integration gives cumulative hazard $H = h_0(e^{a\Delta T} - 1)/(aq)$ and survival e^{-H}. The median solves

$$\Delta T_{1/2} = \frac{1}{a}\log\left[1 + \frac{aq\log 2}{h_0}\right].$$

For $h_0 = 0.010$ per second, $a = 1.0$ per kelvin and $q = 0.10$ K/s the median undercooling is 2.07 K. Doubling cooling rate raises it to 2.70 K. The shift is smaller than one kelvin but large relative to many thermometer precisions. It arises even with exactly the same thermodynamic coexistence temperature and material interfaces. Counting a median freezing temperature at only one ramp rate can therefore mix kinetics with the equilibrium phase boundary.

The probability density per unit undercooling is $[h(\Delta T)/q]e^{-H}$, where the first factor measures the rate conditional on surviving to that temperature. Taking a log derivative shows that an interior density maximum occurs when $h = aq$, provided this exceeds the initial hazard. The median and the most probable temperature are distinct. At slower ramps the maximum can even lie at the beginning of the observation range, depending on preparation and chosen T_0. A histogram peak should be interpreted together with the surviving population, not treated as a direct plot of microscopic rate.

To reconstruct hazard from a measured ramp, estimate the fraction still untransformed at each temperature and inspect minus the log of that fraction. Differentiate smooth cumulative hazards cautiously; differentiation amplifies noise and small late survivor counts. A joint fit of raw first-event and censored records over several ramp rates generally preserves more information than differentiating a single binned histogram. Varying sample volume as well helps separate a distribution of bulk rates from one associated with surfaces.

A committor test examines whether radius is a good reaction coordinate

The capillarity free-energy maximum uses one coordinate, radius. A practical microscopic test begins from candidate configurations near the proposed critical size, randomises thermal momenta or bath noise consistently with the preparation and launches many short independent trajectories. For each starting configuration estimate the fraction reaching a specified stable new-phase basin before returning to the parent phase. This fraction is the forward committor. A good transition-surface coordinate should identify configurations with committor near one half under the specified dynamics.

If configurations at the same radius show committors ranging from almost zero to almost one, that radius alone poorly locates the dynamical transition state. Shape, local order, strain or neighbour concentration may supply missing coordinates. Binomial uncertainty from finite trials also matters: 25 independent launches near one half give a standard error around 0.10 for an individual configuration. Increasing launches distinguishes true variation between configurations from sampling noise.

The committor depends on physically chosen basins and dynamics. An artificially accelerated Monte Carlo kinetic rule may preserve equilibrium free energies while changing which paths reach the new phase first. Use the same dynamics when comparing committor tests with measured waiting times.

Equilibrium barrier heights, kinetic attempt rates and the adequacy of a proposed reaction coordinate are connected questions but require different data.

Distinguish first appearance from transformed volume fraction

A first-event detector stops caring about a sample once the earliest nucleus is registered. A calorimeter or diffraction measurement instead often follows the total fraction of material transformed. Even if nucleation is homogeneous at a constant rate and crystals grow at constant speed, these two signals have different time dependences. Treat an infinite initially untransformed sample, with nuclei placed randomly at rate J per volume and spherical domains growing radially at speed u. To ask whether a particular point remains untransformed at time t, count possible earlier nucleation events whose growth sphere would have reached it.

A nucleus born at time s reaches points within radius $u(t-s)$, corresponding to volume $4\pi u^3(t-s)^3/3$. Integrate the attempt rate over that backwards influence volume from zero to t. For an ideal Poisson space-time nucleation model with this growth rule the probability of no such event is $\exp[-\pi J u^3 t^4/3]$, giving transformed fraction $X = 1 - \exp[-\pi J u^3 t^4/3]$. The result uses the standard extended-volume construction with random nucleation and isotropic constant-rate growth. It is not the simple first-event law $1 - e^{-JVt}$ for a finite observation sample.

For sites all pre-existing at time zero with random density n_0, the same backwards-volume reasoning gives

$$X = 1 - \exp[-(4\pi/3) n_0 u^3 t^3].$$

A log plot of minus the log remaining fraction against log time therefore has ideal slope four for continual three-dimensional nucleation and three for instantaneous site saturation. In two dimensions the powers would both be reduced by one. A measured slope between these values is not automatically evidence for a fractional geometric dimension. Changing growth law, active-site exhaustion, temperature ramps or overlapping different mechanisms can alter the apparent slope.

If measured isothermal transformed fraction is one half at 100 s in the continuous model, the product Ju^3 equals $3\log 2/(\pi\, 100^4) = 6.62 \times 10^{-9}$ per second to the fourth. This identifies the product, not the nucleation rate and growth speed separately. Independent imaging of isolated domain radii versus time, or a set of first-event measurements in small known volumes, can fix the remaining scale. A fit to the entire same ideal fraction curve cannot identify two quantities that enter it only through their product.

Actual calorimetric power is approximately latent enthalpy times dX/dt, convolved with thermal response. Its maximum occurs when transformation rate peaks rather than when the first nucleus forms or exactly when half the volume has transformed. Compare integrated calorimetric energy with the final transformed fraction and deconvolve the calibrated instrument response before interpreting any schedule of peaks as a nucleation-time distribution.

The influence-volume argument also predicts a useful geometry check in a thin film. If domains reach both film surfaces early and then grow as expanding disks in the plane, a three-dimensional spherical

growth model will eventually overestimate newly transformed volume. The apparent exponent can cross from an initially three-dimensional value to a lower late-time value without any change of microscopic nucleation pathway. Compare films of several thicknesses, identify the time at which individual radius reaches the film thickness and then test the predicted crossover in transformed fraction. Altering film thickness usually changes wall-to-volume ratio as well, so site density and growth geometry can vary together. Independent first-event measurements and direct images of individual domains remain necessary if the objective is to separate those two effects.

Locate a critical cluster and its barrier

Vary surface energy and bulk driving force independently. Find the stationary radius before reading the height of the barrier.

Open this experiment online

Try the calculation

Interfacial energy (mJ/m²) 35

Bulk driving density (MJ/m³) 25



Critical radius 2.8 nm; barrier 7.174 eV. The free-energy curve crosses zero again at 4.2 nm.

Tutorial · Extract a rate from records with censored observations

Twelve equal 2.0 microlitre samples are held at constant temperature for at most 20 s. Eight show first events at times whose sum is 64 s; four remain untransformed at the end. Estimate the homogeneous rate and explain the bias from averaging only the observed event times.

▼ **Count events and sum the full time at risk**

Likelihood for independent exponential waits with right censoring is proportional to $h^{N_e} e^{-hT_{\text{exposure}}}$. The exposure is $64 + 4(20) = 144$ s, so $\hat{h} = 8/144 = 0.0556$ per second. Dividing by 2.0×10^{-9} cubic metres gives $J = 2.78 \times 10^{7}$ per cubic metre per second. The observed-only mean of 8 s would instead imply 0.125 per second and bias the rate upward by discarding samples known to survive a long time. With eight events, statistical uncertainty is substantial despite the precise arithmetic.

Tutorial · Distinguish area from volume nucleation

Two cubic samples have side lengths L and $2L$. Their measured first-event hazards are 0.060 and 0.360 per second. Assume homogeneous bulk and wall contributions $b = J_v L^3$, $s = J_s 6L^2$ in the smaller cube. Infer their relative importance and predict the hazard for a cube of side $3L$.

▼ **Use different geometric powers to separate the contributions**

The measurements give $b + s = 0.060$ and $8b + 4s = 0.360$. Subtracting four times the first equation gives $4b = 0.120$, hence $b = s = 0.030$ per second. For side $3L$, hazard is $27b + 9s = 1.08$ per second. This prediction requires the same temperature, wall state and sample history, and negligible detection lag. An unmodelled population of corner-active sites adds a term independent of L, so measurements at a third size test the two-term assumption rather than merely improving its fit.

Tutorial · Propagate interfacial uncertainty into a rate

Suppose the barrier is estimated at $35 k_B T$, with five-percent relative uncertainty in interfacial tension while driving force, temperature and prefactor are treated as known. Estimate its effect on log rate and explain why the rate uncertainty is strongly asymmetric when expressed in ordinary units.

▼ **Differentiate the barrier first and exponentiate last**

The barrier scales as γ^3, so its first-order relative uncertainty is 15 percent. The uncertainty in $\log J$ is therefore about $35(0.15) = 5.25$, corresponding to multiplicative factors of roughly $e^{5.25} = 191$. A symmetric error in interfacial tension need not produce a symmetric additive error in rate. Computing the exact cubed upper and lower tensions produces further asymmetry.

Repeat the inference with a driving-force uncertainty, whose logarithmic coefficient is minus two in the barrier. Tension and driving parameters estimated from common coexistence data may be correlated. A direct comparison of predicted log rates across several undercoolings is often more stable than comparing raw rates spanning many orders of magnitude.

Cambridge Statistical Physics · metastability, phases and transitions

Transport, heat sources and diagnostics

Separate sensible heating from transported or radiated energy. Test the temperature signal against a field loss measurement and a conservative heat discretisation.

1. Control-volume momentum and forces on fittings
2. Boundary layers and momentum thickness
3. Compressible nozzles and choking
4. Time-dependent diffusion
5. Interfaces, contact resistance and Robin data
6. What a periodic grid does to a decaying Fourier mode
7. Measuring magnetic core loss by a hysteresis loop and by calorimetry
8. Thermal radiance, étendue and detector calibration
9. Master equations, detailed balance and probability currents
10. From binary collisions to the Boltzmann equation
11. Transport coefficients from equilibrium fluctuations
12. Work fluctuations and equilibrium free-energy differences
13. Measure the heat capacity of a reservoir from the width of an energy histogram
14. Recover a free-energy difference from a distribution of work

Control-volume momentum and forces on fittings

Fluid mechanics

Use the velocity vector, gauge pressure and a consistent force boundary to find the support load on a pipe bend.

FROM FIRST PRINCIPLES

Relate a support reaction to momentum carried by a stream

Newton's second law applies to a fixed set of fluid particles: external force equals the time derivative of their momentum. In a pipe bend it is easier to observe a stationary control volume that particles continually enter and leave. The rate of change for the material particles consists of storage in that volume plus the net rate at which momentum crosses its boundary.

$$\sum \mathbf{F} = \frac{\partial}{\partial t}\int_V \rho \mathbf{u}\, dV + \oint_{\partial V} \rho \mathbf{u}(\mathbf{u}\cdot\mathbf{n})\, dA.$$

The dot product determines whether mass crosses inwards or outwards, while the other velocity factor retains the direction of the momentum carried. For steady uniform inlet and outlet sections, the flux integral reduces to

$$\mathbf{F}_{\text{wall}} + \mathbf{F}_{\text{pressure}} + \mathbf{W} = \dot{m}(\mathbf{V}_2 - \mathbf{V}_1).$$

Pressure pushes inward normal to each open section, so its contribution must be included before solving for the wall force. The force of the fluid on the bend is opposite to the wall's force on the fluid. Equal inlet and outlet speed can still give a substantial support load when their directions differ.

DERIVATION AND APPLICATIONS

Derive the irreversible head loss across a sudden expansion

A sudden enlargement slows a liquid and partly restores static pressure. Consider steady incompressible flow from area A_1 into area $A_2 > A_1$ in a horizontal pipe. Choose the downstream section after mixing has produced an approximately uniform velocity again. Approximate both station profiles as uniform, neglect wall shear over the short control volume and take the mean pressure on the annular shoulder to be the upstream pressure p_1. Mass conservation makes the two mean speeds V_1, V_2 proportional to inverse area.

$$Q = A_1 V_1 = A_2 V_2,$$
$$p_1 A_2 - p_2 A_2 = \rho Q (V_2 - V_1).$$

The inlet and shoulder together contribute $p_1 A_2$ to the axial pressure force. Rearranging the momentum balance gives a positive downstream pressure rise, despite the separated mixing region.

$$p_2 - p_1 = \rho V_2 (V_1 - V_2).$$

Insert that pressure rise into the mechanical energy balance. The difference between kinetic energy released and pressure recovered is the specific energy dissipated in mixing, expressed below as a head of flowing liquid.

$$h_L = \frac{V_1^2 - V_2^2}{2g} - \frac{p_2 - p_1}{\rho g} = \frac{(V_1 - V_2)^2}{2g}.$$

This result tends to zero quadratically as the area change vanishes. A gradual diffuser would require a different shoulder-pressure and separation model. Pressure recovery can thus coexist with positive mechanical-energy loss and its measurement alone does not provide the loss.

WORKED EXAMPLE

Audit an abrupt pipe enlargement

Water moving at 3.00 m/s enters a pipe section with three times the area. Its final speed is 1.00 m/s. Taking $\rho = 1000$ kg/m³, the pressure rises by $1000(1.00)(2.00) = 2.00$ kPa. The kinetic pressure falls by 4.00 kPa; the missing 2.00 kPa is a head loss of 0.204 m. For 6.00 L/s, the rate of mechanical energy dissipation is $Q\rho g h_L = 12.0$ W.

Change the inputs and inspect this calculation below

SECOND APPLICATION

Recover the enlargement ratio from a head test

Another rig measures $V_1 = 2.50$ m/s and $h_L = 0.0800$ m. Taking the positive speed drop yields

$$V_2 = 2.50 - \sqrt{2(9.81)(0.0800)} = 1.25\,\mathrm{m/s}.$$

The inferred area ratio is $A_2/A_1 = V_1/V_2 = 2.00$ to the stated precision. The expected water pressure rise is about 1.56 kPa, a separately measurable check on the shoulder-pressure assumption.

Interactive calculation · Change the inputs of this example

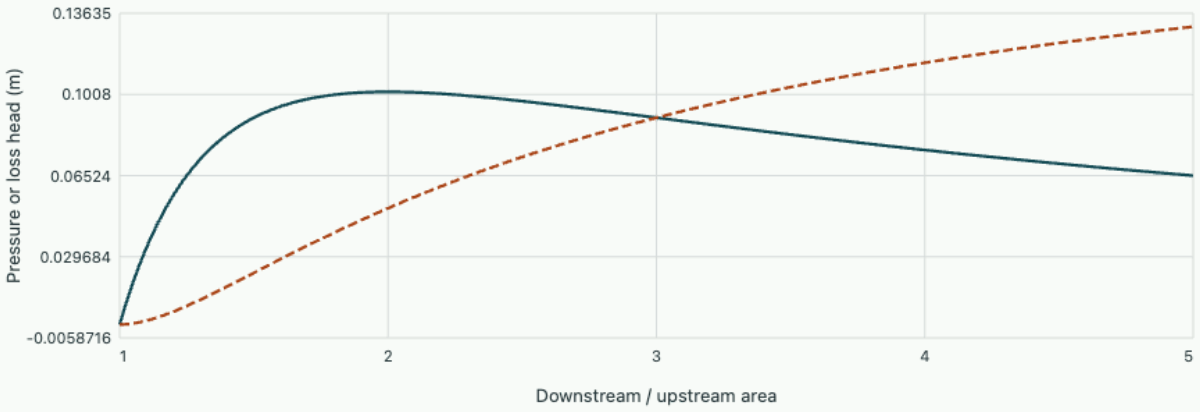
Use the adjustable quantities to test the two worked calculations above. The plot recomputes this derivation's balance or constitutive relation; its numerical readout reports the quantities needed for the forward calculation and the inverse or design question. Keep the fixed assumptions stated in the readout when comparing with the example values.

Open this live calculation

Try the calculation

Upstream speed (m/s) 2

Selected downstream/upstream area 2



At selected area ratio 2, downstream speed=1.000 m/s and loss=0.051 m. The pressure recovery is smaller than the ideal deceleration estimate.

Momentum crosses the boundary with the mass

A pipe bend can exert a substantial force on its supports even when its inlet and outlet speeds agree. Momentum has direction. Turning a stream requires force and gives the fitting an equal and opposite load. Enclose the fluid inside the fitting in a fixed control volume. The external forces include wall traction, pressure on the control surfaces and body forces such as gravity.

A surface element transports mass per second $\rho(\mathbf{u}\cdot\mathbf{n})dA$. Multiply by the local velocity to obtain momentum per second. Newton's second law applied to the control volume reads

$$\sum \mathbf{F}_{\text{ext}} = \frac{d}{dt}\int_{\text{CV}} \rho\mathbf{u}\, dV + \int_{\text{CS}} \rho\mathbf{u}(\mathbf{u}\cdot\mathbf{n})\, dA.$$

For a steady fitting with nearly uniform inlet and outlet velocities this reduces to a sum of outgoing $\dot{m}\mathbf{u}$ minus incoming $\dot{m}\mathbf{u}$. Use one fixed set of axes even as pipe directions change. Pressure always pushes into the control volume, whereas the flux carries the direction of the local velocity, including its sign on the chosen axis.

Include the end pressures explicitly

Take a horizontal bend turning water from the positive x direction to the positive y direction. The inlet and outlet areas are each 0.010 m² and $Q = 0.040\,\mathrm{m^3/s}$. Their speeds are both 4.0 m/s and $\dot{m} = 40$ kg/s. The net outward momentum flux vector is $(-160, +160)$ N. Let the gauge pressures be 90 kPa at inlet and 80 kPa at outlet. Pressure pushes +900 N in x through the inlet and −800 N in y through the outlet.

Let $\mathbf{W}$ be the fitting's wall force on the enclosed fluid, with external atmospheric effects accounted for consistently using gauge pressures. Then $\mathbf{W} + (900, -800) = (-160, 160)$ N, giving $\mathbf{W} = (-1060, 960)$ N. The fluid's net gauge load on the fitting is the opposite, $(1060, -960)$ N, of magnitude 1.43 kN. Supports plus connected pipework must provide the opposite force for a stationary fitting, in addition to carrying the weight of hardware and fluid. The strongest contribution here stems from the pressure terms, despite the conspicuous turn in velocity.

Recover momentum from a nonuniform profile

A flow meter usually measures Q, not $\int u^2 dA$. At a plane normal to a unidirectional profile define $\beta = \int u^2 dA / (A\bar{u}^2)$. Then the axial momentum rate is $\beta \rho Q \bar{u}$. Uniform flow gives $\beta = 1$. For a developed laminar circular pipe $u = 2\bar{u}[1 - (r/R)^2]$; inserting $dA = 2\pi r\, dr$ in the integral gives $\beta = 4/3$. The largest velocities carry a disproportionate share of momentum. If one bend cross-section admits strong swirl or backflow, resolve vector fluxes instead of applying a single profile coefficient.

Recognise an open-stream boundary

A free jet at atmospheric pressure strikes a fixed flat plate normal to its axis and turns into a thin sheet that leaves sideways. Neglecting rebound and friction far from the impact, the outgoing axial momentum flux is zero. For water flowing at 10 m/s through a 20 mm diameter nozzle, $\dot{m} = 3.14$ kg/s, so the axial force on the plate is 31.4 N. Both the incoming and outgoing free-stream control surfaces have zero gauge pressure. The plate force then follows entirely from the change of momentum. A nozzle bolted to a pressurised pipe would require different pressure forces even if it produced exactly the same free jet.

A jet catches a moving vane

Consider a single flat vane moving steadily in the jet direction at speed u. In the vane frame the arriving volume rate is $A(V - u)$ and the incoming streamwise relative velocity is $V - u$. If the jet spreads tangentially across the face, the resulting normal force is $\rho A (V - u)^2$. For water, $A = 2.0\,\mathrm{cm^2}$, $V = 10$

m/s and $u = 4$ m/s, this gives 7.2 N. Its mechanical power on the moving vane is $Fu = 28.8\,\mathrm{W}$. Using the stationary volume rate would count fluid parcels that the moving vane never catches.

Derive the two depths of a hydraulic jump

A hydraulic jump converts shallow fast open-channel flow into deeper slow flow over a short distance. Turbulent mixing dissipates mechanical energy, so applying an inviscid energy balance between the two uniform sections would miss its defining loss. Instead choose a control volume bounded by sections before and after the roller, the horizontal bed and the free surface. Approximate the velocity as uniform and the pressure distribution as hydrostatic at the two end sections. Atmospheric pressure cancels, and over the short reach assume the integral of bed friction is small compared with the pressure and momentum forces.

In a rectangular channel per unit width let the volume rate be q, depths y_1, y_2 and speeds $v_i = q/y_i$. The horizontal pressure resultant per width is $\rho g y_i^2/2$. Equating pressure force difference to momentum-flux change gives

$$\frac{y_1^2}{2} + \frac{q^2}{g y_1} = \frac{y_2^2}{2} + \frac{q^2}{g y_2}.$$

One solution is the unchanged depth. Factoring out $y_2 - y_1$ for the nontrivial solution gives

$$q^2 = g y_1 y_2 (y_1 + y_2)/2.$$

With upstream Froude number $Fr_1 = v_1/\sqrt{g y_1}$, the downstream-to-upstream depth ratio becomes $[\sqrt{1 + 8 Fr_1^2} - 1]/2$. The deeper downstream branch occurs when $Fr_1 > 1$. As $Fr_1 \to 1$, the ratio tends to unity and the finite jump disappears.

Calculate the energy dissipated in a channel transition

A horizontal test channel carries water at $y_1 = 0.20\,\mathrm{m}$ and $v_1 = 4.0\,\mathrm{m/s}$. Its unit-width discharge is $q = 0.80\,\mathrm{m^2/s}$ and Froude number is

$$4/\sqrt{9.81(0.20)} = 2.86.$$

The depth ratio is about 3.57, giving $y_2 = 0.714\,\mathrm{m}$ and $v_2 = 0.80/0.714 = 1.12\,\mathrm{m/s}$. The downstream Froude number is about 0.42. These are conjugate depths for the specified momentum balance; the tailwater has to support the downstream depth for a stationary jump to sit at the chosen location.

Using specific mechanical head $E = y + v^2/(2g)$, the inlet and outlet give 1.015 and 0.778 m respectively. The loss is approximately 0.237 m of water, or $1000(9.81)(0.80)(0.237) = 1.86\,\mathrm{kW}$ dissipated per metre of channel width. Substitution of the momentum relation into the energy difference gives

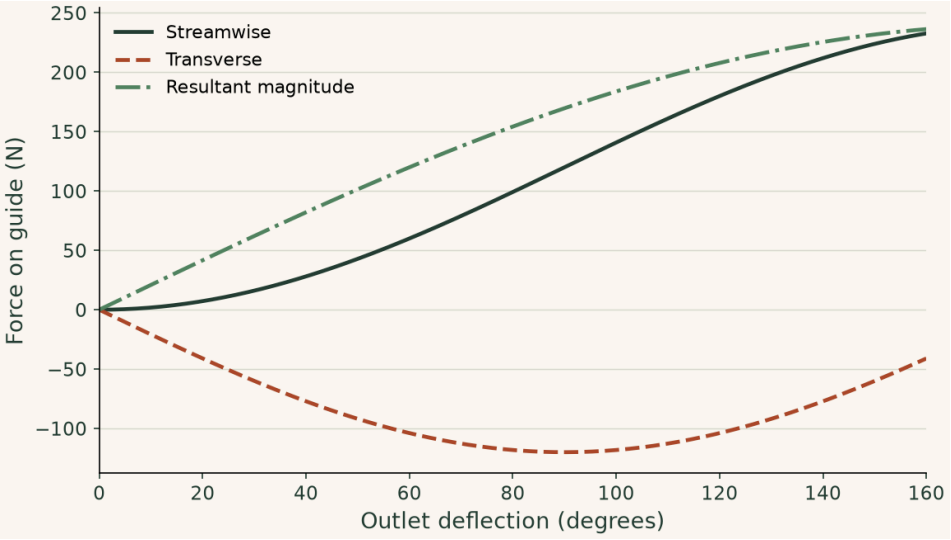
$$\Delta E = (y_2 - y_1)^3 / (4 y_1 y_2),$$

which is positive for the physical jump and approaches zero rapidly for a weak jump. This is a useful cross-check on a measured depth pair. If measurements are taken within the roller, pressure may be nonhydrostatic and velocity markedly nonuniform; using those sections in the simple end-state formula can produce an apparent energy or momentum imbalance.

Another discrimination comes from comparing the two depths with an upstream energy curve. At $q = 0.80$ m²/s, a downstream depth giving exactly the upstream energy would have to satisfy $y + 0.03262/y^2 = 1.015$. The deeper root is about 0.981 m, appreciably deeper than the momentum-predicted 0.714 m. Pairing those equal-energy depths would require an additional horizontal force or a different channel geometry to satisfy momentum. A smooth reversible change over a channel hump can conserve specific head while changing velocity and depth; a jump on a flat bed instead conserves the appropriate momentum function while dissipating head. Laboratory measurements of both levels are sufficient to distinguish the two idealisations.

Resolve the momentum force on a turning guide

A steady free water jet of mass flow 6 kg/s enters a smooth stationary guide at 20 m/s and leaves with the same speed at the indicated angle. Inlet and exit are at ambient pressure. Neglect weight for this short control volume. Reaction on the guide is the negative of the fluid momentum change.



A 90° turn gives +120 N streamwise and −120 N transverse, with resultant about 170 N. A 180° reversal would give 240 N streamwise. Using just the change in speed would predict zero for every turn.

▶ **Read values from the plot**

Interactive calculation · change the inputs of this example

Compare a right-angle turn with a complete reversal. The transverse force requires vector momentum, not just speed magnitude.

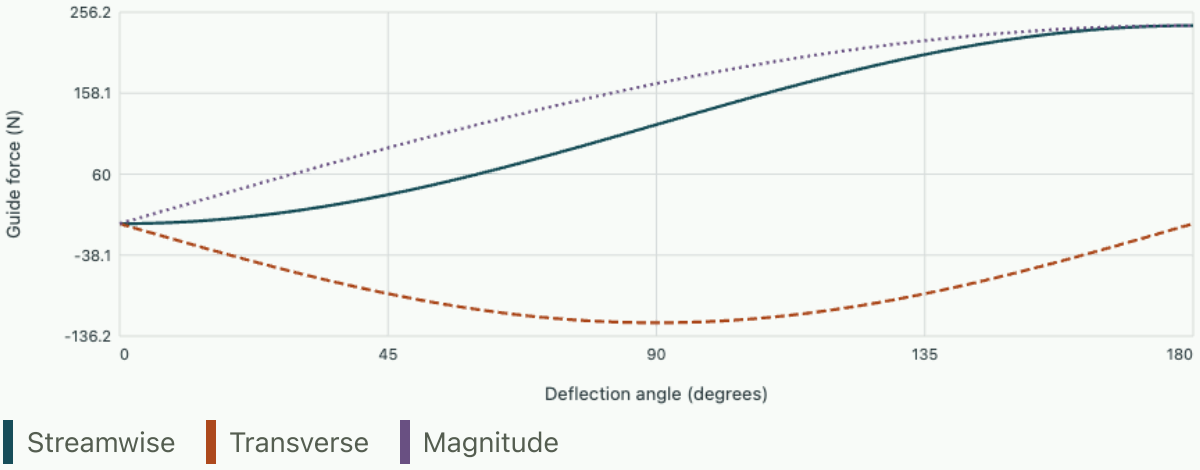
The curves below are recalculated from this section's equations in your browser. The fixed reference plot above remains available for comparison.

Open this live calculation

Try the calculation

Mass flow (kg/s) **6**

Jet speed (m/s) **20**



A 90° turn gives components 120.000 and -120.000 N. Both ports are at ambient pressure.

Boundary layers and momentum thickness

Fluid mechanics

Connect a measured velocity deficit to wall drag and distinguish the physical edge of a boundary layer from its integral thicknesses.

FROM FIRST PRINCIPLES

Obtain wall drag from the streamwise momentum deficit

A thin boundary layer slows fluid next to a solid wall. Consider steady two-dimensional flow over a smooth flat plate with constant external speed U and pressure. Use unit span and constant density ρ. Choose a control volume ending at location x, with its upper surface a streamline outside the layer. Mass conservation relates the uniform inlet flow to the downstream velocity profile. Multiplying the discharged mass in a strip by the streamwise speed it has lost gives that strip's contribution to the momentum deficit.

$$dD = \rho\, u(y)[U - u(y)]\, dy,$$

$$D(x) = \rho U^2 \theta(x),$$

$$\theta = \int_0^\infty \frac{u}{U}\left(1 - \frac{u}{U}\right) dy.$$

The integrated deficit equals the drag per unit span exerted by the wall on the flow in magnitude. Extending the plate by a small dx adds wall drag $\tau_w\, dx$, so differentiating yields

$$\tau_w = \rho U^2 \frac{d\theta}{dx},$$

$$C_f = \frac{\tau_w}{\rho U^2/2} = 2\frac{d\theta}{dx}.$$

These relations assume negligible pressure gradient. The thickness θ weights lost velocity by the mass flux actually present in that part of the profile.

DERIVATION AND APPLICATIONS

Relate displacement thickness to boundary-layer entrainment

Mass conservation requires a weak normal flow at the edge of a developing boundary layer. Consider steady, two-dimensional incompressible flow over an impermeable flat plate with constant external speed U. Let $u(x, y)$ be tangential velocity, $v(x, y)$ normal velocity and $y = 0$ the wall. Integrate continuity from the wall to a fixed measurement height H chosen well outside the layer throughout the region of interest. Since $v(x, 0) = 0$,

$$\partial_x u + \partial_y v = 0,$$

$$v(x, H) = -\frac{d}{dx}\int_0^H u(x, y)\, dy.$$

The integral would equal UH for a uniform velocity profile. The missing volume flux per unit span defines displacement thickness δ^* times U. For constant H, U, differentiating cancels the large uniform part and connects the missing flux directly to the weak external normal velocity.

$$U\delta^* = \int_0^H (U - u)\, dy,$$

$$v(x, H) = U\frac{d\delta^*}{dx}.$$

For an illustrative profile take $u/U = 2\eta - \eta^2$ up to $\eta = y/\delta = 1$ and $u = U$ above it. Its deficit is $(1 - \eta)^2$ and its integral is one third. When δ grows as the square root of distance from the leading edge, the normal velocity follows by differentiation.

$$\delta^* = \frac{\delta}{3},$$

$$\delta = C\sqrt{x} \implies v(x, H) = \frac{U\delta}{6x}.$$

The sign is positive because the low-speed layer occupies an increasing share of the fixed-height section. The value at fixed external height differs from the mass entering a control surface that itself follows the growing edge.

WORKED EXAMPLE

Estimate normal flow above a plate

At $x = 0.500$ m, a measured polynomial layer has $\delta = 6.00$ mm and $U = 4.00$ m/s. If it grows locally as $\sqrt{x}$, the predicted edge velocity is

$$4.00(0.00600)/(6 \times 0.500) = 0.00800\,\mathrm{m/s}.$$

Its ratio to the external tangential speed is only 0.00200. The displacement thickness is 2.00 mm, the height by which a uniform-speed stream would need to be displaced to produce the same missing flux.

Change the inputs and inspect this calculation below

SECOND APPLICATION

Use two profiles to infer the average vertical flux

Independent surveys at $x = 0.200$ and 0.600 m find displacement thicknesses 1.20 and 2.00 mm. At $U = 3.00$ m/s the interval-average normal velocity is

$$3.00(0.000800)/0.400 = 0.00600\,\mathrm{m/s}.$$

Over a plate span of 0.500 m this corresponds to outward volume flow $0.00600(0.400)(0.500) = 1.20$ L/s through the fixed upper face, equal to the growth of the streamwise flux deficit.

Interactive calculation · Change the inputs of this example

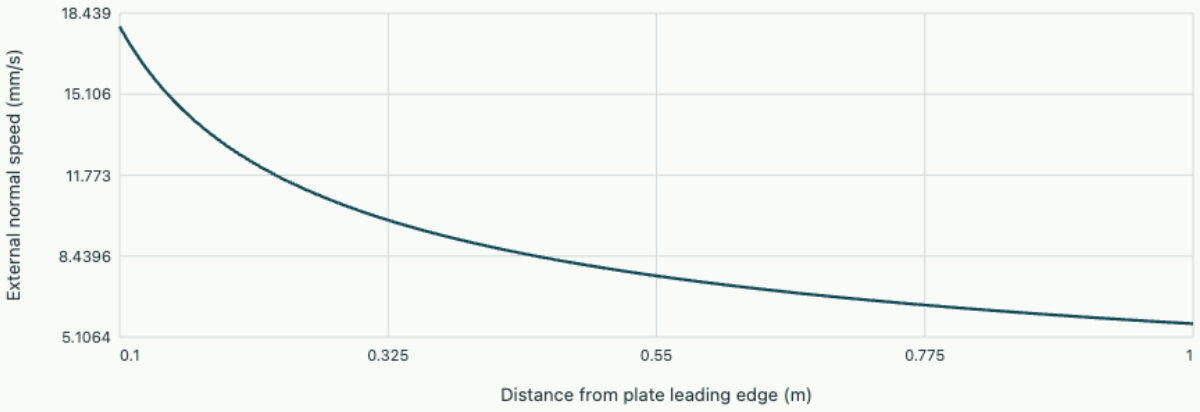
Use the adjustable quantities to test the two worked calculations above. The plot recomputes this derivation's balance or constitutive relation; its numerical readout reports the quantities needed for the forward calculation and the inverse or design question. Keep the fixed assumptions stated in the readout when comparing with the example values.

Open this live calculation

Try the calculation

External speed (m/s) 4

Layer thickness at0.5 m (mm) 6



Quadratic-profile displacement effect

At0.5 m, upward speed 8.000 mm/s. The displacement thickness is one third of the assumed boundary-layer thickness.

Viscosity matters where velocity changes rapidly

Flow past a smooth plate with a large length-based Reynolds number often has nearly uniform speed outside a thin layer next to the wall. No slip requires zero velocity at the stationary surface. Inside the thin region, transverse velocity gradients can be large enough for viscosity to balance streamwise inertia. If the free-stream speed is U, the streamwise distance is x and the layer thickness is δ, the leading inertial and viscous scales are U^2/x and $\nu U/\delta^2$. Their balance gives $\delta/x \sim 1/\sqrt{\mathrm{Re}_x}$, where $\mathrm{Re}_x = Ux/\nu$.

This scaling predicts how a laminar layer grows. Its numerical factor depends on how the thickness is defined. A common measured edge is the position reaching 99% of the external speed. Two integral measures instead quantify flow displaced by the layer and its lost momentum, and are less sensitive to picking a particular edge point.

$$\delta^* = \int_0^\infty (1 - u/U)\, dy,$$

$$\theta_m = \int_0^\infty (u/U)(1 - u/U)\, dy.$$

The thickness δ^* gives the equivalent missing volume flux per unit width divided by U. Momentum thickness θ_m weights that deficit by the actual local speed. Fluid right at the wall contributes strongly to displacement thickness but carries almost no streamwise momentum flux.

Relate the momentum deficit to wall shear

A steady control volume surrounding one face of a flat plate at zero incidence receives fluid from the uniform upstream stream. For constant external speed and pressure, the integrated wall force equals the deficit of streamwise momentum flux downstream after accounting for entrained fluid. Per unit plate width this gives

$$D'(x) = \rho U^2 \theta_m(x).$$

Differentiating gives local wall shear $\tau_w = \rho U^2\, d\theta_m/dx$. A growing deficit implies positive drag on the plate.

Suppose a velocity survey is represented by $u/U = 2\eta - \eta^2$ for $0 \le \eta = y/\delta \le 1$, and uniform flow outside. Integrating yields

$$\delta^*/\delta = \int_0^1 (1-\eta)^2 d\eta = 1/3.$$

Since $\int_0^1 (2\eta - \eta^2) d\eta = 2/3$ and $\int_0^1 (2\eta - \eta^2)^2 d\eta = 8/15$, the momentum thickness is $2\delta/15$.

Infer drag from a measured layer

At the end of a 0.60 m plate of width 0.40 m in air, let $U = 12\,\mathrm{m/s}$, $\rho = 1.20\,\mathrm{kg/m^3}$ and $\delta = 5.0\,\mathrm{mm}$ according to the fitted profile. Then $\theta_m = 0.667\,\mathrm{mm}$ and drag on the measured face is

$$1.2(12)^2(0.000667)(0.40) = 0.0461\,\mathrm{N}.$$

The mean shear over that face is

$$0.0461/(0.60 \times 0.40) = 0.192\,\mathrm{Pa}.$$

It need not equal the local shear at the measuring station. An adverse pressure gradient or separated flow changes the momentum balance; surface pressure forces and external acceleration must then be included before interpreting a wake or boundary-layer survey.

Predict a boundary-layer thickness before viewing a simulation

At the same measuring station the laminar scaling from the first section gives $\delta \propto \sqrt{\nu/U}$. Starting from a measured 35.84 mm layer at $x = 0.50$ m, predict its thickness after separately doubling free-stream speed and doubling kinematic viscosity. Which change should raise the dimensional wall stress if density and viscosity are held fixed while speed increases? Compare your thickness ratios with the two computed runs and distinguish the measured skin-friction coefficient from wall stress in pascals. Compare the flat-plate runs.

Reduce a partial differential equation to one similarity profile

For a laminar zero-pressure-gradient boundary layer along a flat plate, the leading equations are $u_x + v_y = 0$ and $uu_x + vu_y = \nu u_{yy}$. Assume external speed U constant, a stationary no-slip surface and a layer thin relative to streamwise distance. Introduce a streamfunction ψ through $u = \psi_y,\ v = -\psi_x$ to satisfy continuity identically. The scaling $\delta \sim \sqrt{\nu x/U}$ suggests $\eta = y\sqrt{U/(\nu x)}$ and $\psi = \sqrt{\nu U x}\, f(\eta)$. Differentiation gives $u = U f'$ and $v = \tfrac{1}{2}\sqrt{\nu U/x}(\eta f' - f)$.

When these forms enter the momentum equation, the terms containing $\eta f' f''$ from longitudinal and transverse advection cancel. Both the remaining inertial and viscous terms carry U^2/x. Dividing through gives the Blasius boundary-value problem

$$f''' + \tfrac{1}{2} f f'' = 0,$$

$$f(0) = f'(0) = 0,$$

$$f'(\infty) = 1.$$

The two wall conditions impose zero normal and tangential velocity. The outer condition matches the uniform stream. The unknown wall slope of the velocity profile, $f''(0)$, is determined by matching that distant condition; prescribing it independently would overdetermine the third-order equation. Numerical

integration gives approximately 0.3321 for this definition of η. Conventions that include a factor of two in the similarity coordinate assign a different numerical constant but the same physical wall stress.

Predict a force measurement from the similarity solution

The wall shear is $\mu U f''(0)\sqrt{U/(\nu x)}$, giving $c_f(x) = \tau_w/(\rho U^2/2) = 0.664/\sqrt{\mathrm{Re}_x}$. Its integral over a plate of length L yields mean coefficient $\bar{c}_f = 1.328/\sqrt{\mathrm{Re}_L}$. The mean equals twice the local trailing-edge value because the integrable $x^{-1/2}$ shear is highest near the leading edge. Using a single trailing-edge wall gauge for the whole area would therefore underestimate the force.

Take a smooth plate 0.50 m long and 0.25 m wide in air at $U = 6.0\,\mathrm{m/s}$, $\nu = 1.5 \times 10^{-5}\,\mathrm{m^2/s}$, $\rho = 1.20\,\mathrm{kg/m^3}$. The trailing-edge Reynolds number is 200,000, and a nominally laminar calculation gives

$$c_f(L) = 0.00149,\ \bar{c}_f = 0.00297.$$

The one-face drag is

$$0.5(1.20)(6)^2(0.50)(0.25)(0.00297) = 0.00802\,\mathrm{N}.$$

The trailing-edge wall shear is just 0.0321 Pa, and the 99% thickness estimate $5L/\sqrt{\mathrm{Re}_L}$ is 5.59 mm. A support carrying both plate faces would register roughly twice the one-face drag if both flows are comparable, plus fixture and edge drag. A single net force measurement cannot by itself tell whether a discrepancy arose from transition, unequal face conditions or the support geometry.

The similarity solution also predicts how to collapse profiles measured at different stations. At $x = 0.125$ m the same value of η corresponds to half the wall-normal distance it does at 0.50 m, since the distance scales with $\sqrt{x}$. Meanwhile wall shear doubles and local skin-friction coefficient doubles. A traverse resolving only whole millimetres may describe the outer velocity defect at both positions but badly estimate the near-wall slope at the earlier one. Testing both the profile collapse and the independent integral drag provides a stronger check of the assumed laminar flow than comparing a single velocity at the outer edge.

Measure the momentum transported by turbulent fluctuations

In a smooth plane channel driven by a constant pressure drop, time averaging can hide the dominant transport mechanism. Consider steady statistics and a fully developed mean flow between parallel walls. Let $y = 0$ denote the lower wall and h the distance to the centreline. Write $u = U(y) + u'$, $v = v'$. Impermeability and incompressibility give zero mean transverse velocity. The average of the instantaneous flux of streamwise momentum across a horizontal surface is nevertheless $\rho\overline{uv} = \rho\overline{u'v'}$. Averaging the product retains the correlation between the two fluctuations.

For constant density and viscosity the mean streamwise balance becomes

$$0 = -d\bar{p}/dx + d(\mu\, dU/dy - \rho\overline{u'v'})/dy.$$

Set $G = -d\bar{p}/dx > 0$. Integrating from y to the centreline, where symmetry sets the total shear to zero, gives

$$\tau_{\text{total}}(y) = \mu \frac{dU}{dy} - \rho\overline{u'v'} = G(h - y),$$

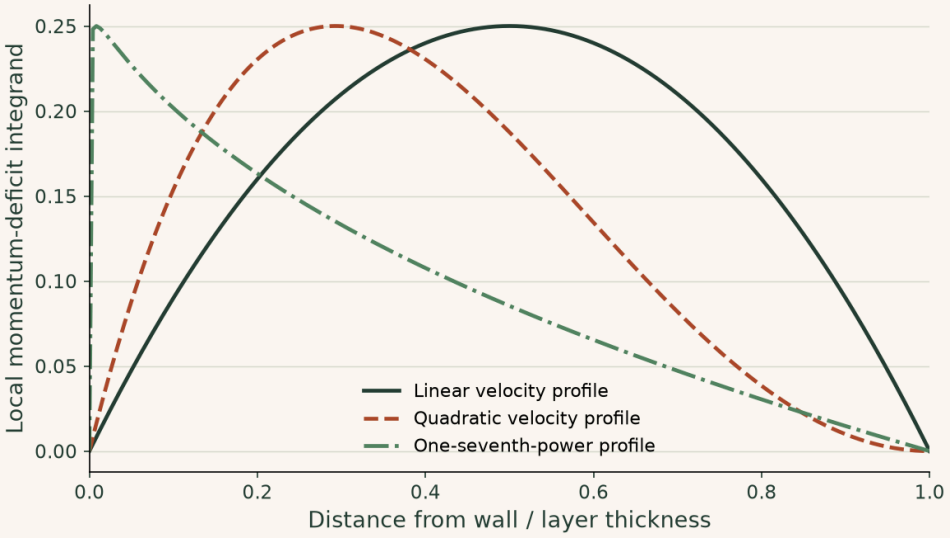
$$\tau_w = Gh.$$

The right side is fixed by the driving pressure gradient. At a no-slip wall both fluctuating velocities vanish, so molecular viscosity carries the full shear there. Farther out a parcel moving rapidly downstream can move towards the lower wall ($u' > 0,\ v' < 0$). A slow parcel leaving that wall has the same negative product. Both transfer streamwise momentum towards the wall. With the signs above the turbulent contribution to the shear is positive.

For water with $\rho = 1000$ kg/m³ and $\mu = 0.0010$ Pa s, take $h = 0.020$ m and $G = 100$ Pa/m. The wall stress is 2.00 Pa. At $y = 0.0050$ m the total shear is 1.50 Pa. A measured slope of 150 s⁻¹ accounts for 0.150 Pa. The remaining 1.35 Pa implies $\overline{u'v'} = -1.35 \times 10^{-3}$ m²/s². The exact force balance gives a quantitative consistency check for velocity measurements or a simulation. Predicting the split from G alone would require a relation for the velocity correlations.

Estimate a boundary layer's momentum deficit

Momentum thickness is $\theta = \int (u/U)(1 - u/U)\, dy$. Use three idealised trial profiles at the same outer thickness. Integrating the ordinate with respect to the plotted dimensionless height gives θ/δ.



The three areas are 1/6, 2/15 and 7/72. All profiles agree at both endpoints, yet they predict different momentum losses. The power-law profile is unsuitable for inferring wall shear from its limiting wall derivative.

▶ **Read values from the plot**

Interactive calculation · change the inputs of this example

Change the exponent of the trial velocity profile. Equal endpoint velocities need not imply equal momentum deficit.

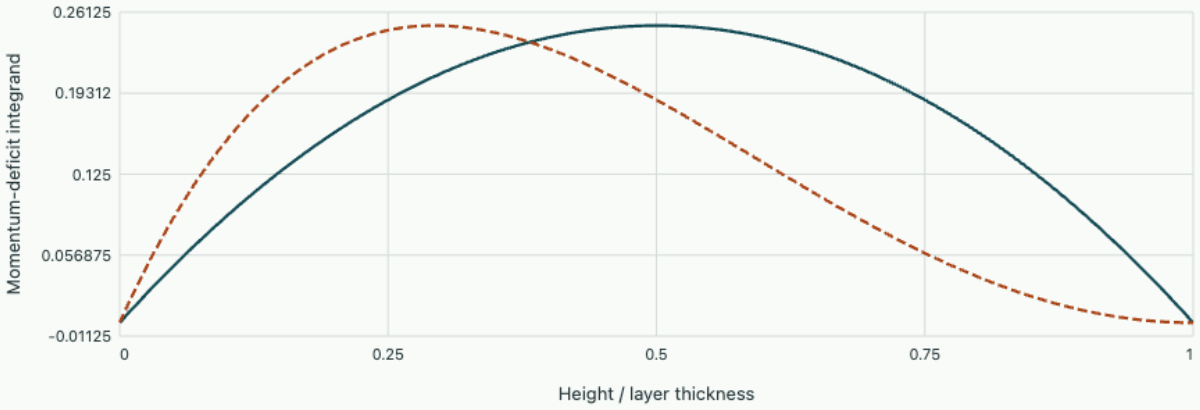
The curves below are recalculated from this section's equations in your browser. The fixed reference plot above remains available for comparison.

Open this live calculation

Try the calculation

Power-law exponent 1

Layer thickness (mm) 5



Selected power profile Quadratic reference

Selected momentum thickness is 0.833 mm. The quadratic reference gives 0.667 mm. For exponent below one the wall derivative diverges.

Compressible nozzles and choking

Fluid mechanics

Relate area change to Mach number and calculate the limiting mass flow through a converging nozzle.

FROM FIRST PRINCIPLES

Derive the area–velocity relation for an isentropic stream

Consider a steady, quasi-one-dimensional flow through a smooth duct of area A. Assume negligible viscosity, heat exchange and elevation change, and no shocks. Density ρ and speed V vary between stations. Mass conservation gives constant $\rho V A$, so taking logarithmic differentials relates three fractional changes.

$$\frac{d\rho}{\rho} + \frac{dV}{V} + \frac{dA}{A} = 0.$$

Streamwise momentum gives

$$dp + \rho V\, dV = 0.$$

For an isentropic change, sound speed satisfies $a^2 = (\partial p/\partial \rho)_s$ so the pressure differential can be eliminated in favour of density. With $M = V/a$ this yields

$$\frac{d\rho}{\rho} = -\frac{V\, dV}{a^2} = -M^2 \frac{dV}{V},$$

$$\frac{dA}{A} = (M^2 - 1)\frac{dV}{V}.$$

Subsonic flow accelerates in a contraction because its density variation is comparatively small. A supersonic stream accelerates in a diverging duct because density falls more rapidly. A smooth transition through $M = 1$ with nonzero velocity gradient requires a stationary area, usually the minimum area of a nozzle. Reservoir state and downstream pressure determine whether that sonic condition is actually reached.

DERIVATION AND APPLICATIONS

Recover subsonic Mach number from stagnation pressure

A forward-facing pressure probe brings a small sample of moving gas nearly to rest. Relating the resulting pressure to speed requires accounting for compressibility. Consider a calorically perfect ideal gas, constant heat-capacity ratio $\gamma > 1$, no heat input or shaft work, and subsonic flow brought reversibly to rest. Neglect the height change. Use $a^2 = \gamma R T$, Mach number $M = V/a$ and $c_p = \gamma R/(\gamma - 1)$. The steady energy balance gives

$$c_p T_0 = c_p T + \tfrac{1}{2} V^2,$$
$$\frac{T_0}{T} = 1 + \frac{\gamma - 1}{2} M^2.$$

For an isentropic ideal gas, p/ρ^γ is constant. Using the gas law to eliminate density yields a pressure ratio as a power of the temperature ratio. This provides a relation between a stagnation tap and a separate static tap at the same upstream state.

$$\frac{p_0}{p} = \left(\frac{T_0}{T}\right)^{\gamma/(\gamma-1)} = \left(1 + \frac{\gamma - 1}{2} M^2\right)^{\gamma/(\gamma-1)}.$$

Solve for the nonnegative speed magnitude. A static temperature measurement is needed to convert dimensionless Mach number into dimensional velocity. Expanding the pressure law for small Mach number recovers the familiar dynamic pressure as its leading term.

$$M = \sqrt{\frac{2}{\gamma - 1}\left[\left(\frac{p_0}{p}\right)^{(\gamma-1)/\gamma} - 1\right]},$$
$$p_0 - p = \tfrac{1}{2}\rho V^2 + O(p M^4).$$

Both pressures in the ratio must be absolute. A supersonic probe can form a shock and its recovered stagnation pressure then differs from the reversible upstream value.

WORKED EXAMPLE

Estimate air speed from two pressure readings

For air take $\gamma = 1.40$, $R = 287$ J/(kg K), $T = 290$ K, $p = 90.0$ kPa and $p_0 = 106$ kPa. Their ratio is 1.178, giving $M = \sqrt{5[(106/90)^{2/7} - 1]} = 0.489$. The sound speed is 341 m/s, so the flow speed is about 167 m/s. The incompressible pressure estimate using static density would be higher because the gas compresses while it decelerates.

Change the inputs and inspect this calculation below

SECOND APPLICATION

Predict the pressure ratio of a calibration stream

A second air stream has measured speed 102 m/s and static temperature 288 K, yielding $a = 340$ m/s and $M = 0.300$. At static pressure 100 kPa, the stagnation rise is

$$100[(1 + 0.2(0.300)^2)^{3.5} - 1] = 6.44\,\mathrm{kPa}.$$

The leading incompressible expression predicts

$$0.5\gamma p M^2 = 6.30\,\mathrm{kPa}.$$

Their 0.14 kPa difference sets a scale for the pressure resolution needed to detect compressibility here.

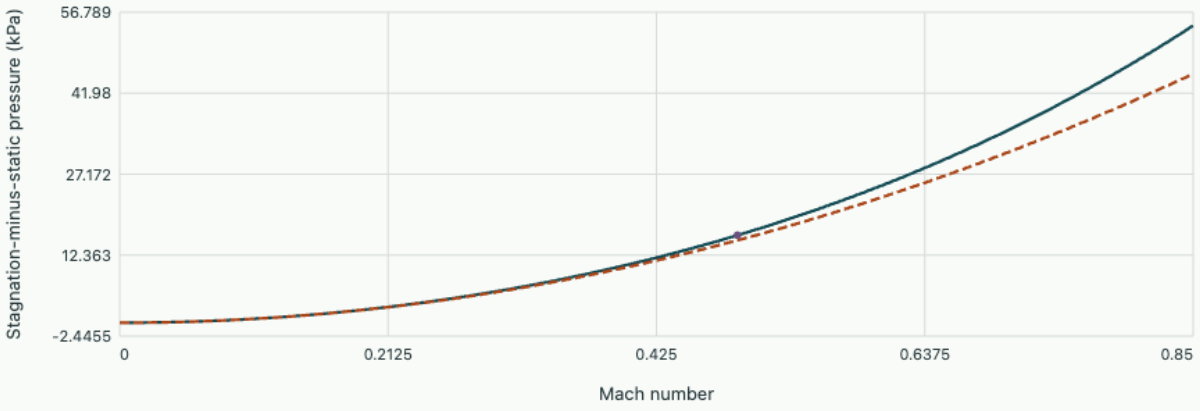
Interactive calculation · Change the inputs of this example

Use the adjustable quantities to test the two worked calculations above. The plot recomputes this derivation's balance or constitutive relation; its numerical readout reports the quantities needed for the forward calculation and the inverse or design question. Keep the fixed assumptions stated in the readout when comparing with the example values.

Open this live calculation

Try the calculation

Static pressure (kPa)	90
Pitot excess pressure (kPa)	16
Static temperature (K)	290



Inferred Mach=0.489; speed=167.0 m/s. The subsonic, shock-free pitot relation is used.

Incorporate expansion work and mass conservation

Gas accelerates through a nozzle by converting enthalpy into kinetic energy. At sufficiently high speeds, density changes appreciably and a constant-density volume balance fails. For steady one-dimensional adiabatic flow without shaft work or appreciable elevation change, $h + V^2/2 = h_0$. A calorically perfect

gas has $h = c_p T$; its local sound speed is $a = \sqrt{\gamma R T}$, and Mach number is $M = V/a$. Evaluate a at local static temperature, rather than at the upstream reservoir temperature.

Assume the acceleration is reversible, so $dp = a^2 d\rho$ along the isentrope. Differential momentum balance gives $dp = -\rho V dV$, hence $d\rho/\rho = -M^2 dV/V$. Inserting this in continuity $d\rho/\rho + dV/V + dA/A = 0$ gives the area–velocity relation

$$\frac{dA}{A} = (M^2 - 1)\frac{dV}{V}.$$

A subsonic stream accelerates through a converging passage. A supersonic stream accelerates through a diverging one. A smooth transition through sonic conditions requires a stationary area, conventionally a minimum-area throat for acceleration from subsonic to supersonic speed. A converging nozzle supplied from rest can first reach Mach one at its smallest area, the exit.

Obtain the throat state from the reservoir

In the upstream reservoir the flow is essentially at rest, so its static pressure and temperature are stagnation values p_0, T_0. Combining $c_p(T_0 - T) = V^2/2$ with $c_p = \gamma R/(\gamma - 1)$ gives

$$T/T_0 = [1 + (\gamma - 1)M^2/2]^{-1}.$$

Apply the ideal-gas isentropic pressure–temperature relation to find the critical pressure and temperature ratios at $M = 1$.

$$T^*/T_0 = 2/(\gamma + 1),$$

$$p^*/p_0 = [2/(\gamma + 1)]^{\gamma/(\gamma - 1)}.$$

With exit area fixed, lowering back pressure increases mass flow until the exit becomes sonic. Further reduction in back pressure does not change the ideal throat state for fixed reservoir conditions. Expansion downstream can then adjust the emerging jet to its surroundings.

Estimate a choked discharge

An insulated converging air nozzle has throat area $50\,\mathrm{mm}^2$ and reservoir conditions 300 kPa absolute and 300 K. Use $R = 287\,\mathrm{J/(kg\,K)}$ and $\gamma = 1.40$. The critical temperature is $300(2/2.4) = 250\,\mathrm{K}$, critical pressure $300(2/2.4)^{3.5} = 158.5\,\mathrm{kPa}$, density $158500/(287 \times 250) = 2.21\,\mathrm{kg/m^3}$ and critical speed $\sqrt{1.4(287)(250)} = 317\,\mathrm{m/s}$.

Atmospheric back pressure of 100 kPa lies below the critical pressure, so the ideal mass flow is

$$\rho^* a^* A^* = 2.21(317)(50 \times 10^{-6}) = 0.0350\,\mathrm{kg/s}.$$

A back pressure of 210 kPa would leave this converging nozzle unchoked, and the exit state must then be calculated from that back pressure. Real contraction and boundary-layer losses can reduce the flow through a calibrated orifice compared with the ideal nozzle prediction; a measured discharge coefficient accounts for those effects within its tested range.

Predict changes at a choked throat

At fixed reservoir pressure, the critical temperature is proportional to T_0, critical density to $1/T_0$ and critical speed to $\sqrt{T_0}$. Their product shows that the choked mass flow scales as $p_0 A^*/\sqrt{T_0}$ for the same gas. Warming the original reservoir from 300 to 360 K at unchanged absolute pressure therefore reduces the ideal mass flow by $\sqrt{300/360} = 0.913$, to about 0.0320 kg/s. Increasing throat area or reservoir pressure has a direct linear effect while the downstream pressure stays low enough to maintain choking.

Derive the admissible jump from conservation laws

Supersonic gas in a duct can decelerate abruptly through a normal shock. The shock is thin compared with the duct dimensions, so treat it as a stationary discontinuity of unchanged area. Take a calorically perfect gas, no heat or shaft work and negligible elevation change. Mass, streamwise momentum and stagnation enthalpy are conserved:

$$\rho_1 V_1 = \rho_2 V_2,$$

$$p_1 + \rho_1 V_1^2 = p_2 + \rho_2 V_2^2,$$

$$c_p T_1 + V_1^2/2 = c_p T_2 + V_2^2/2.$$

Entropy increases through the shock, so the two end states cannot be joined with the isentropic pressure–temperature relation. Set $r = \rho_2/\rho_1 = V_1/V_2$. Momentum gives $p_2/p_1 = 1 + \gamma M_1^2 (1 - 1/r)$ and the gas law gives

$$T_2/T_1 = (p_2/p_1)/r.$$

Substitution into the energy equation yields two roots, one for unchanged flow and the compression solution

$$r = \frac{(\gamma+1)M_1^2}{2 + (\gamma-1)M_1^2},$$

$$M_2^2 = \frac{1 + (\gamma-1)M_1^2/2}{\gamma M_1^2 - (\gamma-1)/2}.$$

For $M_1 > 1$ this second state has higher density and lower Mach number. As $M_1 \to 1^+$, every state ratio tends to unity. For very large upstream Mach number the compression ratio approaches $(\gamma+1)/(\gamma-1)$ in the constant-specific-heat approximation, although real-gas changes eventually limit that approximation.

Calculate a diffuser’s stagnation-pressure loss

Air upstream of a normal shock has $M_1 = 2.0,\ p_1 = 80\,\mathrm{kPa}$ absolute and $T_1 = 250\,\mathrm{K}$. Use $\gamma = 1.4,\ R = 287\,\mathrm{J/(kg\,K)}$. The density ratio is

$$2.4(4)/(2 + 0.4(4)) = 2.667,$$

$$p_2/p_1 = 1 + 1.4(4)(1 - 1/2.667) = 4.50.$$

Hence $p_2 = 360\,\mathrm{kPa}$ and $T_2 = 250(4.50/2.667) = 421.9\,\mathrm{K}$. The downstream Mach number is

$$\sqrt{1.8/(5.6 - 0.2)} = 0.577.$$

Upstream speed is approximately 634 m/s and mass conservation gives about 238 m/s downstream.

The stagnation temperature upstream is $250[1 + 0.2(4)] = 450\,\mathrm{K}$; downstream it is $421.9[1 + 0.2(0.577)^2] = 450\,\mathrm{K}$ as required by energy conservation. Evaluate stagnation pressure separately from each local state by a hypothetical isentropic deceleration on its own side. Upstream $p_{01} = 80(1.8)^{3.5} = 626\,\mathrm{kPa}$, whereas downstream $p_{02} = 360(1.0667)^{3.5} = 451\,\mathrm{kPa}$. Recovery $p_{02}/p_{01} = 0.721$ shows that a large static pressure rise can accompany a 28% loss of stagnation pressure. For the same stagnation temperature the entropy rise is $-R\ln(0.721) = 94\,\mathrm{J/(kg\,K)}$, a positive value consistent with the irreversible deceleration.

A pitot tube facing the upstream supersonic flow forms its own shock ahead of the tip. The final stagnation pressure measured there corresponds ideally to isentropic deceleration after the probe’s shock, about 451 kPa here, rather than to the upstream isentropic stagnation pressure of 626 kPa. A conventional subsonic pitot reduction applied to the reading would therefore infer an incorrect upstream speed. By contrast, a static pressure tap placed sufficiently far upstream can still report the 80 kPa static state. The ratio $451/80 = 5.64$ illustrates why the probe configuration must be stated when comparing a pressure ratio with a compressible-flow prediction.

The downstream temperature calculation can also be checked using mass flux. The upstream density is $80000/[287(250)] = 1.12\,\mathrm{kg/m^3}$ and the measured downstream static values imply $360000/[287(421.9)] = 2.97\,\mathrm{kg/m^3}$. Multiplication by the corresponding velocities gives roughly $707\,\mathrm{kg/(m^2 s)}$ on both sides. A calculation conserving stagnation pressure instead would generally fail one of these basic balances even if its pressure trend looked plausible.

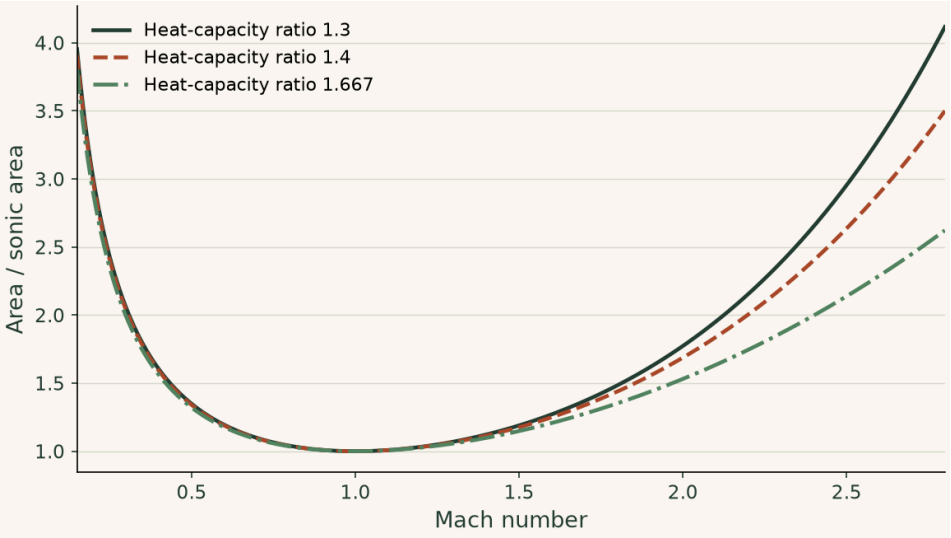
The shock also changes the area needed to choke a downstream passage. For a given gas the choked mass flow scales as $A^* p_0/\sqrt{T_0}$. Stagnation temperature is unchanged but stagnation pressure drops to 72.1% of its upstream value. Passing the same mass flow therefore requires a downstream critical area larger by $1/0.721 = 1.39$. This does not mean the physical duct jumps in area at the normal shock. The critical area is the sonic reference area for an isentropic stream tube drawn from each side, and the two hypothetical reference conditions differ because entropy has increased.

A converging test section downstream that was sized from the upstream stagnation pressure may consequently choke earlier than expected. The entropy and critical-area tests point to the same lost capacity from different directions: one uses the gas state, the other its maximum deliverable mass flux.

Together with conservation of actual mass flux at the shock, they make it possible to distinguish a real shock loss from a simple mismatch between static and total pressure instrumentation.

Find the two states that share a nozzle area ratio

For steady isentropic flow of a calorically perfect gas, combine continuity with stagnation enthalpy and the isentropic pressure law. The critical area A^* is the area giving Mach one for the same mass flow and stagnation state.



At γ = 1.4 and A/A* = 2 the two solutions are approximately Mach 0.306 and 2.20. Geometry alone does not choose the branch. Reservoir and downstream conditions, and possible shock formation, complete the physical specification.

▶ **Read values from the plot**

Interactive calculation · change the inputs of this example

Change heat-capacity ratio and target area. Track both Mach branches and their common sonic limit.

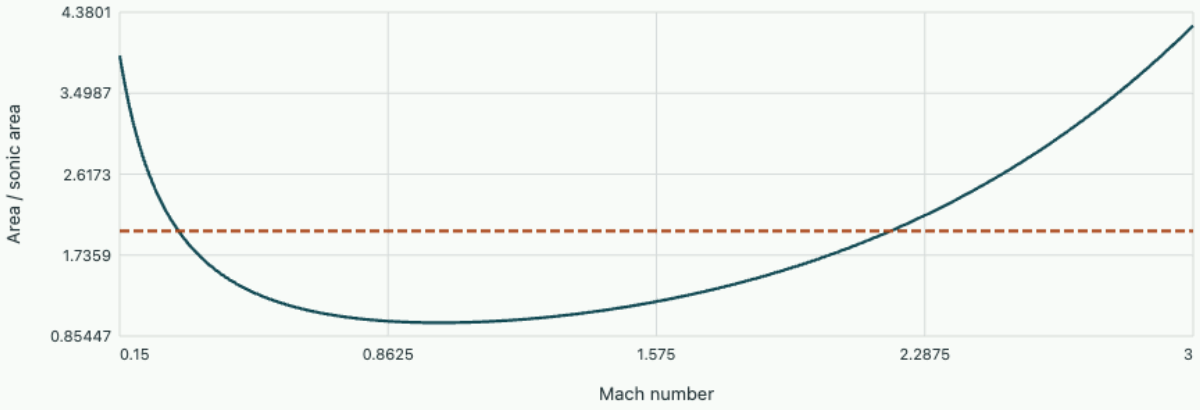
The curves below are recalculated from this section's equations in your browser. The fixed reference plot above remains available for comparison.

Open this live calculation

Try the calculation

Heat-capacity ratio **1.4**

Required area ratio **2**



Isentropic area map Selected area

Subsonic Mach 0.306; supersonic Mach 2.197. Boundary conditions determine which branch can occur.

Time-dependent diffusion

Numerical partial differential equations

Derive the diffusion time-step restriction from two independent arguments and compare explicit and implicit decay of spatial modes.

FROM FIRST PRINCIPLES

Derive the explicit heat step and its grid-dependent time scale

Take the heat equation $T_t = \alpha T_{xx}$ with constant positive diffusivity on a uniform one-dimensional grid. Fourier conduction through each side of a cell produces a temperature update proportional to the difference of the two neighbour gradients. With forward Euler timing, collect those fluxes to obtain

$$T_i^{n+1} = T_i^n + r(T_{i-1}^n - 2T_i^n + T_{i+1}^n),$$
$$r = \alpha \Delta t / h^2.$$

When $0 \le r \le 1/2$ the new temperature is a convex average of three old ones, preserving bounds on a source-free grid. To check whether a higher step could still be stable for all periodic grid disturbances, substitute $T_j^n = g^n e^{ij\theta}$, where j is the node index. The amplification simplifies to

$$g = 1 + r(e^{-i\theta} - 2 + e^{i\theta}) = 1 - 4r\sin^2(\theta/2).$$

The alternating mode at $\theta = \pi$ gives $g = 1 - 4r$, so $|g| \le 1$ again requires $r \le 1/2$. Halving spatial spacing therefore quarters the largest explicit time step for this discretization. Implicit stepping changes this restriction but still has to resolve any transient relevant to the measurement.

DERIVATION AND APPLICATIONS

Can an energy-stable heat step create an alternating transient?

A diffusion mode should decay rather than oscillate. To distinguish decay of a norm from preservation of sign, examine the finite-element or finite-volume semidiscrete system $M\dot{u} + Ku = 0$, with symmetric positive definite mass matrix M and thermal stiffness K positive semidefinite. An eigenvector satisfying $Kv = \lambda M v$ has scalar amplitude $y(t)$ obeying $y' = -\lambda y$. A generalized θ step applied to this mode gives

$$[1 + \theta z]y^{n+1} = [1 - (1-\theta)z]y^n,$$
$$z = \Delta t\, \lambda.$$

For Crank–Nicolson take $\theta = 1/2$. The factor $G = (1 - z/2)/(1 + z/2)$ has $|G| \le 1$ for all $z \ge 0$ but changes sign when $z > 2$. An initial grid-scale component can therefore alternate in time. To test a remedy, replace the first full step by two half backward-Euler steps. Their combined factor is

$$G_{\mathrm{start}} = (1 + z/2)^{-2}.$$

Thereafter return to the second-order Crank–Nicolson update. After n further full steps the mode amplitudes, with and without the initial damping, are

$$y_{\mathrm{damped}} = G_{\mathrm{start}} G^{n} y^{0},$$
$$y_{\mathrm{plain}} = G^{n+1} y^{0}.$$

For small z, both startup strategies agree through first order, while for large z the two half steps strongly suppress the unresolved mode. This argument explains a startup treatment for nonsmooth initial data; sign preservation of every physical grid value is a stronger matrix property than sign preservation of one mode.

WORKED EXAMPLE

Check damping of a near-grid-scale eigenmode

Take $\lambda = 2000$ s⁻¹ and $\Delta t = 0.0100$ s. Then $z = 20$. A single Crank–Nicolson step leaves $-9/11 = -0.818$ of the amplitude; two backward-Euler half steps leave $1/121 = 0.00826$. The exact amplitude factor is

$$e^{-20} = 2.06 \times 10^{-9}.$$

Both discrete results are bounded, but the startup treatment removes almost all of a mode that the time grid cannot resolve.

SECOND APPLICATION

Check the cost for a physically slow component

For the same mesh and step, a slow mode with $\lambda = 2.00$ s⁻¹ has $z = 0.0200$. Crank–Nicolson gives $0.99/1.01 = 0.980198$, the startup gives $1/1.01^{2} = 0.980296$ and the exact factor is 0.980199. The slow component changes by less than 10^{-4} in this comparison. A timestep must still resolve a transient that is itself part of the desired output.

Discretise space and time separately

A heat pulse in a thin insulated rod smooths over time according to $u_t = D u_{xx}$ for constant positive diffusivity D. Use a uniform grid of spacing h, and initially examine interior points or periodic boundaries so the stencil is unchanged everywhere. Replacing the second derivative by a centred difference gives a system of ordinary differential equations for the nodal temperatures. The time method still has to resolve the fastest rate that this spatial operator can generate.

An explicit Euler step yields $u_i^{n+1} = r u_{i-1}^n + (1-2r) u_i^n + r u_{i+1}^n$ with $r = D\Delta t / h^2$. If $0 \le r \le 1/2$, the update is a convex combination of three temperatures. The maximum cannot rise and the minimum cannot fall in a source-free interior step. Starting with one hot node at 1 and neighbours at zero, $r = 0.4$ gives a centre value 0.2 and neighbours 0.4 after a single step. At $r = 0.6$ the centre would become negative although the continuum heat equation preserves nonnegativity.

Identify the spatial modes that set the limit

Insert $u_j^n = g^n e^{\mathrm{i} j \theta}$, where i is the imaginary unit and j indexes the grid. The two neighbouring values contribute factors $e^{\pm \mathrm{i}\theta}$, yielding $g = 1 - 4r \sin^2(\theta/2)$. Long waves have $\sin^2(\theta/2) \simeq \theta^2/4$, so their decay approaches the continuum rate as the mesh is refined. The alternating grid mode $\theta = \pi$ has the most negative amplification and enforces $r \le 1/2$ for bounded magnitude.

At the limiting value $r = 1/2$ the alternating mode flips sign without losing amplitude. A strict margin damps it. For $D = 10^{-4}$ m²/s and $h = 2$ mm, the maximum explicit step is

$$h^2/(2D) = 0.020\,\mathrm{s}.$$

Refining to 1 mm reduces that limit to 0.005 s. The mesh gives a sharper temperature profile yet can cost four times more steps over the same physical interval, besides the increased number of spatial degrees of freedom.

Compare backward Euler and Crank–Nicolson

Writing the semidiscrete system as $u' = -A_h u$, backward Euler solves $(I + \Delta t A_h) u^{n+1} = u^n$. Crank–Nicolson averages the old and new diffusive fluxes, giving $(I + \Delta t A_h/2) u^{n+1} = (I - \Delta t A_h/2) u^n$. For a Fourier mode put $s = 4r \sin^2(\theta/2)$. Their factors are $1/(1+s)$ and $(1 - s/2)/(1 + s/2)$, respectively. Both remain bounded at arbitrarily large s, but Crank–Nicolson may alternate for under-resolved sharp features.

With $s = 20$, backward Euler retains only $1/21$ per step. Crank–Nicolson yields $-9/11$, retaining much more of the fast component and reversing its sign. For a smooth low-frequency mode Crank–Nicolson generally gives more accurate temporal decay because it is second order. Decide whether early sharp transients or later smooth behaviour is the relevant physical quantity. To identify spatial and temporal error separately, refine Δt first on a fixed mesh, then refine space with time error kept well below spatial error. Merely halving both at once cannot reveal which term dominated the original discrepancy.

Apply a prescribed heat input without confusing it with temperature

An insulated periodic rod with uniform heating rate s measured in K/s has an average temperature that must rise by $s\Delta t$ each step. Sum the centred second differences over the periodic indices: each nodal value appears with coefficients $1, -2, 1$, which cancel. The source alone changes the mean. For $s = 3$

K/s and a step of 0.02 s the numerical average must increase by 0.06 K, independently of the diffusivity and initial profile.

A code can pass this global test while still mishandling short-wave damping, since an alternating error has zero spatial mean. Pair a mean-temperature test with a nonconstant mode-decay test. Use a homogeneous periodic run for isolates of diffusion, and a uniform temperature with prescribed source for the storage term. A failure in one test then identifies a different part of the mathematical balance than a failure in the other.

Prove decay of a discrete thermal energy

Let a closed semidiscrete diffusion model be $M\dot{u} + Ku = 0$, with M symmetric positive definite and K symmetric positive semidefinite. Assume $K\mathbf{1} = 0$, expressing zero flux for a uniform temperature. Backward Euler satisfies $M(u^{n+1} - u^n) + \Delta t K u^{n+1} = 0$. Multiplying by the constant vector and using symmetry gives

$$\mathbf{1}^T M u^{n+1} = \mathbf{1}^T M u^n.$$

Thus the capacity-weighted sum, proportional to stored sensible heat for constant capacities and a fixed reference, is conserved for any step size.

Now multiply by u^{n+1}. The identity $a^T M(a-b) = [\|a\|_M^2 - \|b\|_M^2 + \|a-b\|_M^2]/2$ yields the exact balance

$$\tfrac{1}{2}\|u^n\|_M^2 - \tfrac{1}{2}\|u^{n+1}\|_M^2 = \Delta t\,(u^{n+1})^T K u^{n+1} + \tfrac{1}{2}\|u^{n+1} - u^n\|_M^2.$$

Both terms are nonnegative. The quadratic quantity therefore decays although total sensible heat is conserved. Subtracting the conserved mean first yields the same decay statement for temperature variation around equilibrium. The first term represents spatial smoothing measured at the new time, and the second reflects the numerical time update. This derivation gives a stability certificate on nonuniform grids without assuming Fourier modes.

Audit two cells with different heat capacities

Choose capacities 1 and 3 J/K, conductance 2 W/K, and initial temperature deviations $(4, 0)^T$ K. The conserved sensible heat is 4 J, so the final equilibrium in a closed system must be one kelvin above reference. At $\Delta t = 0.5$ s the matrix $M + \Delta t K$ has rows $(2, -1)$ and $(-1, 4)$, and the right-hand side is $(4, 0)^T$. The new temperatures are $(16/7, 4/7)^T$. Checking their weighted sum gives $16/7 + 3(4/7) = 4$, as required. The unweighted average changes, which is physically appropriate for unequal capacities.

The quadratic quantity at the start is $4^2/2 = 8$. After the step it is

$$[(16/7)^2 + 3(4/7)^2]/2 = 152/49.$$

Its drop is $240/49$. The conductive contribution is $(0.5)(2)(12/7)^2 = 144/49$; the time-update contribution is

$$[(-12/7)^2 + 3(4/7)^2]/2 = 96/49.$$

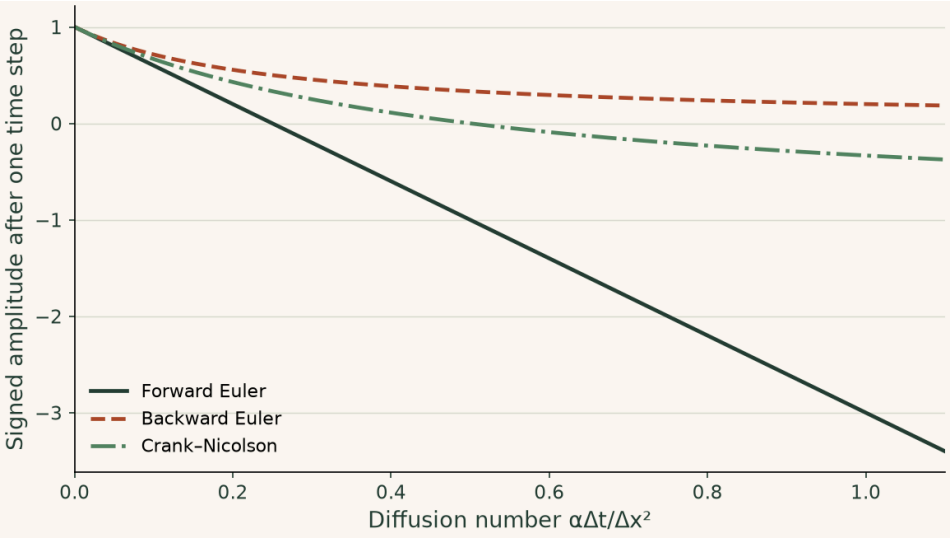
Their sum recovers the measured drop exactly. Neither contribution equals a loss of the 4 J of sensible heat. The quadratic measure depends on squared temperatures and has different physical units.

A positive heat supply or a prescribed exterior temperature introduces a boundary or source work term into these identities. Applying the closed-system decay test to an externally heated experiment would therefore diagnose the wrong physical model. Verify which conserved or dissipated quantity is actually implied by the capacity matrix, boundary conditions and source terms before comparing numerical integrators.

The conserved mean suggests another check after many large steps. The steady state must have zero conductance dissipation, which for two connected cells requires equal temperatures. Together with the weighted heat total this fixes both temperatures at one kelvin above reference. A scheme that instead approaches the unweighted initial average of two would be converging to the wrong conserved quantity despite producing a visually smooth, uniform endpoint. The measured quadratic variation about the correct mean tends to zero; the uncentred squared quantity tends to a positive constant because the remaining uniform state still has temperature above the chosen reference.

What happens to a grid-scale temperature ripple?

The alternating spatial mode on a uniform heat grid has discrete Laplacian eigenvalue $-4/\Delta x^2$. Substitute this mode into each method and compare the amplitude after one step. The plot isolates the shortest resolvable pattern.



Forward Euler reverses the mode above 1/4 and becomes unstable above 1/2. Crank–Nicolson stays bounded for every plotted step but also reverses the mode above 1/2. Smooth initial data can hide this high-frequency diagnostic.

▶ **Read values from the plot**

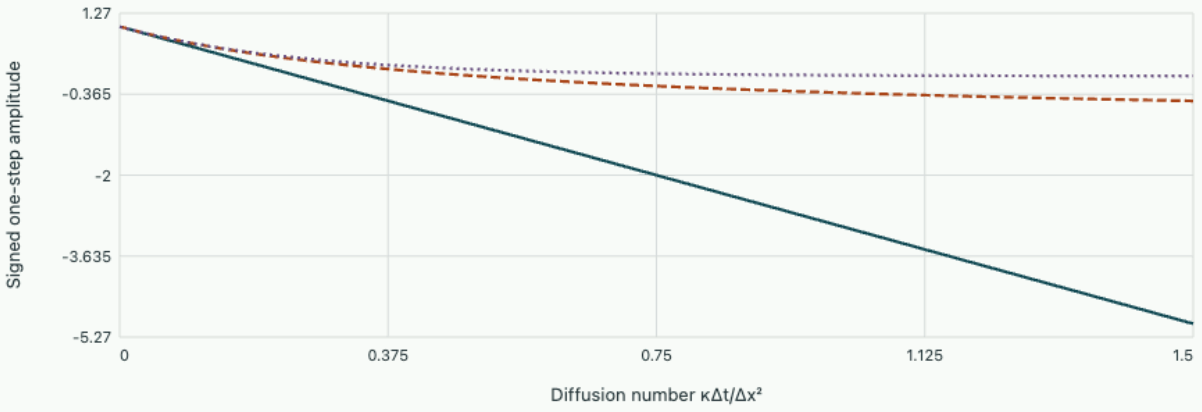
Change parameters · What happens to a grid-scale temperature ripple?

Change one input at a time and compare the recalculated curves with the argument above. The browser evaluates this small model locally; the accompanying numerical values state its assumptions and reference checks. Reset restores the reference inputs, and the data can be downloaded as CSV.

Open this calculation with adjustable inputs

Try the calculation

Highest diffusion number	**1.5**
Mode fraction of Nyquist	**1**



Directional Laplacian magnitude times Δx² is 4. Forward Euler is stable up to r = 0.5 and nonnegative only up to r = 0.25.

Interfaces, contact resistance and Robin data

Numerical partial differential equations

Derive a face flux that transports the same heat through two different materials and through a convective boundary.

FROM FIRST PRINCIPLES

Derive an interface flux from two half-cell temperature drops

Two neighbouring cell centres can lie in materials with sharply different conductivities. Consider a planar interface between them, normal to a one-dimensional steady heat flow. Let d_L, d_R be the distances from the centres to the interface, and suppose there is no interface source, storage or contact resistance. Steady conservation requires one heat flux q, positive to the right, and perfect contact requires one interface temperature T_I. Applying Fourier's law to each half gives

$$q = k_L(T_L - T_I)/d_L = k_R(T_I - T_R)/d_R.$$

Solve each expression for its temperature drop, add the two drops and eliminate T_I. The conductance follows from the actual material distances:

$$q = \frac{T_L - T_R}{d_L/k_L + d_R/k_R},$$

$$k_{\text{eff}} = \frac{d_L + d_R}{d_L/k_L + d_R/k_R}.$$

For equal distances this is the harmonic mean; only equal conductivities recover the same result as the arithmetic mean. If one material approaches zero conductivity with positive thickness, the flux approaches zero. A finite thermal contact resistance adds another temperature drop qR_c, while a heat-generating interface alters the flux-continuity condition itself.

DERIVATION AND APPLICATIONS

Enforce the flux jump at a heat-generating interface

Two materials in series carry the same conductive current only if their common interface stores and generates no energy. Consider two homogeneous slabs with an idealized thin heater between them. Let T_L, T_R be prescribed outer surface temperatures, R_L, R_R the positive thermal resistances per unit area in m² K/W, and s the heater release in W/m². Assume perfect thermal contact, steady conduction and negligible thickness of the heater. Define the interface temperature as T_I and flux q positive to the right.

Fourier's law integrated through each slab relates its own constant flux to its temperature drop. The pillbox balance across the heater requires the right flux to exceed the left one by the generated heat:

$$q_L = (T_L - T_I)/R_L,$$
$$q_R = (T_I - T_R)/R_R,$$
$$q_R - q_L = s.$$

Substitute the two face laws into the balance. With $g_j = 1/R_j$ as the two conductances per area, the interface temperature and individual fluxes are

$$T_I = \frac{g_L T_L + g_R T_R + s}{g_L + g_R},$$
$$q_R = g_R(T_I - T_R),$$
$$q_L = g_L(T_L - T_I).$$

Alternatively eliminate T_I to express the fluxes as through-flow plus a split source. The resistances determine the fraction of added heat directed to each boundary:

$$q_L = \frac{T_L - T_R - sR_R}{R_L + R_R},$$
$$q_R = \frac{T_L - T_R + sR_L}{R_L + R_R}.$$

Setting $s = 0$ restores a single common flux. Equal outside temperatures send more generated heat toward the lower resistance. For an interface at a mesh face the two cells therefore need distinct physical face fluxes linked by the stated jump; assigning exactly the same flux while adding no source anywhere violates energy balance.

WORKED EXAMPLE

Determine both heat flows through an instrumented wall

Let $T_L = 320$ K, $T_R = 300$ K, $R_L = 0.100$, $R_R = 0.200$ m² K/W and $s = 150$ W/m². The interface sees conductances 10.0 and 5.00 W/(m² K) and reaches 323.33 K. The left flux is -33.3 W/m² and the right flux 116.7 W/m². Heat leaves on both sides, including toward the hotter external surface. Their oriented difference equals 150 W/m².

SECOND APPLICATION

Select a heater setting that makes one sensor read zero

Keeping the same temperatures and layers, impose $q_L = 0$. Then

$$T_I = T_L = 320\,\mathrm{K},$$
$$s = (T_L - T_R)/R_R = 100\,\mathrm{W/m^2}.$$

All generated heat exits on the right. Reducing s to zero instead gives 66.7 W/m² from left to right through both layers. A zero left-side flux in the heated experiment therefore does not imply an insulating left layer.

Touch two long solids: which temperature does the interface acquire?

Two long blocks, initially at temperatures $T_L > T_R$, are pressed into ideal thermal contact at $x = 0$. Denote conductivity by k, volumetric heat capacity by C, and diffusivity by $\alpha = k/C$. At early times the penetration distance on each side is small compared with its length. Model the blocks as half-spaces with a common interfacial temperature T_*. This transient contact problem has no fixed layer thickness to insert into a pair of steady thermal resistances.

In the right half-space the similarity coordinate is $x/(2\sqrt{\alpha_R t})$. Substituting a function F of this coordinate into the heat equation gives

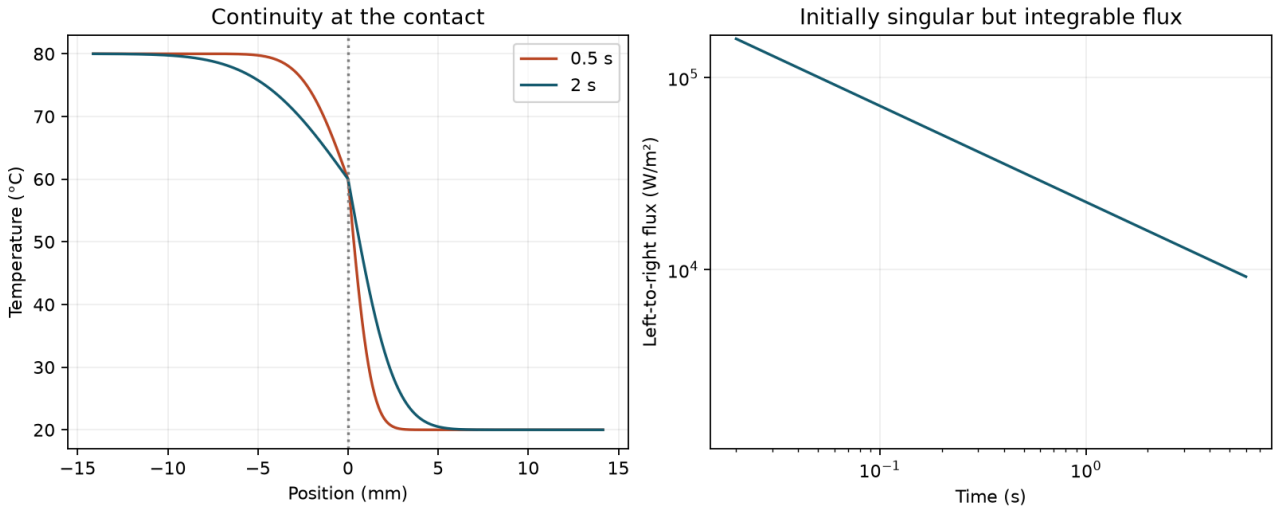
$$F'' + 2\eta F' = 0.$$

Its bounded solution matching a fixed surface and a uniform distant temperature has an error-function shape. Thus $T_R(x,t) = T_* + (T_R - T_*)\,\mathrm{erf}[x/(2\sqrt{\alpha_R t})]$. The left solution uses distance $-x$, its own diffusivity and the initial left temperature. These profiles share their interface temperature by construction; conservation still requires their heat fluxes to match.

Differentiate the error functions at zero and apply Fourier's law. The left flux equals $e_L(T_L - T_*)/\sqrt{\pi t}$, and the right flux is $e_R(T_* - T_R)/\sqrt{\pi t}$, both oriented from left to right. Here $e = \sqrt{kC}$ is thermal effusivity. Equating the fluxes gives

$$T_* = (e_L T_L + e_R T_R)/(e_L + e_R).$$

The interfacial temperature is independent of time within this idealisation, while the penetration layers broaden continuously.



The two temperature gradients differ because flux continuity weights each gradient by its own conductivity. The common temperature depends on heat capacity as well as conductivity.

For the notebook's $k_L = 4,\ k_R = 1\,\mathrm{W/(m\,K)}$ and equal capacities of $10^6\,\mathrm{J/(m^3\,K)}$, the effusivity ratio is two. Between 80 and 20 degrees Celsius the common temperature is 60 degrees Celsius. The initially infinite ideal flux is integrable: the energy transferred by time t equals $2tq(t)$ per unit contact area. The notebook independently integrates temperature loss and gain throughout the two bodies and checks this balance. A real surface resistance, a finite approach time or the far surface of a finite specimen would eventually invalidate different parts of the idealisation.

A finite-volume scheme offers a sensitive check of this benchmark. On a sufficiently long computational domain, sum nodal or cell-average energies on each side at the same time. Internal face fluxes cancel in the total by conservation. A scheme may pass that global check while partitioning energy incorrectly between materials. Compare both individual energy changes, the early flux integrated over the first step and the temperature nearest the interface. This combination tests the material weighting and the time integration independently.

Keep both initial temperatures and conductivities fixed, but multiply the right-hand volumetric heat capacity by four. Predict T_* and the right-hand penetration depth at the same time.

▼ Work through the solution

The right effusivity doubles and now equals the left one. The interface becomes 50 degrees Celsius. At fixed k_R the right diffusivity falls to one quarter, so the penetration distance proportional to $\sqrt{\alpha_R t}$ halves. A larger local capacity can draw more flux while confining the altered temperature to a narrower region.

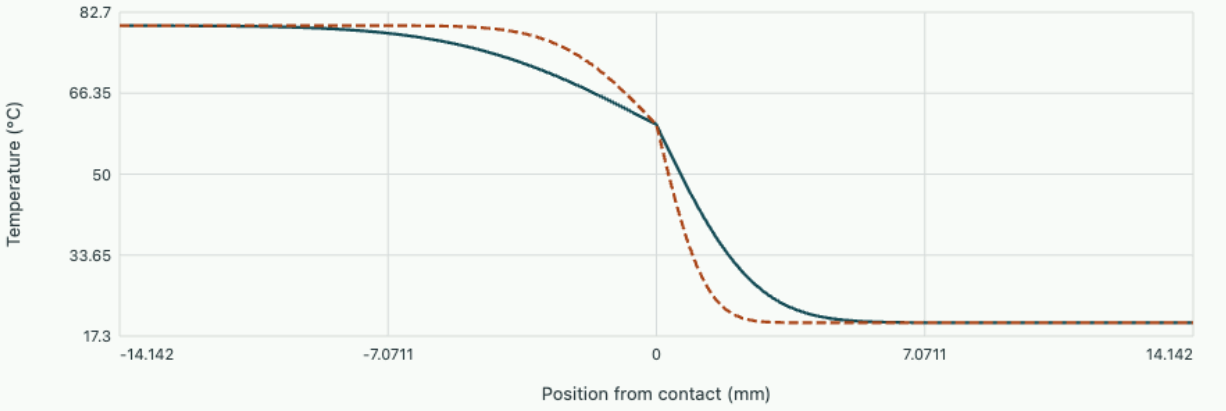
Change parameters · Explore this notebook example

Change one input at a time and compare the recalculated curves with the argument above. The browser evaluates this small model locally; the accompanying numerical values state its assumptions and reference checks. Reset restores the reference inputs, and the data can be downloaded as CSV.

Open this calculation with adjustable inputs

Try the calculation

Left conductivity (W/m/K)	4
Right conductivity (W/m/K)	1
Elapsed time (s)	2



Common interface temperature 60 °C; heat flux 15958 W/m². Both semi-infinite materials have volumetric heat capacity 10⁶ J/m³/K; contact resistance is zero.

Open editable notebook Download .ipynb

One flux through resistances in series

Let $q = -kT'$ be heat flux in the positive coordinate direction, with unit cross-sectional area and no internal source. Between two cell centres separated by an interface, let the left and right centre-to-interface distances be d_L, d_R. Piecewise constant conductivities $k_L, k_R > 0$ make the equilibrium temperature linear in each segment. If the interface has perfect contact, its temperature T_f is continuous. Integrating Fourier's law on each side gives

$$T_L - T_f = q\,d_L/k_L,$$
$$T_f - T_R = q\,d_R/k_R.$$

Adding eliminates the unknown interface temperature.

$$q = \frac{T_L - T_R}{d_L/k_L + d_R/k_R}.$$

For equal distances this amounts to using the harmonic mean conductivity over the full centre spacing. Inserting the same signed q as the outflow of the left cell and the inflow of the right cell produces cancellation when their balances are added. If a material interface stores energy or supplies heat independently, a flux jump belongs in the balance and this steady series calculation must be modified.

Measure the temperature drop

Take $d_L = d_R = 0.01$ m, $k_L = 2,\ k_R = 10$ W/(m K) and $T_L = 30,\ T_R = 18$ °C. The series resistance per unit area is $0.005 + 0.001 = 0.006$ m²K/W. The flux is 2000 W/m² and the interface temperature is $30 - 2000(0.005) = 20$ °C. The lower conductivity carries the larger temperature drop despite carrying the same flux. Using the arithmetic mean 6 with the centre difference over 0.02 m instead predicts 3600 W/m².

Introduce a contact resistance $R_c = 0.002$ m²K/W. Its temperature jump is qR_c; the total resistance rises to 0.008, reducing the heat flux to 1500 W/m². The two sides of the contact now have temperatures 22.5 and 19.5 °C. Imposing a single interface temperature on this model would remove the physically measured 3 K jump.

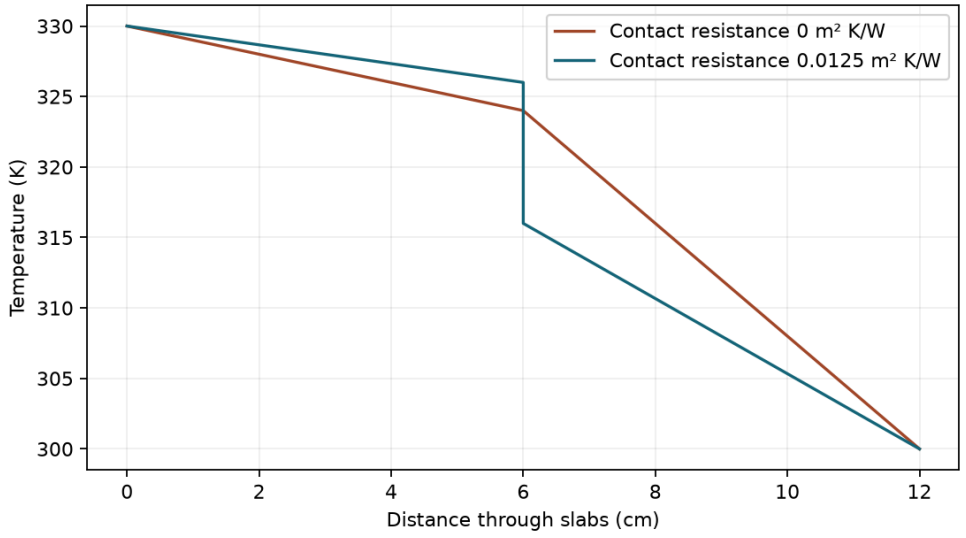
OneSimulation · Resolve a temperature jump at imperfect contact

Two slabs are each 0.06 m thick, with conductivities 12 and 3 W/(m K). Their outer surfaces are held 30 K apart. A contact resistance R_c permits two distinct temperatures at the shared physical location. The flux stays continuous. The face law and the exact reference for the whole assembly are

$$T_{f,L} - T_{f,R} = qR_c,$$
$$q = \frac{T_L - T_R}{L_1/k_1 + R_c + L_2/k_2}.$$

The code adds R_c only to the resistance between the two cells adjacent to the contact. OneSimulation solves the resulting balance system. The flux falls from 1200 to 800 W/m² for $R_c = 0.0125$ m² K/W. Reconstructing the two face temperatures gives 326 K and 316 K. Connecting only the cell-centre temperatures would smear this 10 K jump over one mesh cell. Refining an ideal-contact model would converge to the wrong interface law.

```
"""Find the temperature jump at an imperfect contact between two slabs."""
import matplotlib.pyplot as plt
import numpy as np
from onesim.numerics.linear.matrices.dense_linear_algebra import solve_dense

n = 12  # even number of equal, cell-centred control volumes
length = 0.12  # m
dx = length / n
k = np.repeat([12., 3.], n // 2)  # W/(m K)
left, right = 330., 300.  # K
centres = (np.arange(n) + 0.5) * dx
fig, ax = plt.subplots(figsize=(6.8, 3.8), constrained_layout=True)
for contact, colour in [(0.0, "#a4492b"), (0.0125, "#176879")]:  # m² K/W
    resistance = np.r_[dx/(2*k[0]), dx/(2*k[:-1]) + dx/(2*k[1:]), dx/(2*k[-1])]
    resistance[n//2] += contact
    g = 1 / resistance
    matrix = np.diag(g[:-1]+g[1:]) - np.diag(g[1:-1], 1) - np.diag(g[1:-1], -1)
    rhs = np.zeros(n)
    rhs[0], rhs[-1] = g[0]*left, g[-1]*right
    temperature = solve_dense(matrix, rhs)
    fluxes = -np.diff(np.r_[left, temperature, right]) / resistance
    q = fluxes[0]
    assert np.allclose(fluxes, (left-right)/(length/2/k[0]+contact+length/2/k[-1]))
    hot_face = temperature[n//2-1] - q*dx/(2*k[0])
    cold_face = temperature[n//2] + q*dx/(2*k[-1])
    assert np.isclose(hot_face-cold_face, q*contact)
    x = np.r_[0, centres[:n//2], length/2, length/2, centres[n//2:], length]
    y = np.r_[left, temperature[:n//2], hot_face, cold_face, temperature[n//2:], right]
    ax.plot(100*x, y, color=colour, label=f"Contact resistance {contact:g} m² K/W")
    print(f"Flux {q:.0f} W/m²; interface jump {hot_face-cold_face:.1f} K")
ax.set(xlabel="Distance through slabs (cm)", ylabel="Temperature (K)")
ax.legend()
ax.grid(alpha=0.2)
fig.savefig("partial-differential-equations.png", dpi=160)
```



The vertical segment reconstructs both surface temperatures at the same location. All control-volume faces carry the same steady flux, including the imperfect contact.

Download the Python example

Apply a convective boundary

At an exterior face the outward heat flux obeys $q = h_c(T_s - T_\infty)$, a Robin condition. Eliminate the unknown surface temperature as above to obtain

$$q = (T_P - T_\infty)/(d/k + 1/h_c).$$

If h_c tends to zero, the boundary becomes insulating. With very large h_c, the surface approaches ambient temperature while the finite centre-to-face resistance remains. For $d = 0.01$ m, $k = 2$, $h_c = 100$ W/(m² K), $T_P - T_\infty = 15$ K, the predicted flux is 1000 W/m² and the surface is 10 K above ambient. Multiplying by face area converts this flux to the power used in a cell energy balance. Mixing up these two quantities makes the result depend incorrectly on mesh area.

A jump located inside a cell

If the materials change at a known position inside a segment, the fractions of that segment matter. Moving the interface so that 75 percent of the length has conductivity 2 and 25 percent has conductivity 10 produces effective conductivity $1/(0.75/2 + 0.25/10) = 2.5$ W/(m K). Giving each material half the segment instead would give $10/3$ W/(m K). The harmonic rule depends on actual layer thicknesses. A mesh-refinement study that snaps the interface to different positions at every refinement can mix geometric representation error with flux-discretisation error.

Derive effective transport parallel and normal to layers

A fine laminate of isotropic materials can have a direction-dependent macroscopic conductivity. Consider two perfectly bonded planar layers occupying thickness fractions θ and $1 - \theta$, with conductivities k_1, k_2. Assume steady conduction, no internal sources, and boundary conditions that permit a one-dimensional field for the loading orientation being tested. For transport normal to the layers, continuity forces equal flux. Adding their temperature drops across a total thickness L gives

$$k_\perp = [\theta/k_1 + (1 - \theta)/k_2]^{-1}.$$

For transport parallel to the layers the imposed longitudinal gradient is common and the area-weighted fluxes add, giving $k_\parallel = \theta k_1 + (1 - \theta)k_2$.

The arithmetic and harmonic expressions are thus both correct, but apply to different geometries of imposed transport. Their ordering follows without a conductivity table: multiplying the two positive weighted sums $\theta k_1 + (1 - \theta)k_2$ and $\theta/k_1 + (1 - \theta)/k_2$ gives $1 + \theta(1 - \theta)(k_1 - k_2)^2/(k_1 k_2) \geq 1$. Equality requires identical conductivities or a missing layer. Spatial averaging performed before identifying the flux direction would erase precisely the distinction that determines the coarse physics.

Predict two experiments on the same material sample

Suppose a repeated laminate consists of 40 percent material with conductivity 1 and 60 percent with conductivity 10 W/(m K). Then $k_{\parallel} = 0.4 + 6 = 6.4$ W/(m K) and $k_{\perp} = 1/(0.4 + 0.06) = 50/23 \approx 2.174$ W/(m K). A temperature drop of 4 K over 0.02 m imposes gradient magnitude 200 K/m. Perfectly normal loading produces flux magnitude about 434.8 W/m², whereas the same gradient parallel to the layers yields 1280 W/m². The larger flux is consistent with the fast material providing a continuous path in the latter orientation.

For the normal experiment the low-conductivity phase occupies 0.008 m and accounts for a drop $434.8(0.008)/1 \approx 3.478$ K. The high-conductivity phase occupies 0.012 m and drops only $434.8(0.012)/10 \approx 0.522$ K. Adding recovers the imposed four kelvin. For the parallel experiment individual flux densities are 200 and 2000 W/m². Averaging by their area fractions gives

$$0.4(200) + 0.6(2000) = 1280\,\mathrm{W/m^2}.$$

These local checks use the same two material properties and explain the difference in the observed coarse responses.

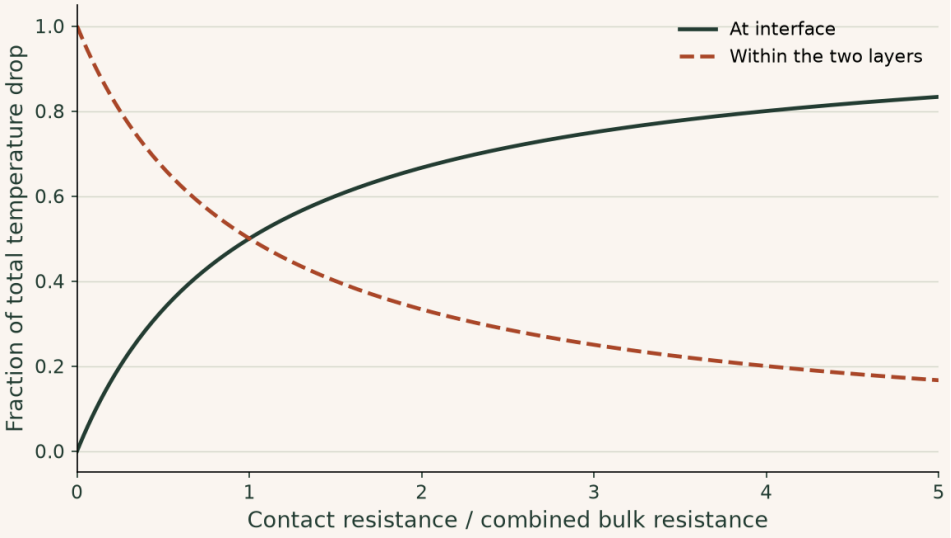
Using the effective tensor assumes the observation scale spans representative layers. A sensor within one layer may respond to its local temperature slope instead. Finite contact resistance adds a temperature jump in the normal direction; if the number of interfaces per macroscopic length changes, the effective normal resistance changes even at identical material fractions. Measuring both orientations can distinguish geometric or contact effects that a single scalar fitted conductivity would conceal.

Rotating a laminate specimen through ninety degrees interchanges the roles of these two effective values without changing its volume fractions. This makes an orientation experiment particularly discriminating: a fit that happened to reproduce the first orientation by adjusting a scalar conductivity predicts the same flux for the rotated specimen, whereas the fine-layer balance predicts a factor of about $6.4/(50/23) = 2.94$ between them.

Repeating the test over multiple layer spacings also probes the assumption of negligible contact resistance.

Separate the temperature jump from the bulk gradient

A steady one-dimensional stack has two positive bulk conduction resistances in series with a contact resistance R_c. Flux is common to all three. Scale R_c by the sum of bulk resistances and predict where probes would measure the imposed end-to-end temperature difference.



When contact resistance equals the bulk sum, half the end-to-end drop occurs at a geometrically thin interface. A numerical scheme that enforces temperature continuity there can converge smoothly to the wrong interface model.

▶ **Read values from the plot**

Change parameters · Separate the temperature jump from the bulk gradient

Change one input at a time and compare the recalculated curves with the argument above. The browser evaluates this small model locally; the accompanying numerical values state its assumptions and reference checks. Reset restores the reference inputs, and the data can be downloaded as CSV.

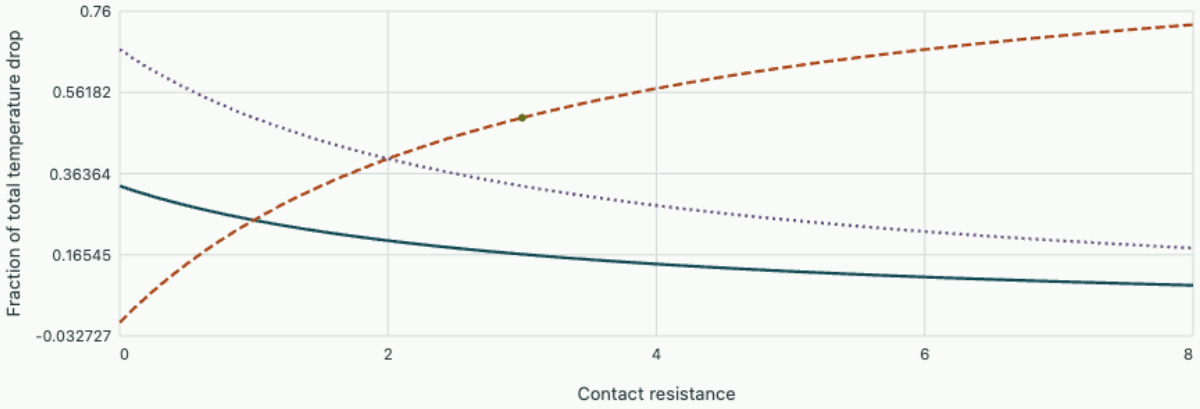
Open this calculation with adjustable inputs

Try the calculation

First bulk resistance 1

Second bulk resistance 2

Selected contact resistance 3



First bulk layer · Contact jump · Second bulk layer · Selected contact fraction

At selected contact resistance, the jump carries 50% of the imposed drop. The three fractions always sum to one.

What a periodic grid does to a decaying Fourier mode

Numerical partial differential equations

A periodic diffusion solver predicts the shape of a sampled wave while its amplitude decays too slowly. Fourier analysis separates spatial differentiation error, an unresolved physical mode and errors caused by the time integrator.

Diagonalise the differential operator on the periodic interval

Use a dimensionless coordinate $0 \le x < 2\pi$ and solve $u_t = \kappa u_{xx}$ with periodic values and periodic derivatives. The complex exponentials e^{imx}, with integer m, satisfy both conditions and are eigenfunctions of the second derivative. Put $u(x,t) = A_m(t)e^{imx}$ and cancel the common exponential. The scalar amplitude equation is

$$A_m' = -\kappa m^2 A_m.$$

Thus every cosine or sine mode decays by $\exp(-\kappa m^2 t)$, with no phase shift, while the m=0 mean is conserved.

On a uniform grid $x_j = 2\pi j/N$, the samples of m and $m + \ell N$ agree for every integer ℓ. A discrete Fourier transform returns coefficients in one representative band. For the real cosine test with even N we can fold a positive physical mode into $k = \min(r, N - r)$, $r = m \bmod N$. A Fourier spectral Laplacian multiplies that coefficient by $-k^2$. For a resolved mode it differentiates the trigonometric interpolant exactly, up to arithmetic. For unresolved data it differentiates the aliased representative; no operation using only those samples can identify the original m.

Apply the centred difference stencil to a single wave

The standard second difference is $(u_{j+1} - 2u_j + u_{j-1})/h^2$, $h = 2\pi/N$. Put the grid wave $u_j = e^{ikjh}$ into the stencil. Factor out the common phase and simplify with $e^{i\xi} + e^{-i\xi} = 2\cos\xi$. This gives the symbol and its small-h expansion:

$$\frac{e^{ikh} - 2 + e^{-ikh}}{h^2} = -\frac{4}{h^2}\sin^2(kh/2) = -k^2 + \frac{k^4 h^2}{12} + O(k^6 h^4).$$

The effective squared wavenumber is smaller than k^2 for every nonconstant mode in the resolved band, so this spatial approximation decays too slowly if the time equation is integrated exactly. The error is modest when $kh \ll 1$, but at the Nyquist mode $kh = \pi$ the ratio is

$$4/\pi^2 = 0.4053.$$

A stable time integrator cannot repair this spatial attenuation error without additional modeling. The experiment exponentiates both discretizations' symbols exactly, isolating space and sampling here.

Measure separate errors with the same sampled initial data

Choose $\kappa = 0.5,\ m = 3,\ N = 16$. The true squared wavenumber is 9 and the difference stencil gives 8.006. Starting at unit amplitude, at $t = 0.20$ exact and spectral amplitudes are 0.40657; finite differences give 0.44906. Doubling N to 32 changes the difference result to 0.41716. The absolute amplitude error falls from 0.04249 to 0.01059, a ratio of 4.01 consistent with the second-order expansion. This is a spatial-convergence study because time evolution introduces no truncation error.

Now take $m = 12,\ N = 16$. Every sampled cosine equals the low mode $k = 4$ at $t = 0$. At $t = 0.02$, the continuum amplitude for the true physical mode is $e^{-1.44} = 0.237$, whereas the spectral grid predicts

$$e^{-0.16} = 0.852.$$

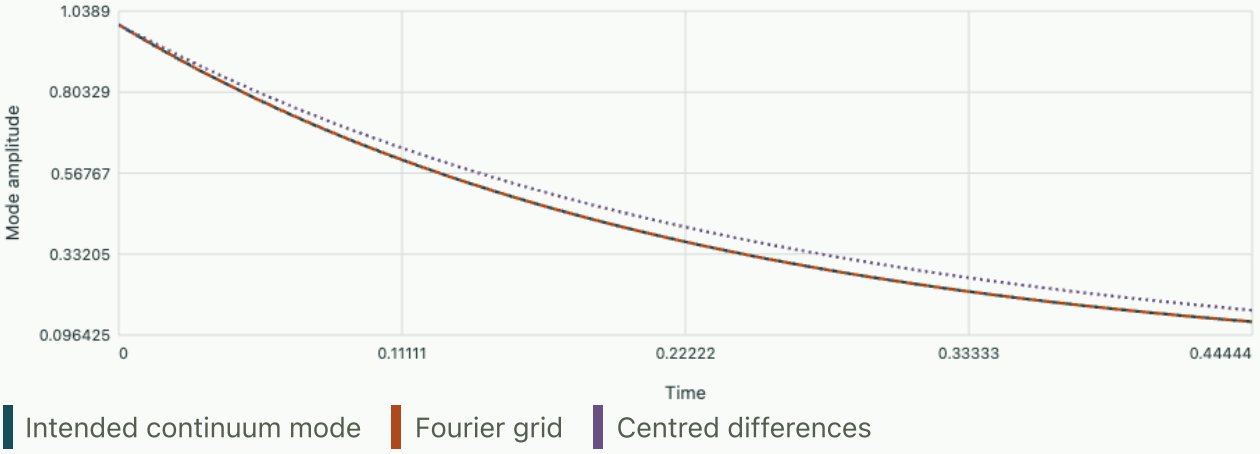
A spectral computation can therefore agree exactly with its trigonometric interpolant and disagree strongly with the evolution of the intended signal. Sampling on a finer grid or obtaining prior bandwidth information resolves that ambiguity. The mean mode is particularly revealing when a nonconstant physical cosine aliases to a constant.

Start at mode 3 and successively double N from 8 to 32. Compute two amplitude-error ratios at the same time. Then use mode 12 on 16 points to observe an alias, and increase N until the Fourier grid represents the intended mode.

Open this experiment online

Try the calculation

Periodic grid points	16
Physical cosine mode	3
Diffusivity	0.5



The grid represents mode 3 as 3. At t = 0.20, amplitudes are 0.40657 for the physical mode, 0.40657 for the Fourier grid and 0.44906 for the difference stencil.

Tutorial · resolve a damping rate to a target tolerance

For a mode well inside the grid bandwidth, estimate the number of samples per wavelength needed for a relative error below 1% in the decay rate with centred differences. Test the estimate for $m = 2$. How could the time integrator contaminate a measurement of this spatial error?

▼ **Show the resolution estimate**

Writing p for samples per wavelength gives

$$kh = 2\pi/p.$$

The relative rate error is approximately $(kh)^2/12$. Requiring at most 0.01 gives $p \gtrsim 18.14$; try 19 samples per wavelength, or $N = 38$ for $m = 2$. The exact ratio $[\sin(\pi/19)/(\pi/19)]^2$ is about 0.991. A forward-Euler step would replace exponential propagation by $1 - \kappa k_{\mathrm{eff}}^2 \Delta t$. Its repeated-step error could partly mask the underdamping of the spatial stencil. Compare with the exponential of the discrete symbol before attributing a measured rate to space alone.

Further reading · Lloyd N. Trefethen, Spectral Methods in MATLAB · periodic grids, Fourier differentiation and downloadable programs

Check the discrete eigenvalue with OneSimulation

A spatial sine also diagonalises the five-point Dirichlet Laplacian on a square. Use the same modified wave-number argument in both coordinate directions. Run `python poisson-mode.py --mode 2 --cells 8 16 32` and compare the observed nodal error with the discrete eigenvalue prediction in the CSV. This is a steady forced problem whose spatial operator can be checked independently of any time integrator. Download the OneSimulation verification model.

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

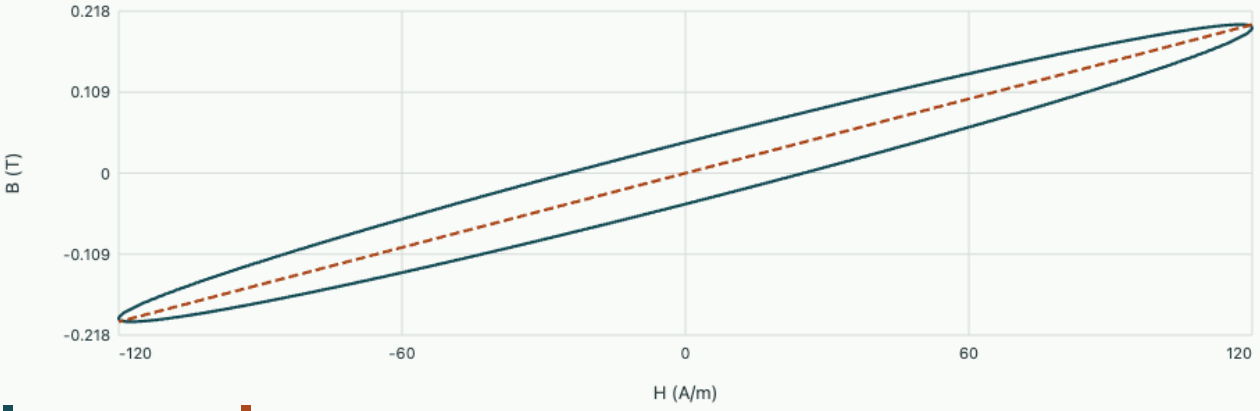
The live experiment reconstructs the H, B trajectory and compares it with a zero-lag reference. It uses the same peak field and induction for the two curves. At fixed peaks changing frequency leaves an assigned phase lag and therefore the prescribed energy per cycle unchanged, while watts rise with the 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.

Thermal radiance, étendue and detector calibration

Electromechanics, optics and measurement

A square millimetre of a heated object emits thermal radiation into an entire hemisphere. A detector sees only the rays accepted by its optics and only wavelengths to which it responds. Derive the equilibrium spectrum, follow it through an optical collection system and infer what temperature or emissivity a calibrated reading can actually determine.

Count optical states in a large equilibrium cavity

Take a cubic airless cavity of volume V in thermal equilibrium at temperature T. For its bulk mode density we can use periodic boundary conditions. Allowed wavevectors are separated by $2\pi/L$ in each coordinate. Count a spherical shell in wavevector space, divide by the volume per state and include two independent transverse polarisations. Substituting $k = 2\pi\nu/c$ gives

$$g(\nu)d\nu/V = 8\pi\nu^2 d\nu/c^3.$$

Using standing waves in an ideally reflecting large box gives the same leading volume density when directions and polarisations are counted consistently.

One oscillator mode of frequency ν can hold $m = 0, 1, 2, \ldots$ photons. Its thermal occupation has probabilities proportional to $e^{-mh\nu/k_BT}$. Summing the geometric series and taking its logarithmic derivative gives

$$\langle m \rangle = 1/(e^{h\nu/k_BT} - 1).$$

Multiplying by the density of modes and energy per photon gives spectral energy density u_ν. The zero-point contribution of a harmonic oscillator is not part of the thermal power exchanged between these black surfaces. At low frequency the result approaches an energy k_BT per mode; extending that classical limit to arbitrarily high frequency would make the integrated power diverge.

An isotropic distribution of energy travelling at speed c produces a normal flux through a small aperture weighted by the direction cosine. Per unit solid angle its radiance is $B_\nu = cu_\nu/(4\pi)$. Transform to equal wavelength intervals using $B_\lambda d\lambda = B_\nu |d\nu|$. The two useful spectral forms are

$$B_\nu = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/k_BT} - 1},$$

$$B_\lambda = \frac{2hc^2}{\lambda^5}\frac{1}{e^{hc/(\lambda k_BT)} - 1}.$$

The wavelength expression has units of watts per square metre per steradian per metre of wavelength. To express it per micrometre multiply by one millionth. At 400 K and 10 micrometres the result is 33.56 W per square metre per steradian per micrometre. At 295 K it is 9.14 in the same units. Equal wavelength and equal frequency bins carry different widths, so their distributions reach their maxima at different converted wavelengths. A spectrum's peak should always name its horizontal variable and spectral-density convention.

Distinguish emitter power from brightness in a chosen direction

For a black surface with angle-independent radiance into its outward hemisphere, integrate $\cos\theta\, d\Omega$. The result is π, not 2π. The emitted spectral power per unit area is πB_λ. Integrate over frequency using $x = h\nu/(k_B T)$. The remaining integral $\int_0^\infty x^3/(e^x - 1)\, dx = \pi^4/15$ follows by expanding $1/(e^x - 1)$ as a sum of exponentials and integrating term by term. This yields Stefan–Boltzmann exitance $M = \sigma T^4$ and identifies the constant $\sigma = 2\pi^5 k_B^4/(15 h^3 c^2)$.

A black patch of 1.0 square millimetre at 400 K emits about 1.45 mW into its entire outward hemisphere over all wavelengths. The same patch at 295 K emits about 0.429 mW. Replacing one by the other changes all-wavelength hemispherical power by about 1.02 mW. This is not the power a small infrared receiver sees through a narrow filter. The integrated emitter number is valuable for energy conservation, but by itself contains no receiver area, direction, passband or optical transmission.

The surface spectrum of a body in thermal equilibrium has directional spectral emissivity equal to its absorptivity in the corresponding reciprocal optical modes under the usual reciprocal-material conditions. An opaque diffuse grey surface with emissivity ϵ in an isotropically illuminated background can be modelled by $\epsilon B_\lambda(T) + (1 - \epsilon) B_\lambda(T_{\mathrm{bg}})$. It emits less than a black surface but also reflects background. A shiny object viewed in a warm room may consequently display a radiance temperature nearer the reflected scene than its own physical temperature. Specular surfaces require explicit directional knowledge of that scene.

Derive the geometric acceptance without counting the angle twice

An area element dA normal to the collection axis emits into $d\Omega$ at angle θ a spectral power $B_\lambda \cos\theta\, dA\, d\Omega$. For a uniformly filled cone of half-angle α and a uniform patch of area A, integrating gives $A\pi \sin^2\alpha$. This is its geometrical throughput in air when there is no vignetting. Optical étendue more generally uses $n^2 dA \cos\theta\, d\Omega$. A lossless passive optical relay in geometrical optics preserves this phase-space measure; correspondingly L/n^2 is conserved along a ray in transparent media under the usual stationary conditions.

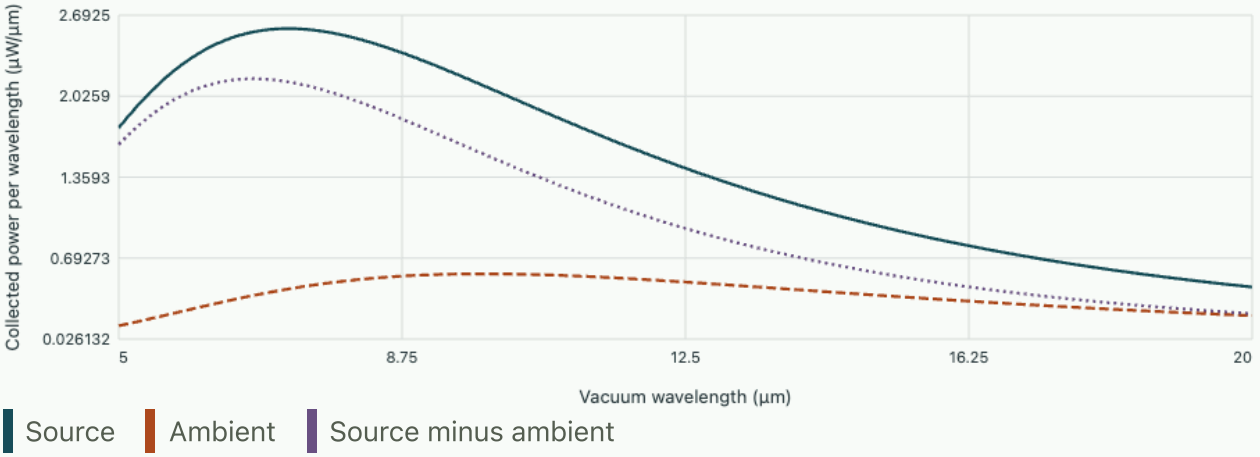
A direct lens example makes the restriction tangible. An image relay enlarging a uniformly bright patch by a factor of two in each transverse direction forms an image four times the area. In a lossless paraxial design filling the original pupil, each image point then occupies one quarter the solid angle. A detector covering the whole enlarged image and accepting that cone can receive the original power. Placing a fixed small sensor at the image plane intercepts less area instead. Higher magnification has not created additional radiance, even if it makes a small detail easier to resolve.

For the same 1.0 square millimetre patch seen within ten degrees, the projected acceptance is 9.47×10^{-8} square metre steradian. A wavelength-independent combined transmission and detector absorption of 0.65 leaves effective acceptance 6.16×10^{-8} square metre steradian. Integrating the Planck curve from 8 to 12 micrometres gives source and background band radiances 133.74 and 35.43 W per square metre per steradian. Their difference yields only 6.05 microwatts in the receiver. Keeping a single midpoint radiance times four micrometres happens to be fairly close here, but there is no fixed universal midpoint correction across temperatures or filters.

Open the live experiment

Try the calculation

Source temperature (K)	400
Ambient temperature (K)	295
Collection half-angle (°)	10



The net 8–12 µm signal is 6.053 µW for a 1 mm² area and 65% throughput. Ideal source peak wavelength is 7.24 µm.

Set source and ambient temperatures equal. The emitted and ambient curves remain finite while the differenced instrument reading vanishes. Increase the cone half-angle from ten to twenty degrees and predict the scale factor $\sin^2 20^\circ / \sin^2 10^\circ = 3.88$. The spectral shape should stay fixed if transmission is separable from angle. If real data shift spectral shape with aperture, inspect filter angle dependence, source nonuniformity and stray light. The integral in the readout is computed on the specified 8–12 micrometre band even though the graph shows wavelengths beyond it.

Convert received power to calibrated voltage or counts

A thermal detector may respond to absorbed power through a temperature rise and an electrical responsivity. A photon-counting detector measures absorbed photons, so its count rate contains a factor $\lambda/(hc)$ multiplying spectral power before integration. Two receivers with identical geometric acceptance and nominal quantum efficiency can respond differently to a changing source temperature if one output is proportional to energy and the other to photon number. Calibrate the actual electrical transfer and make the time interval explicit for accumulated counts.

For an ideal linear power detector let $V = V_0 + gP$, with offset V_0 and gain g. Two calibrated sources spanning the use range determine both. A black reference port at the same ambient temperature can also reject a common background in a chopped measurement. A filter, window or warm lens that absorbs thermal radiation also emits it. If these optics remain unchanged and stable, part of their own emission cancels in a reference difference. A reference traversing a different path generally leaves a residual input that is not captured by an electronic dark measurement.

Suppose effective source-minus-ambient powers of 2.00 and 8.00 microwatts yield 0.120 and 0.420 V. The fitted slope is 0.0500 V per microwatt with an offset of 0.020 V. A reading of 0.320 V implies 6.00 microwatts after offset correction. Substituting the same reading directly into a single-point gain with zero assumed offset would bias the inferred temperature. Rechecking a third reference near six microwatts tests linearity, while repeated zero-input readings at realistic instrument temperature test offset drift.

Invert the radiance and propagate uncertainties

A narrowband blackbody radiance measurement can be inverted analytically by solving Planck's expression for T. In a broad passband use numerical quadrature across the calibrated spectral response and solve a scalar monotone equation when emissivity and background are known. The derivative at a single wavelength is particularly informative for estimating precision. With $x = hc/(\lambda k_B T)$,

$$\frac{\partial \log B_\lambda}{\partial \log T} = \frac{x}{1 - e^{-x}},$$

$$T = \frac{hc/(\lambda k_B)}{\log[1 + 2hc^2/(\lambda^5 B_\lambda)]}.$$

At 10 micrometres and 400 K the logarithmic sensitivity is 3.70. A one-percent relative radiance uncertainty therefore gives about 1.08 K temperature uncertainty if all other quantities are exact. An unknown five-percent emissivity error can dominate that estimate. A two-colour temperature ratio may reduce uncertainty from an emissivity that is constant across the two narrow bands, but cannot cancel an arbitrary spectral emissivity or a large reflected ambient component.

As an example consider a grey object at 400 K and ambient 295 K with emissivity 0.70. At ten micrometres its outgoing radiance is $0.70(33.56) + 0.30(9.14) = 26.24$ in the earlier spectral units. Inverting as a perfect blackbody gives a brightness temperature of about 375 K. The difference is a physical radiance–temperature degeneracy. Repeat at known ambient temperatures while keeping the object thermally stabilised to constrain the reflected fraction. If heating the background changes the actual object temperature, a second independent temperature sensor or a transient thermal model is necessary.

When detector heat capacity sets the measured signal

Consider a thermal receiver of heat capacity C_d coupled to its surroundings with thermal conductance G_d. For small excursions its temperature obeys $C_d\dot{\theta} + G_d\theta = P_{\mathrm{abs}}(t)$, where θ is the excess above its operating point. A voltage thermometer with responsivity r_T converts this into output. Its frequency-dependent power responsivity is $r_T/(G_d - i\omega C_d)$. The geometric and spectral calculations determine absorbed optical power, while this thermal equation determines how much of its modulation appears at the electrical output.

Take $C_d = 2.0$ microjoules per kelvin and $G_d = 0.10$ milliwatt per kelvin. The thermal time constant is 20 ms and ordinary corner frequency is 7.96 Hz. A sinusoidal absorbed-power amplitude of 6.0

microwatts would give a static temperature rise of 0.060 K. At 20 Hz the amplitude is reduced by $1/\sqrt{1+(2\pi f\tau)^2}=0.370$, leaving about 22.2 mK with a phase lag magnitude near 68 degrees. A receiver calibrated with a steady optical power is not automatically calibrated for this chopped experiment.

A square-wave chopper also changes the relation between time-domain peak differences and a lock-in's fundamental reading. For equal hot and reference dwell times, a square wave switching between powers P_1, P_0 has fundamental sinusoidal peak amplitude $2(P_1-P_0)/\pi$, with different factors for rms and peak reports. The thermal transfer attenuates this component at the chopper frequency and the odd harmonics at their respective frequencies. Substituting the full hot-minus-reference power directly into the first-harmonic response would overestimate the expected signal.

Measure a frequency sweep on a stable reference and test whether amplitude and phase share a single time constant. A geometrical error rescales the amplitude but ordinarily leaves this thermal phase unchanged. A second thermal mass or an electronic filter can introduce another pole. Lowering chop frequency increases thermal signal until the static limit but can expose low-frequency environmental drift. Choose the frequency by comparing the calibrated response with the measured noise spectrum.

Finally, a reference and an unknown source must illuminate the same active area if responsivity varies across the receiver. Imaging one source as a small spot and the other as a filled patch can create a calibration bias even when total spectral powers agree. Translating a reference spot across the detector or imaging the same aperture successively at both temperatures tests that spatial uniformity assumption.

If the sensor is used on a source much smaller than the calibrated reference aperture, record an image or a field-stop scan to determine its filling fraction. Otherwise even exact knowledge of filter transmission and thermodynamic temperature leaves an unknown geometrical scale factor. Aperture filling is part of the measurement equation, not an electronic gain setting.

Tutorial · Infer emissivity using two controlled backgrounds

An opaque diffuse sample at fixed temperature is observed at 10 micrometres against two isotropic backgrounds with calibrated spectral radiances 8.00 and 16.00 W per square metre per steradian per micrometre. The resulting outgoing readings are 25.00 and 27.00 in the same units. Infer emissivity and the sample's blackbody-equivalent radiance. Estimate the significance of independent reading errors of 0.10 units.

▼ **Separate the reflected and emitted contributions**

The increased outgoing radiance is 2.00 for an 8.00 increase of background. Thus reflectivity is 0.250 and emissivity 0.750. The first measurement gives intrinsic blackbody radiance $[25.00-0.25(8.00)]/0.75=30.67$. The second returns the same value. With 0.10 independent uncertainty on each outgoing reading and exact backgrounds, the slope uncertainty is $\sqrt{2}(0.10)/8=0.0177$, as is the absolute emissivity uncertainty. Larger background separation can improve this slope estimate if sample temperature stays fixed.

If the sample transmits radiation, the residual fraction does not all correspond to reflection. A controlled change behind the sample helps detect that additional path. A directional specular reflector exposed to a warm enclosure also does not necessarily see the same angular average used in the diffuse model. Without these checks, a formally precise fitted emissivity can describe the experimental surroundings instead of the material.

Tutorial · Design a numerical spectral-integration check

An engineer tabulates spectral radiance per micrometre on a wavelength mesh. Another rewrites the detector model on an equally spaced frequency mesh, obtaining a signal that differs by several percent after independent mesh refinement. Explain what must transform, what stays dimensionless and how to compare a photon-counting channel with a power channel.

▼ **Transform the complete measurement integral**

The differential power $B_\lambda \, d\lambda$ must equal $B_\nu \, |d\nu|$, with $|d\lambda/d\nu| = c/\nu^2$. A dimensionless transmission curve becomes the same physical transmission at $\lambda = c/\nu$; it acquires no additional Jacobian. The integration bounds reverse when increasing wavelength is replaced by increasing frequency. A response specified per unit wavelength, or a calibrated gain integrating a particular finite bin, must instead be interpreted from how it was defined.

Evaluate a unit-transmission test using a very narrow band, then the full 8–12 micrometre band whose radiance values are given above. Refine each grid independently, treating sharp filter edges consistently and converting per-micrometre densities before using metres. For a photon counter divide each spectral power contribution by photon energy before summation. Replacing that energy by a single central-wavelength value is an approximation that should be checked against the full band and source temperature. Close agreement between two different meshes using the same incorrect density is not a valid units check.

NIST radiance-temperature calibrations · spectral radiance, sources and uncertainty

Master equations, detailed balance and probability currents

Statistical mechanics from microscopic counting to material response

A two-state detector can switch at random while displaying reproducible occupation fractions and relaxation times. A three-state detector can maintain exactly constant occupancies while carrying a persistent directed current. Derive the probability dynamics, infer transition rates from timed records and decide what a stationary histogram can and cannot reveal about equilibrium.

From short-time transition probabilities to a generator

Take discrete states i resolved by an instrument. In a short time dt, conditional on current state i, let the probability of one transition $i \to j$ be $w_{ij}dt + o(dt)$. The probability of remaining in i is one minus the sum of departures to this order. Assume time-independent rates and the Markov property, meaning the conditional law of future transitions is determined by the current state. Sum gains and losses to derive

$$\dot{p}_i = \sum_{j \neq i} (p_j w_{ji} - p_i w_{ij}),$$

$$\sum_i \dot{p}_i = 0.$$

Each gain in one state is a loss in another, so total probability is conserved. In a column-vector convention $\dot{\mathbf{p}} = L\mathbf{p}$, off-diagonal entry $L_{ij} = w_{ji}$ and columns sum to zero. In a row-vector convention transpose the matrix. This is not cosmetic in a numerical calculation: using a matrix assembled for column probabilities while evolving rows can lose normalisation or give an incorrect stationary distribution.

For a given state the exit rate is

$$r_i = \sum_{j \neq i} w_{ij}.$$

Its survival probability obeys $\dot{S}_i = -r_i S_i$, so $S_i(t) = e^{-r_i t}$. Conditional on exiting, the destination has probability w_{ij}/r_i. This permits direct event-driven simulation: draw an exponential waiting time, choose a destination with these probabilities, advance the clock and repeat. Sampling one transition per arbitrary fixed time step generally distorts waiting-time statistics unless a controlled approximation or exact finite-time propagator is used.

Determine two rates from one occupancy and one time

For a two-state system name the rates $a = w_{01}$, $b = w_{10}$. Eliminating $p_0 = 1 - p_1$ gives

$$\dot{p}_1 = a - (a + b)p_1.$$

Its stationary value $p_* = a/(a+b)$ is independent of the common rate scale, and deviations decay with time constant $\tau = 1/(a+b)$. Thus one stationary fraction gives a ratio, one relaxation time gives a sum, and the two together identify both rates.

A location sensor records an upper-state fraction 0.30 and relaxation time 0.20 seconds. Then $a = 1.50$ and $b = 3.50$ per second. Mean waiting times are $1/a = 0.667$ s in the lower state and $1/b = 0.286$ s in the upper one. Neither equals the population relaxation time of 0.20 s. A transition rate inferred by taking the reciprocal of the average of these two waiting times estimates a different quantity.

Define the indicator $X = 1$ in state one and zero otherwise. Conditional probabilities from the same linear solution give

$$\langle [X(t) - p_*][X(0) - p_*] \rangle = p_*(1 - p_*)e^{-|t|/\tau}.$$

Consequently a normalised equilibrium autocorrelation determines τ without needing a deliberate preparation away from stationarity. A bandlimited detector may miss short visits, reducing counts and distorting waiting-time distributions even if the mean occupation looks plausible. Compare time sampling with both exit times as well as the relaxation time, or model the measured signal with an explicit observation process.

Detailed balance is stronger than stationarity

Let π_i denote a stationary distribution. The net probability current on an oriented edge is $J_{ij} = p_i w_{ij} - p_j w_{ji}$. Stationarity requires only that currents sum to zero at each vertex. Detailed balance requires $\pi_i w_{ij} = \pi_j w_{ji}$ on every edge individually. For a system in equilibrium with one heat bath, with states even under time reversal and energies E_i, thermal rates satisfy $w_{ij}/w_{ji} = e^{-\beta(E_j - E_i)}$ for nondegenerate states. If states collect multiple microstates, degeneracies or coarse-grained free energies enter their equilibrium weights.

For the previous two-state process at equilibrium with equal state degeneracies, $p_*/(1 - p_*) = 3/7$ implies

$$(E_1 - E_0)/(k_B T) = \log(7/3) = 0.847.$$

Multiplying both rates by ten preserves the energy inference and speeds relaxation tenfold. Thermodynamic weights therefore do not determine the kinetic barrier. A physical sample may change its rate scale with solvent viscosity or transition geometry without changing the equilibrium ratio.

A convenient test on a connected graph with reciprocal edges is the product of rate ratios around each closed cycle. With detailed balance all stationary factors telescope to one. Conversely, if the product is one on every independent loop, consistent state weights can be reconstructed from any reference state by integrating ratios along a path. Path independence makes these weights a detailed-balance stationary distribution after normalisation. A missing reverse edge or unobserved channel requires additional care; a detector that cannot resolve two routes can hide directed currents.

A steady nonequilibrium system with the same state histogram

Consider three equally named sites arranged in a ring, with clockwise rate k_+ and anticlockwise rate k_- on every link. Symmetry gives stationary occupation $1/3$ at each site for all positive rates. Each oriented

clockwise edge carries current $J = (k_+ - k_-)/3$. The cycle affinity in dimensionless units is $\mathcal{A} = \log[(k_+/k_-)^3]$. For $k_+ = 4$, $k_- = 1$ per second, each edge transports net probability one per second, and $\mathcal{A} = 3\log 4 = 4.159$. A perfectly flat histogram conceals the directed operation.

For Markov jumps with resolved reversible edges and their conventional time reversals, the total stochastic entropy production rate is

$$\dot{S}_{\mathrm{tot}} = \frac{k_B}{2}\sum_{ij}(p_i w_{ij} - p_j w_{ji}) \log \frac{p_i w_{ij}}{p_j w_{ji}} \geq 0.$$

Each term is nonnegative under the product $(x - y)\log(x/y)$; the factor of one half prevents counting both directions separately. For the stationary ring this gives $k_B J \mathcal{A} = 4.159 k_B$ per second. An actual heat dissipation rate needs a physical local detailed-balance model identifying environments, work and entropy flows. A collection of counted transitions without that calibration can still yield evidence for broken time-reversal symmetry at its level of resolution.

Stationary Shannon entropy $-k_B \sum p_i \log p_i$ is constant in this ring. The positive total entropy production is compatible with that constancy because an environment sustains the directed current and carries compensating entropy. Equating production with just the derivative of stationary state entropy would predict zero for every driven stationary process.

Why a distribution relaxes toward a stationary law

For a finite irreducible time-independent chain let $\pi_i > 0$ be its unique stationary distribution. The relative entropy $D(p\|\pi) = \sum p_i \log(p_i/\pi_i)$ is nonnegative. To see its contraction over a finite observation interval, write the propagated probability as a stochastic mixture of earlier probabilities and apply the log-sum inequality to each output state. Summing gives

$$D(Pp\|P\pi) \leq D(p\|\pi).$$

Since $P\pi = \pi$, the relative entropy to stationarity cannot grow.

For a detailed-balance chain, substituting the master equation into this derivative gives $-k_B dD/dt = \dot{S}_{\mathrm{tot}}$ in the simple closed heat-bath model. For a driven chain the system may remain at π with constant relative entropy while producing positive steady entropy. Relaxation toward stationarity and dissipation required to maintain that stationary state are distinct measurable quantities.

Compare relaxation inferred from a time trace with the thermodynamic irreversibility experiment

Infer generator entries from residence times

An ideal continuously observed Markov record contains successive states and dwell lengths. Let T_i be accumulated time in state i and N_{ij} the number of resolved $i \to j$ jumps. Its log likelihood contains $\sum_{ij} N_{ij} \log w_{ij} - \sum_i T_i \sum_{j \neq i} w_{ij}$. Differentiation gives

$$\widehat{w}_{ij} = N_{ij}/T_i.$$

Dividing by total experiment time instead estimates the gross stationary flux $\pi_i w_{ij}$, not a conditional transition rate.

For a nearly stationary two-state record with $T_0 = 70$ s, $T_1 = 30$ s and 105 transitions in each direction the estimates are 1.50 and 3.50 per second. Counting uncertainty for each reasonably large Poisson-like count is roughly the inverse square root of the count in relative terms; precise interval estimates condition correctly on exposures and the observation process. Equality of opposite transition counts in a long two-state record is essentially forced by alternation. It cannot by itself establish a thermal energy gap or exclude multiple hidden driven pathways.

Turn a barrier diagram into two independently measurable rates

A thermally activated transition between two metastable configurations often has a forward rate $a = \nu_A e^{-(G^\ddagger - G_A)/(k_B T)}$ and reverse rate $b = \nu_B e^{-(G^\ddagger - G_B)/(k_B T)}$, where $G^\ddagger$ denotes a dividing-surface barrier under a chosen model. If the same transition-state construction applies in both directions with consistent partition factors, their ratio reproduces the equilibrium free-energy difference. The overall barrier controls how rapidly equilibrium is approached. Changing a barrier while leaving both wells unchanged can alter both kinetics by many orders of magnitude without changing stationary occupancy.

As an illustration let the higher-energy state cost $2k_B T$, and put a transition configuration $10k_B T$ above the lower state. With a common attempt rate of 10^6 per second, the upward rate is 45.4 per second and the downward rate 335 per second. The stationary upper fraction is approximately 0.119 and the relaxation time 2.63 ms. Lower the barrier by $2k_B T$ without changing the two state energies. Both rates increase by $e^2 = 7.39$, while occupancy is unchanged. A detector recording only the equilibrium level histogram could miss the intervention entirely.

For a temperature scan, a straight line in the log rate against inverse temperature is evidence for an approximately constant activation energy and prefactor over that interval. A free-energy barrier includes entropic as well as enthalpic contributions; its numerical value can vary with temperature even if the microscopic potential does not. A change of dominant pathway can bend the apparent Arrhenius plot. Measure forward and reverse rates independently, then compare their ratio with an independently determined equilibrium population. A barrier fit explaining one rate but violating the equilibrium ratio usually indicates inconsistent state definitions, missed short visits or an unmodelled pathway.

A useful control is a weak static bias that lowers one well relative to the other. The equilibrium log-odds shift measures the difference between the state couplings to the bias. The individual log-rate shifts additionally report the related dividing-surface response. If the state-occupation ratio moves as expected but only one rate changes appreciably, the effective transition configuration may be much closer in bias response to the other state. Such statements depend on the activated single-path interpretation; the response slopes themselves are more directly observed than an inferred barrier shape.

Use uniformisation to propagate a generator without negative probabilities

The formal finite-time propagator of a time-independent jump process is $P(t) = e^{Lt}$. A forward Euler update $p(t + \Delta t) = (I + L\Delta t)p(t)$ conserves total probability if the generator is assembled correctly, but a large step can yield negative diagonal probabilities. An exact algebraic representation starts with a rate q at least as large as every total exit rate. Then $K = I + L/q$ is a stochastic matrix in the same column convention. Write $L = q(K - I)$; since K and the identity commute,

$$e^{Lt} = e^{-qt} \sum_{n=0}^{\infty} \frac{(qt)^n}{n!} K^n.$$

This can be read as drawing a Poisson number of opportunities at rate q, followed by n transitions of K. At a lightly connected state some opportunities become self transitions. The construction leaves real waiting-time statistics correct while eliminating negative transition probabilities. Its truncation error in total probability is bounded by the omitted Poisson tail because every matrix power is stochastic.

For a two-state model with rates two and six per second choose $q = 6$ per second. Starting in the lower state, the first opportunity after some waiting time moves upward with probability one third and otherwise does nothing. In the upper state the next opportunity moves downward with probability one. At 0.25 s the Poisson mean is 1.5. Retaining terms through order eight leaves a tail below 3×10^{-5}. Compare the resulting upper probability against $(2/8)[1 - e^{-8(0.25)}] = 0.2162$. A small discrepancy due solely to truncation should shrink with the computable tail probability.

The representation also suggests a measured-data sanity check. A fitted generator with negative off-diagonal rates cannot yield a valid K at any positive q. A snapshot transition matrix can itself be nonnegative and stochastic but need not arise from a time-homogeneous continuous-time generator. In the two-state case its nontrivial eigenvalue must be $e^{-(a+b)t} > 0$. A measured negative eigenvalue rules out that simple generator model under perfect sampling, though finite counts and hidden periodic driving can distort empirical estimates. Taking a matrix logarithm blindly may return complex or negative rates rather than a physical kinetic explanation.

Population relaxation and a first-arrival experiment answer different questions

A first-passage measurement stops its clock when a target is first reached, even if that target would subsequently be left many times. To calculate its mean, treat the target as absorbing. For any other state i, wait on average $1/r_i$, then add the mean remaining time from the next state weighted by its conditional transition probability. The backward equations are

$$-1 = \sum_j w_{ij}(t_j - t_i),$$

with $t_{\text{target}} = 0$. Their solution depends on possible returns before the first successful passage. It generally differs from an inverse equilibrium occupation or from a slowest relaxation time.

Consider a three-site linear chain $0 \leftrightarrow 1 \leftrightarrow 2$ with equal rate k on each adjacent directed edge, starting at zero and stopping at two. From zero one must reach one after mean $1/k$. From one the mean wait is $1/(2k)$, then there is probability one half of returning to zero and one half of finishing. Solving $t_0 = 1/k + t_1$, $t_1 = 1/(2k) + t_0/2$ gives

$$t_0 = 3/k.$$

Ignoring returns would yield $1.5/k$, exactly half the correct first-passage mean.

The unmodified closed chain has equal stationary weights, but that fact alone says nothing about the numerical value of $3/k$. Speeding every link by ten leaves occupancies unchanged and reduces passage times tenfold. A detector that registers only visits longer than a threshold makes a brief first arrival invisible and can report even longer apparent passage times. Calibrate detector dead time and distinguish entrance to the target from persistence within it.

For a larger network, solve the sparse backward linear system and compare against many event-driven paths. Check boundary initial conditions by starting simulations exactly at the absorbing target: the first-passage time must then be zero under that convention. Some experiments instead ask for a first return after leaving, which requires a different initial condition. A zero-lag event in one definition is a physically different measurement in the other.

Estimate a cycle affinity with finite counting statistics

A clockwise ring trajectory contains gross counts in both directions on each edge. Estimating each conditional rate as count divided by origin-state exposure, the product of clockwise-to-anticlockwise rates around a complete closed loop cancels the exposures. The estimated cycle affinity is the sum of log forward counts minus log reverse counts, assuming every required count is nonzero and rates were stationary during the record. This is a useful cross-check independent of how the total time was partitioned among states.

If all three edges register roughly 120 clockwise and 80 anticlockwise events, the estimated affinity is

$$3 \log(1.5) = 1.216.$$

For large independent Poisson-like gross counts the approximate variance of this log product is $3(1/120 + 1/80) = 0.0625$, giving standard error 0.25 as a first estimate. Actual path constraints couple counts; uncertainty can be assessed from independent records or a complete stochastic likelihood. The point estimate alone does not tell whether an observed directionality is statistically distinguishable from zero.

One missing reverse event makes the unregularised logarithm infinite, but a finite record cannot establish an infinite physical affinity. Report a confidence bound based on the exposure and upper limit on the unseen reverse rate. Extend the acquisition or change sampling speed to exclude missed short-lived reverse states. Numerical addition of an arbitrary count can help regularise a plotting routine but is not a substitute for a specified prior or measurement model when claiming physical entropy production.

A driven occupation loop can carry charge without changing concentrations

Place three sites in a cycle and drive their transition rates by two controls, for example energies and barriers. Even if each instantaneous frozen generator satisfies detailed balance with a well-defined equilibrium distribution, the physical occupations need not equal that moving equilibrium distribution at each moment. A slow lag can produce a small edge current at every instant and a finite integrated current over a cycle. The current through an edge is time-integrated transfer; it can be nonzero while all occupations return to their initial values after a complete drive period.

A simple piecewise illustration avoids subtleties of adiabatic geometric pumping. In stage one, connect only sites zero and one while strongly favouring one. In stage two connect only one and two while favouring two. In stage three connect only two and zero while favouring zero. Begin with nearly all probability at zero and wait long enough for each enabled pair to relax. Almost the entire population travels once around the loop per cycle. No site accumulates probability over repeated complete cycles, but each edge registers about one forward crossing net per cycle. External control changes the energy and barrier landscape and supplies the required work.

If each programmed transfer succeeds with probability 0.95, it is tempting to announce a current of 0.95^3 per cycle. This is the probability of one fully successful trajectory starting on the prescribed site in a simplified no-return preparation, not generally the long-run edge current of a periodically driven ensemble. Failed transfers leave probability on other sites in later stages. The correct periodic distribution is the unit-eigenvalue vector of the product of the three finite-time propagators. Compute edge currents during each stage using its evolving occupations and integrate, then compare a complete cycle after transients have decayed.

This construction supplies a statistical check on experimental counting. Over a complete cycle the integrated net inflow at a site equals its change in occupancy probability. In a periodically stationary state that change is zero, so all three consistently oriented integrated edge currents must agree. A disagreement between measured edges can reveal a missed transition channel, a start-up transient or incorrect gating. Agreement does not prove absence of drive or detailed balance. A torque-free stationary ring and a three-stage external pump can share nearly the same time-averaged site histogram yet differ clearly in time-resolved edge currents.

Finally compare the work and heat bookkeeping with the master-equation entropy-production expression. During an external change of a site's energy without a transition, the state's energy changes by work. During a transition at frozen controls, assigning heat requires the energy difference and the relevant bath. Lumping an entire driven cycle into a single time-averaged generator can conceal those distinctions and underestimate dissipation.

The ordered protocol can also be run backwards. If the three stages are reversed while preparations are equilibrated in each corresponding periodic state, the preferential current direction should reverse in the ideal transfer picture. A fixed instrumental preference for counting neighbouring sites in one direction would not generally reverse. This distinguishes controlled transport from a direction-dependent detection efficiency. Do not confuse this observed current reversal with exact microscopic time reversal of each individual stochastic trajectory: the reversed experiment has its own ensemble of random jumps. Comparing path probabilities, rather than pairing individual paths as though their noise histories were

reproducible, is the appropriate starting point for entropy-production and fluctuation-relation tests. The distinction becomes particularly visible when a stage transfers almost all the probability in one direction while a few rare trajectories travel against the applied protocol.

Distinguish stationarity from detailed balance

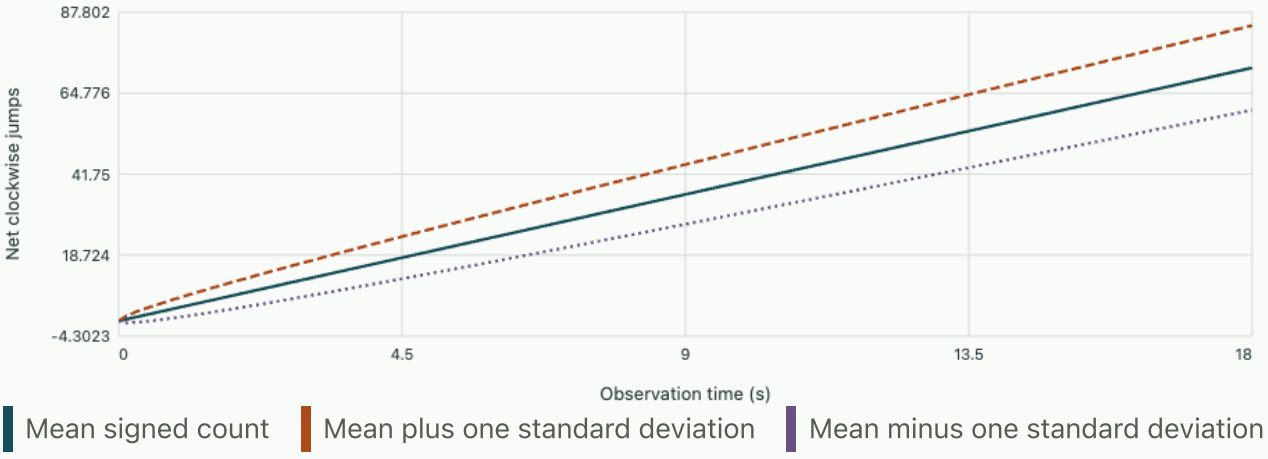
Keep the three populations identical while varying the directed transition rates. Identify the observable that reveals time-reversal asymmetry.

Open this experiment online

Try the calculation

Clockwise bias **0.5**

Total jump rate (s⁻¹) **8**



Every stationary population is 1/3. Net clockwise edge current 1.3333 s⁻¹. Entropy production 4.3944 kB/s.

Tutorial · Recover a generator from snapshots rather than full events

At sampling interval 0.10 s, a two-state record yields $P_{01} = 0.12$ and $P_{10} = 0.28$. Infer the continuous-time rates, and compare with directly dividing each transition probability by 0.10 s.

▼ **Use the exact finite-interval propagator**

The exact transition probabilities are $P_{01} = p_*[1 - e^{-(a+b)\Delta t}]$ and $P_{10} = (1 - p_*)[1 - e^{-(a+b)\Delta t}]$. Their sum is 0.40, so $a + b = -\log(0.60)/0.10 = 5.108$ per second. The stationary upper fraction is

$$0.12/0.40 = 0.30.$$

Thus $a = 1.532$ and $b = 3.576$ per second. Dividing by the interval gives only 1.2 and 2.8 because return transitions inside a sampling interval can erase evidence of the intermediate visit.

Tutorial · Can measured rates arise from one equilibrium potential?

For a three-state triangle suppose clockwise to counterclockwise conditional rate ratios are 2, 3 and 0.25 around the three distinct edges. Could these be equilibrium weights in a common bath? If a single uncertain ratio were corrected, what would it have to be on the third link?

▼ **Apply the closed-loop consistency condition**

The product is $2 \cdot 3 \cdot 0.25 = 1.5$, giving nonzero loop affinity $\log 1.5 = 0.405$. No consistent detailed-balance set of three state weights exists for these exact ratios. The third ratio would have to be $1/6$, giving a telescoping unit product. If counts are noisy, propagate uncertainty in the log ratios or compare a constrained detailed-balance likelihood against an unconstrained model. A value barely inconsistent with unity within one sparse record is not enough to establish driving.

Tutorial · Detect a hidden state with the same observed level

A detector reports two visible levels, but its lower level combines two unresolved substates. The measured lower-state survival probability is $0.5e^{-t} + 0.5e^{-9t}$ with time in seconds. Calculate the initial hazard and its long-time limit. Can a two-state Markov fit reproduce both?

▼ **Differentiate survival before interpreting the tail**

The hazard is minus the derivative of $\log S$. At zero time it is the initial weighted mean of rates, 5 per second. At long times the rapidly departing component has almost disappeared and the hazard tends to 1 per second. A single Markov exit rate would give a constant hazard for all dwell ages. The observed model cannot reproduce both limits with one rate.

A changing apparent hazard may also result from limited detector temporal resolution or variable preparation of the substates. Repeat with shorter acquisition intervals and condition the dwell distribution on the preceding visible transition or an independent preparation marker. A fitted sum of exponentials identifies a need for a richer observation or state model; it does not uniquely identify microscopic structures absent in the measurements.

Cambridge Kinetic Theory · stochastic processes and approach to equilibrium

From binary collisions to the Boltzmann equation

Statistical mechanics from microscopic counting to material response

An initially anisotropic velocity distribution can relax even though every individual collision conserves energy and momentum. Describe a dilute gas with a phase-space density, derive the structure of its collision term and prove which distribution can stop producing entropy. Then examine the assumptions behind that proof and the measurements most likely to reveal their failure.

What the one-particle function counts

Let $f(\mathbf{r}, \mathbf{v}, t) d^3r\, d^3v$ be the expected number of particles in a small phase-space cell. Integrating over velocities gives number density n, and the mean velocity is $\mathbf{u} = n^{-1} \int \mathbf{v} f \, d^3v$. Define the peculiar velocity $\mathbf{c} = \mathbf{v} - \mathbf{u}$; its second moment sets local kinetic energy after removing the organised flow. A gas can have large bulk speed but low temperature.

Follow a particle ballistically over a short interval under acceleration $\mathbf{a} = \mathbf{F}/m$. With no collisions, phase-space probability is advected along its trajectories. For a standard position-dependent force, phase-space flow has zero divergence, giving the streaming operator $\partial_t f + \mathbf{v} \cdot \nabla_r f + \mathbf{a} \cdot \nabla_v f$. Magnetic Lorentz acceleration also has zero velocity-space divergence. A dissipative velocity-dependent drag generally requires the divergence form $\nabla_v \cdot (\mathbf{a} f)$ instead. It should not silently be inserted into the Hamiltonian streaming expression.

We now assume a dilute classical gas of identical particles with short-ranged binary elastic collisions, approximated as local on the spatial scale resolved by f. Different particles approaching collision are assumed uncorrelated apart from their individual one-particle distributions at that position. This molecular-chaos approximation is an assumption about incoming pairs, not a claim that the outgoing particles remain uncorrelated. It is the step that allows the evolution to close on a one-particle function.

Count pairs crossing a scattering area

Consider a tagged incident velocity $\mathbf{v}$ and a possible partner $\mathbf{v}_1$. Their relative speed $g = |\mathbf{v} - \mathbf{v}_1|$ converts a differential scattering cross section $d\sigma/d\Omega$ into the swept interaction volume per unit time. A binary collision rotates the relative velocity in the centre-of-mass frame while preserving its magnitude. With equal masses, set $\mathbf{V} = (\mathbf{v} + \mathbf{v}_1)/2$ and $\mathbf{g} = \mathbf{v} - \mathbf{v}_1$. The partner and tagged postcollision velocities are $\mathbf{V} \pm \mathbf{g}'/2$, where $|\mathbf{g}'| = |\mathbf{g}|$. This parametrisation enforces momentum and kinetic energy conservation before choosing a scattering direction.

The loss term counts incoming pairs with velocities $\mathbf{v}, \mathbf{v}_1$. The gain counts inverse collisions that produce the specified velocity. Microreversibility of elastic scattering and invariance of the integration measure allow the compact expression

$$C[f](\mathbf{v}) = \int d^3v_1 \, d\Omega \, g \frac{d\sigma}{d\Omega} \left[f(\mathbf{v}') f(\mathbf{v}_1') - f(\mathbf{v}) f(\mathbf{v}_1) \right].$$

Both products refer to the same position and time. A change in which pair is called primed merely swaps gain and loss if one reverses the parametrisation consistently. This formula uses a consistent distinct-

partner convention for the full scattering measure. When counting the total number of collision events per volume rather than the loss of tagged particles, divide the sum over all ordered pairs by two. Confusing those two observables gives an erroneous collision frequency.

For hard spheres of molecular diameter d, the total geometric cross section is πd^2. In an equilibrium Maxwellian the mean relative speed is $\sqrt{2}$ times the mean single speed for identical masses. The mean free path under the usual hard-sphere convention is $\ell = 1/(\sqrt{2} n \pi d^2)$. For $d = 0.36$ nm, $n = 2.4 \times 10^{25}$ per cubic metre and molecular mass 28 unified atomic units at 300 K, $\ell = 72.4$ nm and $\bar{v} = 476$ m/s. Their ratio gives a characteristic flight time 0.152 ns.

Take a weak moment of the collision operator

For a test function $\psi(\mathbf{v})$, multiply the collision term by ψ and integrate over tagged velocity. Exchange the identities of both particles and then exchange incoming and outgoing velocities in the gain integral. The result expresses the change of the pair's total test-function value:

$$\int \psi C[f] \, d^3v = \frac{1}{2} \int d^3v \, d^3v_1 \, d\Omega \, g \frac{d\sigma}{d\Omega} f f_1 [\psi' + \psi_1' - \psi - \psi_1].$$

For $\psi = 1$, $m\mathbf{v}$ and $mv^2/2$ the square bracket vanishes collision by collision. Integrating the kinetic equation gives exact local conservation of mass, momentum and energy with appropriate streaming fluxes, regardless of how close the distribution is to Maxwellian. Conservation alone leaves a hierarchy: the momentum flux depends on the pressure tensor and the energy flux on a heat flux, both higher velocity moments.

A numerical collision solver should check all these invariants separately. Equal masses allow a direct test with velocities $(300, 100, 0)$ and $(-100, 100, 0)$ m/s. Their centre velocity is $(100, 100, 0)$ and relative velocity has magnitude 400 m/s. Rotating it from x to y gives outgoing velocities $(100, 300, 0)$ and $(100, -100, 0)$. Total velocity and total squared speed agree before and after, but the directional kinetic energy has changed. A field of such collisions can isotropise the pressure tensor while leaving total energy fixed.

Derive a sign-definite entropy production

Define the local kinetic H density as $\int f \log(f/f_\star) d^3v$ with a constant reference of matching units. The reference only changes a conserved number contribution. For a homogeneous isolated gas the collision contribution to its time derivative follows by applying the weak form with $\psi = \log f$. Average the result with its incoming–outgoing interchange. Setting $A = f f_1$, $B = f' f_1'$, one obtains

$$\frac{dH}{dt} = -\frac{1}{4} \int d^3v \, d^3v_1 \, d\Omega \, g \frac{d\sigma}{d\Omega} (A - B) \log(A/B) \le 0.$$

For positive A, B, logarithm and difference have the same sign; appropriate limits handle zero occupation. Thus the Boltzmann entropy $-k_B H$ does not decrease from these local binary collisions. In

an inhomogeneous region, advection additionally transports entropy through the boundaries. A local cell's kinetic entropy can drop because entropy leaves the cell, while collision production remains nonnegative.

Equality requires $ff_1 = f'f_1'$ for collisions connected by the scattering kernel. Equivalently $\log f + \log f_1$ is a collision invariant. For sufficiently rich elastic scattering in three dimensions the smooth additive collision invariants are linear combinations of particle number, momentum and energy. Therefore $\log f = a + \mathbf{b}\cdot\mathbf{v} + cv^2$, with $c < 0$ for normalisation. Completing the square and matching moments yields the local Maxwellian:

$$f_M = n\left(\frac{m}{2\pi k_B T}\right)^{3/2} \exp[-m|\mathbf{v}-\mathbf{u}|^2/(2k_B T)].$$

A Maxwellian at each location makes the collision term zero locally, but spatial gradients usually leave a nonzero streaming term. Such a distribution can evolve because neighbouring cells exchange particles. Conversely, a truly global equilibrium in an external conservative potential includes the appropriate spatial Boltzmann factor. In one-dimensional equal-mass elastic collisions particles merely exchange velocities, so the assumed richness of three-dimensional scattering fails; an arbitrary velocity histogram need not thermalise through those collisions.

Assess the kinetic and hydrodynamic descriptions separately

Define Knudsen number $Kn = \ell/L$ using the length over which the distribution or boundary condition changes. In the earlier gas $Kn = 7.2 \times 10^{-5}$ for a 1 mm chamber feature but 0.072 for a 1 micrometre channel gap. The same binary collision model can remain dilute in both, yet a no-slip bulk hydrodynamic closure may be much less accurate in the microchannel. A mean free path small relative to the overall device length is insufficient if a sharp boundary layer develops on a much shorter scale.

A useful relaxation-time collision approximation is $C_{\mathrm{BGK}} = -(f - f_M)/\tau$, with f_M computed from the local conserved moments of f. This enforces their conservation and models return to local Maxwellian without evaluating an angular integral. If an externally fixed Maxwellian is used irrespective of actual local density, momentum and energy, the approximation can remove these quantities spuriously. A single fitted relaxation time also need not reproduce both stress and heat transport for a real gas.

A particle simulation resolving binary encounters must choose spatial cells small enough to represent local variation, and event or collision timesteps short against relevant flight or collision times. Statistical noise decreases with independent sampled pairs; refining cells while keeping total simulated particle count fixed can make local collision estimates noisier. Compare conservation, relaxation of anisotropic moments and a transport observable under separate refinements of geometry and sample count. A correct mean pressure alone is a weak test of a collision operator.

Compare the flux-weighted velocities of escaping particles with the bulk velocity distribution

Take the next moment and exhibit the source of viscosity

Write momentum density as $\rho u_i = m \int v_i f \, d^3v$. The flux of i-momentum through a surface with normal in direction j is $m \int v_i v_j f \, d^3v = \rho u_i u_j + P_{ij}$, where $P_{ij} = m \int c_i c_j f \, d^3v$. The scalar pressure is one third of its trace in three dimensions. Multiplying the kinetic equation by $m v_i$ and using collision invariance yields the momentum balance with $\partial_j P_{ij}$ explicitly. No constitutive assumption has yet set the off-diagonal components to zero or made them proportional to strain rate.

For a local Maxwellian, the odd moments of peculiar velocity vanish and $P_{ij} = n k_B T \delta_{ij}$. Now consider slow shear with mean velocity $u_x(y)$, and approximate collisions by a relaxation time τ. In a local coordinate frame moving with the fluid write $f \simeq f_M - \tau v_y \partial_y f_M$ to first order in gradient. At constant density and temperature,

$$\partial_y f_M = (m c_x / k_B T) f_M \partial_y u_x .$$

The off-diagonal momentum flux correction is therefore a fourth Gaussian velocity moment:

$$P_{xy} = -\tau \frac{m^2}{k_B T} \partial_y u_x \int c_x^2 c_y^2 f_M d^3v = -n k_B T \tau \, \partial_y u_x .$$

The integral factorises into variances $k_B T / m$ in the two directions. Under the kinetic pressure-tensor convention the result has a minus sign, because random motion transports high x-momentum toward lower mean velocity. The viscous part of the Cauchy stress is commonly written with the opposite sign convention. State which tensor enters the balance before declaring a predicted shear stress negative. The dynamic viscosity of this simple model is $\eta = p \tau$.

If $p = 1.0 \times 10^5$ Pa and the fitted stress relaxation time is 1.8×10^{-10} s, this gives 1.8×10^{-5} Pa s. At a shear rate of 10^3 per second the dissipative stress magnitude is 0.018 Pa, tiny compared with equilibrium pressure and with dimensionless shear rate $\tau |\partial_y u_x| = 1.8 \times 10^{-7}$. A stress measurement must resolve a small anisotropic contribution on top of a large isotropic pressure. A pressure gauge alone cannot test this kinetic viscosity relation.

In a dilute hard-sphere gas the mean collision time scales inversely with density. The simultaneous pressure factor then makes dynamic viscosity approximately independent of density at fixed temperature in the dilute continuum regime. The kinematic viscosity, η / ρ, still changes strongly. A gas made sufficiently rarefied no longer supports ordinary no-slip continuum flow through a fixed channel, so an apparent pressure-dependent flow resistance can coexist with an essentially density-independent bulk dynamic viscosity.

Infer accommodation from a wall momentum measurement

The interior collision operator does not prescribe the boundary condition at a solid wall. Specular reflection reverses normal velocity while preserving tangential velocity in the wall frame. Diffuse thermal re-emission from a fully accommodating stationary wall instead draws outgoing velocities from a wall-temperature distribution appropriate to the outward flux and erases incident tangential drift. A real clean

or rough surface may lie between these limits. The boundary scattering model matters particularly when particles encounter the wall as often as other particles.

For a simple moving plate in a dilute gas at the same temperature, suppose a fraction α of incident particles accommodates tangentially to plate speed U, and the rest reflects specularly. To leading order in small tangential drift, accommodated particles carry away an average tangential momentum increment mU if incident gas has zero mean tangential speed. The incident number flux on one plate face is $n\bar{v}/4$. Thus the leading tangential drag magnitude on the plate is $\alpha m U n \bar{v}/4$ per area in this ideal free-molecular geometry.

Take $n = 10^{21}$ per cubic metre, molecular mass 4.65×10^{-26} kg, $\bar{v} = 480$ m/s, $U = 2.0$ m/s and $\alpha = 0.80$. The predicted shear is about 8.93×10^{-3} Pa per exposed face. The effect scales linearly with density here, unlike the dilute continuum viscosity at a fixed macroscopic velocity gradient. Both results derive from momentum transport, but the first describes particles crossing a wall-bounded collisionless region and the second local shear sustained by many intermolecular collisions. Compare them only after specifying the Knudsen number and wall geometry.

To test the accommodation model, reverse the plate motion to remove static force offsets, vary density across a regime that remains free molecular and change the exposed surface while keeping area fixed. A single measurement gives the combination of accommodation fraction, incident number flux and velocity difference. It cannot separate a changing incident gas drift from a changing material accommodation coefficient. Measuring the upstream velocity distribution or using opposed plates with known relative motion supplies additional constraints.

A discrete collision quartet demonstrates the H theorem numerically

A single allowed reversible collision connects two incoming velocity cells to two outgoing cells. Suppose cell-volume and scattering factors have been normalised so that its net flux is $K(A - B)$, with $A = f_1 f_2$, $B = f_3 f_4$. It removes particles from cells one and two and adds them to three and four when $A > B$. Conservation requires the representative cell velocities to satisfy the same pair momentum and energy constraints used for a physical collision. Assigning arbitrary four velocities may conserve the number of particles but will generally create energy.

Choose the velocities from the earlier rotated-pair example and occupancies $f_1 = 4, f_2 = 1, f_3 = 2, f_4 = 1$ in a common dimensionless reference. Then $A = 4$, $B = 2$ and the forward flux is $2K$. Its contribution to dH/dt is $-2K \log 2$, negative. The exact continuous rate tends to zero as the products equalise. A forward Euler step larger than $f_2/(2K)$ would make the second occupancy negative and destroy the logarithm's meaning. Nonnegative entropy production for the continuum collision model does not guarantee it for an arbitrary numerical time step.

An implicit or positivity-preserving update can avoid this failure, but still requires independent tests of conserved moments and of the correct equilibrium. For this isolated quartet the condition $f_1 f_2 = f_3 f_4$ alone leaves many possible stationary cell distributions. A network of collision quartets sufficiently connecting the velocity space is needed to enforce a Maxwellian-like equilibrium. On a coarse discrete velocity lattice too few permitted momentum- and energy-conserving quartets may leave artificial

invariants. A code can display exactly conserved energy and monotonically decreasing H while relaxing to the wrong distribution because its collision graph is incomplete.

A velocity-selective aperture tests collision equilibration

A thin aperture feeding a vacuum produces a flux-weighted distribution from a local Maxwellian reservoir. A detector placed downstream generally selects a finite angular cone and may respond differently to fast and slow particles. Consider two narrow speed bins with the same width in speed and the same angular acceptance. For an effusive Maxwellian their flux ratio contains a cube of speed multiplied by a Boltzmann exponential. A bulk sampler instead has a square of speed. Confusing these prefactors shifts a temperature inferred from ratios even if the gas truly is locally equilibrated.

For bins at speeds v and $2v$, and choosing $mv^2/(2k_BT) = 1$, the effusive high-to-low ratio is

$$8e^{-3} = 0.398.$$

The bulk number-density ratio at those speeds would be $4e^{-3} = 0.199$. A downstream detector measuring particles present in a fixed optical observation volume may weight by travel time, introducing an additional approximate inverse speed that moves the response closer to density weighting. Distinguish arrivals counted per time from particles counted during an exposure.

Changing aperture diameter is an independent control. A broad opening or high upstream density can produce collisional expansion and a directed jet, altering both angular and speed distributions. Measure pressure and beam shape across aperture sizes and compare the small-aperture limit before attributing a non-Maxwellian downstream spectrum to a failure of local thermal equilibrium throughout the reservoir.

A flight-time measurement tests free paths and speed selection separately

Imagine a dilute molecular tracer introduced into a uniform bath. For a deliberately simplified constant collision-rate model with rate r, probability of no collision by time t is e^{-rt}. For tracers conditioned to have one speed v, the distribution of collision-free distance is then exponential with mean v/r. If the collision rate itself depends on speed through the average relative velocity against the bath, the conditional mean free path is $v/r(v)$. Mixing all initial speeds need not produce a single exponential in distance, even if every conditional waiting time is exponential.

Place a detector a distance L downstream that accepts only trajectories maintaining their initial direction. For a narrow injected speed class, the unscattered fraction is approximately $e^{-r(v)L/v}$ if any collision removes it from the acceptance. For a class at 500 m/s and mean free path 0.20 m, a detector at 0.10 m receives an unscattered fraction 0.607; at 0.30 m it receives 0.223. A ratio of the two readings removes the unknown injection strength and recovers the attenuation coefficient if detector acceptances and geometric factors match.

If forward-scattered trajectories remain inside the acceptance, the apparent attenuation is smaller than the total collision rate. Angularly narrowing the receiver should approach the total-removal model for

suitable scattering. A finite source pulse and finite time-of-flight bin also select velocities even if the source initially emitted a Maxwellian. Fit survival separately in narrow speed bins first. A nonexponential whole-beam attenuation then need not imply that the microscopic collision process has memory; it can arise purely from an unresolved distribution of rates.

A simulation can distinguish these alternatives by logging every collision event and comparing the full collision count with the detector-defined removed count. First test conditional survival for a prepared monochromatic beam, then add a known distribution of incident speeds and calculate the mixture prediction explicitly. Finally include a finite angular acceptance and integrate the differential cross section over the appropriate retained directions. These steps separate the kinetic collision operator, the initial distribution and the observation rule.

The same logic applies to a relaxation experiment reporting a single characteristic mean free time. The decay time of a directional flux depends on how much momentum is deflected per collision, not simply on how often particles meet. Strongly forward-peaked scattering can have a large total cross section and a much smaller momentum-transfer cross section weighted by $1-\cos\theta$. A measured attenuation and a measured viscosity may consequently constrain different moments of precisely the same differential scattering process.

For example, if every collision deflects a test particle through a small fixed angle of ten degrees in the relevant frame, the momentum-transfer weighting is

$$1-\cos 10^{\circ}=0.0152.$$

The corresponding transport cross section can then be far smaller than the total cross section even when collisions are frequent. A velocity-randomising approximation that uses the total collision time as its stress-relaxation time would predict much faster momentum relaxation than this angular picture supports. Actual stress and energy transport involve additional angular weights and velocity averages, so this numerical comparison is a scale diagnostic rather than an exact viscosity formula. Measure several angular moments of the differential scattering kernel or compare at least two independently sensitive transport observables before assigning a universal relaxation time to a real collisional gas.

Follow an H functional during a reversible elementary collision

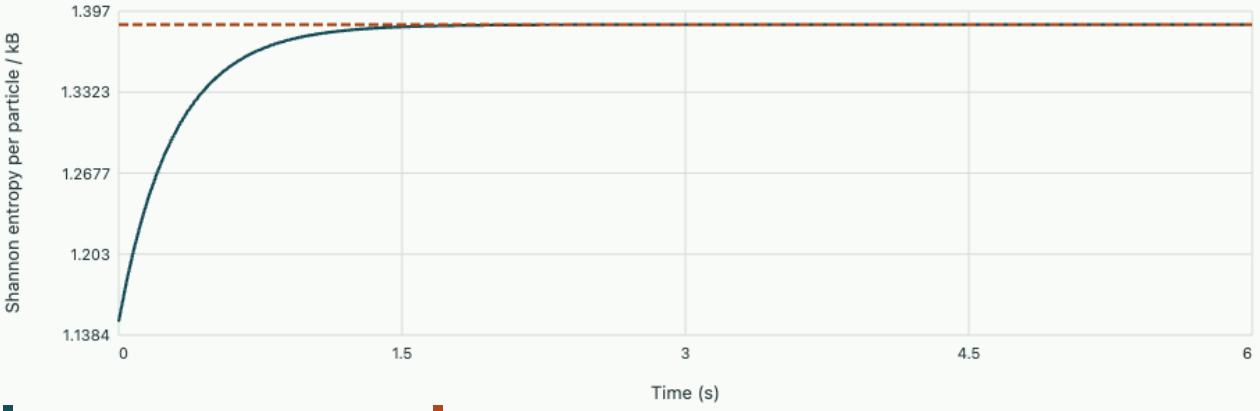
Prepare unequal populations in two velocity pairs. Count forward and reverse collision rates and compare the sign of entropy change with the population imbalance.

Open this experiment online

Try the calculation

Initial fraction in velocity pair A **0.83**

Collision rate coefficient (s⁻¹) **0.8**



Initial clockwise-minus-reverse pair imbalance x²−(1−x)² = 0.66. Initial d(S/kB)/dt = 0.83721 s⁻¹.

Tutorial · Find the equilibrium temperature of an anisotropic preparation

A homogeneous gas has zero mean flow and component kinetic temperatures 600, 300 and 150 K, defined by $m\langle v_i^2\rangle = k_B T_i$. In an isolated elastic gas, predict its equilibrium temperature and pressure tensor change. Could a thermostat holding the velocity variance in one direction at its original value permit the same inference?

▼ **Retain all three conserved energy contributions**

Conservation of total kinetic energy implies $T_f = (600 + 300 + 150)/3 = 350$ K. Initially the diagonal kinetic pressure components are nk_BT_i; ultimately each becomes $nk_B(350\,\mathrm{K})$. Their sum is unchanged. A component singled out by a thermostat can exchange work or heat with its environment. Its final common temperature, if one exists, must be obtained from the full energy balance including that exchange. Isotropisation cannot simply be assumed to retain the initial isolated mean.

Tutorial · Estimate the error in a collision-frequency inference

A detector reports a mean speed of 480 m/s and a geometric collision cross section 4.0×10^{-19} square metres at $n = 2.0 \times 10^{25}$ per cubic metre. Estimate the collision rate per tagged particle and total events per unit volume per second for an equilibrium identical-particle gas. What changes in a molecular beam with almost parallel velocities?

▼ **Use the relative rather than the single-particle speed**

The equilibrium mean relative speed is $\sqrt{2}(480) = 679$ m/s. The rate per tagged particle is $ng\sigma \simeq 5.43 \times 10^{9}$ per second. Multiplying by $n/2$ gives event rate 5.43×10^{34} per cubic metre per second. In a nearly monoenergetic collimated beam the relative speed can be much less than the mean laboratory speed. Evaluate its pair-distribution average explicitly; translating every velocity by a common vector leaves all collision rates unchanged at fixed distribution of relative velocities.

Tutorial · Does reversing all velocities contradict the H theorem?

In a dilute-gas simulation, reverse all particle velocities at a chosen instant with exact precision. An ideal reversible microscopic integrator would retrace the prior trajectory. Explain how this outcome can coexist with a kinetic equation predicting nondecreasing entropy, and propose a comparison that isolates the missing information.

▼ **State which distributions are controlled by the reversal**

The reversed incoming pairs carry precise correlations inherited from earlier outgoing collisions. Their two-particle distribution generally does not factor into the product of their one-particle distributions just before future collisions. The molecular-chaos step used to derive the collision operator therefore fails for this deliberately prepared history. Microscopic Hamiltonian evolution and the coarse one-particle closure make different state specifications.

Prepare a second gas with the same reversed one-particle velocity histogram but with spatial and partner correlations randomised while conserving total momentum and energy. Compare the two subsequent anisotropy histories. The kinetic equation has the same initial one-particle input for both, while a microscopic simulation can distinguish them through higher-order correlations. Imperfect numerical reversal introduces a separate error; test exactness and time-step dependence before interpreting a failure to retrace as a fundamental thermodynamic observation.

Cambridge Kinetic Theory · the Boltzmann equation, collision invariants and hydrodynamics

Transport coefficients from equilibrium fluctuations

Statistical mechanics from microscopic counting to material response

How can a fluid in equilibrium reveal the diffusion constant that would govern a spreading drop, or the viscosity that would resist an applied shear? Derive an exact connection between displacements and velocity correlations, generalise it to stress and heat currents and test how a finite record can bias each number. The assumptions behind convergence matter as much as the prefactor.

Integrate the velocity correlation before taking a long-time limit

Observe one Cartesian coordinate of a tagged particle in a homogeneous equilibrium fluid with zero mean drift. The displacement over time t is the integral of its velocity. Square that integral, average over the equilibrium ensemble and use stationarity to replace the two-time expectation by $C_v(t_1 - t_2)$. For a scalar component in a time-reversal-symmetric equilibrium state this correlation is even. Integrate over the two triangles of the time square:

$$\langle [x(t) - x(0)]^2 \rangle = 2 \int_0^t (t - s) C_v(s)\, ds,$$

$$\frac{1}{2} \frac{d}{dt} \langle \Delta x^2(t) \rangle = \int_0^t C_v(s)\, ds.$$

If the correlation is integrable and boundary effects remain absent, the long-time slope yields a diffusion coefficient $D = \int_0^\infty C_v(s) ds$. In d isotropic dimensions one may use $D = (1/d) \int \langle \mathbf{v}(s) \cdot \mathbf{v}(0) \rangle ds$ or the mean-square-displacement slope divided by $2d$. Using the full vector dot product without division by dimension gives an overestimate. Directional diffusion in an anisotropic material requires tensor components rather than a forced scalar average.

The derivation is kinematic and does not assume a Markovian motion. A fluid with velocity memory or an oscillating solid still obeys the finite-time relation. Those systems need not exhibit the same long-time proportionality to time. A bound oscillator can have persistent or alternating velocity correlations and a bounded displacement. An observation window in which the mean square grows over a fraction of an oscillation is insufficient to claim diffusion.

A solvable benchmark from ballistic to diffusive motion

For one coordinate take the underdamped Langevin equation $m\dot{v} = -\zeta v + \xi$, with zero-mean white noise satisfying $\langle \xi(t) \xi(t') \rangle = 2 k_B T \zeta\, \delta(t - t')$. In equilibrium $\langle v^2 \rangle = k_B T / m$. Condition on $v(0)$ and integrate for positive time. The later random force has no correlation with the initial velocity, giving $C_v = (k_B T / m) e^{-t/\tau_v}$, where $\tau_v = m/\zeta$. Its area and the resulting displacement are

$$D = \frac{k_B T}{\zeta},$$

$$\langle \Delta x^2(t) \rangle = 2D[t - \tau_v (1 - e^{-t/\tau_v})].$$

At $t \ll \tau_v$, the constant and linear terms inside the brackets cancel and displacement is ballistic:

$$\langle \Delta x^2 \rangle \simeq (k_B T/m) t^2.$$

At $t \gg \tau_v$ the slope approaches $2D$. Fitting a line through the origin at times only a few relaxation times overestimates the influence of the negative intercept in this long-time expression and underestimates D. Comparisons at multiple lags distinguish crossover curvature from a true anomalous power law.

For an inert sphere of radius 0.50 micrometres and density 2000 kg per cubic metre in water with $\eta = 1.00$ mPa s at 298 K, its bare mass is 1.05×10^{-15} kg and Stokes drag 9.42×10^{-9} kg/s. The estimate $\tau_v = 0.111$ microseconds is far faster than ordinary video tracking. $D = 0.437$ square micrometres per second and the one-coordinate rms displacement over 10 ms is 0.0935 micrometres. Hydrodynamic added mass and memory affect very short-time behaviour; the simple exponential is an analytical reference, not a claim that a water suspension remains Markovian down to arbitrary times.

Relate a weak imposed force to equilibrium motion

In the same soluble model a small constant force F turns the mean velocity into $F(1 - e^{-t/\tau_v})/\zeta$, so the mobility $b = \lim \langle v \rangle / F = 1/\zeta$. Together with equilibrium diffusion this gives the Einstein relation $D = k_B T b$. A broader equilibrium argument considers a dilute suspension in a weak conservative potential $U(x)$. Its current is

$$j = -D\,\partial_x n + bFn,$$

with $F = -\partial_x U$. The Boltzmann distribution $n \propto e^{-U/k_B T}$ must carry zero current; substitution gives the same relation without assuming exponential velocity relaxation.

Linear response connects the full weak-force impulse response to an equilibrium correlation. For a perturbation coupled as $-F(t)x$, classical equilibrium response of velocity has causal kernel $C_v(t)/(k_B T)$ for $t > 0$ under the usual microreversibility assumptions. Integrating the response kernel for a constant weak force gives

$$b = (1/k_B T) \int_0^\infty C_v(t) dt,$$

which recovers the kinematic diffusion relation. Strong forcing, an active bath or a nonequilibrium steady state need not obey this simple equilibrium identity.

A direct check compares independently measured weak sedimentation or calibrated drift with freely diffusing trajectories. For the previous bead a force 10 fN would produce a mean drift about 1.06 micrometres per second in the ideal Stokes regime. During a 10 ms frame its mean drift is about 0.011 micrometres, much smaller than rms Brownian motion, so many frames or many particles are needed. A single short jagged record can appear to move against the force. Use equal positive and negative applied forces to cancel constant stage drift in the estimated mobility.

Work through diffusion and mobility in the random-walk chapter

Stress and heat current carry their own normalisations

In a homogeneous isotropic laboratory fluid, define P_{xy} as a spatially averaged off-diagonal microscopic stress, with units Pa. Its equilibrium mean is zero in an isotropic state without imposed shear. A weak velocity gradient couples to momentum transport. The Green–Kubo expression for dynamic shear viscosity is

$$\eta = \frac{V}{k_B T} \int_0^\infty \langle P_{xy}(t) P_{xy}(0) \rangle \, dt.$$

If a program outputs an extensive stress moment $J_{xy} = V P_{xy}$, the prefactor is $1/(V k_B T)$. Multiplying an already extensive current by volume again is a common orders-of-magnitude error. Check dimensions: pressure squared times time times volume divided by energy gives pascal seconds. Kinetic and interaction contributions to microscopic stress must both be included according to the model; a calculation with ideal kinetic flux alone need not reproduce the viscosity of a dense fluid.

Similarly for a component of total heat current $J_{Q,x}$ with units energy times length per time, the thermal conductivity in an isotropic equilibrium single-component system is

$$\kappa = \frac{1}{V k_B T^2} \int_0^\infty \langle J_{Q,x}(t) J_{Q,x}(0) \rangle \, dt.$$

Here $k_B T^2$ means k_B times T^2. If a heat-flux density is used instead, the volume prefactor moves as for stress. Heat current must exclude energy carried by a net mass flow as appropriate to the chosen thermodynamic constraint. In mixtures several conserved species currents can couple, and the measured conductivity at zero mass flux can differ from one unprojected current coefficient. The meaning of the measured flux and the imposed condition must be specified before comparing coefficients.

For an illustrative volume 10^{-24} cubic metres at 300 K, assume one stress autocorrelation $C_P(t) = (2.0 \times 10^6 \ \mathrm{Pa})^2 e^{-t/(2.0 \ \mathrm{ps})}$. Its area is 8.0 square pascal seconds. Multiplication by $V/(k_B T)$ gives viscosity 1.93×10^{-3} Pa s. Doubling the simulated volume of an otherwise similar homogeneous fluid generally reduces the variance of its volume-averaged stress, keeping the physical transport coefficient intensive. Holding the same numerical stress variance while merely changing the prefactor would describe a different equilibrium fluctuation model.

Find a correlation integral's useful time window

An estimate at lag s averages over pairs of samples separated by that lag. In a record of duration T_{rec}, the number of available pairs decreases with lag. Overlapping time origins strongly correlate the estimates, so plotting one error bar per raw pair exaggerates precision. Integrating an estimated correlation to the maximum possible lag frequently accumulates late random noise rather than useful signal. Conversely stopping too early omits a physical tail.

For a positive single-exponential reference, integrating to three correlation times retains $1 - e^{-3} = 95.0$ percent of the area. Five times retain 99.3 percent. A correlation containing a small, much slower

component can look essentially flat on the main decay scale while contributing appreciably to the area. Divide a long run into blocks, compare running integrals and check several window lengths. Report the variation between independent blocks and any explicit tail model used beyond the measured interval.

At short lags, finite exposure averages correlated positions or velocities. Uncorrelated localisation errors in position increase nonzero-lag displacement variance by a constant $2s_x^2$ per coordinate. Numerically differentiating those same noisy positions creates large artificial velocity variance and negative correlations between adjacent increments. An apparent negative velocity lobe can therefore be physical backscattering or a measurement artefact. Comparing a raw displacement fit with the velocity integral after independently characterising detector noise is more informative than differentiating first and interpreting every feature as fluid memory.

A forced shear wave gives a second viscosity measurement

An equilibrium stress correlation estimates viscosity without imposing any flow. A separate driven measurement can test that coefficient against a directly observed penetration depth. Take a Newtonian fluid filling $y > 0$, driven by a laterally oscillating plate at $y = 0$ with speed $U_0 \cos \omega t$. Neglect boundaries at large y, compressibility and nonlinear convective effects in this spatially uniform tangential shear. The momentum equation reduces to

$$\partial_t u = \nu \partial_y^2 u,$$

with kinematic viscosity $\nu = \eta / \rho$.

Seek $u = \mathrm{Re}[U_0 e^{-\lambda y - i\omega t}]$ and choose the square root with positive real part. Then $\lambda = (1 - i)/\delta$,

$$\delta = \sqrt{2\nu/\omega}.$$

Both amplitude decay and phase lag carry information: at distance one penetration depth the amplitude is e^{-1} of the wall speed and the lag magnitude is one radian. A single complex velocity map can determine viscosity from either slope and use their agreement to test the simple model.

For a fluid of density 1000 kg per cubic metre and measured penetration depth 0.126 mm at 20 Hz, $\nu = \omega \delta^2 / 2 \simeq 1.00 \times 10^{-6}$ square metres per second and $\eta \simeq 1.00$ mPa s. A stress-fluctuation integral yielding a value twice as large would predict a penetration depth larger by $\sqrt{2}$, approximately 0.178 mm. Such a comparison can discriminate a missing component or volume factor in the equilibrium calculation if density, temperature and fluid composition match.

At high enough frequency the fluid may have a complex frequency-dependent viscosity rather than the zero-frequency coefficient from a Green–Kubo area. In linear response that generalisation is the one-sided Fourier transform of the equilibrium stress correlation with appropriate prefactors. A single-exponential stress memory gives a frequency-dependent response with a characteristic roll-off and phase lag. Comparing a high-frequency penetration measurement directly to a zero-frequency area without checking the timescale can therefore produce a genuine difference, not a statistical failure.

Remove a centre-of-mass mode only when the observable requires it

In a finite periodic particle simulation total momentum may be exactly conserved. If the initial centre of mass has a random velocity and no thermostat removes it, every particle's measured laboratory coordinate includes the common drift $V_{\mathrm{CM}}t$. Its mean-squared displacement gains a ballistic term $\langle V_{\mathrm{CM}}^2\rangle t^2$. In a canonical ensemble the centre-of-mass velocity variance in one direction is k_BT/M_{tot}. Though it falls with system size, at sufficiently long times its quadratic contribution exceeds the linear diffusive part.

For 1000 equal particles of mass m, an equilibrium centre-of-mass mode contributes one thousandth of the single-particle instantaneous velocity variance. In a long-time autocorrelation this small constant never integrates to a finite diffusion coefficient. Subtract centre-of-mass displacement or work in a zero-total-momentum ensemble before taking the long-time integral of an internal tracer's motion. In a real sample with externally imposed drift, decide whether the physical quantity sought is diffusion about that drift or absolute laboratory displacement; the subtraction has a physical meaning rather than merely a cosmetic improvement.

Conversely, indiscriminately subtracting the mean velocity of a tiny observation subregion separately at every time can remove genuine hydrodynamic long-wavelength motion. A strongly applied single-particle thermostat can alter current correlations and suppress transport even while maintaining the correct static temperature histogram. Check the thermostat strength and system size; compare an equilibrium-prepared weakly thermostatted trajectory with an energy-conserving production segment where appropriate. Maintaining the correct one-time distribution is not sufficient to preserve all dynamical correlation functions.

The distinction also matters when removing an estimated mean stress before integrating its fluctuations. The true shear mean may vanish by symmetry, but a short noisy record has a nonzero sample mean. Subtracting a slowly varying trend on the same timescale as a genuine stress tail can erase physical viscosity. Use independent samples, separately measured drifts and explicit sensitivity to the subtraction window. A persistent nonzero off-diagonal stress under nominal isotropic conditions may indicate incomplete equilibration, a deformed box or an unrelaxed solid rather than a fluid transport coefficient.

Heat and particle currents can share fluctuations

Consider two thermodynamic fluxes J_1, J_2 driven weakly by conjugate forces X_1, X_2, chosen so entropy production density is $J_1X_1 + J_2X_2$. Linear response writes $J_i = \sum_j L_{ij}X_j$. Near time-reversal-symmetric equilibrium, with flux parities chosen appropriately and no magnetic-field reversal complications, reciprocal relations give

$$L_{12} = L_{21}.$$

The symmetric part must be positive semidefinite so that entropy production is nonnegative for every small force. Thus $L_{11}, L_{22} \geq 0$ and $L_{12}^2 \leq L_{11}L_{22}$.

Suppose a thermal measurement enforces $J_1 = 0$ by allowing a compensating force X_1 to develop, as when a concentration or electrochemical gradient cancels a coupled particle flux. Solving the first row gives

$$X_1 = -L_{12}X_2/L_{11}.$$

Substitution into the second yields response coefficient $L_{22} - L_{21}L_{12}/L_{11}$. The measured constrained heat response can therefore be smaller than the coefficient at fixed $X_1 = 0$. The difference has a simple operational origin in the boundary condition enforced during the experiment.

Take consistent normalised coefficients $L_{11} = 4, L_{22} = 3, L_{12} = L_{21} = 2$. At zero first force the heat flux slope is three. At zero first flux it is $3 - 4/4 = 2$. Both are positive and consistent with the same equilibrium correlation matrix. Comparing a simulated heat-current autocorrelation corresponding to the unconstrained coefficient against a zero-particle-flux experiment requires the mixed-current correlation too. Apparent disagreement by a factor of one and a half can arise entirely from this distinction.

Green–Kubo cross-correlations must use properly defined conjugate fluxes and units. Redefining heat current by adding a constant multiple of particle current changes the matrix entries but must leave a physically defined zero-particle-flux heat measurement invariant when the forces are transformed consistently. This reference invariance is analogous to subtracting a per-particle energy reference in grand-canonical covariance. It provides a powerful test of a multicomponent transport implementation without relying on any empirical transport value.

A lattice trajectory provides an exactly known correlation integral

For a one-dimensional nearest-neighbour walk with steps of length a every interval Δt, suppose successive step directions have correlation $\langle s_{j+k}s_j\rangle = r^k$ with $|r| < 1$, and zero mean. Squaring the sum of n steps gives

$$a^2[n + 2\sum_{k=1}^{n-1}(n-k)r^k].$$

In the large-n limit divide by $2n\Delta t$ to obtain

$$D = a^2(1+r)/[2\Delta t(1-r)].$$

The uncorrelated random-walk value is recovered at $r = 0$.

With $a = 1$ micrometre and $\Delta t = 0.10$ s, uncorrelated diffusion is five square micrometres per second. Persistence $r = 0.5$ raises it to fifteen, while anticorrelation $r = -0.5$ lowers it to $5/3$. All three walks have the same step-size variance, so an estimate based only on one-frame displacement misses the entire correlation effect. The finite-lag autocorrelation sum and a long-time displacement fit should recover the same value on a long simulated record.

The discrete formula also clarifies the special zero-lag weight in numerical integration. An integral approximated from sampled symmetric correlations gives half weight to the zero lag on a one-sided

trapezoidal rule. A careless double count of the instantaneous self term predicts twice the uncorrelated diffusion. Benchmark a numerical estimator with this exact walk before applying it to molecular data with unknown late tails and noisy velocity derivatives.

Measure an off-diagonal diffusion coefficient after rotating the sample

In an anisotropic sheet a tracer's two-dimensional displacement covariance need not align with the camera axes. Let its long-time principal diffusion coefficients be D_1, D_2 and the first principal axis make angle θ with laboratory x. In matrix form the laboratory tensor is $R\,\mathrm{diag}(D_1, D_2)R^{\mathsf{T}}$, with R the planar rotation. At sufficiently large lag in an effectively unbounded drift-free region, the displacement covariance matrix tends to $2Dt$. An off-diagonal covariance is a measurable correlation between displacement components even if the noise is independent in the material principal directions.

Take $D_1 = 2.0$ and $D_2 = 0.50$ square micrometres per second with $\theta = 30^\circ$. Laboratory components are $D_{xx} = 1.625$, $D_{yy} = 0.875$ and $D_{xy} = 0.650$, in the same units. At 0.10 s the cross covariance is 0.130 square micrometres. Rotating the physical sheet by ninety degrees interchanges the two diagonal values and reverses the sign of the off-diagonal one. An instrument-fixed camera readout artefact would not generally follow this material rotation.

To infer the principal axes from data, fit each covariance component versus lag using a common drift subtraction and independent position calibration. Diagonalise the symmetric fitted diffusion matrix. Positive semidefiniteness requires

$$D_{xy}^2 \le D_{xx} D_{yy}.$$

A strong violation indicates incompatible estimates, unit conversion errors or statistical uncertainty, not a physically permissible strongly anisotropic diffusion tensor. Near-degenerate principal values make the orientation poorly determined even when the mean diffusion coefficient is precise. Report uncertainties on eigenvalues and orientation separately.

A finite field of view introduces a further bias. Fast particles along the high-diffusion direction leave the image sooner, and a tracking pipeline retaining only long trajectories can favour slow or fortuitously confined paths. Compare different fields of view and minimum track durations, or fit a measurement model incorporating censoring. A physical material anisotropy should persist after correcting these effects; apparent anisotropy aligned with the short direction of a rectangular observation window demands particular scrutiny.

A weak calibrated force along different axes supplies an independent check through the tensor Einstein relation under equilibrium conditions. If the diffusion tensor predicts off-axis mean drift, the mobility experiment should observe its sign and magnitude after reversing the force. A discrepancy could reveal active forcing, nonstationarity, an incorrect temperature or a mismatch between the mobility and displacement observables. It cannot be diagnosed from a scalar mean square alone.

Another rotational invariant is the trace of the diffusion tensor. In the example its value stays 2.50 square micrometres per second for every sample orientation, while the determinant stays one square

micrometre to the fourth per second squared. A comparison of two camera orientations therefore provides numerical checks independent of the fitted angle. A drift mistakenly left in the displacement covariance can violate the expected linear dependence on lag, making the inferred effective eigenvalues grow with the selected fitting window. Subtract the ensemble mean displacement at each lag and record that mean separately as a flow measurement. Conversely do not subtract an individual best-fit line from every short trajectory without modelling the resulting loss of long-time fluctuations; that procedure can bias both diffusion coefficients downward.

Choose an integration window for a transport coefficient

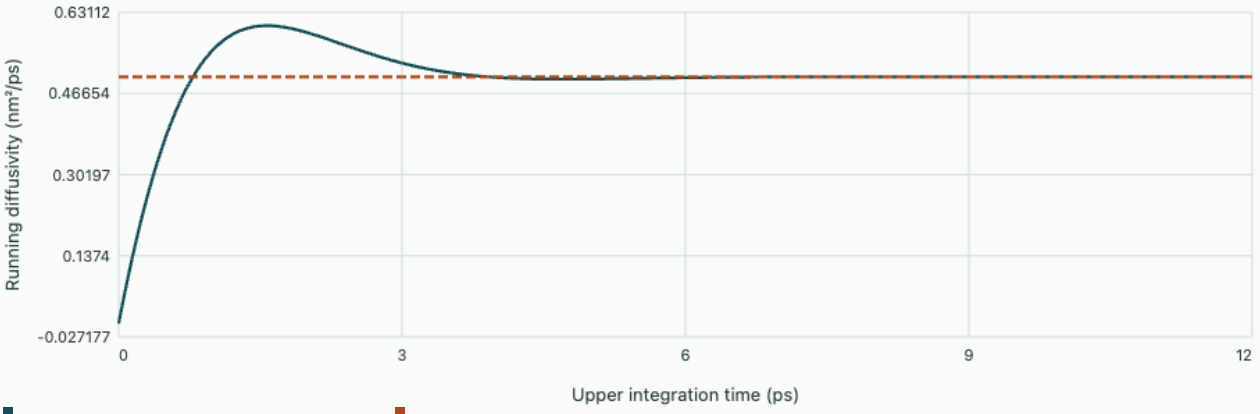
Integrate an oscillatory velocity correlation over a finite record. Compare a short apparent plateau with the independently computed infinite-time result.

Open this experiment online

Try the calculation

Correlation decay time τ (ps) 1

Oscillation frequency Ω (rad/ps) 1



With C(0) = 1 nm²/ps² the exact 1D diffusion coefficient is 0.5 nm²/ps. At 2τ the integral gives 0.58969 nm²/ps.

Tutorial · Invert diffusion measurements at two lags

In a one-dimensional diffusive window a measured displacement variance is 0.024 square micrometres at 10 ms and 0.064 at 50 ms. Assume long-time Brownian displacements plus independent frame-localisation variance per position measurement. Estimate D and localisation precision.

▼ **Fit slope and intercept separately**

The variance slope is $(0.064 - 0.024)/(0.050 - 0.010) = 1.00$ square micrometre per second, so $D = 0.500$ square micrometres per second. The intercept is 0.014 square micrometres, equal to twice the single-frame localisation variance. The implied rms localisation error is $\sqrt{0.007} = 0.0837$ micrometres. Forcing the fit through the origin at the shortest lag would give $D = 1.20$, more than twice the inferred value.

Acquire more lags to test the linear model. A static drift contributes a term growing as t^2 if its mean displacement is not removed, while finite exposure can lower the physically measured short-lag variance. Two points always define a line and cannot independently validate either noise model.

Tutorial · Estimate a hidden slow tail in viscosity

A stress correlation is modelled as $A[e^{-t/\tau} + 0.02e^{-t/(40\tau)}]$. What fraction of its true area comes from the low-amplitude tail? How much does a cutoff at 5τ capture?

▼ **Weight a correlation by its duration**

The fast and slow areas are $A\tau$ and $0.80A\tau$. The slow contribution is 44.4 percent of total. Integration to five fast times captures $A\tau[(1 - e^{-5}) + 0.80(1 - e^{-5/40})] \simeq 1.087A\tau$, or about 60.4 percent. Visual disappearance of the dominant peak therefore does not establish convergence of the transport coefficient.

The slow contribution must be supported by independent long blocks or a physically motivated model. A fitted tiny constant baseline can generate arbitrarily large apparent area if extrapolated indefinitely. Subtracting a sample mean estimated from the same short record can conversely suppress a genuine slow mode. Compare records much longer than the proposed tail time before trusting the extrapolation.

Tutorial · Is an algebraic tail compatible with ordinary diffusion?

Suppose a measured long-time velocity correlation remains positive and scales as $t^{-\alpha}$. Determine when its Green–Kubo area exists and infer the displacement growth for $0 < \alpha < 1$ and for $\alpha = 1$. Why would a finite periodic simulation require care?

▼ **Integrate the time-dependent slope before naming a diffusion coefficient**

For $\alpha > 1$ the tail integral converges. For $0 < \alpha < 1$, the running area grows as $t^{1-\alpha}$, so displacement variance grows as $t^{2-\alpha}$ under the stated persistent positive tail. At $\alpha = 1$, the running area grows logarithmically and the displacement variance is of order $t \log t$. Sign oscillations, finite-size termination or another late-time crossover can alter these conclusions.

A periodic box admits only discrete long-wavelength modes and may cut off a hydrodynamic tail. Repeating at larger sizes and comparing the range before signals traverse the box helps identify a bulk regime. Unwrap particle coordinates crossing periodic boundaries before calculating a displacement. A

wrapped coordinate can falsely saturate and imitate a confined particle even in a freely diffusing periodic fluid.

Cambridge Kinetic Theory · correlations, response and transport

Work fluctuations and equilibrium free-energy differences

Statistical mechanics from microscopic counting to material response

Repeated rapid changes of a laboratory control can give a broad distribution of work even when every run follows the same prescribed protocol. Derive why a particular exponential average retrieves an equilibrium free-energy difference, why the average work is generally larger, and why a finite experiment can miss the rare trajectories that make the equality true.

Specify the system, preparation and measured work

Let a classical system have Hamiltonian $H(z,\lambda)$, with microstate z and an externally set parameter λ. Initially it is sampled from the canonical equilibrium distribution at λ_A and temperature T. A prescribed protocol $\lambda(t)$ carries it to λ_B. Work done on the system along one trajectory is the inclusive parameter work

$$W = \int_0^{t_f} \dot{\lambda}\, \partial_\lambda H(z(t),\lambda(t))\, dt,$$

$$\Delta F = F(\lambda_B,T) - F(\lambda_A,T).$$

If λ is a trap centre, the generalised force in this convention is the derivative of energy with respect to that centre. A transducer reporting the opposite force exerted by the system on the actuator needs a sign change. For a stochastic subsystem in contact with a heat bath, a change of subsystem energy also includes heat. It is not generally equal to the measured protocol work. Specify the tip, trap, sample and bath boundary before turning a force trace into a work distribution.

Every run must begin from the same equilibrium preparation, not simply from whichever end state the previous protocol produced. The final state immediately after a fast drive need not be equilibrium. The free energy being inferred refers to equilibrium endpoint Hamiltonians at the same temperature. Equilibrating after the drive exchanges heat but involves no additional protocol work when the control is held fixed. Changing control during that relaxation would define a different protocol.

A Hamiltonian proof using phase-space volume

First treat an isolated driven Hamiltonian system after canonical preparation. Work along a trajectory equals $H_B(z_f) - H_A(z_i)$. Multiply $e^{-\beta W}$ by the initial canonical weight and integrate over initial phase space. The dependence on initial energy cancels. Hamiltonian dynamics preserves phase-space volume and is invertible, so change variables from initial to final point to obtain

$$\left\langle e^{-\beta W} \right\rangle_A = \frac{1}{Z_A} \int dz_i\, e^{-\beta H_B(z_f)} = \frac{Z_B}{Z_A} = e^{-\beta \Delta F}.$$

Thus the Jarzynski equality follows without assuming the driving is slow. This proof requires the mapping to cover the relevant final phase space with the stated measure; idealised instantaneous removal of infinitely high constraints can create subtleties of support. For a weakly coupled thermal environment include bath variables in the Hamiltonian derivation or use a stochastic dynamics satisfying appropriate

local detailed balance. Arbitrary feedback, nonequilibrium initial preparation or an unmodelled strong bath coupling does not automatically obey this simple subsystem formula.

Jensen's inequality gives

$$\langle W \rangle \geq \Delta F,$$

with equality for a deterministically reversible work distribution in the relevant limit. An individual run may give less work than ΔF, sometimes even work of the opposite sign. This does not violate the ensemble inequality. The exponential average weights such small-work events more strongly than their raw frequency indicates. The larger the typical dissipated work, the more difficult a direct estimate often becomes.

Solve a complete finite-time example analytically

Take one overdamped particle equilibrated in a one-dimensional harmonic trap $H(x, \lambda) = k(x - \lambda)^2/2$ at $\lambda = 0$. Suddenly shift the centre to d while the particle position cannot respond. The endpoint free energy is unchanged by translation in an unbounded uniform bath. Initial x is Gaussian with mean zero and variance $k_B T/k$. During the instantaneous shift

$$W = \frac{k}{2}[(x - d)^2 - x^2] = \frac{kd^2}{2} - kdx.$$

Work is Gaussian with mean $kd^2/2$, variance $kd^2 k_B T$, and $\Delta F = 0$. Evaluating the Gaussian moment generates $e^{-\beta\langle W\rangle + \beta^2 \operatorname{var} W/2} = 1$, verifying the equality exactly at arbitrary shift distance. The mechanism is visible in the trajectories: particles already on the downstream side can lose potential energy in the sudden control shift, contributing negative work and disproportionately large exponential weights.

For $d = 2\sqrt{k_B T/k}$, the mean work is two thermal energies and the standard deviation is also two thermal energies. The probability of negative work is about 0.159. For $d = 5\sqrt{k_B T/k}$, mean work is $12.5 k_B T$, standard deviation $5 k_B T$, and negative work occurs only about 0.0062 of the time. Typical runs can all show substantial positive work even though the correct free-energy change for the whole translation is zero.

The most important work values for the exponential average are not simply all values below zero. Multiplying a Gaussian $P(W)$ by $e^{-\beta W}$ shifts its peak to $\langle W \rangle - \beta \operatorname{var} W$, which for this trap shift is $-\langle W \rangle$. In the five-width shift example that important region is five standard deviations below the ordinary work mean. Observing a few negative work values can still leave the estimate badly biased if the most heavily weighted part remains unsampled.

Compare forward and reverse trajectory probabilities

Define the reverse protocol by reversing the order of control values, and reverse momenta and any time-odd external parameters as required by microscopic time reversal. Prepare its initial system

independently in equilibrium at B. Under microscopic reversibility the probability ratio of a forward path and its time-reversed counterpart has a boundary weight from the initial ensembles and a conditional path weight tied to heat exchange. These combine to give $e^{\beta(W-\Delta F)}$. Grouping paths by work produces the Crooks relation

$$\frac{P_F(W)}{P_R(-W)} = e^{\beta(W-\Delta F)}.$$

Where the two work densities coincide, $W = \Delta F$. Integrating the relation against $e^{-\beta W}$ recovers Jarzynski. The reverse work axis must be reflected: a crossing of two histograms plotted with the same unflipped work sign generally has no such interpretation. A finite bin estimates an integral over a work interval, so discretisation and sparse counts modify a naive pointwise log-ratio fit.

For a more general measured Gaussian forward work distribution with mean $5.0k_BT$ and variance $4.0(k_BT)^2$, the equality predicts $\Delta F = 3.0k_BT$ if the Gaussian law including tails is justified. With equal-variance reverse statistics, the corresponding mean of $-W_R$ is $1.0k_BT$. The two densities cross at $3.0k_BT$. Plotting $\log[P_F(W)/P_R(-W)]$ against W should give slope β, providing a check on independent force and temperature calibration. Good agreement over a tiny overlap range does not constrain remote tails strongly.

Use the finite estimator honestly

From independent forward works W_i, estimate ΔF as minus k_BT times the logarithm of their mean exponential weight. The sample mean of weights is unbiased under correct sampling, but its logarithm is nonlinear. Jensen implies the finite-sample free-energy estimate is biased upward on average under the ideal assumptions. A single dataset can fluctuate either way. An ordinary mean-work error bar does not describe uncertainty in an exponential moment.

For Gaussian work the squared coefficient of variation of one exponential weight is $e^{\beta^2 \operatorname{var} W} - 1$. With $d = 2$ thermal widths it equals

$$e^4 - 1 = 53.6.$$

To estimate the mean weight to ten percent relative standard error by independent averaging would require roughly 5360 runs under that distribution. At five thermal widths the required count is dramatically higher. The nominal number of runs is not a sufficient convergence diagnostic: inspect the fraction of weight carried by the largest events and compare independent batches.

For numerical stability shift all dimensionless logarithmic weights by their largest value before exponentiating, then add it back after taking the logarithm. This prevents overflow without repairing missing rare events. Bootstrapping a dataset with no observations in the relevant tail cannot generate them. More gradual driving, intermediate equilibrium windows or a bidirectional estimator using forward and reverse overlap can reduce the difficulty. Each approach requires its own equilibrium preparations and sufficient overlap, not just a different final formula applied to the same deficient sample.

Discretise a real force trace without changing the work definition

For a measured optical trap centre λ_j and positions x_j, actual control updates and acquisition clocks may be staggered. If the control moves instantaneously from λ_j to λ_{j+1} at fixed x_j, its exact contribution is $H(x_j, \lambda_{j+1}) - H(x_j, \lambda_j)$. A smoothly moving control instead uses an integral of the measured generalised force, with a quadrature matched to timestamps and sampling bandwidth. A midpoint convention applied to an unresolved abrupt jump can miss a finite quadratic term.

For the harmonic trap the instantaneous work above is exactly $k[(\lambda_j + \lambda_{j+1})/2 - x_j](\lambda_{j+1} - \lambda_j)$. Taking only the initial force would give a difference $k(\lambda_{j+1} - \lambda_j)^2/2$ per jump. Over many short smooth increments this bias may vanish under refinement, but for a deliberate quench it is the entire mean work if the initial mean displacement is zero. The distinction can be checked by reversing the jump and independently measuring the equilibrium spatial variance.

Compare irreversible work and relaxation with the course's independently defined entropy-balance experiment

Derive the dissipated work for a constant-speed pull

Between infinitely slow displacements and a sudden quench lies an exactly soluble finite-duration example. For a harmonic trap translated at constant velocity u in an overdamped bath, the mean coordinate obeys $\zeta\, d\langle x\rangle/dt = -k(\langle x\rangle - \lambda)$. Begin equilibrated at zero and set $\lambda(t) = ut$ for $0 < t < t_f$. The mean lag $\lambda - \langle x\rangle$ obeys a first-order equation and grows from zero toward $u\tau$, with $\tau = \zeta/k$. The mean force exerted by the actuator in this sign convention is $k(\lambda - \langle x\rangle)$. Integrating the product of this force with control speed gives

$$\langle W\rangle = \zeta u^2[t_f - \tau(1 - e^{-t_f/\tau})].$$

The free-energy change for translating this uniform trap is zero. At a fixed displacement $d = ut_f$, taking $t_f \ll \tau$ gives $kd^2/2$, precisely the sudden-quench mean. As $t_f \gg \tau$ the main term is $\zeta d^2/t_f$, tending to zero for a quasistatic pull. Simply keeping speed fixed while taking a longer time is a different limit because the endpoint then moves farther.

The fluctuating part of $x(t)$ around its deterministic mean remains an equilibrium Ornstein–Uhlenbeck process in this additive linear model. Work is a linear functional of that Gaussian process and is therefore Gaussian. Its variance follows by integrating its covariance twice, giving $\operatorname{var} W = 2k_BT\langle W\rangle$. This recovers the exponential equality at every pull duration and supplies a nontrivial benchmark for a stochastic integrator. The probability of extracting work can be computed from the resulting Gaussian tail without assuming a second ensemble.

For a pull of two thermal widths completed in one relaxation time, use $\zeta d^2/t_f = kd^2 = 4k_BT$. The bracketed dimensionless fraction is $1 - (1 - e^{-1}) = e^{-1}$, so mean work is $1.47k_BT$ and standard deviation about $1.72k_BT$. A sudden pull of the same distance has higher mean $2k_BT$. A pull lasting five relaxation times reduces the mean to $4k_BT[5 - (1 - e^{-5})]/25 = 0.641k_BT$. These changes predict which protocol will yield a usable work-overlap region for a finite number of trials.

A measurement can test this crossover without fitting free energy at all. Calibrate stiffness from an equilibrium position distribution, obtain the relaxation time from a correlation or driven response and predict the entire mean-work curve versus duration. Then assess Jarzynski estimates independently. Failure of the predicted variance-to-mean relation can reveal nonharmonic confinement, non-Gaussian bath forces, actuator timing offsets or insufficient equilibration. It need not invalidate the fluctuation identity itself.

A two-outcome work distribution shows how rare events restore equality

A particularly transparent work distribution comes from a sudden change in a two-level system. Initially its ground energy is zero and upper energy $2k_BT$. In a sudden protocol move the upper level to zero while leaving the initial ground energy unchanged. The initial upper probability is

$$p = e^{-2}/(1+e^{-2}) = 0.1192.$$

Work can take only two values: zero for an initially ground system and minus two thermal energies for a particle initially upper. Both new energy levels are degenerate, so the exact equilibrium partition ratio is

$$2/(1+e^{-2}) = 1.762.$$

The exponential average equals $(1-p)+pe^{2} = 1.762$, giving $\Delta F = -0.566k_BT$. Mean work is only $-0.238k_BT$, greater than this equilibrium free-energy change as required. With five independent realisations the probability of observing no upper preparation is

$$(1-p)^{5} = 0.530.$$

More than half such tiny samples would report zero work every time and infer a free-energy change of zero. A bootstrap resampling those five identical observations would report no empirical uncertainty at all, spectacularly missing the known sampling limitation.

Increase the initial gap to $6k_BT$ before the same quench to degeneracy. The initial upper probability falls to 0.00247, but its exponential weight rises to e^{6}. Its mean contribution to the exponential estimator remains near one. Even one hundred runs then miss this branch with probability about 0.78. Equality relies on the true preparation probability of a rare event combined with its enormous statistical weight. A finite dataset cannot certify absence of the branch by having no observations of it.

The reverse experiment begins with two equally populated degenerate states and raises one level. It samples the formerly rare branch far more easily. Reflect its raw works onto the forward work axis and compare point masses, not unrelated smooth histogram densities. A forward and reverse data combination then constrains the free-energy difference much more effectively. A fitted reverse sample must still begin at equilibrium in the degenerate endpoint; carrying the low excited-state fraction over from the forward final state would destroy its intended initial ensemble.

Instrumental work noise is amplified by the exponential

Suppose each genuine trajectory work is recorded with an independent Gaussian error of mean zero and variance s_W^2. The arithmetic mean work stays unbiased. The mean exponential weight is multiplied by $\exp(\beta^2 s_W^2/2)$, so a naive Jarzynski estimate of free energy shifts downward by $\beta s_W^2/2$ in energy units in the infinite-record limit. This systematic shift is distinct from the finite-record logarithm bias, which tends upward on average for ideal correctly measured weights.

For rms measurement error $0.5k_BT$, this Gaussian-noise effect is $-0.125k_BT$. A small positive finite-sampling bias can accidentally cancel it, producing an apparently correct number with both the work measurement and ensemble estimate wrong. Vary the number of independent trajectories and the instrument noise separately. Calibrate work reconstruction on repeated known control or force signals to estimate the error distribution rather than assuming it is Gaussian from its rms alone.

If measurement error correlates with the trajectory—for example a saturated force transducer clips the largest restoring forces—the factorisation above fails. Errors in low-work tails can matter most even if a force trace looks accurate over the bulk of ordinary values. Compare independent sensors or forward and reverse force calibrations across the full visited range. A threshold discarding negative work as unphysical would systematically delete exactly the values the equality needs.

A related problem occurs when temperature itself is inferred from the same finite fluctuation record. Temperature uncertainty affects both the exponent and conversion of the final logarithm to energy. A Gaussian work model gives an analytic sensitivity, but outside that model propagate uncertainty by re-evaluating the full weighted estimator over independently plausible temperatures and work calibrations. Treat correlations between calibration and work estimates appropriately; merely quoting the precision of the bath set point does not establish the temperature of a fast, weakly equilibrated preparation.

Verify a stochastic work protocol on a two-state network

A small simulated system with discrete energies and heat-bath jumps supplies an independent test of a coupled dynamics and work implementation. At each step change a control value instantaneously while preserving current state; record the corresponding energy difference as parameter work. Then allow a period of fixed-control stochastic relaxation whose transition probabilities preserve the new canonical distribution. Energy change during that relaxation is heat, not additional protocol work. Define the reverse process by precisely reversing the sequence of controlled changes and using an independent equilibrium start at its endpoint.

An exact proof for such a process does not require a continuum limit. After one energy change, multiplying the previous equilibrium weights by the exponential work converts them into unnormalised canonical weights of the new levels. A stochastic transition preserving these new weights leaves that weighted distribution unchanged. Apply the same argument successively and sum the final weighted probabilities to retrieve the ratio of endpoint partition functions. Each trajectory can jump many times, but the weighted identity follows from a short sequence of matrix operations.

This proof is a practical verification tool. Enumerate all paths of a two-state system over a few control steps and compare the exact weighted sum with a Monte Carlo estimate. Test a deliberately coarse but

precisely specified sequence before refining the time mesh. If direct enumeration fails, inspect signs, equilibrium preparation, order of control and relaxation, and whether the transition kernel truly preserves the prescribed endpoint weights. If enumeration succeeds while sampled averages disagree, inspect rare weights and independent batch variability before modifying the underlying dynamics.

Why postselection changes a work experiment

The standard work equality assumes a specified forward protocol and its initial equilibrium ensemble. Suppose an operator begins with an equilibrium distribution in the harmonic trap but keeps only trajectories whose initial displacement already points toward the intended final centre. Those retained starts are not a canonical ensemble for the original unrestricted trap. The conditional work distribution is systematically shifted toward smaller work, so inserting it into the ordinary exponential average generally produces a biased apparent free-energy change. This selection problem arises even if no macroscopic control action visibly changes between the retained runs.

For a sudden positive shift d, work decreases linearly with initial x. Conditioning on $x > 0$ deletes precisely the high-work half of a symmetric Gaussian preparation. More generally, groups selected with probabilities depending on the measured initial state should be recombined using the true group frequencies if the aim is to estimate the unconditional work moment. An equal-weight average of two conditionally estimated free energies is not equivalent to averaging their exponential weights, because the logarithm is nonlinear.

If the feedback actually chooses the direction of a subsequent shift after measuring the initial position, the protocol itself is now correlated with the state. For example, shifting the trap toward the initial displacement on each run reduces average mechanical work relative to a fixed-direction protocol. The additional information and the measurement/control apparatus are part of the thermodynamic accounting. A direct substitution into the no-feedback equalities derived above would address the wrong experiment. A full information-thermodynamic treatment needs specified measurement probabilities and a well-defined corresponding reverse procedure.

A simple experimental guard is to record every attempted run, including rejected starts, along with selection and trigger conditions. Verify the initial position and energy distributions both before and after data-quality cuts. A cut based on electronic saturation can remove a physically important work tail even if it has no deliberate feedback purpose. Conversely, discarding runs for an independent external electronics failure can be statistically benign when independence is demonstrated. The fact that a cut is operationally convenient is not evidence for such independence.

The same reasoning applies to chained protocols. A sequence of small equilibrium windows can improve overlap, but each initial distribution must be re-equilibrated or appropriately reweighted. Starting each window immediately after the previous rapid drive changes its preparation. A complete path-level method may still treat the sequence consistently; multiplying independent endpoint estimates derived under a false equilibrium assumption generally does not. An explicit record of state preparation, control timing and retained trajectories is as central to a fluctuation-relation measurement as an accurate mechanical force calibration.

A useful stationary control keeps the trap centre fixed but processes position and force data through the same analysis pipeline as a moving-trap run. Correct inclusive parameter work is then zero on every trajectory because the control never changes, even though the particle exchanges random heat with the bath and its potential energy fluctuates. A processed nonzero work distribution in this test reveals that the analysis integrated a different quantity, confused particle displacement with actuator displacement or introduced artificial control noise. Conversely a real translating-trap experiment can finish with the same mean potential energy it had initially after equilibration, while having delivered positive net work and dissipated heat. Tracking both control displacement and particle displacement makes these statements testable trajectory by trajectory where the specified stochastic energy convention permits.

Determine which rare work values dominate a free-energy estimate

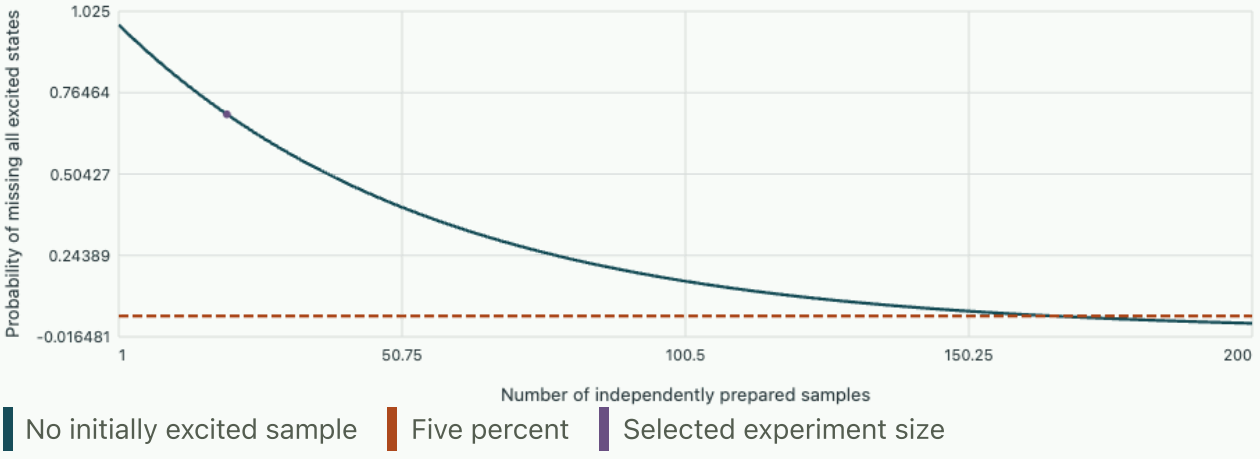
Change the preparation gap and the number of repetitions. Compare the average work, the exact free-energy change and the chance of missing the rare branch entirely.

Open this experiment online

Try the calculation

Initial gap (kBT) **4**

Independent preparations **20**



P(no rare work sample) = 0.69559. Mean work / kBT = -0.071945, ΔF/kBT = -0.675.

Tutorial · A stiffness quench changes the equilibrium free energy

An equilibrated classical harmonic trap has stiffness k. It is suddenly changed to $2k$ at the same centre. Derive mean work, exact free-energy difference and verify the exponential equality without assuming Gaussian work.

▼ **Integrate the Gaussian position rather than a guessed Gaussian work law**

Work is $W = kx^2/2$, so mean work is $k_BT/2$. The configurational partition function scales as $k^{-1/2}$, giving $\Delta F = (k_BT/2)\log 2 = 0.3466 k_BT$. The exponential-weight average is the integral of $e^{-\beta k x^2/2}$ against the original position density proportional to $e^{-\beta k x^2/2}$. The combined Gaussian integral yields

$$1/\sqrt{2} = e^{-\beta \Delta F}.$$

The work is nonnegative and chi-square distributed, not Gaussian. Its mean excess above ΔF is about $0.1534 k_BT$.

Tutorial · Estimate a crossing with uncertainty

Measured forward and sign-reflected reverse densities have equal-variance Gaussian shapes with means 6.0 and 2.0 thermal energies. Their standard deviations are both 2.0 thermal energies. Find the crossing, verify the expected logarithmic slope and explain how an unmodelled force offset would affect a displacement protocol.

▼ **Complete the square and track the offset through both protocols**

The ratio of the two Gaussian densities has logarithmic slope $(6-2)/4 = 1$ per thermal energy and vanishes at their midpoint $4.0 k_BT$. This is consistent with a free-energy change of four thermal energies. If fitted means have independent standard errors 0.2 and 0.3 thermal energies, their midpoint has standard error $\sqrt{0.2^2 + 0.3^2}/2 = 0.18 k_BT$ under the equal-variance model. Unknown unequal variances or poor histogram overlap require a more general fit.

An additive force calibration offset integrated over a known displacement adds a constant work shift proportional to that displacement. The reverse protocol has opposite displacement and hence opposite offset in its raw work, which becomes the same shift after reflecting its work axis. The crossing moves with both histograms and their difference need not reveal the calibration error. Independent force-zero checks matter.

Tutorial · Why an unchanged average dissipated work is insufficient

Two protocols with the same endpoints have the same mean work above equilibrium free-energy change. One has a narrow distribution near that mean and a very rare low-work tail; the other has a broader nearly Gaussian shape. Can equal mean dissipation guarantee comparable convergence of a unidirectional free-energy estimate?

▼ **Compare the quantities that govern the estimator**

The mean work determines only the first moment. The exponential estimator depends on the entire distribution and its variance depends on the second exponential moment $\langle e^{-2\beta W} \rangle$, which can differ greatly at fixed $\langle W \rangle$ and fixed Jarzynski average. A rare low-work tail can dominate both the true mean weight and its sampling error. Identical mean dissipation does not determine its scale or frequency.

Compare forward and reverse overlap over the actual important work interval, analyse independent batches and vary protocol speed. In a proposed experiment verify initial equilibrium at each endpoint, acquisition of the full work history and calibration of temperature and control force. Convergence of a mean mechanical work with repeated trials answers a different question from convergence of an exponential estimate of equilibrium free energy.

Jarzynski · original nonequilibrium equality for free-energy differences

Crooks · entropy production and the forward/reverse work relation

Measure the heat capacity of a reservoir from the width of an energy histogram

Counting, correlations and thermal measurements

Two small Einstein solids share a fixed number of excitation quanta. A calorimeter records how many are in the first solid. Count the allowed microstates, compare the exact distribution with a reservoir approximation and see how energy conservation correlates the readings.

Count the ways to store an integer number of quanta

Let solid A contain a independent harmonic modes and solid B contain b, all with energy spacing ε. Omit their fixed zero-point energy. The total excitation count Q is fixed. In one solid a configuration is an ordered list of nonnegative mode occupations. An arrangement of q identical excitation symbols and $a-1$ separators specifies such a list. The multiplicity is $\binom{q+a-1}{a-1}$. The count for distinguishable packets would be different, because exchanging two excitations between identical occupations introduces no new configuration here.

$$P(q) = \frac{\binom{q+a-1}{a-1}\binom{Q-q+b-1}{b-1}}{\binom{Q+a+b-1}{a+b-1}},$$

$$q = 0, \ldots, Q.$$

The denominator counts all occupations of the $a+b$ modes; summing the numerator over every partition must reproduce it. As a simple check, set $a=2, b=1, Q=2$. The unnormalised weights for $q=0,1,2$ are 1, 2 and 3. The most probable count is 2, but the mean is $4/3$. “Thermal equilibrium” does not require every observed energy to coincide with the most likely energy.

Separate the mean from the fluctuations

By permutation symmetry each individual mode holds mean $Q/(a+b)$. Therefore $\langle q \rangle = Qa/(a+b)$. For the second moment, introduce an auxiliary variable x marking quanta in solid A. The coefficient of z^Q in $G(x,z) = (1-xz)^{-a}(1-z)^{-b}$ is the unnormalised probability generating function. Differentiating once in x weights by q, twice by $q(q-1)$. At $x=1$ the second derivative is $a(a+1)z^2(1-z)^{-(a+b+2)}$. Divide its coefficient by the total multiplicity and combine first and second moments. Writing $s = a+b$ gives

$$\operatorname{var}(q) = \frac{Qab(s+Q)}{s^2(s+1)}.$$

The second solid contains exactly $Q-q$. Its variance is identical, and the covariance of the two measured energies is minus their common variance in energy units. A pair of independently fitted canonical distributions cannot reproduce this anticorrelation. At $Q=24,\ a=4,\ b=8$ the mean number in A is 8 and the variance is about 14.77. These are fluctuations of a perfectly isolated combined system, not a leak in the calorimeter.

Background · Can two noisy calorimeters expose the true energy variance?

The two solids exchange energy while their sum is fixed. If the first energy fluctuates upward by δE, the other fluctuates downward by exactly the same amount. Let the two recorded energies be $X = E_1 + n_1$ and $Y = E_2 + n_2$. Suppose each readout noise is independent of the physical energies and of the other noise. The same-channel variance and the cross covariance then measure different combinations.

$$\operatorname{var}(X) = \operatorname{var}(E_1) + \operatorname{var}(n_1),$$
$$\operatorname{cov}(X, Y) = -\operatorname{var}(E_1).$$

A single histogram that is too broad can mimic a reservoir with different heat capacity. Simultaneous measurements in the two channels allow the exchanged-energy variance to be recovered from the negative covariance. A drifting bath or common electronic pickup violates the assumptions. The measured variance of $X + Y$ is a direct diagnostic for such extra fluctuations when the readout noise levels are known.

Compare with a canonical ensemble at the same mean

In a canonical description each mode has a geometric distribution. Match its mean to $\bar{n} = Q/s$. A single mode then has variance $\bar{n}(1 + \bar{n})$, and a independent modes have variance $a\bar{n}(1 + \bar{n})$. Dividing the exact microcanonical variance by this canonical variance leaves $b/(s + 1)$. At the example point this is $8/13$. Increasing a along with b at fixed fraction does not make a macroscopic subsystem independent of the reservoir; the ratio stays below one. Holding a fixed, taking b large and keeping Q/s fixed does recover the canonical width.

The negative-binomial canonical distribution has support on every nonnegative integer, even those above Q. This is another sign that it describes another experiment. Merely deleting the tail without changing its other weights does not recover the true finite-reservoir conditional probability.

Calculate the distribution without enormous factorials

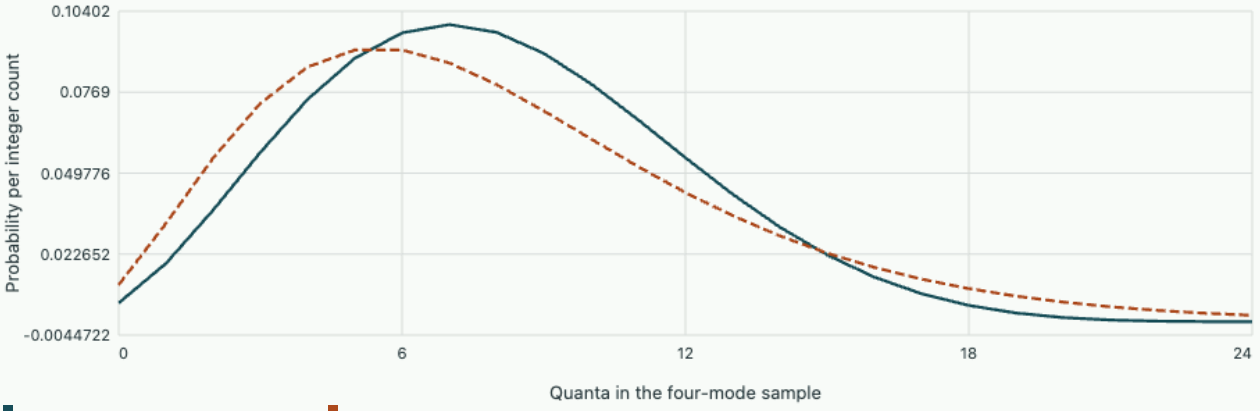
The ratio of successive unnormalised weights is $(q + a)(Q - q)/[(q + 1)(Q - q + b - 1)]$. Start with a positive arbitrary initial weight, march through the allowed q and normalise once. The graph compares this exact probability to a canonical law at the same mean. Adjust reservoir size and total energy; use the displayed first two moments to distinguish a shift in the average from a change in the width. The plotted canonical values cover the physically accessible count interval, so their displayed partial sum can be less than one.

Open this experiment online

Try the calculation

Reservoir modes 8

Excitation quanta 24



Mean 8 quanta; closed-pair variance 14.769. Canonical variance 24.

Tutorial

If A and B contain identical mode counts, is $Q/2$ necessarily the most likely observed value?

▼ **Work through the solution**

The distribution is symmetric. For odd Q there is no integer value at exactly $Q/2$; two equally likely central counts are possible. With one mode in each solid, every admissible value is equally likely. Symmetry fixes the mean exactly even when it does not identify a unique most probable observation.

For $a = 4, b = 8, Q = 24$ what variance would an unconditioned canonical fit predict?

▼ **Work through the solution**

Here $\bar{n} = 2$. The canonical count variance is $4(2)(3) = 24$. Multiplying by $8/13$ returns 14.769 for the closed pair. Count-to-energy conversion multiplies variances by ε^2.

An experiment finds a noisy sum of the two reported energies. What can one infer before relaxing the energy-conservation model?

▼ **Work through the solution**

In the ideal model $E_A + E_B$ is constant. Check both readouts' measurement noise and energy calibration. Independently fluctuating detector errors add variance in the sum even for a closed sample.

Record a reference configuration or estimate detector noise from repeated measurements of the same state.

The finite-reservoir chapter connects this exact count to the curvature of the entropy.

Recover a free-energy difference from a distribution of work

Counting, correlations and thermal measurements

A two-state system starts in equilibrium, then a control jumps its energy gap. Individual work readings differ although every run uses the same control protocol. Use the elementary model to derive an exponential work identity, compare it with the average work and estimate how many runs it takes to see the events that dominate the estimate.

Specify how work is measured

Set the lower energy to zero and the upper energy initially to ϵ_i. Prepare a canonical ensemble at inverse temperature β, isolate it and suddenly change the gap to ϵ_f without changing the energy eigenvectors. A run that started low stays low and requires zero work. A run that started high stays high and requires work $\epsilon_f - \epsilon_i$, positive if the external control gives energy to the sample. No heat flows during the jump. Returning the system to equilibrium afterwards can exchange heat, but it is a separate part of the protocol.

The initial high-state probability is

$$p_i = e^{-\beta\epsilon_i}/(1 + e^{-\beta\epsilon_i}).$$

Therefore

$$\langle e^{-\beta W}\rangle = (1 - p_i) + p_i e^{-\beta(\epsilon_f - \epsilon_i)} = \frac{1 + e^{-\beta\epsilon_f}}{1 + e^{-\beta\epsilon_i}} = \frac{Z_f}{Z_i}.$$

Taking a negative logarithm and multiplying by $1/\beta$ gives the equilibrium free-energy difference of two ensembles at the same temperature. The actual final population after the jump is still p_i, which generally differs from the final equilibrium occupation. This distinction explains why an equilibrium state function can be recovered from a nonequilibrium final distribution.

Compare the two averages

The mean work is $p_i(\epsilon_f - \epsilon_i)$. Convexity of the exponential gives

$$\langle W\rangle \geq \Delta F.$$

In units of $k_B T$, take an initial gap of 4 and a final gap of 0. The high state occurs with probability 0.01799 and gives work −4 when it occurs. Mean work is about −0.07194, yet $\Delta F = -\ln[2/(1 + e^{-4})] \simeq -0.6750$. Both numbers are negative. The greater number is the one closer to zero. Checking only the sign of average work would not reveal the large difference.

Each rare upper-state event contributes e^{4} to the exponential average. Its probability times its weight is nearly one, comparable to the entire contribution of the common lower state. This is why replacing an

average of exponentials by an exponential of an average loses essential information. Repeating the same protocol more slowly with contact to a bath would describe a different path through state space.

Ask what a finite dataset can support

For N independent preparations, the probability of seeing no initial upper state is $(1 - p_i)^N$. With 20 runs in the example this is about 0.696. In such a dataset every measured work is zero and the empirical free-energy difference is zero. This is fully compatible with the model. To have at least a 95 percent chance of seeing even one upper-state event requires $N \geq \ln(0.05) / \ln(1 - p_i)$, or about 166 runs. Seeing only one such event still gives a noisy estimate of its frequency.

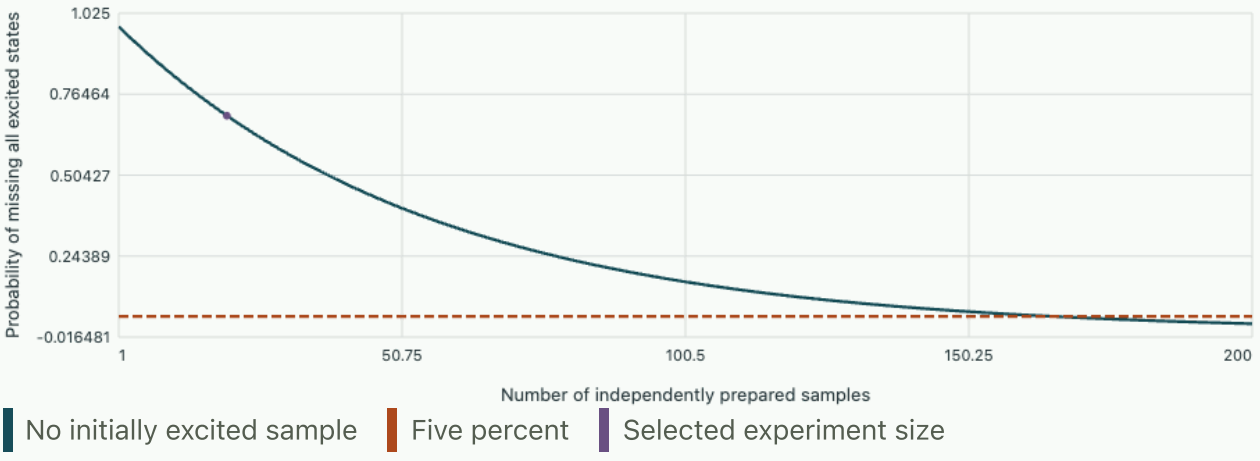
The graph holds the final gap at zero. Change the initial gap and ensemble size. It shows the probability of missing the rare branch at each sample size, together with the five-percent level. The reported exact free energy and mean work use all initial states. These predictions can be checked by explicitly enumerating a two-outcome histogram, without drawing a random sample and silently treating its fluctuations as a new law.

Open this experiment online

Try the calculation

Initial gap (kBT) **4**

Independent preparations **20**



P(no rare work sample) = 0.69559. Mean work / kBT = -0.071945, ΔF/kBT = -0.675.

Tutorial

For an initial gap zero and a final gap $2k_BT$, calculate $\langle W \rangle$ and ΔF.

▼ **Work through the solution**

Initially $p_i = 1/2$. The average work is $k_B T$. The free-energy difference is $k_B T \ln[2/(1+e^{-2})] \simeq 0.5662 k_B T$. Some energy supplied by the control is unavailable as recoverable equilibrium free energy at this temperature.

A preparation procedure always starts the sample in its lower state independent of the specified initial gap. Can the same equality be used by changing only the sample count?

▼ **Work through the solution**

No. The canonical initial probabilities were used algebraically to turn the weighted sum into Z_f/Z_i. With deterministic preparation the exponential average is one in this protocol. Repeated measurements approach that different expectation.

At fixed initial and final gaps, what does reducing temperature do to the number of runs required to see an initial excitation?

▼ **Work through the solution**

The excitation probability is approximately $e^{-\beta\epsilon_i}$ for a large positive reduced gap. Required sample count grows approximately as $1/p_i$ for any fixed detection probability. Controlling preparation and understanding tails become more important as temperature falls.

Connect this model to the partition function and to the role of free energy in a reversible isothermal process.

Applied study

Size a heat exchanger from energy balances and determine the temperature limits of the design.

Follow energy through an open system

A heat exchanger recovers energy from a warm process stream to preheat an incoming cold stream. A turbine also cools a stream, yet delivers shaft work. A steady-flow control volume distinguishes the mechanisms. Neglect changes in kinetic and potential energy and let heat into the control volume and shaft work out be positive. For one inlet and one outlet the first law reads $\dot{Q} - \dot{W} = \dot{m}(h_{\text{out}} - h_{\text{in}})$, where h is specific enthalpy. For an ideal gas with constant specific heat, or a liquid stream with constant specific heat and negligible pressure change,

$$h_{\text{out}} - h_{\text{in}} = c_p(T_{\text{out}} - T_{\text{in}}).$$

Draw a boundary around an insulated exchanger containing both streams. With no shaft work, any heat given up by the hot stream must appear in the cold stream. Define heat-capacity rates $C_h = \dot{m}_h c_{p,h}$ and $C_c = \dot{m}_c c_{p,c}$, measured in W/K. The positive rate of internal heat transfer is

$$\dot{Q}_{\text{ex}} = C_h(T_{h,i} - T_{h,o}) = C_c(T_{c,o} - T_{c,i}).$$

The smaller capacity rate limits the possible transfer to $\dot{Q}_{\max} = C_{\min}(T_{h,i} - T_{c,i})$. Energy balance alone gives one equation for the two outlet temperatures. The missing information is how much surface and thermal conductance connect the streams.

Derive the counterflow temperature profile

Let x increase in the hot-stream direction. The cold stream enters at the other end. Per unit length, a strip of wall of width w transfers $Uw(T_h - T_c)$. Assume steady operation, constant overall coefficient U, no external losses and negligible axial conduction. A slice balance gives

$$dT_h/dx = -Uw(T_h - T_c)/C_h,$$
$$dT_c/dx = -Uw(T_h - T_c)/C_c.$$

Both slopes are negative in this coordinate. A cold parcel warms as it moves towards decreasing x.

Take $C_h = C_{\min}$, $C_r = C_h/C_c$ and $\text{NTU} = UA/C_h$, with transfer area $A = wL$. Subtract the slice balances. The local temperature difference at L divided by its value at zero is $\exp[-\text{NTU}(1 - C_r)]$. Substitute the two stream balances above to eliminate the outlet temperatures. The resulting effectiveness is

$$\varepsilon = \frac{\dot{Q}_{\text{ex}}}{\dot{Q}_{\max}} = \frac{1 - e^{-\text{NTU}(1-C_r)}}{1 - C_r e^{-\text{NTU}(1-C_r)}}.$$

With equal capacity rates the difference is constant along the wall, and the limiting effectiveness is $\text{NTU}/(1 + \text{NTU})$. With a vanishing area there is no exchange. These limits provide simple checks on

the algebra.

Worked example · Two water channels

Use $T_{h,i} = 350$ K, $T_{c,i} = 290$ K, $C_h = 0.0054$ W/K and $C_c = 0.0090$ W/K. The wall and fluid films together give $U = 90$ W/(m² K) over $A = 0.045(0.002) = 9.0 \times 10^{-5}$ m². Thus $C_r = 0.60$, $\mathrm{NTU} = 1.50$, and the exponential factor is

$$e^{-0.60} = 0.5488.$$

The effectiveness is 0.6727 and $\dot{Q}_{\mathrm{ex}} = 0.6727(0.0054)(60) = 0.2180$ W.

The hot outlet is $350 - 0.2180/0.0054 = 309.64$ K, and the cold outlet is $290 + 0.2180/0.0090 = 314.22$ K. These outlets lie at opposite ends. The local hot-to-cold differences at the ends are $350 - 314.22 = 35.78$ K and $309.64 - 290 = 19.64$ K, both positive. For a body cooling in time against a reservoir, the same conductance enters a different energy balance. Study transient cooling of a solid.

Transfer problem · A compressor and turbine

In an ideal gas-turbine cycle, a compressor raises the pressure by a factor of 8, heat is supplied at constant pressure and the turbine expands the stream back to its starting pressure. Let air enter the compressor at 300 K and the turbine at 1200 K. Assume ideal-gas air with constant $\gamma = 1.4$ and $c_p = 1.005$ kJ/(kg K), reversible adiabatic compression and expansion, and negligible pressure losses. For either isentropic machine $T_{\mathrm{out}}/T_{\mathrm{in}} = (p_{\mathrm{out}}/p_{\mathrm{in}})^{(\gamma-1)/\gamma}$.

The pressure factor raised to the temperature exponent is

$$8^{0.4/1.4} = 1.811.$$

The compressor therefore discharges at 543.4 K and the turbine at 662.5 K. By the steady-flow first law the turbine delivers $c_p(1200 - 662.5) = 540.2$ kJ/kg of work, while compression requires $c_p(543.4 - 300) = 244.6$ kJ/kg. The net output is 295.6 kJ/kg. Dividing by the heat input $c_p(1200 - 543.4) = 659.9$ kJ/kg gives an ideal thermal efficiency of 44.8%. The exhaust is hotter than the compressed air by about 119 K, so those two streams offer an opportunity for recuperation before further heat is supplied.

Exercises

1. Size the exchanger at fixed flow

Keep both original capacity rates and U fixed. What area gives $\varepsilon = 0.80$? How long would the 2 mm wide wall need to be?

▼ **Show the sizing calculation**

Inverting the effectiveness formula gives

$$\mathrm{NTU} = \ln[(1 - \varepsilon C_r)/(1 - \varepsilon)]/(1 - C_r) = \ln(2.6)/0.4 = 2.389.$$

Then

$$A = \mathrm{NTU}\, C_h/U = 1.43 \times 10^{-4}\,\mathrm{m}^2,$$
$$L = A/w = 71.7\,\mathrm{mm}.$$

The extra area supplies diminishing gains as the end temperatures approach one another.

---

2. Check the second law

For the original insulated exchanger assume negligible pressure drops and incompressible streams. Use the logarithmic entropy change of each stream to find the entropy-generation rate.

▼ **Show the second-law balance**

At steady state $\dot{S}_{\mathrm{gen}} = C_h \ln(T_{h,o}/T_{h,i}) + C_c \ln(T_{c,o}/T_{c,i})$. The hot stream loses entropy and the cold stream gains more. Substitution of 309.638 K and 314.217 K gives $6.02 \times 10^{-5}\,\mathrm{W/K}$. The positive result is consistent with transfer through a finite temperature difference.

ONESIMULATION LABORATORY

Explain why greater exchanger effectiveness can give less power

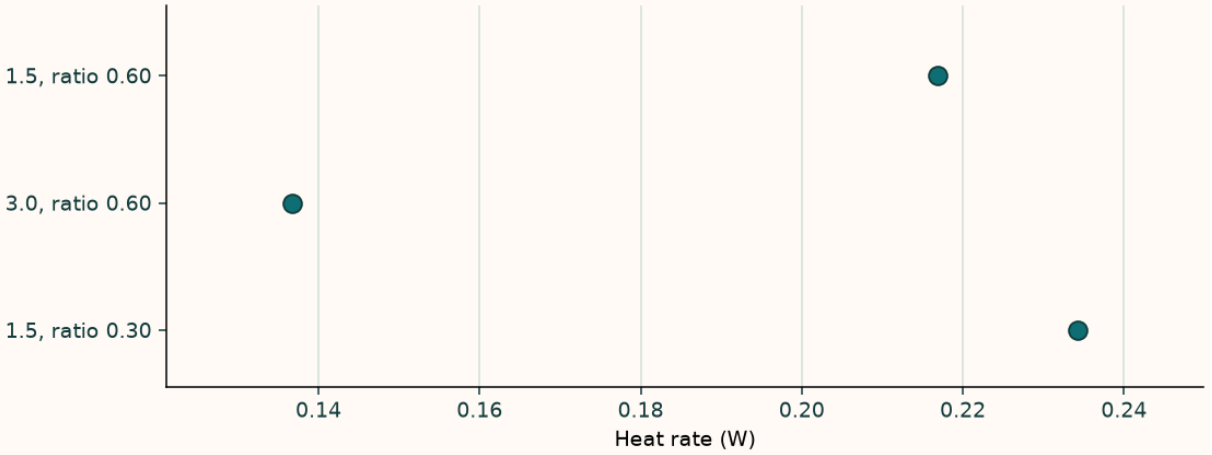
Compare recorded OneSimulation runs on this page. To test your own inputs, download the notebook and model files and run the model in an activated OneSimulation environment.

Run the heat-exchanger example at `ntu=1.5`, `capacity_ratio=0.6` and inlet temperatures 350 and 290 K. Record both cross-section-averaged outlet temperatures and the power lost by the hot stream. In this code the geometry and conductance are fixed. Changing `ntu` changes the computed capacity rates and inlet speeds.

1. Set `ntu=3.0` at the original capacity ratio. Both flows halve. The one-dimensional model predicts an effectiveness of 0.853 and a hot outlet near 298.8 K. The heat rate should fall, despite the greater effectiveness, since there is now less hot water to cool every second.
2. Restore `ntu=1.5` and set `capacity_ratio=0.3`. The hot capacity stays fixed and the cold capacity doubles. Expect a lower cold outlet (about 303.1 K) and an increased heat rate (about 0.235 W). Check both energy balances against the numerical solution.

Computed exchanger outlets and heat released by the hot stream

NTU and capacity ratio	Hot outlet (K)	Cold outlet (K)	Heat rate (W)
1.5, ratio 0.60	309.84	314.20	0.2168
3.0, ratio 0.60	299.36	320.79	0.1367
1.5, ratio 0.30	306.62	303.11	0.2343



Doubling NTU here lowers transferred heat from 0.2168 to 0.1367 W because the capacity rates fall. Lowering only the capacity ratio at the original NTU increases the heat rate.

As a numerical balance check, also calculate the heat gained by the cold stream. The two outlet-based estimates differ by about 0.5% in the baseline and 1.4% at NTU 3. Refining the mesh lets you test how much of this discrepancy comes from the spatial approximation.

View the OneSimulation counterflow model and its temperature profiles.

Download experiment files (ZIP)

Extract the files together and open `counterflow_heat_exchanger_ntu.ipynb` in a local OneSimulation environment.

Interactive reference for the OneSimulation laboratory

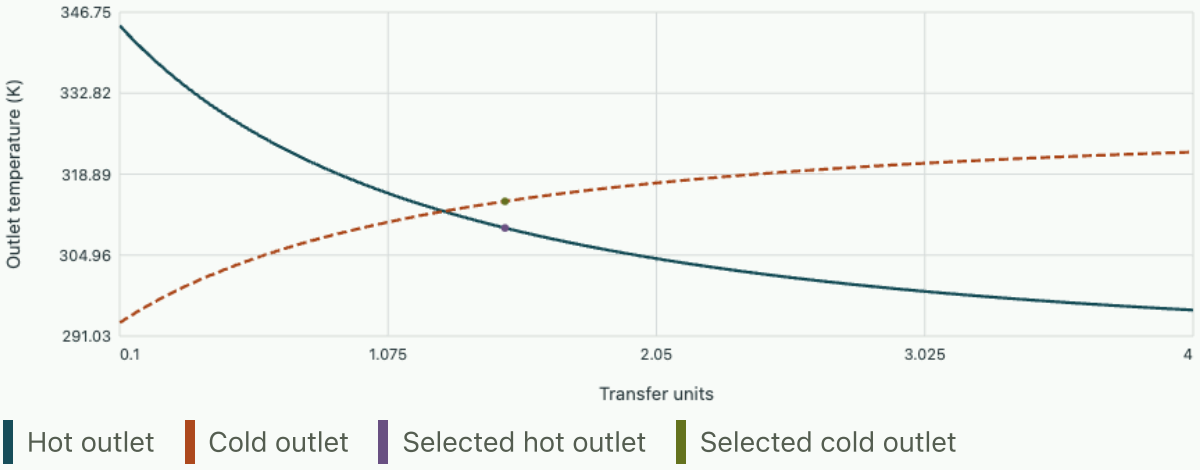
Change capacity rates through NTU at fixed conductance and check both outlet temperatures. The browser evaluates the1D counterflow energy-balance reference.

Open the live laboratory reference

Try the calculation

NTU **1.5**

Hot / cold capacity rate **0.6**



Selected heat transfer 0.2180 W; hot capacity 0.00540 W/K. Fixed UA0.0081 W/K means increasing NTU reduces flow. This1D energy-balance reference can be compared with the finite-dimensional OneSimulation model below.

Problem sheet 6 · Recover an exchanger capacity rate from outlet data

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.

Biot number

$$\mathrm{Bi} = hL_c/k$$

Here h is a surface heat-transfer coefficient and k is conductivity. A small Biot number supports a nearly uniform specimen temperature for the specified boundary condition.

Derivation · Thermal diffusion and transient surface cooling

Entropy generation

$$\dot{S}_{\mathrm{gen}} = \dot{S}_{\mathrm{system}} - \sum \dot{Q}_j/T_j - \dot{S}_{\mathrm{mass,in}} + \dot{S}_{\mathrm{mass,out}}$$

Use boundary temperatures at which heat actually crosses. All flows, storage and units must refer to the same control volume.

Derivation · Entropy, irreversibility and useful-work loss

Exergy destruction

$$\dot{X}_{\mathrm{dest}} = T_0 \dot{S}_{\mathrm{gen}}$$

The environmental reference temperature enters in kelvin. Irreversibility destroys useful-work potential even when the total energy balance closes.

Derivation · Useful work and exergy destruction

Heat-exchanger effectiveness

$$\epsilon = \dot{Q}/[C_{\min}(T_{h,i} - T_{c,i})]$$

Heat-capacity rate is mass flow times specific heat for a single-phase stream. The smaller rate limits the maximum possible sensible heat transfer.

Derivation · Heat-exchanger effectiveness and finite capacity rates

Specific enthalpy

$$h = u + pv$$

Specific quantities are per unit mass. Enthalpy accounts for internal energy and flow work when mass passes through a control surface.

Derivation · Open systems, enthalpy and steady-flow devices

Vapour quality

$$x = m_v/(m_v + m_l)$$

This mass fraction applies in a two-phase equilibrium region. Mole fraction, mass fraction and fraction of volume occupied by vapour can be very different.

Derivation · Phase equilibrium and steam quality

Further reading

- Trace states around the Rankine and Brayton cycles. Reconstruct the work and heat transfers from the property data before comparing cycle efficiencies. MIT · Thermodynamics and Propulsion
- Extend the heat-exchanger work by separating wall, film and contact resistances. Identify the assumption that allows a single overall heat-transfer coefficient. Lienhard and Lienhard · A Heat Transfer Textbook

Compare specific topics with the university sources