

PART ONE

PROPERTIES OF MATTER AND THERMAL PHYSICS

Matter is not the silent backdrop against which physics happens. Matter has a structure, a temperament, and a set of habits of its own, and the Form Five physicist who learns to listen to those habits will find that an enormous portion of the syllabus suddenly speaks with a single voice. Steel bends and remembers, glass shatters, rubber stretches and warms, railway tracks lengthen in the sun, ice floats while stones sink, and a candle flame transfers heat across an empty room: every one of these is matter telling us something about its molecules. This part of the book is dedicated to learning that language. We will study how solids deform under load, how heat moves through substances and changes their state, and how the molecular picture quietly underwrites all of it.

Chapter 1

ELASTICITY

INTRODUCTION

Kipanga had taken many daladalas in his life, but this particular Tuesday afternoon would later become, in his own private filing system, the day a broken bench taught him advanced physics.

The route was Mwenge to Mbagala. The vehicle was the usual kind, with paint that had been refreshed so many times that you could no longer tell what colour it had started as. Kipanga boarded near the front, paid his fare to the conductor, and squeezed himself onto the long bench seat behind the driver, the one that runs the full width of the vehicle and is engineered, in theory, with a hinged cushion above two long springs.

In practice, only half the bench was still working. The right half was already occupied by a serious-looking woman with a basket of mangoes, who did not look like somebody in the mood to share bench physics. So Kipanga took the left half.

He sat. There was a quiet metallic sigh from beneath the cushion, and then nothing. The cushion did not move. It did not rise up to meet him, did not press back against him, did not behave in any way like a bench should behave. It was flat. Permanently, philosophically, unconditionally flat.

He bounced experimentally. The cushion did not bounce. Kipanga was sitting, essentially, on a metal pancake.

Across the aisle, a small boy of perhaps eight years old was on the bench opposite, swinging his legs and bouncing happily with every pothole the daladala hit, which was many. That bench, clearly, still had opinions. It compressed when the boy came down, and it pushed him back up again. Two benches. Same daladala. Presumably the same factory, the same kind of spring underneath. One was alive and one was dead.

The daladala arrived at Ubungo bus stop.

New passengers poured in. A man with a roll of foam mattresses squeezed past the conductor. A woman with a chicken in a box. Three secondary school students in uniform. Two office workers who had clearly missed the earlier bus. By the time the doors finally closed, every seat was full, the aisle was full, the space behind the driver was full, and several people were standing in positions the daladala engineers had probably never imagined.

One of those people was **Kipute**.

She found no seat, and ended up standing right next to Kipanga, one hand gripping the overhead rail, her school bag wedged awkwardly against the driver's seat, her physics notebook somehow still open in her free hand, because of course it was. From where she was standing, she could look straight down at the top of Kipanga's head.

"You look like a man with a question," she said.

Kipanga looked up. *"Kipute! When did you get on?"*

"Ubungo. About thirty seconds ago. You were staring at your bench as if it had insulted your mother."

"This bench is broken," said Kipanga. "But that one isn't." He pointed across the aisle at the bouncing boy.

"It is not broken," said Kipute. "It has been overloaded."

"What is the difference?"

"The difference is the chapter we start tomorrow."

"What chapter?"

"Elasticity. Mr. Akilikubwa has been threatening it for two weeks. He says it is the chapter where physics finally tells the truth about why some things spring back and others give up."

The daladala hit a pothole. Every standing passenger swayed as one body. Kipute braced herself against the overhead rail. Kipanga, sitting on his dead bench, felt absolutely nothing because a pancake cannot bounce.

“Why?” he said, with the kind of urgency only a bad bench can provide. “Why does one bench remember its shape and the other one forgets? They are both made of metal. They were both pressed by the same kinds of bottoms. Why does one come back and the other one stay flat?”

Kipute smiled, and closed her notebook, because she knew when an answer was going to be more useful if it came from Mr. Akilikubwa rather than from her.

“Tomorrow,” she said, looking down at him from her standing position. “Bring the question.”

The next morning, Mr. Akilikubwa walked into Form Five Physics carrying two short metal springs. He had cut them from identical bicycle pumps the previous evening. One was neatly coiled and looked ready for work. The other had clearly been stretched far beyond its limit and now hung from his fingers like a piece of dead rope. He placed both on the demonstration table and waited for the class to settle.

Kipanga’s hand went up before the bell had finished ringing.

“Sir, I have a question from a daladala.”

“The best questions usually are.”

Kipanga described the two benches. The class laughed in the way Tanzanian students always laugh at daladala stories, because every one of them had a similar memory waiting somewhere in the back of their mind. Mr. Akilikubwa let the laughter run its full course; he was never in a hurry to interrupt laughter, and then he picked up the two springs and held them up side by side.

“Two springs,” he said. “Same metal. Same factory. Same length, originally. One of them I bent gently and let go, and it came back. The other I stretched until it could not return, and it stayed stretched. The metal did not break. Every iron atom that was in this spring yesterday is still in it today. And yet one of them is alive, and one of them is dead.”

He set the springs down.

“This is the chapter where physics finally tells you the truth about what is happening inside that metal. We are going to climb all the way down to the level of individual atoms, look at the forces between them, and from those forces we are going to derive every elastic thing in your syllabus. Hooke’s law. Young’s modulus. The reason a glass rod shatters and a steel wire bends. The reason a rubber band gets warm when you stretch and release it many times. The reason railway tracks have gaps. The reason an artery is stretchy and a tendon is not. All of it, from one picture.”

He turned to the board and wrote four words in deliberate capitals: THE INTERMOLECULAR POTENTIAL ENERGY CURVE.

“By the end of this chapter, when somebody hands you a piece of dead spring on a daladala bench, you will not just shake your head and accept it as fate. You will know exactly which atoms moved past which other atoms, and why they are never coming back. Welcome to Chapter 1.”

Kipanga, sitting in the front row, wrote a note in the margin of his exercise book that read: *“Daladala bench; atoms moved.”* He underlined it twice.

INTERMOLECULAR FORCES AND THE POTENTIAL ENERGY CURVE

Mr. Akilikubwa’s promise was bold, and he is about to make it good. Every elastic phenomenon in this chapter; every spring constant, every modulus, every thermal stress, every reason a daladala bench either remembers its shape or forgets it, flows from a single curve. We will draw that curve once, label it carefully, and then return to it again and again as the chapter unfolds. Every time some new elastic behaviour appears in this book, we will discover that it was already living, quietly, somewhere on this curve.

Forces between molecules

Two molecules in a solid do not just sit beside each other passively. They interact. Between every pair of neighbouring molecules there is a force that depends only on how far apart they are. This is the central physical fact on which everything else in this chapter is built, and it deserves to be stated slowly.

When the molecules are far apart, the force between them is attractive: each molecule pulls the other gently towards it. This is the force that holds a solid together at all. Without it, every block of metal in the world would simply fall apart into a cloud of atoms.

When the molecules are pushed too close together, however, the force becomes strongly repulsive: they shove each other away with great vigour. This repulsion is what stops a solid from being compressed indefinitely. It is the reason you cannot squeeze a steel ball into a smaller steel ball no matter how hard you press.

At one particular separation, and one only, the attractive and repulsive forces exactly cancel and the net force is zero. We give this separation a name and a symbol. We call it the **equilibrium separation**, and we write it r_0 . It can be defined as:

The distance between two neighbouring molecules at which the attractive and repulsive intermolecular forces exactly cancel, so that the net force between them is zero and their potential energy is at its minimum.

A pair of bonded molecules will always settle, if left alone, at this separation. It is where they are most comfortable, and where they will return if disturbed gently.

From force to potential energy

It is often more useful to think about the energy of a pair of molecules than about the force directly, because energy is a single number while force has a direction. The two descriptions are completely equivalent: from one we can always get the other. The potential energy of a pair of molecules separated by a distance r , which we shall write $U(r)$, is the work done by the intermolecular force as the molecules are brought from infinity to that separation. Since the molecules attract each other from afar, the force does positive work on them as they approach, and the potential energy decreases. Since they repel each other strongly when very close, the potential energy rises sharply at small separations. Somewhere in between, the energy reaches its lowest possible value. That lowest value sits exactly at r_0 , because that is where the force is zero and there is nothing left to push the energy down further.

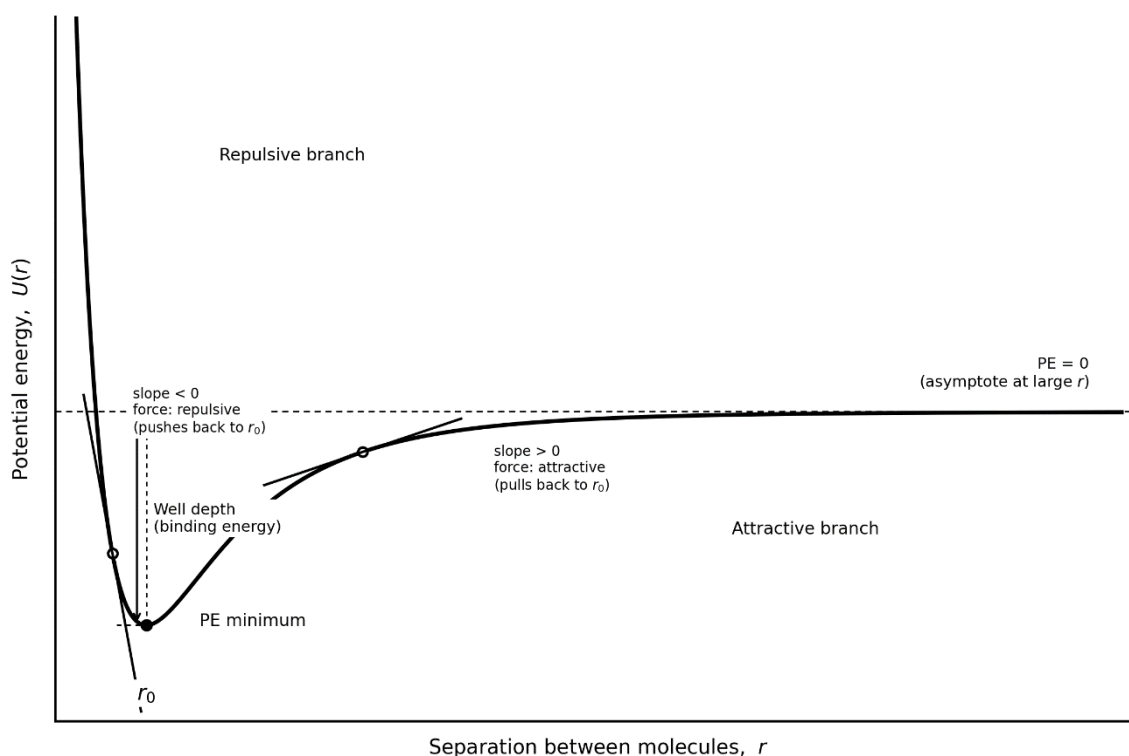


Figure: The intermolecular potential energy $U(r)$ for a pair of bonded molecules as a function of their separation r . The curve reaches its minimum at the equilibrium separation r_0 , where the attractive and repulsive forces exactly balance and the net force is zero. The depth of the well below the horizontal asymptote at $U = 0$ is the binding energy of the pair. At separations smaller than r_0 the steep repulsive branch dominates, and at separations larger than r_0 the gentler attractive branch rises towards the asymptote, which represents two free molecules infinitely far apart. The slope of the curve at any point gives the magnitude and direction of the force on the molecules, and that force always points back towards r_0 .

Reading the curve

This curve is one of the most important pictures in the entire book. Every feature of it has a physical meaning, and we will now go through them one by one. Spend a few minutes with the diagram before reading on;

almost everything that follows in this chapter is just a careful re-description of something already visible there.

The deep well at the centre. The curve has a clear minimum at the equilibrium separation. The depth of this well, measured from the dashed asymptote down to the bottom, is called the binding energy of the pair of molecules. It is the amount of energy you would have to supply to pull the two molecules apart from rest at equilibrium and carry them off to infinity. A deeper well means a more strongly bound pair of molecules, and therefore a more rigid material. A shallower well means a softer, more easily deformed material.

The repulsive branch on the left. To the left of the minimum, the curve climbs steeply and almost vertically. This is the regime in which the molecules are being pushed unwillingly into each other. Even a tiny decrease in separation costs an enormous amount of potential energy, because the slope of the curve is very steep here. A steep slope means a large force, and the direction of the force opposes the compression; the molecules push back hard. This is why solids resist compression so strongly: there is a wall of energy waiting on the left, and it gets steeper the further you try to climb.

The attractive branch on the right. To the right of the minimum, the curve rises again, but much more gently, and approaches the horizontal asymptote at $PE = 0$ from below. This is the regime in which the molecules are being pulled apart. Energy must be supplied to stretch them, but the slope of the curve is shallower here than on the left, so the restoring force is weaker for the same displacement. The curve eventually flattens out at large separations: once the molecules are far enough apart, they no longer feel each other at all. The asymptote represents two free, independent molecules that have completely escaped one another's influence.

Small displacements behave like a spring. Look very carefully at the bottom of the well, in a small region just to the left and right of the minimum. In that small region the curve looks almost like a parabola; a smooth, symmetrical U-shape. This is no accident. Any smooth curve, when viewed close enough to its minimum, looks like a parabola, and the potential energy of a parabola corresponds to a force that is proportional to displacement. In other words: *for small displacements from equilibrium, the restoring force on a pair of molecules is directly proportional to how far they have been displaced.* This is exactly the statement of **Hooke's law**. We have just discovered the molecular origin of the most famous law in elasticity, and we have not written a single equation to do it. Later, we will give the law its proper algebraic form; for now it is enough to know that Hooke's law is not a separate principle imposed from outside. It is the bottom of the intermolecular well, viewed from close up.

Large displacements do not behave like a spring. Now look further out from the minimum, on either side. The curve is no longer a parabola. The left branch climbs faster than a parabola; the right branch climbs slower. The further you stretch or compress, the worse Hooke's law becomes as an approximation. This is why every spring in the world has a limit beyond which it stops behaving like a spring, and why Kipanga's daladala bench, having been pushed so far to the right of the well that it crested the slope and came down on the far side, never returned to its original shape.

The well is not symmetric. There is one more feature of the curve that will become very important later, on thermal expansion. Notice that the well is not symmetric: the repulsive wall on the left is much steeper than the attractive slope on the right. This asymmetry means that when a molecule vibrates back and forth in the well, it spends slightly more of its time on the right (the attractive side) than on the left. As the temperature rises and the vibrations become more energetic, this lopsidedness becomes more pronounced, and the average separation drifts to the right. The molecules move further apart on average, and the material expands. Thermal expansion, which has puzzled students for centuries, turns out to be nothing more than the asymmetry of the well, made visible. We will return to this in the thermal expansion section with the curve in front of us.

One curve, three classes of behaviour

The shape of this same curve, drawn for different materials, explains why those materials behave so differently. The picture is universal; the depth and width of the well are not. Three contrasting cases are worth carrying with you through the rest of the chapter.

1) A deep, narrow well (applies for a rigid solid such as steel or diamond)

The well is so deep that it takes an enormous amount of energy to pull the molecules apart, and so steep on both sides that even a small displacement from equilibrium produces a huge restoring force. Materials of this kind have very high binding energies, very high stiffness, and very high melting points. They resist both compression and extension fiercely.

2) A shallower, broader well (applies for a softer solid such as lead or copper)

The well is still a well, so the material still holds together, but the slopes are gentler. The same displacement from equilibrium produces a smaller restoring force. The material is more easily stretched, more easily compressed, more easily dented, and melts at a lower temperature.

3) No bound state at all (applies for a gas)

In a gas, the molecules are too far apart and moving too fast for the attractive part of the curve to capture them. They are not sitting at the bottom of any well; they are flying past each other out at the flat asymptote on the right. The gas as a whole feels essentially no restoring force when you try to stretch it, because there is nothing pulling its molecules back. This is exactly why a gas cannot resist being pulled apart: there is no spring to restore it. But push the molecules of a gas closer together, and they start to climb the repulsive wall on the left, and they push back. This is why a gas can be compressed against a force (think of pumping a bicycle tyre, even though it offers no resistance to expansion).

We now have the picture. Two molecules sit in a well, comfortable at the bottom, restored gently when nudged a little, restored more roughly when nudged further, and finally unable to return at all when nudged past the edge. Everything else in this chapter: the algebra of stress and strain, the three moduli, the energy stored in a stretched wire, the failure of an overloaded rail, the heat of a snapping rubber band, the thermal stress in a clamped composite rod, is the same picture, dressed in different clothes. Before we put any of those clothes on, however, we need a little practice reading the curve out loud. The two examples that follow are not calculations. They are exercises in seeing.

BINDER Example 1

Use the intermolecular potential energy curve to explain why a solid resists both compression and extension, while a gas resists only compression.

Solution

In a solid, the molecules sit close to the equilibrium separation r_0 at the bottom of the intermolecular potential energy well. Any attempt to compress the solid pushes neighbouring molecules to the left of the minimum, where the curve climbs the steep repulsive branch and produces a large restoring force pushing them apart again. Any attempt to stretch the solid pulls the molecules to the right of the minimum, onto the attractive branch, where the curve again rises and produces a restoring force pulling them back together. Either way, displacing the molecules from r_0 costs energy, and the energy gradient gives a restoring force. A solid therefore resists both compression and extension.

In a gas, the molecules are not sitting in the well at all. They are far out on the flat asymptote on the right, where the curve is essentially horizontal. Pulling them further apart produces no change in potential energy, and therefore no restoring force as there is nothing to pull the molecules of a gas back together once they have been separated. A gas offers no resistance to being stretched. Pushing the molecules of a gas closer together, however, eventually drives them down the asymptote and onto the rising part of the curve, where the slope is no longer zero and a restoring force appears. The gas resists compression.

The asymmetry of behaviour between solids and gases is therefore not a separate fact about each state of matter. It is one fact about the same curve, read in two different places.

Making Sense of the Answer: *The argument never invoked any special property of “solids” or “gases” as such. It used only the shape of the intermolecular potential energy curve and the rule that force is minus the slope of the energy. The same logic, applied at different points along the curve, gave the behaviour of two different states of matter. This is the kind of unification physicists live for.*

Think Like a Physicist: *Whenever a question asks you to compare the elastic behaviour of two materials or two states of matter, your first move should be to draw the curve and ask, “which part of the curve are the molecules sitting on?” Almost every qualitative question in this chapter can be answered by pointing at the right part of the curve. Reach for the picture before reaching for the words.*

REAL Example 2

Kipute is sitting in the front row, frowning at the curve on the board. After a long pause, she raises her hand.

Kipute: *“Sir, you said that molecules attract each other when they are far apart. But the air in this room is made of molecules, and they are far apart. Why doesn’t all the air collapse into a puddle on the floor? If they attract each other, surely they should pull themselves into a tiny ball at my feet.”*

Mr. Akilikubwa: “That is exactly the question this curve was waiting for somebody to ask. The answer is on the diagram; you just have to look at it the right way.”

Use the intermolecular potential energy curve, together with the kinetic energy of the molecules, to explain why the air in the room does not collapse onto the floor under its own intermolecular attraction.

Solution

Two ingredients matter, and the answer is a comparison between them. The first ingredient is the depth of the well which reflects the binding energy that two air molecules would gain if they fell into the bottom of it. The second ingredient is the kinetic energy each molecule already has by virtue of being at room temperature, the energy of its restless thermal motion.

If the kinetic energy of an air molecule is much smaller than the depth of the well, then any two molecules that drift close enough to feel each other’s attraction will be captured by it. They will fall into the well, lose their freedom, and form a bound pair. Many such pairs together make a liquid or a solid. This is exactly what happens to water vapour when it cools to a temperature at which kinetic energies are comparable to the depth of the water-water well: the vapour condenses into liquid water.

If the kinetic energy is much larger than the depth of the well, however, the situation is completely different. Two molecules that drift towards each other do feel the attractive part of the curve, but the attractive force is far too weak to capture them. They fall a little way into the well, gather a bit of extra speed on the way down, and then sail straight out the other side, still moving fast enough to escape. The well is still there. The attraction is still there. But the molecules have so much kinetic energy that the well looks, from their point of view, like a tiny dimple on the floor of an enormous room. They walk right over it without falling in.

At room temperature, the average kinetic energy of an air molecule is roughly a hundred times larger than the depth of the well between two air molecules. The attraction exists, but it is completely overpowered by the thermal motion. The air therefore stays as a gas, with its molecules flying past each other on the flat asymptote of the curve, never bound, never collapsing. Cool the air enough, say down to about -196°C for nitrogen, and the kinetic energy will fall below the well depth, and at that point the air will indeed condense into a puddle on the floor. Liquid nitrogen is exactly that puddle.

Making Sense of the Answer: *The molecules of a gas are not free because they have no attraction. They are free because they have so much kinetic energy that the attraction is, for all practical purposes, invisible. Whenever you want to know whether a substance will be a solid, a liquid, or a gas, the question to ask is not “is there a force?” (there is always a force). The question to ask is “is the kinetic energy big enough to overcome (escape) it?”*

Think Like a Physicist: *Two energies, compared. Almost every question in physics about whether something will stay bound or escape; a satellite in orbit, an electron in an atom, a molecule in a liquid, a footballer trying to leave the pitch, reduces to the same comparison: the depth of the well versus the kinetic energy of the thing trying to escape it. Once you start noticing this pattern, you will see it in chapter after chapter for the rest of your physics life.*

The curve is now in our hands, and we can read it. The next section takes the bottom of the well and gives it the algebraic dress that engineers and examination boards prefer. We will define the two quantities: **stress** and **strain**, that physicists use to describe loaded materials, and we will see that they are nothing more than careful, fair ways of measuring how far we have pushed the molecules from r_0 , and how hard we had to push to do it.

STRESS AND STRAIN

At the bottom of the intermolecular potential energy well introduced in the previous section, the restoring force on a pair of molecules is proportional to how far they have been displaced from their equilibrium separation r_0 . That is the molecular truth. We now have to dress that truth in the algebra of engineering, because in the laboratory we never load a single pair of molecules. We load whole wires, whole rods, whole rails, and whole bones, all of which contain countless molecular pairs being pulled or pushed at the same time. Two questions immediately arise. *First, how hard are we pushing? Second, how far have we gone?* The answers to these two questions are the two quantities at the heart of every elasticity problem you will ever solve. They are **stress** and **strain**.

Both quantities have to be defined in a way that is fair across wires of different thicknesses, different lengths, and different materials. In physics, the word “fair” almost always means “as a ratio”, and we are about to see why.

A small note before we begin. Readers who paid careful attention to the chapter on dimensional analysis in *Foundations of Physics* will remember that strain was listed there as one of the dimensionless quantities, alongside angles in radians. That listing was a preview of what we are about to define. The same word “elastic” also appeared in *Foundations of Physics*, in the chapter on collisions, but its meaning there was different. An elastic collision is one in which kinetic energy is conserved, whereas in this chapter **elastic** refers to *the ability of a material to return to its original shape after a load is removed*. Same word, two different physical uses, and the context will always make clear which one is meant.

Stress: how hard the molecules are being pushed

Imagine two steel wires of the same material and the same length, hanging vertically from a beam. From each wire we hang a 20N load. One wire is thick, with about the diameter of a ballpoint pen refill. The other is thin, with about the diameter of a strand of hair. The same 20N pulls on both. But the load reaches a very different number of molecules in each case. In the thick wire, the 20N is shared between an enormous number of molecular bonds across a wide cross section, and each bond carries only a tiny share. In the thin wire, the same 20N is shared between far fewer bonds, and each one of them is being pulled much harder. The thin wire feels the load more intensely than the thick one, even though the total force is identical in the two cases.

The total force, on its own, is therefore not a good measure of how hard the molecules are being pushed. What we really want to measure is the force per molecule in the cross section, which is the same as the force per unit area of that cross section. We give this ratio a name and a symbol. We call it **tensile stress**, and we write it σ :

$$\sigma = \frac{F}{A}$$

where F is the magnitude of the force pulling along the wire, and A is the cross-sectional area of the wire measured perpendicular to the direction of the force.

By definition, **tensile stress** is the tensile force acting on a body divided by the cross-sectional area of the body measured perpendicular to that force. Its symbol is σ and its SI unit is the pascal (Pa), where 1Pa is equal to 1 N/m².

The pascal is a small unit. A wire being loaded hard enough to be of any practical interest is almost always under a stress of millions of pascals, so the megapascal (1 MPa = 10⁶ Pa) and the gigapascal (1 GPa = 10⁹ Pa) appear in almost every numerical answer in this chapter. Always be ready to convert between pascals, megapascals, and gigapascals.

Stress is not a vector in the everyday sense, because it does not have a single direction, but it does have a sign convention.

- When the force is **pulling the material apart**, the stress is called tensile and is taken as **positive**.
- When the force is **squeezing the material together**, the stress is called compressive and is taken as **negative**.

By definition; **compressive stress** is the stress that arises when the applied force is pushing the material together rather than pulling it apart

The same algebraic formula $\sigma = \frac{F}{A}$ applies to both, with F carrying the appropriate sign. In this chapter we shall be working mostly with tensile loads and shall write everything as positive, but the convention is worth holding on to for the day when a railway track or a roof beam is in compression.

Tensile stress and compressive stress, taken together, are sometimes called **linear stress** or **longitudinal stress**, because in both of them the force is directed along a single line perpendicular to the cross section of the body. Later in this chapter we shall meet two other kinds of stress that act in different geometric ways, namely shear stress, in which the force is parallel to the surface rather than perpendicular to it, and bulk stress, in which the force is a pressure squeezing the body equally from every direction at once. For now, every stress in our calculations will be of the linear kind, which is what most of the syllabus and most examination questions are concerned with.

Cross-sectional area of a circular wire

Almost every wire in this book has a circular cross section, and the area of a circle is needed in nearly every numerical example. The two equally useful forms are:

$$A = \pi r^2 = \frac{\pi d^2}{4}$$

where r is the radius and d is the diameter. The diameter form is the one to use whenever an examination question quotes a diameter directly, because it removes one division and one chance to make a mistake. The radius form is the one to use whenever the question quotes a radius. Always convert millimetres to metres before substituting, because the SI unit of stress is built on the metre and not on the millimetre, and a single forgotten conversion will produce an answer that is wrong by a factor of a million.

Strain: how far the molecules have moved

Now consider the second question. The 20N load has produced an extension. *How far has the wire actually stretched?* Suppose it has lengthened by half a millimetre. Is half a millimetre a lot, or is it a little? The answer depends entirely on how long the wire was to begin with. Half a millimetre on a wire that started two metres long is a feather of an extension and is almost nothing. Half a millimetre on a wire that started two millimetres long is a strain of one quarter, which is far beyond what any metal could survive, and is a catastrophe. The absolute extension on its own tells us very little. What we really want to know is the fractional extension, which is the extension expressed as a ratio of the original length. We give this ratio its own name and symbol. We call it **tensile strain**, and we write it ϵ :

$$\epsilon = \frac{\Delta L}{L}$$

where L is the original (unloaded) length of the wire and ΔL is the change in length produced by the load.

By definition; **tensile strain** is the change in length of a body divided by its original length. Its symbol is ϵ . Because it is a length divided by a length, strain is a pure number with no unit.

Strain is therefore the dimensionless quantity you have already met under that name in the dimensional analysis discussion mentioned earlier. A bare number like 0.00025 is a perfectly good strain. It is also common to express strain as a percentage by multiplying by a hundred, so the same value would be reported as 0.025%, which sometimes makes it easier to picture how much the wire has actually changed.

The same sign convention applies as for stress. When the wire has lengthened, ΔL is positive and the strain is called tensile. When the wire has shortened under compression, ΔL is negative and the strain is called compressive. Same formula, signed inputs.

So **compressive strain** is the strain that arises when the change in length of a body is a decrease rather than an increase.

Tensile strain and compressive strain share an important feature. In both cases, the change we are measuring is a change in length along a single direction, the direction of the applied force. For this reason, the two are grouped together under a single umbrella name. They are together called **linear strain**, because the deformation happens along one line.

So **linear strain** is the change in length of a body, in a chosen direction, divided by its original length in that direction. Both tensile strain and compressive strain are the two cases of linear strain, the first when the body lengthens and the second when the body shortens.

With the language of stress and strain in our hands, we are now ready to do some calculations. The three examples that follow are deliberately straightforward. They build the muscle memory of converting millimetres to metres, computing circular cross sections, dividing forces by areas, and recognising stress and strain when they appear in the wild.

BINDER Example 3

A steel wire of radius 0.5mm hangs from a fixed support and carries a load of 20N at its lower end. Calculate the tensile stress in the wire.

Solution

Converting the radius to metres:

$$r = 0.5\text{mm} = 0.5 \times 10^{-3} \text{ m} = 5 \times 10^{-4} \text{ m}$$

The cross-sectional area of the wire is:

$$A = \pi r^2 = \pi \times (5 \times 10^{-4} \text{ m})^2 = 7.854 \times 10^{-7} \text{ m}^2$$

The tensile stress is the force divided by this area:

$$\sigma = \frac{F}{A} = \frac{20\text{N}}{7.854 \times 10^{-7} \text{ m}^2}$$

$$\sigma = 2.546 \times 10^7 \text{ Pa} \approx 25.5 \text{ MPa}$$

The tensile stress in the wire is approximately 25.5MPa.

Making Sense of the Answer: Twenty-five megapascals sounds like an enormous number, but for steel it is comfortably small. The elastic limit of structural steel is about 250MPa, ten times this value, and the breaking stress is several times higher again. The wire is therefore nowhere near in danger, and is doing exactly the kind of routine work for which steel was invented. The fact that the stress is millions of pascals is not a sign of trouble. It is the natural scale on which engineering metals operate.

Think Like a Physicist: Whenever a question gives you a radius or a diameter in millimetres, your very first move is to convert it to metres. The radius enters the area as a square, so a single missing factor of a thousand becomes a missing factor of a million in the answer. Convert first, substitute second, and check the units of the final result.

BINDER Example 4

A copper rod of original length 2.0m extends by 0.5mm when a tensile load is applied to it. Calculate the tensile strain in the rod and express it as a percentage.

Solution

Converting the extension to metres so that the units cancel cleanly:

$$\Delta L = 0.5\text{mm} = 0.5 \times 10^{-3} \text{ m} = 5 \times 10^{-4} \text{ m}$$

The tensile strain is the fractional extension:

$$\epsilon = \frac{\Delta L}{L} = \frac{5 \times 10^{-4} \text{ m}}{2.0 \text{ m}}$$

$$\epsilon = 2.5 \times 10^{-4}$$

Expressed as a percentage:

$$\epsilon (\%) = 2.5 \times 10^{-4} \times 100\% = 0.025\%$$

The strain in the rod is 2.5×10^{-4} , or 0.025%.

Making Sense of the Answer: A strain of 0.025% means the rod has lengthened by twenty-five parts per hundred thousand. The metres in the numerator and the metres in the denominator have cancelled completely, leaving a bare number with no unit, just as the definition demands. Strains in the elastic range of common metals are almost always smaller than 0.1%, which is the source of one of the perpetual frustrations of experimental elasticity. The quantity you most want to measure is always smaller than the quantity you can comfortably see. A later part of this chapter, on the experimental determination of Young's modulus, will return to this problem and show how Searle's apparatus is built specifically to defeat it.

Think Like a Physicist: When you compute a strain, the very first thing to check is that the answer has no units. If you find a stray metre or millimetre attached to your strain, you have either forgotten to convert one of the lengths or used inconsistent units. Strain is, by definition, a pure number, and that unit-cancellation is your first line of defence against arithmetic mistakes.

REAL Example 5

A piano tuner stretches three steel strings between the same two pegs of a piano frame. The three strings are made of identical steel and are tightened to exactly the same tension, but their thicknesses are different. One

is thick, one is medium, and one is thin. As the tuner raises the tension, the thinnest string is always the first to snap. Explain, using the definition of stress, why the thinnest string fails first even though all three strings carry the same applied force.

Solution

Steel breaks when the stress inside it reaches a definite value called the breaking stress. This breaking stress is a property of steel itself and not of any particular string. Every piece of the same kind of steel will fail at the same value of σ . The question is therefore not which string carries the largest force, because all three strings carry the same force by the conditions of the problem, but which string reaches the breaking stress first.

From the definition of stress,

$$\sigma = \frac{F}{A}$$

we see that for a fixed force F , the stress depends on the cross-sectional area A . A thick string has a large area, so the same force is shared between many molecules and the stress on each molecule is small. A thin string has a small area, so the same force is shared between far fewer molecules and the stress on each molecule is large. The thinner the string, the larger the stress, even though the total force pulling it has not changed at all.

As the tuner gradually increases the tension, the stress in all three strings rises together. But it rises fastest in the thin string, because the thin string starts at the highest stress and adds the largest stress increment for any given increase in tension. The thin string therefore reaches the breaking stress of steel first, while the medium and thick strings are still well below it. The thin string snaps, the tuner sighs, and physics smiles!

Making Sense of the Answer: Notice that the answer never required a number. It rested entirely on the algebraic shape of the equation $\sigma = \frac{F}{A}$. When the same force is divided by a smaller area, the result is a larger stress. The physics of why thin things break first is exactly the physics of why high heels damage soft floors more than flat shoes do, why a sharp knife cuts more easily than a blunt one, and why an elephant's footprint sinks less into the ground than a stiletto's does. In every case the trick is the same. The same force, concentrated onto a smaller area, becomes a much larger stress.

Think Like a Physicist: Whenever a problem describes "the same force" applied to objects of different shapes, your first instinct should be to compare not the forces but the stresses. The force is what is given to you, and the stress is what the molecules actually feel. Examination questions love to confuse the two, and students who confuse them lose marks they should not be losing.

Stress and strain are now defined and we can compute them. But we still cannot connect them to one another. Knowing the stress in a wire does not yet tell us the strain it will produce, because we have not yet asked the wire how stiff it is. The relationship between stress and strain is governed by a single number that captures the elastic personality of the material, and that number is the subject of the next part of this chapter. We are about to meet Young's modulus, the most quoted material constant in all of engineering.

YOUNG'S MODULUS

Stress and strain are now sitting on the table in front of us, and we know how to calculate each of them. But we still cannot connect them to one another. *If we apply a certain stress to a wire, what strain will we get in return?* The answer depends on what the wire is made of. A steel wire and a rubber band subjected to the same stress will produce wildly different strains, and the reason is nothing more than the difference in stiffness between steel and rubber at the molecular level. To capture that stiffness in a single number, one quantity for each material, we need a new idea. That idea was first written down carefully by the English physicist Thomas **Young** in the year 1807, and it has carried his name ever since.

From the PE curve to a single number

The bottom of the intermolecular potential energy well is a parabola, as we saw in the section on the PE curve. A parabolic well gives a restoring force that is proportional to displacement, which is Hooke's law at the molecular level. What Young's modulus does is translate that molecular fact into the language of stress and strain, so that we can use it for whole wires and whole rods without having to count individual molecules.

The translation is disarmingly simple. *For any material being loaded within the elastic range, the stress is directly proportional to the strain.* That is **Hooke's law**, restated in the new language. The constant of proportionality is what we call **Young's modulus**, and we write it **E** (sometimes denoted by **Y** in older texts):

$$\sigma = E\varepsilon$$

Equivalently, and more usefully for calculations, Young's modulus is the ratio of stress to strain:

$$E = \frac{\sigma}{\varepsilon}$$

By definition; **Young's modulus** of a material is the ratio of the tensile stress acting on the material to the resulting tensile strain, provided the material is within its elastic range. Its symbol is **E** and its SI unit is the pascal (Pa), the same as stress, because strain is dimensionless.

Two consequences of this definition are worth stating immediately, because they are the single largest source of confusion in the whole topic.

First, Young's modulus is a property of the material and not of the particular wire or rod that has been cut from it. A thick steel cable and a thin steel hair have the same Young's modulus, because they are made of the same steel. Their stiffnesses as objects are very different, because one is thick and one is thin, but the stiffness per unit of material is the same. This is exactly the property we were hunting for in the previous section when we insisted that stress and strain be defined as ratios rather than as raw forces and raw extensions. Young's modulus is the reward for that discipline.

Second, the SI unit of Young's modulus is the same as the SI unit of stress, because strain has no unit at all. A pascal of stress divided by a dimensionless strain is still a pascal. In practice, the pascal is far too small for engineering materials. Steel has a Young's modulus of roughly two hundred billion pascals, which is why nearly every number you will see in this section is quoted in gigapascals, where 1GPa is equal to one billion pascals, or 10^9 Pa.

The expanded form for laboratory use

The definition $E = \frac{\sigma}{\varepsilon}$ is compact and elegant, but in the laboratory we almost never measure stress and strain directly. We measure force, area, original length, and extension. Substituting $\sigma = \frac{F}{A}$ and $\varepsilon = \frac{\Delta L}{L}$ into the definition gives the expanded form:

$$E = \frac{\frac{F}{A}}{\frac{\Delta L}{L}} = \frac{FL}{A\Delta L}$$

Every one of its four inputs (force, original length, cross-sectional area, and extension) is something you can read off a laboratory bench.

The same equation can be rearranged to give each of its variables in turn. The four useful rearrangements are:

$$F = \frac{EA\Delta L}{L}$$

$$\Delta L = \frac{FL}{EA}$$

$$A = \frac{FL}{E\Delta L}$$

$$L = \frac{EA\Delta L}{F}$$

You will use the first two of these constantly. The first one tells you the force needed to produce a chosen extension in a wire of known material and known dimensions. The second tells you the extension produced in a wire of known material and known dimensions by a given force. Between them, they solve most of the numerical elasticity problems you will ever meet.

Young's modulus values for common materials

Young's modulus values span an enormous range, from gigapascals for the hardest engineering metals down to megapascals for rubber and soft tissue, a factor of roughly one hundred thousand from the stiffest material in the table to the softest. The table below lists values you will meet throughout this chapter. You are not expected to memorise the numbers, but you should develop a feel for their order of magnitude, because a great many examination questions can be sanity-checked simply by noticing whether the answer you computed sits in the right neighbourhood for the material concerned.

Material	Young's modulus E
Diamond	1100GPa
Steel	200GPa
Copper	120GPa
Brass	100GPa
Aluminium	70GPa
Glass	70GPa
Bone (compact)	20GPa
Concrete	25GPa
Wood (along grain)	10GPa
Tendon	1.5GPa
Rubber	0.005 to 0.1GPa

Three patterns in the table are worth noticing before we move on.

- 1) The hardest engineering metals (steel, copper, brass, aluminium) all sit between 70 and 200 gigapascals, and diamond at the top of the list is an outlier with a value several times higher.
- 2) Among the biological tissues, bone is remarkably stiff for a living material, sitting at 20 gigapascals, which places it in the same general neighbourhood as ordinary structural concrete. This is one of the reasons the human skeleton can support the body against gravity for decades. Tendon is softer than bone by a factor of more than ten, which is exactly what you want from a tissue whose job is to stretch slightly under load and store elastic energy.
- 3) Rubber is softer still, by another factor that runs from roughly fifteen for the stiffer rubbers up to several hundred for the softest ones, which is why a rubber band can be stretched to several times its natural length without breaking, something no metal in the table could ever do.

From Young's modulus to the spring constant of a wire

Before we can compute the spring constant of a wire, we need to introduce two ideas that have been quietly waiting in the background of the chapter. The first is the spring constant of an ordinary spring. The second is the law that governs how that spring behaves under load, which is called **Hooke's law**.

Take an ordinary helical spring of the kind found inside a daladala bench, a kitchen weighing scale, or a retractable ballpoint pen. Hang it vertically from a fixed support, attach a small load to its lower end, and measure how much the spring stretches. Now double the load and measure again. The extension also doubles. Triple the load, and the extension triples. Provided you do not load the spring so heavily that it deforms permanently, the extension is always proportional to the applied force. This proportionality is one of the oldest experimental laws in physics. It was first published by the English natural philosopher Robert Hooke in 1678. The law states that:

Provided a spring is not loaded beyond its elastic limit, the extension of the spring from its natural length is directly proportional to the applied force.

Mathematically:

$$\mathbf{F = kx}$$

where F is the force applied to the spring, x is the extension of the spring measured from its natural unloaded length, and k is a positive constant called the **spring constant**.

So the spring constant of a spring *is the constant of proportionality between the applied force and the extension in Hooke's law*. Its SI unit is newtons per metre (N/m). It tells you how much force is needed to stretch the spring by one unit of extension. It is a property of the particular spring, depending on the material it is made of and on the way it has been wound, and not a property of the load placed on it. A stiff spring has a large spring constant, because a large force is needed to produce even a small extension. A soft spring has a small spring constant, because a small force already produces a large extension.

Hooke's law is not a fundamental law of physics in the way Newton's laws are. It is an empirical regularity that holds very well for most materials over a small range of extensions and breaks down for large extensions, large compressions, and any loading that permanently deforms the spring. Inside that small range, however, it is one of the most useful relationships in all of physics, and the spring constant is the single number that captures it for any particular spring.

A wire loaded along its length behaves in exactly the same way as a spring, because a wire is just a very thin, very straight spring made of a single continuous piece of metal. It has a spring constant of its own, and we can compute that spring constant from the Young's modulus of the material it is made of, together with its length and its cross-sectional area.

Starting from the expanded form $E = \frac{FL}{A\Delta L}$, we rearrange for F :

$$F = \frac{EA\Delta L}{L}$$

$$F = \left(\frac{EA}{L}\right)\Delta L$$

Comparing this with Hooke's law in the form $F = k\Delta L$ (where the extension of the wire plays the role of x), we can read off the spring constant of the wire directly. The constant sitting in front of the extension is the wire's own spring constant:

$$\mathbf{k = \frac{EA}{L}}$$

So the spring constant of a wire loaded along its length *is the product of its Young's modulus and its cross-sectional area, divided by its original length*. It tells you how much force is needed to produce a unit extension in that particular wire.

Three features of this formula deserve to be stated out loud:

- 1) Making the wire longer reduces its spring constant ($k \propto \frac{1}{L}$). A 2m wire has half the spring constant of an otherwise identical 1m wire made of the same material. Intuitively this is because the longer wire has more molecules in series between the anchor and the load, and each pair contributes a tiny extension of its own, so a longer stack of such pairs produces a larger total extension for the same force.
- 2) Making the wire thicker increases its spring constant ($k \propto A$). A wire with twice the cross-sectional area has twice the spring constant. Intuitively this is because the thicker wire has more molecular bonds in parallel across its cross section, and each one carries a share of the load, so the load each bond has to carry is smaller and the extension each bond contributes is smaller.
- 3) Young's modulus itself is unchanged by any of this. Cutting a steel wire in half, or splitting it into two parallel halves, does nothing to the stiffness of the steel itself. The spring constants of the resulting pieces are different from the original, because the dimensions have changed, but E is still exactly what it was before. This is the same material-property-versus-object-property distinction we drew when we first defined Young's modulus as the ratio of stress to strain, showing up again in a new place. *The spring constant of a wire is an object property, but the Young's modulus of the metal the wire is made of is a material property*, and they should never be confused.

Stiffness is not the same as strength

Young's modulus measures stiffness, which is a material's resistance to being deformed. It does not measure strength, which is a material's resistance to being broken. These two ideas are constantly confused by students.

The cleanest way to see the distinction is with the Kipanga dialogue that follows, in which Kipanga finally asks the question that has been bothering him since he first saw a glass rod break cleanly on the demonstration table. The dialogue will not stop at the simple stiffness-versus-strength contrast, however. It will push one step further, into a third quantity called toughness, because the everyday experience of dropping a glass bottle and watching it shatter is not really a question about breaking stress at all. It is a question about how much energy a material can absorb before its molecular bonds give way, and that is a slightly more subtle thing than either stiffness or strength on its own.

The stiffness of a material or of an object is its resistance to elastic deformation under load. For a material, stiffness is measured by Young's modulus E . For a wire or rod, stiffness is measured by its spring constant k .

The strength of a material, also called its **breaking stress**, is the maximum stress the material can sustain before it fractures. Its SI unit is the pascal, the same as stress itself. Stiffness and strength are two independent numbers, and a material can rank high on one scale and low on the other.

We shall say much more about strength when we reach the stress-strain curve later in this chapter and see how a material's elastic region, plastic region, and ultimate tensile strength all live on the same graph. For now, the important point is simply that stiffness and strength are two different questions with two different answers, and Young's modulus measures only the first of them.

With the theory now in place, we are ready for the first real numerical work of the chapter.

BINDER Example 6

A steel wire of length 3m and radius 0.6mm hangs from a ceiling and carries a load of 50N at its lower end. Young's modulus for steel is 2×10^{11} Pa. Calculate the extension of the wire.

Solution

Converting the radius to metres:

$$r = 0.6\text{mm} = 6 \times 10^{-4}\text{m}$$

The cross-sectional area of the wire is:

$$A = \pi r^2 = \pi \times (6 \times 10^{-4}\text{m})^2 = 1.131 \times 10^{-6}\text{m}^2$$

Using the rearrangement of Young's modulus that gives the extension:

$$\Delta L = \frac{FL}{EA}$$

$$\Delta L = \frac{50\text{N} \times 3\text{m}}{(2 \times 10^{11}\text{Pa})(1.131 \times 10^{-6}\text{m}^2)} = 6.63 \times 10^{-4}\text{m} \approx 0.66\text{mm}$$

The wire extends by about 0.66mm under the 50N load.

Making Sense of the Answer: A 3m steel wire stretches by about two thirds of a millimetre under a 50N load. That is an extension you would struggle to see with the naked eye, and yet the wire is not doing anything remarkable. It is behaving exactly the way any steel wire of those dimensions would behave under a load of that size. This is why the elastic behaviour of metals is almost invisible in everyday life. The strains are small, the extensions are tiny, and the whole subject would be much harder to study if the deformations were not so well hidden by their own modest size.

Think Like a Physicist: The extension formula has four inputs. Any question that gives you any three of them will let you solve for the fourth. The fastest way to spot which rearrangement to use is to ask which variable is missing from the question, because that is always the variable you need to isolate. In this example the missing variable was the extension, so we used $\Delta L = FL/(EA)$. When the missing variable is F , the rearrangement is $F = EA\Delta L/L$. When the missing variable is E , the rearrangement is $E = FL/(A\Delta L)$. Memorise the expanded form once and you get all four rearrangements for free.

BINDER Example 7

A copper cable of diameter 4mm and length 5m extends by 2.1mm when a load F is applied to it. Young's modulus for copper is 1.2×10^{11} Pa. Calculate the load F .

Solution

Convert the diameter to metres and compute the cross-sectional area:

$$d = 4\text{mm} = 4 \times 10^{-3}\text{m}$$

$$A = \frac{\pi d^2}{4} = \frac{\pi \times (4 \times 10^{-3}\text{m})^2}{4} = 1.257 \times 10^{-5} \text{ m}^2$$

Using the rearrangement that gives the force:

$$F = \frac{EA\Delta L}{L}$$

$$F = \frac{(1.2 \times 10^{11}\text{Pa})(1.257 \times 10^{-5} \text{ m}^2)(2.1 \times 10^{-3}\text{m})}{5\text{m}} = 633.5\text{N} \approx 634\text{N}$$

The load on the cable is approximately 634N.

Making Sense of the Answer: Six hundred and thirty-four newtons is roughly the weight of a 65kg adult standing on the cable. A 5m copper cable of 4mm diameter can easily support that weight while stretching by only 2.1mm, which is less than the thickness of a fingernail. Copper is softer than steel (its Young's modulus is about 60 percent that of steel), so the same cable made of steel would have stretched even less.

Think Like a Physicist: Whenever a question gives you a diameter rather than a radius, use the diameter form of the area formula directly. Squaring a diameter and dividing by four is one operation. Halving the diameter, then squaring, then multiplying by π is three operations, which is three chances to make a small arithmetic mistake. Fewer operations means fewer errors.

REAL Example 8

Kipute is given two wires of unknown material in the physics laboratory. The first wire is 2m long with a diameter of 1mm. When she hangs an 80N load from it, the wire extends by 1.02mm. The second wire is 3m long with a diameter of 0.8mm. When she hangs a 50N load from this second wire, it extends by 2.49mm. Which of the two wires is stiffer, and by what factor?

Solution

The two wires differ in length, diameter, and applied load, so we cannot compare their stiffness directly from the extensions. We must compute Young's modulus for each wire and then compare the two values.

For the first wire, the cross-sectional area is:

$$A_1 = \frac{\pi d^2}{4} = \frac{\pi \times (1 \times 10^{-3}\text{m})^2}{4} = 7.854 \times 10^{-7} \text{ m}^2$$

Young's modulus of the first wire:

$$E_1 = \frac{FL}{A_1\Delta L_1}$$

$$E_1 = \frac{(80\text{N})(2\text{m})}{(7.854 \times 10^{-7} \text{ m}^2)(1.02 \times 10^{-3}\text{m})} = 2.00 \times 10^{11}\text{Pa} \approx 200\text{GPa}$$

For the second wire, the cross-sectional area is:

$$A_2 = \frac{\pi \times (0.8 \times 10^{-3}\text{m})^2}{4} = 5.03 \times 10^{-7} \text{ m}^2$$

Young's modulus of the second wire:

$$E_2 = \frac{(50\text{N})(3\text{m})}{(5.03 \times 10^{-7} \text{ m}^2)(2.49 \times 10^{-3}\text{m})} = 1.20 \times 10^{11}\text{Pa} \approx 120\text{GPa}$$

The first wire is stiffer because its Young's modulus is higher. The ratio of the two moduli is:

$$\frac{E_1}{E_2} = \frac{200\text{GPa}}{120\text{GPa}} \approx 1.67$$

The first wire is about 1.67 times stiffer than the second.

Making Sense of the Answer: The two wires differ in length, in thickness, and in the load they carry, so a direct comparison of their extensions would be meaningless. Only Young's modulus can strip away the geometry and leave a number that belongs to the material itself. The stiffer wire is close to steel at 200GPa and the softer wire is close to copper at 120GPa, which the student can confirm by glancing at the table earlier in this section.

Think Like a Physicist: Young's modulus exists precisely so that wires of different shapes loaded by different forces can still be compared fairly. Any question that asks which of several wires is stiffer, when the wires differ in more than one dimension or load, must be answered by computing Young's modulus for each wire separately. The raw extensions are meaningless on their own because they mix material properties with geometry.

HOT Example 9

A 60N load hangs from a combination of two wires joined end to end. The upper wire is copper, with length 2m and cross-sectional area $1.5 \times 10^{-6} \text{ m}^2$. The lower wire is steel, with length 2m and cross-sectional area $1 \times 10^{-6} \text{ m}^2$. Young's modulus for copper is $1.2 \times 10^{11} \text{ Pa}$, and for steel is $2 \times 10^{11} \text{ Pa}$. Find (a) the tensile stress in each wire, (b) the extension of each wire, and (c) the total extension of the combination.

Solution

When two wires are joined end to end and a single load hangs from the lower wire, the same force is transmitted through both wires. Each wire therefore carries the full 60N, just as every section of a single long rope carries the full weight hanging from it. This is the key physical insight for the problem, and it is the one students most often miss.

(a) Stresses in the two wires:

For the copper wire:

$$\sigma_{\text{Cu}} = \frac{F}{A_{\text{Cu}}} = \frac{60\text{N}}{1.5 \times 10^{-6} \text{ m}^2} = 4 \times 10^7 \text{ Pa} = 40\text{MPa}$$

For the steel wire:

$$\sigma_{\text{St}} = \frac{F}{A_{\text{St}}} = \frac{60\text{N}}{1 \times 10^{-6} \text{ m}^2} = 6 \times 10^7 \text{ Pa} = 60\text{MPa}$$

The stresses are different even though the force is the same, because the cross-sectional areas are different. This is exactly the piano-string logic in the REAL Example 5.

(b) Extensions of the two wires:

For the copper wire, using the extension form of Young's modulus:

$$\begin{aligned} \Delta L_{\text{Cu}} &= \frac{FL}{E_{\text{Cu}}A_{\text{Cu}}} \\ \Delta L_{\text{Cu}} &= \frac{(60\text{N})(2\text{m})}{(1.2 \times 10^{11}\text{Pa})(1.5 \times 10^{-6} \text{ m}^2)} \\ \Delta L_{\text{Cu}} &= 6.67 \times 10^{-4} \text{ m} \approx 0.67\text{mm} \end{aligned}$$

For the steel wire:

$$\begin{aligned} \Delta L_{\text{St}} &= \frac{(60\text{N})(2\text{m})}{(2 \times 10^{11}\text{Pa})(1 \times 10^{-6} \text{ m}^2)} \\ \Delta L_{\text{St}} &= 6 \times 10^{-4} \text{ m} = 0.6\text{mm} \end{aligned}$$

(c) Total extension

The wires are joined end to end, so their extensions add directly:

$$\Delta L_{\text{total}} = \Delta L_{\text{Cu}} + \Delta L_{\text{St}} = 0.67\text{mm} + 0.6\text{mm} = 1.27\text{mm}$$

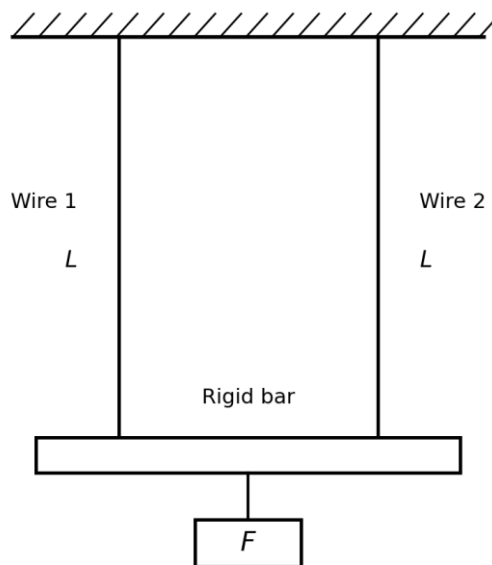
The combined wire extends by approximately 1.27mm under the 60N load.

Making Sense of the Answer: The copper wire stretches slightly more than the steel wire (0.67mm versus 0.6mm), even though the copper is thicker and carries a smaller stress. The reason is that copper's Young's modulus is only about 60 percent of steel's, and that softness more than compensates for the larger cross section. This is a good reminder that in composite problems you cannot predict which piece will stretch more just by looking at the dimensions. You have to do the arithmetic with both the dimensions and the Young's modulus of each material, because both contribute to the answer.

Think Like a Physicist: In any series combination of wires or springs, the load on each element is the same (the whole load passes through each piece one after the other), and the total extension is the sum of the individual extensions. In any parallel combination of wires or springs, the extension is the same (both pieces are forced to stretch by the same amount because they share the same two endpoints), and the total load is the sum of the individual loads. These two rules between them handle nearly every composite problem in elasticity, and they are the exact twins of the rules for resistors in series and parallel that you will meet later when you study electric circuits.

HOT Example 10

A rigid horizontal bar is suspended from a fixed ceiling by two vertical wires attached to its two ends. Both wires have the same length of 2m. The first wire is steel with cross-sectional area $2 \times 10^{-6} \text{ m}^2$ and Young's modulus 200GPa. The second wire is copper with cross-sectional area $3 \times 10^{-6} \text{ m}^2$ and Young's modulus 120GPa. A 100N load is hung from the exact centre of the bar. Find (a) the extension of each wire, (b) the tension in each wire, and (c) confirm that the two tensions add up to the applied load.



Solution

Because the bar is rigid and the two wires have the same length, both wires must stretch by exactly the same amount when the load is applied. The bar cannot tilt, so whatever extension happens in one wire must happen in the other. This is the defining condition of a parallel arrangement of springs or wires: equal extension, and the applied load shared between the two wires in amounts that depend on their individual stiffnesses.

The spring constant of each wire, from the formula $k = \frac{EA}{L}$, is:

$$k_1 = \frac{(200 \times 10^9 \text{ Pa})(2 \times 10^{-6} \text{ m}^2)}{2 \text{ m}} = 2 \times 10^5 \text{ N/m}$$

$$k_2 = \frac{(120 \times 10^9 \text{ Pa})(3 \times 10^{-6} \text{ m}^2)}{2 \text{ m}} = 1.8 \times 10^5 \text{ N/m}$$

(a) Extension of each wire (common extension):

In a parallel arrangement, the two spring constants add to give the effective stiffness of the combination:

$$k_{\text{total}} = k_1 + k_2 = (2 + 1.8) \times 10^5 \text{N/m} = 3.8 \times 10^5 \text{N/m}$$

Using $F = kx$ or $x = \frac{F}{k}$, the common extension is then:

$$x = \frac{F}{k_{\text{total}}} = \frac{100\text{N}}{3.8 \times 10^5 \text{N/m}} = 2.63 \times 10^{-4} \text{m} \approx 0.263 \text{mm}$$

The extension of each wire is 0.263mm.

(b) Tension in each wire:

Applying Hooke's law $F = kx$ to each wire separately:

$$F_1 = k_1 x = (2 \times 10^5 \text{N/m})(2.63 \times 10^{-4} \text{m}) = 52.6 \text{N}$$

$$F_2 = k_2 x = (1.8 \times 10^5 \text{N/m})(2.63 \times 10^{-4} \text{m}) = 47.4 \text{N}$$

The steel wire carries a tension of about 52.6N and the copper wire carries a tension of about 47.4N.

(c) Check that the tensions add up to the load.

$$F_1 + F_2 = 52.6 \text{N} + 47.4 \text{N} = 100 \text{N}$$

The sum of the two tensions equals the applied load, as expected.

Making Sense of the Answer: The steel wire takes the larger share of the load, about 53 percent of it, even though both wires have stretched by exactly the same amount. The reason is that steel is stiffer than copper, and a stiffer wire requires a larger force to produce the same extension. This is the opposite pattern from the series case we just met, where the softer wire stretched more and the stresses were different. In the parallel case, the extensions are forced to be the same and it is the forces that differ. The stiffer element always pulls its weight a bit harder than its softer neighbour, which is why the load-bearing cables in real structures are almost always made of the stiffest practical material the engineer can afford.

Think Like a Physicist: The spring constant of the parallel combination is the sum of the individual spring constants, which is the same addition rule as for conductances in parallel electrical circuits. The series case had extensions adding rather than spring constants adding, which corresponds to resistances adding in series electrical circuits. These parallels between elasticity and electricity are not a coincidence. Both are linear-response theories, where a driving quantity (force or voltage) produces a response (extension or current) proportional to itself, and the combination rules for series and parallel elements follow from the same underlying linearity regardless of the physical context.

Young's modulus is now in your hands. You can compute it from raw laboratory measurements, you can use it to predict the extension of a wire under load, you can use it to compute the spring constant of a wire from its dimensions, and you can tell the difference between stiffness and strength. But so far we have stayed inside the elastic region, where loading and unloading a material returns it to its original shape with no memory of what happened in between. Real materials do not stay elastic forever. Push them hard enough and they begin to deform permanently. Push them harder still and they break. The full journey from gentle loading to final fracture is recorded on a single graph that every engineer in the world has committed to memory, and that graph is the subject of the next part of this chapter. Welcome, in advance, to the stress-strain curve.

THE STRESS-STRAIN CURVE

Young's modulus tells us how much a material deforms under a small load, provided the load is small enough for Hooke's law to apply. But Hooke's law does not apply forever. Pull on a metal wire hard enough and it will eventually stop being linear. Pull harder still and it will begin to deform permanently, so that when you release the load it does not return to its original length. Pull harder still and it will narrow locally, then snap. All of this behaviour, from the gentlest initial stretch to the final fracture, is captured on a single graph that every engineer in the world has committed to memory. The graph is called the stress-strain curve, and it is the subject of this section.

The curve is plotted with stress on the vertical axis and strain on the horizontal axis, using the definitions of both quantities that we established earlier in this chapter. To build the curve experimentally, you take a

standard specimen of the material you want to test, you load it in tension along its length, you measure the extension as the load increases, and you compute the stress and strain at each measurement. Plotting the results gives a curve whose shape depends only on the material, not on the specimen dimensions. This is the same intensive-versus-extensive principle that gave us Young's modulus, applied now across the full range of loading.

The journey along a ductile stress-strain curve

We shall walk through the curve point by point, starting at the origin and ending at fracture. The figure below shows a typical stress-strain curve for a ductile material such as mild steel, with six key points labelled O, A, B, C, D, and E along the way.

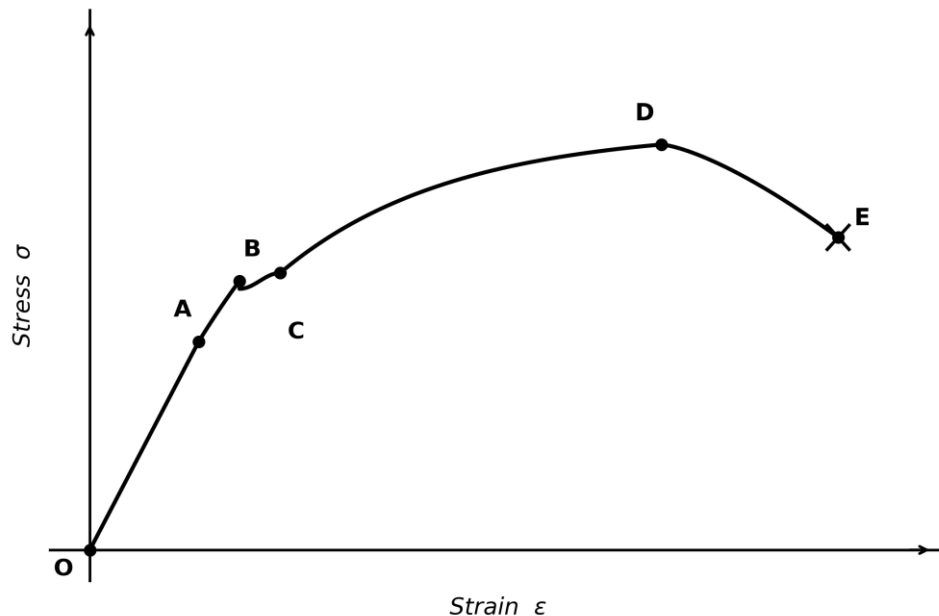


Figure: The stress-strain curve for a typical ductile material such as mild steel. The curve rises linearly from the origin O to the proportional limit A, curves slightly to the elastic limit B, then passes the yield point C where plastic deformation begins. From C to D the material work-hardens, reaching its ultimate tensile strength at D. Beyond D the specimen necks down and the stress falls until fracture finally occurs at E, shown by the small cross.

At the origin O, the material is unloaded. As we start to pull, the curve rises in a perfectly straight line from O up to a point marked A. In this region the stress is directly proportional to the strain, which is just Hooke's law in the language of stress and strain. Young's modulus of the material is the slope of this straight line, and it is the same number we would get from any elastic measurement on any piece of the same metal. Point A is the end of this linear behaviour. It is called the **proportional limit**, because beyond it the proportionality between stress and strain begins to break down. It can be defined as *the largest stress for which the strain remains exactly proportional to the stress*. Up to this point the material obeys Hooke's law, and the stress-strain curve is a straight line whose slope is the Young's modulus of the material.

Just beyond A, the curve is no longer a straight line, but the material is still elastic. If we remove the load at any point between O and B, the specimen returns to its original length and no permanent change has been made. Point B marks the very last stress for which this full recovery is possible. It is called the **elastic limit**, which can be defined as *the largest stress the material can sustain without suffering any permanent deformation*. Below the elastic limit, removing the load returns the specimen to its original length. Above it, some permanent deformation remains even after the load is removed. The elastic limit is marked as point B on the curve, and it lies a little above the proportional limit A for most ductile materials.

Once the stress exceeds the elastic limit, something new happens inside the metal. The molecular bonds that were stretched elastically in the earlier region are now beginning to break and reform in new positions. Planes of atoms slide past one another. Point C on the curve marks the moment when this sliding becomes widespread enough to produce a sudden yielding: the material gives way, and the strain increases rapidly even though the stress has barely risen (and, in some metals, has actually dipped slightly). Point C is called the **yield point**. Beyond it, the material has entered the plastic region, where deformation is permanent.

By definition, the **yield point** of a material is *the stress at which the material suddenly begins to deform rapidly with little or no further increase in stress*. It marks the boundary between the elastic region (where deformations are recoverable) and the plastic region (where deformations become permanent).

So **Plastic deformation** can be defined as *the permanent change in shape of a material under a load that has exceeded the elastic limit*. Unlike elastic deformation, plastic deformation remains after the load is removed, because the molecular bonds have reorganised into new positions that no longer correspond to the original unloaded shape.

From the yield point **C** onwards, the curve rises again, but in a different way. The slope is much gentler than in the elastic region, and the rise is concave, gradually levelling off. In this region the metal is strengthening itself as it deforms. Each new slip between planes of atoms leaves behind defects that make further slipping harder, so more and more stress is needed to produce each additional increment of strain. This self-strengthening process is called work hardening, and it is *the reason a blacksmith's hammer can turn a soft iron rod into a harder tool, and the reason a paperclip becomes stiffer the more you bend it back and forth before it finally snaps*.

So **work hardening** (also called **strain hardening**) is defined as *the strengthening of a ductile material caused by plastic deformation*. As the material is deformed beyond its yield point, defects accumulate in its atomic structure that resist further deformation, so the stress required to continue stretching the material rises. Work hardening is responsible for the rising portion of the stress-strain curve between **C** and **D**.

The work-hardening region ends at point **D**, which is the highest stress the specimen ever reaches. This point is called the **ultimate tensile strength** of the material. Beyond **D** the curve actually starts to drop, which sounds strange because the specimen has not yet broken. The drop happens because the specimen begins to narrow locally, forming a thin neck somewhere along its length. Inside the neck the cross-sectional area is smaller than anywhere else, so for the same load the local stress in the neck rises sharply even though the total force on the specimen is no longer increasing. The plot, however, is drawn using the original (undeformed) cross-sectional area of the specimen, so it shows a falling stress, which is an artefact of using the original area rather than the true current area. Physically, the wire is being pulled apart at the neck.

Concisely, **ultimate tensile strength** of a material is *the maximum stress the material can sustain before it begins to neck down and fail*. It is the highest point reached on the stress-strain curve, marked as point **D**. It is also called the **breaking stress** in some textbooks, though the true fracture actually happens a little beyond this point at **E**.

The neck continues to thin until the material finally fractures at point **E**. On the graph, fracture is marked with a small cross to indicate the sudden end of the curve. The specimen is now in two pieces, and the experiment is over.

So the **fracture point** of a material can be defined as *the stress at which the specimen actually breaks into two pieces*. It occurs a little beyond the ultimate tensile strength because of the necking that takes place after **D**.

Three material classes

The stress-strain curve we have just walked through is typical of a ductile metal such as mild steel or copper. Not all materials behave this way. Some fracture much earlier, with almost no plastic deformation at all. Others deform enormously before they fracture, with almost no truly linear region at the start. The figure below shows the stress-strain curves of three materials drawn on the same axes for comparison, so that you can see the differences at a glance.

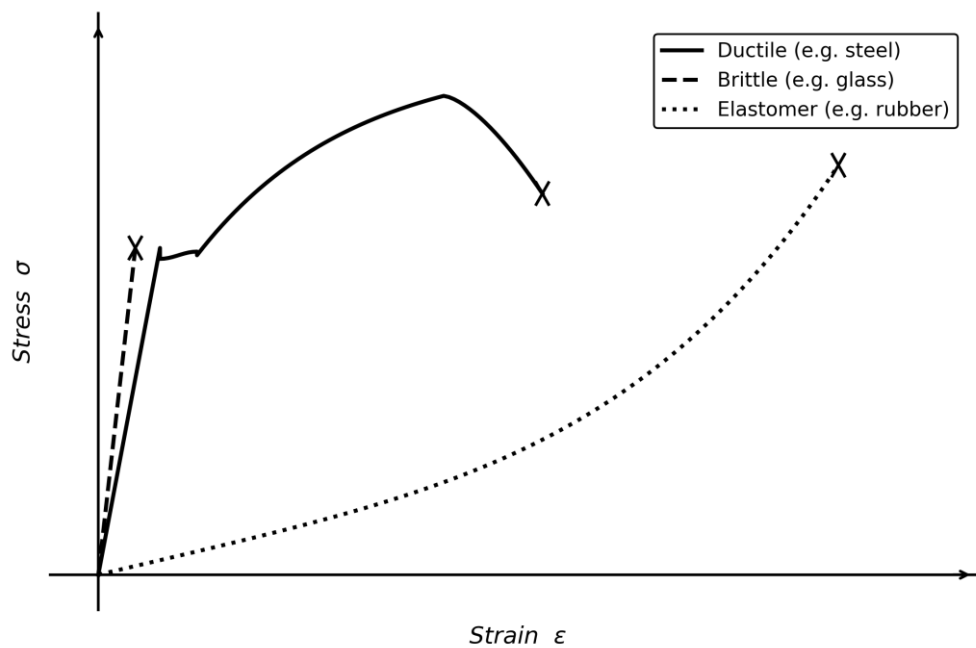


Figure: Stress-strain curves of three material classes drawn on the same axes for comparison. The solid line shows a ductile material such as steel, with a long plastic region and eventual necking before fracture. The dashed line shows a brittle material such as glass, which fractures at a small strain with almost no plastic deformation. The dotted line shows an elastomer such as rubber, which has a very low initial slope but stiffens as the strain grows and eventually fractures at a very large strain. The small crosses indicate the fracture point of each material.

The solid line shows the ductile curve we have already studied. Steel and copper behave this way, with a clear linear region, a clear yield point, a long plastic region with work hardening, and eventual fracture after substantial permanent deformation. Such materials are called **ductile**.

Concisely, a **ductile material** is one that undergoes substantial plastic deformation before it fractures. Its stress-strain curve has a long plastic region beyond the yield point, and the total strain at fracture is large compared with the strain at the elastic limit. Examples include mild steel, copper, aluminium, and most of the engineering metals.

The dashed line shows a **brittle** material such as glass or cast iron. A brittle material is stiff: its stress-strain curve rises very steeply in the elastic region. But it fractures almost immediately after the elastic limit, with almost no plastic region and almost no warning. The total strain at fracture is tiny, and the break happens suddenly. A glass rod pulled in a laboratory tensile test will snap cleanly at a small extension, with the two halves still fitting back together almost perfectly because no plastic deformation occurred to change their shapes.

Concisely, a **brittle material** is one that fractures at or very soon after its elastic limit, with little or no plastic deformation. Its stress-strain curve is almost entirely straight and stops abruptly at fracture. The total strain at fracture is small. Examples include glass, cast iron, ceramics, and many stones.

The dotted line shows an **elastomer** such as natural rubber. An elastomer is very different from both the ductile and the brittle classes. It has an extremely low initial slope, which means it is very soft: a small force produces a large strain. But the curve is not straight at all, and the slope steepens as the strain increases, so the material stiffens the more it is stretched. Eventually the curve becomes very steep and then fractures, often at a strain of several hundred percent, which is far beyond anything any metal could survive. A rubber band can be stretched to three or four times its natural length before it finally breaks.

Concisely, an **elastomer** is a material that can undergo very large elastic strains (typically several hundred percent) and return to its original shape when the load is removed, but whose stress-strain relationship is strongly non-linear even at small strains. Rubber and most soft polymers are elastomers. Elastomers do not have a proportional limit in the ordinary sense because they never have a truly linear region.

Toughness: the area under the curve

We now meet the third of the three quantities that describe the mechanical personality of a material. We already have **Young's modulus**, which measures stiffness, and **breaking stress**, which measures strength. The third quantity answers a different question altogether: *how much energy can the material absorb before it fractures?* The answer to that question is called **toughness**.

Toughness is not the same as strength, and it is not the same as stiffness. A tough material is one that absorbs a lot of energy per unit volume before giving way, regardless of whether that energy is absorbed through a small strain at high stress or through a large strain at modest stress. *On a stress-strain curve, the energy absorbed per unit volume is represented by the area under the curve from zero up to the point where fracture occurs. The tougher the material, the larger that area.*

This deserves a small pause, because we are introducing a new interpretation of the stress-strain curve. In the chapter on momentum in *Foundations of Physics* you met the idea that the area under a force-time graph equals the total impulse delivered, because impulse is defined as force multiplied by time. The same kind of reasoning applies here. The area under a stress-strain curve turns out to equal the work done per unit volume of the material, because stress is force per area and strain is extension per length, and the product of the two has the dimensions of work per volume. We shall derive this result carefully in a later part of the chapter when we study elastic potential energy, but for now the key fact is that the area under the curve is energy, and the area up to the fracture point is the total energy the material absorbed before it gave way.

Concisely, the **toughness of a material** is the total energy per unit volume that the material can absorb before it fractures. It is measured by the area under the stress-strain curve from zero strain up to the fracture point. Its SI unit is the joule per cubic metre (J/m^3). A tough material is not the same as a strong material: a material with a modest breaking stress but a large strain at fracture can be tougher than a material with a high breaking stress but a tiny strain at fracture.

The three material classes we just introduced have very different toughnesses, and the differences are visible directly on the combined figure. Glass, with its steep but very short curve, encloses only a very small thin (narrow) area before fracture, so its toughness is low. Steel, with its long work-hardening region stretching out to substantial strains before fracture, encloses a much larger area, so its toughness is high. Rubber, with its curve reaching out to strains of several hundred percent, encloses a truly enormous area despite its modest breaking stress, so its toughness is also very high. These three different toughnesses are exactly what Kipanga was trying to ask about in the dialogue that follows.

REAL Example 11

Mr. Akilikubwa has just finished drawing the three stress-strain curves on the board and has defined toughness as the area under the curve up to fracture. Kipanga, who has been silent for the whole demonstration, finally raises his hand.

Kipanga: “Sir, I have been waiting for this lesson for weeks, and I think I now understand something that has been bothering me. Glass has a higher Young's modulus than rubber. Glass has a higher breaking stress than rubber. So by both stiffness and strength, glass wins. And yet when I drop a glass bottle on the floor it smashes, and when I drop a rubber ball on the floor it bounces. The glass is the winner on both scales and also the loser in the real world. How is that possible?”

Mr. Akilikubwa: “Kipanga, this is exactly the right question, and the answer is sitting on the board in front of you. Look at the three curves and tell me which one encloses the most area before fracture.”

Kipanga walks to the board and runs his finger along each curve.

Kipanga: “The glass curve is very steep but stops almost immediately. The area under it is tiny. The steel curve is less steep but goes further across before it stops. Its area is bigger. The rubber curve is almost flat at the start but it stretches all the way across the board before it finally stops, so even though it is low, the total area under it is enormous.”

Mr. Akilikubwa: “Correct. And that area is the total energy per unit volume the material can absorb before it fractures. When you drop a glass bottle on the floor, the impact delivers a certain amount of energy to the glass very quickly. That energy has to go somewhere. In a ductile or elastomeric material, it would be absorbed by the elastic and plastic deformation of the material itself, because the area under the curve provides plenty of room to absorb it. In glass, the area under the curve is so small that the incoming energy is more than the material can absorb elastically, and the only way out is to break the bonds that hold the

molecules together. The crack starts at the first flaw on the surface of the bottle and races through it in microseconds, and the bottle is in pieces before you finish hearing it hit the floor.”

Mr. Akilikubwa: *“When you drop a rubber ball, the same impact delivers roughly the same amount of energy. But the rubber has a gigantic area under its curve, so the incoming energy is absorbed easily by a brief elastic deformation. The rubber compresses, stores the energy as elastic potential energy, and then releases it again by springing back to its original shape. The ball bounces, and the story ends well for the rubber.”*

Kipanga: *“So the word I was missing all along was toughness, not strength.”*

Mr. Akilikubwa: *“Yes. Glass is stiff and strong but not tough. Rubber is soft and weak but very tough. Steel is stiff, strong, and tough, which is why we build bridges out of it.”*

Explain in your own words why a glass bottle dropped on a hard floor shatters while a rubber ball dropped on the same floor bounces back unharmed, using Young's modulus, breaking stress, and toughness as the three key quantities.

Solution

The everyday word “hard” secretly describes three independent properties, and the paradox dissolves once we give them their proper names. Young's modulus measures stiffness. Breaking stress measures strength. Toughness measures how much energy per unit volume the material can absorb before fracture. A material can rank high on any combination of these three scales, and there are materials at every corner of the three-dimensional map.

Glass is stiff (high Young's modulus) and reasonably strong (breaking stress around 50MPa) but not tough, because its stress-strain curve is almost a vertical line that stops abruptly at a very small strain. The area under the glass curve is tiny, so glass can absorb only a tiny amount of energy per unit volume before its bonds break. Dropping a bottle on the floor delivers more energy than the glass can absorb, and the only escape route is fracture.

Rubber is soft (low Young's modulus) and weaker than glass in breaking stress (around 20MPa) but is far tougher, because its stress-strain curve extends out to strains of several hundred percent before fracture. The area under the rubber curve is therefore enormous, and the rubber can absorb the energy of a normal impact without any damage at all. It deforms elastically, stores the energy briefly as elastic potential energy, and releases it again when the ball bounces back.

The everyday observation that glass breaks and rubber bounces is therefore an observation about toughness, not about stiffness or strength. Without the three separate quantities, the observation looks like a paradox. With the three separate quantities, it is an obvious consequence of where each material sits on the toughness axis.

Making Sense of the Answer: *Notice that the resolution of the paradox did not require any of the three properties to be denied. Glass really is stiffer than rubber, and glass really does have a higher breaking stress than rubber. Those two facts remain true. What was missing was the third quantity that the everyday word “hard” was silently conflating with the other two. Once toughness is added to the vocabulary, all three materials fit naturally into a consistent picture, and the puzzle of why a glass bottle breaks and a rubber ball does not stops being a puzzle at all.*

Think Like a Physicist: *Whenever everyday intuition seems to contradict a physical principle, ask whether the everyday intuition is actually using a single word to describe two or three separate physical quantities. The same trap appears with “heat” and “temperature” in thermodynamics, with “weight” and “mass” in mechanics, and with “hardness” in materials science. In each case, the resolution is not to argue about the intuition but to replace the ambiguous word with the precise physical quantities it was secretly covering.*

Elastic hysteresis and elastic fatigue

The stress-strain curves we have drawn so far all assumed that we were increasing the load from zero up to fracture. But real engineering materials are often loaded and unloaded repeatedly rather than monotonically. A spring inside a daladala bench is compressed and released thousands of times a day. A car tyre is squeezed and released with every rotation of the wheel. A rubber engine mount absorbs vibration by being cyclically loaded at high frequency. To describe these situations, we need to follow what happens when the load is removed and re-applied, and for that we need a slightly different kind of graph.

If you load a piece of rubber, then release the load slowly and watch the stress-strain curve trace itself out, you find that the curve does not come back along the same path it went up by. The unloading curve lies below the loading curve, and the two together form a closed loop as shown in the figure below. This loop is called an elastic hysteresis loop, and it tells us something important about the energy that went into stretching the rubber.

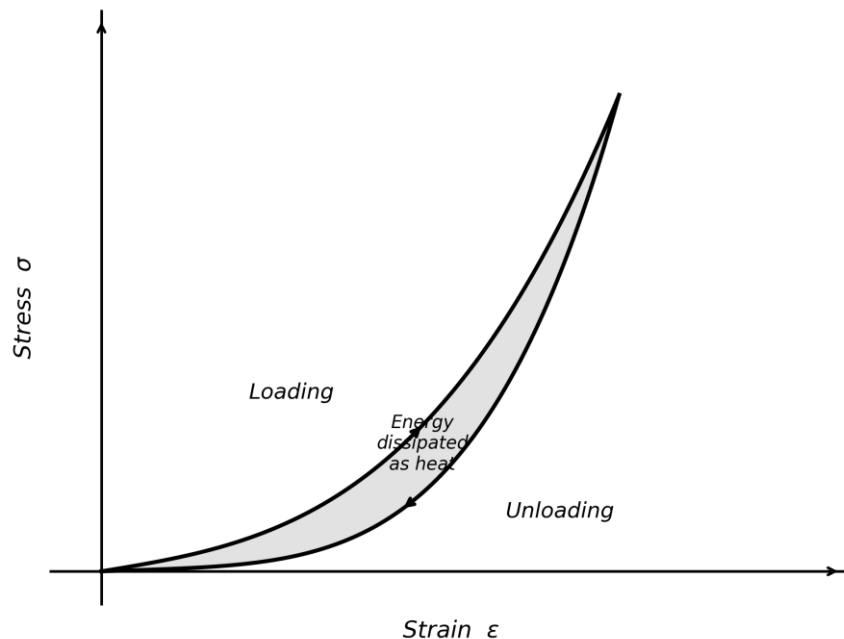


Figure: The elastic hysteresis loop of a rubber-like material. The loading curve lies above the unloading curve, and the two together enclose a shaded area. The area under the upper curve is the energy per unit volume delivered to the rubber during stretching, and the area under the lower curve is the energy per unit volume returned when the load is removed. The shaded difference between the two areas is the energy per unit volume dissipated as heat during each complete loading and unloading cycle.

Because the area under a stress-strain curve is energy per unit volume, the area under the upper (loading) curve is the energy per unit volume that was delivered to the rubber while it was being stretched. The area under the lower (unloading) curve is the energy per unit volume that the rubber gave back when the load was removed. These two areas are not equal: the loading area is larger. The difference between them is the shaded region enclosed by the loop, and that difference is the energy per unit volume that was lost during the cycle. The lost energy does not disappear. It is converted into heat inside the rubber, and if you load and release the rubber rapidly enough, the temperature of the rubber actually rises in a way you can feel with your hand. This is why a rubber band snapped repeatedly between your fingers becomes warm, and why the tyres of a car driven at high speed can get hot enough to damage themselves on a long journey.

Concisely, **elastic hysteresis** is the phenomenon in which the unloading curve of a material lies below the loading curve, so that energy is lost during each loading and unloading cycle. The lost energy appears as heat inside the material. The amount of energy lost per unit volume per cycle is equal to the area enclosed by the hysteresis loop on the stress-strain diagram. Rubber and other elastomers show strong hysteresis, while most metals show very little at loads below their elastic limit.

A related but distinct phenomenon occurs when a material is loaded repeatedly at stresses well below its breaking stress. One might expect such a material to survive forever, because each individual loading is not enough to cause damage. But real materials often fail after a large but finite number of cycles, even when the stress per cycle was always safely below the breaking stress. The failure is caused by tiny cracks that begin at microscopic flaws in the material and grow a little with each cycle, until eventually they are large enough to cause sudden fracture. This gradual weakening under cyclic loading is called **elastic fatigue**, which can be defined as *the progressive weakening of a material under repeated cycles of loading, even when each individual load is well below the breaking stress*. Elastic fatigue is responsible for many real-world engineering failures in aircraft wings, bridge cables, engine parts, and railway tracks, and is the reason such structures are routinely inspected for cracks and retired long before the number of cycles they have survived would suggest.

Elastic fatigue is also the reason a piece of wire that has been bent back and forth many times eventually snaps at the bend, even though the single bending force is much less than the force that would snap the wire if applied only once. Each bend does not break the wire, but each bend grows the tiny cracks inside the metal a little bit, until after enough cycles the cracks are large enough to cause failure. A student who wants to snap a thick wire without access to wire cutters knows this trick instinctively, even without knowing the physics: bend the wire back and forth many times, and eventually it will break.

The ideas have finished cooking; now let us sit down and enjoy them through three worked examples. The first attaches a number to the hysteresis loop just discussed, and the next two return to the stress-strain curve itself.

BINDER Example 12

A rubber engine mount is designed to absorb vibration by being cyclically compressed and released. Its stress-strain hysteresis loop encloses an area of $4 \times 10^4 \text{ J/m}^3$ per cycle. The mount has a volume of 50 cm^3 and is cycled 25 times per second by the running engine. Calculate (a) the energy dissipated as heat per cycle, and (b) the rate at which heat is generated inside the rubber.

Solution

(a) The energy dissipated as heat per cycle is the area enclosed by the hysteresis loop multiplied by the volume of the rubber. Calling the energy per cycle Q_1 (where Q_1 stands for the heat per cycle), the loop area u and the volume V :

$$Q_1 = u \times V$$

Substituting in SI units and computing:

$$Q_1 = (4.0 \times 10^4 \text{ J/m}^3)(50 \times 10^{-6} \text{ m}^3) = 2 \text{ J}$$

Therefore, the rubber dissipates 2J of heat in each compression-and-release cycle.

(b) The rate of heat generation is the energy per cycle multiplied by the number of cycles per second. Calling the cycling frequency f and the rate of heat generation P :

$$P = Q_1 \times f$$

Substituting and computing:

$$P = (2.0 \text{ J})(25 \text{ s}^{-1}) = 50 \text{ W}$$

Therefore, the rubber generates heat at a rate of 50W.

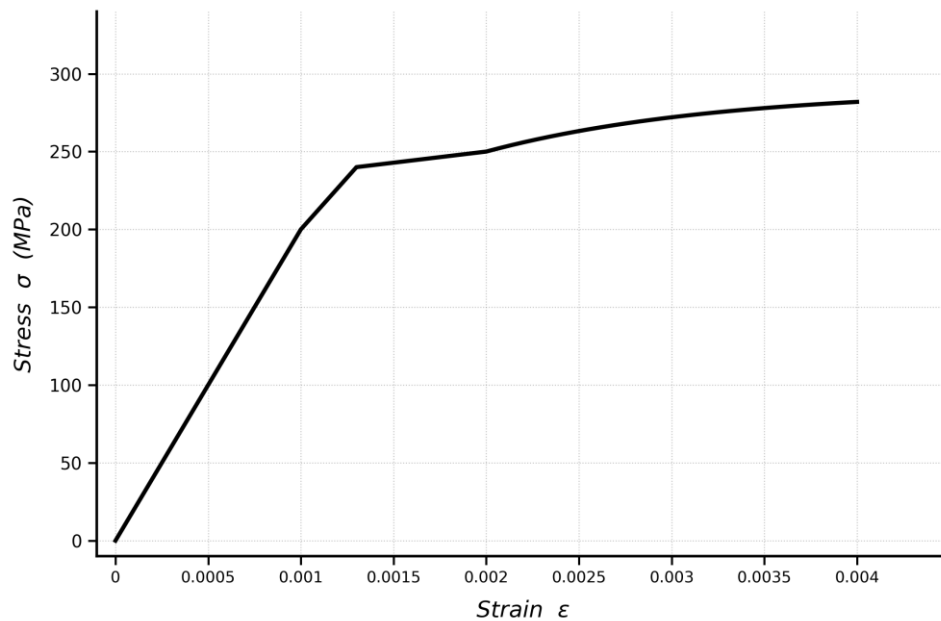
Making Sense of the Answer: The number 50W tells us why rubber engine mounts have to be designed with cooling in mind. The heat produced per cycle is small, only 2J, but the engine delivers that small heat twenty-five times every second, and the rubber is a poor conductor that cannot shed the heat as quickly as it is being made. An undersized mount can therefore fail by overheating long before it fails by mechanical fatigue, even though it never approaches its breaking stress. Hysteresis is a quiet but persistent thief of mechanical energy, and the area inside the loop is the receipt for everything it has stolen.

Think Like a Physicist: Whenever a material is loaded cyclically and there is no obvious external heat source, look for hysteresis. The same calculation applies, with the same three numbers in the same three places, to brake pads on a long descent, to the rubber bushes of a daladala suspension on a Dar es Salaam pothole, and to the tyres of a long-distance lorry on a hot afternoon. In each case, the heat is made by the material from itself, and the rate is loop area times volume times frequency. Once the formula is in your toolkit, you will start to notice it everywhere mechanical work is being repeatedly done and undone.

With a number now attached to the hysteresis loop, we return to the stress-strain curve itself and learn to read it carefully.

BINDER Example 13

The figure below shows the early portion of the stress-strain curve of a ductile metal loaded in tension. From the figure, read off (a) the proportional limit, (b) the elastic limit, (c) the stress at the yield point, and (d) estimate Young's modulus of the metal.



Solution

(a) The curve is a perfect straight line from the origin up to the point where it first begins to bend. Reading across from that point to the vertical axis gives a stress of approximately 200MPa, and reading down to the horizontal axis gives a strain of approximately 0.001.

The proportional limit of the metal is therefore about 200MPa.

(b) Just beyond the proportional limit the curve continues to rise but with a gentle curvature. The elastic limit is the highest stress for which the strain is fully recovered when the load is removed, and on the graph it sits slightly above the proportional limit, at approximately 240MPa at a strain of about 0.0013.

The elastic limit of the metal is therefore about 240MPa.

(c) The yield point is the stress at which the curve flattens noticeably, indicating that the material has begun to flow plastically. On the graph this happens at approximately 250MPa at a strain of about 0.002.

The yield point is therefore about 250MPa.

(d) Young's modulus is the slope of the initial linear portion of the curve, measured between any two points in the straight-line region. Using the origin and the proportional limit itself:

$$E = \frac{\Delta\sigma}{\Delta\epsilon} = \frac{200\text{MPa} - 0}{0.001 - 0}$$

Substituting in SI units and computing:

$$E = \frac{200 \times 10^6 \text{Pa}}{0.001} = 2 \times 10^{11} \text{Pa} = 200\text{GPa}$$

Young's modulus of the metal is 200GPa.

Making Sense of the Answer: Every part of this question was answered by reading values directly off the graph. No extra information was needed, because the stress-strain curve contains the full elastic personality of the material in a single compact picture. This is why the curve is so valuable to engineers. Once you have it, you can find the proportional limit, the elastic limit, the yield point, Young's modulus, the ultimate tensile strength, and the breaking strain all from the same figure, without having to run any further experiments.

Think Like a Physicist: When you estimate Young's modulus from the slope of a graph, always use two points that are as far apart as possible within the straight-line region. A slope computed between two points that are close together is extremely sensitive to small reading errors. A slope computed between two points that are far apart averages those errors out. In this example the straight-line region happened to extend all the way from the origin to the proportional limit, so the origin and the proportional limit were the natural two points to use.

HOT Example 14

Three unknown materials, labelled X, Y, and Z, are tested in tension in a laboratory. For each material, the load and extension are measured at several points up to the moment of fracture. From the measurements, the stress and strain are computed at each point. The tabulated results are shown below. Classify each material as ductile, brittle, or elastomeric, and justify your classification with specific reference to the shape of its stress-strain relationship.

Material X

Strain ϵ	Stress σ (MPa)
0.0005	100
0.001	200
0.0012	235
0.002	250
0.020	350
0.080	420
0.150	410
0.200	fracture

Material Y

Strain ϵ	Stress σ (MPa)
0.0002	14
0.0004	28
0.0006	42
0.0008	56
0.001	fracture

Material Z

Strain ϵ	Stress σ (MPa)
0.5	0.6
1.0	1.4
1.5	2.8
2.0	5.5
2.5	11.0
3.0	20.0
3.2	fracture

Solution

Material X. The first three data points lie on a straight line, because the stress rises by equal amounts for each equal step of strain: from 0 to 200MPa as the strain increases from 0 to 0.001, which is a perfect linear relationship. The fourth and fifth points show a small rise in stress for a relatively large jump in strain (235MPa to 250MPa while strain increases tenfold), which is the yield region. The remaining points show a gradual work hardening up to about 420MPa at a strain of 0.08, followed by a slight drop as the specimen necks down before fracturing at a strain of 0.20. The curve therefore has a clear linear region, a clear yield point, a long plastic region with work hardening, and substantial strain at fracture. This is the signature of a ductile material like steel.

Material Y. All four measured points lie on a straight line, because the stress rises by exactly 14MPa for each 0.0002 step of strain, all the way up to fracture at 0.001 strain. There is no yield point, no plastic region,

and no work hardening. The specimen fractures abruptly at the end of the linear region, at a very small total strain. This is the signature of a brittle material like glass.

Material Z. The stress-strain relationship is non-linear from the very first measurement. A strain of 0.5 (which is already 50 percent) gives only 0.6MPa of stress, while an additional strain of 0.5 (up to 1.0, or 100 percent) gives only another 0.8MPa of stress. But from there the curve steepens sharply: each successive step of strain produces a larger increase in stress than the previous one. Fracture occurs at a strain of about 3.2, which is 320 percent of the original length. No metal could possibly survive such a strain. This is the signature of an elastomer like a natural rubber or a similar polymer.

Making Sense of the Answer: *The three materials fit the three classes cleanly, with no ambiguity. The key observations were: Material X had a clear linear region followed by a long plastic region and large strain at fracture. Material Y had only a linear region and fractured almost immediately. Material Z was non-linear from the start and stretched to several times its original length before fracturing. Each class is recognisable from the shape of the data alone, without needing to know what the material is.*

Think Like a Physicist: *When you are given tabulated laboratory data and asked to classify a material, the first thing to check is whether the first few data points lie on a straight line, which tells you whether the material has a linear elastic region at all. If it does, you can compute the slope to get Young's modulus and compare with the table of known values. Then look at how far the curve extends past the linear region before fracture: a short distance indicates brittle behaviour, a long distance indicates ductile behaviour. If there is no straight-line region at all and the strain values are of order one rather than one thousandth, the material is an elastomer.*

The stress-strain curve is now in your toolkit, along with the three quantities that describe the mechanical personality of any material: stiffness measured by Young's modulus, strength measured by breaking stress, and toughness measured by the area under the curve. But we have so far only discussed one kind of stress and one kind of strain, namely the linear tensile kind in which force is pulled along the length of a wire. Real materials can be loaded in two other geometrically different ways: they can be sheared sideways, and they can be squeezed equally from all directions at once. Each of these other loading modes has its own stress, its own strain, and its own modulus. The next part of this chapter introduces the shear modulus and the bulk modulus, which together with Young's modulus make up the three moduli of elasticity that form the complete linear elastic description of a solid material.

SHEAR MODULUS AND BULK MODULUS

So far every stress we have met has been tensile: a force pulled on a wire along its length, and the wire stretched a little. But pulling is only one of the three geometrically distinct ways you can apply a force to a solid body. You can also push sideways on a block with a force parallel to one of its faces, and the block will skew out of shape while keeping the same volume. Or you can squeeze a body equally from all directions at once, and it will shrink in volume while keeping the same shape. Each of these three loading modes produces its own kind of stress, its own kind of strain, and its own modulus of elasticity. Tensile loading gives the Young's modulus we have already met. Sideways loading gives the **shear modulus**. All-round squeezing gives the **bulk modulus**. Between them, Young's modulus, the shear modulus, and the bulk modulus describe the complete linear elastic behaviour of any solid material.

Shear stress and shear strain

Imagine a rectangular block of material resting on a horizontal table, with its base glued firmly to the table. Now push the top face of the block with a horizontal force that is parallel to the face, not perpendicular to it. The block does not stretch or shrink in any direction, because the force is not pulling it apart or squeezing it together. Instead, the block skews out of shape. The top face slides sideways a little, the base stays where it is because it is glued, and the originally-rectangular block becomes a parallelogram. This kind of deformation is called shear, and the following figure shows what it looks like.

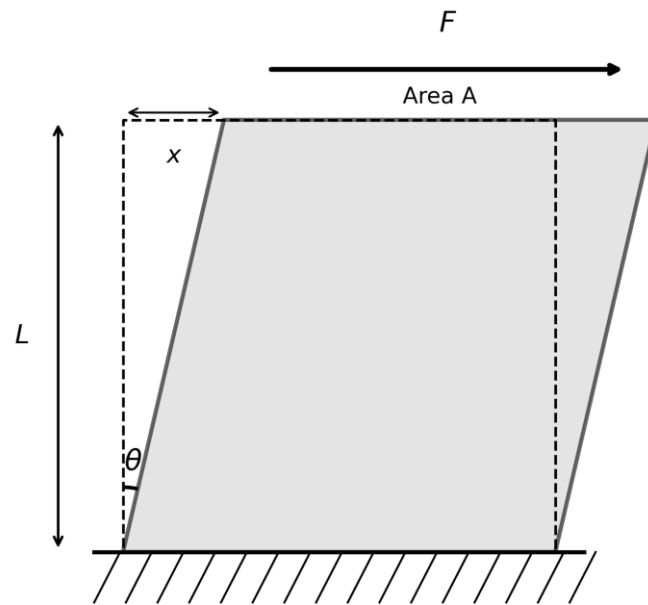


Figure: A rectangular block of original height L with its base fixed to a rigid surface, subjected to a horizontal force F parallel to its top face. The top face, of area A , displaces sideways by a distance x while the base stays where it is. The originally rectangular block deforms into a parallelogram. The small angle θ between the original vertical side and the new slanted side is the geometric tilt that we shall use to define the shear strain in the prose below. The dashed outline shows the undeformed shape for comparison.

To describe this deformation quantitatively we need two new quantities. The **first** is the stress produced by the **tangential force**. Just as tensile stress was defined as the normal force divided by the area of the face it acts on, **shear stress** is defined as the ratio of the tangential force applied to a surface to the area of that surface. Its symbol is τ (tau), and thus:

$$\tau = \frac{F}{A}$$

Its SI unit is the pascal (Pa), the same as tensile stress. It measures how hard each unit of area is being pushed sideways, rather than pulled or squeezed.

The **second** quantity is the strain produced by the deformation. When the top face of the block slides sideways by a distance x while the base remains fixed, the originally vertical sides of the block tilt away from the vertical by a small angle. Call that angle θ , just as you would call any other angle in geometry. From the right triangle formed by the original vertical edge, the new slanted edge, and the horizontal displacement of the top, the tangent of θ is the sideways displacement divided by the original height of the block:

$$\tan \theta = \frac{x}{L}$$

For small deformations, which is the only regime where Hooke's law applies anyway, the tangent of a small angle in radians is essentially equal to the angle itself, so we can drop the tangent and write the angle directly:

$$\theta = \frac{x}{L}$$

This small angle θ is the geometric description of the deformation, and any student of geometry could have written it down without knowing any elasticity at all. But in elasticity it has a special name and a special symbol, because it plays the same role for shear loading that the ordinary tensile strain ϵ plays for tensile loading. The elastic-theory name for this angle is the shear strain, and its standard symbol is γ (gamma). The shear strain is identical in magnitude to the tilt angle:

$$\gamma = \theta = \frac{x}{L}$$

So the **shear strain** can be defined as *the small angle (in radians) by which the originally vertical sides of a sheared block tilt under the applied tangential force*. Equivalently, it is the ratio of the sideways displacement of the top face of the block to the original perpendicular distance between the top face and the fixed base. Its symbol is γ and it has no units.

The shear modulus

Now, just as tensile stress divided by tensile strain gave us Young's modulus, shear stress divided by shear strain gives us a new material property called the shear modulus. Its symbol is G and its mathematical definition is:

$$G = \frac{\tau}{\gamma}$$

Substituting the definitions of τ and γ gives the expanded form:

$$G = \frac{\frac{F}{A}}{\frac{x}{L}} = \frac{FL}{Ax}$$

The **shear modulus** of a material can be defined as *the ratio of the shear stress acting on the material to the resulting shear strain, provided the material is within its elastic range*. Its SI unit is the pascal (Pa). It measures how strongly the material resists being skewed out of shape. A material with a large shear modulus is hard to shear; a material with a small shear modulus is easy to shear. The shear modulus is also known as the **modulus of rigidity**, and some older textbooks use the symbol η (eta) instead of G .

The shear modulus of a typical metal is roughly one third of its Young's modulus. For steel, Young's modulus is about 200GPa and the shear modulus is about 80GPa. For copper, the corresponding values are 120GPa and 45GPa. For aluminium, 70GPa and 26GPa. The rough one-third ratio is not a coincidence: it can be derived from the theory of linear elasticity for an isotropic solid, together with one additional material parameter called Poisson's ratio, which we shall not pursue further in this chapter.

Bulk stress and bulk strain

The third and final loading mode is all-round compression. Imagine taking a small cube of material and squeezing it equally from every direction at once, the way the water pressure at the bottom of the Indian Ocean squeezes every cubic centimetre of a submarine pressure hull. The cube does not change shape, because the squeezing is equal in all directions. But it does shrink in volume, because every molecule in the interior is being pushed slightly closer to its neighbours. The following figure shows a cube subjected to such a uniform pressure.

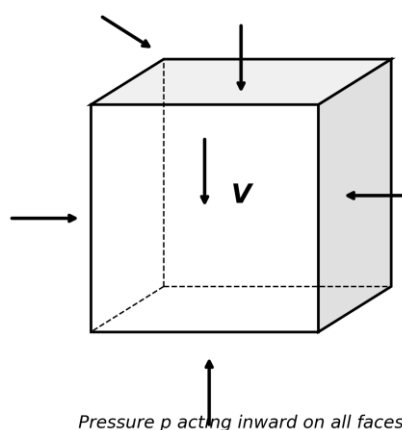


Figure: A cube of material of volume V subjected to a uniform pressure p acting inward on all six of its faces simultaneously. The cube shrinks slightly in volume by an amount ΔV while retaining its original shape. This kind of loading is called bulk compression, and unlike tensile loading or shear loading it can be applied to gases and liquids as well as to solids.

The stress in this situation is the pressure itself, which is the normal force per unit area acting on any face of the cube. The strain is the fractional change in volume of the cube. If the original volume of the cube is V and the volume change (a decrease, because the cube shrinks) is ΔV , then the bulk strain is:

$$\text{bulk strain} = -\frac{\Delta V}{V}$$

The negative implies that when the pressure on the cube increases, the volume shrinks. The bulk strain is dimensionless, just like tensile strain and shear strain. The bulk strain of solids is usually a small number. A bulk strain of -0.001 means the volume has decreased by one part in a thousand under the applied pressure, and for most solids under ordinary conditions the strain is this small or smaller in magnitude.

Concisely:

Bulk stress is the pressure acting uniformly and inward on all faces of a body. Its symbol is p (for pressure) and its SI unit is the pascal (Pa). Unlike tensile stress and shear stress, bulk stress is the only kind of stress that can be applied to a liquid or a gas, because liquids and gases cannot sustain tensile or shear loads at rest.

Bulk strain is the fractional change in volume of a body under a uniform pressure. It is dimensionless, and for small deformations (the only regime where Hooke's law applies) it is always a small negative number.

The bulk modulus

The ratio of bulk stress to bulk strain is called the bulk modulus, K , and it has its own subtlety. An increase in pressure always produces a decrease in volume, so the volume change ΔV is inherently negative when the pressure change is positive. To keep the bulk modulus itself a positive quantity (which is what we want, because a positive modulus describes a stiff material and a negative modulus would describe something that expands when you squeeze it, which no ordinary material does), we include an explicit minus sign in the definition:

$$K = \frac{\text{Bulk stress}}{\text{Bulk strain}} = -\frac{\Delta p}{\frac{\Delta V}{V}} = -V \frac{\Delta p}{\Delta V}$$

The SI unit of K is the pascal, the same as Young's modulus and the shear modulus.

A material with a large bulk modulus is hard to squeeze; a material with a small bulk modulus is easy to squeeze. Water has a bulk modulus of about 2.2 GPa, which is why water at the bottom of the deepest ocean trench is only about 4 percent denser than water at the surface, despite the enormous pressure.

The reciprocal of the bulk modulus has its own name and symbol. It is called the **compressibility** of the material, and it measures how easily the material can be squeezed rather than how strongly it resists. Its symbol is κ (kappa) and its SI unit is the inverse pascal ($1/\text{Pa}$ or Pa^{-1}).

$$\kappa = \frac{1}{K} = \frac{\text{Bulk strain}}{\text{Bulk stress}}$$

A highly compressible material is one whose volume changes substantially under a modest pressure. A barely compressible material is one whose volume is almost unchanged even under large pressures. Gases are highly compressible, liquids are barely compressible, and solids are almost completely incompressible by comparison.

Values of E , G , and K for common materials

The three elastic moduli of a material are related, but they are not equal to one another, and for any given material the three values are usually of the same order of magnitude while differing by factors of two or three. The table below lists typical values for the materials most commonly met in this chapter. Notice that for each metal, the shear modulus is roughly one third of Young's modulus, and the bulk modulus is slightly larger than Young's modulus. For rubber, the shear modulus is tiny but the bulk modulus is comparable to water, because rubber is soft in shear but nearly incompressible.

Material	E (GPa)	G (GPa)	K (GPa)
Steel	200	80	160
Copper	120	45	140
Aluminium	70	26	76
Glass	70	28	40
Water	—	—	2.2
Rubber	0.01	0.003	2.0

Water and rubber are the two unusual entries in the table. Water has no Young's modulus and no shear modulus because it is a fluid: it cannot sustain a tensile pull or a sideways push when at rest, because it simply flows out of the way. It does, however, have a perfectly well-defined bulk modulus, because squeezing it equally from all directions cannot be avoided by flowing. Rubber, at the other end, has a shear modulus so small (three megapascals, about a thousand times smaller than glass) that it deforms dramatically under even the gentlest sideways push, but its bulk modulus is almost as large as that of water, which makes it almost incompressible in volume. This is why a rubber ball squashed between your hands changes shape enormously but barely changes volume at all.

Why fluids have only a bulk modulus

A solid sitting on a table does not flow, because its molecules are held in fixed positions by strong intermolecular forces. When you push sideways on the top of a solid block, those forces resist the sideways push, and the block skews into a parallelogram whose shape is determined by the shear modulus. But the molecules of a fluid (either a liquid or a gas) are not held in fixed positions. They can slide past one another freely, and they will do so whenever the slightest sideways force is applied, at a rate determined by the viscosity of the fluid rather than by any elastic property. A fluid at rest therefore cannot sustain a shear stress: if you try to shear it, it flows out of the way rather than deforming elastically. The same argument applies to tensile loading: a liquid in an open container cannot be pulled apart from its surface, because the surface simply retreats from whatever is pulling on it. Only bulk compression survives, because equal inward pressure from all directions cannot be avoided by any kind of flow. This is why the table above lists a bulk modulus for water but leaves the other two columns blank, and it is why fluid mechanics needed its own chapter in *Foundations of Physics*, with its own concepts of flow, viscosity, and pressure, separate from the elastic theory of solids that we are developing here.

For now, take a short breath and set theory aside, let us warm up with some carefully chosen worked examples.

BINDER Example 15

A rubber block of height 5cm and square top face of side 10cm has its base glued to a workbench. A horizontal force of 30N is applied to the top face parallel to one of its edges, causing the top face to displace sideways by 2mm. Calculate the shear modulus of the rubber.

Solution

The area of the top face is:

$$A = (0.1\text{m})(0.1\text{m}) = 1 \times 10^{-2} \text{ m}^2$$

The shear stress is:

$$\tau = \frac{F}{A} = \frac{30\text{N}}{1 \times 10^{-2} \text{ m}^2} = 3000\text{Pa} = 3\text{kPa}$$

The shear strain is:

$$\gamma = \frac{x}{L} = \frac{2 \times 10^{-3}\text{m}}{5 \times 10^{-2}\text{m}} = 0.04$$

The shear modulus is:

$$G = \frac{\tau}{\gamma} = \frac{3000\text{Pa}}{0.04} = 7.5 \times 10^4\text{Pa} = 75\text{kPa}$$

The shear modulus of the rubber block is about 75kPa, or 0.000075GPa.

Making Sense of the Answer: Seventy-five kilopascals is an extraordinarily small shear modulus. Steel is about a million times larger at 80 gigapascals. This is why a rubber block is so easily twisted and skewed by hand, while a steel block of the same dimensions is utterly rigid to anything short of industrial machinery. Softness in shear is one of the defining features of rubber as a material, and it is exactly why rubber is used to absorb vibration in car engine mounts, machinery foundations, and earthquake-resistant buildings. The softness allows the rubber to deform freely under sideways loads, dissipating the vibrational energy harmlessly as heat instead of transmitting it to the structure above.

Think Like a Physicist: The shear modulus formula has exactly the same structure as the Young's modulus formula, with τ playing the role of σ and γ playing the role of ϵ . Any rearrangement you learned for Young's modulus works unchanged for the shear modulus, and the laboratory procedure for measuring G is essentially the same: apply a known force, measure the displacement of a known face, compute stress and strain, and divide. The only thing that changes is the geometry of how the force is applied.

BINDER Example 16

A solid metal cube of volume 1000cm^3 is lowered into the deep part of a swimming pool, where the water pressure is $1.5 \times 10^5\text{Pa}$ above atmospheric. The cube shrinks in volume by 1.2mm^3 . Calculate the bulk modulus of the metal.

Solution

The cube shrinks, so the volume change is negative. Converting to SI units:

$$V = 1000\text{cm}^3 = 1 \times 10^{-3}\text{m}^3$$

$$\Delta V = -1.2\text{mm}^3 = -1.2 \times 10^{-9}\text{m}^3$$

The bulk strain is negative:

$$\frac{\Delta V}{V} = \frac{-1.2 \times 10^{-9}\text{m}^3}{1 \times 10^{-3}\text{m}^3} = -1.2 \times 10^{-6}$$

Applying the bulk modulus formula with the explicit minus sign:

$$K = -\frac{\Delta p}{\frac{\Delta V}{V}} = -\frac{1.5 \times 10^5\text{Pa}}{-1.2 \times 10^{-6}}$$

$$K = 1.25 \times 10^{11}\text{Pa} = 125\text{GPa}$$

The bulk modulus is 125GPa.

Making Sense of the Answer: Notice how small the fractional volume change was: about one part in a million under the pressure of a few metres of water. For a metal cube of ordinary size the actual change in volume is barely measurable. This is why metals are treated as almost incompressible in most engineering calculations, and why the bulk modulus only becomes practically important in situations involving enormous pressures such as the deep ocean, high-pressure hydraulic systems, or the interior of planets.

Think Like a Physicist: The minus sign in the bulk modulus formula is doing real work. The volume change ΔV is genuinely negative when the pressure change is positive, so the ratio $\Delta V/V$ is also negative. Multiplying by the explicit minus sign flips the answer back to positive, which is what we want K to be for any ordinary material. In numerical problems you can track the signs carefully and the two minus signs will cancel, or you can substitute the magnitudes of Δp and ΔV directly and get the same positive answer. Either way works, provided you understand what each minus sign is doing and do not drop any of them by accident. If you ever get a negative bulk modulus as your final answer, you have made a sign error somewhere.

REAL Example 17

Kipute is reading the back of her uncle's diving wristwatch, which he uses on weekend dives off the coast of Mafia Island. The case is engraved with the words "water resistant to 200m". The case is made of stainless steel with a bulk modulus of 160GPa, and its internal volume is 5cm^3 . Take the density of seawater as 1025kg/m^3 and $g = 9.8\text{m/s}^2$. Find the change in the internal volume of the watch case at a depth of 200m.

Solution

The gauge pressure at a depth of 200m below the surface is:

$$\Delta p = \rho gh = (1025 \text{ kg/m}^3)(9.8 \text{ m/s}^2)(200 \text{ m}) = 2.01 \times 10^6 \text{ Pa}$$

Rearranging the bulk modulus formula for the bulk strain gives $\frac{\Delta V}{V} = -\frac{\Delta p}{K}$, where the minus sign tells us the volume decreases when the pressure rises. Computing:

$$\frac{\Delta V}{V} = -\frac{2.01 \times 10^6 \text{ Pa}}{160 \times 10^9 \text{ Pa}} = -1.26 \times 10^{-5}$$

The actual change in the internal volume of the case is:

$$\Delta V = V \times \frac{\Delta V}{V} = (-1.26 \times 10^{-5})(5 \text{ cm}^3) = -6.3 \times 10^{-5} \text{ cm}^3$$

The negative sign confirms that the case shrinks, and the magnitude of the shrinkage is $6.3 \times 10^{-5} \text{ cm}^3$ at a depth of 200m, which is sixty nanolitres, or roughly the volume of one tenth of a single drop of water.

Making Sense of the Answer: Two hundred metres of seawater is a serious depth. The pressure at that depth is twenty times atmospheric, and the diver feels it on every square centimetre of skin. Yet the internal volume of the steel watch case shrinks by only six hundredths of a millionth of a litre, which is so small that no instrument the diver carries could measure it directly. Steel is so stiff against bulk compression that the case is, for all practical purposes, rigid. The gut feeling that the case might be crushed by the pressure is wrong by many orders of magnitude. What actually limits the depth rating of a diving watch is not the compressibility of the steel case but the seal between the case and the glass face, which is what the manufacturer is really certifying when they engrave “water resistant to 200m” on the back.

Think Like a Physicist: When you meet a problem about something being crushed, broken, or deformed by water pressure at depth, always compute the actual fractional change before you trust your intuition. Water pressure feels enormous when you experience it on your eardrums, because biological tissue is much more compressible than metal. But for a metal object of any reasonable size, even the pressure at the deepest part of the Indian Ocean produces a fractional volume change of less than half a percent, which is far smaller than the precision of most engineering tolerances. The deepest point of the **Mariana Trench** (the deepest part of the world's oceans, found in Pacific Ocean), where the pressure exceeds a thousand atmospheres, would shrink a steel cube by only about half a percent of its volume.

HOT Example 18

A copper cube of side 5cm is subjected simultaneously to two experiments in the same laboratory. In the first experiment, its base is glued to a bench and a horizontal force of 2000N is applied to its top face, parallel to one edge. In the second experiment, it is placed in a high-pressure chamber and subjected to a uniform pressure of $1 \times 10^8 \text{ Pa}$ above atmospheric. Take the shear modulus of copper as 45GPa and the bulk modulus of copper as 140GPa. Find (a) the sideways displacement of the top face in the shear experiment, and (b) the fractional change in volume of the cube in the bulk experiment. Comment on why the two answers are so different.

Solution

(a) Shear experiment.

The area of the top face is:

$$A = (0.05 \text{ m})(0.05 \text{ m}) = 2.5 \times 10^{-3} \text{ m}^2$$

The shear stress is:

$$\tau = \frac{F}{A} = \frac{2000 \text{ N}}{2.5 \times 10^{-3} \text{ m}^2} = 8 \times 10^5 \text{ Pa}$$

The shear strain is:

$$\gamma = \frac{\tau}{G} = \frac{8 \times 10^5 \text{ Pa}}{45 \times 10^9 \text{ Pa}} = 1.78 \times 10^{-5}$$

The sideways displacement of the top face is:

$$x = \gamma L = (1.78 \times 10^{-5})(0.05 \text{ m}) = 8.9 \times 10^{-7} \text{ m}$$

The top face moves sideways by about 0.89 micrometres, which is less than a thousandth of a millimetre.

(b) Bulk experiment.

The fractional change in volume is:

$$\frac{\Delta V}{V} = -\frac{p}{K} = -\frac{1 \times 10^8 \text{ Pa}}{140 \times 10^9 \text{ Pa}} = -7.1 \times 10^{-4}$$

The negative sign confirms that the cube shrinks. In magnitude, the volume decreases by about 0.07 percent, which is a far larger fractional deformation than the shear strain of part (a).

Comment. The two experiments were carried out on the same cube of the same metal, yet the fractional deformation in the bulk experiment was about forty times $\left(\frac{7.1 \times 10^{-4}}{1.78 \times 10^{-5}}\right)$ larger than the strain in the shear experiment. The reason is that the pressure applied in the bulk experiment ($1 \times 10^8 \text{ Pa}$, or 100MPa) was much larger than the shear stress in the shear experiment (0.8MPa). A small force applied tangentially produces a modest shear stress; an enormous pressure applied equally from all directions produces an enormous bulk stress.

Making Sense of the Answer: *The two loading modes are of comparable physical importance for copper, but they show up in different circumstances. Shear loading matters when a metal part is being twisted, bolted, or punched, which is almost every engineering situation involving metal. Bulk loading matters only when the surrounding pressure is extreme: deep ocean, hydraulic cylinders, the interiors of planets, or the compression chambers of diesel engines. For everyday structural calculations, the shear modulus is the one you will meet far more often.*

Think Like a Physicist: *Whenever a problem involves two independent loading modes on the same object, solve each mode as if the other did not exist. The linearity of elastic theory guarantees that the two responses do not interfere with each other, so the shear displacement in this example is exactly the same whether the cube is under pressure or not, and the volume change is exactly the same whether the cube is being sheared or not. This independence is the principle of superposition, and it holds for any linear response theory in physics, including electromagnetism, small-oscillation mechanics, and low-amplitude wave motion.*

Three moduli, three loading modes, one complete description of how a solid responds to small forces. What happens to the energy delivered to the material while it is being loaded? Where does it go, and how much of it comes back when the load is removed? The answer is in the next part of this chapter.

ELASTIC POTENTIAL ENERGY

Pull on a rubber band, and you can feel the rubber band pulling back. Let it go, and it snaps to its original length with enough energy to sting your fingers if they are in the wrong place. The energy that was delivered to the rubber band while you were stretching it did not vanish into the air and it did not turn into heat. It was stored inside the rubber, waiting. The moment you let go, the rubber returned every last joule of it.

This is the question we closed the previous part of the chapter with: *where does the energy delivered to a stretched material go while the material is stretched, and how much of it comes back when the load is removed?* The answer is that the energy is stored inside the material as elastic potential energy, and for any material within its elastic region all of it comes back. Nothing is lost, nothing is wasted, and the formula for how much is stored turns out to be one of the cleanest results in the whole of mechanics.

Deriving the energy formula

Take a spring of spring constant k and stretch it slowly from its natural length up to a final extension x . By “slowly” we mean unhurriedly enough that nothing accelerates and no kinetic energy is involved, so that every joule of work done by our hand goes straight into the spring and stays there. At an intermediate extension x' the applied force is $F = kx'$ by Hooke's law, and extending the spring by a further dx' does work $dW = F dx' = kx' dx'$. The total work from zero extension to the final x is:

$$W = \int_0^x kx' dx' = k \left[\frac{x'^2}{2} \right]_0^x = \frac{1}{2} kx^2$$

Because the process was slow, no kinetic energy appeared along the way, and every joule of this work is now sitting inside the spring as **elastic potential energy U** . That gives us our first and most important result:

$$U = \frac{1}{2} kx^2$$

There is a geometric way to see the same result, and it is worth a moment because it will reappear in other corners of physics. Plot the applied force against the extension. Because the spring obeys Hooke's law, the plot is a straight line through the origin with slope k , as shown in the figure below. The work done in stretching the spring from zero to the final extension x is the area under this line (because work is force times distance, and the area of any thin vertical strip of the graph is exactly the small work $F dx'$ done over that

strip of extension). The line is straight, so the area is a simple triangle of base x and height kx ($\frac{y}{x} = k$). Its area is:

$$\frac{1}{2}(x)(kx) = \frac{1}{2}kx^2$$

The same result we obtained from the integral.

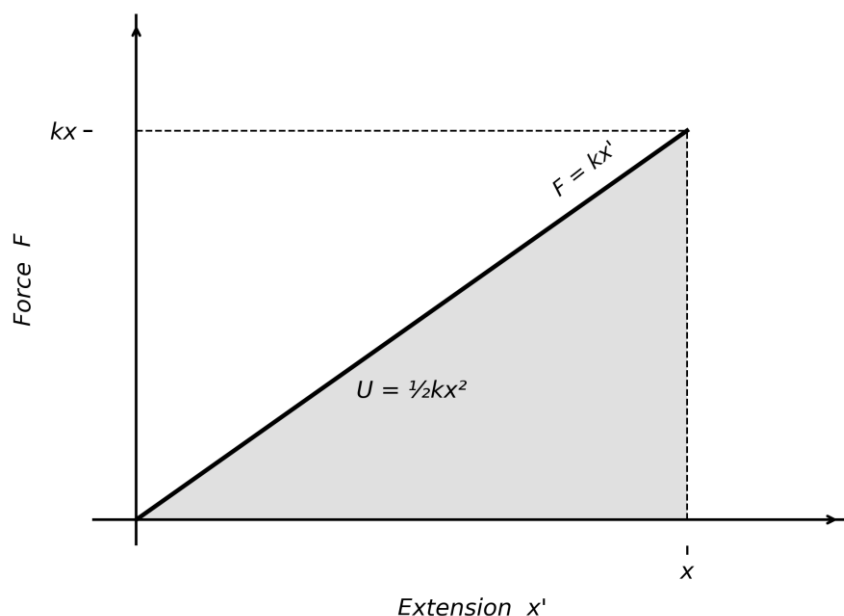


Figure: The force required to hold a Hookean spring at extension x' is $F = kx'$, plotted as a straight line through the origin. The work done in stretching the spring from zero to a final extension x is the area under the line, which is the shaded triangle of base x and height kx . Its area $(1/2)kx^2$ is the same result obtained by integration and represents the elastic potential energy stored in the spring.

That triangle is not unique to springs. Any linear restoring system, when slowly loaded from rest, produces exactly the same picture: a straight-line graph, a triangular area, and a factor of one half in the stored energy formula. You will meet the same triangle again when we come to the energy stored in a charged capacitor, and again in the energy stored in a current-carrying inductor, and again in any other linear-response system where a driving quantity pulls gently against a restoring quantity. Once you recognise the triangle, half the energy problems in physics are already half solved.

Now, we can define the elastic potential energy stored in a body that obeys Hooke's law as *the work done against the restoring force in stretching it from its natural length to its present extension*.

Three equivalent forms

The formula $U = \frac{1}{2}kx^2$ is the most common form, but a problem does not always hand you the extension as the thing you know. Sometimes you know the final force instead, and sometimes you know both. Substituting $F = kx$ into the energy formula in two different ways gives two further useful forms:

$$U = \frac{1}{2}kx^2 = \frac{1}{2}Fx = \frac{F^2}{2k}$$

Use the first when the extension is known, the second when both force and extension are known, and the third when only the force is known.

The final form is derived from $U = \frac{1}{2}kx^2$.

Multiplying by $\frac{k}{k}$:

$$U = \frac{1}{2}kx^2 \times \left(\frac{k}{k}\right) = \frac{(kx)^2}{2k} = \frac{F^2}{2k}$$

Stretched wire and energy per unit volume

Everything we have done so far applies to a spring, and you might worry that a wire is a different beast. It is not. Earlier in this chapter we showed that a wire of length L , cross-sectional area A , and Young's modulus E behaves as a spring of constant $k = \frac{EA}{L}$, so the spring formula carries over word for word once the spring constant is rewritten in the language of the wire. Substituting gives the energy stored in the wire when it is stretched by ΔL :

$$U = \frac{1}{2} \left(\frac{EA}{L} \right) (\Delta L)^2$$

This tells us how much energy a whole wire stores, but it depends on the dimensions of that particular wire. A more useful quantity is the energy stored per unit volume of the material, which is a pure property of the material itself and does not care about whether your sample is a short thick rod or a long thin filament. Dividing by the volume $V = AL$ and tidying up gives:

$$\frac{U}{V} = \frac{1}{2} \left(\frac{EA}{L \times AL} \right) (\Delta L)^2 = \frac{1}{2} \times E \times \left(\frac{\Delta L}{L} \right)^2 = \frac{1}{2} E \epsilon^2 = \frac{1}{2} \sigma \epsilon = \frac{\sigma^2}{2E}$$

Look at the form $\frac{U}{V} = \frac{1}{2} \sigma \epsilon$. That is the area of a triangle with base ϵ and height σ , which is exactly the area under a linear stress-strain curve from zero strain up to the current strain. Earlier, when we first met the idea of toughness, we claimed that the area under a stress-strain curve represents the energy stored per unit volume, and we promised a proper derivation later in the chapter. This is that promise kept. Toughness is just this same quantity evaluated at the fracture strain, and the area-equals-energy statement that we once had to take on trust has now earned its place honestly.

Concisely, **strain energy per unit volume** of an elastically stretched material *is the elastic potential energy stored in each unit of volume*. For a material within its linear elastic region it equals $\frac{1}{2} \sigma \epsilon$, where σ is the current stress and ϵ is the current strain. Its SI unit is the joule per cubic metre (J/m^3).

One formula, three forms, and a family of problems that all yield to the same triangle. Some come dressed as springs, some as wires, and some as catapults, and one or two will pretend the elastic energy has nothing to do with kinetic energy. Do not be fooled. The triangle is always the same triangle. Pick up your pen for worked examples.

BINDER Example 19

A spring of spring constant 200N/m is stretched from its natural length by 0.15m . Calculate the elastic potential energy stored in the spring.

Solution

$$U = \frac{1}{2} kx^2 = \frac{1}{2} (200\text{N/m})(0.15\text{m})^2 = 2.25\text{J}$$

The spring stores 2.25J at this extension.

Making Sense of the Answer: *Two and a quarter joules is roughly the kinetic energy of a cricket ball travelling at 7 metres per second, or the work done lifting a half-kilogram mass to waist height. The spring holds exactly that much and releases all of it the moment it is allowed to contract.*

Think Like a Physicist: *The x in $U = (1/2)kx^2$ is always the extension from the natural length, never from some other starting point such as the equilibrium position of a hanging mass. A spring hanging from a ceiling with a mass on the end has a natural length and a stretched length, and x is the difference between them.*

REAL Example 20

Kipanga has built a catapult by tying a rubber band between the two prongs of a Y-shaped stick. The rubber band behaves approximately as a Hookean spring of spring constant 80N/m . He loads the catapult with a small stone of mass 20g , pulls the rubber band back by 0.12m , and releases it. Find the speed of the stone as it leaves the catapult.

Solution

Elastic potential energy stored in the rubber band at the moment of release:

$$U = \frac{1}{2}kx^2$$

All of this converts to kinetic energy of the stone:

$$\frac{1}{2}mv^2 = \frac{1}{2}kx^2$$

$$v = \sqrt{\frac{kx^2}{m}} = x\sqrt{\frac{k}{m}} = 0.12\text{m}\sqrt{\frac{80\text{N/m}}{0.020\text{kg}}} = 7.59\text{m/s}$$

The stone leaves the catapult at approximately 7.6m/s.

Making Sense of the Answer: Seven and a half metres per second is a realistic catapult speed, fast enough to cause a noticeable impact but slow enough to dodge if you see it coming. The entire energy budget of the shot is less than one joule, converted cleanly from elastic to kinetic by the same conservation-of-energy principle that governs every collision and every projectile.

Think Like a Physicist: A physicist does not compute intermediate numerical values unless the problem specifically asks for them. Writing the energy balance $\frac{1}{2}kx^2 = \frac{1}{2}mv^2$ and cancelling the halves before substituting numbers is always cleaner than computing U first and then plugging it into the kinetic energy formula, because it keeps the final answer in terms of the quantities the problem actually gave you. Symbolic first, numerical last, is the default approach. It also makes the physics visible. The formula $v = x\sqrt{k/m}$ immediately tells you that doubling the extension doubles the speed, that a stiffer rubber band gives a faster shot, and that a heavier stone goes slower, without any arithmetic at all. Numbers hide that kind of insight behind a wall of digits; symbols put it in plain view.

HOT Example 21

A steel wire of length 2m, cross-sectional area $1 \times 10^{-6} \text{ m}^2$, and Young's modulus 200GPa is stretched by 1mm. Using both the spring-analogue method and the energy-density method, calculate the elastic potential energy stored in the wire and confirm that the two methods agree.

Solution

By using the spring-analogue method:

The wire is a spring in disguise, with spring constant $k = \frac{EA}{L}$, so the cleanest route to the energy is the spring formula $U = \frac{1}{2}kx^2$ with $x = \Delta L$. Substituting the wire's spring constant directly:

$$U = \frac{1}{2} \frac{EA}{L} (\Delta L)^2$$

Substituting the numbers in SI units:

$$U = \frac{1}{2} \frac{(200 \times 10^9 \text{Pa})(1 \times 10^{-6} \text{m}^2)}{2\text{m}} (1 \times 10^{-3} \text{m})^2$$

$$U = 0.05\text{J} = 50\text{mJ}$$

The wire stores 0.05J of elastic potential energy at this extension.

By using the energy-density method:

This is done by using the energy-per-unit-volume formula $\frac{U}{V} = \frac{1}{2}\sigma\epsilon$.

Computing the stress and strain from the given loading:

$$\sigma = E\epsilon = E \frac{\Delta L}{L} = (200 \times 10^9 \text{Pa}) \times \frac{1 \times 10^{-3} \text{m}}{2\text{m}} = 1 \times 10^8 \text{Pa}$$

$$\text{and the strain is } \epsilon = \frac{\Delta L}{L} = \frac{1 \times 10^{-3} \text{m}}{2\text{m}} = 5 \times 10^{-4}$$

So the energy per unit volume is:

$$\frac{U}{V} = \frac{1}{2}\sigma\epsilon = \frac{1}{2}(1 \times 10^8 \text{Pa})(5 \times 10^{-4}) = 2.5 \times 10^4 \text{J/m}^3$$

Multiplying by the wire's volume $V = AL = (1 \times 10^{-6} \text{ m}^2)(2\text{m}) = 2 \times 10^{-6} \text{ m}^3$:

$$U = \frac{U}{V} \times V = (2.5 \times 10^4 \text{ J/m}^3)(2 \times 10^{-6} \text{ m}^3) = 0.05\text{J}$$

Hence the two routes agree, giving 0.05J of elastic potential energy.

Making Sense of the Answer: 0.05J packed into two cubic millimetres of wire works out to twenty five thousand joules per cubic metre. A car tyre at normal pressure stores about a kilojoule of elastic energy, which is why a ruptured tyre can injure anyone standing nearby. Elastic energy density is one reason pressure vessel design is a serious discipline.

Think Like a Physicist: When a calculation can be done by two different formulas that ought to agree, doing both is not double work. It is the cheapest check available, and it catches errors that a single calculation would leave hidden. More deeply, the agreement of the two methods reinforces the link between the object-level description (the total energy in this particular wire) and the material-level description (the energy per unit volume of steel, multiplied by the wire's volume). The first depends on what you built; the second depends only on what you built it from. This intensive-versus-extensive distinction appears in every later chapter of this book, and recognising it early pays dividends long after you forget the details of any particular wire.

So far, we have three moduli, four strain formulas, and one energy equation written in six different ways. What we do not yet have is a laboratory apparatus that lets a Form Five student actually measure Young's modulus for themselves. The good news is that the apparatus is waiting for us in the next part of this chapter.

EXPERIMENTAL DETERMINATION OF YOUNG'S MODULUS

So far, Young's modulus has been a number that appears in formulas. We have used it to predict how much a wire will stretch under a given load, to compute the spring constant of a wire from its dimensions, and to calculate the elastic potential energy stored in a loaded rod. But every number in physics that deserves to be called a material property has to be measurable, and a student who has used Young's modulus repeatedly without ever having determined it for a real wire has been handed a gift without being shown the price tag. The purpose of this section is to show how that price is paid in the laboratory.

The measurement challenge

At first sight the measurement looks straightforward. Hang a known weight from a wire of known length and cross-section, measure how much the wire extends, and rearrange the equation for Young's modulus to obtain it in one step. The difficulty is hidden in the extension. For a steel wire two metres long and one millimetre in diameter, a load of fifty newtons produces an extension of less than a millimetre, often only a few tenths of a millimetre. Measuring such a small length with a ruler is impossible. Measuring it with a vernier caliper is only just possible. Measuring it accurately enough for a Young's modulus calculation requires an apparatus that was designed with this problem in mind.

The apparatus in question is Searle's two-wire arrangement. Its essential trick is that the quantity we actually need to measure is not the length of the wire itself but the change in the length of the wire. If we could somehow measure the change directly without first measuring the full length and then subtracting the initial length, most of the error would disappear. Searle's method does exactly that, and the rest of this section is devoted to understanding how it works and why every one of its features is there for a reason.

Searle's two-wire apparatus

Two identical wires are suspended side by side from the same rigid support overhead. Both wires are made from the same material and are cut from the same spool, so that they are as alike as a pair of wires can be. One of them is the test wire, which will carry the load whose effect we want to measure. The other is the reference wire, which is kept under a constant small load just heavy enough to keep it straight. A light rigid frame is attached to the lower end of each wire, and the two frames are connected to each other by a sensitive measuring device, traditionally a spirit level fitted with a micrometer screw gauge, though modern laboratory versions use a dial gauge or a digital displacement sensor in place of the spirit level.

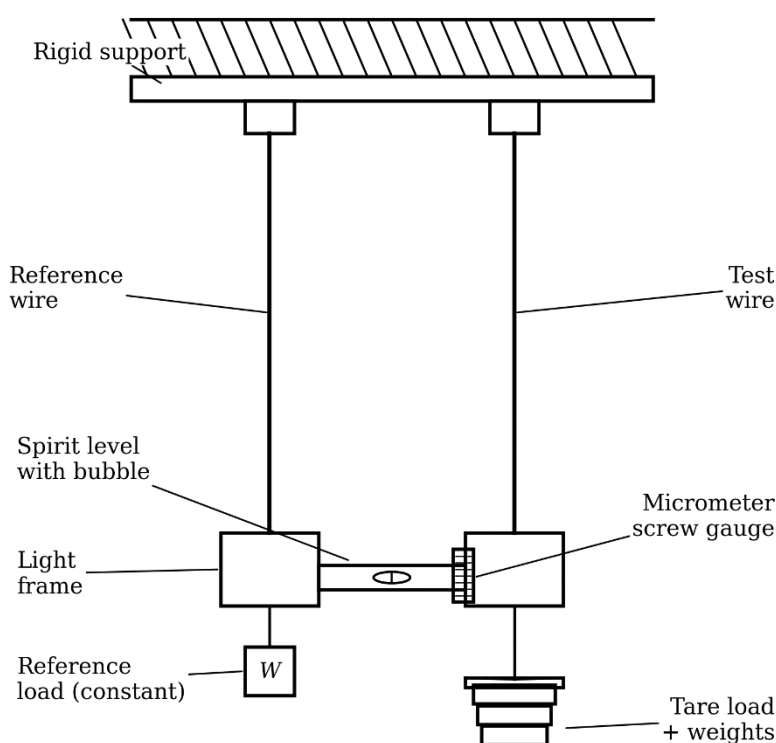


Figure: Searle's two-wire apparatus for measuring Young's modulus of a wire. Two identical wires hang side by side from the same rigid support. The left wire is the reference wire, kept under a constant small load W just heavy enough to keep it straight. The right wire is the test wire, carrying the tare load and any additional weights whose effect on its extension is to be measured. A light rigid frame is attached to the lower end of each wire, and the two frames are connected by a sensitive spirit level fitted with a micrometer screw gauge. Because both wires hang in the same air at the same temperature, any thermal expansion affects them equally and cancels out of the difference in their lengths. The spirit level therefore records only the extra extension of the test wire caused by the load, which is exactly the quantity needed for the calculation of Young's modulus.

The key insight behind the two-wire arrangement is thermal compensation. Whatever expansion or contraction happens to the reference wire during the experiment happens equally to the test wire, because the two are hanging in the same air, at the same temperature, at the same moment. If the room warms up by half a degree during the experiment, both wires lengthen by the same small amount, and the spirit level is not disturbed at all. The only thing that can make the two wires differ in length is the load applied to the test wire, and that is exactly what we want to measure. Thermal effects, which would otherwise contaminate the measurement of a single wire, cancel exactly.

The reference wire plays a second, quieter role as well. It absorbs any small unintended movements of the support beam from which both wires are suspended, for example from vibration or from the slow settling of the clamp during a long experiment. Whatever happens to the beam happens to the tops of both wires equally, and therefore does not appear in the difference.

Taking the readings

The procedure begins with a small constant load, called the tare load, already hanging from the test wire. The tare load is there to remove any initial slack or kinks in the wire. Without it, the first few increments of load would go into straightening the wire rather than stretching it, and the early data would sag below the straight line we expect. With the tare load in place, the spirit level is adjusted until the bubble sits in the middle of its scale, and the initial reading on the micrometer is recorded as the zero of the experiment.

Known weights are then added to the test wire in equal increments, typically half a kilogram or one kilogram per step. After each weight is added, the spirit level tips slightly because the test wire has stretched more than the reference wire. The micrometer is turned until the bubble is centred again, and the new reading is recorded. The difference between the new reading and the initial reading is the extension produced by the total load so far. Five or six such increments give a table of load and extension values, and the experiment is concluded by carefully removing the weights and checking that the wire has returned to its original reading.

If it has, the entire experiment was conducted within the elastic limit, and the data are trustworthy. If it has not, the wire was loaded beyond the elastic limit at some point, and the measurements must be repeated with smaller weights on a fresh wire.

From gradient to Young's modulus

The data are plotted with the load F on the vertical axis and the extension on the horizontal axis. Within the elastic limit, the graph is a straight line through the origin, because Hooke's law holds for the wire as a whole: the load is proportional to the extension. The gradient of the line is the force required to produce unit extension, which is exactly what we called the spring constant of the wire in the section on Young's modulus, and which we showed to equal:

$$\text{Gradient} = \frac{EA}{L}$$

Rearranging this relation for Young's modulus gives the result we are after:

$$E = \frac{\text{Gradient} \times L}{A}$$

The power of this method lies in the fact that Young's modulus is extracted from the slope of a best-fit straight line through several data points, not from a single load-extension measurement. A single measurement could be wildly wrong because of a sticky spirit-level bubble, a slipped weight, or a misread micrometer. A best-fit line through six points averages out such random errors and gives a value of Young's modulus that is far more reliable than any single-point calculation. This is why the experiment is always done as a graph, not as a single division.

Precautions and dominant sources of uncertainty

Four precautions matter more than all the others put together.

- 1) The initial tare load must always be present before the first reading is taken, so that the wire is being stretched rather than straightened when the extra weights begin to arrive.
- 2) The diameter of the wire must be measured in several places and in two perpendicular orientations at each place, and an average taken, because real wires are never perfectly uniform and never perfectly circular.
- 3) The experiment must never exceed the elastic limit of the wire, because once plastic deformation has begun, the wire will not return to its original length and all subsequent readings will be contaminated.
- 4) Parallax must be avoided when reading the spirit level, by looking at the bubble from directly above.

One of these precautions carries far more weight than the others, and it is not obvious which one until the error analysis is carried out. The reason is that the cross-sectional area of the wire depends on the square of the diameter:

$$A = \pi \left(\frac{d}{2}\right)^2 = \frac{\pi d^2}{4}$$

A small fractional error in the diameter therefore produces twice as large a fractional error in the area, and hence twice as large a fractional error in Young's modulus. When the individual uncertainties from the length, extension, and diameter are combined, the result is:

$$\frac{\Delta E}{E} = \frac{\Delta L}{L} + \frac{\Delta(\Delta L)}{\Delta L} + 2 \times \frac{\Delta d}{d}$$

Notice the factor of two in front of the diameter term. Even when the diameter is measured more precisely than any other quantity in the experiment, its contribution to the total uncertainty in Young's modulus is doubled before it is added to the sum. This is why students doing Searle's experiment are told, often with some irritation, to measure the diameter with a micrometer screw gauge and not with a vernier caliper. The reduced uncertainty in the diameter is worth more than an equivalent reduction in any of the other quantities, because the diameter enters the formula squared.

A note on the shear modulus

The experiment we have just described gives Young's modulus of a wire by a static method: weight is added, the extension is measured, the graph is plotted, and the slope is converted to Young's modulus. The

corresponding experiment for the shear modulus is not static but dynamic. It uses a wire suspended vertically from a rigid support with a horizontal disc attached at the bottom, and it measures the period of torsional oscillations of the disc as the wire twists back and forth. The shear modulus is then extracted from the period. This is a simple harmonic motion experiment, and it belongs naturally to the chapter on simple harmonic motion later in this book. We shall meet it again there, once simple harmonic motion has given us the language in which the period of a torsional oscillation can be properly written.

Three worked examples now put the apparatus and the error analysis to use.

BINDER Example 22

A Form Five student performs Searle's experiment on a steel wire of length 2.00m and diameter 0.80mm. After adding a tare load and zeroing the spirit level, the student records the following five (load, extension) values:

Load F (N)	10	20	30	40	50
Extension ΔL (mm)	0.20	0.38	0.60	0.78	1.00

Plot a suitable graph, determine its gradient, and use the gradient to calculate Young's modulus of the steel wire. Compare your value with the standard value of 2.0×10^{11} Pa for steel.

Solution

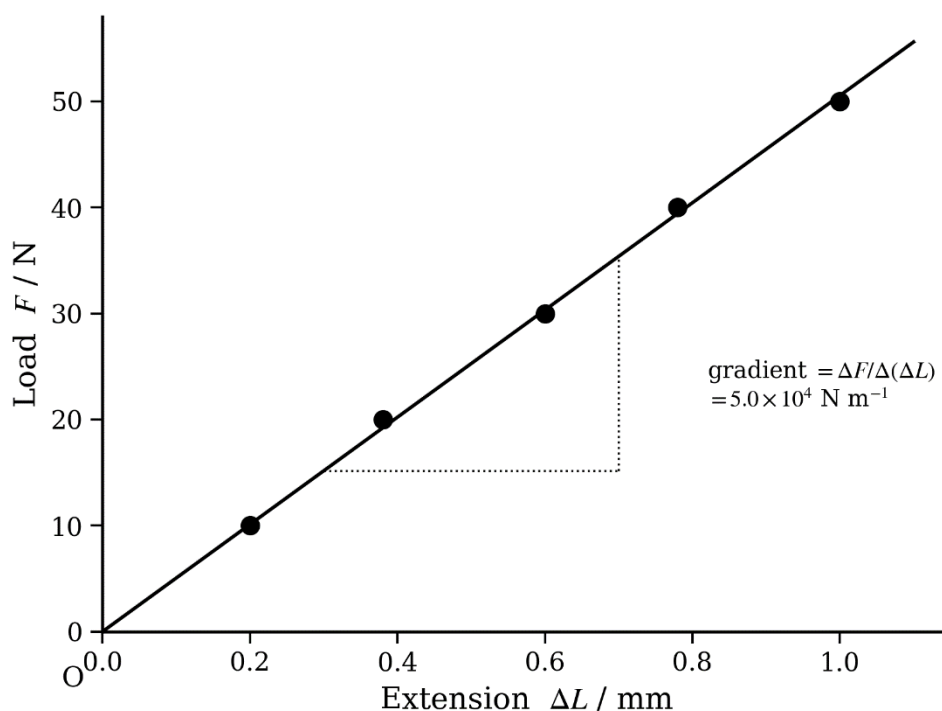


Figure: Student's plot of load F against extension ΔL from Searle's experiment on a steel wire of length 2.00m and diameter 0.80mm.

$$\text{Gradient} = \frac{(50 - 0)\text{N}}{(1.00 - 0)} \times 10^{-3} \text{ m} = 5.0 \times 10^4 \text{ N/m}$$

The cross-sectional area of the wire is calculated from its diameter using the formula for the area of a circle.

$$A = \frac{\pi d^2}{4} = \frac{\pi \times (0.80 \times 10^{-3} \text{ m})^2}{4} = 5.03 \times 10^{-7} \text{ m}^2$$

Young's modulus is then obtained by multiplying the gradient by the length of the wire and dividing by the cross-sectional area:

$$E = \frac{\text{gradient} \times L}{A} = \frac{5.0 \times 10^4 \text{ N/m} \times 2.00 \text{ m}}{5.03 \times 10^{-7} \text{ m}^2} = 1.99 \times 10^{11} \text{ Pa}$$

Therefore, Young's modulus of the steel wire is approximately 2.0×10^{11} Pa, which agrees with the standard value for steel to within the precision of the experiment.

Making Sense of the Answer: *The agreement between the experimental value and the standard value is the strongest possible evidence that the apparatus, the procedure, and the arithmetic are all working correctly. Notice that the final answer comes from the slope of the straight line, not from any single data point, so a bad reading at one or two increments would not have wrecked the result. This is the quiet magic of doing physics with graphs rather than with individual measurements.*

Think Like a Physicist: *Whenever a quantity can be extracted from the slope of a straight line fitted to several data points, prefer the slope method over a single-point calculation even when the single-point answer would seem easier to compute. The slope method automatically averages out random errors in individual readings and uses every data point as evidence, whereas the single-point method stakes the entire answer on one lucky measurement. This is one of the most transferable habits in all of experimental physics, and you will use it for every experiment in this book from now on.*

REAL Example 23

In Searle's experiment for measuring Young's modulus, a small constant weight known as the tare load is always hung from the test wire before the first reading is taken. A careful student might object that this practice is unfair, because the wire is already deformed when the so-called zero of the measurement is recorded on the micrometer. Explain the purpose of the tare load, describe what the load-extension data would look like if the tare load were omitted and the experiment were started from a bare wire, and state why the extension caused by the tare load itself does not need to be known or recorded for the final calculation of Young's modulus to be correct.

Solution

The purpose of the tare load is to straighten out any initial kinks, bends, or slack in the test wire before the measurement begins. A wire that has just been taken off a spool and clamped at both ends is almost never perfectly straight, and when a small load is first applied, most of that load is used to straighten the wire rather than to stretch it. The apparent extension in the first few increments of load is therefore much larger than Hooke's law would predict for a wire that is only being stretched, because the wire is simultaneously unbending.

If the tare load were omitted and the experiment were started from the bare wire, the early data points on a load-extension graph would lie well above the true Hooke's-law line, because the extensions recorded there would include both the stretching and the unbending. The early part of the graph would curve away from the straight line expected from Hooke's law, and only after several increments, when the wire was finally straight, would the graph settle into the correct linear behaviour. The gradient extracted from such a graph would be contaminated by the curved early section and would not correspond to the true spring constant of the wire, so that Young's modulus calculated from it would be too small by a significant amount.

The tare load solves this problem by doing the unbending in advance, before any readings are taken. Once the wire is straight, every additional load produces a genuine elastic extension, and the load-extension graph is a clean straight line from the very first data point. The extension caused by the tare load itself is not part of the measurement and does not need to be known, because the zero of the micrometer is defined with the tare load already in place. Everything that the experiment records is the extension beyond that zero, and everything that the calculation uses is the slope of the straight line, both of which are independent of the unknown initial extension caused by the tare load itself.

Making Sense of the Answer: *Kipanga's suspicion was not foolish, and it is worth honouring rather than dismissing. A careful student who looks at the apparatus and asks why a weight is hanging on the wire before any measurements are taken is behaving exactly as a physicist should. The resolution is that the tare load is not part of the measurement but part of the preparation, like cleaning a test tube before pouring a chemical into it. The measurement begins only after the preparation is complete, and the zero of the measurement is the state of the wire at the end of the preparation, not the state of the wire before it.*

Think Like a Physicist: *Every measurement in physics requires a zero, and the choice of zero is almost always more delicate than it looks. In Searle's experiment the zero is the position of the micrometer after the tare load has straightened the wire. In a thermometer the zero is the reading at a well-defined reference temperature. In an electrical measurement the zero is the reading with the circuit open. Whenever you encounter a new experimental apparatus, ask yourself first: what defines the zero, and what preparation has*

to be done before the zero can honestly be called zero? The answer to that question often contains more physics than the rest of the procedure put together.

HOT Example 24

A student performs Searle's experiment with the following measurements and their uncertainties: the length of the test wire is $(2.50 \pm 0.01)\text{m}$, the diameter of the wire is $(0.92 \pm 0.03)\text{mm}$ measured with a vernier caliper, and the extension produced by a load of 50N (known exactly) is $(0.90 \pm 0.05)\text{mm}$. Calculate (a) Young's modulus of the wire without uncertainty (b) the percentage uncertainty in Young's modulus and its absolute uncertainty (giving the final result in the standard notation), and (c) identify which of the three measured quantities contributes the most to the final uncertainty, and explain why.

Solution

(a) Young's modulus is given by Young's-modulus formula rearranged so that the area of the wire is written in terms of its diameter:

$$E = \frac{FL}{A \times \Delta L} = \frac{4FL}{\pi d^2 \times \Delta L}$$

Substituting the measured values in SI units: with the diameter squared as $(0.92 \times 10^{-3} \text{ m})^2 = 8.464 \times 10^{-7} \text{ m}^2$, and computing step by step:

$$E = \frac{4 \times 50\text{N} \times 2.50\text{m}}{\pi \times (0.92 \times 10^{-3} \text{ m})^2 \times 0.90 \times 10^{-3} \text{ m}} = 2.09 \times 10^{11} \text{ Pa}$$

Therefore, the value of Young's modulus, before the uncertainty is attached, is $2.09 \times 10^{11} \text{ Pa}$.

(b) Since the diameter enters the formula squared, its fractional uncertainty is doubled before it is added to the sum. The force is known exactly, so its contribution is zero:

$$\frac{\Delta E}{E} = \frac{\Delta L}{L} + \frac{\Delta(\Delta L)}{\Delta L} + 2 \times \frac{\Delta d}{d}$$

Substituting the individual fractional uncertainties:

$$\frac{\Delta E}{E} = \frac{0.01}{2.50} + \frac{0.05}{0.90} + 2 \times \frac{0.03}{0.92} = 0.125 = 12.5\%$$

The absolute uncertainty in Young's modulus is therefore:

$$\Delta E = 0.125 \times 2.09 \times 10^{11} \text{ Pa} \approx 0.26 \times 10^{11} \text{ Pa}$$

Rounding the value and its uncertainty to the same decimal place gives the final result:

$$E = (2.1 \pm 0.3) \times 10^{11} \text{ Pa}$$

(c) The three fractional contributions to the total uncertainty are 0.4 percent from the length, 5.6 percent from the extension, and 6.5 percent from the diameter (after doubling). The diameter is therefore the largest contributor, even though its raw fractional uncertainty of 3.3 percent was smaller than the 5.6 percent uncertainty in the extension. The reason is the factor of two that comes from squaring the diameter in the formula for the area of the wire.

Making Sense of the Answer: The final answer of $(2.1 \pm 0.3) \times 10^{11} \text{ Pa}$ is consistent with the standard value of Young's modulus for steel, which is about $2.0 \times 10^{11} \text{ Pa}$.

Think Like a Physicist: When a quantity in a formula is raised to a power, its fractional uncertainty is multiplied by that power before being added to the total. A quantity raised to the second power therefore contributes twice as heavily as a quantity that appears only once, and a quantity raised to the third power contributes three times as heavily. Whenever you design an experiment, identify the quantity with the highest power in the formula and spend your measurement effort on it first. A small improvement there is worth a much larger improvement in any of the linear quantities, and often a small improvement in one powered quantity is enough to turn an embarrassing experiment into a respectable one.

Searle's apparatus has shown us how the most quoted material constant in engineering is actually extracted from a real wire in a real laboratory. But wires in real laboratories and real bridges do not only experience loads. They also experience temperature changes, and a wire that is heated expands by an amount that is usually small but is never zero. If the wire is free to expand, the expansion is harmless. If the wire is clamped

between rigid supports and cannot expand, the temperature change produces a stress that can be large enough to buckle a railway line or crack a bridge beam. The next part of this chapter shows how to calculate the expansion produced by a given temperature change, the stress produced when the expansion is prevented, and what both calculations mean in the molecular picture of the potential energy curve with which this chapter began.

THERMAL EXPANSION

Linear expansion and the coefficient α

Almost every solid expands when it is heated and contracts when it is cooled. The effect is small enough to be invisible at the scale of a wristwatch but large enough at the scale of a railway bridge to rip the bridge apart if the engineer has not designed for it. For a rod of initial length L_0 that is heated through a temperature change ΔT , the increase in length ΔL is found experimentally to be proportional both to the original length and to the temperature change, and the constant of proportionality is a property of the material called its coefficient of linear expansion, α :

$$\Delta L = L_0 \alpha \Delta T$$

The coefficient α has the units of reciprocal temperature and is quoted in either K^{-1} or $^{\circ}C^{-1}$, which are numerically identical because a change of one kelvin equals a change of one degree Celsius.

The molecular origin of thermal expansion is already familiar from the very first section of this chapter. The potential energy curve between two atoms is not a symmetric parabola around its minimum but is steeper on the repulsive side than on the attractive side. As the temperature rises, the atoms vibrate with greater amplitude, and because the curve is asymmetric, the mean separation between atoms grows a little with every rise in temperature. The whole solid therefore becomes a little larger.

Area and volume expansion

When a flat plate is heated, it does not only become longer; every dimension in its plane grows, and the area grows with it. When a solid cube is heated, every edge grows and the volume grows. We do not need any new measurements to find the area and volume coefficients of expansion. We only need to apply the linear formula to the sides of the shape and keep the dominant terms.

Consider a square plate of side L_0 at the initial temperature. After heating through ΔT , each side has become longer by the amount $L_0 \alpha \Delta T$, so the new side length is:

$$L = L_0 + L_0 \alpha \Delta T = L_0(1 + \alpha \Delta T)$$

The new area of the plate is therefore the square of this new side:

$$A = L^2 = (L_0(1 + \alpha \Delta T))^2 = L_0^2 (1 + \alpha \Delta T)^2$$

Expanding the square $(1 + \alpha \Delta T)^2$ using the identity $(a + b)^2 = a^2 + 2ab + b^2$:

$$(1 + \alpha \Delta T)^2 = 1 + 2(\alpha \Delta T) + (\alpha \Delta T)^2$$

Now comes the key step. The product $\alpha \Delta T$ is an extremely small number for the temperature changes we meet in the laboratory or in ordinary engineering. For a steel bar heated by 50K, for instance, $\alpha \Delta T = (12 \times 10^{-6} K^{-1})(50K) = 6 \times 10^{-4}$, which is a tiny fraction. Squaring such a small number gives something even tinier: $(\alpha \Delta T)^2 \approx 3.6 \times 10^{-7}$, which is four thousand times smaller than $2(\alpha \Delta T)$ itself. To the precision of any realistic measurement we can therefore drop the $(\alpha \Delta T)^2$ term, keeping only the first-order term:

$$(1 + \alpha \Delta T)^2 \approx 1 + 2(\alpha \Delta T)$$

Substituting this back into the expression for the new area:

$$A = L_0^2(1 + 2(\alpha \Delta T)) = L_0^2 + 2L_0^2(\alpha \Delta T)$$

The first term on the right, L_0^2 , is the original area A_0 . The second term is therefore the change in area:

$$\Delta A = A - A_0 = 2A_0(\alpha \Delta T)$$

This has the same form as the linear expansion formula, with a new coefficient that is exactly twice the linear coefficient. We call this new coefficient the **coefficient of area** (or **superficial**) expansion and give it the symbol β :

$$\Delta A = A_0 \beta \Delta T, \text{ where } \beta = 2\alpha$$

The same reasoning applied to a cube of side L_0 gives the volume expansion. Each side becomes $L_0(1 + \alpha\Delta T)$ after heating, so the new volume is:

$$V = L^3 = (L_0(1 + \alpha\Delta T))^3 = L_0^3 (1 + \alpha\Delta T)^3$$

Expanding the cube using the binomial identity $(1 + x)^3 = 1 + 3x + 3x^2 + x^3$ with $x = \alpha\Delta T$:

$$(1 + \alpha\Delta T)^3 = 1 + 3(\alpha\Delta T) + 3(\alpha\Delta T)^2 + (\alpha\Delta T)^3$$

The terms $3(\alpha\Delta T)^2$ and $(\alpha\Delta T)^3$ are negligibly small next to $3(\alpha\Delta T)$ for the same reason as before, so we drop them and keep only the first-order term:

$$(1 + \alpha\Delta T)^3 \approx 1 + 3(\alpha\Delta T)$$

Substituting and simplifying:

$$V = L_0^3(1 + 3(\alpha\Delta T)) = V_0 + 3V_0(\alpha\Delta T)$$

$$\Delta V = 3V_0(\alpha\Delta T) = V_0 \gamma \Delta T, \text{ where } \gamma = 3\alpha$$

The two relations $\beta = 2\alpha$ and $\gamma = 3\alpha$ are worth memorising, not as isolated facts but as direct consequences of the linear formula. If you ever forget them in an examination, you can recover them in half a minute by applying the linear formula to a square or a cube and dropping the small terms, just as we have done here.

Thermal stress when expansion is prevented

If a rod is free to expand, its length changes according to the linear formula and no stress develops inside it. But suppose the rod is clamped rigidly between two walls and then heated. The rod wants to grow longer, but the walls will not allow it to. *What happens inside the rod?*

The argument has two-steps, and it is worth following slowly.

Step one: imagine for a moment that the walls are not there. The rod would expand freely by the amount:

$$\Delta L_{\text{free}} = L_0 \alpha \Delta T$$

Step two: now put the walls back. The walls must somehow force the rod to stay at its original length L_0 , even though the rod wanted to grow by ΔL_{free} . The only way the walls can do this is by applying an inward compressive force that squeezes the rod back by exactly the amount it wanted to grow. The compression produced by this force is given by the usual Young's modulus formula, rearranged for the change in length:

$$\Delta L_{\text{comp}} = \frac{FL_0}{AE}$$

For the rod to fit between the walls, the compression must exactly cancel the would-be expansion, so $\Delta L_{\text{comp}} = \Delta L_{\text{free}}$:

$$\frac{FL_0}{AE} = L_0 \alpha \Delta T$$

The original length L_0 appears on both sides of this equation and cancels. Dividing both sides by L_0 gives:

$$\frac{F}{AE} = \alpha \Delta T$$

Multiplying both sides by E and recognising that F/A is the thermal stress σ in the rod:

$$\sigma = \frac{F}{A} = E\alpha\Delta T$$

This is the thermal stress formula. It is compressive (negative in the sign convention that calls tension positive) and has the striking property that the length of the rod has completely disappeared from the answer. A steel rod clamped between rigid walls develops exactly the same thermal stress for a given temperature change whether the rod is one centimetre long or one hundred metres long, because both the wanted expansion and the strain that results from preventing it are proportional to L_0 , so the length simply cancels.

Composite rods and differential expansion

Two rods of different materials placed side by side and heated through the same temperature change expand by different amounts, because their coefficients of linear expansion differ. Two particular problems involving this differential expansion appear frequently in examinations and deserve careful derivation.

1) The constant-length-difference problem

Two rods, one of steel and one of copper, are placed side by side. At a reference temperature, the steel rod has length L_s and the copper rod has length L_c . We want the difference between their lengths to remain exactly the same at every temperature, so that the end of the steel rod is always a fixed distance away from the end of the copper rod. When the temperature changes by ΔT , the steel rod expands by $L_s \alpha_s \Delta T$ and the copper rod expands by $L_c \alpha_c \Delta T$. For the difference to remain constant, these two expansions must be exactly equal:

$$L_s \alpha_s \Delta T = L_c \alpha_c \Delta T$$

The common factor ΔT cancels from both sides, leaving a condition that involves only the lengths and the coefficients:

$$L_s \alpha_s = L_c \alpha_c$$

Rearranging for the ratio of the two lengths:

$$\frac{L_s}{L_c} = \frac{\alpha_c}{\alpha_s}$$

This is the key result. The two rod lengths must be in the inverse ratio of their coefficients of expansion: the rod with the smaller coefficient (steel) must be made longer, and the rod with the larger coefficient (copper) must be made shorter, so that they expand by equal amounts. This is exactly the working principle of the compensated clock pendulum, whose precision depends on the gap between its two supporting rods staying constant through all seasons of the year.

2) The clamped-composite problem

Two rods of different materials are joined end to end and the whole assembly is clamped rigidly between two walls. When the assembly is heated by ΔT , neither rod can expand, because the walls hold the two ends fixed. Our task is to find the thermal stress developed in the composite. Assume that the two rods have the same cross-sectional area A . We build the argument in four short steps.

Step 1. Compute the free expansion. If the walls were absent, each rod would expand freely according to the linear formula. The total free expansion of the composite is the sum of the two individual expansions:

$$\Delta L_{\text{free}} = L_1 \alpha_1 \Delta T + L_2 \alpha_2 \Delta T$$

Step 2. Recognise that the walls transmit the same force through both rods. A compressive force F applied by the walls at the outer ends must also act at the joint between the two rods, because the joint is in equilibrium. Since both rods have the same cross-sectional area A , the stress $\sigma = F/A$ is the same in both rods, even though the strain in each is different.

Step 3. Compute the compression of each rod under this stress. Using Young's modulus for each material, the strain in rod 1 is σ/E_1 and the strain in rod 2 is σ/E_2 . The compression of each rod is its strain times its length:

$$\Delta L_1 = \left(\frac{\sigma}{E_1}\right) L_1 = \frac{\sigma L_1}{E_1}$$

$$\Delta L_2 = \left(\frac{\sigma}{E_2}\right) L_2 = \frac{\sigma L_2}{E_2}$$

Step 4. Set the total compression equal to the total free expansion, because the walls force the net change in length to be zero:

$$\frac{\sigma L_1}{E_1} + \frac{\sigma L_2}{E_2} = L_1 \alpha_1 \Delta T + L_2 \alpha_2 \Delta T$$

Factoring out σ on the left and solving for it:

$$\sigma = \frac{(L_1 \alpha_1 + L_2 \alpha_2) \Delta T}{\frac{L_1}{E_1} + \frac{L_2}{E_2}}$$

As a consistency check, set both materials equal: $E_1 = E_2 = E$ and $\alpha_1 = \alpha_2 = \alpha$. The numerator becomes $(L_1 + L_2)\alpha\Delta T$ and the denominator becomes $(L_1 + L_2)/E$, so the stress reduces to $\sigma = E\alpha\Delta T$,

the single-material formula from the previous subsection. The composite formula is therefore a clean generalisation of the simpler result.

Bimetallic strips and applications

A bimetallic strip is a thin bar formed by bonding two layers of different metals along their entire length. When the strip is heated, the metal with the larger coefficient of expansion tries to grow more than the metal with the smaller coefficient, but the bond between them prevents either layer from sliding over the other. The strip therefore bends into an arc, with the metal of larger α on the outside of the curve (because the outside of a curve is longer than the inside) and the metal of smaller α on the inside. On cooling below the reference temperature, the strip bends the other way. This simple device is the heart of almost every mechanical thermostat: the bending strip opens or closes an electrical contact as the temperature crosses a set point, and by choosing the two metals and the thickness of the strip, the designer can tune the set point over a wide range.

Thermal expansion shows up in many other pieces of everyday engineering. Expansion joints are built into long railway lines, concrete slab roads, and bridges to give the material somewhere to go when it grows. The Selander Bridge in Dar es Salaam uses finger joints between its concrete deck spans for exactly this reason. Hot rivets were historically driven through holes in steel plates and allowed to cool in place, so that as they contracted they pulled the plates tightly together with an enormous clamping force. Shrink-fitting exploits the same effect: a steel ring is heated until it expands enough to slip over a shaft, then cools and grips the shaft so tightly that it can transmit large torques without keys or welds. Railway wheels, turbine discs, and gear rings are all assembled by shrink-fitting.

Ideas, like rails, expand when heated; and these ones are ready to buckle. Let us clamp them down with a few worked examples.

BINDER Example 25

A length of steel railway rail is installed with its two ends rigidly clamped so that it cannot change its length. The rail is laid at 20°C and the ambient temperature then rises to 65°C during the hottest part of the afternoon. Taking the Young's modulus of steel as $2 \times 10^{11}\text{Pa}$ and its coefficient of linear expansion as $12 \times 10^{-6}\text{K}^{-1}$, calculate the compressive thermal stress developed in the rail.

Solution

The temperature change is $\Delta T = 65^\circ\text{C} - 20^\circ\text{C} = 45\text{K}$. Because the rail cannot change its length, the thermal expansion that would naturally occur is cancelled by an equal compressive strain, and the resulting thermal stress is given by the formula $\sigma = E\alpha\Delta T$. Substituting:

$$\sigma = E\alpha\Delta T = (2 \times 10^{11}\text{Pa})(12 \times 10^{-6}\text{K}^{-1})(45\text{K}) = 1.08 \times 10^8\text{Pa}$$

The thermal stress developed in the clamped rail is therefore $1.08 \times 10^8\text{Pa}$, (or 108MPa).

Making Sense of the Answer: *The answer depends only on the Young's modulus, the coefficient of linear expansion, and the temperature change. Notice that the length of the rail has vanished from the calculation: a clamped steel rail develops the same thermal stress of 108MPa for a 45K temperature rise whether the rail is a centimetre long or a kilometre long. What changes with length is not the stress but the tendency to buckle sideways when that stress is applied.*

Think Like a Physicist: *Whenever a thermal stress calculation gives a large number, ask whether the actual failure mode will match what the calculation assumed. Direct crushing is only one way a clamped rod can give way. Lateral buckling, joint slippage, and fracture at pre-existing flaws can all occur at stresses far below the raw compressive strength, especially in long slender pieces. A stress well below the crushing limit is not a guarantee of safety, only a single box that has been ticked.*

REAL Example 26

Expansion gaps are built into every long railway line and into the concrete slab roads of the Morogoro highway. Explain in physical terms why these gaps are necessary, what would go wrong if they were omitted, and why, despite their presence, the gaps do not produce a noticeably bumpy ride for trains or for vehicles travelling along the road at normal ambient temperatures.

Solution

A continuous rail or continuous concrete slab that is rigidly held at its ends cannot expand freely when the temperature rises, and instead develops a compressive thermal stress of magnitude $\sigma = E\alpha\Delta T$. For steel rails heated from a cool morning temperature of 20°C to an afternoon temperature of 60°C, this thermal stress is of the order of 10^8 Pa , which is not enough to crush the steel but is easily enough to cause lateral buckling of a slender rail fixed to its sleepers. A buckled rail twists sideways into a sinuous shape and can derail a train. Concrete slabs subjected to the same kind of thermal stress crack in compression during the hottest part of the day, and can also crack in tension on cold nights if they are bonded to neighbouring slabs, because concrete is much weaker in tension than in compression.

Expansion gaps cure both problems by giving the material somewhere to grow. When the rail warms up, it expands into the gap and no thermal stress develops. The size of the gap is chosen so that it just closes at the highest temperature ever expected on that stretch of line, with a small safety margin. A gap that is too small will still allow the rail to buckle on the hottest day; a gap that is too large will widen uncomfortably on cold nights, when the rail contracts.

The gaps do not cause a bumpy ride under normal conditions because they are very small compared with the size of a train wheel or a vehicle tyre. A typical expansion gap on a railway line is only a few millimetres wide at normal temperatures, and the contact patch between the wheel and the rail is larger than this, so the wheel always has most of itself resting on solid rail as it rolls across the gap. A small but audible click is sometimes heard when the wheels cross each joint, especially at low temperatures when the gap is wider, but the jolt is too small to feel in a moving carriage. Concrete slab roads use the same principle: the gaps are thin enough that tyres bridge them easily, and the flexible sealing compound that fills them further smooths the transition.

Making Sense of the Answer: *The design of an expansion gap is a quiet compromise between two different mechanical failures. Without any gap at all, the material fails in compression at high temperature. With too large a gap, the wheels begin to jolt and the joints themselves become damaged by repeated impact. The correct size is the one that just closes on the hottest expected day and never opens dangerously wide on the coldest expected night.*

Think Like a Physicist: *Whenever a civil engineering structure carries a large load over a long distance, look for the thermal compromise hidden somewhere in its geometry. Bridges have expansion joints at their ends or in the middle. Railway lines have gaps between rails. Long pipelines have zigzag bends or coils. Power lines sag deliberately so that they can contract on cold nights without snapping. In every case, the same physical question is being answered: where will the thermal expansion go, and how can we let it go there without damaging anything?*

HOT Example 27

At a temperature of 20°C, a steel rod is exactly 5.0cm longer than a copper rod placed beside it. The rods are to be used in a precision instrument in which the difference in their lengths must remain constant at every temperature. Given that $\alpha_s = 12 \times 10^{-6} \text{ K}^{-1}$ and $\alpha_c = 17 \times 10^{-6} \text{ K}^{-1}$, calculate the actual length of each rod at 20°C.

Solution

For the difference in length to remain the same at every temperature, both rods must expand by exactly the same absolute amount whenever they are heated by the same ΔT . Setting the two linear expansions equal:

$$L_s \alpha_s \Delta T = L_c \alpha_c \Delta T$$

The common factor ΔT cancels from both sides, leaving:

$$L_s \alpha_s = L_c \alpha_c$$

Substituting the numerical values of the coefficients:

$$L_s \times 12 = L_c \times 17$$

Rearranging gives the length ratio $\frac{L_s}{L_c} = \frac{17}{12}$. The second condition given in the question, that the steel rod is 5.0cm longer than the copper rod, provides a second equation:

$$L_s - L_c = 5.0 \text{ cm}$$

Substituting $L_s = \left(\frac{17}{12}\right) L_c$ from the first equation into the second:

$$\left(\frac{17}{12}\right) L_c - L_c = 5.0\text{cm}$$

Solving for L_c :

$$L_c = 12\text{cm}$$

And therefore $L_s = L_c + 5.0\text{cm} = 12\text{cm} + 5.0\text{cm} = 17\text{cm}$.

The copper rod is 12cm long and the steel rod is 17cm long at 20°C.

Making Sense of the Answer: Both rods expand when the instrument is heated, but because the copper rod is shorter and the steel rod is longer in just the right proportion, they expand by identical absolute amounts. The 5.0cm gap between their free ends therefore remains exactly 5.0cm at every temperature. The particular length difference of 5.0cm was incidental; any desired gap can be arranged by scaling both rods up or down while keeping their length ratio at 17:12.

Think Like a Physicist: The general lesson of this problem is that whenever two effects must cancel at every temperature, it is their rates of change that must match, not their absolute values. Setting rates equal at the outset converts a physics problem about expansion into an algebra problem about ratios, and the algebra is usually easier than the physics.

HOT Example 28

A solid steel cube of side 5.0cm at 20°C is held inside a perfectly rigid steel jig that prevents the cube from expanding in any direction. The cube is then heated uniformly to 70°C. The bulk modulus of steel is $1.6 \times 10^{11}\text{Pa}$ and its coefficient of linear expansion is $12 \times 10^{-6}\text{K}^{-1}$. (a) Derive an expression for the hydrostatic compressive pressure that the jig must exert on the cube to prevent it from expanding, in terms of the bulk modulus B , the coefficient of linear expansion α , and the temperature change ΔT . (b) Calculate the numerical value of this pressure. (c) Calculate the total inward force exerted on each face of the cube by the jig.

Solution

(a) The temperature change is $\Delta T = 70^\circ\text{C} - 20^\circ\text{C} = 50\text{K}$. The argument has two-steps, exactly parallel to the linear case but applied in three dimensions instead of one.

Step 1: Find the volume the cube would occupy if it were free to expand

If the jig were not there, the cube would expand in all three directions according to the volume coefficient of expansion $\gamma = 3\alpha$. The fractional change in volume would be:

$$\frac{\Delta V_{\text{free}}}{V_0} = \gamma \Delta T = 3\alpha \Delta T$$

Step 2: Find the pressure required to compress the cube back to its original volume

The jig holds the cube at exactly its original volume V_0 , so the cube cannot grow at all. The fractional volume change must therefore be cancelled by an equal and opposite fractional volume change produced by an inward hydrostatic pressure. From the definition of the bulk modulus:

$$B = -\frac{\Delta P}{\frac{\Delta V}{V}}$$

The pressure required to compress the cube by a fractional volume of $-\Delta V_{\text{free}}/V_0 = -3\alpha \Delta T$ is therefore:

$$\Delta P = -B \times \frac{\Delta V}{V_0} = -B(-3\alpha \Delta T) = 3B\alpha \Delta T$$

So the hydrostatic compressive pressure exerted by the jig on the cube is:

$$\boxed{\Delta P = 3B\alpha \Delta T = B\gamma \Delta T}$$

This is the three-dimensional analogue of the linear thermal stress formula $\sigma = E\alpha \Delta T$ that appeared in the example 25. The Young's modulus E has been replaced by the bulk modulus B , and the linear coefficient α has been replaced by the volume coefficient $\gamma = 3\alpha$, which is exactly what we should expect when the geometry generalises from one axis to three.

(b) Substituting the numerical values with units carried at every step:

$$\Delta P = 3(1.6 \times 10^{11} \text{Pa})(12 \times 10^{-6} \text{K}^{-1})(50 \text{K}) = 2.88 \times 10^8 \text{Pa}$$

The hydrostatic pressure that the jig must exert on the cube is therefore approximately $2.88 \times 10^8 \text{Pa}$, or about 288MPa.

(c) The area of one face of the cube is $A = (5.0 \times 10^{-2} \text{m})^2 = 2.5 \times 10^{-3} \text{m}^2$.

The total inward force exerted on each face is the pressure multiplied by the area:

$$F = \Delta P \times A = (2.88 \times 10^8 \text{Pa})(2.5 \times 10^{-3} \text{m}^2) = 7.2 \times 10^5 \text{N}$$

So each of the six faces of the cube is being pushed inward by a force of approximately $7.2 \times 10^5 \text{N}$.

Making Sense of the Answer: The pressure of 288MPa is about three times larger than the 108MPa thermal stress we found for a clamped steel rail in the linear case under the same temperature rise. The factor of three is not a coincidence: it appears because the cube is being prevented from expanding in three independent directions, not just one. The bulk modulus B relates a hydrostatic pressure to a fractional volume change, and the volume change for the same temperature rise is three times larger than the length change, so the pressure required to cancel it is correspondingly larger. The formula $\Delta P = 3B\alpha\Delta T$ encodes both effects at once.

Think Like a Physicist: When a solid is constrained against expansion, the right modulus to use depends on which dimensions are being held fixed. If only one length is held fixed (a clamped rod), use Young's modulus and the linear coefficient: $\sigma = E\alpha\Delta T$. If all three dimensions are held fixed (a clamped block, or a fluid sealed inside a rigid container), use the bulk modulus and the volume coefficient: $\Delta P = B\gamma\Delta T = 3B\alpha\Delta T$. The two formulas have the same algebraic shape because the underlying physics is the same because prevented expansion produces stress proportional to the modulus that resists that kind of deformation. Whenever a thermal stress problem appears, the first question to ask is how many dimensions of the solid are constrained. The answer tells you which modulus to reach for.

HOT Example 29

A composite rod is made by joining a copper rod of length 30cm to a steel rod of length 20cm end to end. The two rods have the same cross-sectional area, and the composite rod is clamped rigidly between two walls at 20°C . The whole assembly is then heated to 70°C . Calculate the thermal stress developed in the composite, given that for copper $E_c = 1.2 \times 10^{11} \text{Pa}$ and $\alpha_c = 17 \times 10^{-6} \text{K}^{-1}$, and for steel $E_s = 2.0 \times 10^{11} \text{Pa}$ and $\alpha_s = 12 \times 10^{-6} \text{K}^{-1}$.

Solution

The temperature change is $\Delta T = 70^\circ\text{C} - 20^\circ\text{C} = 50 \text{K}$.

The total free expansion that would occur if the walls were absent is the sum of the free expansions of the copper and steel sections:

$$\Delta L_{\text{free}} = L_c \alpha_c \Delta T + L_s \alpha_s \Delta T$$

Substituting the numerical values with units:

$$\Delta L_{\text{free}} = (0.30 \text{m})(17 \times 10^{-6} \text{K}^{-1})(50 \text{K}) + (0.20 \text{m})(12 \times 10^{-6} \text{K}^{-1})(50 \text{K}) = 3.75 \times 10^{-4} \text{m}$$

The walls transmit the same compressive force through both rods, so the stress $\sigma = F/A$ is the same in both, even though the strain in each is different. So the compression of each rod, using Young's modulus:

$$\Delta L_c = \frac{\sigma L_c}{E_c}, \quad \Delta L_s = \frac{\sigma L_s}{E_s}$$

The total compression must equal the total free expansion, because the walls hold the ends of the composite fixed:

$$\frac{\sigma L_c}{E_c} + \frac{\sigma L_s}{E_s} = \Delta L_{\text{free}}$$

Making σ the subject:

$$\sigma = \frac{\Delta L_{\text{free}}}{\frac{L_c}{E_c} + \frac{L_s}{E_s}}$$

Substituting:

$$\sigma = \frac{3.75 \times 10^{-4} \text{m}}{\frac{0.30 \text{m}}{1.2 \times 10^{11} \text{Pa}} + \frac{0.20 \text{m}}{2.0 \times 10^{11} \text{Pa}}} = 1.07 \times 10^8 \text{Pa}$$

The thermal stress in the composite rod is therefore approximately $1.07 \times 10^8 \text{Pa}$, (or about 107MPa), compressive throughout both sections.

Making Sense of the Answer: *The thermal stress in this two-material composite is almost identical to the 108MPa we found for a single clamped steel rail in Example 26. This is not a coincidence. Copper expands more than steel but has a smaller Young's modulus, so its extra expansion is roughly cancelled by its greater willingness to be compressed. The simpler single-material formula $\sigma = E\alpha\Delta T$ is therefore a reasonable first approximation for any composite whose components are not wildly different, and the more careful composite formula refines that first approximation rather than overturning it.*

Think Like a Physicist: *The key step in composite-rod problems is recognising that when two pieces share the same cross-section and the same clamping force, the stress is the same in both pieces even though the strain is not. Stress is force per unit area, which is transmitted unchanged from one piece to the next across the joint. Strain is how much each material yields to that stress, which depends on its stiffness. Keeping these two quantities separate is the difference between a clean solution and a paragraph of confusion.*

Thermal expansion now joins Young's modulus, the stress-strain curve, and Searle's apparatus on the list of ideas that form the working physics of elastic solids. The last remaining task for this chapter is to step back from the derivations and calculations, gather together the engineering and biological situations in which elasticity is quietly doing its work, and ask what the whole chapter has been preparing us to understand.

APPLICATIONS OF ELASTICITY

The ideas built up in the previous sections are not abstract. Every modulus, every stress, every strain, and every coefficient of expansion has a job to do in the world outside the physics classroom. This final section of the chapter gathers those jobs together and shows how the single body of elasticity theory underwrites the design of buildings and lifts, the shape of the beams that hold up a floor, the mechanical operation of the human body, and the slow failure of springs and cables through years of use. None of these applications requires new physics. They require only the careful application of what has already been learned.

1. Safety factors in engineering design

Every engineering structure is designed to carry a maximum expected load, and every material has a stress beyond which it will yield or fracture. In a perfect world, a structure designed to the exact yield stress of its material would work correctly at its design load and fail by only the tiniest margin. In the real world this is never done, because the real world contains hidden sources of error: material flaws beneath the surface, loads that fluctuate unpredictably, temperature changes, corrosion, manufacturing variations, and the occasional careless operator. To protect against all these sources of uncertainty, engineers divide the material's yield stress (or ultimate tensile strength) by a number called the safety factor, and use the result as the working stress that the structure must never exceed in normal service:

$$\sigma_{\text{working}} = \frac{\sigma_{\text{yield}}}{n}$$

Here n is the safety factor, typically between 2 and 10 depending on the application. For ordinary structural steel beams in a building, safety factors of 2 or 3 are common. For lift and elevator cables, which must carry people, the factor is usually between 4 and 6. For aircraft components, whose failure can be fatal and whose load history is uncertain, safety factors of 8 or more are standard, and for crane cables lifting overhead loads above workers, factors of up to 10 are required by law. The exact value chosen depends on the consequences of failure, the reliability of the load estimate, the quality of the material, and the cost of adding extra strength. A higher safety factor is always safer but always more expensive.

The procedure for using a safety factor in design is always the same. Start from the worst-case load F that the component must be able to carry. Compute the corresponding working stress σ_{working} from the yield stress and the chosen safety factor. Rearrange the definition of stress to find the minimum cross-sectional

area $A = \frac{F}{\sigma_{\text{working}}}$ needed to keep the stress at or below the working value. If the component is a circular wire or cable, convert the area to a diameter using $d = 2\sqrt{\frac{A}{\pi}}$. The result is the smallest diameter that will survive the worst-case load with the chosen margin of safety.

2. Structural geometry and the I-beam

The stress inside a loaded beam is not uniform across its cross-section. When a horizontal beam bends under a downward load, the material at the top of the beam is compressed, the material at the bottom is stretched, and the material along the middle horizontal line (called the neutral axis) experiences no stress at all. The stress at any point in the cross-section grows in proportion to its distance from the neutral axis, which means that the material near the top and bottom of the beam is doing most of the work of resisting the load, while the material along the middle is almost idle.

Engineers exploit this fact by removing the idle material from the middle and concentrating it at the top and bottom, where it earns its place. The result is the I-beam, whose cross-section resembles the capital letter I: two broad horizontal flanges at the top and bottom, connected by a thin vertical web in the middle. An I-beam uses much less material than a solid rectangular beam of equivalent bending strength, which is why railway rails, bridge girders, floor joists, and the frames of tall buildings are almost always I-shaped. The rectangular beam is simpler but wasteful. The I-beam is a direct application of stress analysis to the economy of structural design.

3. Biological elasticity: four tissues, four stiffnesses

The body of every animal is a collection of elastic structures, each one tuned by evolution to a specific mechanical task. Four tissues in the human body illustrate the range of Young's moduli found in living material, spanning nearly five orders of magnitude from the softest to the stiffest.

Tendons connect muscles to bones and must transmit the full force of a contracting muscle without stretching significantly, because any stretch would waste energy that the muscle meant to deliver to the bone. Tendons therefore have a high Young's modulus of about 1.5GPa, which is stiff enough to behave almost like a rigid rope under normal loads. A tendon stretching by even 4% is close to rupture, which is why tendon injuries are so common in athletes.

Arteries must do the opposite of what tendons do. They must stretch and contract with every heartbeat to smooth the pulsatile flow of blood from the ventricles into a steady flow in the smaller vessels. Arteries therefore have a Young's modulus of only about 0.5MPa, roughly three thousand times smaller than that of tendons, and strains of 10% to 15% during each cardiac cycle are normal. With age, the walls of arteries become progressively stiffer, a condition called arteriosclerosis, and the artery loses its ability to buffer the pressure pulse. This is one reason why elderly patients are more prone to high blood pressure and to the rupture of small blood vessels in the brain.

Cartilage covers the ends of bones at every joint and acts as a shock absorber during walking, running, and jumping. Its Young's modulus of about 1MPa is intermediate between tendon and artery, giving it enough firmness to bear load and enough compliance to cushion impacts. Cartilage has no blood supply of its own and therefore heals slowly once damaged, which is why knee and hip cartilage injuries often become chronic.

Bone is the stiffest biological tissue, with a Young's modulus of about 20GPa for compact cortical bone. This is as stiff as concrete and a little stiffer than wood along the grain. The high stiffness of bone allows it to form the rigid skeleton that supports the body, and its strength per unit mass is actually better than that of most metals, which is one reason a bird's hollow bones can be so light and yet carry the whole weight of the bird in flight.

Put together, the four tissues span nearly five orders of magnitude in stiffness: artery at 0.5MPa, cartilage at 1MPa, tendon at 1.5GPa, and bone at 20GPa. Each stiffness is matched to a mechanical job. A body built entirely from tendon would have no cushioning at the joints and no pressure buffering in the arteries. A body built entirely from artery-grade tissue would collapse under its own weight. Evolution has chosen the right material for each part of the body by picking the right value of Young's modulus.

4. Elastic fatigue in everyday life

Elastic fatigue is the quiet failure mode that catches engineers and students by surprise. A material that has been loaded many times at a stress well below its breaking point will eventually fail, because tiny cracks at

microscopic flaws grow a little with each loading cycle until one of them becomes large enough to cause sudden fracture. The failure is not preceded by any visible warning, which is why elastic fatigue has been responsible for many historical engineering disasters and continues to determine the inspection schedules of aircraft, bridges, and pressure vessels today.

A spring balance in daily use at a market loses its zero progressively over months and years, not because the spring has ever been overloaded in any single use but because thousands of loading and unloading cycles have accumulated microscopic plastic deformation at internal crack tips. A spring balance that reads two hundred grams when empty is not broken; it is suffering from the same condition in slow motion that kills aircraft wings in years and bridge cables in decades. The difference is only the number of cycles required for the damage to become visible.

Aircraft wings flex slightly with every takeoff, landing, and gust of turbulence. The fatigue life of a wing spar is measured in flight cycles, not flight hours, and the number of cycles at which fatigue cracks become detectable is what determines how often the aircraft must be inspected and eventually retired. The first generation of commercial jet aircraft in the 1950s included the de Havilland Comet, three of which broke up in flight because stress concentrations at the corners of their square windows caused fatigue cracks to grow and propagate catastrophically. The disaster taught engineers to round off every corner and changed aircraft design forever.

Human bone is also vulnerable to elastic fatigue. A runner who suddenly increases training mileage can develop stress fractures in the tibia or metatarsals after weeks or months of cumulative microscopic bone damage, even though no single training session exceeded what the bone could withstand. The same physics that kills a bridge cable in fifty years can break an athlete's shin in three months, and the treatment is the same in both cases: stop loading, let the cracks heal, and do not repeat the pattern that caused them.

REAL Example 30

A spring balance used daily in a busy Kariakoo market reads 200g when empty, even though it read zero when it was new several years ago. The trader has never hung a load on the balance that was anywhere near its rated maximum, so the spring has never been overloaded in any single use. Explain the physical process responsible for the progressive loss of the zero reading, identify how it differs from ordinary overloading, and suggest one practical step that the trader could take to slow the rate at which the error develops.

Solution

The progressive loss of the zero reading is caused by elastic fatigue in the steel spring of the balance. Every time a load is hung on the balance and then removed, the spring is stretched and then released. On any single cycle, the stretch is well within the elastic limit of the steel, and the spring appears to return exactly to its original length. In reality, microscopic flaws on the inside and outside surfaces of the wire grow a tiny amount with each cycle. The growth is too small to see or to detect in a single measurement, but it accumulates slowly over thousands of cycles into a permanent plastic deformation of the spring, which is why the unloaded length of the spring gradually grows longer and the pointer slowly drifts away from zero.

This is completely different from the damage caused by ordinary overloading. Overloading means stretching the spring beyond its elastic limit in a single event, which produces an immediate and obvious permanent deformation. Elastic fatigue, by contrast, produces no damage that any single measurement can detect, and is caused by repeated loading well within the elastic limit. Elastic fatigue damage accumulates cycle by cycle over a long period of time, while overloading damage happens all at once. The two processes produce similar-looking results at the end, but they arise from very different causes, and the first is much harder to guard against than the second.

To slow the rate at which the zero error develops, the trader could remove the load from the balance whenever it is not in active use, instead of leaving heavy goods hanging on the hook between customers. Each hour that the spring sits under load is an hour in which the crack tips inside the metal are being stressed, even though no measurement is being taken. Unloading the balance between uses reduces the total number of stress-hours the spring experiences over its life and extends the time before fatigue becomes significant. A second step, just as valuable, is to recalibrate the balance regularly against a known standard mass and to apply a correction to every reading until the day comes when the error becomes too large to correct, at which point the spring must be replaced.

Making Sense of the Answer: *The steady growth of the zero error is a direct analogue of the cracks that grow in an aircraft wing spar or a bridge cable. Markets, aircraft, and bridges all fail by the same mechanism, separated only by the number of cycles required for the damage to become visible. The spring*

balance has a short fatigue life because it is cycled many times each day; the bridge cable has a long fatigue life because it is cycled only a few times per minute as traffic passes over it. Both are on the same clock, just set to different speeds.

Think Like a Physicist: Whenever an instrument is said to be 'in need of calibration' after long use, ask whether the underlying cause is elastic fatigue rather than damage. Elastic fatigue cannot be prevented by any amount of careful use; it can only be slowed down by reducing the number of loading cycles, reducing the amplitude of each cycle, or replacing the component before the damage becomes catastrophic. The trader's habit of unloading the balance between customers is in the first category. The inspection schedule of an aircraft fleet is in the third. Both are expressions of the same underlying idea: material fatigue is something you manage, not something you cure.

HOT Example 31

A passenger lift is to be designed to carry a maximum live load of 600kg (including passengers and goods), supported by four identical parallel steel cables. The yield stress of the steel used for the cables is $2.5 \times 10^8 \text{ Pa}$, and the safety factor required by the local building code is 4. Taking $g = 9.8 \text{ m/s}^2$, calculate (a) the total weight of the maximum live load, (b) the maximum force that each cable must support, (c) the maximum working stress allowed in each cable, and (d) the minimum diameter of each cable.

Solution

(a) The total weight W of the maximum live load is the mass times the gravitational field strength:

$$W = mg = (600\text{kg})(9.8\text{m/s}^2) = 5880\text{N}$$

The maximum live load therefore exerts a downward weight of 5880N on the four cables combined.

(b) The four cables share the load equally, so the force on each one is one quarter of the total:

$$F = \frac{W}{4} = \frac{5880\text{N}}{4} = 1470\text{N}$$

Each cable must therefore support a force of 1470N in normal service with the maximum live load.

(c) The maximum working stress is the yield stress divided by the safety factor:

$$\sigma_{\text{working}} = \frac{\sigma_{\text{yield}}}{n} = \frac{2.5 \times 10^8 \text{ Pa}}{4} = 6.25 \times 10^7 \text{ Pa}$$

Each cable is therefore allowed to experience a stress of at most $6.25 \times 10^7 \text{ Pa}$ in normal service. This is the stress that corresponds to the full live load divided by four and still leaves a factor of four in hand before the steel begins to yield.

(d) The minimum cross-sectional area of each cable is the force on the cable divided by the maximum allowed stress:

$$A = \frac{F}{\sigma_{\text{working}}} = \frac{1470\text{N}}{6.25 \times 10^7 \text{ Pa}} = 2.352 \times 10^{-5} \text{ m}^2$$

Each cable is circular, so its cross-sectional area is related to its diameter by $A = \pi d^2/4$. Rearranging for the diameter:

$$d = \sqrt{\frac{4A}{\pi}} = \sqrt{\frac{4 \times 2.352 \times 10^{-5} \text{ m}^2}{\pi}} = 5.47 \times 10^{-3} \text{ m}$$

Therefore each cable must have a diameter of at least $5.47 \times 10^{-3} \text{ m}$ (or about 5.5mm), to carry the maximum live load with the required safety factor of four.

Making Sense of the Answer: A diameter of 5.5mm sounds thin for a lift cable, and in fact real lift cables are almost always thicker than this; typically, 8mm to 12mm for a small passenger lift. The difference comes from three considerations that the simple calculation above did not include. First, the mass of the lift car itself is additional dead weight that has to be supported in addition to the live load of passengers. Second, dynamic effects when the lift starts, stops, or changes direction add extra forces on top of the static weight, often as much as 50 percent more. Third, modern lifts use counterweights to reduce the net force on the cables, which can partially offset the first two effects. The 5.5mm answer should therefore be understood as

the absolute minimum dictated by the safety factor for the stated live load alone, not as a practical design specification.

Think Like a Physicist: Design-feasibility problems always have the same four-step structure: (i) find the worst-case load, (ii) divide it among the load-bearing members to find the force on each, (iii) compute the working stress from the yield stress and the safety factor, and (iv) rearrange the definition of stress to find the minimum cross-section or diameter needed. Keeping the four steps separate prevents the easy mistake of dividing by the safety factor twice or forgetting to divide at all. Whenever you are asked to size a cable, a rod, or a column to carry a given load, recognise the pattern and walk through the steps in order.

The chapter has walked us from the PE curve to the cross-section of a lift cable, and along the way it has let stress and strain, the three moduli, Searle's apparatus, and rails buckling in the afternoon sun all introduce themselves one at a time. They will no longer be polite. The miscellaneous worked examples that follow mix Young's modulus with thermal stress, hysteresis with safety factors, and hanging wires with breaking points, the way real NECTA questions do.

MISCELLANEOUS WORKED EXAMPLES ON ELASTICITY

Example 32

(a) Distinguish between the proportional limit, the elastic limit, and the yield point of a metal rod under tensile loading.

(b) A steel rod of length 2.0m and diameter 2.0mm is subjected to an increasing tensile force. At a load of 500N the rod extends by 1.60mm. When the load is increased to 1200N the rod extends by 3.80mm, and beyond 1200N the rod begins to deform plastically. (i) Calculate Young's modulus of the steel from the first pair of readings. (ii) Calculate the stress in the rod at the elastic limit. (iii) Calculate the elastic potential energy stored in the rod at the elastic limit.

Solution

(a) The proportional limit is the highest stress up to which stress is directly proportional to strain, that is, up to which Hooke's law holds. The elastic limit is the highest stress at which the rod still returns exactly to its original length when the load is removed, and lies just above the proportional limit. The yield point is the stress at which the rod begins to flow plastically, giving a large increase in strain for a small increase in stress, and lies above the elastic limit.

(b) (i) Cross-sectional area of the rod:

$$A = \frac{\pi d^2}{4} = \frac{\pi(2.0 \times 10^{-3}\text{m})^2}{4} = 3.14 \times 10^{-6}\text{m}^2$$

Young's modulus from the first pair of readings:

$$E = \frac{FL}{A\Delta L} = \frac{(500\text{N})(2.0\text{m})}{(3.14 \times 10^{-6}\text{m}^2)(1.60 \times 10^{-3}\text{m})} = 1.99 \times 10^{11}\text{Pa}$$

(ii) Stress at the elastic limit (load 1200N):

$$\sigma = \frac{F}{A} = \frac{1200\text{N}}{3.14 \times 10^{-6}\text{m}^2} = 3.82 \times 10^8\text{Pa}$$

(iii) Elastic potential energy stored at the elastic limit:

$$U = \frac{1}{2}F\Delta L = \frac{1}{2}(1200\text{N})(3.80 \times 10^{-3}\text{m}) = 2.28\text{J}$$

Example 33

(a) A steel wire and a copper wire of equal length and equal cross-section are joined end to end and stretched by the same tensile force. Given that the Young's modulus of steel is greater than that of copper, explain which wire extends more, and state the physical reason.

(b) A composite wire is formed by joining a steel wire of length 2.0m to a copper wire of length 3.0m end to end. Both wires have the same cross-sectional area of $1.5 \times 10^{-6}\text{m}^2$. A tensile force of 450N is applied to the free ends. Given $E_s = 2.0 \times 10^{11}\text{Pa}$ and $E_c = 1.2 \times 10^{11}\text{Pa}$, calculate (i) the tensile stress in each wire, (ii) the extension of each section, and (iii) the total extension of the composite.

Solution

(a) The copper wire extends more. With equal cross-section and equal force, both wires carry the same stress. Strain is stress divided by Young's modulus, so the wire with the smaller Young's modulus (copper) has the larger strain. With equal lengths, the larger strain produces the larger extension.

(b) (i) The two wires are in series, so they carry the same force and have the same cross-section. The stress in both is therefore equal:

$$\sigma = \frac{F}{A} = \frac{450\text{N}}{1.5 \times 10^{-6}\text{m}^2} = 3.0 \times 10^8\text{Pa}$$

(ii) Extension of the steel section:

$$\Delta L_s = \frac{\sigma L_s}{E_s} = \frac{(3.0 \times 10^8\text{Pa})(2.0\text{m})}{2.0 \times 10^{11}\text{Pa}} = 3.0 \times 10^{-3}\text{m} = 3.0\text{mm}$$

Extension of the copper section:

$$\Delta L_c = \frac{\sigma L_c}{E_c} = \frac{(3.0 \times 10^8\text{Pa})(3.0\text{m})}{1.2 \times 10^{11}\text{Pa}} = 7.5 \times 10^{-3}\text{m} = 7.5\text{mm}$$

(iii) Total extension of the composite:

$$\Delta L = \Delta L_s + \Delta L_c = 3.0\text{mm} + 7.5\text{mm} = 10.5\text{mm}$$

Example 34

(a) Explain why a horizontally rotating wire carrying a mass at its free end can break at a rotation speed at which an equivalent static tensile load would be perfectly safe.

(b) A copper wire of length 1.0m and diameter 0.80mm has a mass of 0.50kg attached to its free end. The other end is held fixed and the wire is rotated in a horizontal circle. The breaking stress of the copper wire is $2.5 \times 10^8\text{Pa}$. Ignoring the mass of the wire itself, calculate (i) the maximum tension the wire can withstand just before breaking, (ii) the maximum angular velocity that the wire can survive, and (iii) the corresponding linear speed of the mass at this maximum.

Solution

(a) A rotating wire is subjected to repeated dynamic stressing accompanied by small bending and vibrational stresses that a pure static test does not reproduce. Elastic fatigue accumulates microscopic cracks at surface flaws and lowers the effective breaking stress well below its static value. The wire therefore fails sooner under rotation than under an equivalent static load.

(b) (i) Cross-sectional area of the wire:

$$A = \frac{\pi d^2}{4} = \frac{\pi(0.80 \times 10^{-3}\text{m})^2}{4} = 5.03 \times 10^{-7}\text{m}^2$$

Maximum tension:

$$T_{\text{max}} = \sigma_{\text{break}} \cdot A = (2.5 \times 10^8\text{Pa})(5.03 \times 10^{-7}\text{m}^2) = 125.7\text{N}$$

(ii) The tension provides the centripetal force $T = m\omega^2 r$, with $r = L = 1.0\text{m}$:

$$\omega_{\text{max}} = \sqrt{\frac{T_{\text{max}}}{mr}} = \sqrt{\frac{125.7\text{N}}{(0.50\text{kg})(1.0\text{m})}} = 15.9\text{rad/s}$$

(iii) Linear speed of the mass at this angular velocity:

$$v = \omega r = (15.9\text{rad/s})(1.0\text{m}) = 15.9\text{m/s}$$

Example 35

(a) A long rubber band and a short steel wire of the same cross-sectional area are each stretched until they store the same amount of elastic potential energy per unit volume. Which of the two is strained more, and state the physical reason.

(b) A steel wire of length 1.50m, cross-sectional area $2.0 \times 10^{-7} \text{m}^2$, and Young's modulus $2.0 \times 10^{11} \text{Pa}$ is stretched by an axial force until its strain reaches 1.0×10^{-3} . (i) Calculate the tensile stress in the wire. (ii) Calculate the extension of the wire. (iii) Calculate the elastic energy per unit volume stored in the wire.

Solution

(a) The rubber band is strained more. The elastic energy per unit volume is $u = \frac{1}{2} E \epsilon^2$. For the two materials to store the same energy per unit volume, the product $E \epsilon^2$ must be the same. Since steel is much stiffer, the rubber's Young's modulus is many orders of magnitude smaller than steel's; so the rubber's strain must be many times larger to compensate.

(b) (i) Tensile stress in the wire:

$$\sigma = E \epsilon = (2.0 \times 10^{11} \text{Pa})(1.0 \times 10^{-3}) = 2.0 \times 10^8 \text{Pa}$$

(ii) Extension of the wire:

$$\Delta L = \epsilon L = (1.0 \times 10^{-3})(1.50 \text{m}) = 1.50 \times 10^{-3} \text{m} = 1.50 \text{mm}$$

(iii) Elastic energy per unit volume:

$$u = \frac{1}{2} \sigma \epsilon = \frac{1}{2} (2.0 \times 10^8 \text{Pa})(1.0 \times 10^{-3}) = 1.0 \times 10^5 \text{J/m}^3$$

Example 36

(a) Explain why a long vertical wire can break under its own weight even when no external load is attached to its lower end.

(b) A steel wire of density 7800kg/m^3 and breaking stress $8.0 \times 10^8 \text{Pa}$ is suspended vertically from a rigid support. Take $g = 9.8 \text{m/s}^2$. (i) Derive an expression for the maximum length of such a wire that can hang vertically without breaking at its point of suspension. (ii) Calculate the numerical value of this maximum length.

Solution

(a) The point of suspension at the top of the wire has to carry the entire weight of the wire hanging below it. As the length of the wire increases, the weight below the top increases in proportion, and the tensile stress at the top increases with it. When the stress at the top reaches the breaking stress of the material, the wire snaps at that point. The rest of the wire, which carried less weight, remains intact.

(b) (i) Consider a wire of length L , cross-sectional area A , and density ρ . The total weight of the wire is:

$$W = \rho A L g$$

All weight acts on the top cross-section. The stress at the top is therefore:

$$\sigma_{\text{top}} = \frac{W}{A} = \rho L g$$

Setting σ_{top} equal to the breaking stress σ_{break} and solving for the maximum length:

$$L_{\text{max}} = \frac{\sigma_{\text{break}}}{\rho g}$$

(ii) Substituting the values:

$$L_{\text{max}} = \frac{8.0 \times 10^8 \text{Pa}}{(7800 \text{kg/m}^3)(9.8 \text{m/s}^2)} = 1.05 \times 10^4 \text{m}$$

Example 37

(a) Explain why a fluid cannot support a static shear stress in equilibrium, while a solid of the same volume can.

(b) A cube of steel of side 5.0cm is subjected to a shearing force of $2.5 \times 10^5 \text{N}$ applied tangentially along one of its faces while the opposite face is held fixed. The shear modulus of steel is $8.0 \times 10^{10} \text{Pa}$. Calculate (i) the shear stress on the face, (ii) the shear strain in the cube, and (iii) the lateral displacement of the top face relative to the bottom face.

Solution

(a) In a fluid, the layers have no binding force resisting tangential displacement relative to one another. When a shear stress is applied, the layers slide past each other continuously, and the fluid flows rather than settling into equilibrium. In a solid, the intermolecular forces exert restoring forces against tangential displacement, and the solid reaches a static equilibrium at a finite shear strain determined by the shear modulus.

(b) (i) Area of the sheared face:

$$A = (5.0 \times 10^{-2} \text{m})^2 = 2.5 \times 10^{-3} \text{m}^2$$

Shear stress:

$$\tau = \frac{F}{A} = \frac{2.5 \times 10^5 \text{N}}{2.5 \times 10^{-3} \text{m}^2} = 1.0 \times 10^8 \text{Pa}$$

(ii) Shear strain:

$$\gamma = \frac{\tau}{G} = \frac{1.0 \times 10^8 \text{Pa}}{8.0 \times 10^{10} \text{Pa}} = 1.25 \times 10^{-3}$$

(iii) Lateral displacement of the top face:

$$x = \gamma \cdot h = (1.25 \times 10^{-3})(5.0 \times 10^{-2} \text{m}) = 6.25 \times 10^{-5} \text{m} = 0.0625 \text{mm}$$

Example 38

(a) State the physical meaning of the compressibility of a fluid, and explain why gases have much higher compressibilities than liquids at ordinary pressures.

(b) A volume of 1.025m^3 of sea water is taken from the surface to a depth where the pressure is $5.0 \times 10^7 \text{Pa}$ greater than at the surface. The bulk modulus of sea water is $2.2 \times 10^9 \text{Pa}$, and the density of sea water at the surface is 1025kg/m^3 . Calculate the density of the sea water at this greater depth.

Solution

(a) Compressibility is the fractional change in volume of the fluid per unit increase in applied pressure; it is the reciprocal of the bulk modulus. Gases have much higher compressibilities than liquids because their molecules are far apart and there is large empty space between them that can be squeezed out by an applied pressure. In liquids the molecules are almost in contact and their mutual repulsion makes further compression extremely difficult, so liquids are nearly incompressible.

(b) Calculating the fractional change in volume:

$$\frac{\Delta V}{V} = -\frac{\Delta P}{B} = -\frac{5.0 \times 10^7 \text{Pa}}{2.2 \times 10^9 \text{Pa}} = -2.27 \times 10^{-2}$$

Calculating the actual change in volume:

$$\Delta V = -2.27 \times 10^{-2} \times 1.025 \text{m}^3 = -2.33 \times 10^{-2} \text{m}^3$$

The mass of the sea water does not change with compression. Mass at the surface:

$$m = \rho_0 V_0 = (1025 \text{kg/m}^3)(1.025 \text{m}^3) = 1050.6 \text{kg}$$

Calculating new volume at the depth:

$$V' = 1.025 \text{m}^3 - 2.33 \times 10^{-2} \text{m}^3 = 1.0017 \text{m}^3$$

New density at the depth:

$$\rho' = \frac{m}{V'} = \frac{1050.6 \text{kg}}{1.0017 \text{m}^3} = 1049 \text{kg/m}^3$$

Alternative solution

For constant mass, $\rho \propto \frac{1}{V}$. So fractional decrease in volume = fractional **increase** in density

$$\rho' = \rho_0 + \Delta \rho = 1025 \text{kg/m}^3 + 2.27 \times 10^{-2} \times 1025 \text{kg/m}^3 = 1048 \text{kg/m}^3$$

Example 39

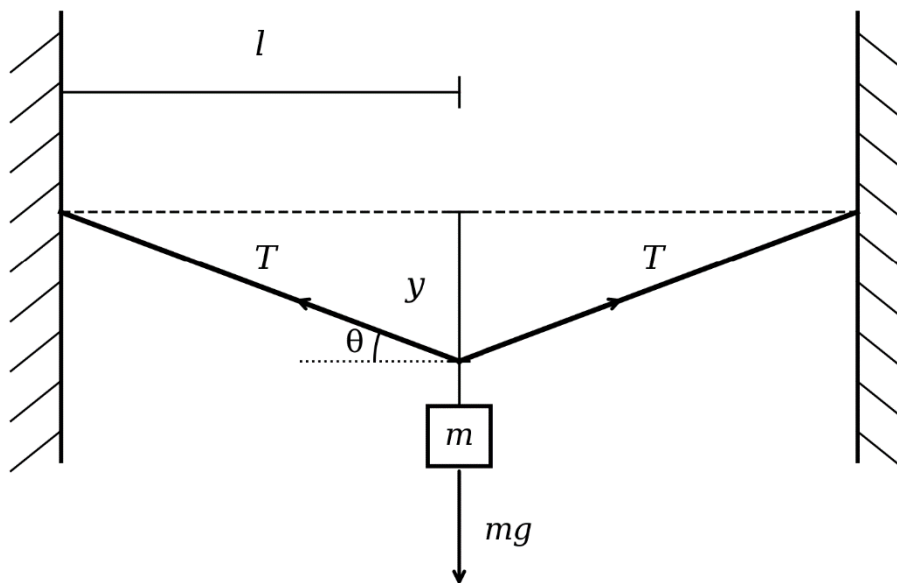
(a) Explain why a horizontal wire supporting a small load at its midpoint cannot remain perfectly horizontal, however short the wire or however small the load.

(b) A mass of 1.0kg is suspended from the midpoint of a steel wire of total length 2.0m stretched horizontally between two rigid supports. The wire has cross-sectional area $2.0 \times 10^{-6} \text{m}^2$ and Young's modulus $2.0 \times 10^{11} \text{Pa}$. Take $g = 9.8 \text{m/s}^2$. (i) Derive, using the small-angle approximation, an expression relating the midpoint depression y to the mass m , the half-length l of the wire, the cross-sectional area A , and the Young's modulus E . (ii) Calculate the numerical value of the midpoint depression.

Solution

(a) A perfectly horizontal wire can only exert tension along its own length, which is horizontal. A horizontal tension has no vertical component, and cannot support the downward weight of the suspended load. The wire must therefore tilt downward slightly at its midpoint so that its tension acquires a small vertical component, which is just large enough to balance the weight of the load. The midpoint must sag.

(b) (i) Consider the following diagram:



(b) (i) Two methods of derivation are presented below; you are invited to study both and decide which one feels more natural.

Method 1: Pythagoras and the binomial approximation

Finding the new length of each half

After the wire sags, each half runs from a fixed support down to the midpoint, forming the hypotenuse of a right-angled triangle whose horizontal side is l and whose vertical side is y . By Pythagoras' theorem, the new length of each half is:

$$L = \sqrt{l^2 + y^2}$$

Applying the small-sag approximation by factoring out l

Since y is much smaller than l , we factor l out of the square root to bring the small quantity into a recognisable form:

$$L = \sqrt{l^2 \left(1 + \frac{y^2}{l^2} \right)} = l \sqrt{1 + \frac{y^2}{l^2}}$$

For any small quantity x , the binomial approximation gives $\sqrt{1+x} \approx 1 + \frac{1}{2}x$. Applying this with $x = \frac{y^2}{l^2}$ (which is very small because $y \ll l$):

$$L \approx l \left(1 + \frac{1}{2} \times \frac{y^2}{l^2} \right) = l + \frac{y^2}{2l}$$

Finding the extension of each half

The extension of each half is its new length minus its original length l :

$$\Delta l = L - l = \left(l + \frac{y^2}{2l} \right) - l = \frac{y^2}{2l}$$

Finding the strain in each half

$$\varepsilon = \frac{\Delta l}{l} = \frac{\frac{y^2}{2l}}{l} = \frac{y^2}{2l^2}$$

Finding the tension in each half from Young's modulus

The stress in each half is $\sigma = E\varepsilon = \frac{E y^2}{2l^2}$, and the tension is the stress times the cross-sectional area:

$$T = \sigma A = \frac{E A y^2}{2l^2}$$

Applying equilibrium at the midpoint

The midpoint is in equilibrium under three forces: the two tensions pointing along the wire from the midpoint outward toward the supports, and the weight mg pointing downward. The two horizontal components of tension cancel each other by symmetry. The two vertical components must together support the weight. Each vertical component is $T \sin \theta$, where θ is the small angle each half of the wire makes with the horizontal. From the geometry, $\sin \theta \approx \tan \theta = \frac{y}{l}$ for small θ , so:

$$2T \times \frac{y}{l} = mg$$

Substituting T from $T = \sigma A = \frac{E A y^2}{2l^2}$ and solve for y

$$2 \times \frac{E A y^2}{2l^2} \times \frac{y}{l} = mg$$

$$\frac{E A y^3}{l^3} = mg$$

$$y^3 = \frac{m g l^3}{E A}$$

$$y = l^3 \sqrt{\frac{m g}{E A}}$$

Method 2: Direct small-angle method

Defining the small angle and write its trigonometric ratios

Let θ be the small angle that each half of the wire makes with the horizontal after sagging. From the geometry:

$$\tan \theta = \frac{y}{l}$$

For a small angle θ , the standard small-angle approximations are:

$$\sin \theta \approx \tan \theta \approx \theta, \quad \cos \theta \approx 1 - \frac{\theta^2}{2}$$

In particular, $\sin \theta \approx \frac{y}{l}$

Finding the new length of each half from the cosine

The horizontal half-length l is the projection of the new (stretched) half-length L onto the horizontal direction:

$$l = L \cos \theta$$

Solving for L:

$$L = \frac{l}{\cos \theta}$$

For a small angle, $\frac{1}{\cos \theta} \approx 1 + \frac{\theta^2}{2}$, so:

$$L \approx l \left(1 + \frac{\theta^2}{2} \right) = l + \frac{l\theta^2}{2}$$

Finding the extension and the strain

$$\Delta l = L - l = \frac{l\theta^2}{2}$$

$$\varepsilon = \frac{\Delta l}{l} = \frac{\theta^2}{2}$$

Substituting $\theta \approx \frac{y}{l}$:

$$\varepsilon = \frac{1}{2} \left(\frac{y}{l} \right)^2 = \frac{y^2}{2l^2}$$

Finding the tension in each half from Young's modulus

$$T = \sigma A = E \varepsilon A = \frac{EAy^2}{2l^2}$$

Applying equilibrium at the midpoint

The two vertical components of the tensions must support the weight mg :

$$2T \sin \theta = mg$$

Substituting $\sin \theta \approx y/l$ from Step 1:

$$2T \cdot \frac{y}{l} = mg$$

Substituting T and solve for y

$$2 \cdot \frac{EAy^2}{2l^2} \cdot \frac{y}{l} = mg$$

$$\frac{EAy^3}{l^3} = mg$$

$$y = l \sqrt[3]{\frac{mg}{EA}}$$

The two methods reach the same expression by different routes. Method 1 stays in pure algebra and applies the small-quantity approximation to a square root. Method 2 stays in pure trigonometry and applies the small-angle approximations. Both are equally valid; whichever feels more natural is the one to use in the examination room.

(ii) Substituting:

$$y = (1.0\text{m}) \sqrt[3]{\frac{(1.0\text{kg})(9.8\text{m/s}^2)}{(2.0 \times 10^{11}\text{Pa})(2.0 \times 10^{-6}\text{m}^2)}} = 2.90 \times 10^{-2}\text{m} \approx 2.9\text{cm}$$

The midpoint of the wire therefore sags by approximately 2.9cm under the suspended kilogram mass.

Example 40

(a) Two wires of different materials are clamped between the same pair of rigid walls so that they lie side by side, each running the full distance between the walls. Explain what happens to each wire as the temperature of the whole system is lowered below the reference temperature at which the wires were originally fitted.

(b) A steel wire and a copper wire, each of length 50.0cm at 20°C, are clamped between the same two rigid walls. The system is then cooled to 0°C. Given $E_s = 2.0 \times 10^{11}\text{Pa}$, $E_c = 1.2 \times 10^{11}\text{Pa}$, $\alpha_s = 12 \times 10^{-6}\text{K}^{-1}$, and $\alpha_c = 17 \times 10^{-6}\text{K}^{-1}$, calculate (i) the tensile thermal stress developed in each wire, and (ii) state which wire experiences the greater thermal stress and identify the physical reason.

Solution

(a) When the temperature is lowered, each wire tries to contract by an amount proportional to its length, its coefficient of linear expansion, and the temperature drop. The walls hold the ends of both wires fixed, so neither can shorten. Each wire therefore develops a tensile stress large enough to stretch the material elastically by exactly the amount it wanted to contract thermally. Because the two wires have different Young's moduli and different coefficients of expansion, they develop different tensile stresses even though they experience the same temperature drop.

(b) (i) The magnitude of the temperature change is $|\Delta T| = 20\text{K}$. Thermal stress in each wire is $\sigma = E\alpha|\Delta T|$.

For the steel wire:

$$\sigma_s = E_s \alpha_s |\Delta T| = (2.0 \times 10^{11}\text{Pa})(12 \times 10^{-6}\text{K}^{-1})(20\text{K}) = 4.8 \times 10^7\text{Pa}$$

For the copper wire:

$$\sigma_c = E_c \alpha_c |\Delta T| = (1.2 \times 10^{11}\text{Pa})(17 \times 10^{-6}\text{K}^{-1})(20\text{K}) = 4.08 \times 10^7\text{Pa}$$

(ii) The steel wire experiences the greater thermal stress ($4.8 \times 10^7\text{Pa}$ compared with $4.08 \times 10^7\text{Pa}$ for copper). Although copper has a larger coefficient of expansion, steel has a much larger Young's modulus, and the product $E\alpha$ for steel exceeds that of copper, so steel is the stiffer partner in the pair.

Example 41

(a) Explain why, in the design of a long railway line, the length of individual continuous rail sections is limited not by the breaking stress of the steel but by the maximum expected temperature change at the location.

(b) A steel rod is clamped rigidly between two walls at 15°C. The maximum compressive stress that the rod can safely withstand before it is judged to buckle is $1.5 \times 10^8\text{Pa}$. Given $E = 2.0 \times 10^{11}\text{Pa}$ and $\alpha = 12 \times 10^{-6}\text{K}^{-1}$, (i) derive an expression for the maximum temperature rise above 15°C at which the rod remains safe, (ii) calculate this maximum temperature rise numerically, and (iii) state the highest safe ambient temperature for the rod.

Solution

(a) A clamped rail develops a thermal stress $\sigma = E\alpha\Delta T$ that is independent of its length. However, a long slender rail does not fail by direct compression at the full crushing strength of steel: it fails by sideways buckling at stresses much smaller than the crushing strength. The longer the rail, the lower the buckling stress. Expansion gaps must therefore be sized so that they always close before buckling begins, and this requirement limits the length of any continuous rail section based on the maximum temperature change the rail can experience between the coldest night and the hottest afternoon.

(b) (i) Setting $\sigma = E\alpha\Delta T$ equal to the maximum safe compressive stress and solving for ΔT :

$$\Delta T_{\max} = \frac{\sigma_{\max}}{E\alpha}$$

(ii) Substituting the values:

$$\Delta T_{\max} = \frac{1.5 \times 10^8\text{Pa}}{(2.0 \times 10^{11}\text{Pa})(12 \times 10^{-6}\text{K}^{-1})} = 62.5\text{K}$$

(iii) The highest safe ambient temperature is:

$$T_{\max} = 15^\circ\text{C} + 62.5^\circ\text{C} = 77.5^\circ\text{C}$$

Example 42

(a) Starting from the definition of work done by a variable force, derive the expression $U = \frac{1}{2}Fx$ for the elastic potential energy stored in a wire that obeys Hooke's law, where F is the final applied force and x is the corresponding extension.

(b) A brass wire of length 0.50m, cross-sectional area $1.5 \times 10^{-6} \text{m}^2$, and Young's modulus $9.0 \times 10^{10} \text{Pa}$ is stretched by a force of 80N. Calculate (i) the extension of the wire, and (ii) the elastic potential energy stored in the wire.

Solution

(a) For a wire obeying Hooke's law, the force required to produce an extension x' is $F' = kx'$, where k is the spring constant of the wire. The work done in increasing the extension from x' to $x' + dx'$ against this variable force is $dW = F' dx' = kx' dx'$. Integrating from the unstretched state $x' = 0$ to the final extension $x' = x$:

$$W = \int_0^x kx' dx' = \frac{1}{2} kx^2$$

The spring constant can be eliminated using $k = \frac{F}{x}$ at the final extension, giving $W = \frac{1}{2} \left(\frac{F}{x}\right) x^2 = \frac{1}{2} Fx$.

Since all the work done against the elastic force is stored as elastic potential energy, $U = W$, therefore:

$$U = \frac{1}{2} Fx$$

(b) (i) Extension of the wire:

$$\Delta L = \frac{FL}{AE} = \frac{(80\text{N})(0.50\text{m})}{(1.5 \times 10^{-6} \text{m}^2)(9.0 \times 10^{10} \text{Pa})} = 2.96 \times 10^{-4} \text{m}$$

(ii) Elastic potential energy stored:

$$U = \frac{1}{2} F\Delta L = \frac{1}{2} (80\text{N})(2.96 \times 10^{-4} \text{m}) = 1.19 \times 10^{-2} \text{J}$$

Example 43

(a) Explain what is meant by stress concentration in an engineering component, give one example of its consequences in practice, and state the design modification commonly used to reduce it.

(b) A steel wire of uniform diameter 1.0mm is used to suspend a mass vertically. The ultimate tensile strength of the steel is $8.0 \times 10^8 \text{Pa}$ and the required safety factor is 5. Take $g = 9.8 \text{m/s}^2$. Calculate (i) the maximum working stress in the wire, (ii) the maximum mass that the wire may safely suspend, and (iii) the new maximum safe mass if a surface defect reduces the effective diameter of the wire to 0.9mm at one point, assuming the stress is concentrated at the narrowest cross-section.

Solution

(a) Stress concentration is the localised amplification of stress at a geometric discontinuity in a component, such as a sharp corner, a notch, a hole, or a sudden change in cross-section. The stress at the discontinuity can be several times larger than the nominal stress elsewhere, so the component can fail at that point while the average stress is still well below the breaking stress. The classic example is the de Havilland Comet aircraft of the 1950s, three of which broke up in flight because fatigue cracks developed at the sharp corners of their rectangular windows. The design modification is to round off every sharp corner and to smooth any sudden change in cross-section, so that stress varies gradually rather than jumping at an edge.

(b) (i) Maximum working stress:

$$\sigma_{\text{working}} = \frac{\sigma_{\text{UTS}}}{n} = \frac{8.0 \times 10^8 \text{Pa}}{5} = 1.6 \times 10^8 \text{Pa}$$

(ii) Original cross-sectional area:

$$A = \frac{\pi d^2}{4} = \frac{\pi(1.0 \times 10^{-3} \text{m})^2}{4} = 7.85 \times 10^{-7} \text{m}^2$$

Maximum force the wire may carry:

$$F_{\text{max}} = \sigma_{\text{working}} \cdot A = (1.6 \times 10^8 \text{Pa})(7.85 \times 10^{-7} \text{m}^2) = 125.7 \text{N}$$

Maximum mass:

$$m_{\text{max}} = \frac{F_{\text{max}}}{g} = \frac{125.7 \text{N}}{9.8 \text{m/s}^2} = 12.8 \text{kg}$$

(iii) Reduced area at the defect:

$$A' = \frac{\pi(0.9 \times 10^{-3}\text{m})^2}{4} = 6.36 \times 10^{-7}\text{m}^2$$

New maximum force and mass:

$$F'_{\max} = (1.6 \times 10^8\text{Pa})(6.36 \times 10^{-7}\text{m}^2) = 101.8\text{N}$$

$$m'_{\max} = \frac{101.8\text{N}}{9.8\text{m/s}^2} = 10.4\text{kg}$$

Example 44

(a) Two vertical rods of different materials support a rigid horizontal beam side by side. Explain why, when a load is suspended from the beam, the two rods carry different shares of the load even though the beam is rigid and horizontal.

(b) A copper rod of cross-sectional area $2.0 \times 10^{-6}\text{m}^2$ and a steel rod of cross-sectional area $1.0 \times 10^{-6}\text{m}^2$, each of length 1.5m, support a rigid horizontal beam in parallel. A total load of 1200N is shared between the two rods. $E_c = 1.2 \times 10^{11}\text{Pa}$, $E_s = 2.0 \times 10^{11}\text{Pa}$. Assume the beam remains horizontal. Calculate (i) the force supported by each rod, and (ii) the common extension of the two rods.

Solution

(a) Because the beam is rigid and horizontal, both rods must extend by the same amount. The force required to produce a given extension in a rod is proportional to its Young's modulus and its cross-sectional area (for the same length), so the rod with the larger product EA carries a proportionally larger share of the load. The two rods therefore carry different forces even though they experience the same extension, and the distribution of the load is determined by the relative stiffness of the rods.

(b) (i) Let F_c and F_s be the forces in the copper and steel rods. Equal extension requires:

$$\frac{F_c L}{A_c E_c} = \frac{F_s L}{A_s E_s}$$

The common length L cancels, giving:

$$\frac{F_s}{F_c} = \frac{A_s E_s}{A_c E_c} = \frac{(1.0 \times 10^{-6}\text{m}^2)(2.0 \times 10^{11}\text{Pa})}{(2.0 \times 10^{-6}\text{m}^2)(1.2 \times 10^{11}\text{Pa})} = 0.833$$

Combined with $F_c + F_s = 1200\text{N}$:

$$F_c(1 + 0.833) = 1200\text{N} \Rightarrow F_c = 654.6\text{N}, \quad F_s = 545.4\text{N}$$

(ii) Common extension, calculated from the copper rod:

$$\Delta L = \frac{F_c L}{A_c E_c} = \frac{(654.6\text{N})(1.5\text{m})}{(2.0 \times 10^{-6}\text{m}^2)(1.2 \times 10^{11}\text{Pa})} = 4.09 \times 10^{-3}\text{m} = 4.09\text{mm}$$

You may check with the steel rod, and will give the same value, confirming consistency.

Example 45

(a) Explain the physical principle underlying the shrink-fitting process, in which a metal ring is joined to a metal shaft without bolts, welds, or keys.

(b) A brass ring of internal diameter 25.00mm at 20°C is to be fitted over a steel shaft of external diameter 25.05mm at the same temperature. The coefficient of linear expansion of brass is $19 \times 10^{-6}\text{K}^{-1}$. (i) Calculate the minimum temperature to which the brass ring must be heated so that its internal diameter just equals the diameter of the shaft. (ii) Describe what happens as the ring then cools back to 20°C while in position on the shaft.

Solution

(a) A metal ring expands when heated and contracts when cooled. By heating the ring until its internal diameter is slightly larger than the external diameter of the shaft, the ring can be slipped onto the shaft without force. As the ring cools back to ambient temperature, it tries to contract back to its original smaller internal diameter, but the shaft prevents the inner surface from shrinking. The ring therefore develops a large tensile hoop stress and grips the shaft with a radial compressive force that is strong enough to transmit torques without any keys, bolts, or welds.

(b) (i) The internal diameter of the ring must increase from 25.00mm to 25.05mm, an increase of 0.05mm. Using the linear expansion formula for the diameter (the internal diameter expands as a solid piece of brass would):

$$\Delta d = d_0 \alpha \Delta T \Rightarrow \Delta T = \frac{\Delta d}{d_0 \alpha}$$

Substituting the numerical values:

$$\Delta T = \frac{0.05 \times 10^{-3} \text{m}}{(25.00 \times 10^{-3} \text{m})(19 \times 10^{-6} \text{K}^{-1})} = 105.3 \text{K}$$

The ring must therefore be heated to:

$$T = 20^\circ\text{C} + 105.3^\circ\text{C} \approx 125^\circ\text{C}$$

(ii) As the ring cools back to 20°C while seated on the shaft, it tries to return to its original smaller diameter, but the shaft prevents the inner surface of the ring from contracting below 25.05mm. The ring therefore remains stretched by 0.05mm in diameter relative to its natural size, and develops a tensile hoop stress that presses its inner surface radially inward against the shaft with a large compressive force. This force holds the two components together permanently.

Example 46

(a) A steel wire is simultaneously loaded with a tensile force and heated. Explain which contribution to the total strain obeys Hooke's law and which does not, and state whether the two strains add or subtract when the wire is heated.

(b) A steel wire of length 2.0m and cross-sectional area $1.0 \times 10^{-6} \text{m}^2$ is stretched by a load of 100N at 20°C , and the system is then heated to 50°C while the load is maintained. $E_{\text{steel}} = 2.0 \times 10^{11} \text{Pa}$, $\alpha_{\text{steel}} = 12 \times 10^{-6} \text{K}^{-1}$. Calculate (i) the elastic strain produced by the load alone at 20°C , (ii) the thermal strain produced by the temperature rise, and (iii) the total extension of the wire at 50°C .

Solution

(a) The elastic strain, which is produced by the applied tensile force, obeys Hooke's law: it is proportional to the stress through Young's modulus, $\epsilon_{\text{elastic}} = \sigma/E$. The thermal strain, produced by the temperature change, does not obey Hooke's law: it is given by $\epsilon_{\text{thermal}} = \alpha \Delta T$ and is independent of the applied stress. When a wire is both stretched and heated, both effects cause the wire to lengthen, so the two strains add together.

Understand! If the wire were cooled instead of heated, the thermal strain would be negative and would subtract from the elastic strain.

(b) (i) Tensile stress and elastic strain from the load:

$$\sigma = \frac{F}{A} = \frac{100 \text{N}}{1.0 \times 10^{-6} \text{m}^2} = 1.0 \times 10^8 \text{Pa}$$

$$\epsilon_{\text{elastic}} = \frac{\sigma}{E} = \frac{1.0 \times 10^8 \text{Pa}}{2.0 \times 10^{11} \text{Pa}} = 5.0 \times 10^{-4}$$

(ii) Thermal strain from the temperature rise of $\Delta T = 30\text{K}$:

$$\epsilon_{\text{thermal}} = \alpha \Delta T = (12 \times 10^{-6} \text{K}^{-1})(30\text{K}) = 3.6 \times 10^{-4}$$

(iii) Total strain and total extension at 50°C :

$$\epsilon_{\text{total}} = \epsilon_{\text{elastic}} + \epsilon_{\text{thermal}} = 5.0 \times 10^{-4} + 3.6 \times 10^{-4} = 8.6 \times 10^{-4}$$

$$\Delta L = \epsilon_{\text{total}} \cdot L = (8.6 \times 10^{-4})(2.0\text{m}) = 1.72 \times 10^{-3} \text{m} = 1.72 \text{mm}$$

Forty-six worked examples now lie behind you, and forty fresh questions wait in the Digging Deeper Exercise that follows: twelve to bind the basics, ten to take the ideas into the real world, and eighteen designed to make your pen earn its keep.

DIGGING DEEPER EXERCISE 1

EXERCISE 1A: BINDER QUESTIONS

Question 1

A wire that obeys Hooke's law extends by ΔL when a force F is applied. Explain in physical terms why doubling the force doubles the extension only as long as the stress remains below a certain limit, and identify what that limit is called.

Question 2

A long thin wire and a short thick wire are made from the same material and stretched by the same force. Explain which wire extends more and identify the two geometric factors that determine the answer.

Question 3

Distinguish between Young's modulus, the shear modulus, and the bulk modulus of a material. For each modulus, identify the type of deformation it describes and the geometric situation in which it is the relevant material property.

Question 4

Explain why a fluid has a finite bulk modulus but no shear modulus, and use your explanation to identify the essential physical difference between a solid and a fluid.

Question 5

A rubber band and a steel wire of the same dimensions are each stretched by the same small force. Explain which one extends more, and explain which one stores more elastic potential energy.

Question 6

Distinguish between the elastic potential energy stored in a stretched wire and the energy stored per unit volume of the wire. Explain why the energy per unit volume is the more useful quantity when comparing two materials.

Question 7

Explain what is meant by the elastic fatigue of a material, and describe how it differs from ordinary overloading. Give one example of an everyday object whose useful life is limited by elastic fatigue rather than by overloading.

Question 8

Sketch a typical stress-strain curve for a ductile metal and label on it the proportional limit, the elastic limit, the yield point, the ultimate tensile strength, and the fracture point. Explain in one sentence what is happening to the metal between each pair of adjacent labelled points.

Question 9

A bimetallic strip is formed by bonding a strip of brass to a strip of steel of equal length. Explain why the strip bends when heated, identify which metal lies on the outer side of the curve and explain why, and describe one practical device whose operation depends on this bending.

Question 10

Explain the working principle of Searle's two-wire apparatus for determining the Young's modulus of a wire, and identify two advantages of using a two-wire arrangement over a single-wire experiment.

EXERCISE 1B: REAL QUESTIONS

Question 11

Mama Amina buys a new spring balance from a hardware shop in Kariakoo and uses it daily to weigh sugar in her shop. After two years, she notices that the balance reads 50g even when nothing is hanging from it, although the spring still works smoothly and visibly returns to its same position whenever she removes a load. Explain in physical terms what has happened to the spring, and explain why simply applying a correction of -50g to every reading is not a complete solution to the problem.

Question 12

Explain why expansion joints are built into long suspension bridges in coastal cities.

Question 13

A car tyre is much softer than a metal wheel rim, and engineers deliberately design tyres that way rather than making them as stiff as the rim. Explain in physical terms what role the elastic compliance of the tyre plays in the everyday operation of a car, and explain why a car with completely rigid wheels would be intolerable to drive even on a smooth road.

Question 14

Explain why an old riveted steel bridge creaks under load while a similar welded bridge does not.

Question 15

Explain why steel reinforcing bars in a concrete beam are placed near the bottom of the beam rather than near the top.

Question 16

Explain why a motorcycle helmet must contain a layer of crushable foam in addition to its hard outer shell.

Question 17

Mr. Akilikubwa has just sent Kipanga to the staff room to fetch a textbook. Kipanga returns a few minutes later, looking pleased with himself.

Kipanga: “Sir, somebody has broken your window in the staff room! There is a long crack running from the corner all the way to the middle of the glass. It must have been one of those Form Three boys.”

Kipute: “Are you sure, Kipanga? I was in the staff room yesterday and the crack was already there. Did you look carefully at where it started?”

Kipanga: “It started from the corner. So what?”

Kipute: “And tell me, when the carpenter installed that window last term, did he put rubber spacers around the edges of the glass before fixing the frame, or did he press the glass straight against the wood?”

Kipanga: “I do not remember. Why does it matter?”

Explain why the crack started at the corner of the window, and identify the role that the missing rubber spacers played in producing it.

Question 18

Kipanga goes fishing at Kunduchi beach with a stiff nylon line. He hooks a large fish, and as he pulls the rod sharply upward to set the hook, the line snaps even though the breaking strength of the line is well above the weight of the fish. Explain why a sudden jerk on a stiff line is more likely to break it than a steady pull of the same average force.

Question 19

Mr. Akilikubwa is sitting in the staff room at break time, marking a stack of test papers, when Kipute comes in looking troubled.

Kipute: “Sir, my grandmother fell at home last week and broke her hip. The doctor at the hospital said it is very common in elderly people. But I have been reading about elasticity in the textbook, and it says that bone has a high Young’s modulus, comparable to that of concrete. If bone is so stiff and so strong, how can it break from such a small fall? And another thing Sir, why was it her hip and not her wrist or her arm? She was holding a teacup when she fell, but the teacup did not break and her arm was not injured.”

Mr. Akilikubwa puts down his red pen and looks at Kipute thoughtfully for a long moment before answering.

Explain why an elderly person’s bone can fracture under a load that a young person’s bone of the same Young’s modulus would survive.

EXERCISE 1C: HOT QUESTIONS**Question 20**

A long steel cable used in a colliery winding gear has cross-sectional area $4.0 \times 10^{-4} \text{ m}^2$ and supports a cage of mass 1500kg that descends at 3.0m/s and is brought to rest in 0.40s by the action of the cable. Taking $g = 9.8 \text{ m/s}^2$, calculate the maximum stress developed in the cable during the deceleration.

Question 21

A brass ring with internal diameter 30.00mm at 20°C is to be shrink-fitted onto a steel shaft of external diameter 30.06mm at the same temperature. Given the coefficient of linear expansion of brass as $19 \times 10^{-6} \text{ K}^{-1}$, calculate the temperature to which the brass ring must be heated so that it just slips over the shaft.

Question 22

A vertical steel wire just breaks under its own weight when it has length $1.05 \times 10^4 \text{ m}$. Taking $g = 9.8 \text{ m/s}^2$ and the density of steel as 7800 kg/m^3 , calculate the breaking stress of the steel.

Question 23

A copper wire of length 1.0m and diameter 0.80mm is rotated in a horizontal circle about one of its ends. A mass of 0.40kg is attached to the free end of the wire, and the breaking stress of the copper is $2.5 \times 10^8 \text{ Pa}$. Calculate the maximum linear speed of the mass at the moment the wire breaks.

Question 24

Two vertical wires hang side by side from the same rigid support. Wire A is steel with $E_s = 2.0 \times 10^{11} \text{ Pa}$, length 1.0m and cross-sectional area $1.0 \times 10^{-6} \text{ m}^2$. Wire B is brass with $E_b = 9.0 \times 10^{10} \text{ Pa}$, length 1.5m and cross-sectional area $2.0 \times 10^{-6} \text{ m}^2$. The same mass is hung from each wire in turn. Calculate the ratio of the elastic potential energy stored in wire A to that stored in wire B.

Question 25

A solid steel cube of side 8.0cm at 25°C is held inside a perfectly rigid jig that prevents it from expanding in any direction. The cube is heated to 95°C . Given the bulk modulus of steel as $1.6 \times 10^{11} \text{ Pa}$ and $\alpha = 12 \times 10^{-6} \text{ K}^{-1}$, calculate the total compressive force exerted by the jig on each face of the cube.

Question 26

A brass ring is heated from 20°C to a maximum safe temperature of 200°C in order to be shrink-fitted onto a steel shaft. The coefficient of linear expansion of brass is $19 \times 10^{-6} \text{ K}^{-1}$. If the internal diameter of the ring at 20°C is 25.00mm, calculate the maximum diameter of steel shaft that the heated ring can be slipped over.

Question 27

A steel rod is clamped rigidly between two walls at 20°C . The Young's modulus of steel is $2.0 \times 10^{11} \text{ Pa}$, its coefficient of linear expansion is $12 \times 10^{-6} \text{ K}^{-1}$, and its ultimate compressive strength is $5.0 \times 10^8 \text{ Pa}$. Calculate the maximum temperature to which the rod can be heated before it crushes.

Question 28

A passenger lift uses 6 identical steel cables in parallel, each of diameter 5.0mm. The yield stress of the steel is $2.5 \times 10^8 \text{ Pa}$ and the safety factor required by the building code is 5. Taking $g = 9.8 \text{ m/s}^2$, calculate the maximum live load (in kilograms) that the lift may safely carry.

Question 29

A horizontal steel wire of total length 2.0m and cross-sectional area $1.5 \times 10^{-6} \text{ m}^2$ is stretched between two rigid supports, and a mass is suspended from its midpoint. The midpoint is observed to sag by 5.0cm below its original horizontal position. Given $E = 2.0 \times 10^{11} \text{ Pa}$ and $g = 9.8 \text{ m/s}^2$, calculate the mass of the suspended load.

Question 30

A copper wire of length 80cm and diameter 0.50mm is attached to a 0.20kg mass and used as a conical pendulum, with the wire making a steady angle of 30° with the vertical. Given that $E_c = 1.2 \times 10^{11} \text{Pa}$ and $g = 9.8 \text{m/s}^2$, calculate the extension of the wire as the pendulum rotates steadily.

Question 31

A diver descends below the surface of a lake to a depth at which the density of the water has increased to 1003.0kg/m^3 from its surface value of 1000kg/m^3 . The bulk modulus of the water is $2.2 \times 10^9 \text{Pa}$, and $g = 9.8 \text{m/s}^2$. Calculate the depth of the diver below the surface.

Question 32

A 2.0kg steel ball is attached to the lower end of a vertical steel wire of length 3.0m and cross-sectional area $1.0 \times 10^{-6} \text{m}^2$, the upper end of the wire being fixed to a rigid support. The ball is held so that the wire is just taut and at its natural length, and is then released so that it falls and is caught by the wire. Treating the ball-and-wire system as conservative and taking $E = 2.0 \times 10^{11} \text{Pa}$ and $g = 9.8 \text{m/s}^2$, calculate the maximum extension of the wire during the subsequent motion.

Question 33

A rubber block of dimensions $2.0 \text{cm} \times 2.0 \text{cm} \times 5.0 \text{cm}$ is glued by one of its $2.0 \text{cm} \times 2.0 \text{cm}$ faces to a fixed wall. A tangential force of 0.30N is applied to the opposite face along the long direction of the block. The shear modulus of the rubber is $5.0 \times 10^5 \text{Pa}$. Calculate the lateral displacement of the loaded face relative to the fixed face.

Chapter 2

TEMPERATURE AND HEAT**INTRODUCTION**

It is a cold Miono morning and the school gate is closed. Kipanga arrives first, wrapped in a jumper two sizes too big, and touches the metal bars to push them open.

Kipanga: *“Aa! This gate is freezing! Somebody has left it in the fridge overnight.”*

He stumbles backward, sits down on the wooden bench by the gate, and lets out a theatrical sigh.

Kipanga: *“But this bench, this bench is warm. I should have come and sat here first. The gate is definitely colder than the bench. Anyone who disagrees is a fool.”*

Kipute, who has been standing three steps away with her hands in her pockets, produces a small thermometer from her bag and holds it first against the metal gate, then against the wooden bench. She looks at it for ten seconds in silence, then holds it up for Kipanga to read.

Kipute: *“Fourteen degrees. Both of them. Same reading to the nearest tenth.”*

Kipanga stares at the thermometer. He stares at the gate. He stares at the bench. He stares at Kipute. He stares at the thermometer again.

Kipanga: *“The thermometer is broken.”*

Kipute: *“The thermometer is not broken.”*

Kipanga: *“Then my hand is broken.”*

Kipute: *“Your hand is fine, Kipanga. It is your understanding of what ‘cold’ means that is broken.”*

At that moment Mr. Akilikubwa arrives, carrying his briefcase in one hand and his morning tea in the other. He has been listening from the corner since the word “fridge”. He sets down his briefcase, looks at the gate, looks at the bench, looks at Kipanga sitting on the bench and still squinting at the thermometer, and smiles the smile of a teacher who has just been handed a two-week chapter on a silver platter.

Mr. Akilikubwa: *“Congratulations, Kipanga. You have just discovered why thermometers were invented, why your fingers are not a reliable substitute for them, and why the next fifty pages of your physics book exist. Your hand is not broken and the thermometer is not broken. The thing flowing between your hand and the metal gate is not the same thing the thermometer is reading, and sorting those two things out is the work of this entire chapter. Come inside. The bench can wait.”*

Kipanga follows him in, still holding the thermometer, still looking at the gate as if it has personally insulted him. Somewhere in the back of his mind, a small voice is asking whether the bench has been lying to him all these years too.

Chapter 1 taught us that solids change their length when they get hot. The SGR railway between Dodoma and Dar is still expanding as you read this sentence, a thermostat is clicking somewhere in a fridge, and a brass ring somewhere is grudgingly fitting over a steel shaft it was heated to accept. In all of these, something called **temperature** is driving the show, and yet we have been using the word without ever asking what it really means. *What exactly is temperature? How is it different from heat? How do we measure it reliably when our own fingers, as Kipanga has just discovered, cannot be trusted?* And once we can measure it, what can we calculate, how much heat does it take to boil a kettle, to melt a block of ice, to cool a hospital ward, to freeze a lake from the top down?

This chapter answers those questions in order. The opening sections build the machinery of temperature. They explain what temperature is, how the Zeroth Law gives thermometers their logical foundation, what counts as a thermometric property, how different thermometers are constructed, and why the thermodynamic scale is the only one that does not depend on the liquid inside the glass. The later sections then switch on the calculator and develop specific heat capacity, latent heat, heating curves, the ice on a pond, and the refrigerator in the kitchen. By the end, you will know why metal feels colder than wood at the same temperature, although Kipanga will have to wait for Chapter 3 for the full answer, why a kettle takes longer to boil than an equivalent mass of ice takes to melt, and why your own body’s sweating is one of the most efficient cooling systems on the planet.

Welcome to Chapter 2. Your fingers have been lying to you all your life, and it is time to meet the instruments that tell the truth.

THE CONCEPT OF TEMPERATURE AND THE ZEROth LAW

Temperature is not heat

Before we can measure temperature, we have to know what it is, and what it is not. The two most common mistakes students make with thermal physics are both versions of the same confusion, treating temperature and heat as if they were the same thing. They are not, and the distinction is worth getting right before another paragraph passes.

Temperature is a scalar measure of the degree of hotness of a body. It tells us *how hot* the body is, but it says nothing by itself about *how much energy* the body contains. A teaspoon of boiling water and a swimming pool of boiling water are at exactly the same temperature (100 °C at standard pressure), but the energy contained in the pool is millions of times greater than the energy contained in the teaspoon. If you spilled the teaspoon on your hand, you would yelp. If you fell into the swimming pool, the consequences would be considerably less comic. Same temperature, vastly different amounts of stored energy.

Heat, by contrast, is a flow of energy from one body to another because of a temperature difference. Heat is never at rest. It is always going somewhere, from the hotter body to the cooler body, until the two bodies are at the same temperature and the flow stops. The gate in the morning story was not transferring its “coldness” to Kipanga’s hand; it was accepting heat **from** Kipanga’s hand (because Kipanga at 37 °C was hotter than the gate at 14 °C), and **Kipanga’s hand felt the loss as a sensation of cold**. The wooden bench was accepting heat from Kipanga’s trousers at a much slower rate, which is why the bench felt warmer. Both surfaces were at the same temperature. The thermometer, being a more honest judge of hotness than Kipanga’s skin, reported the truth.

Thermal equilibrium

Two bodies placed in contact with each other exchange heat until they reach a state in which there is no further net flow of energy between them. At that point, we say the two bodies are in **thermal equilibrium**. The thermometer pressed against the gate reached thermal equilibrium with the gate after a few seconds, then the same thermometer pressed against the bench reached thermal equilibrium with the bench a few seconds later. Both times it showed 14 °C, because both times it arrived at the same final state.

Thermal equilibrium has a simple physical meaning. The hotter body loses energy until its hotness drops, the colder body gains energy until its hotness rises, and the two meet in the middle. The process ends when neither body has any more reason to give or receive. Whatever quantity it is that determines this “meeting in the middle” is what we call temperature, and two bodies at the same temperature are defined as being in thermal equilibrium with each other.

The Zeroth Law of Thermodynamics

Now here is the beautiful statement that makes thermometers logically possible. Suppose Body A is in thermal equilibrium with Body C, and Body B is, separately, also in thermal equilibrium with Body C. A and B have never been in contact with each other, not even briefly. Can we conclude anything about A and B?

The **Zeroth Law of Thermodynamics** answers: yes, we can. *If A is in thermal equilibrium with C, and B is in thermal equilibrium with C, then A and B are also in thermal equilibrium with each other, and therefore at the same temperature.*

This sounds almost obvious. It is not. It is an experimental law of nature that happens to be true for the universe we live in, but there is no logical reason it had to be this way. The Zeroth Law is what guarantees that temperature is a well-defined, transitive property of bodies, rather than a relationship that only makes sense two bodies at a time. It is the law that justifies the existence of thermometers.

To see why, think about what a thermometer actually does. A thermometer is a small body (the “C” of our argument) that we bring into contact with the body we want to measure (the “A”) until the thermometer and A reach thermal equilibrium. We read the thermometer’s scale and record a number, say 14 °C. Later, we bring the same thermometer (having cooled down) into contact with another body B, reach thermal equilibrium, and again read 14 °C. The Zeroth Law now lets us conclude, without ever bringing A and B

into contact, that A and B are in thermal equilibrium with each other, that is, at the same temperature. Without the Zeroth Law, this reasoning would not be valid, and a thermometer would not be a meaningful instrument.

Why is it called the “zeroth” law?

Historically, the First and Second Laws of Thermodynamics were formulated and named before the Zeroth Law. When physicists later realised that the transitivity of thermal equilibrium was more fundamental than either of them, because the First and Second Laws both assume you already know what temperature is, they faced an awkward naming problem. Renumbering the First and Second Laws would have broken a century of textbook references. Ralph Fowler proposed the solution in the 1930s: give the new law the number zero, placing it logically before the First Law without renumbering anything. The name stuck. The Zeroth Law sits at the foundation of thermodynamics and is called “zeroth” for the same reason a physicist starts counting rows at zero rather than one, because sometimes the thing that comes first is best labelled as coming before the thing we used to call first.

Theory and you have been heating each other up for a while. Time for a little thermal equilibrium, in the form of two worked examples.

BINDER Example 1

Two metal blocks X and Y are brought into contact with each other and left until they reach thermal equilibrium. The two blocks are then separated, and block X is brought into contact with a third, cooler block Z, and the two are left until they reach thermal equilibrium. (a) What can you conclude about the relative temperatures of Y and Z without placing Y and Z in contact? (b) Which physical principle allows you to make this prediction? (c) Explain how the same principle guarantees that a thermometer is a logically valid instrument for comparing the temperatures of two different bodies.

Solution

(a) Block Y is at a higher temperature than block Z. Blocks X and Y reached thermal equilibrium first, so they share a common temperature, call it T_{XY} . Block X was then brought into contact with the cooler block Z and reached a new thermal equilibrium at some temperature T_{XZ} lower than T_{XY} , because X was the hotter of the two and lost heat to Z. Therefore Y, still at T_{XY} , is hotter than Z, which is now at $T_{XZ} < T_{XY}$. We reach this conclusion without Y and Z ever having touched each other.

(b) The principle is the **Zeroth Law of Thermodynamics**, which states that *if two bodies are each in thermal equilibrium with a third body, they are in thermal equilibrium with each other*. It is the Zeroth Law that allows us to use X as a common reference in the comparison between Y and Z.

(c) A thermometer is a body that is brought into thermal equilibrium first with one object and then with another. The Zeroth Law guarantees that two objects giving the same thermometer reading must be at the same temperature, and that two objects giving different readings are at different temperatures. Without this law, the thermometer’s reading for object A would tell us nothing about how object A compares with object B, and the instrument would have no measuring power. The Zeroth Law is therefore not just a footnote in the theory of heat, it is the law that makes thermometry possible at all.

Making Sense of the Answer: *The argument works because temperature is transitive: if X shares a temperature with Y, and later with Z, then the three temperatures can be compared by using X as a go-between. Notice that we never needed to quote a number for any of the three temperatures. The physical ordering was enough. That is also how any real thermometer works when we compare two rooms, two patients, or two buckets of water without ever bringing them into the same place.*

Think Like a Physicist: *Whenever you are asked about temperatures of bodies that have been in contact in different sequences, draw a chain: A equilibrates with B, B equilibrates with C, and so on. The Zeroth Law lets you follow the chain with confidence that every link preserves equality of temperature. The only way the chain breaks is if a body is heated or cooled between contacts, which will always be stated in the question.*

REAL Example 2

Mr. Akilikubwa is on break duty in the staff room when Kipanga appears at the door, looking frustrated.

Kipanga: *“Sir, I have just been to the clinic because I thought I had a fever, and the nurse put a thermometer in my mouth. I thought she was going to leave it there forever. Why did she not read it immediately? I know my own body, and if I am feeling hot, I am hot. A thermometer is just a piece of glass. It measures its own temperature, not mine. The moment she put it in my mouth, it started cooling down my tongue, so it was obviously going to read lower than I really was.”*

Mr. Akilikubwa invites Kipanga to sit down and explain to Kipanga: (a) why the nurse waited before taking the reading, (b) why the thermometer's reading eventually gave the correct body temperature despite being, as Kipanga correctly noted, "just a piece of glass", and (c) whether Kipanga's worry that the thermometer cooled his tongue is actually a problem for the measurement.

What was Mr. Akilikubwa's response?

Solution

(a) The nurse waited because the thermometer and Kipanga's mouth were not initially in thermal equilibrium. When the thermometer was first placed in Kipanga's mouth, it was at room temperature (about 25 °C) and his mouth was at body temperature (about 37 °C). Heat began to flow from the mouth to the thermometer, and the thermometer's mercury column rose as its temperature climbed. Until thermal equilibrium had been reached, the thermometer was reading its own current temperature, which was somewhere between 25 °C and 37 °C. Reading it immediately would have given a value below the true body temperature.

(b) Kipanga is correct that the thermometer measures its own temperature. But the Zeroth Law guarantees that once the thermometer has reached thermal equilibrium with his mouth, the thermometer's temperature equals the mouth's temperature. At that moment, "the thermometer's temperature" and "Kipanga's body temperature" are numerically the same value, and reading the thermometer's scale reveals the body temperature.

(c) Kipanga's concern is real but small. The thermometer does indeed cool his tongue slightly, by drawing a little heat from his mouth on its way to equilibrium. But the amount of heat required to warm the thermometer's small mercury column from 25 °C to 37 °C is tiny compared with the heat content of his whole mouth and the blood flowing through it, so the mouth's temperature drops by a negligible fraction of a degree. The thermometer essentially reaches the mouth's original temperature, not the other way round. This is one of the reasons clinical thermometers are made with very small mercury bulbs: the smaller the thermometer's thermal capacity, the less it disturbs the body it is measuring.

We will meet thermal capacity formally later in this chapter; for now, it is enough to know that the nurse's instrument is designed to read the patient, not to change the patient.

Making Sense of the Answer: *Kipanga's three objections each contained a grain of truth, and Mr. Akilikubwa's answers did not dismiss any of them. The thermometer really does measure its own temperature, it really does cool the tongue slightly, and the nurse really did have to wait. What Kipanga missed was that thermal equilibrium resolves all three worries at once. Once the thermometer and the mouth are at the same temperature, there is nothing left to cool, nothing left to wait for, and the two temperatures are numerically equal. The Zeroth Law quietly guarantees that the single number on the thermometer is also the number for Kipanga's body.*

Think Like a Physicist: *Whenever a measurement involves putting a small instrument in contact with a larger system, ask two questions. First, how long until thermal equilibrium? That tells you how long to wait before reading. Second, how much does the instrument disturb the system it is measuring? That tells you how large your instrument is allowed to be. Clinical thermometers have tiny bulbs for exactly this reason. Laboratory calorimeters have water jackets for the same reason. Every good measuring instrument is a compromise between reading quickly and reading honestly.*

With the concept of temperature secured and the Zeroth Law standing behind every thermometer ever made, we are ready to ask a more practical question. *Any physical property that changes with temperature can in principle be used to build a thermometer, but what makes one property good and another useless?* The next section separates the contenders from the pretenders.

THERMOMETRIC PROPERTIES AND EMPIRICAL TEMPERATURE SCALES

The Zeroth Law tells us that thermometers are logically possible. It does not yet tell us how to build one. A working thermometer needs some physical quantity inside it that changes in a trustworthy way as temperature changes, and it needs a scale printed alongside that quantity so that a reader can turn the change into a number. This section is about what a trustworthy physical quantity looks like, and about how a simple scale is built from it by choosing two reference points and drawing a straight line between them.

Thermometric properties

A **thermometric property** is any physical property of a substance that varies with temperature in a way that we can read off and turn into a temperature measurement. In principle, almost any physical quantity that depends on temperature could be used. This include:

- the length of a metal rod,
- the volume of a liquid,
- the pressure of a gas held at constant volume,
- the electrical resistance of a wire,
- the electromotive force generated at the junction of two dissimilar metals,
- the colour of a heated piece of iron.

The whole of thermometry is an exercise in choosing among these candidates the ones that are reliable and practical, and discarding the ones that look attractive until you try to use them.

Three characteristics separate a good thermometric property from a bad one.

1) The property must vary continuously with temperature

A property that changes smoothly as the temperature rises gives a single, unambiguous reading at every temperature. A property that jumps suddenly at a particular point, or that has flat regions where it refuses to change, is useless as a thermometer because it cannot distinguish between the temperatures in those flat regions. The length of a mercury column meets this test; the colour of a metal rod (which is black at room temperature and stays black until several hundred degrees) **fails** it in the lower range.

2) The property must be measurable with sufficient precision

A thermometer is only as accurate as the reading we can take off its scale. If a property changes by a huge amount for a given temperature change, the thermometer is sensitive and a small temperature difference can be resolved. If the property changes by only a tiny amount, the thermometer is insensitive and small temperature differences go undetected. The pressure of a gas held at constant volume changes substantially with temperature and is easy to read off a manometer. The speed of sound in a solid also changes with temperature, but only by a fraction of a percent per degree; using it as a thermometric property would demand extraordinary equipment to detect each degree.

3) The property must be reproducible

The same substance, brought to the same temperature under the same conditions, must give the same value of the property every time. A substance that behaves differently on Monday from how it behaves on Tuesday, or that gives different readings in different laboratories, cannot anchor a scientific instrument. Pure mercury in a sealed capillary is fully reproducible. The colour of a roast chicken at a given temperature is not, for a great many reasons.

A property that passes all three tests can be used to build a thermometer. A property that fails **any** one of them cannot. Most of the candidates listed earlier (mercury volume, gas pressure, platinum resistance, thermocouple electromotive force) pass all three and are exploited in real instruments. Some others (length of a metal rod, specific heat, thermal conductivity) pass in principle but are too insensitive or too inconvenient to build a practical thermometer around. And a few (colour of a heated rod, rate of sound travel) fail outright in one way or another.

Building a temperature scale

Even a trustworthy thermometric property is useless on its own. The property gives us a number, such as “the mercury column is 43 mm long”, or “the gas pressure is 1.27 times its ice-point value”, and by itself this number is not a temperature. To convert the property into a temperature we need a **scale**: a rule that assigns a temperature value to every possible reading of the property.

The simplest rule is a linear one. Pick two fixed points at which the temperature is known, measure the thermometric property at each of these points, and assume the property varies linearly with temperature between them. Any reading in between can then be converted into a temperature by the straight-line formula.

The Celsius scale is built in exactly this way. Its two fixed points are the **ice point**, defined as *the temperature of pure melting ice at standard atmospheric pressure and assigned the value 0°C*, and the **steam point**, defined as *the temperature of pure water boiling at standard atmospheric pressure and assigned the value 100°C*.

Let X_0 be the value of the thermometric property at the ice point, X_{100} its value at the steam point, and X_θ its value at an unknown Celsius temperature θ .

The assumption of linearity between the two fixed points means that a graph of the property X against the temperature θ is a straight line passing through the two calibration points $(0, X_0)$ and $(100, X_{100})$. On a straight line, any two segments between the endpoints have the same slope. Taking the segment from the ice point up to the unknown temperature and comparing it with the segment from the ice point all the way up to the steam point:

$$\frac{X_\theta - X_0}{\theta - 0} = \frac{X_{100} - X_0}{100 - 0}$$

Both sides of this equation are just the slope of the line X versus θ , computed two different ways, and linearity guarantees they are equal. Simplifying the denominators:

$$\frac{X_\theta - X_0}{\theta} = \frac{X_{100} - X_0}{100}$$

Cross-multiplying to clear the fractions:

$$100 (X_\theta - X_0) = \theta (X_{100} - X_0)$$

Dividing both sides by $(X_{100} - X_0)$ to isolate θ :

$$\theta = \frac{100 (X_\theta - X_0)}{X_{100} - X_0}$$

Writing this in the more compact form:

$$\theta = \left(\frac{X_\theta - X_0}{X_{100} - X_0} \right) \times 100^\circ\text{C}$$

This is the **interpolation formula** for the Celsius scale, and it lies behind every empirical thermometer calibrated on ice and steam. Substituting $X = \ell$ (length of the mercury column) gives the Celsius temperature from a mercury-in-glass thermometer; substituting $X = P$ (gas pressure at constant volume) gives it from a gas thermometer; substituting $X = R$ (electrical resistance) gives it from a platinum resistance thermometer. The formula does not care what the property is. It only assumes that the property varies linearly with temperature between the two fixed points.

Why different thermometers disagree between fixed points

Here is the catch. The interpolation formula **assumes** linearity. It does not guarantee it. Real thermometric properties vary **approximately** linearly over short temperature ranges, but not exactly, and the departure from linearity is different for different substances. The mercury column does not expand with perfect proportionality to temperature; neither does the gas pressure, neither does the platinum resistance. Each of them follows its own slightly curved graph of property against temperature, and each of them is forced by calibration to agree at the two fixed points.

Between those two fixed points, however, the three curves pass through different regions, and the three thermometers give three different readings for the same physical temperature. A mercury thermometer and a platinum resistance thermometer, both calibrated at 0°C and 100°C , might both read 49.3°C and 50.2°C for a water bath that a gas thermometer reads as 50.0°C . None of the three is broken. Each is simply reporting what its own thermometric property tells it, on the assumption that the property varies linearly, when in fact the property's graph has a slight curvature that the formula cannot see. The thermometers agree at 0°C and 100°C by design. Between them, they disagree because the physics of their properties disagrees.

This is not a failure of the interpolation method. It is a feature of empirical temperature scales. An **empirical scale** is one that depends on the behaviour of a particular thermometric substance, and different substances give different scales. Later in this chapter we will meet a way around this problem: the thermodynamic scale, which is independent of the substance used in the thermometer and to which all real thermometers can be compared. For the moment, the lesson is this. Two thermometers calibrated at the same fixed points will agree at those fixed points exactly and at every other temperature only approximately, and the size of the disagreement reveals the **non-linearity** of one or both of their thermometric properties.

Faulty thermometers

There is one more quick application of the interpolation formula before we leave this section. Suppose a mercury-in-glass thermometer has been manufactured carelessly, and its “0°C” mark has been engraved where the mercury actually sits at 4°C, and its “100°C” mark has been engraved where the mercury actually sits at 98°C. The mercury column itself is blameless. It still expands by the same amount per degree as it always did, because its physics has not changed. What has changed is only the labels painted on the outside of the glass, and the labels are wrong by a fixed amount at each end. *If the faulty thermometer reads θ_f , what is the true Celsius temperature?*

The key insight is that the faulty reading θ_f is itself a valid thermometric property of the liquid being measured. It varies continuously with temperature, it can be read off the scale with precision, and it is reproducible from one measurement to the next (it is, after all, simply a misprinted mercury column). So we can treat the faulty reading exactly as we would treat any other thermometric property X , and apply the interpolation formula with the faulty readings as the X -values.

To use the formula, we need three numbers: the faulty thermometer’s reading at the true ice point, at the true steam point, and at the unknown temperature. Let θ_0^f be the faulty reading when the thermometer is immersed in pure melting ice (at a true temperature of 0°C), and θ_{100}^f be the faulty reading when the thermometer is immersed in boiling water (at a true temperature of 100°C). By construction, θ_0^f and θ_{100}^f are simply the true temperatures at which the faulty “0°C” and “100°C” marks have been engraved. In our example above, $\theta_0^f = 4^\circ\text{C}$ and $\theta_{100}^f = 98^\circ\text{C}$.

Substituting $X_\theta \rightarrow \theta_f$, $X_0 \rightarrow \theta_0^f$, and $X_{100} \rightarrow \theta_{100}^f$ into the interpolation formula gives the true temperature θ from the faulty reading:

$$\theta = \left(\frac{\theta_f - \theta_0^f}{\theta_{100}^f - \theta_0^f} \right) \times 100^\circ\text{C}$$

The formula is identical in structure to the calibration formula; only the labels on the quantities have changed. A faulty thermometer is just an ordinary interpolation problem in disguise.

Theory and you have been heating each other up again. Let us reach thermal equilibrium with three worked examples before moving on to the zoo of real thermometers in the next section.

BINDER Example 3

A student proposes using the colour intensity of a glowing metal rod as a thermometric property, arguing that the rod goes from dull red to bright red to orange to white as the temperature rises. Evaluate this proposal by checking it against the three characteristics of a good thermometric property. Compare your conclusions with the corresponding properties of the electrical resistance of a platinum wire.

Solution

To evaluate this, we must answer three crucial questions:

1) Does the colour intensity vary continuously with temperature across the useful range?

The colour intensity of a heated rod varies continuously with temperature only above a threshold of roughly 500°C, below which the rod emits no visible light at all and appears simply as dark metal. Below this threshold, the property is not just insensitive, it is absent: the rod is black at 100°C, black at 200°C, and black at 400°C. A thermometer based on colour intensity would be blind across this entire range. The electrical resistance of a platinum wire, by contrast, varies smoothly from below room temperature upward with no threshold, and a single wire can be used from below the ice point to several hundred degrees above the steam point.

2) Can it be measured with sufficient precision at every temperature in the range?

Colour intensity is hard to measure precisely because human eyes disagree with each other and because the perceived colour depends on ambient lighting, on the angle from which the rod is viewed, and on the chemical composition of the rod’s surface. Instruments called optical pyrometers do exist for measuring colour-temperature in industrial settings, but their precision even at best is several degrees, and at low temperatures they cannot operate at all. A platinum resistance measurement, by contrast, can resolve a hundredth of a degree with a Wheatstone bridge and is limited only by the precision of the electrical instruments, not by the human observer.

3) Is it reproducible across different samples and different laboratories?

Reproducibility across samples is poor for colour intensity because different metals emit different spectra at the same temperature, and the same metal with a tarnished or oxidised surface emits differently from a freshly polished sample. The same piece of iron heated to 800°C in Dar es Salaam and in Dodoma would be reported with slightly different colour readings by identical optical pyrometers if the surfaces had weathered differently. Platinum resistance is fully reproducible because high-purity platinum has a resistance-temperature relationship that is stable across samples to better than a part in ten thousand and is not affected by surface tarnishing.

On all three counts, colour intensity fails and platinum resistance succeeds. Colour intensity can be used as a rough qualitative indicator in a blacksmith's forge, but not as a scientific thermometer. Platinum resistance is the standard by which all other precision thermometers are compared.

Making Sense of the Answer: *The three characteristics we checked are not a bureaucratic list. Each of them rules out a different kind of failure. Continuity rules out thermometers that can only read in part of the range, precision rules out thermometers that cannot resolve small differences even where they do read, and reproducibility rules out thermometers that give different answers in different laboratories. A property can pass two tests and fail the third, in which case it is still unsuitable. Colour intensity fails all three at once, which is why blacksmiths use it for judgement and physicists use platinum.*

Think Like a Physicist: *When somebody proposes a new thermometric property, ask the three questions in order. Does the property change across the range you care about? Can you measure the change precisely? Does the same substance give the same reading every time? If the answer to any one of the three is no, walk away, however clever the idea might otherwise be. The graveyard of physics is full of beautiful thermometers that fail one of these tests in an inconvenient place.*

BINDER Example 4

A mercury-in-glass thermometer has been manufactured carelessly. Its “0°C” mark has been engraved where the mercury actually sits at 3°C, and its “100°C” mark has been engraved where the mercury actually sits at 97°C. A student uses this thermometer to measure the temperature of a warm liquid, and the thermometer reads 59°C. What is the true Celsius temperature of the liquid?

Solution

The mercury column itself expands linearly between the two physical endpoints, which are at true temperatures of 3°C and 97°C. The engravings on the glass divide this physical interval into 100 equal parts, but the parts span only $97 - 3 = 94$ true degrees rather than the nominal 100 degrees. The thermometer therefore under-reads at low temperatures and over-reads at high temperatures, and a single reading in the middle of the scale must be corrected.

Applying the interpolation formula with the faulty readings as the X-values:

$$\theta = \left(\frac{\theta_f - \theta_0^f}{\theta_{100}^f - \theta_0^f} \right) \times 100^\circ\text{C}$$

With $\theta_f = 59^\circ\text{C}$, $\theta_0^f = 3^\circ\text{C}$, and $\theta_{100}^f = 97^\circ\text{C}$:

$$\theta = \left(\frac{59 - 3}{97 - 3} \right) \times 100^\circ\text{C} = 59.57^\circ\text{C}$$

The true temperature of the liquid is approximately 59.6°C, slightly higher than the faulty thermometer suggested.

Making Sense of the Answer: *The correction of 0.6°C looks small, and in a kitchen it would not matter. In a chemistry laboratory where a reaction is controlled to within one degree, a faulty thermometer that under-reads by half a degree at 60°C is an expensive piece of glass. The lesson is that calibration errors are not romantic; they accumulate quietly into bad science. A good physics student, meeting a new thermometer, always measures the ice and steam points before trusting anything else it says.*

Think Like a Physicist: *Whenever a thermometer is suspected of being faulty, ask yourself what its readings at the true ice and true steam points really are. Those two readings, substituted as θ_0^f and θ_{100}^f in the interpolation formula, convert any other reading into its true value. The same formula works whether the mistake is in the engraving, in the mercury mass, or in the glass capillary bore. Empirical interpolation is remarkably forgiving of manufacturing errors, provided you know what the errors are.*

REAL Example 5

A mercury-in-glass thermometer and a platinum resistance thermometer are both carefully calibrated at the ice point (0°C) and the steam point (100°C). A student immerses both thermometers in a water bath that a precision gas thermometer reports to be at exactly 50.0°C , waits for thermal equilibrium, and reads the two thermometers. The mercury thermometer reads 49.3°C and the platinum resistance thermometer reads 50.2°C . The student exclaims that one of the thermometers must be broken, since they were calibrated identically. Evaluate this claim and explain what the disagreement actually reveals about the two thermometric properties.

Solution

Neither thermometer is broken. The student has overlooked the single assumption on which the interpolation formula rests: the thermometric property varies **linearly** with temperature. In reality, the variation is only approximately linear, and the departure from linearity is different for different properties. The three readings in the question are enough to reveal what is happening.

Take the gas thermometer's reading of 50.0°C as the true temperature, since an ideal gas at constant volume is known to vary very nearly linearly with temperature. The mercury thermometer, applying the interpolation formula to its own measured column length, returns 49.3°C , which is 0.7°C below the true value. The formula does its arithmetic faithfully, so the under-reading can only be produced by the mercury column at 50°C being slightly shorter than a linear expansion between the two fixed points would predict. This means the mercury's volume expansion falls **below** a straight-line relationship in the middle of the interval. The platinum thermometer, by the same argument, returns 50.2°C , which is 0.2°C above the true value. Its resistance at 50°C is therefore slightly greater than a linear rise between the fixed points would predict, which means platinum resistance rises **above** a straight-line relationship in the middle of the interval.

The disagreement between the two thermometers therefore reveals that the two thermometric properties are **non-linear in opposite directions**: mercury volume expansion is slightly **sub-linear** between the fixed points, while platinum electrical resistance is slightly **super-linear**. The mercury's non-linearity is also larger in magnitude (0.7°C versus 0.2°C), which is consistent with the fact that platinum resistance is widely used as a sub-standard thermometer while mercury is used only for routine work.

Making Sense of the Answer: *Two thermometers calibrated at the same fixed points must agree at those fixed points by design, but they are under no obligation to agree at any other temperature. The student's reaction is a common one: if two instruments disagree, one of them must be lying. A more experienced physicist would reach a different conclusion: if two instruments disagree and neither is broken, the disagreement is telling you something about the instruments themselves, not about the thing they are both measuring.*

Think Like a Physicist: *Whenever you calibrate a thermometer on two fixed points, you are promising the reader only that the thermometer will agree with itself at those two points. Between them, the reading depends on the shape of the property-versus-temperature curve, which is different for every thermometric property. Two thermometers based on different properties will disagree slightly between fixed points in a way that no amount of careful calibration can fix. The only way to eliminate the disagreement is to escape empirical scales altogether, which is what the thermodynamic scale was invented to do.*

We now have the two ingredients that every working thermometer needs: a reliable physical property to respond to temperature, and a simple scale to turn the property's reading into a number. But we have been speaking of thermometers in the abstract, as though any thermometer will do. In real laboratories and real hospitals, different thermometers are chosen for different jobs, and the differences between them are not cosmetic. The mercury thermometer, the gas thermometer, the platinum resistance thermometer, and the thermocouple have very different strengths and very different weaknesses, and knowing which is which can be the difference between a measurement that matters and a number that misleads. It is time to meet them individually.

TYPES OF THERMOMETERS

The previous section gave us the formula that every empirical thermometer uses to turn a physical property into a temperature. What it did not give us is the thermometers themselves. Different jobs demand different thermometers, and the choice between a mercury bulb, a platinum coil, a thermocouple, and an optical pyrometer is not a matter of taste. Each instrument has a range in which it works well, a range in which it works poorly, and a range in which it does not work at all. The purpose of this section is to meet each of them in turn, see how it is built, see what it is good for, and see where it fails.

1. Liquid-in-glass thermometers

The liquid-in-glass thermometer is the one most students have actually held in their hand. It consists of a thin-walled glass bulb at the bottom, a fine capillary tube rising from the bulb, and a graduated scale etched or printed alongside the capillary. A small quantity of liquid fills the bulb and rises partway up the capillary. When the bulb is placed in contact with a warmer body, the liquid expands and the column rises up the capillary. When placed in a cooler body, the liquid contracts and the column falls. **The length of the column is the thermometric property.**

Mercury is the most common working liquid for laboratory and clinical use. Pure mercury expands uniformly over a very wide range of temperatures, from its freezing point at -39°C to its boiling point at 357°C , and across almost all of this range *it expands so nearly linearly* with temperature that the small non-linearities we met in the previous section contribute less than a degree of error over the whole span from ice point to steam point. Mercury is *opaque and silvery*, so the top of the column is *easy to see against the graduated scale*. It *does not wet the glass*, which means the *column rises and falls cleanly without leaving droplets behind* that would distort the reading.

Mercury's strengths come with three real limitations. **First**, mercury freezes at -39°C , so a mercury thermometer cannot be used at all in polar or cryogenic conditions. **Second**, mercury is extremely toxic if spilled or inhaled as vapour, which makes a broken mercury thermometer in a classroom a small disaster that requires specialist cleanup. **Third**, mercury is expensive, and a laboratory-grade mercury thermometer is one of the more costly pieces of glassware in a school stockroom.

Alcohol (usually ethanol coloured with a red dye to make the meniscus visible) is used where mercury cannot go. Alcohol freezes at -115°C , so *it remains fluid deep below the mercury cut-off*, and it is therefore the liquid of choice for thermometers used to measure atmospheric temperatures in cold regions or the temperatures inside domestic freezers. Its weaknesses are the mirror image of mercury's strengths. Alcohol *expands less uniformly with temperature than mercury*, so alcohol thermometers are slightly less accurate across any given range. Alcohol also boils at 78°C , which puts a *strict upper limit* on the temperatures it can measure. And because ethanol wets glass, the column tends to leave a film behind when it falls, which introduces a small *readout error* that must be corrected for.

The clinical thermometer is a specialised mercury-in-glass instrument designed for measuring human body temperature. It has two features that distinguish it from a laboratory thermometer. **First**, its scale is restricted to a narrow window around body temperature, typically 35°C to 42°C , which allows each degree to be spread across several millimetres of capillary for easy reading. **Second**, and more importantly, the capillary has a **constriction** just above the bulb, a short section narrower than the rest of the tube. When the bulb is warmed during use, the mercury expands and is forced through this constriction into the main capillary. When the thermometer is withdrawn and begins to cool, the mercury in the bulb contracts and tries to pull the column back down, but surface tension at the constriction breaks the column at that point. The mercury above the constriction cannot return to the bulb, and stays in place until the nurse has had time to read the scale. To reset the thermometer, the nurse gives it a sharp downward flick, which uses centrifugal force to drive the upper mercury back through the constriction into the bulb.

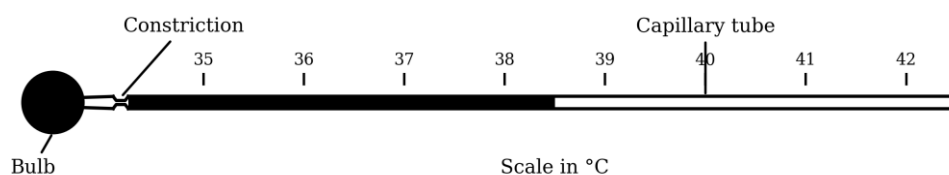


Figure: *Clinical thermometer* showing the bulb, the constriction just above the bulb, and the narrow capillary tube carrying a scale calibrated for body temperature.

2. Constant-volume gas thermometer

The constant-volume gas thermometer uses the **pressure of a fixed mass of gas held at constant volume as its thermometric property**. The gas is sealed inside a rigid glass bulb connected by a thin tube to a mercury manometer, and the temperature of the gas is changed by immersing the bulb in the body whose temperature is being measured. As the gas heats up, its pressure rises; the difference in mercury levels in the two arms of the manometer gives the pressure, and the pressure is substituted into the interpolation formula to give the temperature on the gas scale:

$$\theta_{\text{gas}} = \left(\frac{P_{\theta} - P_0}{P_{100} - P_0} \right) \times 100^{\circ}\text{C}$$

A sibling variant (**constant-pressure gas thermometer**) exists in which the pressure is held constant and the volume is allowed to change; the corresponding formula is $\theta_{\text{gas}} = \frac{V_{\theta} - V_0}{V_{100} - V_0} \times 100^{\circ}\text{C}$, but the constant-volume form is the one almost always used in practice.

The gas thermometer is often called the **ideal empirical thermometer**, because *the pressure of a dilute gas varies very nearly linearly with temperature over an enormous range*. The departure from linearity that we saw affecting the mercury and platinum thermometers is much smaller for a gas, and it becomes smaller still as the gas is made more and more dilute. In the limit where the amount of gas in the bulb tends to zero, the gas thermometer agrees exactly with the thermodynamic scale that we will meet in the next section. This is why the gas thermometer is used as the primary standard against which all other thermometers are calibrated.

The strength of the gas thermometer is its *high accuracy*. A carefully constructed laboratory gas thermometer can resolve temperatures to a thousandth of a degree. Its weaknesses are practical rather than physical. The apparatus is *large, fragile, and slow to respond*, because a large mass of gas takes time to reach thermal equilibrium with the body being measured. It is also unsuited to measuring rapidly changing temperatures, and it cannot be taken into the field. For these reasons the gas thermometer lives in the national metrology laboratory, not in the clinic or the classroom.

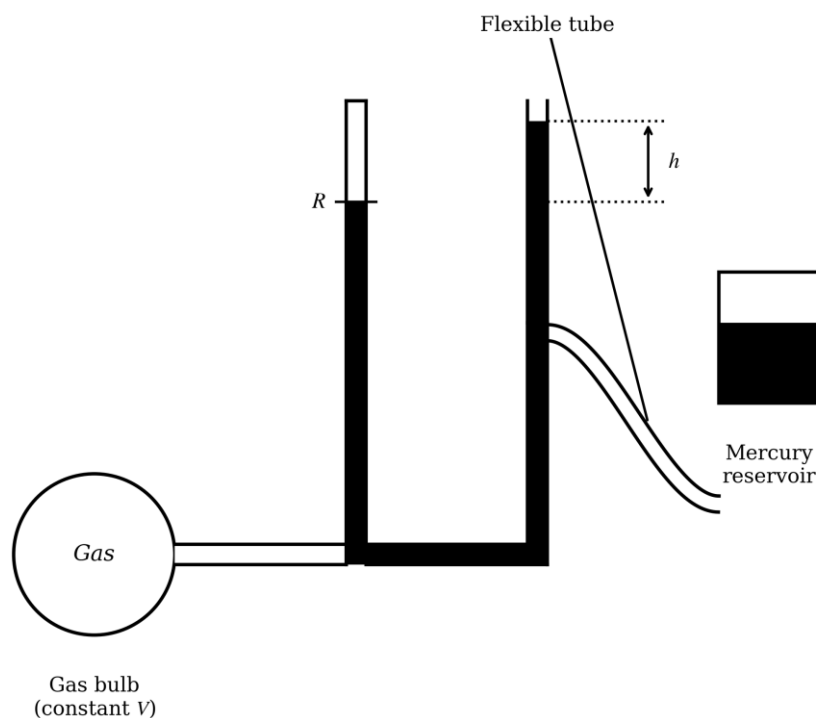


Figure: Constant-volume gas thermometer. The gas is held in a rigid bulb whose temperature is to be measured. A flexible tube connects the right arm of the mercury manometer to a reservoir whose height is adjusted so that the mercury in the left arm always reaches the reference mark *R*, keeping the gas volume constant. The height difference *h* between the two arms is then a direct measure of the gas pressure.

3. Platinum resistance thermometer

The platinum resistance thermometer uses the **electrical resistance of a pure platinum wire as its thermometric property**. Platinum is chosen because it has a *high melting point* (1773°C), is *chemically inert* across an enormous range of temperatures, *resists corrosion*, and has a *resistance that changes by a substantial and reproducible fraction* as the temperature changes. The wire is wound as a coil on a mica former to give a compact, mechanically stable element, and the coil is protected inside a quartz or ceramic sheath. The resistance is measured by connecting the coil as one arm of a Wheatstone bridge, with the other three arms provided by known resistors. At each temperature, the bridge is balanced and the coil's resistance is read off directly.

The resistance at an unknown temperature θ is substituted into the interpolation formula in the same way as the mercury column length or the gas pressure:

$$\theta_R = \left(\frac{R_\theta - R_0}{R_{100} - R_0} \right) \times 100^{\circ}\text{C}$$

For calibration work that demands a functional form, the resistance of platinum as a function of true temperature is frequently written as $R_\theta = R_0(1 + a\theta + b\theta^2)$, a parabolic form in which a and b are small constants determined by the particular wire. An alternative form is $R_\theta = R_0(1 + 800b\theta - b\theta^2)$, which is algebraically equivalent to the first after a redefinition of a . The platinum resistance thermometer is also the thermometric instrument of choice for many sub-standard calibrations, and a variant called the **thermistor**, made from semiconducting material whose resistance **decreases** with temperature, is used for applications where small size and fast response matter more than absolute accuracy.

The platinum resistance thermometer's strengths are its *high accuracy* (second only to the gas thermometer), its *wide working range* (from below -200°C to roughly 1000°C), and its *stability* across years of use without recalibration. Its weaknesses are its *slow response time* (the coil has a larger thermal mass than a thermocouple junction), its *need for a stable current source and a bridge circuit*, and the fact that *self-heating* by the measuring current itself can introduce a small error at high precision.

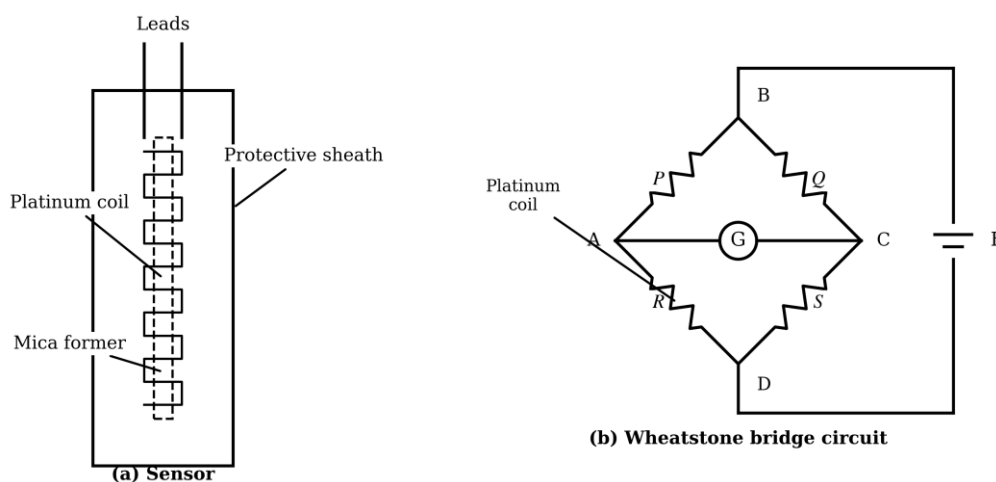


Figure: Platinum resistance thermometer. The sensor shown in (a) consists of a platinum coil wound on a mica former and enclosed in a protective sheath. The measurement circuit shown in (b) connects the coil as one arm of a Wheatstone bridge, with known resistors P , Q , S forming the other three arms, a galvanometer G across the diagonal, and a battery E providing the current.

4. Thermocouple thermometer

The thermocouple exploits a phenomenon discovered by Thomas Seebeck in 1821. When two wires of different metals are joined at both ends to form a closed loop and the two junctions are held at different temperatures, an electromotive force (EMF) appears around the loop, and a current flows if the loop is completed. This **Seebeck EMF** depends only on the pair of metals used and on the temperatures of the two junctions, not on the shape of the wires or the lengths between the junctions. A thermocouple thermometer consists of two such wires welded together at one end (the **hot junction**, placed in the body being measured) and connected to a high-resistance voltmeter at the other end (the voltmeter connections serve as the **cold junction**, held at a known reference temperature, usually 0°C in ice water). The voltmeter reads the **Seebeck EMF**, which is the **thermometric property**.

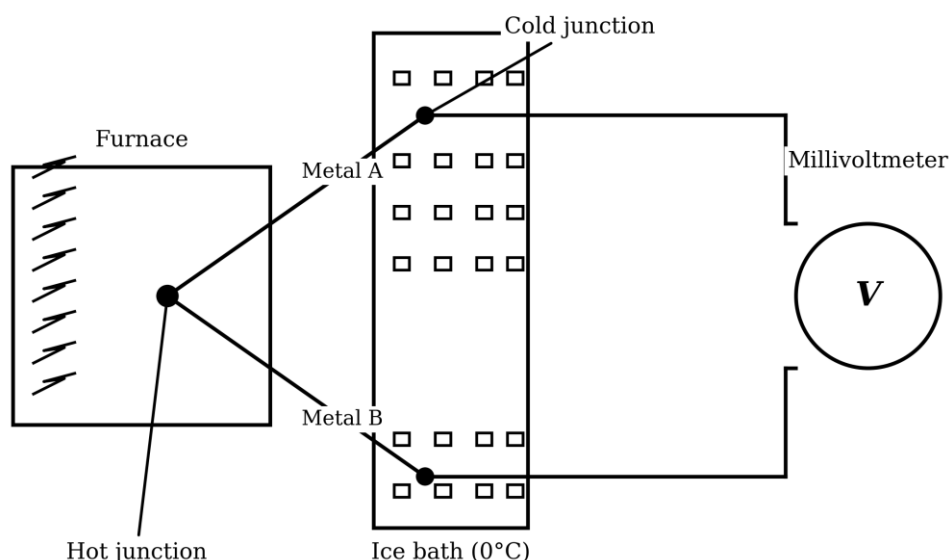


Figure: Thermocouple thermometer. Two wires of dissimilar metals (Metal A and Metal B) are welded together at a hot junction placed inside the furnace whose temperature is being measured. A cold junction is formed where each wire meets a copper lead in an ice bath held at 0°C . The millivoltmeter reads the thermoelectric EMF generated around the loop.

The Seebeck EMF is not a linear function of the hot-junction temperature. For a typical thermocouple with the cold junction at 0°C , the EMF varies with the hot-junction temperature θ according to a parabolic relation, which is commonly written in one of two algebraically equivalent forms:

$$E = a\theta + b\theta^2 \quad \text{or} \quad E = a\theta + \frac{1}{2}b\theta^2$$

The two forms arise from different conventions in the literature, and a student should read any given problem carefully to see which form the constants a and b are fitted to. Both a and b are specific to the pair of metals used; a is always positive and b is always small and negative, which means the EMF rises rapidly with temperature at first, reaches a peak at some specific temperature, and then falls as the temperature continues to climb.

Like every other empirical thermometer, the thermocouple also has an interpolation formula built from the general form of the previous section by substituting $X = E$:

$$\theta_{tc} = \left(\frac{E_\theta - E_0}{E_{100} - E_0} \right) \times 100^\circ\text{C}$$

When the cold junction is held at 0°C , E_0 is zero by construction, and the formula reduces to:

$$\theta_{tc} = \left(\frac{E_\theta}{E_{100}} \right) \times 100^\circ\text{C}$$

This is the formula a student uses when given tabulated EMF readings at the fixed points and an EMF reading at the unknown temperature. The Seebeck parabolic equation introduced above, by contrast, describes the underlying non-linear physics of the thermocouple. The two formulas serve different purposes: the interpolation formula gives the thermocouple-scale reading a calibrated instrument would display, while the parabolic Seebeck equation gives the true EMF produced by the physics of the junction. A student must always check which formula a question is asking for before substituting numbers.

The temperature at which the EMF peaks is called the **neutral temperature**, and is denoted θ_N . Physically, the **neutral temperature** is the hot-junction temperature at which the rate of increase of EMF with temperature has fallen to zero. Beyond the neutral temperature, the EMF actually decreases as the hot junction gets hotter, so the same value of E can correspond to two different hot-junction temperatures: one below θ_N and one above. A thermocouple used at or near its neutral temperature is therefore ambiguous, and a thermocouple used beyond its neutral temperature is useless as a measuring instrument unless other information pins down which side of the peak the hot junction is on.

To find the neutral temperature from the first form, we differentiate E with respect to θ and set the result to zero, because a smooth curve is flat at its peak.

Differentiating $E = a\theta + b\theta^2$ with respect to θ :

$$\frac{dE}{d\theta} = a + 2b\theta$$

Setting $\frac{dE}{d\theta} = 0$ at the neutral temperature θ_N :

$$a + 2b\theta_N = 0$$

Solving for θ_N :

$$\theta_N = -\frac{a}{2b}$$

The maximum EMF is the value of E at $\theta = \theta_N$. Substituting $\theta_N = -\frac{a}{2b}$ back into $E = a\theta + b\theta^2$:

$$E_{\max} = a\left(-\frac{a}{2b}\right) + b\left(-\frac{a}{2b}\right)^2 = -\frac{a^2}{2b} + \frac{a^2}{4b} = -\frac{a^2}{4b}$$

The minus sign in $E_{\max} = -\frac{a^2}{4b}$ does not mean the EMF itself is negative; it cancels against the negative value of b to give a positive maximum. Because $b < 0$, the quantity $-\frac{a^2}{4b}$ is positive.

For the second form of the Seebeck equation, $E = a\theta + \frac{1}{2}b\theta^2$, the same procedure gives:

$$\frac{dE}{d\theta} = a + b\theta, \Rightarrow \theta_N = -\frac{a}{b}, \Rightarrow E_{\max} = -\frac{a^2}{2b}$$

The result suggests that the neutral temperature is twice as large in the second form, and the maximum EMF is twice as large. However, this is not a physical difference between two different thermocouples; it is the same thermocouple described by two different constants fitted to two different forms of the same parabolic relationship. A student must always check which form a question uses before substituting numbers.

Related to the neutral temperature is the **inversion temperature** θ_i , defined as *the hot-junction temperature at which the EMF returns to the same value it had at the cold junction, namely zero* (assuming the cold junction is at 0°C). Setting $E = 0$ in the first form gives $\theta(a + b\theta) = 0$, whose non-zero root is:

$$\theta = -\frac{a}{b} = 2\theta_N$$

When the cold junction is held at a non-zero temperature θ_{cold} , the EMF curve shifts horizontally but remains symmetric about the neutral temperature, and the inversion temperature becomes:

$$\theta_i = 2\theta_N - \theta_{\text{cold}}$$

The Seebeck EMF, its neutral temperature, and its inversion temperature are summarised in the characteristic parabolic graph below.

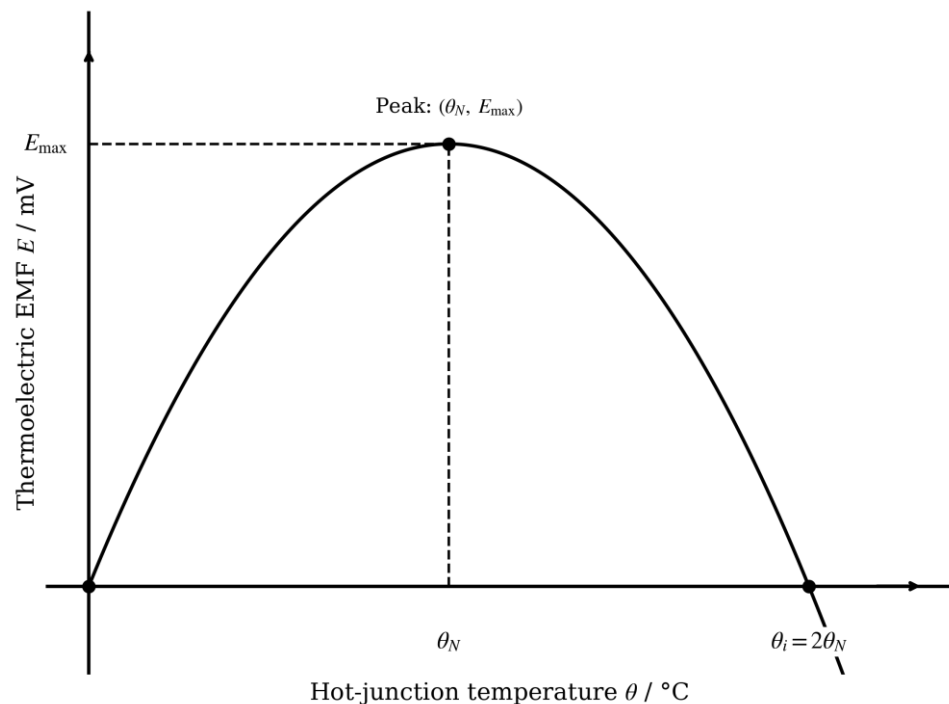


Figure: The thermoelectric EMF E as a function of hot-junction temperature θ , for a thermocouple whose cold junction is held at 0°C . The EMF rises from zero at the cold-junction temperature, reaches a maximum $E_{\max}$ at the neutral temperature θ_N , and returns to zero at the inversion temperature $\theta_i = 2\theta_N$. Beyond θ_i , the EMF reverses sign, and the thermocouple is then said to be inverted.

The thermocouple's strengths are its *small size*, its *fast response* (because the junction has almost no thermal mass), its *wide range* (from -200°C to over 1500°C depending on the metal pair), and its *ability to measure the temperature at a single point* rather than over a volume. Its weaknesses are its relatively *low accuracy* (typically one or two degrees, not thousandths), its *non-linearity* (forcing the parabolic formula above), and its *need for a stable cold-junction reference*. Thermocouples are the instrument of choice in industrial applications where speed, ruggedness, and high temperature matter more than ultimate precision.

5. Pyrometers

For temperatures above roughly 1500°C , where even the best thermocouple wires begin to melt or drift, thermometry moves to a different principle altogether. A **pyrometer** *measures temperature by detecting the radiation emitted by a hot body, without making any physical contact with the body itself*. The hottest furnaces, the interiors of blast furnaces, and the surface temperature of the sun itself are measured with pyrometers. The physics behind them is **Wien's displacement law**, which relates the wavelength of peak thermal emission to the absolute temperature, and **Stefan's law**, which relates the total power radiated to the fourth power of the absolute temperature. Both laws are derived formally in Chapter 3; here we simply note that the colour and brightness of a glowing body encode its temperature, and that an optical pyrometer reads that code. Pyrometers are expensive, specialist instruments used chiefly in metallurgy and astrophysics.

Comparing the thermometers at a glance

The five thermometer types we have met differ in range, sensitivity, response time, and intended use. A summary table is the fastest way to see all of them side by side.

Thermometer	Thermometric property	Approximate range	Accuracy	Response time	Typical use
Mercury-in-glass	Volume (column length) of mercury	-39°C to 357°C	$\pm 0.1^{\circ}\text{C}$	Seconds	Laboratory and clinical work
Alcohol-in-glass	Volume (column length) of alcohol	-115°C to 78°C	$\pm 0.5^{\circ}\text{C}$	Seconds	Low-temperature and domestic use
Constant-volume gas	Pressure of a fixed mass of gas	-270°C to 1500°C	$\pm 0.001^{\circ}\text{C}$	Minutes	Metrological standard
Platinum resistance	Electrical resistance of platinum	-200°C to 1000°C	$\pm 0.01^{\circ}\text{C}$	Seconds	Sub-standard calibration; precision laboratory use
Thermocouple	Seebeck EMF between two metals	-200°C to 1500°C	$\pm 1^{\circ}\text{C}$	Milliseconds	Industrial use; rapid or point measurements
Optical pyrometer	Wavelength or brightness of emitted radiation	500°C to $3000^{\circ}\text{C}+$	$\pm 5^{\circ}\text{C}$	Instant	Furnaces, molten metal, astrophysics

The table above is not a memorisation exercise. Every entry in it can be recovered from the physics of the thermometer it describes. A thermocouple responds in milliseconds because a junction has almost no thermal mass to warm up. A gas thermometer needs minutes because a gas bulb is a reservoir of molecules that must all reach equilibrium. An alcohol thermometer cannot reach 78°C because ethanol boils there, and the column simply disappears. When in doubt, think about the physics, and the numbers in the table will fall out naturally.

Theory and you have built thermometers together for several pages now. Time for thermal equilibrium with five worked examples to see the instruments in action.

BINDER Example 6

A constant-volume gas thermometer has a gas pressure of $0.900 \times 10^5 \text{ Pa}$ at the ice point and $1.230 \times 10^5 \text{ Pa}$ at the steam point. When the bulb is immersed in a water bath, the gas pressure is $1.073 \times 10^5 \text{ Pa}$. Calculate the temperature of the bath on the gas scale.

Solution

Substituting $P_0 = 0.900 \times 10^5 \text{ Pa}$, $P_{100} = 1.230 \times 10^5 \text{ Pa}$, and $P_\theta = 1.073 \times 10^5 \text{ Pa}$ into the gas-thermometer formula:

$$\theta_{\text{gas}} = \left(\frac{P_\theta - P_0}{P_{100} - P_0} \right) \times 100^\circ\text{C}$$

$$\theta_{\text{gas}} = \left(\frac{1.073 \times 10^5 - 0.900 \times 10^5}{1.230 \times 10^5 - 0.900 \times 10^5} \right) \times 100^\circ\text{C}$$

$$\theta_{\text{gas}} = 52.4^\circ\text{C}$$

The temperature of the bath is 52.4°C .

Making Sense of the Answer: The factor of 10^5 cancels from top and bottom of the fraction, so the pressures could equally well have been given in any other consistent units, and the answer would have been the same. That is the quiet freedom of interpolation formulas: they care only about the ratio of differences, not about the units in which those differences are measured.

Think Like a Physicist: A gas-thermometer calculation is always the same three-step operation: identify the three pressures (at the ice point, at the steam point, and at the unknown temperature), substitute, and compute. The answer will always be a number between the two fixed-point temperatures (assuming the unknown lies between them), so if the calculation produces a value below 0°C or above 100°C , check whether the unknown pressure really lies in the expected range.

BINDER Example 7

A platinum resistance thermometer is calibrated at the ice point (0°C) and the steam point (100°C), with measured resistances of $R_0 = 50.00\Omega$ and $R_{100} = 69.50\Omega$. The variation of platinum resistance with true temperature is known to follow $R_\theta = R_0(1 + a\theta + b\theta^2)$ with $a = 3.98 \times 10^{-3}\text{K}^{-1}$ and $b = -5.87 \times 10^{-7}\text{K}^{-2}$. When the thermometer is immersed in a liquid whose true temperature (as measured by a gas thermometer) is 50°C , calculate:

- (a) the actual resistance of the platinum coil at this temperature, and
- (b) the temperature the platinum thermometer reports on its own empirical scale.

Solution

- (a) Substituting $\theta = 50^\circ\text{C}$ into $R_\theta = R_0(1 + a\theta + b\theta^2)$:

$$R_{50} = 50.00 (1 + (3.98 \times 10^{-3})(50) + (-5.87 \times 10^{-7})(50)^2)\Omega = 59.88\Omega$$

The actual resistance of the platinum coil is 59.88Ω .

- (b) Substituting into the interpolation formula for the platinum scale:

$$\theta_R = \left(\frac{R_\theta - R_0}{R_{100} - R_0} \right) \times 100^\circ\text{C} = \left(\frac{59.88 - 50.00}{69.50 - 50.00} \right) \times 100^\circ\text{C}$$

$$\theta_R = \left(\frac{9.88}{19.50} \right) \times 100^\circ\text{C} = 50.67^\circ\text{C}$$

The platinum scale reads 50.67°C when the true temperature is 50°C .

Making Sense of the Answer: The platinum thermometer over-reads by about two-thirds of a degree at the mid-point of its calibration range. This is exactly the kind of disagreement we studied in the previous section, and the numerical size of the disagreement (less than 1°C out of 100°C) confirms that platinum resistance is only mildly non-linear. A mercury thermometer in the same bath would under-read by a similar amount in the opposite direction, and a gas thermometer would sit very close to the true value.

Think Like a Physicist: When a problem gives a functional form such as $R_\theta = R_0(1 + a\theta + b\theta^2)$, resist the temptation to use it as the thermometer's scale. That functional form tells you the true physical resistance at the true temperature, which is a different question from the reading the thermometer produces. The reading is always obtained by substituting into the linear interpolation formula. The two give different answers precisely because the real physics is non-linear.

REAL Example 8

Kipanga is in the physics laboratory, measuring the temperature of a hot metal block with a thermocouple. The thermocouple has $a = 40\mu\text{V}/^\circ\text{C}$ and $b = -0.04\mu\text{V}/^\circ\text{C}^2$ in the form $E = a\theta + b\theta^2$, with the cold junction held in an ice bath at 0°C . Kipanga reports a hot-junction EMF of 9.90 mV and proudly announces that the block is at about 470°C . Kipute, watching from the next bench, puts down her own work and frowns.

Kipute: “Kipanga, before you write that number down, do you know what the neutral temperature of that thermocouple is?”

Kipanga: “It is where the EMF is maximum. I remember that. Why?”

Kipute: “Compute it from the values of a and b that are stamped on the side of your apparatus.”

Kipanga computes the neutral temperature and finds that it is 500°C , only 30°C above his answer. Kipute then asks him to consider three specific problems with his measurement. What was Kipute's argument? Address:

- (a) why a reading taken near the neutral temperature is ambiguous,
- (b) why the sensitivity of the thermocouple near the neutral temperature is poor, and
- (c) why, in view of these two issues, it may be wiser to use a different thermometer entirely for this particular measurement.

Solution

(a) Near the neutral temperature, the EMF-versus-temperature curve is flat (because $dE/d\theta = 0$ at θ_N), which means a small change in the hot-junction temperature produces only a **tiny** change in the EMF. More dangerously, the same EMF corresponds to **two** different hot-junction temperatures, one below θ_N and one above. A reading of 9.90 mV could equally well represent a hot junction at 470°C or one at 530°C (the two temperatures that are symmetric about the neutral temperature 500°C), and the thermocouple alone cannot tell Kipanga which is the correct one.

(b) The sensitivity of a thermocouple is $dE/d\theta$, which is the change in EMF per change in temperature. At the neutral temperature, $dE/d\theta = 0$, which means the thermocouple is completely unable to distinguish temperatures at that point. Slightly off the neutral temperature, $dE/d\theta$ is small but non-zero, so a small temperature change produces only a small EMF change. Kipanga's reading lies in a region where even a very careful EMF measurement corresponds to a wide range of possible temperatures, and the instrument's precision is badly degraded.

(c) Given both problems, the thermocouple is the wrong tool for this measurement. A platinum resistance thermometer would have no ambiguity (its resistance rises monotonically with temperature over this range) and excellent sensitivity (of order a tenth of an ohm per degree). If the block is hot enough that a platinum coil cannot be used safely, an optical pyrometer, which reads colour temperature directly and has no neutral-temperature problem at all, is the obvious alternative. Kipanga's confident announcement of 470°C is therefore based on a measurement the thermocouple cannot actually make with the precision he claims, and the sensible response is to repeat the measurement with a different instrument.

Making Sense of the Answer: Kipute's three points are not abstract worries. Each one is a different failure mode of the same underlying physics, namely that the Seebeck EMF is a parabolic function of temperature and therefore has a peak. Below the peak, the EMF rises and the thermometer works; above the peak, the EMF falls and the thermometer lies about the direction of change; near the peak, the EMF is almost constant and the thermometer lies about the size of the change. The right response to all three failures is the same: pick an instrument whose thermometric property is monotonic in the temperature range of interest.

Think Like a Physicist: Any thermometric instrument has a working range, and the working range ends well before the property stops changing altogether. For a thermocouple, the working range ends some distance below the neutral temperature, precisely to keep the sensitivity $dE/d\theta$ comfortably above zero. Whenever you are asked to use a thermocouple near its neutral temperature, the answer is almost always

that you should not. A physicist who knows the shape of a curve knows which parts of the curve are safe to measure on, and which parts are not.

HOT Example 9

A thermocouple has its cold junction held in an ice-water bath at 0°C and its thermoelectric EMF follows $E = a\theta + \frac{1}{2}b\theta^2$, with $a = 52.0\mu\text{V}/^\circ\text{C}$ and $b = -0.040\mu\text{V}/^\circ\text{C}^2$. Calculate (a) the neutral temperature, (b) the maximum EMF, and (c) the inversion temperature.

Solution

(a) The neutral temperature is where $\frac{dE}{d\theta} = 0$. Differentiating the given form of E :

$$\frac{dE}{d\theta} = a + b\theta$$

Setting $\frac{dE}{d\theta} = 0$ at $\theta = \theta_N$:

$$a + b\theta_N = 0 \Rightarrow \theta_N = -\frac{a}{b}$$

$$\theta_N = -\frac{52.0\mu\text{V}/^\circ\text{C}}{-0.040\mu\text{V}/^\circ\text{C}^2} = 1300^\circ\text{C}$$

The neutral temperature is 1300°C .

(b) The maximum EMF is E evaluated at $\theta = \theta_N$. Substituting $\theta_N = -\frac{a}{b}$ back into the given form:

$$E_{\max} = a\left(-\frac{a}{b}\right) + \frac{1}{2}b\left(-\frac{a}{b}\right)^2 = -\frac{a^2}{b} + \frac{a^2}{2b} = -\frac{a^2}{2b}$$

Substituting numerical values:

$$E_{\max} = -\frac{(52.0\mu\text{V}/^\circ\text{C})^2}{2(-0.040\mu\text{V}/^\circ\text{C}^2)} = 3.38 \times 10^4 \mu\text{V} = 33.8\text{mV}$$

The maximum EMF is 33.8mV .

(c) The inversion temperature is where the EMF returns to 0. With the cold junction at 0°C , setting $E = 0$:

$$a\theta + \frac{1}{2}b\theta^2 = 0 \Rightarrow \theta\left(a + \frac{1}{2}b\theta\right) = 0$$

The non-zero root is:

$$\theta_i = -\frac{2a}{b} = 2\theta_N$$

$$\theta_i = 2(1300^\circ\text{C}) = 2600^\circ\text{C}$$

The inversion temperature is 2600°C .

Making Sense of the Answer: A neutral temperature of 1300°C is well above the melting point of most metals a Form 5 student has met, which is exactly where the thermocouple's usefulness lies. The inversion temperature at 2600°C is higher than any practical thermocouple can actually survive, because the metal wires themselves would have melted long before the hot junction reaches that temperature. The inversion temperature is therefore an idealised mathematical root of the parabolic fit, not an operational limit; the real working limit is set by whichever wire melts first.

Think Like a Physicist: When a problem gives you a and b together with a form of E , the first move is always to identify which form (with or without the factor of $1/2$) the constants are fitted to. The same pair of numerical values of a and b produces answers that differ by a factor of two between the two forms, and mixing forms is one of the commonest mistakes in thermocouple problems. Read the formula given in the question before reaching for any memorised shortcut.

HOT Example 10

At an unknown temperature, a constant-volume gas thermometer gives a pressure reading of $P_\theta = 1.120 \times 10^5\text{Pa}$ (where the ice and steam point pressures are $P_0 = 0.800 \times 10^5\text{Pa}$ and $P_{100} = 1.200 \times 10^5\text{Pa}$), and a platinum resistance thermometer in the same bath reads $R_\theta = 58.50\Omega$ (where $R_0 = 50.00\Omega$ and $R_{100} = 70.00\Omega$).

(a) Calculate the temperature of the bath on the gas scale and on the platinum resistance scale.

(b) Using the gas-scale value as the true temperature, calculate the percentage discrepancy of the platinum-scale reading.

(c) Assess the reliability of the platinum resistance thermometer reading and state what action a physicist should take.

Solution

(a) Substituting into the interpolation formula for each thermometer.

Gas scale:

$$\theta_{\text{gas}} = \left(\frac{P_{\theta} - P_0}{P_{100} - P_0} \right) \times 100^{\circ}\text{C} = \left(\frac{1.120 - 0.800}{1.200 - 0.800} \right) \times 100^{\circ}\text{C} = 80.0^{\circ}\text{C}$$

The temperature of the bath on the gas scale is 80.0°C .

Platinum resistance scale:

$$\theta_{\text{R}} = \left(\frac{R_{\theta} - R_0}{R_{100} - R_0} \right) \times 100^{\circ}\text{C} = \left(\frac{58.50 - 50.00}{70.00 - 50.00} \right) \times 100^{\circ}\text{C} = 42.5^{\circ}\text{C}$$

The temperature of the bath on the platinum resistance scale is 42.5°C .

(b) Percentage discrepancy of the platinum reading relative to the true gas-scale value:

$$\text{Discrepancy} = \frac{\theta_{\text{R}} - \theta_{\text{gas}}}{\theta_{\text{gas}}} \times 100\% = \frac{42.5 - 80.0}{80.0} \times 100\% = -46.9\%$$

The percentage discrepancy is 46.9% below the true temperature.

(c) The platinum-scale reading is 46.9% below the true temperature. A discrepancy of this size is far too large to result from ordinary non-linearity of platinum resistance, which for a properly constructed coil would be less than a few percent over this range. The most likely explanation is that the platinum coil has been damaged, contaminated, or badly calibrated. A physicist in this situation would send the platinum resistance thermometer for recalibration or replacement.

Making Sense of the Answer: *The numbers in this problem are chosen to make the two thermometers disagree visibly. In a real laboratory, a 50% disagreement between a gas thermometer and a platinum resistance thermometer would not be a textbook illustration of non-linearity; it would be a sign that something is wrong with the instrument. Real platinum resistance thermometers agree with gas thermometers to better than a degree over their working range, and the small difference between the two can be calibrated out for high-precision work.*

Think Like a Physicist: *When two independent thermometers disagree by a few tenths of a degree, suspect their physics (non-linearity of the properties). When they disagree by several degrees, suspect their calibration. When they disagree by tens of degrees or more, suspect the instrument itself. Different sizes of disagreement point to different causes, and a physicist's first act on meeting a large discrepancy is not to invoke subtle theory but to check whether the apparatus is actually working.*

We have now met every instrument a student is likely to encounter in a laboratory or on an examination paper. Each one uses a different thermometric property, each one has its own range and its own non-linearity, and each one agrees with the others only at the fixed points on which it was calibrated. The disagreement between them is not a flaw of the thermometers, it is a flaw of the very idea of calibrating a scale on two points and assuming linearity in between. If we want a single temperature scale that every thermometer can be made to agree with, a scale that does not depend on the liquid inside the glass or the metal inside the wire, we have to escape the empirical framework altogether. The next section shows how physicists did exactly that, by choosing a single fixed point defined by the physics of matter itself, and by building a scale on it that is universal across every substance in nature.

THE THERMODYNAMIC SCALE AND THE TRIPLE POINT

Every empirical thermometer we have met so far is a good instrument with a small private flaw. The flaw is that it reports its own physics. A mercury thermometer reports how mercury expands, a platinum thermometer reports how platinum's resistance changes, a gas thermometer reports how a fixed mass of gas responds. Calibrating any of these instruments on two fixed points produces a scale that agrees with the other scales only at those two points. Between the fixed points, each thermometer tells a slightly different story, and nothing in the empirical approach can say which story is the right one.

This is not a small inconvenience. A temperature of " 50°C on the mercury scale" is a legitimate quantity, but so is " 50°C on the platinum scale", and the two numbers refer to physically different states. If the world contained only one thermometric substance we would be fine; since it does not, we are forced to ask a harder

question. Is there a temperature scale that does not depend on the thermometric substance at all? A scale that every properly constructed thermometer would converge towards in the limit of perfect construction? This section answers that question with a clear yes, and builds the scale explicitly.

The problem with two fixed points

The empirical Celsius scale uses two fixed points, the ice point and the steam point. Both are sensitive to external conditions in ways that the casual reader does not always appreciate. The **ice point** (*the temperature at which pure water and pure ice coexist at equilibrium*) depends on the applied pressure, because changing the pressure on ice shifts the melting temperature slightly. The **steam point** (*the temperature at which pure water boils*) depends very strongly on the applied pressure, because **boiling** is the point where the vapour pressure of water equals the atmospheric pressure above it. A laboratory 2000m above sea level and a laboratory at sea level will therefore measure two slightly different ice points and two significantly different steam points, unless both are carefully corrected to standard atmospheric pressure.

Even worse, the “ice point” as normally prepared in a laboratory contains dissolved air in the water, which depresses the freezing temperature by about 0.01°C relative to pure water with no air dissolved in it. This is small, but it is also uncorrectable without careful preparation, and a scale that depends on how carefully the ice-water mixture was prepared is a scale with an unwelcome human fingerprint on it.

The thermodynamic scale escapes all of this by refusing to use either the ice point or the steam point as a reference. Instead, it uses a single fixed point defined by a physical state that is unique in nature, and from that single point it builds the entire scale.

The triple point of water

A **phase diagram** for water plots pressure on the vertical axis and temperature on the horizontal axis, and marks the curves along which two different phases of water coexist in equilibrium. Along the fusion curve, solid and liquid coexist; along the vapourisation curve, liquid and vapour coexist; along the sublimation curve, solid and vapour coexist. The three curves are not independent: they must meet at a single point, and at that point all three phases of water (ice, liquid water, and water vapour) coexist simultaneously in equilibrium. This single point is called the **triple point of water**.

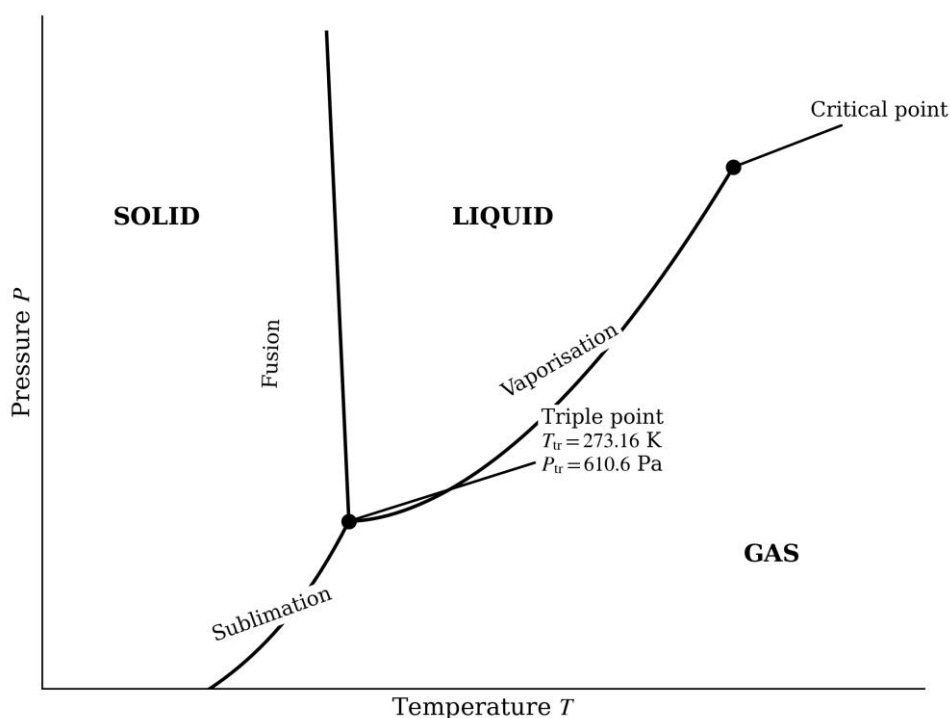


Figure: The phase diagram of water. The three phase-boundary curves (fusion, vapourisation, sublimation) meet at the triple point, at which solid, liquid, and vapour coexist in equilibrium at a unique temperature and pressure. The vapourisation curve ends at the critical point, beyond which liquid and vapour become indistinguishable.

The triple point of water occurs at a unique combination of temperature and pressure, specifically at $T_{tr} = 273.16\text{K}$ and $P_{tr} \approx 610.6\text{Pa}$ (about 0.006atm). These two numbers are not adjustable. They are properties of water itself, determined by the intermolecular forces in the water molecule, and any laboratory anywhere in the universe that correctly prepares a triple-point cell (a sealed vessel containing pure water with ice, liquid, and vapour in equilibrium) will measure the same temperature and the same pressure to within the precision of its instruments.

Compare this with the ice point. An ice-point bath depends on the purity of the water, the presence of dissolved air, and the applied atmospheric pressure, and each of these introduces a small correction that must be handled carefully. A triple-point cell depends on none of them. Once the cell is sealed, the pressure is fixed by the vapour pressure of water in contact with ice, the temperature is fixed by the simultaneous coexistence of all three phases, and nothing outside the cell can change either value without breaking the triple-point condition. The triple point is therefore *reproducible to an unprecedented precision*, and it is the only fixed point used to define the modern thermodynamic scale.

Why the ice point sits at 273.15 K, not 273.16 K

A small numerical detail is worth clearing up before we go further. The triple point of water is at exactly 273.16K , by definition of the kelvin. The ice point of pure water at standard atmospheric pressure sits at 273.15K , which is exactly 0.01K (i.e. 0.01°C) below the triple point. The 0.01K difference comes from two small corrections.

First, the ice point uses pure water and ice at standard atmospheric pressure ($1.013 \times 10^5\text{Pa}$), not at the triple-point pressure of 610.6Pa , and the higher pressure lowers the freezing temperature slightly.

Second, the ice point uses water that has been in contact with air, and the dissolved air lowers the freezing point by a further small amount.

Adding the two corrections gives approximately the 0.01K gap between the two reference temperatures.

The thermodynamic-scale formula

The thermodynamic scale is built on the observation that the pressure of an ideal gas held at constant volume is directly proportional to its absolute temperature. For an ideal gas at constant volume, Gay-Lussac's law gives

$$P \propto T \quad \Rightarrow \quad \frac{P}{T} = \text{constant}$$

This relation becomes exact in the limit of vanishing gas density and is very nearly exact for any dilute real gas. Now imagine a constant-volume gas thermometer whose bulb is first immersed in a triple-point cell and gives a pressure reading P_{tr} , and then immersed in a body at an unknown temperature T where it gives a pressure reading P_θ . Applying the constant-ratio relation at both temperatures:

$$\frac{P_\theta}{T} = \frac{P_{tr}}{T_{tr}}$$

Solving for T :

$$T = T_{tr} \times \frac{P_\theta}{P_{tr}}$$

Substituting the defined value $T_{tr} = 273.16\text{K}$:

$$T = 273.16\text{K} \times \frac{P_\theta}{P_{tr}}$$

This single equation defines the thermodynamic temperature T of any body, using only the pressure reading of an ideal gas thermometer at two states: the state of the body, and the triple point of water.

Notice that the whole scale has been built from a single fixed point. There is no ice point, no steam point, no interpolation. The scale is absolute, in the sense that $T = 0$ corresponds to $P_\theta = 0$, which is the state an ideal gas would reach if its pressure could be brought to zero at constant volume. This temperature is called **absolute zero**, and it is the lowest temperature physically permitted by the laws of thermodynamics.

Absolute zero and the Celsius–Kelvin conversion

Absolute zero is not a temperature that any real gas thermometer can actually reach, because *every real gas liquefies or solidifies long before its pressure approaches zero*. But by measuring the pressure of a gas at several high temperatures and extrapolating the straight line $P \propto T$ backwards to $P = 0$, the horizontal-axis intercept gives the temperature of absolute zero. The result, common to every real gas in the dilute limit, is $T = 0\text{K}$, which on the Celsius scale corresponds to $\theta = -273.15^\circ\text{C}$.

The Celsius scale and the thermodynamic (kelvin) scale are therefore related by a simple additive shift:

$$T(\text{K}) = \theta(^{\circ}\text{C}) + 273.15$$

A change of 1K and a change of 1°C are numerically equal, because the two scales have the same unit size; only their zero points differ. The ice point is 273.15K on the absolute scale and 0°C on the Celsius scale. The steam point is 373.15K and 100°C . The triple point is 273.16K and 0.01°C . Absolute zero is 0K and -273.15°C . These correspondences, together with the older Fahrenheit scale still used in some countries, are summarised in the diagram below.

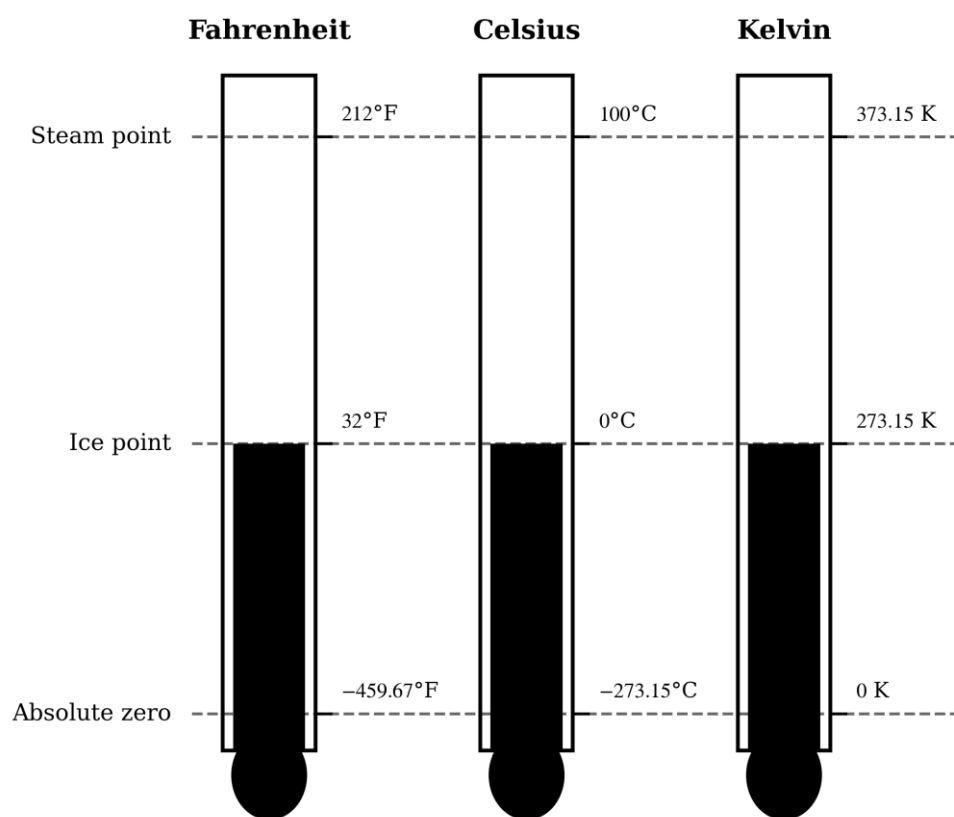


Figure: The three temperature scales side by side. The Fahrenheit scale places the ice point at 32°F and the steam point at 212°F . The Celsius scale uses 0°C and 100°C . The thermodynamic (Kelvin) scale uses 273.15K and 373.15K, with absolute zero as its true zero. Notice that absolute zero is not labelled on the Celsius or Fahrenheit scales as a defining point; it is merely a value those scales happen to reach.

Theory and you have travelled from empirical scales to a substance-independent absolute scale in a handful of pages. Time for thermal equilibrium with two worked examples to settle the ideas.

BINDER Example 11

The modern thermodynamic scale uses only the triple point of water as its fixed reference, while the older Celsius scale used two fixed points, the ice point and the steam point.

(a) Explain why the triple point is a more reliable fixed reference than the ice point.

(b) Show, using the constant-volume gas thermometer equation, how a single fixed point combined with the ideal-gas relation is sufficient to define an entire temperature scale.

Solution

(a) The ice point depends on three external conditions that are difficult to control. First, the temperature at which pure water and pure ice coexist depends on the applied pressure, and any deviation from standard atmospheric pressure shifts it slightly. Second, water in contact with air absorbs dissolved oxygen and nitrogen, which depress the freezing point by roughly 0.01°C . Third, the purity of the water itself affects the freezing temperature. A laboratory that wishes to use the ice point as a reference must correct for all three, and the corrections introduce small laboratory-to-laboratory variations.

The triple point, by contrast, depends on none of these. At the triple point, pure water, pure ice, and water vapour coexist in a single sealed vessel at a temperature and pressure that are both determined uniquely by the physics of the water molecule itself. The vessel is sealed, so no air can dissolve in the water. The pressure inside the cell is fixed at 610.6Pa by the vapour pressure of water in contact with ice, so no external atmosphere can disturb it. The temperature inside the cell is fixed at 273.16K by the requirement that all three phases coexist in equilibrium. Any laboratory that prepares a triple-point cell correctly will measure the same temperature and the same pressure, regardless of altitude, weather, or local atmospheric composition. This is what makes the triple point “fully reproducible” in the physicist’s sense: the same physical state gives the same number everywhere.

(b) The thermodynamic scale is built on the constant-volume gas relation $P \propto T$, which for an ideal gas gives $\frac{P}{T} = \text{constant}$. Applying this at the triple point and at an unknown temperature T :

$$\frac{P_{\text{tr}}}{T_{\text{tr}}} = \frac{P_{\theta}}{T} \Rightarrow T = T_{\text{tr}} \times \frac{P_{\theta}}{P_{\text{tr}}}$$

With the defined value $T_{\text{tr}} = 273.16\text{K}$:

$$T = 273.16\text{K} \times \frac{P_{\theta}}{P_{\text{tr}}}$$

One fixed point (the triple point, which fixes T_{tr}) and one physical law (the ideal gas relation) together determine T at every other pressure. No second fixed point is needed, because the ideal gas law itself fixes the shape of the P -versus- T relation. The empirical Celsius scale needed two fixed points because it made no assumption about the shape of its thermometric property’s variation with temperature; it had to calibrate that shape at two points and interpolate linearly. The thermodynamic scale needs only one fixed point because the ideal-gas law tells us in advance that the shape is linear through the origin.

Making Sense of the Answer: *The move from two fixed points to one is not just a cosmetic simplification. It is the move from an empirical to a fundamental temperature scale, because the ideal-gas relation is a law of physics and not a property of a particular substance. Every thermometer ever built, if used correctly, will converge on the thermodynamic temperature in the limit of vanishing gas density. The triple point is simply the anchor the world has agreed to peg the scale to, and 273.16K is the number that makes the new scale numerically consistent with the older Celsius scale at the ice point.*

Think Like a Physicist: *When you are told that a scale has been defined by a single reference point and a physical law, look for the law carefully. The law is doing half the work. In the thermodynamic scale, the law is Gay-Lussac’s $P \propto T$; in other contexts it might be $E = mc^2$ or $\lambda T = \text{constant}$. The reference point alone is never enough to build a scale; it is the combination of a reference point and a law that makes the construction possible.*

BINDER Example 12

A constant-volume gas thermometer has a pressure of $1.500 \times 10^4\text{Pa}$ at the triple point of water. The bulb is then immersed in a pan of boiling water at an unknown temperature, and the pressure rises to $2.050 \times 10^4\text{Pa}$.

(a) Find this temperature on the Kelvin scale.

(b) Convert the answer to the Celsius scale.

(c) Comment on whether the result is consistent with the boiling point of water at standard atmospheric pressure.

Solution

(a) Substituting into the thermodynamic-scale formula $T = 273.16\text{K} \times \frac{P_\theta}{P_{tr}}$:

$$T = 273.16\text{K} \times \frac{2.050 \times 10^4\text{Pa}}{1.500 \times 10^4\text{Pa}} = 373.3\text{K}$$

The temperature on the Kelvin scale is 373.3K.

(b) Converting to Celsius:

$$\theta = T - 273.15 = 373.3 - 273.15 = 100.1^\circ\text{C}$$

The temperature on the Celsius scale is 100.1°C.

(c) The boiling point of water at standard atmospheric pressure is 100°C, and the thermometer reads 100.1°C. The two values agree to within 0.1°C, which is consistent with the thermometer measuring a body of boiling water at close to standard pressure. The small positive discrepancy could come from a slightly elevated atmospheric pressure in the laboratory (which would raise the boiling point above 100°C), or from a small non-ideal-gas correction in the thermometer, or from ordinary experimental error. In any case, the reading is physically sensible for boiling water.

Making Sense of the Answer: *The answer matches a temperature the student already recognises, which is the whole purpose of the exercise. The thermodynamic formula is not a mystery: it is the ideal-gas relation rearranged to give a temperature from a pressure ratio. When the ratio is roughly 1.37, the temperature is roughly 1.37 times the triple-point temperature, which puts it close to the boiling point of water. Whenever a calculation produces a number close to a physically meaningful temperature like 273K or 373K, pause and check which physical state the body is probably in.*

Think Like a Physicist: *A gas thermometer calculation on the thermodynamic scale is always the same two-step operation: take the ratio of the two pressures, and multiply by 273.16K. The whole calculation is one multiplication, yet it is one of the most powerful measurements in physics, because it measures an absolute temperature without reference to any other substance. If a problem gives you a pressure ratio and asks for a thermodynamic temperature, do not reach for the Celsius interpolation formula; the answer is always 273.16K times the ratio.*

We have now escaped the empirical framework and reached a temperature scale that every laboratory in the world can measure against the same reference. But the thermodynamic scale is not the end of the story. In practice, physicists still use mercury thermometers, platinum coils, and thermocouples, all of which are empirical instruments. *When exactly do these agree with the thermodynamic scale, and when do they diverge?* The next section takes the four thermometers we have built, sets them side by side at the same temperature, and works out how their readings compare.

COMPARISON AND LIMITATIONS OF THERMOMETRIC SCALES

We now have a zoo of thermometric scales at our disposal. Every empirical thermometer from the previous sections gives its own scale, and alongside them we have the thermodynamic (Kelvin) scale, which was derived from first principles and depends on no particular substance. The empirical scales agree with each other, and with the thermodynamic scale, only at the fixed points on which they were calibrated. Everywhere else, they disagree by small but systematic amounts. This section examines the pattern of agreement and disagreement, identifies which thermometer sits closest to the thermodynamic truth, and states the limitations that remain even for the best empirical instrument.

Where the scales agree and where they diverge

Consider four thermometers: a mercury-in-glass thermometer, a platinum resistance thermometer, a constant-volume gas thermometer, and a thermocouple. Suppose all four are carefully calibrated at the ice point and the steam point, and all four are then immersed in a water bath at some intermediate temperature. What readings do they give?

At the ice point and the steam point, all four thermometers give the same reading by construction. This is trivial, because calibration forces every thermometer to agree at its two calibration temperatures. The interesting question is what happens between them.

At every intermediate temperature, the four thermometers give four slightly different readings. The mercury thermometer reads slightly below the true temperature (because the **expansion of mercury is sub-linear** in temperature between the fixed points), the platinum resistance thermometer reads slightly above (because the **resistance of platinum is super-linear**), the thermocouple reads an amount that depends on where the measurement falls on its parabolic EMF curve, and the gas thermometer reads very nearly correctly, closer to the truth than any of the other three. The pattern is not random. Each thermometer's deviation reflects the shape of its own thermometric property, and *the size of the deviation reflects how non-linear that property is*.

Why the gas thermometer is closest to the truth

Among all empirical thermometers, the constant-volume gas thermometer enjoys a special status because its thermometric property (the pressure of a fixed mass of gas at constant volume) is almost exactly linear in temperature. For an ideal gas, the relation $P \propto T$ is exact, and the gas-scale temperature equals the thermodynamic temperature identically. Real gases deviate slightly from ideal behaviour because their molecules attract each other weakly, and this causes the gas-thermometer scale to differ from the thermodynamic scale by a small correction.

The correction shrinks as the amount of gas in the bulb is reduced, because in the low-density limit the gas molecules are so far apart that their interactions become negligible and the ideal-gas law becomes exact. In the limit as the gas quantity tends to zero, the gas thermometer and the thermodynamic scale become identical. This is why the thermodynamic scale is said to be **defined** by the ideal-gas relation taken in the low-density limit, and it is why the gas thermometer is used as the primary standard against which all other empirical thermometers are calibrated.

Limitations that remain

Even for the gas thermometer, three limitations remain.

- 1) The low-density limit can only be approached, never reached, because a thermometer with zero gas in its bulb would produce no pressure and no reading. In practice, multiple measurements are taken at successively lower gas densities and the results are extrapolated to zero density, which introduces a small but finite error.
- 2) The gas thermometer is slow, delicate, and impractical for routine use, which means most laboratory and industrial measurements use faster thermometers (mercury, platinum, thermocouple) that are in turn calibrated against the gas thermometer.
- 3) At very high temperatures (above roughly 1000°C) no gas thermometer can survive, and the thermodynamic scale must be extended using other physical principles, such as the Stefan law of blackbody radiation or the Wien displacement law that we will meet in Chapter 3.

For the other empirical thermometers, the principal limitation is non-linearity. A mercury thermometer calibrated at the ice point and the steam point will read a few tenths of a degree away from the truth at the mid-point of its range, and the error is systematic rather than random. For most laboratory work, this error is small enough to ignore; for precision measurements, a correction table is constructed by comparing the thermometer with a gas thermometer at many intermediate temperatures, and the correction is applied to every reading.

Practical consequences for the student

For a student, two practical consequences follow from the comparison above.

The first is that a question asking for “the temperature” of a body requires the student to identify which scale is being used.

- When the question specifies a thermometer by name, the answer is that thermometer's empirical reading, obtained by substituting into its interpolation formula.
- When the question speaks of “the true temperature” or “the thermodynamic temperature”, the answer is the gas-thermometer reading (or equivalently the Kelvin temperature computed from the triple-point formula).

Treating the two as interchangeable is one of the commonest mistakes in thermometry questions.

The second is that when two thermometers are calibrated at the same fixed points and give different readings at an intermediate temperature, neither of them is broken. The disagreement is a consequence of their different physics, and the only legitimate response is to quote both readings alongside the name of the scale. A single “temperature” without a scale is an incomplete quantity, much like a length without a unit.

Theory and you have surveyed every thermometer in this chapter from a high vantage point. Two worked examples now bring the comparison down to ground level.

BINDER Example 13

A platinum resistance thermometer and a constant-volume gas thermometer are both calibrated at the ice point (0°C) and the steam point (100°C). A student argues: “*Since both thermometers use the same two fixed points, they must agree at every temperature between them.*” Explain precisely why this student is wrong, and identify the single condition under which the two scales would agree at every intermediate temperature.

Solution

The student’s argument is wrong because calibration at two fixed points forces the two thermometers to agree only at those two points, not everywhere in between. The interpolation formula assumes the thermometric property varies linearly with temperature, but the actual variation is linear for **neither** thermometer exactly. The platinum resistance is slightly super-linear with temperature, so the platinum-scale reading bulges above the straight line, while the gas pressure at constant volume is very nearly linear with temperature and sits very close to the straight line. The two scales therefore disagree at every intermediate temperature, by amounts that reflect the shape of each thermometric property’s graph.

The single condition under which the two scales would agree at every temperature is that both thermometric properties must be **exactly linear** functions of the true thermodynamic temperature, with the same slope per unit of calibration range. In practice, this is true only in an idealised limit: if the gas in the gas thermometer were truly ideal (no inter-molecular forces) and the resistance of platinum were truly linear in temperature (no higher-order terms in θ), the two scales would coincide. Real gases approach the ideal limit at low density, but real platinum does not become perfectly linear by any such limit, so the two scales always retain a small non-zero discrepancy.

Making Sense of the Answer: *Two fixed points are the minimum requirement for calibrating a scale, but they are not enough to make two different scales agree everywhere. The formula that interpolates between them is a straight line, and two curved graphs that share the same two endpoints can still differ everywhere in between. The student’s mistake is confusing agreement at the endpoints with agreement everywhere, which is a very common error for anyone who has not yet thought carefully about what linear interpolation actually promises.*

Think Like a Physicist: *Whenever two thermometers are calibrated at the same fixed points but show different readings at an intermediate temperature, the disagreement reveals the non-linearity of at least one of the two thermometric properties. This is not a weakness of the instruments; it is information about the physics of the substances inside them. A physicist who sees two thermometers disagree does not declare one of them broken; they record both readings, note the disagreement, and use it to test their understanding of the underlying material properties.*

HOT Example 14

The resistance of a platinum coil at the ice point, at the steam point, and in a certain room is $R_0 = 63.0\Omega$, $R_{100} = 75.0\Omega$, and $R_\theta = 65.09\Omega$ respectively. A constant-volume gas thermometer at the same three temperatures gives pressures $P_0 = 8.00 \times 10^6\text{Pa}$, $P_{100} = 1.10 \times 10^7\text{Pa}$, and $P_\theta = 8.51 \times 10^6\text{Pa}$.

- Calculate the room temperature on the platinum resistance scale and on the gas scale.
- Explain why the two readings differ.
- State which of the two readings is closer to the true thermodynamic temperature and justify your choice.

Solution

- Applying the interpolation formula to each thermometer.

Platinum scale:

$$\theta_R = \left(\frac{R_\theta - R_0}{R_{100} - R_0} \right) \times 100^\circ\text{C} = \left(\frac{65.09 - 63.0}{75.0 - 63.0} \right) \times 100^\circ\text{C} = 17.42^\circ\text{C}$$

The room temperature on the platinum resistance scale is 17.42°C .

Gas scale:

$$\theta_{\text{gas}} = \left(\frac{P_{\theta} - P_0}{P_{100} - P_0} \right) \times 100^{\circ}\text{C} = \left(\frac{8.51 \times 10^6 - 8.00 \times 10^6}{1.10 \times 10^7 - 8.00 \times 10^6} \right) \times 100^{\circ}\text{C} = 17.00^{\circ}\text{C}$$

The room temperature on the gas scale is 17.00°C .

(b) The two readings differ by 0.42°C because the two thermometric properties are not perfectly linear with true temperature. The resistance of platinum is slightly super-linear between the fixed points, which means the platinum resistance at the true temperature of 17°C is slightly higher than a linear interpolation from the fixed points would predict. Substituting this slightly elevated resistance into the interpolation formula returns a reading of 17.42°C , which is above the true value. The gas pressure at constant volume, by contrast, is very nearly linear with true temperature, so the gas reading of 17.00°C sits very close to the truth. The disagreement between the two scales is therefore a genuine consequence of the non-linear physics of platinum, not a calibration error.

(c) The gas-scale reading of 17.00°C is closer to the true thermodynamic temperature. The pressure of a fixed mass of gas at constant volume is nearly proportional to the thermodynamic temperature, with deviations from ideal behaviour that become negligible in the low-density limit. The resistance of platinum, by contrast, has no comparable principle of approach to the thermodynamic scale; its non-linearity is a property of the metal itself and does not disappear with any limiting procedure. The gas thermometer is therefore the instrument of choice when the true thermodynamic temperature is needed, and the platinum reading must be corrected against a gas thermometer before it can be trusted for precision work.

Making Sense of the Answer: *The two readings differ by less than half a degree out of roughly seventeen, which is less than three percent. In a physics laboratory this is a large and noticeable disagreement; in a hospital it would be undetectable; in a kitchen it would be invisible. The size of the disagreement is therefore a guide to which thermometer is appropriate for which task. Precision work demands the gas thermometer or a platinum thermometer carefully corrected against it; routine work is perfectly well served by either.*

Think Like a Physicist: *Whenever a problem gives you readings of two thermometers at the same three states, the calculation is always two parallel interpolation formulas. The interesting physics lies in explaining the difference between the two answers, not in the arithmetic of the two substitutions. Practise looking at the size of the difference and identifying which thermometric property is responsible. A disagreement of a few tenths of a degree almost always reveals a non-linearity, a disagreement of a few degrees reveals a miscalibration, and a disagreement of tens of degrees reveals a broken instrument.*

We have now closed the book on thermometry. Every thermometer we will ever need has been built, every scale we will ever quote has been defined, and every limitation has been made explicit. What remains is to leave temperature behind and move on to the other great thermal quantity, heat itself, and to learn how to compute the heat absorbed or released when a body changes its temperature. The next section introduces the specific heat capacity, which tells us how much heat it takes to warm one kilogram of a substance by one degree, and turns the thermometer from a passive reader of temperature into a tool for quantitative thermal calculation.

SPECIFIC HEAT CAPACITY

We have spent five sections learning how to measure temperature honestly. It is time now to put the thermometer to work, because temperature on its own is only half the thermal story. The other half is heat, the flow of energy from one body to another because of a temperature difference, and the question we have not yet asked is the most practical one of all. *When heat flows into a body, by how much does the body's temperature rise? When heat flows out, by how much does it fall?* The answer depends on the substance, and the quantity that encodes the answer is called the **specific heat capacity**.

Defining specific heat capacity

Experiments show that the heat ΔQ required to raise the temperature of a body is proportional to two quantities: the mass m of the body, and the temperature rise ΔT . The proportionality is written as

$$\Delta Q = mc\Delta T$$

where c is a substance-dependent constant called the specific heat capacity. Rearranging:

$$c = \frac{\Delta Q}{m\Delta T}$$

In words, the **specific heat capacity** is the heat required to raise the temperature of one kilogram of the substance by one kelvin (or equivalently one degree Celsius, since the two intervals are numerically equal). Its SI unit is $\text{J kg}^{-1}\text{K}^{-1}$.

The specific heat capacity is a property of the substance, not of the particular sample. Two kilograms of water and two hundred kilograms of water both have the same specific heat capacity of $4200\text{J kg}^{-1}\text{K}^{-1}$; the difference is that warming the larger sample by one degree takes a hundred times more heat, because it has a hundred times more mass. The formula $\Delta Q = mc\Delta T$ handles this automatically: the heat required increases with the mass for a fixed temperature rise, and with the temperature rise for a fixed mass.

Thermal capacity

For a particular sample of matter, it is sometimes convenient to bundle the mass and the specific heat capacity together into a single quantity called the **thermal capacity**:

$$C = mc$$

The **thermal capacity** (unit J K^{-1}) is the heat required to raise the whole sample's temperature by one kelvin, and the heat equation then reads $\Delta Q = C\Delta T$. The specific heat capacity belongs to the substance; the thermal capacity belongs to the sample. Two saucepans made of identical cast iron have the same specific heat capacity but different thermal capacities if one is heavier than the other. Thermal capacity is the quantity that actually decides how much heat your kitchen needs to warm something, and it reappears constantly in the worked examples that follow.

A quick note on notation: For solids and liquids, the specific heat capacity is almost independent of how the heating is carried out, whether at constant volume or constant pressure, because solids and liquids barely change their volume with temperature. For gases, however, the two differ noticeably, and physicists distinguish between c_p (specific heat at constant pressure) and c_v (specific heat at constant volume). We will meet this distinction properly in Chapter 4 when we develop the kinetic theory of gases; for the chapter in hand, the word “specific heat capacity” without subscript means the common value used for solids and liquids.

Measuring the specific heat capacity: the method of mixtures

The oldest and most intuitive way to measure the specific heat capacity of a solid is the **method of mixtures**. A known mass of the substance in question is heated to a known high temperature, then dropped into a calorimeter containing water at a known lower temperature, and the final equilibrium temperature is recorded. The principle behind the method is the conservation of energy: in an insulated system, the heat lost by the hot body equals the heat gained by the cold water plus the heat gained by the calorimeter itself.

Let m_1 be the mass of the hot substance, c_1 its (unknown) specific heat capacity, T_1 its initial high temperature, and T_f the final equilibrium temperature. Let m_2 be the mass of the water, $c_2 = 4200\text{J kg}^{-1}\text{K}^{-1}$ its specific heat capacity, T_2 its initial temperature (equal to the calorimeter's initial temperature). Let m_{cal} be the mass of the calorimeter, c_{cal} its specific heat capacity (typically copper, $c_{\text{cal}} = 400\text{J kg}^{-1}\text{K}^{-1}$).

Writing the heat balance:

$$m_1 c_1 (T_1 - T_f) = m_2 c_2 (T_f - T_2) + m_{\text{cal}} c_{\text{cal}} (T_f - T_2)$$

The left-hand side is the heat lost by the hot substance as it cools from T_1 to T_f . The right-hand side is the heat gained by the water plus the heat gained by the calorimeter, both of which warm from T_2 to T_f . Solving for c_1 :

$$c_1 = \frac{(m_2 c_2 + m_{\text{cal}} c_{\text{cal}})(T_f - T_2)}{m_1 (T_1 - T_f)}$$

The experimental arrangement is shown in the following figure. The calorimeter sits inside an insulating jacket to minimise heat loss to the surroundings. A thermometer measures the water temperature continuously. A stirrer keeps the water well mixed so that the thermometer reads a uniform temperature rather than a hot spot near the dropped block.

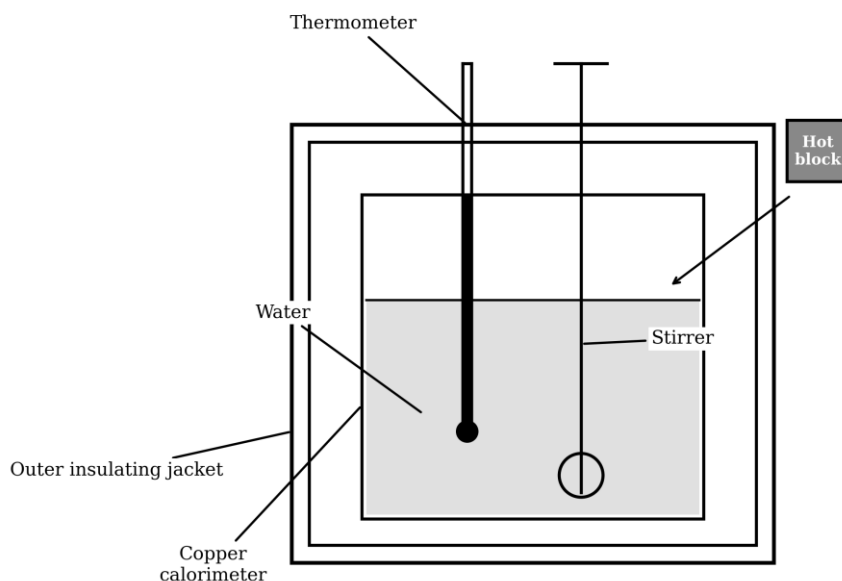


Figure: The method of mixtures. A calorimeter containing water sits inside an outer insulating jacket. A thermometer records the water temperature, a stirrer keeps the water well mixed, and the hot substance (shown here as a metal block) is lowered into the water to allow heat to flow from it to the surrounding water and calorimeter until equilibrium is reached.

Two sources of systematic error limit the accuracy of the method of mixtures.

- 1) **Heat loss from the calorimeter to the surroundings** during the experiment, which removes some of the heat that should have gone into warming the water. This is minimised by using an insulating jacket and by working quickly.
- 2) **Incomplete mixing**, in which the water near the hot block reaches a higher local temperature than the water at the thermometer and the thermometer reads below the true equilibrium value. This is minimised by thorough stirring.

A third, smaller source of error is the heat absorbed by the thermometer and stirrer themselves, which is usually neglected for rough work and corrected by an additional water-equivalent term for precision work.

Measuring specific heat capacity: the electrical method

The method of mixtures works well for solids but is awkward for liquids, because heating a liquid to a known high temperature in one vessel and transferring it without loss into another is very difficult. For liquids, the **electrical heating method** is usually preferred. A heater of known electrical power is immersed in the liquid inside an insulated calorimeter, switched on for a measured time, and the temperature rise is recorded. The heat supplied by the electrical heater equals the electrical energy dissipated:

$$Q = Pt = VIt$$

where V is the voltage across the heater, I the current through it, and t the time for which it is switched on. Setting this equal to the heat absorbed by the liquid and calorimeter:

$$VIt = (m_{\text{liq}}c_{\text{liq}} + m_{\text{cal}}c_{\text{cal}}) \Delta T$$

Solving for c_{liq} :

$$c_{\text{liq}} = \frac{VIt - m_{\text{cal}}c_{\text{cal}} \Delta T}{m_{\text{liq}} \Delta T}$$

The experimental arrangement is shown in the following figure. An ammeter in series with the heater measures the current, a voltmeter across the heater measures the voltage, and the product VI is the power dissipated in the heater.

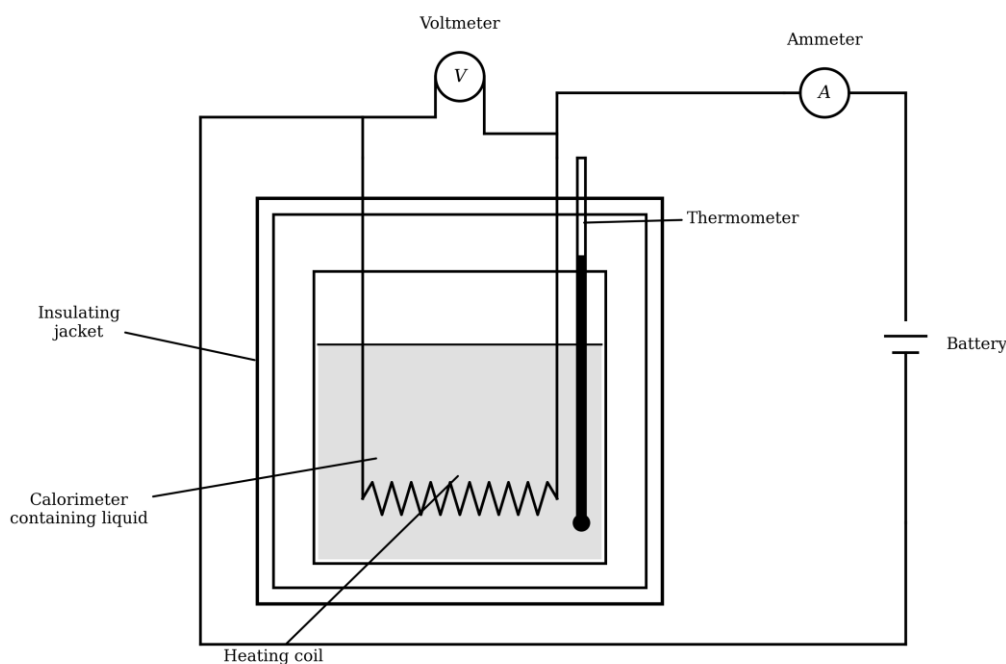


Figure: The electrical heating method for measuring the specific heat capacity of a liquid. The heating coil is immersed in the liquid and is powered by a battery through an ammeter (in series) and a voltmeter (in parallel with the coil). The thermometer records the rise in temperature of the liquid, and the total electrical energy supplied by the coil is calculated as VIt .

The electrical method is more accurate than the method of mixtures for two reasons. First, the electrical energy supplied can be measured with much greater precision than the heat transferred during a physical mixing. Second, the heater is inside the liquid throughout the experiment, so no transfer of a hot body is required and there is no risk of losing heat during the transfer.

Applications of specific heat capacity

Four familiar phenomena all trace back to specific heat capacity, each one more revealing than the last.

1. Cooking utensils

A cook in a hurry wants a pot that heats up quickly, and she therefore prefers a utensil with a low specific heat capacity (and low mass, since what matters is the thermal capacity $C = mc$). An aluminium *sufuria* ($c = 900 \text{ J kg}^{-1} \text{ K}^{-1}$) warms up much faster than an equally heavy cast-iron *sufuria* ($c = 450 \text{ J kg}^{-1} \text{ K}^{-1}$) when both are placed over the same flame, because the aluminium absorbs less heat per degree even though its specific heat capacity is higher; what matters is that the aluminium *sufuria* is typically much thinner and lighter, so its thermal capacity is lower overall. The flip side of the same principle is that a cook who wants food to stay hot during a long meal prefers a high thermal capacity, which keeps the pot warm for longer after the flame is switched off. A heavy cast-iron pot retains heat for an hour after cooking ends; a thin aluminium pan cools in minutes.

2. Water as a coolant

Water is used as the coolant in car engines, nuclear reactors, and power stations because its specific heat capacity of $4200 \text{ J kg}^{-1} \text{ K}^{-1}$ is the highest of any common liquid. A kilogram of water flowing through an engine can absorb large amounts of heat for only a small rise in temperature, so a relatively small mass of water at a relatively low flow rate is enough to keep the engine from overheating. No other fluid in the cook's or engineer's everyday supply comes close. Oil, used where water cannot be used because of its freezing point or its electrical conductivity, has a specific heat capacity of about $2000 \text{ J kg}^{-1} \text{ K}^{-1}$, which is less than half that of water.

3. Fish in frozen ponds

On a cold night in the highlands, the surface of a pond can freeze over while the fish beneath survive the winter. Three properties of water collaborate to make this possible. **The first** is water's large specific heat capacity, which means the pond takes a very long time to cool, and any heat stored in the lower layers is

released slowly as the surface cools. **The second** is water's anomalous expansion near freezing: water reaches its maximum density at 4°C, not at 0°C, so the densest water settles to the bottom at 4°C and colder water floats above it. **The third** is that ice, once it forms on the surface, has a much lower thermal conductivity than water (we will meet thermal conductivity in Chapter 3), so the ice layer acts as an insulating blanket between the cold air above and the 4°C water beneath. The fish live in the 4°C reservoir at the bottom of the pond, insulated from the freezing air by the ice above and from the falling temperature by water's enormous capacity to store heat.

4. Sea breeze and land breeze

A coastal town experiences a distinctive daily pattern: a cool breeze blows in from the sea during the day, and a warmer breeze blows out toward the sea at night. The cause is the difference in effective thermal capacity between sea and land. The sea absorbs the sun's heat into a layer many metres deep, mixed by currents and waves, and warms very slowly because water's specific heat capacity is large and a large mass is involved. The land absorbs heat only into its top few centimetres and warms much more quickly. By midday, the land is hotter than the sea, and the air above the hot land rises; cooler air flows in from the sea to replace it, producing the sea breeze. At night, the land cools more quickly than the sea, the situation reverses, and a land breeze blows out toward the sea. Coastal towns are therefore milder year-round than inland towns at the same latitude, because the nearby sea acts as a giant thermal reservoir that warms slowly in summer and cools slowly in winter.

Theory and you have just defined specific heat capacity, measured it two different ways, and watched it power four familiar phenomena. Time for thermal equilibrium with three worked examples to put the formulas to work.

BINDER Example 15

An electric kettle rated 2.5kW contains 1.0L of water at 25°C. Taking the specific heat capacity of water as 4200J kg⁻¹K⁻¹ and the density of water as 1000kg m⁻³, calculate the time the kettle needs to heat the water to 100°C, assuming no heat is lost to the surroundings.

Solution

The mass of water is $m = \rho V = 1000\text{kg m}^{-3} \times 1.0 \times 10^{-3}\text{m}^3 = 1.0\text{kg}$.

The heat required to raise the temperature of the water from 25°C to 100°C is

$$\Delta Q = mc\Delta T = 1.0\text{kg} \times 4200\text{J kg}^{-1}\text{K}^{-1} \times 75\text{K} = 3.15 \times 10^5\text{J}$$

The kettle supplies electrical energy at a rate of $P = 2.5\text{kW} = 2500\text{J s}^{-1}$. The time required is therefore

$$t = \frac{\Delta Q}{P} = \frac{3.15 \times 10^5\text{J}}{2500\text{J s}^{-1}} = 126\text{s} \approx 2.1\text{minutes}$$

Making Sense of the Answer: Two minutes is about how long a 2.5kW kettle actually takes to boil a litre of water in practice, which is a reassuring confirmation that the calculation is reasonable. Real kettles lose some heat to the surroundings and have their own thermal capacity to warm up, so the actual time is typically a few percent longer than our calculation suggests, but not dramatically so. The dominant physics is the heating of the water itself, and the corrections are secondary.

Think Like a Physicist: Any time a problem specifies a power and asks for a time, or specifies a time and asks for a power, the quantity linking them is the total heat $\Delta Q = mc\Delta T$. The heat is always computed first from the thermal properties of the substance being heated, and the time or power is then extracted by simple division. The calculation is exactly the same for a kettle heating water, a heater warming a room, or a rocket engine warming its exhaust: three quantities, one equation, one division.

BINDER Example 16

A copper calorimeter of mass 100g contains 200g of water at 15°C. A 150g block of iron, initially at 100°C, is dropped into the water. Assuming no heat is lost to the surroundings and taking $c_{\text{Cu}} = 400\text{J kg}^{-1}\text{K}^{-1}$, $c_{\text{Fe}} = 450\text{J kg}^{-1}\text{K}^{-1}$, and $c_{\text{water}} = 4200\text{J kg}^{-1}\text{K}^{-1}$, calculate the final equilibrium temperature.

Solution

Let T_f be the final equilibrium temperature in °C. Applying the heat balance (heat lost by iron equals heat gained by water plus heat gained by calorimeter):

$$m_{\text{Fe}}c_{\text{Fe}}(T_1 - T_f) = m_{\text{water}}c_{\text{water}}(T_f - T_2) + m_{\text{Cu}}c_{\text{Cu}}(T_f - T_2)$$

Substituting the given values in SI units:

$$(0.150)(450)(100 - T_f) = (0.200)(4200)(T_f - 15) + (0.100)(400)(T_f - 15)$$

$$67.5(100 - T_f) = 840(T_f - 15) + 40(T_f - 15)$$

$$67.5(100 - T_f) = 880(T_f - 15)$$

$$T_f = \frac{19950}{947.5} = 21.1^\circ\text{C}$$

The final equilibrium temperature is approximately 21.1°C .

Making Sense of the Answer: The iron block cools by nearly 80°C (from 100 down to 21.1) while the water warms by only 6°C (from 15 up to 21.1). The iron's temperature change is more than thirteen times larger than the water's, because water has about nine times the specific heat capacity of iron and the water mass is slightly larger than the iron mass, and nine times nine is roughly eighty. The numbers are consistent with the ratio of thermal capacities, which is reassuring.

Think Like a Physicist: A method-of-mixtures problem is always a heat-balance equation: the heat lost by the hot body equals the heat gained by everything else. Write the equation carefully before substituting numbers, listing every object that changes temperature (the hot substance, the liquid, the calorimeter, and any other solid in thermal contact). If the calorimeter is made of the same material as the hot substance, you can still include it; its term simply has a different mass and a different initial temperature. Never drop the calorimeter term without checking whether it matters; for small hot masses and large calorimeters, the calorimeter's contribution can dominate the right-hand side.

REAL Example 17

Kipanga has just come home from the market carrying two things: a new aluminium sufuria that his mother wanted for cooking ugali, and a sharp pedagogical disagreement with their neighbour Msemakweli. Mr. Akilikubwa is repairing his gate next door, and Kipanga stops to complain.

Kipanga: "Sir, my mother sent me to buy an aluminium sufuria, saying it heats up fast so she can cook ugali quickly. But Msemakweli saw me walking with it and laughed. He said his cast-iron sufuria is much better because it keeps the food warm for hours after he switches off the stove. Now my mother and Msemakweli cannot both be right, can they? One of them has to be wrong about which sufuria is superior."

Mr. Akilikubwa sets down his hammer, looks at the shiny aluminium sufuria Kipanga is holding up, and smiles.

He then clearly explained the following to Kipanga: (a) why Kipanga's mother is correct that the aluminium sufuria heats up quickly, (b) why Msemakweli is correct that the cast-iron sufuria stays warm longer, and (c) why both can be right at the same time without contradicting each other.

What was Mr. Akilikubwa's explanation?

Solution

(a) Kipanga's mother is correct because the aluminium sufuria has a low thermal capacity. An aluminium sufuria of the typical thickness sold in a Tanzanian market is both thin-walled and light, which means its mass is small. Although aluminium's specific heat capacity ($900\text{J kg}^{-1}\text{K}^{-1}$) is actually twice that of iron ($450\text{J kg}^{-1}\text{K}^{-1}$), the mass of the aluminium pot is so much smaller than the mass of an equivalent cast-iron pot that the product $C = mc$, the thermal capacity of the whole pot, is lower for the aluminium. A smaller thermal capacity means less heat is needed to warm the pot itself before the pot starts warming the food inside it, so the burner's heat goes into the food quickly and the ugali cooks fast.

(b) Msemakweli is correct because the cast-iron sufuria has a large thermal capacity. A cast-iron pot is heavy (high density) and thick-walled, so its mass is large. Even though iron's specific heat capacity is lower than aluminium's, the huge mass more than makes up for it, and the thermal capacity $C = mc$ of the cast-iron pot is much larger than that of the aluminium pot. A pot with a large thermal capacity holds a large amount of heat at a given temperature, so when the heat source is off the pot continues to release this stored heat into the food for a long time before its own temperature drops significantly. The ugali therefore stays hot longer after cooking, not because the pot is a better conductor but because it stores more heat.

(c) Both observations are correct because they are about two different physical quantities that happen to point in opposite directions for the two materials. Kipanga's mother is commenting on how fast the pot reaches cooking temperature, which is governed by the thermal capacity (a small C means fast heating). Msemakweli is commenting on how long the pot keeps its contents warm, which is also governed by the thermal capacity (a large C means slow cooling). The two statements describe opposite virtues of different pots, and both are right within their own context. Neither speaker is wrong; they simply want different things from their pots, and the market has provided them with a range of choices that trade one virtue for the other.

The physics does not force us to pick a winner between the two sufurias. Instead, it tells us what each pot is good for. A small aluminium sufuria is ideal for cooking ugali quickly in the morning when time is short. A heavy cast-iron pot is ideal for slow stews and for keeping food warm during a long meal where many guests arrive at different times. A wise cook owns both and uses each for what it does best.

Making Sense of the Answer: *The apparent paradox in Kipanga's question dissolves as soon as we distinguish between two closely related ideas. Fast heating and slow cooling are not opposites; they are two consequences of the same underlying property, the thermal capacity. A pot with low C heats fast and cools fast; a pot with high C heats slowly and cools slowly. Which is better depends entirely on what the cook needs. The disagreement between Kipanga's mother and Msemakweli is therefore not a disagreement about physics but a disagreement about kitchen strategy, and both of them are correct within their own strategies.*

Think Like a Physicist: *Whenever you hear two people disagree about which of two objects is "better" for some thermal purpose, ask what each speaker really means by "better".*

Specific heat capacity has taken us a long way. Every calculation in the examples above shared a common assumption, namely that warming a substance by one degree requires a certain quantity of heat, and that this quantity of heat scales in a simple linear way with both the mass and the temperature rise. But the assumption hides something that any cook already knows. When water at 100°C is heated further, it does not simply continue to get hotter. Instead, it boils. During the boiling, vast amounts of heat continue to flow into the water, and yet the thermometer refuses to climb above 100°C. *Where is all the heat going, if it is not raising the temperature?* The next section answers this question with a new and different quantity, the latent heat, which accounts for the energy absorbed or released whenever a substance changes from one phase of matter to another. Get the kettle back on the stove. We have one more act of thermal physics to watch unfold.

LATENT HEAT

Put a saucepan of ice cubes on a flame and watch a thermometer in the ice. At first the temperature climbs steadily as the ice warms, following exactly the specific-heat-capacity story of the previous section. At zero degrees Celsius, the climbing stops. The flame is still on. The saucepan is still getting heat. The thermometer, placed in the melting ice, reads stubbornly zero degrees for minute after minute. Only after the very last piece of ice has melted does the thermometer begin to rise again, and the water in the saucepan heats towards the boiling point. At one hundred degrees Celsius, exactly the same thing happens. The temperature stops climbing, the water boils steadily, the thermometer refuses to move above one hundred degrees, and the saucepan slowly empties into the steam above. Only when the last drop of water has boiled away does the steam, if contained, begin to rise in temperature again.

Where did the heat go during the two flat periods? The flame never stopped. The saucepan never cooled. Yet the thermometer sat still for minutes at a stretch, as if the flame had temporarily been switched off. The answer is that the heat did not disappear. It was used to change the substance from one phase of matter to another, from ice to liquid water at zero degrees and from liquid water to steam at one hundred degrees. A phase change requires a specific amount of energy per kilogram, and during the phase change that energy is absorbed without raising the temperature of the substance at all. The quantity of energy required is called the **latent heat**.

Defining latent heat

The latent heat of a phase change is defined as *the heat required, per unit mass, to change a substance from one phase to another at the phase-change temperature without changing the temperature itself.*

$$L = \frac{Q}{m} \quad \Rightarrow \quad Q = mL$$

The SI unit is J kg⁻¹. Two different latent heats appear in ordinary thermal physics:

- **Latent heat of fusion, L_f ,** *the heat required to melt a kilogram of a solid at its melting point into a kilogram of liquid at the same temperature.*
- **Latent heat of vapourisation, L_v ,** *the heat required to vapourise a kilogram of a liquid at its boiling point into a kilogram of vapour at the same temperature.*

For water, the two latent heats are:

$$L_f = 3.34 \times 10^5 \text{ J kg}^{-1} \quad L_v = 2.26 \times 10^6 \text{ J kg}^{-1}$$

The vapourisation value is about seven times larger than the fusion value. A single kilogram of water at one hundred degrees requires over two million joules of heat to boil away completely, enough energy to lift a one-tonne lorry a vertical height of about two hundred and thirty metres! The fusion of a single kilogram of

ice at zero degrees requires considerably less, but still an amount comparable to the heat needed to raise the same kilogram of water by eighty degrees.

Why L_v is so much larger than L_f

The molecular explanation of the two latent heats is simple and physical, and it explains the factor of seven cleanly. In the solid phase, water molecules are held rigidly in place by strong intermolecular forces and can vibrate only about their fixed positions. In the liquid phase, the molecules are still close to each other but are free to slide past one another; the intermolecular forces have been loosened, not broken. In the vapour phase, the molecules are widely separated, the intermolecular forces have been completely broken, and the molecules move independently in all directions.

Melting takes a kilogram of ice and turns it into a kilogram of liquid water. The intermolecular forces are only partially broken, because the molecules of the liquid are still in contact with each other most of the time. The heat required to accomplish this partial loosening is L_f .

Boiling takes a kilogram of liquid water and turns it into a kilogram of vapour. Now the intermolecular forces must be broken completely, and the molecules must be given enough kinetic energy to fly apart and occupy a volume roughly 1700 times larger than the liquid they left behind. A further contribution comes from the external work done by the expanding vapour against the atmospheric pressure pushing down on it. All of this energy must be supplied by the heat source, and the total is L_v .

The factor of seven between L_v and L_f is therefore the energy cost of the difference between loosening the intermolecular forces and breaking them entirely. If water molecules were held together only weakly, L_f and L_v would be closer in size. If they were held together very strongly, both latent heats would be larger and so would their ratio. The observed ratio of about seven is a direct fingerprint of how tightly the water molecule clings to its neighbours.

The heating curve for water

All of the ideas above are summarised in a single graph: the **heating curve** for water, showing the temperature of one kilogram of ice as heat is supplied steadily, from a starting temperature of -20°C until the water has been fully converted to steam at 120°C .

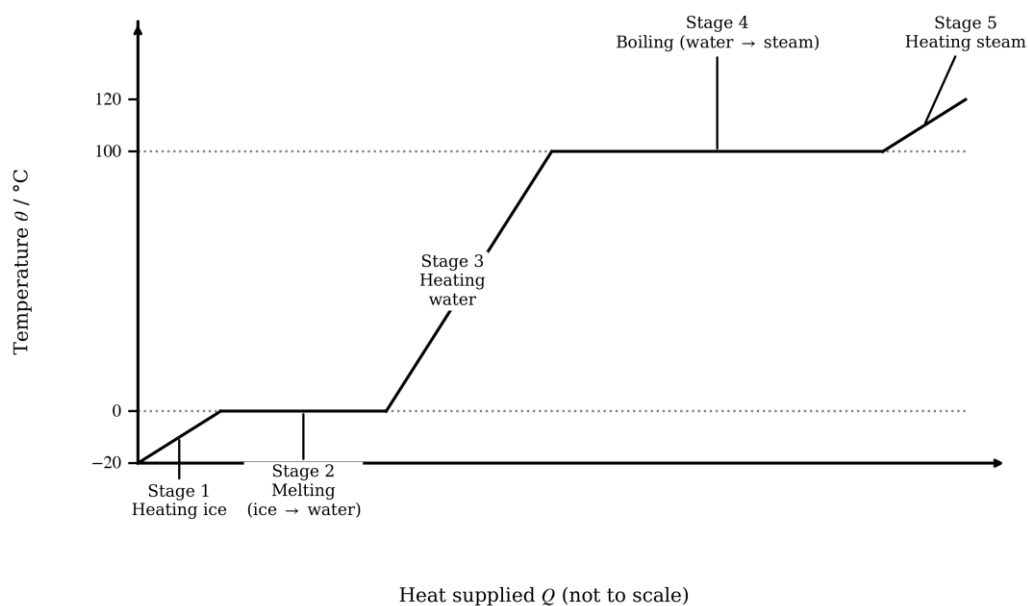


Figure: The heating curve for one kilogram of ice, warmed steadily from -20°C to 120°C . Five stages are visible. Stage 1: the ice warms from -20°C to 0°C , a sloped segment whose slope depends on the specific heat capacity of ice. Stage 2: the ice melts at 0°C , a flat segment whose width is L_f . Stage 3: the liquid water warms from 0°C to 100°C , a sloped segment whose slope depends on the specific heat capacity of liquid water. Stage 4: the water boils at 100°C , a flat segment whose width is L_v . Stage 5: the steam warms from 100°C to 120°C , a sloped segment whose slope depends on the specific heat capacity of steam. The width of Stage 4 is much larger than the width of Stage 2, because L_v is about seven times L_f .

The three sloped segments and the two flat segments each tell a different story. On a sloped segment, heat is being used to raise the temperature of a substance that is not changing phase. The slope $d\theta/dQ$ is inversely proportional to the specific heat capacity, $\frac{d\theta}{dQ} = \frac{1}{mc}$, so steeper slopes correspond to substances with smaller c . The slope of Stage 5 (heating steam) is steeper than the slope of Stage 3 (heating water) because $c_{\text{steam}} \approx 2000 \text{ J kg}^{-1} \text{ K}^{-1}$ is less than half of $c_{\text{water}} = 4200 \text{ J kg}^{-1} \text{ K}^{-1}$, so a given amount of heat raises the steam's temperature more than twice as much as it raises the water's.

On a flat segment, heat is being absorbed but the temperature is not rising. All of the incoming heat is going into breaking or loosening intermolecular bonds, a process that changes the molecular potential energy of the substance without changing the molecular kinetic energy (and temperature is, at root, a measure of average molecular kinetic energy). The width of the flat segment is equal to the total heat absorbed during the phase change, which is mL_f for melting and mL_v for boiling. The ratio of the two widths is the ratio $\frac{L_v}{L_f} \approx 6.8$, which is why Stage 4 in the figure is visibly much wider than Stage 2.

Evaporation and boiling are not the same thing

A common confusion is between evaporation and boiling. They both turn liquid water into water vapour, but they are very different processes, and the difference matters.

Evaporation occurs at the free surface of a liquid, at any temperature above its freezing point. The molecules at the surface have a distribution of kinetic energies, and a small fraction of them at any instant have enough energy to escape into the air above. Those escaping molecules carry away kinetic energy from the liquid, which cools the liquid slightly (this is the cooling effect of evaporation, familiar from sweating skin and wet laundry hanging in the sun). Evaporation is a slow process, limited to the surface, and driven by the random thermal motion of molecules at that surface. It does not require the liquid to reach the boiling temperature; a puddle on the ground at 25°C will evaporate away in a few hours.

Boiling occurs throughout the volume of a liquid, at the specific temperature where the vapour pressure of the liquid equals the atmospheric pressure above it. At this temperature, bubbles of vapour can form anywhere inside the liquid and rise to the surface; the liquid visibly churns. Boiling is a fast process, not limited to the surface, and it requires heat to be supplied from outside the liquid to keep the phase change going. Boiling temperature depends on atmospheric pressure: at sea level, water boils at 100°C ; on a mountain top where the atmospheric pressure is lower, water boils at a lower temperature; and in a pressure cooker where the pressure is higher, water boils at a temperature above 100°C .

Both processes absorb the same latent heat of vapourisation, because both convert liquid to vapour and the molecular cost per molecule is the same. The difference is entirely in where and how fast the phase change happens.

The first law of thermodynamics and phase changes

When a kilogram of water at 100°C is boiled away into steam at 100°C and standard atmospheric pressure, the heat absorbed is $Q = L_v = 2.26 \times 10^6 \text{ J}$ per kilogram. Not all of this heat ends up inside the steam as internal energy. Part of it is used by the steam to push back the atmosphere as it expands from the volume of the water (about one litre per kilogram) to the volume of the steam (about 1670 litres per kilogram). The work done by the steam against the atmospheric pressure is $W = P\Delta V$, where ΔV is the increase in volume.

The remaining energy stays inside the steam as an increase in its internal energy, ΔU . The three quantities are related by the first law of thermodynamics,

$$\Delta U = Q - W$$

which we will derive carefully in Chapter 5. For the present chapter, it is enough to know that the first law says the heat absorbed is partitioned into two contributions: some raises the internal energy of the substance, and the rest does external work against the surroundings. For a phase change from liquid to vapour at constant atmospheric pressure, W is the work done by the expanding vapour and ΔU is the part of Q that actually ends up as internal molecular energy of the steam.

Heat conduction and the ice pond

One final application of latent heat is the growth of an ice layer on the surface of a frozen pond. When the air above a pond drops below 0°C , the water at the surface freezes to form an ice layer, and the ice layer grows thicker over time as more water freezes at its lower surface. The rate at which the layer grows is

controlled by two physical processes: the rate at which heat is conducted through the ice layer from the water below to the cold air above, and the rate at which the water at the bottom of the ice layer freezes.

The rate of heat conduction through a layer of thickness x with cross-section area A , thermal conductivity k , and a temperature difference ΔT across it is given by Fourier's law:

$$\frac{dQ}{dt} = \frac{kA\Delta T}{x}$$

We will derive Fourier's law carefully in Chapter 3, but for now we need it as a working tool. The heat conducted through the ice is the heat released by the water beneath as it freezes, and if an additional thin layer of water of thickness dx freezes, the heat released is $dQ = \rho A dx L_f$, where ρ is the density of water (approximately equal to the density of ice, for this calculation).

Setting the rate of heat conducted equal to the rate at which this freezing-heat is produced:

$$\frac{kA\Delta T}{x} = \rho A L_f \frac{dx}{dt}$$

The area A cancels from both sides:

$$\frac{k\Delta T}{x} = \rho L_f \frac{dx}{dt} \Rightarrow \frac{dx}{dt} = \frac{k\Delta T}{\rho L_f x}$$

This equation tells us that the growth rate of the ice layer is inversely proportional to the layer's current thickness. When the layer is thin, the ice grows quickly; as the layer thickens, the growth slows because the same temperature difference is driving heat through a thicker insulating barrier. This is why a thin ice layer on a pond is untrustworthy (it is still growing fast and its thickness changes rapidly), while a thick layer is stable (its growth has slowed to a crawl). Worked examples later in this section will put concrete numbers into this equation.

Theory and you have just defined latent heat, explained why vapourisation requires so much more energy than fusion, traced a full heating curve, and met two important forward-references (the first law for phase changes and Fourier's law for heat conduction). Time for thermal equilibrium with five worked examples.

BINDER Example 18

A 200g sample of an unknown solid is heated steadily from -40°C through its melting point at 40°C up to a final temperature of 80°C . A plot of temperature against total heat supplied shows three distinct segments. The first segment (heating the solid) absorbs 3600J and spans the temperature range -40°C to 40°C . The second segment is a horizontal region at 40°C , absorbing 16000J without any temperature change. The third segment (heating the liquid) absorbs 4000J and spans 40°C to 80°C . Calculate (a) the specific heat capacity of the solid phase, (b) the specific latent heat of fusion, (c) the specific heat capacity of the liquid phase.

Solution

Using $m = 0.200\text{kg}$ throughout.

(a) The first segment is a sloped region in which 3600J raises the solid's temperature by:

$$\Delta T_1 = 40^\circ\text{C} - (-40^\circ\text{C}) = 80^\circ\text{C}$$

Applying $Q = mc\Delta T$ and solving for c :

$$c_{\text{solid}} = \frac{Q_1}{m \Delta T_1} = \frac{3600\text{J}}{0.200\text{kg} \times 80^\circ\text{C}} = 225\text{J kg}^{-1}\text{C}^{-1}$$

The specific heat capacity of the solid phase is $225\text{J kg}^{-1}\text{C}^{-1}$.

(b) The second segment is a horizontal region at the melting point, in which 16 000J is absorbed at constant temperature. Applying $Q = mL_f$ and solving for L_f :

$$L_f = \frac{Q_2}{m} = \frac{16000\text{J}}{0.200\text{kg}} = 8.0 \times 10^4\text{J kg}^{-1}$$

The specific latent heat of fusion is $8.0 \times 10^4\text{J kg}^{-1}$.

(c) The third segment is a sloped region in which 4000J raises the liquid's temperature by:

$$\Delta T_3 = 80^\circ\text{C} - 40^\circ\text{C} = 40^\circ\text{C}$$

Applying $Q = mc\Delta T$ again:

$$c_{\text{liquid}} = \frac{Q_3}{m \Delta T_3} = \frac{4000\text{J}}{0.200\text{kg} \times 40^\circ\text{C}^{-1}} = 500\text{J kg}^{-1}\text{C}^{-1}$$

The specific heat capacity of the liquid is $500\text{J kg}^{-1}\text{C}^{-1}$.

Making Sense of the Answer: The three numerical values are all smaller than the corresponding values for water. The solid's specific heat capacity is about one-tenth of ice's, the latent heat of fusion is about one-quarter of ice's, and the liquid's specific heat capacity is about one-eighth of water's. The substance is therefore nothing like water; it might be a metal such as lead or bismuth, for which the quoted values are roughly the right order of magnitude. A heating curve, read carefully, tells the physicist not only how a substance behaves thermally but also gives clues about what kind of substance it is.

Think Like a Physicist: A heating curve can be read in either direction. Given a substance and its thermal properties, you can sketch the curve; given a curve and its numerical segment sizes, you can recover the thermal properties. The two directions are the same physics seen from opposite sides. Whenever a problem gives you energies absorbed on specific segments of a heating curve, identify which segment corresponds to which phase or phase change, and apply either $Q = mc\Delta T$ (for sloped segments) or $Q = mL$ (for flat segments) with the appropriate value of ΔT or m .

BINDER Example 19

Calculate the total heat required to convert 500g of ice at -10°C into steam at 100°C . Take $c_{\text{ice}} = 2100\text{J kg}^{-1}\text{K}^{-1}$, $c_{\text{water}} = 4200\text{J kg}^{-1}\text{K}^{-1}$, $L_f = 3.34 \times 10^5\text{J kg}^{-1}$, $L_v = 2.26 \times 10^6\text{J kg}^{-1}$.

Solution

The process has four distinct stages, each contributing a separate amount of heat. Let $m = 0.500\text{kg}$ throughout.

Stage 1. Heating the ice from -10°C to 0°C :

$$Q_1 = m c_{\text{ice}} \Delta T = 0.500\text{kg} \times 2100\text{J kg}^{-1}\text{K}^{-1} \times 10\text{K} = 1.05 \times 10^4\text{J}$$

Stage 2. Melting the ice at 0°C into water at 0°C :

$$Q_2 = m L_f = 0.500\text{kg} \times 3.34 \times 10^5\text{J kg}^{-1} = 1.67 \times 10^5\text{J}$$

Stage 3. Heating the water from 0°C to 100°C :

$$Q_3 = m c_{\text{water}} \Delta T = 0.500\text{kg} \times 4200\text{J kg}^{-1}\text{K}^{-1} \times 100\text{K} = 2.10 \times 10^5\text{J}$$

Stage 4. Boiling the water at 100°C into steam at 100°C :

$$Q_4 = m L_v = 0.500\text{kg} \times 2.26 \times 10^6\text{J kg}^{-1} = 1.13 \times 10^6\text{J}$$

The total heat is the sum of all four contributions:

$$Q_{\text{total}} = Q_1 + Q_2 + Q_3 + Q_4$$

$$Q_{\text{total}} = (1.05 \times 10^4 + 1.67 \times 10^5 + 2.10 \times 10^5 + 1.13 \times 10^6)\text{J} = 1.52 \times 10^6\text{J} = 1.52\text{MJ}$$

Making Sense of the Answer: Three-quarters of the total energy went into Stage 4, the vapourisation of the water, even though that stage did not change the temperature by a single degree. This is a direct consequence of L_v being so much larger than every other contribution in the calculation. If you were designing a solar cooker or a gas boiler, knowing that roughly seventy-five per cent of the energy needed to go from ice to steam is spent during boiling alone is essential, because it tells you where your appliance's heat budget actually goes.

Think Like a Physicist: Whenever a multi-stage heat calculation crosses one or more phase changes, break the calculation into separate stages before attempting the arithmetic. One stage for each segment of the heating curve. The specific heat capacity applies on sloped segments, the latent heat applies on flat segments, and the total heat is the sum of the contributions from all stages. Working this way makes it almost impossible to forget a phase change or use the wrong specific heat.

REAL Example 20

In rural areas of **Tanzania** where electric fans and air conditioners are not available, a traditional remedy for a feverish child is to wipe the child's forehead with a damp cloth. The cloth has been dipped in cool water and wrung out, and it is left resting on the child's forehead until it dries out, at which point it is dipped again. Parents observe that this simple procedure brings the fever down noticeably within minutes, while simply fanning the child with a dry hand has much less effect. Using the concepts of specific heat capacity and latent heat of vapourisation, explain why the damp cloth is so effective at cooling the child, and estimate the quantity of heat removed from the forehead when 5g of water evaporates from the cloth. Take $L_v = 2.26 \times 10^6 \text{ J kg}^{-1}$ for water at body temperature and $c_{\text{water}} = 4200 \text{ J kg}^{-1} \text{ K}^{-1}$.

Solution

The cloth cools the forehead by two separate mechanisms, one of which is much more powerful than the other.

The first mechanism is straightforward *temperature equilibration*. The cool water in the cloth is initially at, say, 25°C , while the forehead is at about 38°C in a feverish child. Heat flows from the forehead to the water by conduction, warming the water and cooling the forehead. The heat that must be absorbed to raise 5g of water from 25°C to 38°C is

$$Q_1 = mc \Delta T = 0.005 \text{ kg} \times 4200 \text{ J kg}^{-1} \text{ K}^{-1} \times 13 \text{ K} = 273 \text{ J}$$

This contribution, about 273 joules, accounts for the cloth feeling cool to the touch when first applied, but it is not the dominant cooling mechanism.

The second mechanism is *evaporation*. Once the water in the cloth has warmed to body temperature, it begins to evaporate from the surface of the cloth into the air. Each molecule of water that escapes into the air carries away energy equal to the latent heat of vapourisation per molecule, and that energy is drawn from the forehead. The heat absorbed when all 5g of water in the cloth evaporates is

$$Q_2 = mL_v = 0.005 \text{ kg} \times 2.26 \times 10^6 \text{ J kg}^{-1} = 1.13 \times 10^4 \text{ J}$$

This is about 11300J, more than forty times larger than the temperature-equilibration contribution. The damp cloth is an effective cooler because evaporation of water is a hugely energy-expensive process, and every gram of water that evaporates from the cloth draws over two thousand joules of heat out of whatever surface it was in contact with.

Fanning with a dry hand, by contrast, only accelerates the loss of heat by convection. It does not trigger any evaporation from the skin unless the skin is already sweating, and even then the cooling is limited to the small amount of sweat that is actually on the skin. The damp cloth brings a fresh supply of water to the forehead and holds it in intimate contact with the skin long enough for nearly all of it to evaporate, delivering cooling equivalent to forty or fifty times what a dry fan could provide.

The total heat removed from the forehead by the damp cloth is therefore

$$Q_{\text{total}} = Q_1 + Q_2 = 273 \text{ J} + 1.13 \times 10^4 \text{ J} \approx 1.16 \times 10^4 \text{ J}$$

Making Sense of the Answer: *The dominant contribution is the evaporation term, and this is not an accident. Every cooling system that relies on water, from the sweat on a marathon runner's skin to the cooling towers of a power station, exploits the same huge latent heat of vapourisation. The reason parents have trusted the damp-cloth remedy for generations is that the physics is on their side: water's latent heat of vapourisation is one of the largest in nature, and evaporative cooling is a far more effective heat-removal mechanism than ordinary conduction or convection at the same temperatures.*

Think Like a Physicist: *Whenever a cooling process involves water and the temperatures are near body temperature, check whether evaporation is occurring. If it is, the latent heat contribution almost always dominates over the specific heat contribution, because L_v for water is numerically much larger than c_{water} multiplied by any reasonable temperature change. This is why sweating is effective at cooling the body, why wet washing dries faster on a windy day, and why spilling water on a hot stove cools the stove surface dramatically.*

HOT Example 21

A mass of 1.0g of water at 100°C is converted into steam at 100°C at the standard atmospheric pressure $P = 1.013 \times 10^5 \text{ Pa}$. The volume of 1.0g of water is 1.0 cm^3 , and the volume of 1.0g of steam at 100°C and standard pressure is 1671 cm^3 . Take $L_v = 2.260 \times 10^6 \text{ J kg}^{-1}$. Calculate:

- the heat absorbed during the phase change,
- the external work done by the expanding steam against the atmospheric pressure, and
- the increase in internal energy of the water.

Solution

- (a) The heat absorbed during the phase change is $Q = mL_v$:

$$Q = 1.0 \times 10^{-3} \text{ kg} \times 2.260 \times 10^6 \text{ J kg}^{-1} = 2260 \text{ J}$$

- (b) The external work done by the steam against the atmospheric pressure is $W = P\Delta V$, where ΔV is the change in volume from liquid water to steam:

$$\Delta V = V_{\text{steam}} - V_{\text{water}} = 1671 \text{ cm}^3 - 1.0 \text{ cm}^3 = 1670 \text{ cm}^3 = 1.670 \times 10^{-3} \text{ m}^3$$

$$W = P\Delta V = 1.013 \times 10^5 \text{ Pa} \times 1.670 \times 10^{-3} \text{ m}^3 = 169.2 \text{ J} \approx 169 \text{ J}$$

- (c) From the first law of thermodynamics, $\Delta U = Q - W$:

$$\Delta U = 2260 \text{ J} - 169 \text{ J} = 2091 \text{ J}$$

Of the 2260J absorbed, 169J is spent on pushing back the atmosphere, and the remaining 2091J goes into the internal energy of the steam.

Making Sense of the Answer: The partition of the absorbed heat is about 93% to internal energy and 7% to external work. The internal-energy contribution is the dominant one, and it represents the energy stored in the widely separated steam molecules, which have broken free from their neighbours in the liquid. The external-work contribution is smaller but not negligible, and it is the reason the latent heat of vapourisation of water at atmospheric pressure is noticeably larger than it would be if the steam were allowed to form in a vacuum (where W would be zero and ΔU would equal Q directly).

Think Like a Physicist: Whenever a phase change occurs at constant atmospheric pressure, the heat absorbed is partitioned into two parts: one part pushes back the atmosphere ($W = P\Delta V$) and one part increases the internal energy ($\Delta U = Q - W$). **For condensed phases (solid to liquid), ΔV is small and W is negligible, so $\Delta U \approx Q$.** For phase changes involving a gas (liquid to vapour), ΔV is very large and W is appreciable; it must always be computed separately and subtracted from Q to get ΔU .

HOT Example 22

On a very cold night in the Tanzanian highlands, a shallow pond develops a layer of ice on its surface. The air temperature above the ice is -10°C and the water immediately beneath the ice is at 0°C . Take the thermal conductivity of ice as $k = 2.2 \text{ W m}^{-1} \text{ K}^{-1}$, the density of ice as $\rho = 920 \text{ kg m}^{-3}$, and the latent heat of fusion as $L_f = 3.34 \times 10^5 \text{ J kg}^{-1}$. At the moment when the ice layer is 10cm thick, calculate:

- the rate of heat loss per square metre of the ice surface,
- the rate at which the thickness of the ice layer is increasing, and
- the time it would take for the ice layer to grow by an additional 1mm at this rate.

Solution

- (a) The rate of heat flow per unit area through the ice, using Fourier's law $\frac{dQ}{dt} = \frac{kA\Delta T}{x}$ and dividing both sides by A :

$$\frac{1}{A} \frac{dQ}{dt} = \frac{k\Delta T}{x} = \frac{2.2 \text{ W m}^{-1} \text{ K}^{-1} \times 10 \text{ K}}{0.10 \text{ m}} = 220 \text{ W m}^{-2}$$

The rate of heat loss is 220 W m^{-2} at this instant.

- (b) The growth rate of the ice layer comes from the relation derived earlier in this chapter,

$$\frac{dx}{dt} = \frac{k \Delta T}{\rho L_f x}$$

Substituting the numerical values:

$$\frac{dx}{dt} = \frac{2.2 \text{ W m}^{-1} \text{ K}^{-1} \times 10 \text{ K}}{920 \text{ kg m}^{-3} \times 3.34 \times 10^5 \text{ J kg}^{-1} \times 0.10 \text{ m}}$$

$$\frac{dx}{dt} = 7.16 \times 10^{-7} \text{ m s}^{-1}$$

The thickness of the ice layer is increasing at the rate of $7.16 \times 10^{-7} \text{ m s}^{-1}$.

This is about $0.72 \mu\text{m s}^{-1}$, or equivalently 2.6 mm h^{-1} , or 62 mm d^{-1} if the conditions remained constant.

(c) At the instantaneous rate computed in part (b), an additional 1mm of ice ($dx = 1 \times 10^{-3} \text{ m}$) would form in a time

$$dt = \frac{dx}{\frac{dx}{dt}} = \frac{1 \times 10^{-3} \text{ m}}{7.16 \times 10^{-7} \text{ m s}^{-1}} = 1397 \text{ s} \approx 23 \text{ min}$$

It would take 1397s or approximately 23min.

This is slightly under twenty-four minutes, assuming the ice thickness does not change appreciably during this interval (which is a reasonable assumption since 1mm is only 1% of the current 10cm layer).

Making Sense of the Answer: *The ice grows at a rate of about 2.6mm per hour at the moment when the layer is 10cm thick. Once the layer reaches 20cm, the growth rate would drop to half that value, because $dx/dt \propto 1/x$. Once the layer reaches 1m thick (a rare but physically achievable state in Arctic conditions), the growth rate would drop by a factor of ten, explaining why even the thickest Arctic ice sheets take years rather than months to form. The inverse relationship between thickness and growth rate is the physics behind why a thick ice layer is stable and safe, while a thin ice layer is rapidly changing and treacherous.*

Think Like a Physicist: *Whenever a heat-conduction problem involves a moving phase boundary (ice growing into water, steam condensing onto a cold surface, frost forming on a window), the rate of phase change is governed by the balance between two physical processes: the rate at which heat is conducted away from the phase boundary, and the amount of heat per unit mass released (or absorbed) by the phase change itself. Set the two rates equal to each other and solve for the velocity of the boundary. The same technique works for every moving-boundary heat problem you will ever meet.*

With specific heat capacity and latent heat now behind us, every thermal quantity in Chapter 2 has been defined, measured, and tested against real physical situations. The calculations in the preceding sections covered kettles, calorimeters, cooking pots, clothes drying in the sun, ponds freezing over in the highlands, and the conversion of water into steam under an atmosphere of pressure. What remains is to step back and collect the various applications of thermal physics into a single narrative: where do these ideas reappear in the world of everyday life and engineering, and what other phenomena have we not yet touched that belong in the story? The next section is our closing tour of that terrain.

APPLICATIONS OF THERMAL PHYSICS

Every idea in this chapter has a job waiting for it in the world outside the classroom. Thermometers measure temperature in hospitals, factories, and kitchens. Specific heat capacity determines the size of a cooking pot, the flow rate of coolant in an engine, and the capacity of a solar water heater. Latent heat drives refrigerators, pressure cookers, steam autoclaves, and the sweating body. This closing section takes six of the most important applications of thermal physics and examines each in turn, showing how the quantities we have built fit together to solve real engineering and biological problems.

1. The refrigerator and the air conditioner

A refrigerator is a box that keeps its interior cooler than the room around it. Since heat naturally flows from hot to cold, the refrigerator has to actively pump heat the other way, from the cold interior to the warm exterior. The physics of this pumping is powered by latent heat of vapourisation.

Inside the cooling coils that run along the back or the underside of the refrigerator interior, a special working fluid is held at a pressure at which it boils at a low temperature, typically around -25°C . The fluid evaporates

in these coils, drawing heat from the surrounding air inside the refrigerator as it does so, because evaporation requires latent heat L_v that must be supplied from somewhere. The air inside the refrigerator provides that heat, and in doing so the air cools, which is what keeps the food cold.

The vapour is then pumped by an electric compressor to the coils on the back of the refrigerator, where it is compressed to a higher pressure at which its boiling point rises above room temperature, say to 45°C . At this higher pressure, the vapour condenses back into liquid, releasing the same latent heat L_v it had absorbed inside. The heat released flows into the warmer room air, which is now cooler than the compressed vapour. The liquid then returns through an expansion valve to the interior coils, where the pressure drops, the boiling point falls back to -25°C , and the cycle repeats.

The refrigerator is therefore a clever use of the fact that a fluid's boiling point depends on pressure. By alternating the same fluid between two different pressures, the engineer arranges for the fluid to absorb heat at the low-pressure boiling point (cold side) and release it at the high-pressure boiling point (hot side). An air conditioner works identically; the only difference is that the "interior" is a room of the building and the "exterior" is the open air outside.

2. The pressure cooker

The reverse application of the same pressure-boiling-point relation is the pressure cooker. A sealed pot with a safety valve is heated on a flame. As the water inside boils, the steam produced cannot escape, and it accumulates above the water until the pressure inside the pot rises well above atmospheric. The elevated pressure raises the boiling point of the water, so the water continues to boil at a temperature higher than 100°C .

In a typical domestic pressure cooker, the internal pressure is held at about 2atm (twice atmospheric pressure) by the safety valve, and the corresponding boiling point of water is approximately 120°C . Food cooked in this hot water reaches cooking temperatures much faster than it would in an ordinary open pot at 100°C , because the rate of heat transfer into the food depends on the temperature difference between the water and the food. A cook who can run the water at 120°C instead of 100°C cuts cooking times by roughly a third for most foods, and by half for tougher meats that require long cooking.

The pressure cooker is therefore not a magic gadget that adds energy to the cooking process. It is a straightforward exploitation of the physics that raising the pressure on a boiling liquid raises its boiling point. The same flame delivers the same heat, but the water reaches a higher temperature before boiling away, and the food consequently cooks faster.

3. Solar water heaters

In Tanzania, where sunshine is abundant and electricity for water heating is expensive, solar water heaters are a common feature on the roofs of homes, hotels, and institutions. A solar water heater consists of two components: a dark absorber panel exposed to the sun, and an insulated storage tank connected to the panel by a pair of pipes. Water flows by natural convection from the tank through the absorber, warming as it passes through, and back into the tank where it remains until a tap is opened below.

Three properties of water make the solar water heater work. **The first** is water's large specific heat capacity. A tank containing 200L (and therefore 200kg) of water can absorb about 840 000J of heat for every degree of temperature rise, which is a substantial buffer against fluctuations in sunlight during the day. **The second** is water's ready availability and low cost; no more expensive working fluid would be justified for domestic use. **The third** is the fact that water flows readily and mixes well, so hot water from the absorber distributes evenly through the tank by convection.

A well-designed solar water heater in Dar es Salaam can warm 200L of water from an inlet temperature of 20°C to a delivery temperature of 50°C or more during a single sunny day, and the insulation of the storage tank keeps most of that heat until the evening when hot water is actually wanted for bathing. The physics of the system is entirely the physics we have already developed: heat absorbed by the panel is used to raise the temperature of the water, the quantity is $Q = mc\Delta T$, and the storage time is set by how well the tank is insulated.

4. Cooling systems for engines and generators

Every internal combustion engine produces more heat than useful work, and the excess heat must be removed continuously if the engine is not to overheat and seize. In almost all modern engines, the heat is removed by circulating water through passages in the engine block, then through a radiator where the heat is released to

the outside air. Water is chosen for the same reason it is chosen as a solar-heating fluid: its specific heat capacity is the highest of any common liquid, and a small mass of water flowing at a moderate rate can carry away a very large amount of heat.

In a typical car engine producing 50kW of waste heat, a coolant flow rate of about 1kg s^{-1} is sufficient to keep the engine block at a safe operating temperature, provided the water returns to the engine from the radiator at a temperature about 12K below the block temperature. The calculation is direct: $Q/t = \dot{m}c\Delta T$, which gives $\dot{m} = 50\,000\text{W}/(4200\text{J kg}^{-1}\text{K}^{-1} \times 12\text{K}) = 1.0\text{kg s}^{-1}$ of water. Were the coolant something with a smaller specific heat capacity, a much larger flow rate would be needed, and the pump and radiator would have to be correspondingly larger. Water is therefore not just a convenient coolant; it is the best practical coolant for the job, and its dominance in engine design is a direct consequence of its thermal properties.

Large diesel generators in Tanzanian factories, hospitals, and off-grid communities use the same principle, scaled up. The cooling fluid is usually water plus additives to prevent corrosion and to lower the freezing point, and the radiator is correspondingly larger than a car's, but the physics is identical.

5. Steam as a heat-transfer agent

Every kilogram of steam that condenses on a cold surface releases the latent heat of vapourisation $L_v = 2.26 \times 10^6\text{J kg}^{-1}$ directly into that surface. This is far more heat per unit mass than any non-phase-change heat-transfer process can deliver, and it is the reason steam is used as a heat-delivery fluid in a remarkable range of industrial applications.

In hospitals, autoclaves use steam at about 120°C to sterilise surgical instruments. The steam penetrates into every cavity of the instruments, condenses on their cold surfaces, and in condensing delivers the latent heat that heats the metal to sterilising temperature. The combination of high temperature and high heat-delivery rate kills bacteria and viruses far more reliably than dry heating, which cannot easily reach the interiors of complex instruments.

In food processing plants, steam is used for cooking, pasteurising, and sterilising because it can be distributed cheaply through pipes to any location where heat is needed and delivers a predictable, controllable amount of heat wherever it condenses. In industrial heating systems for factory floors and large buildings, the same principle applies: steam generated in a central boiler is piped to the areas that need heating, condenses there, and the condensate returns to the boiler to be re-evaporated. The latent heat of vapourisation effectively acts as a thermal capacity reservoir that is orders of magnitude larger than the specific-heat-capacity reservoir of any non-phase-change fluid.

6. Regulation of body temperature

The human body is a thermal system that maintains an internal temperature of about 37°C with remarkable precision, despite external conditions that can range from -20°C on a cold mountain to 40°C in the tropics. Three mechanisms from this chapter work together to make this regulation possible.

The first is the body's large thermal capacity. An adult human is roughly 60% water by mass, and the specific heat capacity of water is the largest of any common substance, so the body's thermal capacity is comparable to that of a bucket of water. This large thermal capacity buffers the body against rapid temperature swings: a sudden blast of cold or heat takes tens of seconds to change the body's core temperature measurably, giving the body's control systems time to respond.

The second is sweating, which exploits the latent heat of vapourisation of water. A person doing moderate physical work in a hot environment can sweat at a rate of about 1L h^{-1} of water, and the evaporation of each litre of that sweat absorbs about $2.26 \times 10^6\text{J}$ from the skin, cooling the body very efficiently. This is the dominant cooling mechanism at high external temperatures, and it is why dehydration is dangerous in hot climates: a body that cannot replenish its water supply loses its ability to sweat and can overheat within hours.

The third is shivering, which is a metabolic heating mechanism. When the body's core temperature drops, rapid involuntary muscle contractions are triggered, and these contractions convert metabolic energy directly into heat. Shivering can raise the body's heat production by a factor of four or five for short periods, which is usually sufficient to restore core temperature until more permanent measures (warm clothing, shelter) can be found.

The human body is therefore a living thermal system that illustrates nearly every idea in this chapter at once: its temperature is measured by thermometers, its behaviour is stabilised by its specific heat capacity, its cooling uses the latent heat of vapourisation of water, and its heating uses the conversion of metabolic chemical energy into heat.

Theory and you have travelled across six applications of thermal physics and watched every quantity in this chapter appear at least once. Time for thermal equilibrium with two worked examples before miscellaneous worked examples.

BINDER Example 23

A domestic pressure cooker operates at an internal gauge pressure of 1.0atm, so that the absolute pressure inside is $P = 2.0\text{atm} = 2.03 \times 10^5\text{Pa}$. Under these conditions, the boiling point of water rises to approximately 120°C . A cook prepares a dish that takes 60min to cook in an open pot of water boiling at 100°C . As a first approximation, assume that cooking time is roughly inversely proportional to the difference between the water's boiling temperature and room temperature (25°C). Calculate (a) the cooking time in the pressure cooker at 120°C , and (b) the percentage time saved.

Solution

(a) The temperature difference driving the cooking is $\Delta T_{\text{cook}} = T_{\text{boil}} - T_{\text{room}}$. Let t_1 and t_2 be the cooking times in the open pot and the pressure cooker respectively, and let ΔT_1 and ΔT_2 be the corresponding temperature differences.

Open pot: $\Delta T_1 = 100^\circ\text{C} - 25^\circ\text{C} = 75^\circ\text{C}$, $t_1 = 60\text{min}$.

Pressure cooker: $\Delta T_2 = 120^\circ\text{C} - 25^\circ\text{C} = 95^\circ\text{C}$.

Under the first-approximation assumption that cooking time is inversely proportional to ΔT :

$$\frac{t_2}{t_1} = \frac{\Delta T_1}{\Delta T_2}$$

$$t_2 = t_1 \times \frac{\Delta T_1}{\Delta T_2} = 60\text{min} \times \frac{75^\circ\text{C}}{95^\circ\text{C}} = 47.4\text{min}$$

The cooking time in the pressure cooker is roughly 47.4min.

(b) The time saved is $t_1 - t_2 = 60\text{min} - 47.4\text{min} = 12.6\text{min}$. As a percentage of the original cooking time:

$$\text{Percentage saved} = \frac{12.6\text{min}}{60\text{min}} \times 100\% = 21\%$$

Under the given approximation, the cook saves about 21% of the cooking time by using the pressure cooker.

Making Sense of the Answer: The approximation used here treats cooking as a simple heat-transfer process whose rate scales linearly with the temperature difference between the water and room temperature. In reality, cooking involves chemical reactions (denaturation of proteins, gelatinisation of starches, and many others) whose rates depend on temperature in more complicated ways, usually exponentially rather than linearly. The 21% figure is therefore an order-of-magnitude estimate rather than an exact prediction; real pressure cookers often save between 25% and 50% of cooking time depending on the dish, with the larger savings for dishes whose cooking relies on chemical changes that accelerate strongly with temperature.

Think Like a Physicist: When a problem gives you a proportionality assumption, check whether the assumption is physically reasonable and what its limits are. Inverse proportionality to a temperature difference is a reasonable first approximation for some processes (simple heat conduction, for example) but a poor one for others (chemical reactions, which typically follow Arrhenius's exponential law). Noting the assumption explicitly and discussing its limits is part of honest physics; silently using a questionable assumption is not.

HOT Example 24

A solar water heater in Dar es Salaam has an absorber panel of area 2.0m^2 , oriented to face the midday sun. On a sunny day, the solar power incident on the panel averages 600Wm^{-2} for 8.0hours between 09:00 and 17:00. The panel absorbs 70% of this incident energy, and the rest is lost by reflection and re-radiation. The absorbed energy is used to warm water from an inlet temperature of 20°C to a delivery temperature of 50°C

in a well-insulated storage tank. Take $c_{\text{water}} = 4200 \text{ J kg}^{-1} \text{ K}^{-1}$ and the density of water as 1000 kg m^{-3} . Calculate:

- the total solar energy incident on the panel during the eight-hour period,
- the energy actually absorbed by the panel, and
- the maximum volume of water, in litres, that can be heated from 20°C to 50°C using the absorbed energy.

Solution

(a) Total solar energy incident on the panel is the product of the incident power, the panel area, and the duration:

$$E_{\text{incident}} = (\text{power per area}) \times (\text{area}) \times (\text{time})$$

$$E_{\text{incident}} = 600 \text{ W m}^{-2} \times 2.0 \text{ m}^2 \times 8.0 \text{ h} \times 3600 \text{ s h}^{-1}$$

$$E_{\text{incident}} = 600 \text{ W m}^{-2} \times 2.0 \text{ m}^2 \times 2.88 \times 10^4 \text{ s} = 3.456 \times 10^7 \text{ J}$$

(b) The panel absorbs 70% of the incident energy:

$$E_{\text{absorbed}} = 0.70 \times E_{\text{incident}} = 0.70 \times 3.456 \times 10^7 \text{ J} = 2.42 \times 10^7 \text{ J}$$

(c) The absorbed energy heats water from 20°C to 50°C , a temperature rise of $\Delta T = 30\text{K}$.

Applying $E = mc\Delta T$ and solving for m :

$$m = \frac{E_{\text{absorbed}}}{c_{\text{water}} \Delta T} = \frac{2.42 \times 10^7 \text{ J}}{4200 \text{ J kg}^{-1} \text{ K}^{-1} \times 30 \text{ K}} = 192 \text{ kg}$$

Since water has a density of $1000 \text{ kg m}^{-3} = 1.0 \text{ kg L}^{-1}$, the volume is

$$V = \frac{m}{\rho} = \frac{192 \text{ kg}}{1.0 \text{ kg L}^{-1}} = 192 \text{ L}$$

The solar water heater can therefore warm about 192L of water from 20°C to 50°C during a single sunny day (which is enough to supply a family of four or five with hot water for bathing and washing).

Making Sense of the Answer: The 192L figure is consistent with the storage capacities of commercial solar water heaters sold in Tanzania, which commonly range from 100L for a small household to 300L for a larger family. The sizing calculation is straightforward energy balance: solar power in, absorber efficiency, heat required to warm the water, and the answer is the volume. Engineers designing a solar water heater for a particular site run this calculation with local sunshine data and local water demand to decide the panel size and the tank capacity.

Think Like a Physicist: A solar-heating sizing calculation is always a three-step operation: total energy incident, efficiency correction to find the absorbed energy, and division by $c \Delta T$ to find the mass of water that can be heated. Every step uses quantities we have already developed in this chapter. The only additional input the engineer needs is the local solar insolation, which varies with latitude, season, and weather, and which is published in engineering handbooks for every major city on the planet.

We have now reached the end of Chapter 2. Chapter 2 has handed you a thermometer, a calorimeter, a pressure cooker, and a phase diagram. Before we let you out of the kitchen, the chapter would like to see what you can do with all four at once. The miscellaneous worked examples are next.

MISCELLANEOUS WORKED EXAMPLES ON TEMPERATURE AND HEAT

Example 25

- Explain why two identical kettles filled with water at the same initial temperature and placed on identical flames take different times to reach boiling if one of them contains twice as much water as the other.
- A 10kg block slides from rest down a rough incline of length 5.0m inclined at 30° to the horizontal. The coefficient of kinetic friction between the block and the incline is 0.40. Assuming that all the heat generated by friction is absorbed by the block itself, and taking the specific heat capacity of the block as $900 \text{ J kg}^{-1} \text{ K}^{-1}$ and $g = 9.81 \text{ m s}^{-2}$, calculate (i) the heat generated by friction over the full length of the incline, (ii) the temperature rise of the block.

Solution

(a) The heat required to raise the water to boiling is $Q = mc\Delta T$. With the same initial temperature and the same final temperature (boiling point), ΔT and c are the same for both kettles, so Q increases directly with m . The kettle holding twice as much water requires twice as much heat. Since both flames deliver heat at the same rate, the larger kettle takes twice as long to reach boiling.

(b) The normal force on the block is $N = mg\cos\theta$

The friction force is $f = \mu N = \mu mg\cos\theta = 0.40 \times 10\text{kg} \times 9.81\text{ms}^{-2} \times \cos 30^\circ = 33.98\text{N}$.

(i) Heat generated by friction over the full length:

$$Q = fL = 33.98\text{N} \times 5.0\text{m} = 169.9\text{J}$$

(ii) Temperature rise of the block:

$$\Delta T = \frac{Q}{mc} = \frac{169.9\text{J}}{10\text{kg} \times 900\text{J kg}^{-1}\text{K}^{-1}} = 1.89 \times 10^{-2}\text{K}$$

The block's temperature rises by about 0.019K, a tiny change reflecting the large thermal capacity of a 10kg metal block.

Example 26

(a) Explain why a mercury-in-glass thermometer and a platinum resistance thermometer, both carefully calibrated at the ice and steam points, give slightly different readings when placed in the same water bath at 50°C , and state which of the two readings can be trusted more for precise work.

(b) A thermocouple has the Seebeck relation $E = a\theta + b\theta^2$ with $a = 40\mu\text{V}/^\circ\text{C}$ and $b = -0.040\mu\text{V}/^\circ\text{C}^2$, with the cold junction held at 0°C . The measured EMF is 9.00mV . Find the two possible hot-junction temperatures consistent with this reading, and state what additional information would resolve the ambiguity.

Solution

(a) The interpolation formula assumes that the thermometric property varies linearly with temperature between the two fixed points, but in reality the expansion of mercury and the resistance of platinum each have a slight curvature. Forced to agree at the ice point and at the steam point, the two thermometers disagree at every intermediate temperature by amounts that reflect the shape of their respective graphs. The platinum resistance reading is generally trusted more for precise scientific work because the resistance of pure platinum has a smaller departure from linearity than the volume of mercury, and because platinum thermometers can be read with an electrical bridge to far higher precision than a mercury column can be read by eye.

(b) Setting $E = 9000\mu\text{V}$ in the Seebeck equation:

$$\begin{aligned} -0.040\theta^2 + 40\theta &= 9000 \\ 0.040\theta^2 - 40\theta + 9000 &= 0 \end{aligned}$$

Solving the quadratic by calculator gives:

$$\theta_1 = 341.9^\circ\text{C} \quad \text{and} \quad \theta_2 = 658.1^\circ\text{C}$$

The ambiguity arises because the EMF-versus-temperature curve is parabolic and peaks at the neutral temperature $\theta_N = -a/(2b) = 500^\circ\text{C}$, which lies between the two roots. The same EMF corresponds to one temperature below the peak and one above. An additional measurement would resolve the ambiguity; for example, a rough check with a separate pyrometer, **or** knowledge of whether the hot junction is being heated up or cooled down (since the EMF changes monotonically on each side of the peak).

Example 27

(a) Explain why coastal cities like Dar es Salaam experience a smaller daily variation in temperature than inland cities like Dodoma, using a single thermal property of water to anchor your answer.

(b) Two thermocouples are connected in series. The first has $a_1 = 40\mu\text{V}/^\circ\text{C}$ and $b_1 = -0.030\mu\text{V}/^\circ\text{C}^2$ in the form $E = a\theta + b\theta^2$. The second has $a_2 = 60\mu\text{V}/^\circ\text{C}$ and $b_2 = -0.045\mu\text{V}/^\circ\text{C}^2$ in the same form. The cold junctions are at 0°C and both hot junctions are at $\theta = 200^\circ\text{C}$. Calculate the total EMF measured by a voltmeter connected across the series combination.

Solution

(a) The sea has a much larger thermal capacity than the surrounding land, because water's specific heat capacity ($4200 \text{ J kg}^{-1} \text{ K}^{-1}$) is several times that of sand or rock, and because the sun's heat penetrates many metres down into the sea while only a few centimetres down into the land. The sea therefore warms slowly by day and cools slowly by night, and the air above it tracks the sea's steady temperature. A coastal city, sitting next to this huge thermal reservoir, experiences moderated day-night swings. An inland city, surrounded by land that warms and cools rapidly, experiences larger swings.

(b) The EMF of each thermocouple at $\theta = 200^\circ\text{C}$:

$$E_1 = a_1\theta + b_1\theta^2 = (40\mu\text{V}/^\circ\text{C})(200^\circ\text{C}) + (-0.030\mu\text{V}/^\circ\text{C}^2)(200^\circ\text{C})^2 = 6800\mu\text{V}$$

$$E_2 = a_2\theta + b_2\theta^2 = (60\mu\text{V}/^\circ\text{C})(200^\circ\text{C}) + (-0.045\mu\text{V}/^\circ\text{C}^2)(200^\circ\text{C})^2 = 10200\mu\text{V}$$

The total EMF in series is $E_{\text{total}} = E_1 + E_2$:

$$E_{\text{total}} = 6800\mu\text{V} + 10200\mu\text{V} = 17000\mu\text{V} = 17.0\text{mV}$$

Example 28

(a) A nurse dips a clinical thermometer in a cold sterilising bath before placing it in a patient's mouth. Explain what physical processes occur between the thermometer and the mouth during the first thirty seconds after the thermometer is inserted, and explain why the nurse must wait before taking the reading.

(b) A laboratory requires thermometers for the following four measurements. For each task, state which type of thermometer (mercury-in-glass, platinum resistance, thermocouple, or optical pyrometer) is most suitable, and give one physical reason supporting your choice: (i) rapidly changing temperatures of exhaust gas leaving a car engine, (ii) the temperature inside a blast furnace at 1600°C , (iii) the precise temperature of a calibration bath at 50°C required to the nearest 0.01°C , (iv) a child's body temperature at a clinic.

Solution

(a) The thermometer is initially at the temperature of the sterilising bath, which is lower than the mouth's body temperature. When placed in the mouth, heat flows from the mouth (hotter) to the thermometer (colder) by conduction. The thermometer's mercury absorbs this heat and expands, and the mercury column rises steadily up the capillary. The mouth loses a tiny amount of heat to the thermometer in the process, but because the thermometer's thermal capacity is very small compared with the mouth's, the mouth's temperature barely changes. The nurse waits because the thermometer has not yet reached thermal equilibrium with the mouth; a premature reading would give a value below the true body temperature. Only after thermal equilibrium has been reached, when heat flow ceases, does the thermometer's temperature equal the mouth's temperature, and the reading is valid.

(b)

(i) **Thermocouple**, because its small junction has a negligible thermal mass and responds in milliseconds to temperature changes.

(ii) **Optical pyrometer**, because no metal-based thermometer can survive temperatures above about 1500°C , and a pyrometer reads temperature from radiation emitted by the hot body without making physical contact.

(iii) **Platinum resistance thermometer**, because it provides the highest precision of any standard laboratory thermometer at moderate temperatures, with resolution of 0.01°C readily achievable using a Wheatstone bridge.

(iv) **Mercury-in-glass clinical thermometer**, because the range of 35°C to 42°C is ideal for body temperature and the small mercury bulb delivers adequate precision without disturbing the patient.

Example 29

(a) Explain why the thermodynamic temperature scale does not depend on the thermometric substance used, whereas the Celsius scale defined in the nineteenth century does depend on the substance, and explain why the modern kelvin is defined using a single fixed point rather than two.

(b) A mercury thermometer is faulty: its " 0°C " mark has been engraved where the mercury actually sits at 5°C , and its " 100°C " mark has been engraved where the mercury actually sits at 102°C . The thermometer

is placed in a beaker of water being heated by a 200W heater, and reads 75°C after the heater has been on for some time. Calculate the true temperature of the water at the moment of reading.

Solution

(a) The thermodynamic scale is built on the ideal-gas relation $P \propto T$ at constant volume, a physical law that every gas obeys exactly in the low-density limit and approximately at ordinary densities. Since the relation is the same for every gas, the scale derived from it is the same regardless of which gas is used. The empirical Celsius scale, by contrast, is built on linear interpolation between two fixed points using the thermometric property of a particular substance (mercury expansion, platinum resistance, or gas pressure), and each of these properties has its own slight non-linearity with temperature. Different substances give slightly different scales between the fixed points.

One fixed point is enough for the thermodynamic scale because the ideal-gas law already fixes the shape of the P-T relation as a straight line through the origin. Only the scale of T remains to be set, and this requires a single reference point. The triple point of water ($T_{tr} = 273.16\text{K}$) is used because it is more reproducible than any other accessible physical state.

(b) Applying the faulty-thermometer interpolation formula with $\theta_f = 75^\circ\text{C}$, $\theta_{of} = 5^\circ\text{C}$, and $\theta_{100f} = 102^\circ\text{C}$:

$$\theta = \left(\frac{\theta_f - \theta_{of}}{\theta_{100f} - \theta_{of}} \right) \times 100^\circ\text{C}$$

$$\theta = \left(\frac{75^\circ\text{C} - 5^\circ\text{C}}{102^\circ\text{C} - 5^\circ\text{C}} \right) \times 100^\circ\text{C} = \left(\frac{70^\circ\text{C}}{97^\circ\text{C}} \right) \times 100^\circ\text{C} = 72.16^\circ\text{C}$$

The true temperature of the water is approximately 72.2°C.

Example 30

(a) A drop of water falls on a hot iron at 200°C and is observed to dance across the surface for several seconds before evaporating. Explain in terms of latent heat why evaporation takes so long, and explain why a drop of alcohol would disappear almost instantly under the same conditions.

(b) A laboratory has three thermometers all calibrated at the ice and steam points. A mercury-in-glass thermometer measures column lengths $\ell_0 = 25\text{cm}$, $\ell_{100} = 45\text{cm}$, and $\ell_\theta = 30\text{cm}$ in a room. A platinum resistance thermometer in the same room measures $R_0 = 60\Omega$, $R_{100} = 75\Omega$, $R_\theta = 63.7\Omega$. A thermocouple measures $E_0 = 0\text{mV}$, $E_{100} = 5.5\text{mV}$, $E_\theta = 1.50\text{mV}$. Calculate the room temperature on each of the three scales, and state which value the laboratory should record as the official room temperature, giving a brief reason.

Solution

(a) Even at 200°C the hot iron delivers heat at a finite rate. For each gram of water to evaporate, the iron must supply the latent heat of vapourisation L_v of water, which is a very large quantity of energy per unit mass. The time for the drop to evaporate is proportional to the energy required divided by the rate at which heat is delivered. Alcohol has a substantially smaller latent heat of vapourisation than water, so an alcohol drop of similar size needs much less energy to evaporate completely. With the same hot iron delivering heat at the same rate, the alcohol drop disappears in a much shorter time than the water drop, which is close to instant under the same conditions.

(b) Applying the interpolation formula to each thermometer.

Mercury scale:

$$\theta_{\text{Hg}} = \left(\frac{\ell_\theta - \ell_0}{\ell_{100} - \ell_0} \right) \times 100^\circ\text{C} = \left(\frac{30\text{cm} - 25\text{cm}}{45\text{cm} - 25\text{cm}} \right) \times 100^\circ\text{C} = 25.00^\circ\text{C}$$

Platinum scale:

$$\theta_R = \left(\frac{R_\theta - R_0}{R_{100} - R_0} \right) \times 100^\circ\text{C} = \left(\frac{63.7\Omega - 60.0\Omega}{75.0\Omega - 60.0\Omega} \right) \times 100^\circ\text{C} = 24.67^\circ\text{C}$$

Thermocouple scale:

$$\theta_{\text{tc}} = \left(\frac{E_\theta - E_0}{E_{100} - E_0} \right) \times 100^\circ\text{C} = \left(\frac{1.50\text{mV} - 0\text{mV}}{5.5\text{mV} - 0\text{mV}} \right) \times 100^\circ\text{C} = 27.27^\circ\text{C}$$

The laboratory should record the platinum resistance thermometer's reading of 24.67°C , because among the three instruments the platinum thermometer is the most accurate and its reading is closest to the true thermodynamic temperature.

Example 31

(a) Explain why a thin aluminium *sufuria* heats up quickly over a flame while a heavy cast-iron pot takes much longer, even though aluminium has a higher specific heat capacity than iron.

(b) In a method-of-mixtures experiment, a copper calorimeter of mass 80g contains 200g of water. The water temperature is monitored by a thermocouple with Seebeck equation $E = a\theta + \frac{1}{2}b\theta^2$ where $a = 50\mu\text{V}/^{\circ}\text{C}$ and $b = -0.040\mu\text{V}/^{\circ}\text{C}^2$, cold junction at 0°C . The thermocouple initially reads $E_1 = 1.20\text{mV}$. A 200g block of an unknown metal is heated to 120°C and dropped into the calorimeter. At thermal equilibrium the thermocouple reads $E_2 = 1.60\text{mV}$. Take $c_{\text{water}} = 4200\text{J kg}^{-1}\text{K}^{-1}$, $c_{\text{Cu}} = 400\text{J kg}^{-1}\text{K}^{-1}$. Calculate the specific heat capacity of the unknown metal.

Solution

(a) The time a pot takes to heat up depends on its thermal capacity $C = mc$, not on its specific heat capacity alone. Aluminium has a smaller density than iron, so for a *sufuria* of a given size the aluminium version has less mass than the iron version. Although aluminium's specific heat capacity ($900\text{J kg}^{-1}\text{K}^{-1}$) is twice that of iron ($450\text{J kg}^{-1}\text{K}^{-1}$), the mass difference more than compensates, so the product mc is smaller for the aluminium *sufuria*. A smaller thermal capacity means less heat is needed per degree of temperature rise, so the aluminium *sufuria* reaches cooking temperature faster.

(b) Converting the two thermocouple readings to temperatures. Setting $E = a\theta + \frac{1}{2}b\theta^2$ and solving for θ .

For $E_1 = 1200\mu\text{V}$:

$$50\theta + \frac{1}{2}(-0.040)\theta^2 = 1200$$

$$0.020\theta^2 - 50\theta + 1200 = 0$$

$$\theta = 24.23^{\circ}\text{C} \quad (2475.77 \text{ is not practical})$$

Since the calorimeter was at laboratory temperature, the smaller root is practically valid in this case.

The initial temperature of calorimeter, $\theta_1 = 24.23^{\circ}\text{C}$.

For $E_2 = 1600\mu\text{V}$:

$$0.020\theta^2 - 50\theta + 1600 = 0$$

$$\theta = 34.42^{\circ}\text{C} \quad (2467.58 \text{ is not practical})$$

Since the metal block started at 120°C and is cooling as it loses heat to the calorimeter, the equilibrium temperature θ must be less than 120°C . So the larger root is a mathematical artefact of the quadratic and has no practical meaning.

The final temperature of calorimeter, $\theta_2 = 34.42^{\circ}\text{C}$.

Applying the heat balance. Heat lost by the metal block = heat gained by water + heat gained by calorimeter:

$$m_M c_M (T_{\text{block}} - \theta_2) = (m_w c_w + m_{\text{cal}} c_{\text{Cu}})(\theta_2 - \theta_1)$$

$$c_M = \frac{(m_w c_w + m_{\text{cal}} c_{\text{Cu}})(\theta_2 - \theta_1)}{m_M (T_{\text{block}} - \theta_2)}$$

$$c_M = \frac{(0.200\text{kg} \times 4200\text{J kg}^{-1}\text{K}^{-1} + 0.080\text{kg} \times 400\text{J kg}^{-1}\text{K}^{-1})(34.42^{\circ}\text{C} - 24.23^{\circ}\text{C})}{0.200\text{kg} \times (120^{\circ}\text{C} - 34.42^{\circ}\text{C})} = 407.8\text{J kg}^{-1}\text{K}^{-1}$$

Specific heat capacity of the metal is $407.8\text{J kg}^{-1}\text{K}^{-1}$

Example 32

(a) Explain why a thermometer placed in a beaker of boiling water at standard atmospheric pressure reads 100°C continuously for several minutes even though the flame is still supplying heat, and state what happens to the heat during this period.

(b) **Kipute** determines the specific latent heat of fusion of ice by the method of mixtures, using a faulty mercury thermometer that reads 2°C at the true ice point and 98°C at the true steam point. She places 300g

of water in a copper calorimeter of mass 100g; the thermometer initially reads 24°C. She then adds 20g of ice at the true ice point and stirs until all the ice has melted. The final thermometer reading is 18°C. Taking $c_{\text{water}} = 4200 \text{ J kg}^{-1} \text{ K}^{-1}$ and $c_{\text{Cu}} = 400 \text{ J kg}^{-1} \text{ K}^{-1}$, calculate the **correct** value the specific latent heat of fusion of ice from Kipute's data.

Solution

(a) When water is at its boiling point, all the heat supplied by the flame is used to change the phase of the water from liquid to vapour rather than to raise its temperature. During the boiling, the molecules acquire the latent heat of vapourisation L_v , which is used to break the intermolecular forces that hold them in the liquid phase and do work against the atmospheric pressure. The heat therefore appears as an increase in molecular potential energy (and a small amount of external work), not as an increase in molecular kinetic energy. Since temperature is a measure of the average molecular kinetic energy, it stays constant until the last drop of water has been converted to steam.

(b) Converting the thermometer readings to true temperatures using the faulty-thermometer formula.

Initial reading 24°C:

$$T_1 = \left(\frac{24^\circ\text{C} - 2^\circ\text{C}}{98^\circ\text{C} - 2^\circ\text{C}} \right) \times 100^\circ\text{C} = 22.92^\circ\text{C}$$

Final reading 18°C:

$$T_2 = \left(\frac{18^\circ\text{C} - 2^\circ\text{C}}{98^\circ\text{C} - 2^\circ\text{C}} \right) \times 100^\circ\text{C} = 16.67^\circ\text{C}$$

Applying the heat balance. Heat lost by the water and calorimeter (cooling from T_1 to T_2) equals heat gained in melting the ice at 0°C and then warming the resulting water from 0°C to T_2 :

$$(m_w c_w + m_{\text{cal}} c_{\text{Cu}})(T_1 - T_2) = m_i L_f + m_i c_w T_2$$

$$L_f = \frac{(m_w c_w + m_{\text{cal}} c_{\text{Cu}})(T_1 - T_2) - m_i c_w T_2}{m_i}$$

Where:

$$(m_w c_w + m_{\text{cal}} c_{\text{Cu}})(T_1 - T_2) = (0.300 \text{ kg} \times 4200 \text{ J kg}^{-1} \text{ K}^{-1} + 0.100 \text{ kg} \times 400 \text{ J kg}^{-1} \text{ K}^{-1})(22.92^\circ\text{C} - 16.67^\circ\text{C}) = 8125 \text{ J}$$

$$m_i c_w T_2 = 0.020 \text{ kg} \times 4200 \text{ J kg}^{-1} \text{ K}^{-1} \times 16.67^\circ\text{C} = 1400 \text{ J}$$

Substituting the numerical values:

$$L_f = \frac{8125 \text{ J} - 1400 \text{ J}}{0.020 \text{ kg}} = 3.36 \times 10^5 \text{ J kg}^{-1}$$

The correct value of is $3.36 \times 10^5 \text{ J kg}^{-1}$.

Example 33

(a) Explain why a wet finger held in the breeze feels cooler than a dry finger, using the concept of latent heat of vapourisation.

(b) A mercury-in-glass thermometer has a mercury bulb of mass 20g (treated as pure mercury, $c_{\text{Hg}} = 140 \text{ J kg}^{-1} \text{ K}^{-1}$) and is at room temperature 25°C. It is dropped into an insulated beaker containing 100g of water at 80°C. Taking $c_{\text{water}} = 4200 \text{ J kg}^{-1} \text{ K}^{-1}$ and ignoring the heat capacity of the glass, calculate the final equilibrium temperature. Show what this implies about the design of precision thermometers.

Solution

(a) Water on the finger evaporates into the breeze. Each kilogram of water that evaporates draws the latent heat of vapourisation ($L_v = 2.26 \times 10^6 \text{ J kg}^{-1}$) from the finger, and this heat loss is what produces the cooling sensation. A dry finger loses heat only by conduction to the slightly cooler air and by a small amount of radiation, neither of which is as efficient as the evaporative process. The breeze increases the rate of evaporation by carrying away the saturated air near the finger, exposing the water to fresh dry air, which speeds up the process and intensifies the cooling.

(b) Applying the heat balance. Heat lost by the water equals heat gained by the thermometer:

$$m_w c_w (T_{\text{water},i} - T_{\text{eq}}) = m_t c_t (T_{\text{eq}} - T_{\text{therm},i})$$

Substituting values with $T_{\text{water},i} = 80^\circ\text{C}$ and $T_{\text{therm},i} = 25^\circ\text{C}$:

$$0.100\text{kg} \times 4200\text{J kg}^{-1}\text{K}^{-1} \times (80^\circ\text{C} - T_{\text{eq}}) = 0.020\text{kg} \times 140\text{J kg}^{-1}\text{K}^{-1} \times (T_{\text{eq}} - 25^\circ\text{C})$$

Solving for T_{eq} gives:

$$T_{\text{eq}} = 79.64^\circ\text{C}$$

The water has dropped from 80.00°C to 79.64°C , a change of 0.36°C . The thermometer, in reading this final temperature, has itself disturbed the water it was supposed to measure. This is why precision thermometers are designed with very small bulbs: the smaller the mass (and hence thermal capacity) of the thermometer's working substance, the less it disturbs the body being measured, and the closer its equilibrium reading comes to the body's original temperature. A clinical thermometer with a bulb of only 0.2g instead of 20g would disturb the water by only a hundredth as much.

Example 34

(a) Explain why a pressure cooker cooks food faster than an open pot, using your knowledge of the pressure-dependence of the boiling point of water.

(b) The specific latent heat of vapourisation of water is determined by an experiment in which steam at 100°C is passed into a calorimeter containing water. A copper calorimeter of mass 80g contains 300g of water, and the water temperature is monitored by a platinum resistance thermometer with $R_0 = 50.0\Omega$ at the ice point and $R_{100} = 70.0\Omega$ at the steam point. The platinum thermometer initially reads $R_1 = 53.0\Omega$. Steam is then passed in until 20g of it has condensed, at which point the platinum reads $R_2 = 60.0\Omega$. Taking $c_{\text{water}} = 4200\text{J kg}^{-1}\text{K}^{-1}$ and $c_{\text{Cu}} = 400\text{J kg}^{-1}\text{K}^{-1}$, calculate the experimental value of L_v obtained from the data.

Solution

(a) The boiling point of water rises with the pressure applied to its surface. In a pressure cooker the steam produced by boiling cannot escape, so it accumulates until the internal pressure has risen well above atmospheric, typically to about 2atm. At this elevated pressure, water boils at a temperature approximately 120°C , twenty degrees higher than in an open pot. The food inside the cooker is therefore surrounded by hot water at 120°C instead of 100°C , and the higher temperature speeds up the chemical reactions that constitute cooking. The cooker does not add energy to the food; it allows the water to reach a higher cooking temperature before boiling away.

(b) Convert the platinum readings to temperatures:

$$T_1 = \left(\frac{R_1 - R_0}{R_{100} - R_0} \right) \times 100^\circ\text{C} = \left(\frac{53.0\Omega - 50.0\Omega}{70.0\Omega - 50.0\Omega} \right) \times 100^\circ\text{C} = 15.0^\circ\text{C}$$

$$T_2 = \left(\frac{R_2 - R_0}{R_{100} - R_0} \right) \times 100^\circ\text{C} = \left(\frac{60.0\Omega - 50.0\Omega}{70.0\Omega - 50.0\Omega} \right) \times 100^\circ\text{C} = 50.0^\circ\text{C}$$

Applying the heat balance. Heat released by steam (condensing at 100°C and then cooling as liquid water from 100°C to T_2) equals heat absorbed by the original water plus the calorimeter (warming from T_1 to T_2):

$$m_s L_v + m_s c_w (100^\circ\text{C} - T_2) = (m_w c_w + m_{\text{cal}} c_{\text{Cu}}) (T_2 - T_1)$$

Substituting the numerical values:

$$\begin{aligned} &0.020\text{kg} \times L_v + 0.020\text{kg} \times 4200\text{J kg}^{-1}\text{K}^{-1} \times (100^\circ\text{C} - 50.0^\circ\text{C}) \\ &= (0.300\text{kg} \times 4200\text{J kg}^{-1}\text{K}^{-1} + 0.080\text{kg} \times 400\text{J kg}^{-1}\text{K}^{-1}) (50.0^\circ\text{C} - 15.0^\circ\text{C}) \end{aligned}$$

Solving for L_v gives:

$$L_v = 2.051 \times 10^6\text{J kg}^{-1}$$

Ten examples down. You have cooked food in pressure cookers, dropped iron into ice water, read three thermometers in the same room, and determined the latent heat of ice with a faulty thermometer that the physics quietly sneaked past. The heavy lifting is complete. What remains is Digging Deeper, where the questions keep coming but the physics stops teaching and starts testing.

DIGGING DEEPER EXERCISE 2

EXERCISE 2A: BINDER QUESTIONS

Question 1

Kipanga says: “If body *A* is in thermal equilibrium with body *B*, and body *B* is in thermal equilibrium with body *C*, then *A* is obviously in thermal equilibrium with *C*. This is just common sense. I don’t see why we need to call it the Zeroth Law.”

Explain what the Zeroth Law guarantees that is not obvious from everyday reasoning, and state why this guarantee is needed before any thermometer can be trusted.

Question 2

A student claims: “A bucket of water at 20°C contains more heat than a cup of water at 80°C, because the bucket has more molecules.” Identify and correct the mistake in this statement.

Question 3

- (a) Distinguish between a **thermometric property** and an ordinary physical property of a substance.
- (b) Give two examples of physical properties that cannot be used as thermometric properties, and state for each why not.

Question 4

The modern kelvin is defined using a single fixed point (the triple point of water) rather than the two fixed points (ice point and steam point) used by the nineteenth-century Celsius scale. State two distinct disadvantages of the ice point that make it unsuitable as the primary reference in the modern definition.

Question 5

A student dips a thermometer in a bath and waits for the reading to stabilise before recording it.

- (a) State what physical condition has been reached when the reading stops changing.
- (b) Explain why this condition must be reached before the reading can be trusted.

Question 6

- (a) State the property of water, together with its numerical value, that makes water the preferred coolant in car engines and power stations.
- (b) Oil has a specific heat capacity of about $2000 \text{ J kg}^{-1} \text{ K}^{-1}$. Explain why oil is not used as a coolant when water is available.

Question 7

Explain, using molecular reasoning, why heat is absorbed when ice melts but released when water freezes, even though both processes occur at exactly 0°C.

Question 8

- (a) State three distinct differences between evaporation and boiling.
- (b) State one thing that the two processes have in common.

Question 9

Explain why precision thermometers are designed with very small bulbs, and state what would happen to the accuracy of a measurement if the bulb’s thermal capacity were comparable to that of the body being measured.

EXERCISE 2B: REAL QUESTIONS**Question 10**

Kipanga and Kipute are arguing on a hot daladala ride from Chalinze to Dar es Salaam. **Kipanga**, sweating heavily, says: “*It is hotter today than yesterday. I know this because I am sweating more than I did yesterday.*” **Kipute** replies: “*Sweating tells you about your body, not about the air. Yesterday I was eating ugali that raised my body heat, and I sweated just as much as you are sweating now.*”

- (a) State which speaker is using temperature correctly as a physical quantity, and why the other is not.
- (b) Explain what a thermometer measures that makes it more reliable than human sweating as a temperature indicator.

Question 11

Nurse Maria at a Bagamoyo clinic is measuring the temperatures of ten children queued for vaccination. She dips her clinical thermometer in cool antiseptic between patients, removes it, shakes it hard to reset the mercury column, and immediately places it in the next child’s mouth. A doctor observing her procedure says the readings will be systematically too low.

- (a) State two physical reasons why the nurse’s procedure produces systematically low readings.
- (b) State two changes she could make that would give more reliable readings.

Question 12

A family in Kibaha is deciding between a charcoal stove and a gas cooker for daily cooking. The mother notes that a pot of beans takes about 90min on charcoal and 45min on gas. The father argues: “*The gas flame must be twice as hot as the charcoal flame.*” Explain why the father’s reasoning is wrong, and state the physical property of the two cookers that actually differs.

Question 13

Mama Juma lives in the cool highlands near Mufindi and uses a hot water bottle to keep her bed warm at night. A salesperson offers her a new bottle of the same size filled with hot cooking oil instead of hot water, claiming “oil stays hot longer”. Taking $c_{\text{water}} = 4200 \text{ J kg}^{-1} \text{ K}^{-1}$ and $c_{\text{oil}} = 2000 \text{ J kg}^{-1} \text{ K}^{-1}$, and assuming both bottles cool at the same rate of heat loss to the surroundings, advise Mama Juma which bottle will keep her warmer for longer, and explain the physics behind your advice.

Question 14

A long-distance bus driver notices his engine is overheating on a climb to Kitonga. He considers replacing the water in the cooling system with used engine oil, which he has available cheaply.

- (a) Explain why this would be a bad idea, using specific heat capacity as the centrepiece of your answer.
- (b) State one property other than specific heat capacity that matters for an engine coolant.

Question 15

A Form Five physics student is performing a calorimetry experiment and notices that three thermometers in the laboratory give slightly different readings when placed in the same beaker of warm water: a mercury thermometer reads 48.0°C , a platinum resistance thermometer reads 48.5°C , and a thermocouple reads 47.2°C . The lab technician tells her to “*just use whichever one you have calibrated most recently.*”

- (a) Explain whether the technician’s advice is physically sensible.
- (b) State which of the three readings the student should record as the best estimate of the true temperature, and justify your choice.

Question 16

A sambusa vendor in Kariakoo fries her sambusas in a large pot of hot oil. On a quiet day she fries five sambusas at a time. On a busy day she drops twenty sambusas into the same pot at once, and she observes that the oil temperature drops noticeably after adding the sambusas, with the first batch taking longer to cook than usual. Explain this observation using the concept of thermal capacity, and recommend one practical change she could make to avoid the problem.

Question 17

Kipanga wants to cool a bottle of soda quickly. He has two options: place the bottle in a basin of iced water at 0°C , or place it in a basin of water that has been cooled to 4°C and contains no ice. Kipute claims the iced water will cool the soda faster even though the water in both basins is already very cold. Explain why Kipute is right.

Question 18

A cook in a hurry leaves the lid off her pot of boiling stew, saying: *“The flame is the same either way, so the time to boil will be the same whether or not I use the lid.”* Explain why the cook is wrong.

Question 19

A domestic pressure cooker is designed with a safety valve that releases steam when the internal pressure exceeds a set limit. A student observes that the valve stays silent for the first ten minutes after the cooker is placed on a strong flame, then begins to hiss and continues to hiss until the flame is switched off.

- Explain why the valve stays silent for the first ten minutes and then begins to hiss.
- Explain what the continued hissing tells us about the rate of heat input compared with the rate of steam loss.
- State whether the cooker’s internal temperature drops if the user reduces the flame to a lower setting after the valve has begun to hiss, provided the valve continues to hiss.

EXERCISE 2C: HOT QUESTIONS**Question 20**

Water falls from a 200m-high waterfall in the Southern Highlands. Assuming all gravitational potential energy is converted to heat absorbed by the water itself, calculate the temperature rise of the water at the bottom of the fall. Take $g = 9.81\text{m s}^{-2}$ and $c_{\text{water}} = 4200\text{J kg}^{-1}\text{K}^{-1}$. State two physical reasons why the actual temperature rise in a real waterfall is smaller than this calculation predicts.

Question 21

A 500g iron block at 200°C is dropped into a copper calorimeter of mass 100g that contains 300g of water and 50g of ice, all initially at 0°C . Take $c_{\text{Fe}} = 450\text{J kg}^{-1}\text{K}^{-1}$, $c_{\text{Cu}} = 400\text{J kg}^{-1}\text{K}^{-1}$, $c_{\text{water}} = 4200\text{J kg}^{-1}\text{K}^{-1}$, and $L_f = 3.34 \times 10^5\text{J kg}^{-1}$. Calculate the final equilibrium temperature of the system.

Question 22

A copper block of mass 800g at 250°C is dropped into an open copper calorimeter of mass 100g that contains 150g of water at 100°C . Take $c_{\text{Cu}} = 400\text{J kg}^{-1}\text{K}^{-1}$, $c_{\text{water}} = 4200\text{J kg}^{-1}\text{K}^{-1}$, and $L_v = 2.26 \times 10^6\text{J kg}^{-1}$. Calculate the mass of water that boils away before the system reaches equilibrium, and state the final temperature of the contents.

Question 23

A thermocouple is calibrated by placing its hot junction in melting ice (0°C , $\text{EMF} = 0$), boiling water at standard atmospheric pressure (100°C , $\text{EMF} = 3.800\text{mV}$), and steam at higher pressure (120°C , $\text{EMF} = 4.512\text{mV}$). Assume the Seebeck relation has the form $E = a\theta + 1/2 b\theta^2$ with the cold junction at 0°C . Determine the constants a and b , and predict the EMF when the hot junction is at 50°C .

Question 24

A shallow pond in the Southern Highlands has an ice layer on its surface. The air above the ice is at -8°C and the water beneath the ice is at 0°C . Take the thermal conductivity of ice as $k = 2.2\text{W m}^{-1}\text{K}^{-1}$, the density of ice as $\rho = 920\text{kg m}^{-3}$, and the latent heat of fusion as $L_f = 3.34 \times 10^5\text{J kg}^{-1}$. Calculate the time required for the ice layer to grow in thickness from 6.0cm to 7.0cm, assuming the air and water temperatures remain constant during this period.

Question 25

A solar water heater in Iringa has an absorber panel of area 2.5m^2 and heats a tank containing 180L of water. On a partly cloudy day, the solar power incident on the panel averages 350W m^{-2} for 9.0hours, and the panel absorbs 65% of the incident energy. The water in the tank starts at 22°C . Take $c_{\text{water}} =$

$4200\text{J kg}^{-1}\text{K}^{-1}$ and the density of water as 1000kg m^{-3} . Calculate the final temperature of the water in the tank.

Question 26

A platinum resistance thermometer is supposed to give $R_0 = 50.00\Omega$ at the ice point and $R_{100} = 65.00\Omega$ at the steam point, but because of calibration errors it actually gives $R_0 = 50.50\Omega$ at the true ice point and $R_{100} = 64.80\Omega$ at the true steam point. The same thermometer is placed in a bath and reads 58.00Ω .

- Calculate the temperature the student would report if she assumed the thermometer was properly calibrated.
- Calculate the true temperature of the bath.
- Express the difference between (a) and (b) as a percentage of the true temperature.

Question 27

Steam at 100°C is passed into an insulated tank containing 10.0L of water at 20°C , and the steam condenses fully as it enters. Taking $L_v = 2.26 \times 10^6\text{J kg}^{-1}$, $c_{\text{water}} = 4200\text{J kg}^{-1}\text{K}^{-1}$, and the density of water as 1000kg m^{-3} , calculate the mass of steam that must condense to bring the water in the tank to 60°C .

Question 28

A pressure cooker holds 2.0L of water that is heated from room temperature 25°C . In an open pot, the water needs only to reach 100°C to begin cooking; in the pressure cooker, the operating temperature is 120°C . Take $c_{\text{water}} = 4200\text{J kg}^{-1}\text{K}^{-1}$ and the density of water as 1000kg m^{-3} . The flame supplies heat at a steady rate of 2.0kW .

- Calculate the additional heat required to bring the cooker to its operating temperature, compared with simply reaching ordinary boiling.
- Calculate the additional time the flame must operate to deliver this additional heat.

Question 29

A wet-and-dry bulb hygrometer has two identical mercury thermometers placed side by side. The dry bulb reads air temperature 30°C ; the wet bulb, whose mercury is wrapped in a damp cloth, reads 24°C in still air. The wet bulb is in a steady state in which the rate of heat lost by evaporation of water from the cloth is equal to the rate of heat gained from the surrounding air, which is 0.50W . Taking $L_v = 2.26 \times 10^6\text{J kg}^{-1}$, calculate the rate at which water must be evaporating from the cloth, expressed in grams per hour.

Question 30

A 100g sample of an unknown solid is heated steadily by a 50W electrical heater. The temperature-versus-time graph shows three segments: from $t = 0$ to $t = 60\text{s}$ the temperature rises linearly from -60°C to -20°C ; from $t = 60\text{s}$ to $t = 140\text{s}$ the temperature stays constant at -20°C ; from $t = 140\text{s}$ to $t = 200\text{s}$ the temperature rises linearly from -20°C to 50°C . Calculate the specific heat capacity of the solid phase, the latent heat of fusion at -20°C , and the specific heat capacity of the liquid phase.

Question 31

A lead bullet of mass 5.0g moving at 350m s^{-1} strikes and embeds in a wooden block. Assume that all the bullet's kinetic energy is converted to heat absorbed by the bullet alone (the block absorbs none). The bullet's initial temperature is 25°C . Take the specific heat capacity of lead as $c_{\text{Pb}} = 128\text{J kg}^{-1}\text{K}^{-1}$, the melting point of lead as 327°C , and the latent heat of fusion of lead as $L_{\text{f,Pb}} = 2.5 \times 10^4\text{J kg}^{-1}$. Determine whether the bullet melts on impact, and if so, calculate the mass of lead that melts.

Question 32

A 50g ice cube at 0°C is dropped from a height of 100m into a still pond at 0°C . Take $g = 9.81\text{m s}^{-2}$ and $L_{\text{f}} = 3.34 \times 10^5\text{J kg}^{-1}$. Assuming all the gravitational potential energy is converted to heat absorbed by the ice cube alone, calculate the mass of ice that melts on impact, and express it as a percentage of the cube's original mass.

Chapter 3

HEAT TRANSFER

INTRODUCTION

The school kitchen at Miono Secondary had two rules. The first was that no student was allowed inside without supervision. The second was that no student was allowed to touch the cooking equipment, because Mama Salha, the school cook, had once caught a Form Five student tasting her bean stew, and the public lecture that followed had been so fierce that the boy avoided the kitchen forever.

None of this applied when Mr. Akilikubwa was involved. He had been teaching physics at Miono for twenty-three years, and the entire school simply trusted him with things. If he wanted to borrow the kitchen for a lesson, Mama Salha would sigh, wipe her hands on her apron, and say, *“Take it, but if that boy with the big voice breaks my pot, you are buying the replacement.”* The boy with the big voice was, of course, Kipanga.

On this Wednesday morning, Mr. Akilikubwa had arrived before the first bell. Sitting on one of the kitchen’s large commercial gas burners was an aluminium *sufuria* of water that had been boiling hard for ten minutes, sending a column of steam toward the corrugated iron ceiling. Inside the *sufuria*, resting against the rim with their handles sticking out above the waterline, were two utensils. On the left: a stainless steel cooking spoon, the kind with a long metal handle that runs all the way from the bowl to the tip without interruption. On the right: a wooden *mwiko*, with a thick round handle Mama Salha had been stirring *ugali* with since before any of these students were born.

Both had been sitting in the same boiling water, in the same pot, on the same gas burner, for the same ten minutes. This detail is important, and Mr. Akilikubwa would make sure it was repeated until nobody could pretend they had missed it.

The Form Five physics class filed in through the side door. The kitchen smelled of steam and yesterday’s bean stew. They arranged themselves in a semicircle around the burner. Kipanga, who always stood at the front on the grounds that education worked best at close range, positioned himself directly between the two utensil handles. Kipute took a spot two rows back, where she had a clear view of both handles and, more importantly, a clear view of Kipanga.

Mr. Akilikubwa let the silence build. He was excellent at this. Other teachers filled silence with words; Mr. Akilikubwa filled it with expectation, and the expectation always paid a higher dividend.

“Two utensils,” he said, when every eye was on him. *“One steel. One wood. Both have been in the same boiling water for the same ten minutes. The same temperature. The same heat available. The same everything. I would like two volunteers. Each will pick up one utensil by the handle, hold it for three seconds, then put it down. Who would like to go first?”*

Kipanga’s hand went up with the speed and enthusiasm of a man who had learned absolutely nothing from the previous two chapters.

“Choose your utensil, Kipanga.”

Kipanga did not hesitate. He did not pause to ask himself why a physics teacher might be offering this choice in front of thirty witnesses. He simply reached for the stainless steel spoon, because it was shinier, because it looked more professional, and because Kipanga, in the long and colourful history of his physics education, had never once made a decision by thinking about it first.

He grabbed the steel handle.

For approximately one quarter of a second, nothing happened. This was the quarter-second in which the thermal energy stored in the steel, conducted at roughly 50 watts per metre per kelvin, was still travelling from his palm to his pain receptors. If physics had a sense of humour, this was the pause before the punchline.

Then the quarter-second ended.

Kipanga released the spoon with a sound that the English language does not have an adequate word for, but which the Swahili language would describe as something between a yelp, a prayer, and a very short autobiography of regret. The spoon hit the rim of the *sufuria*, bounced once with a bright metallic ring, and clattered onto the concrete floor. Kipanga shook his hand in the air with the frantic energy of a man trying to detach his own fingers from his body, then stared at his palm as though expecting to find a written explanation branded into the skin.

“That,” he said, with the wounded dignity of a man who feels the universe owes him an apology, *“is not a fair experiment.”*

Mr. Akilikubwa did not respond. He simply turned to the rest of the class. *“Volunteer number two?”*

Kipute walked forward without raising her hand, because Kipute did not compete for things she had already worked out. She picked up the wooden *mwiko* by its handle. She held it for three seconds. She held it for ten. She could have held it all morning. The handle was warm, the way a cup of tea is warm, but it was not hot. It was not painful. It was not even uncomfortable. It had been sitting in exactly the same boiling water, at exactly the same temperature, for exactly the same ten minutes as the steel spoon that had just reorganised Kipanga’s entire relationship with kitchen utensils. She placed the *mwiko* back in the pot and returned to her spot without a word.

Kipanga, still holding his right hand with his left as though it were an injured bird, stared at her with the expression of a man watching someone walk calmly through a door that had just electrocuted him.

“How,” he said, in a voice quiet enough to confirm that he was genuinely confused rather than performing confusion for the audience. *“Same water. Same temperature. Same time. The only difference between that spoon and that mwiko is the material. And you are telling me the material is the reason I just lost all feeling in my right hand?”*

“Yes,” said Kipute. *“Remember the metal gate two weeks ago? You touched it at 14 degrees and it felt freezing, then sat on the wooden bench at the same 14 degrees and it felt warm. I told you then that your hand was not measuring temperature. It was measuring the rate at which heat flowed out of it. The steel spoon has the same story. It is not hotter than the mwiko. It conducts heat into your skin faster.”*

Kipanga opened his mouth. He closed it. He looked at his palm, where a faint pink stripe was forming where the steel handle had pressed. He looked at the *mwiko*, still resting peacefully in the pot. And then he said something nobody in the Form Five physics class had ever heard Kipanga say before.

“That actually makes sense.”

Three students in the back row applauded. Kipanga ignored them with considerable effort.

Mr. Akilikubwa stepped forward, picked the steel spoon off the floor, wiped it, and placed it back in the *sufuria*. *“Kipute is exactly right. What your hand felt was not temperature. It was the rate at which thermal energy flowed from the spoon into your skin. Steel has a high thermal conductivity; it delivers heat fast. Wood has a very low thermal conductivity; it delivers the same available heat so slowly your hand barely registers the transfer. The temperature was identical. What was completely different was the speed of delivery. And that single idea, the rate of heat transfer, is the subject of this entire chapter.”*

He turned to the small portable whiteboard propped against the tiles and wrote three words in a vertical column:

CONDUCTION CONVECTION RADIATION

“Chapter 2 asked how much heat. How many joules to raise water by one degree, to melt ice, to boil a kettle. All those numbers lived inside $Q = mc\Delta T$ and $Q = mL$.”

*“Chapter 3 asks a different question. It asks how fast that heat moves and which road it travels. Does it travel through the material itself, atom by atom, without the material moving? That is **conduction**, and it is the reason Kipanga’s hand is still stinging. Does it travel because the material itself moves, carrying heat the way a daladala carries passengers from Mwenge to Mbagala? That is **convection**, and it is the reason the ceiling of your classroom is warmer than the floor. Or does it travel as electromagnetic radiation, needing no material at all, crossing the vacuum of space from the Sun to your face? That is **radiation**, and it is the reason the Earth is not a frozen rock, even though it is surrounded by the coldest emptiness in the universe.”*

“Every thermal phenomenon you will meet in the next seventy pages is one of these three mechanisms doing its work, or being deliberately stopped from doing its work. By the end of this chapter, you will know why a blanket keeps you warm in July and why the same blanket wrapped around a block of ice slows its melting. Why birds puff out their feathers and animals curl into a ball. Why ice on a pond grows more slowly as it gets thicker. How a forensic scientist estimates time of death by measuring a body’s temperature. And why Kipanga’s hand hurts and Kipute’s does not.”

Kipanga held up his palm for the class to see, turning it slowly so that even the back row could appreciate the evidence.

“Scientific contribution,” he announced, with the dignity of a man who has decided that his suffering has advanced human knowledge and therefore was not entirely in vain.

The class laughed. Kipute closed her notebook. And just like that, Chapter 3 had begun.

Chapter 2 showed how much heat is needed to change temperature ($Q = mc\Delta T$) and how much is needed for a phase change ($Q = mL$). It established the energy bookkeeping of thermal physics: if you know the mass, the material, and the temperature change, you can calculate the joules. But it deliberately left one question unanswered. When Kipanga touched the metal gate at 14°C and it felt freezing, while the wooden bench at the same 14°C felt warm, Chapter 2 said only that the explanation involved the rate of heat flow and that the full answer would come later. This is later.

Chapter 3 supplies the missing physics. Fourier's law of conduction, derived properly in the conduction section, gives the rate of heat flow through a solid in terms of the material's thermal conductivity, the cross-sectional area, and the temperature gradient. This is the same equation that was borrowed as an unproved tool in Chapter 2's ice pond calculation; it is now derived from first principles and explored in full generality, including radial heat flow through spheres and cylinders. Newton's law of cooling, developed later in this chapter, provides the exponential temperature decay of a body losing heat to its surroundings, the same mathematical form that will reappear in Chapter 11 when we model the discharge of a capacitor through a resistor, because the underlying differential equation is identical. The Stefan-Boltzmann radiation law connects forward to Chapter 5, where the distinction between an isothermal process and an adiabatic process depends entirely on knowing how quickly a system exchanges heat with its surroundings. If the exchange is fast (good conductor, large surface area, long time available), the process is isothermal. If the exchange is slow (poor conductor, small surface, rapid compression), the process is adiabatic. That judgement requires the tools of this chapter.

THERMAL CONDUCTION

Of the three heat-transfer mechanisms introduced in our kitchen, **conduction** is the one that taught Kipanga his lesson. The steel spoon and the wooden *mwiko* were sitting in the same boiling water at the same temperature, yet only one of them sent a stripe across his palm. The water did not move from the pot to his hand. The spoon did not move. The handle simply stood there in the air above the rim, and somehow the heat travelled along it from the boiling tip to the cool grip in less than a second. That mysterious through-the-metal journey is what physicists call conduction, and the next several pages will turn it from a mystery into a calculation.

Conduction is the process in which heat flows from the hotter region of a material to the cooler region without any net movement of the material itself. Picture the steel spoon at the molecular level. Down at the bowl end, the molecules are immersed in 100°C water and are vibrating with great energy. Up at the handle end, the molecules are surrounded by cool morning air and are vibrating gently. Each vibrating molecule jostles its immediate neighbours. The energetic ones at the bottom shake their neighbours harder; those neighbours shake the next neighbours; and the disturbance ripples up the spoon as a wave of increasingly active vibration. Heat has moved from the bowl to the handle, but no atom of steel has actually travelled with it. The molecules vibrate in place. They simply pass the energy along, like a row of seated people clapping a rhythm down the line.

This passing-along is what makes conduction so different from the other two mechanisms. In convection, the hot fluid itself rises and physically carries the energy with it. In radiation, no medium is needed at all. But in conduction, the material stays where it is, and the energy migrates through it from atom to atom. This explains immediately why **solids conduct heat best** (their atoms are packed close together and pass vibrations efficiently), **liquids conduct moderately** (their molecules are close but can shift around, dissipating the wave), and **gases conduct very poorly** (their molecules are so far apart that vibrations rarely reach a neighbour at all).

Fourier's Law of Heat Conduction

Saying that heat flows from hot to cold is true but useless. To do physics with conduction, we need a number. We need to know exactly *how many joules of energy will travel through a slab of material every second, given how thick the slab is, how big its faces are, what it is made of, and how big the temperature difference is between its two faces*. The relationship that gives us this number is called **Fourier's law**, and we will build it from observations rather than memorise it from a book.

Imagine a uniform slab of material with **cross-sectional area A** and **thickness L** . One face is held at the **higher temperature θ_1** and the opposite face is held at the **lower temperature θ_2** . After waiting long enough for the slab to settle into a **steady state** (this means *waiting until the temperature at every internal point has stopped changing with time*), heat flows steadily through the slab from the hot face to the cold face. The geometry is shown in the following figure.

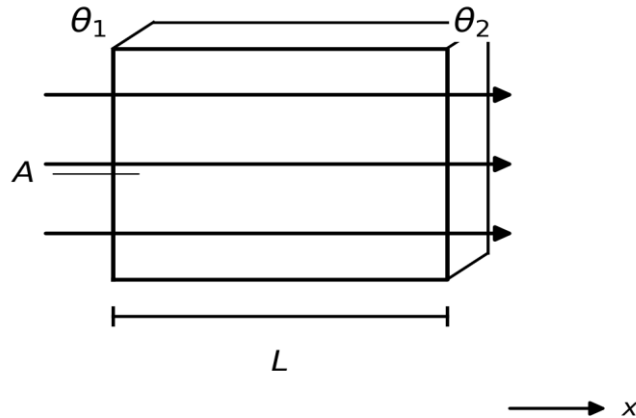


Figure: Heat conduction through a uniform slab of cross-sectional area A and thickness L . The hot face on the left is at temperature θ_1 and the cold face on the right is at θ_2 . Heat flows in the $+x$ direction, from hot to cold.

Let dQ be the small quantity of heat that crosses the slab in a small time dt . Experiment shows that dQ depends on four factors, each in a way that should match your intuition.

First, dQ is directly proportional to the area A of the slab. A bigger glass window loses more heat than a smaller one of the same thickness, just as more lanes on a highway can carry more cars per minute than fewer lanes.

Second, dQ is directly proportional to the time dt for which the heat is allowed to flow. Twice as long, twice as much heat. This is so obvious it almost does not need stating, but Fourier was a careful man and did state it.

Third, dQ is directly proportional to the temperature difference $(\theta_1 - \theta_2)$ between the two faces. Doubling the temperature difference doubles the rate at which heat flows. This is why a window leaks more heat on a freezing night than on a mild evening.

Fourth, dQ is inversely proportional to the thickness L of the slab. A thick wall conducts heat more slowly than a thin one, which is why a 30 cm brick wall keeps a house cooler in midday Dar es Salaam than a 5 mm sheet of glass would.

Combining all four observations into a single proportionality:

$$dQ \propto \frac{A(\theta_1 - \theta_2) dt}{L}$$

To turn the proportionality into an equality, we introduce a constant of proportionality. This constant cannot be universal, because different materials clearly conduct heat at very different rates: a copper bar feels almost as hot as the boiling water at its far end within seconds, whereas a wooden stick of the same dimensions in the same water remains almost cool. The constant must therefore be a property of the specific material being used. We call it the coefficient of thermal conductivity and give it the symbol k . With this constant inserted:

$$dQ = \frac{k A (\theta_1 - \theta_2) dt}{L}$$

Dividing both sides by dt gives the rate of heat flow, which is exactly the quantity we want:

$$\frac{dQ}{dt} = \frac{k A (\theta_1 - \theta_2)}{L}$$

This is **Fourier's law** for steady-state conduction through a uniform slab. The left-hand side is the rate at which heat crosses the slab, measured in joules per second, which is watts. The right-hand side is the product of four things you can measure: the thermal conductivity of the material, the area of the slab, the temperature difference between the faces, and the thickness of the slab. Once you know these four numbers, the rate of heat flow follows immediately.

In words the law can be stated as: *the rate at which heat is conducted through a uniform slab in steady state is directly proportional to the cross-sectional area of the slab and to the temperature difference between its two faces, and inversely proportional to the thickness of the slab.*

The Temperature Gradient and the Negative Sign

The form of Fourier's law we have just written is the one that students use most often, but it hides a small subtlety that becomes important the moment we move beyond a uniform slab. Heat does not really care about the temperature difference between the two faces. Heat cares about the rate at which temperature changes with distance, which is to say it cares about the slope of the temperature curve at the point where the heat is passing through.

To see this clearly, imagine cutting a thin slice of thickness dx out of the slab at some depth x , where the temperature on the hot side of the slice is θ and the temperature on the cold side is $(\theta + d\theta)$. Note that $d\theta$ here is a negative quantity, because temperature decreases as we move from hot to cold along the $+x$ direction. The temperature change across the thin slice is $d\theta$, the thickness is dx , and the rate of change of temperature with distance, called the temperature gradient, is $d\theta/dx$. Applying Fourier's law to this thin slice:

$$\frac{dQ}{dt} = -kA \frac{d\theta}{dx}$$

The minus sign appears for a precise physical reason that is worth stating slowly. We have chosen the $+x$ direction to point from the hot face toward the cold face, so heat flows in the $+x$ direction; we want dQ/dt to be a positive number when heat is flowing the way we have set up our axes. But $d\theta/dx$ is negative, because temperature is dropping as x increases. Multiplying by a minus sign converts the negative gradient into a positive heat flow rate. The minus sign is therefore not a mathematical inconvenience. It is the equation's way of telling you that heat always flows **down** the temperature gradient, from high to low. If you ever find yourself having computed a **negative** rate of heat flow in a problem, the equation is not broken; it is simply telling you that the heat is actually flowing in the opposite direction to the one you assumed.

In a uniform slab at steady state, the temperature falls linearly from θ_1 at one face to θ_2 at the other, so the gradient $d\theta/dx$ is constant everywhere inside the slab and equals $-(\theta_1 - \theta_2)/L$. Substituting this into the Fourier law immediately recovers our earlier formula, with the minus signs cancelling neatly. The two forms of the law are therefore the same physics, written for two different situations: the gradient form is general and applies even when the temperature varies non-linearly with position (such as inside a sphere or a cylinder, which we will meet later in this section), while the slab form is the special case for uniform geometry.

The Thermal Conductivity, k

The constant k that appeared in our derivation is the most important number in this chapter. It is the coefficient of thermal conductivity, and it is the single number that distinguishes a copper saucepan from a clay one, a glass window from a brick wall, and Kipanga's steel spoon from Kipute's wooden *mwiko*. Before we tabulate values for different materials, let us extract a careful definition of k from the equation it lives in.

Rearranging Fourier's law to make k the subject:

$$k = \frac{\frac{dQ}{dt}}{\frac{A(\theta_1 - \theta_2)}{L}}$$

Look at what this equation is telling us. The **thermal conductivity of a material** is the rate of heat flow per unit area per unit temperature gradient. Equivalently, k is the quantity of heat that flows per second through a slab of unit thickness and unit cross-sectional area when the temperature difference between its two faces is one kelvin. A material with a large k delivers a lot of heat per second under modest conditions; a material with a small k delivers very little. The unit of k is watts per metre per kelvin ($\text{W m}^{-1} \text{K}^{-1}$), which can also be written as joules per second per metre per kelvin ($\text{J s}^{-1} \text{m}^{-1} \text{K}^{-1}$). The two are identical, since one watt is one joule per second.

To a very good approximation, the thermal conductivity of any material is independent of temperature over the range of temperatures we will meet in this book. This is a useful simplification: we can treat k as a fixed number for each material, look it up in a table, and use it without worrying that the value might shift midway through a calculation. (At very high temperatures, or for some exotic materials, k does drift; A-level physics ignores this.)

A table of thermal conductivities

The following values cover almost every material you will meet in this chapter and the next. Notice the enormous range, more than four orders of magnitude from silver at the top to air at the bottom.

Material	$k / \text{W m}^{-1} \text{K}^{-1}$	Conductor type
Silver	420	excellent conductor
Copper	380	excellent conductor
Aluminium	240	very good conductor
Brass	109	good conductor
Iron	80	good conductor
Lead	35	moderate conductor
Concrete	1.3	poor conductor
Glass	1.0	poor conductor
Brick	0.7	insulator (mild)
Water	0.6	insulator
Wood (<i>mwiko</i>)	0.15	insulator
Felt	0.04	insulator (good)
Air (still)	0.026	insulator (excellent)

A few patterns leap out of this table. Metals occupy the top half (silver, copper, aluminium, brass, iron, lead) with conductivities in the tens to hundreds. Non-metallic solids and liquids occupy the middle (concrete, glass, brick, water, wood) with conductivities below 2. Trapped gases sit at the bottom (felt and still air) with conductivities below 0.05. Felt and still air are not chemically magical: they are just gases that have been trapped in a way that prevents them from convecting. We will return to this trapped-air idea repeatedly in this chapter, because it explains everything from blankets to bird feathers to the design of vacuum flasks.

One entry deserves a second look. Stainless steel does not appear in the table; its thermal conductivity is roughly $50 \text{ W m}^{-1} \text{K}^{-1}$ (the value we quoted in the kitchen as Kipanga's palm was being introduced to physics). Wood is roughly 0.15. The ratio of these two numbers is more than 300, which is why his hand experienced the steel spoon and the *mwiko* as utterly different objects even though they were sitting in the same boiling water. The temperature was identical. The conductivity was 300 times apart. His skin measured the conductivity, not the temperature.

Three short worked examples follow, each chosen to drill a different muscle. None of them will hurt as much as the steel spoon. We promise.

BINDER Example 1

A glass window in a school laboratory has thickness 8.0mm and area 2.0m^2 . The inner surface of the glass is at 20°C and the outer surface is at 0°C . The thermal conductivity of glass is $1.0\text{Wm}^{-1}\text{K}^{-1}$. Calculate the rate of heat loss through the window.

Solution

Using Fourier's law:

$$\frac{dQ}{dt} = \frac{k A (\theta_1 - \theta_2)}{L}$$

Substituting the given values:

$$\frac{dQ}{dt} = \frac{(1.0 \text{ Wm}^{-1}\text{K}^{-1})(2.0 \text{ m}^2)(20^\circ\text{C} - 0^\circ\text{C})}{8.0 \times 10^{-3} \text{ m}} = 5000 \text{ W} = 5.0 \text{ kW}$$

The rate of heat loss through the window is 5.0 kW.

Making Sense of the Answer: Five kilowatts is the power of five small electric heaters running flat out, and it is leaking out of a single laboratory window. This is why glass windows are such a poor choice in cold climates and why double-glazing exists. The number is large because glass is a much worse insulator than its solid appearance suggests, and because 8mm is genuinely thin.

Think Like a Physicist: The first thing to do in any Fourier-law problem is to convert all lengths to metres. A thickness in millimetres dropped into the formula without conversion will give an answer that is wrong by a factor of a thousand, which is the kind of error that turns a 5kW leak into a 5W trickle, and a passing grade into a failing one.

BINDER Example 2

A copper rod of length 2.0m and uniform cross-sectional area 4.0cm² is well lagged so that no heat escapes from its sides. One end is held at 100 °C and the other end is held at 20 °C. Take the thermal conductivity of copper as 380Wm⁻¹K⁻¹. Find the rate of heat flow along the rod, and find the temperature gradient inside the rod.

Solution

Using Fourier's law for a uniform lagged rod:

$$\frac{dQ}{dt} = \frac{k A (\theta_1 - \theta_2)}{L}$$

Substituting:

$$\frac{dQ}{dt} = \frac{(380 \text{ Wm}^{-1}\text{K}^{-1})(4.0 \times 10^{-4} \text{ m}^2)(100^\circ\text{C} - 20^\circ\text{C})}{2.0 \text{ m}} = 6.08 \text{ W}$$

The rate of heat flow along the rod is 6.08W.

For the temperature gradient: in a lagged uniform rod the temperature drops linearly from one end to the other. The gradient is therefore the same everywhere along the rod and equals the total temperature change divided by the length:

$$\frac{d\theta}{dx} = -\frac{(\theta_1 - \theta_2)}{L} = -\frac{(100 - 20)\text{K}}{2.0 \text{ m}} = -40 \text{ K m}^{-1}$$

The temperature gradient inside the rod is -40K per metre.

The minus sign confirms that temperature decreases as we move from the hot end (x = 0) toward the cold end (x = 2 m).

Making Sense of the Answer: A 6W flow along a thin rod is small in absolute terms but large compared with what a wooden rod of identical dimensions would deliver. With k for wood near 0.15, the same rod made of wood would carry only about 0.0024 W under the same conditions, which is roughly 2500 times less.

Think Like a Physicist: In any uniform lagged rod, the temperature gradient is just the total temperature drop divided by the length. You almost never need calculus to find it. Save the dθ/dx machinery for situations where the geometry forces the gradient to vary, such as in radial heat flow through a sphere, which we will meet a few pages from now.

REAL Example 3

A classroom in a Dar es Salaam secondary school has a single louvre-glass window of total area 1.5m^2 and thickness 4.0mm . On a hot afternoon, the outer surface of the glass reaches 38°C while the inner surface, cooled by a ceiling fan blowing air across it, sits at 30°C . Take k for glass as $1.0\text{Wm}^{-1}\text{K}^{-1}$.

- (a) Calculate the rate at which heat is conducted into the classroom through the window.
 (b) If the lesson lasts 80 minutes, calculate the total quantity of heat delivered into the classroom through that single window during the lesson.

Solution

(a) Heat flows from the hotter outside to the cooler inside. Applying Fourier's law:

$$\frac{dQ}{dt} = \frac{kA(\theta_1 - \theta_2)}{L} = \frac{(1.0\text{Wm}^{-1}\text{K}^{-1})(1.5\text{m}^2)(38^\circ\text{C} - 30^\circ\text{C})}{4.0 \times 10^{-3}\text{m}} = 3000\text{W} = 3.0\text{kW}$$

The rate of heat flow into the classroom is 3.0kW .

(b) Total heat delivered in 80 minutes.

Time $t = 80 \times 60\text{s} = 4800\text{ s}$. Since the rate of heat flow is constant during the lesson:

$$Q = \frac{dQ}{dt} \times t = 3000\text{Js}^{-1} \times 4800\text{s} = 1.44 \times 10^7\text{ J} = 14.4\text{MJ}$$

The total quantity of heat delivered through that single window is $1.44 \times 10^7\text{ J}$.

That 14.4MJ of heat delivered through one window in one lesson, equivalent to the energy released by burning roughly 350 g of wood.

Making Sense of the Answer: *Three kilowatts coming through one window is roughly the heat output of three electric kettles boiling continuously. Multiply this by the four or five windows in a typical Tanzanian classroom and you have ten or fifteen kilowatts of solar heat being conducted in by the glass alone, which is exactly why classrooms in hot regions like Dar es Salaam grow steadily warmer through the morning and why architectural choices about glazing and shading actually matter for student comfort.*

Think Like a Physicist: *When a problem asks for total heat over a period of time, never forget to multiply the heat-flow rate by the time. Students often produce a beautifully correct rate in watts and then write that as the answer for the total energy. Watts are joules per second, not joules. Always finish the multiplication.*

With Fourier's law in our hands and the meaning of k properly understood, we are ready to look at what happens when the slab is not a single block of one material, when its sides are unable to keep heat in, or when its geometry is something other than a flat rectangle. Each of these complications introduces a new wrinkle, but the underlying law remains the same. Heat will keep flowing down the temperature gradient, at a rate determined by the conductivity, the geometry, and the temperature difference. The next subsection looks at the difference between a lagged bar and an unlagged bar, where the same rod can produce two completely different temperature profiles depending on whether or not its sides are insulated.

Lagged and Unlagged Bars

Fourier's law in the plain slab form assumes that all the heat that enters the hot face also leaves the cold face, with none escaping along the way. This is true for a slab whose sides are insulated, but it is dramatically false for a bare bar standing in open air. The same metal rod, with the same temperature applied at one end and the same temperature applied at the other, can produce two completely different temperature profiles depending on whether or not its sides are wrapped against the surrounding air. The two cases must be recognised on sight, because they obey different rules and demand different sketches.

The lagged bar

A bar is said to be **lagged** when its sides are wrapped in a poor conductor (asbestos, wool, foam, or any insulating sleeve) so that no heat can escape sideways. All the heat that enters the hot end is forced to travel along the bar and leave through the cold end. The lagging plays the same role as the four insulated walls of a pipe in fluid mechanics: **it confines the flow to a single path.**

Because no heat is lost from the sides, the rate of heat flow at every cross-section along the bar is the same. If we pick any cross-section A along the length of the bar and ask how many joules per second are crossing

it, the answer is identical to the answer at any other cross-section B. The bar acts as a one-dimensional conductor: heat enters at one end, travels straight through, and exits at the other end with no leakage.

This constancy of heat flow has a direct consequence. Applying Fourier's law between any two points on the bar:

$$\left(\frac{dQ}{dt}\right)_A = \left(\frac{dQ}{dt}\right)_B$$

Because the conductivity k is a property of the material, the area A is constant for a uniform bar, and the rate of heat flow is constant along the bar, the temperature gradient $d\theta/dx$ must also be constant everywhere inside the bar. A constant gradient means a straight-line temperature profile: temperature falls uniformly from θ_1 at the hot end to θ_2 at the cold end, with no curvature anywhere along the way.

Figure 3.2 shows both the lagged bar and the resulting temperature profile.

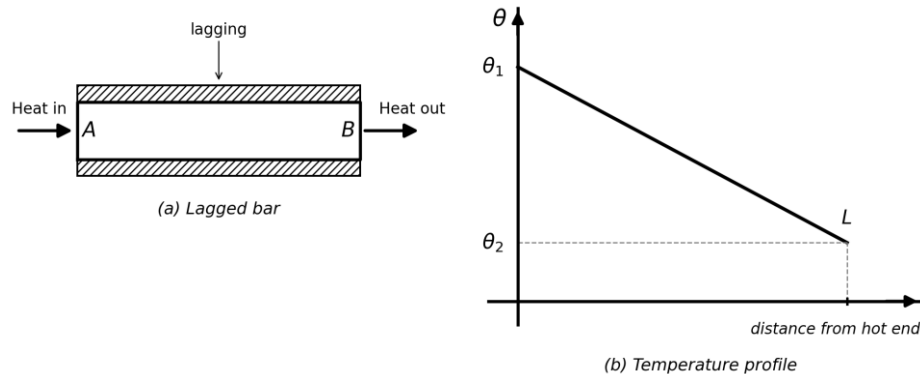


Figure: (a) A lagged bar with insulation along its sides forces all heat to travel along the bar from end A to end B. (b) The temperature falls linearly from θ_1 to θ_2 along the length L of the bar, because the heat-flow rate (and hence the temperature gradient) is constant at every cross-section.

Notice that the linear graph is what makes calculations on a lagged bar so straightforward. *Once the temperatures at the two ends are known, the temperature at any internal point can be read directly from the graph, or calculated from a single proportion:* a point one-third of the way from the hot end has a temperature one-third of the way from θ_1 to θ_2 , and so on. No calculus is needed at any stage.

The unlagged bar

Now strip the insulation off the bar. The same metal, the same length, the same end temperatures, but this time the sides are bare and exposed to the surrounding air. Heat now has more than one road to travel. Some of it still flows along the bar from the hot end to the cold end, exactly as before, but some of it leaves through the sides at every cross-section, escaping into the air by convection and radiation. The bar is leaking heat continuously along its entire length.

This sideways leakage has two important consequences.

- 1) The heat-flow rate is no longer constant along the bar. At the hot end, where the bar is hottest, a large amount of heat enters; further along, after some heat has escaped from the sides, less heat is travelling through the bar; by the time the cold end is reached, only the small fraction that has not yet leaked out remains. The heat-flow rate decreases continuously from the hot end to the cold end.
- 2) The temperature gradient is no longer constant either. Because the heat-flow rate at any cross-section equals $-kA(d\theta/dx)$, and because the rate is now decreasing along the bar, the magnitude of the gradient must also be decreasing. Near the hot end, the bar carries a lot of heat and the gradient is steep; near the cold end, it carries very little and the gradient is gentle. The temperature profile is therefore not a straight line. It is a curve that falls steeply at the start and levels off toward the cold end. This shape is called **concave-up**: it bows downward and flattens as the cold end is approached.

The following figure shows the unlagged bar and its characteristic curved temperature profile, side by side with the lagged bar so the contrast can be seen at a glance.

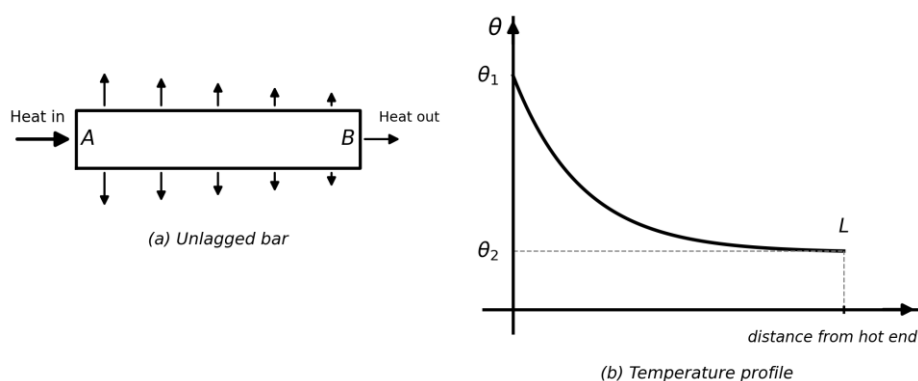


Figure: (a) An unlagged bar loses heat from its sides along its entire length, so less heat reaches the cold end than enters the hot end. (b) The temperature profile is no longer linear; it falls steeply near the hot end and levels off toward the cold end.

The contrast between the two bars is one of the most reliable indicators of physical understanding. A student who sketches a straight line for an unlagged bar has missed the point of the question, because the straight line is exactly the wrong answer. The straight line belongs to the lagged bar; the curve belongs to the unlagged one. Whenever a problem mentions that a bar has bare sides, or simply does not say it is lagged, the temperature profile must be drawn as a concave-up decay, never as a straight line.

Side-by-side comparison

The two cases are summarised in the table below for quick reference.

Feature	Lagged bar	Unlagged bar
Sides	wrapped in insulator	bare, exposed to air
Heat lost from sides	none	continuous along length
Rate of heat flow dQ/dt	constant at every cross-section	decreases from hot end to cold end
Temperature gradient $d\theta/dx$	constant	magnitude decreases along bar
Temperature profile	linear (straight line)	concave-up (curved decay)

Time to put the two bars to work. Bring out your pen and let physics speak for itself.

BINDER Example 4

A copper bar of length 50cm and uniform cross-sectional area 4.0cm^2 is well lagged. The hot end is maintained at 100°C and the cold end at 30°C . Take the thermal conductivity of copper as $380\text{Wm}^{-1}\text{K}^{-1}$. Find the temperature at a point 20cm from the hot end.

Solution

Because the bar is lagged, the temperature profile is a straight line. The temperature at any internal point therefore lies on the straight line joining $(0, 100^\circ\text{C})$ and $(50\text{cm}, 30^\circ\text{C})$.

Given:

Length $L = 50\text{cm} = 0.50\text{m}$

Hot end $\theta_1 = 100^\circ\text{C}$, cold end $\theta_2 = 30^\circ\text{C}$

Distance from hot end $x = 20\text{cm} = 0.20\text{m}$

The total temperature drop is:

$$\theta_1 - \theta_2 = 100^\circ\text{C} - 30^\circ\text{C} = 70^\circ\text{C}$$

Because the gradient is constant, the temperature drop over the first 20cm of the bar is in the same ratio as the distance covered:

$$\Delta\theta = \frac{x}{L} \times (\theta_1 - \theta_2) = \frac{0.20\text{m}}{0.50\text{m}} \times 70^\circ\text{C} = 28^\circ\text{C}$$

Subtracting this drop from the hot-end temperature gives the temperature at $x = 0.20\text{m}$:

$$\theta = \theta_1 - \Delta\theta = 100^\circ\text{C} - 28^\circ\text{C} = 72^\circ\text{C}$$

The temperature at the point 20cm from the hot end is 72°C .

Making Sense of the Answer: The point lies two-fifths of the way along the bar, so the temperature should drop by two-fifths of the total temperature drop. Two-fifths of 70°C is 28°C , leaving 72°C . The proportional answer matches the algebraic answer, which is the kind of double-check that catches arithmetic slips before they reach the marker.

Think Like a Physicist: In any lagged bar, the temperature at any internal point can be found by simple proportion, never by calculus. The straight-line profile is the gift that lagging gives you. Always check at the start of a heat-transfer problem whether the bar is lagged or not, because the answer changes completely between the two cases.

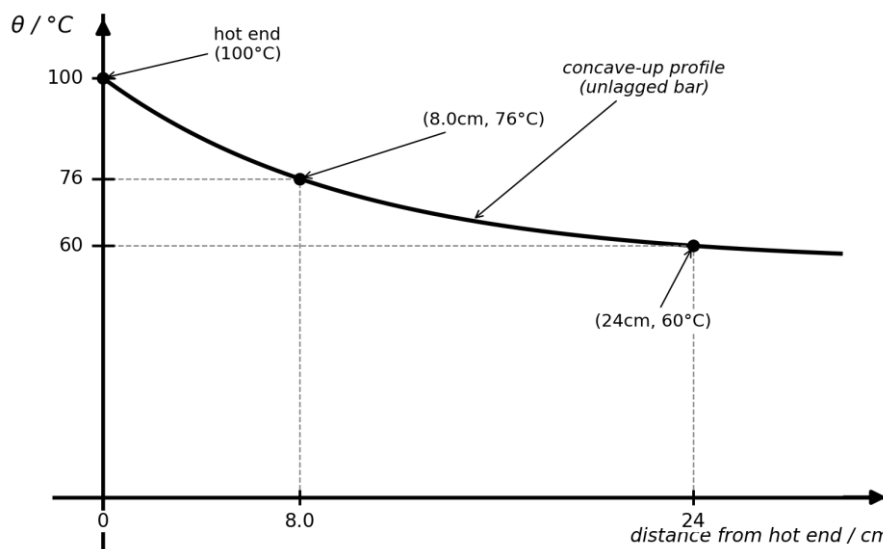
HOT Example 5

A uniform unlagged metal bar has its hot end maintained at 100°C . At a steady state, the temperature 8.0cm from the hot end is 76°C , and the temperature 24cm from the hot end is 60°C .

- Sketch the temperature profile of the bar.
- Calculate the average temperature gradient in the first region (0 to 8.0cm) and in the second region (8.0cm to 24cm), and compare them.
- Explain physically why the average gradient is steeper in the first region than in the second.

Solution

- The bar is unlagged, so the temperature profile is expected to be concave-up as sketched below.



- The average temperature gradient over a region is the temperature drop in that region divided by the length of the region.

First region (0 to 8.0cm):

$$\left(\frac{d\theta}{dx}\right)_1 = \frac{100^\circ\text{C} - 76^\circ\text{C}}{8.0\text{cm}} = \frac{24^\circ\text{C}}{8.0\text{cm}} = 3.0^\circ\text{C/cm}$$

Second region (8.0cm to 24cm), which is 16cm long:

$$\left(\frac{d\theta}{dx}\right)_2 = \frac{76^\circ\text{C} - 60^\circ\text{C}}{16\text{cm}} = \frac{16^\circ\text{C}}{16\text{cm}} = 1.0^\circ\text{C/cm}$$

The average gradient in the first region is $3.0^{\circ}\text{C}/\text{cm}$, three times steeper than the $1.0^{\circ}\text{C}/\text{cm}$ in the second region.

(c) Near the hot end, very little heat has yet escaped from the sides, so almost all the heat that entered at the hot end is still travelling along the bar at the 8.0cm mark. A large rate of heat flow demands a large temperature gradient (because $dQ/dt = -kA d\theta/dx$, and k and A are fixed for a uniform bar). Further along, the heat that was once travelling along the bar has been progressively leaking out of the sides, so by the time we reach the 8.0cm to 24cm region the bar is carrying much less heat. A smaller heat-flow rate requires a smaller temperature gradient, which is exactly what we calculated. The factor-of-three drop in average gradient between the two regions is direct evidence of the heat that has escaped sideways from the first 8.0cm of the bar.

Making Sense of the Answer: Each region of the unlagged bar reports how much heat is still travelling along it at that point. A steep gradient near the hot end says “a lot of heat is still here”. A gentle gradient further along says “most of the heat has already escaped sideways”. The concave-up shape of the temperature profile is the bar’s autobiography, written in the slope of its temperature curve.

Think Like a Physicist: When a problem asks about the gradient or the rate of heat flow at different positions along an unlagged bar, never assume they are equal. They are equal in a lagged bar; they are unequal in an unlagged one. The single word “unlagged” in a problem is a complete change of physical regime, and the entire approach to the problem must change with it.

With the lagged and unlagged cases now firmly in hand, we are ready for the next complication. What if the bar is lagged but is not made of a single material? What if it is two different conductors joined end to end? This is the situation of a composite conductor, which the next subsection takes up.

Composite Conductors

So far we have applied Fourier’s law to a single block of one material. Real walls, real roofs, and real cooking utensils are seldom this simple. The wall of a Tanzanian house may be built of brick on the outside, an air gap in the middle, and plaster on the inside; a cooking pot may have a thin steel base bonded to a thicker aluminium core; a domestic radiator is iron with a layer of paint on the outside. Each of these is a composite conductor, made of two or more materials joined together. The heat does not care that the materials are different. It simply flows from hot to cold, passing through whichever materials lie in its path. Our job is to predict the rate at which it does so.

Composite conductors come in two basic arrangements, and the analysis of each is genuinely beautiful because of how closely it parallels the analysis of resistors in an electric circuit. The two arrangements are slabs in **series** (one after the other, joined end to end so that *the heat must pass through them in turn*) and slabs in **parallel** (placed side by side so that *the heat can flow through either one independently*). Both arrangements are sketched in the following figure.

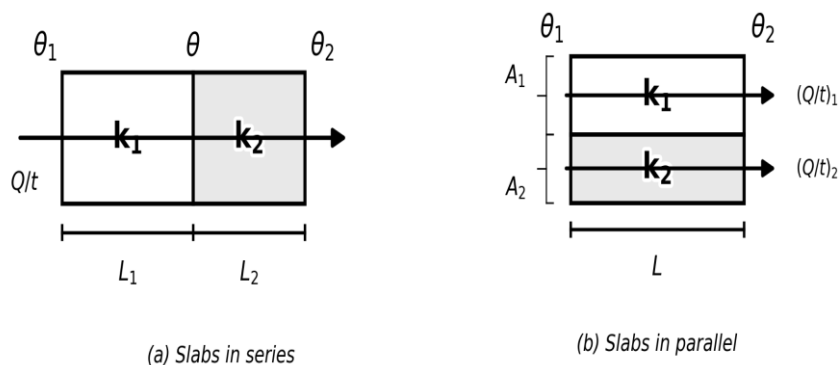


Figure: (a) Two slabs of conductivities k_1 and k_2 , lengths L_1 and L_2 , joined end to end in series. The heat-flow rate Q/t is the same through both slabs. The interface temperature is denoted θ . (b) Two slabs of conductivities k_1 and k_2 , cross-sectional areas A_1 and A_2 , placed side by side in parallel. Each slab carries its own heat-flow rate, and the total heat flow is the sum of the two.

Slabs in series

Consider two well-lagged slabs, both with the same uniform cross-sectional area A , joined end to end as shown in the figure. The first slab has thermal conductivity k_1 and length L_1 ; the second has thermal conductivity k_2 and length L_2 . The outer face of the first slab is held at the higher temperature θ_1 and the outer face of the second slab is held at the lower temperature θ_2 . We do not know the temperature at the interface where the two slabs meet, but we will call it θ and let the physics determine its value.

Because the slabs are well lagged, no heat escapes from the sides. All the heat that enters the first slab must therefore leave the second slab, and the rate of heat flow through the first slab must equal the rate of heat flow through the second slab. This is the key physical insight, and the entire derivation rests on it.

Let Q/t be the rate of heat flow through the composite. Applying Fourier's law to slab 1, where the temperature drops from θ_1 to θ across length L_1 :

$$\frac{Q}{t} = \frac{k_1 A (\theta_1 - \theta)}{L_1}$$

Applying Fourier's law to slab 2, where the temperature drops from θ to θ_2 across length L_2 :

$$\frac{Q}{t} = \frac{k_2 A (\theta - \theta_2)}{L_2}$$

These two equations express the heat-flow rate in two different ways, but the physics has insisted they must be equal. Setting them equal will eventually give us the interface temperature θ , but a more useful approach is to rearrange each equation so that the temperature drop sits alone, then add them.

Rearranging the first equation to isolate the temperature drop across slab 1:

$$\theta_1 - \theta = \left(\frac{Q}{t}\right) \frac{L_1}{k_1 A}$$

Rearranging the second equation to isolate the temperature drop across slab 2:

$$\theta - \theta_2 = \left(\frac{Q}{t}\right) \frac{L_2}{k_2 A}$$

Adding the two equations side by side, the unknown interface temperature θ cancels:

$$\theta_1 - \theta_2 = \left(\frac{Q}{t}\right) \frac{1}{A} \times \left(\frac{L_1}{k_1} + \frac{L_2}{k_2}\right)$$

Solving for the heat-flow rate:

$$\frac{Q}{t} = \frac{A (\theta_1 - \theta_2)}{\frac{L_1}{k_1} + \frac{L_2}{k_2}}$$

This is the working formula for two slabs in series. It generalises immediately: for any number n of slabs in series, the rate of heat flow through the composite is:

$$\frac{Q}{t} = \frac{A (\theta_1 - \theta_2)}{\sum \frac{L_i}{k_i}}$$

where the sum runs over all slabs, and θ_1 and θ_2 are the outer-face temperatures of the entire composite (the hot and cold ends).

Recovering the interface temperature

Once the heat-flow rate Q/t is known, the interface temperature θ follows immediately from either of the original Fourier equations. Substituting Q/t back into the rearranged form for slab 1:

$$\theta = \theta_1 - \left(\frac{Q}{t}\right) \frac{L_1}{k_1 A}$$

This will give a numerical value of θ that lies somewhere between θ_1 and θ_2 . *The natural follow-up question is: closer to which temperature? Closer to the hot-end value θ_1 , or closer to the cold-end value θ_2 ?*

A useful piece of physical insight: the interface temperature is closer to whichever side has the smaller temperature drop, and the smaller temperature drop occurs in whichever slab has the lower thermal resistance (high conductivity, short length). In short, the interface temperature is dragged toward the side that conducts heat more easily.

Thermal Resistance and the Electrical Analogy

The series formula we have just derived has a remarkable structure, and it deserves its own name and its own definition. Before we write them down, let us notice that this is not the first time we have met a situation where something flows, something drives the flow, and something opposes it. The same three-part story governs every electric circuit ever built, and physics rewards us for recognising the pattern.

In an electric circuit, a battery applies a potential difference V across a conductor. That potential difference drives a current I through the conductor. The conductor resists the flow, allowing only a limited current for each volt applied; the ratio V/I is called the resistance R , and the entire relationship is captured by Ohm's law $V = IR$. The units of resistance (volts per ampere, named the ohm) tell you the price in voltage you must pay to push one ampere through the device.

Now read that paragraph again, but replace every electrical word with its thermal counterpart. A **temperature difference $\theta_1 - \theta_2$** plays the role of **potential difference**: it is *the quantity that drives the flow*. The **rate of heat flow Q/t** plays the role of **current**: it is *the quantity that flows in response*, measured in joules per second, which is watts. And whatever plays the role of **resistance** should be the ratio of driving **temperature difference** to resulting **heat current**.

$$\text{Thermal resistance, } R = \frac{\text{Temperature difference}}{\text{heat current}} = \frac{\theta_1 - \theta_2}{\frac{Q}{t}}$$

So **thermal resistance R** of a slab is defined as *the ratio of the temperature difference across the slab to the rate of heat flow through it*.

Thermal resistance is therefore *a measure of opposition to heat current*, in exactly the sense that electrical resistance is a measure of opposition to electric current. A high- R slab demands a large temperature difference to produce a modest heat flow; a low- R slab lets heat flow freely for only a small temperature difference.

From Fourier's law:

$$\frac{Q}{t} = \frac{k A (\theta_1 - \theta_2)}{L}$$

Rearranging:

$$\frac{\theta_1 - \theta_2}{\frac{Q}{t}} = \frac{L}{k A} = R$$

So **thermal resistance** of a slab can also be defined as *the ratio of its thickness to the product of its thermal conductivity and cross-sectional area*:

$$R = \frac{L}{k A}$$

The unit of R follows directly from the formula: metres divided by ($\text{W m}^{-1} \text{K}^{-1} \times \text{m}^2$), which simplifies to **KW^{-1}** . A slab with high resistance impedes heat flow strongly; a slab with low resistance lets heat flow easily.

Every dependence in this formula matches what you would predict from the electrical analogy. A longer slab (larger L) has higher thermal resistance, just as a longer wire has more electrical resistance; heat must travel further and pushes against more molecules along the way. A broader slab (larger A) has lower thermal resistance, because heat has more parallel channels to flow through, just as a thicker wire carries current more easily. A material with higher thermal conductivity k has lower thermal resistance, in the same way a

material with higher electrical conductivity has lower resistivity. Every intuition you already have about resistors in wires transfers, symbol by symbol, to slabs in heat flow.

The correspondence between heat and electricity is summarised in the table below.

Electrical quantity	Thermal counterpart
Potential difference V (volt)	Temperature difference $\theta_1 - \theta_2$ (kelvin)
Current I (ampere, $= \text{Cs}^{-1}$)	Heat current Q/t (watt, $= \text{Js}^{-1}$)
Resistance $R = V/I$ (ohm, $= \text{VA}^{-1}$)	Thermal resistance $R = (\theta_1 - \theta_2)/(Q/t)$ (KW^{-1})
Ohm's law: $V = IR$	Thermal Ohm's law: $(\theta_1 - \theta_2) = (Q/t)R$

The two columns are so perfectly parallel that every technique you have already learned for electrical circuits, including series addition of resistances, parallel reciprocals, voltage-divider reasoning, and Kirchhoff's laws, works without modification in heat-flow problems. The thermal version of a problem is simply the electrical version with relabelled symbols.

The combination rules inherit from Ohm's law

Because the mathematics of the thermal and electrical cases are identical, the combination rules are identical too. The series formula we derived earlier now reads, in resistance form:

$$R_{\text{total}} = R_1 + R_2 + \dots + R_n$$

and the parallel rule, which we shall derive from Fourier's law in the next subsection, will have the familiar electrical reciprocal form:

$$\frac{1}{R_{\text{total}}} = \frac{1}{R_1} + \frac{1}{R_2} + \dots + \frac{1}{R_n}$$

In either arrangement, the rate of heat flow through the composite is then given by thermal Ohm's law:

$$\frac{Q}{t} = \frac{\theta_1 - \theta_2}{R_{\text{total}}}$$

This is the same equation you would write for the current in an electrical circuit, with V replaced by $\theta_1 - \theta_2$ and I replaced by Q/t . Anyone who has solved a resistor-network problem has already solved every heat-flow problem they are likely to meet.

Thermal Insulance: the R-value of a Material

Thermal resistance, as we have defined it, depends on three things: the material (through k), the thickness (through L), and the cross-sectional area (through A) of the specific slab in question. The area dependence is convenient when we are working with a specific object such as the brass rod in the experiment or the copper base of a particular cooking pot, but it becomes inconvenient the moment we want to compare materials at a more general level.

Consider an insulation manufacturer preparing a product catalogue for fibreglass sheets. If they quoted a thermal resistance R , the number would only apply to pieces of a specific area; a customer cutting a larger or smaller sheet would get a different R . What the manufacturer actually wants to publish is a single number that characterises the insulating ability of the material-and-thickness combination by itself, independent of how big a piece the customer happens to order. That number is called **thermal insulance**.

The **thermal insulance r** of a material of given thickness *is a measure of its intrinsic ability to resist the transfer of heat, independent of the cross-sectional area of the specific sample*. For pure conduction through a uniform slab, its mathematical definition is:

$$r = \frac{L}{k}$$

Its unit is $\text{m}^2 \text{K W}^{-1}$.

Insulance and thermal resistance are related by the cross-sectional area of the actual slab:

$$R = \frac{r}{A}$$

So thermal insulance is a pure material-and-thickness property: it tells you how much resistance content is packed into one unit of cross-section. Divide by the actual cross-section of the piece you have to recover the total thermal resistance R . A high-insulance material is a strong insulator; a low-insulance material lets heat through easily. The same piece of fibreglass insulation cut to different sizes has different thermal resistance but the same thermal insulance, which is exactly the property the catalogue wants to describe.

Why insulance captures the full heat-transfer story

We derived $r = L/k$ for pure conduction through a single solid slab, but the concept generalises in a powerful way that justifies the broader name. The insulating performance of a real building wall is built up from several contributions in series: the conduction insulance of each solid layer (L/k for the brick course, L/k for the plaster, L/k for the foam board, and so on), plus additional insulance terms contributed by the thin stagnant-air films that cling to the wall's two external surfaces. At those surfaces, heat crosses partly by convection (air currents carrying energy away) and partly by radiation (infrared waves emitted by the warm surface). These surface insulance terms are tabulated in standard references, in the same units $\text{m}^2 \text{K W}^{-1}$, and added to the conduction layers in series.

This is why the “R-value” quoted in insulation catalogues and building codes can always be added layer by layer to predict the performance of a wall without ever having to know the wall's area. The catalogue R-value is a **total thermal insulance**: a quantitative measure of how effectively a material or assembly can resist the transfer of heat through all three mechanisms (conduction, convection, and radiation) combined. Thermal insulance is the single currency in which the insulating performance of any material, assembly, or system is quoted; thermal resistance is simply that insulance applied to a specific area.

Thermal resistance and thermal insulance at a glance

Quantity	Formula	Units	What it tells you
Thermal resistance R	$L/(kA)$	K W^{-1}	Resistance of a specific slab of area A
Thermal insulance r	L/k	$\text{m}^2 \text{K W}^{-1}$	Material-and-thickness property, independent of area

The relationship between the two quantities is $R = r/A$, and the Fourier's-law formula for the heat-flow rate can be written in either currency. Using thermal resistance:

$$\frac{Q}{t} = \frac{\theta_1 - \theta_2}{R}$$

or, equivalently, using thermal insulance:

$$\frac{Q}{t} = \frac{A(\theta_1 - \theta_2)}{r}$$

Both forms are correct and give the same number. Use R when the specific slab dimensions are fixed up front and you want the Ohm's-law form; use r when the material property comes from a catalogue and you need to multiply by the wall area to get a heat flow.

Slabs in Parallel

Now consider two slabs of the same length L but possibly different cross-sectional areas A_1 and A_2 , placed side by side as shown in figure drawn earlier (page 137). Both slabs span between the same two outer surfaces, which are held at temperatures θ_1 and θ_2 . The heat now has two parallel paths to choose from: it can flow through slab 1 with conductivity k_1 , or through slab 2 with conductivity k_2 . In practice some heat takes each path, and the total heat-flow rate through the composite is the sum of the heat-flow rates through the two individual slabs.

Applying Fourier's law to each slab separately, with the same temperature difference $(\theta_1 - \theta_2)$ acting across both:

$$\left(\frac{Q}{t}\right)_1 = \frac{k_1 A_1 (\theta_1 - \theta_2)}{L}$$

$$\left(\frac{Q}{t}\right)_2 = \frac{k_2 A_2 (\theta_1 - \theta_2)}{L}$$

The total heat-flow rate is the sum:

$$\left(\frac{Q}{t}\right)_{\text{total}} = \left(\frac{Q}{t}\right)_1 + \left(\frac{Q}{t}\right)_2 = \frac{(k_1 A_1 + k_2 A_2) (\theta_1 - \theta_2)}{L}$$

In terms of thermal resistance, the parallel rule mirrors the electrical case exactly. The reciprocals of the individual resistances add to give the reciprocal of the total resistance:

$$\frac{1}{R_{\text{total}}} = \frac{1}{R_1} + \frac{1}{R_2} + \dots + \frac{1}{R_n}$$

The two arrangements obey the same rules as their electrical counterparts: resistances in series add directly, and reciprocals of resistances in parallel add. Anyone who has wired a circuit can analyse a composite wall.

Summary of the Two Arrangements

The two arrangements and their working formulas are summarised below for quick reference.

Feature	Series	Parallel
Common quantity	heat-flow rate Q/t	temperature difference $(\theta_1 - \theta_2)$
What the slabs share	the same heat passes through each in turn	the two faces are at the same pair of temperatures
Total resistance	$R_{\text{tot}} = R_1 + R_2 + \dots + R_n$	$1/R_{\text{tot}} = 1/R_1 + 1/R_2 + \dots + 1/R_n$
Effective conductivity (equal lengths and areas)	$k_e = k_1 k_2 / (k_1 + k_2)$	$k_e = k_1 + k_2$
Electrical analogue	resistors in series	resistors in parallel

Three examples. Bring your pen, your arithmetic, and your common sense. The physics will do the rest.

BINDER Example 6

A 10cm brass bar ($k = 109 \text{ Wm}^{-1}\text{K}^{-1}$) is joined end-to-end to a 10cm copper bar ($k = 380 \text{ Wm}^{-1}\text{K}^{-1}$), forming a composite of cross-section 6.0 cm^2 . The brass end is held at 100°C and the copper end at 20°C , and the entire arrangement is well lagged. Find: (a) the rate of heat flow along the composite, (b) the temperature at the brass-copper interface.

Solution

Two slabs in series between two fixed temperatures. This is a textbook Ohm's-law problem in thermal clothing.

Given:

Brass: $k_1 = 109 \text{ Wm}^{-1}\text{K}^{-1}$, $L_1 = 10 \text{ cm} = 0.10 \text{ m}$

Copper: $k_2 = 380 \text{ Wm}^{-1}\text{K}^{-1}$, $L_2 = 10 \text{ cm} = 0.10 \text{ m}$

Cross-section $A = 6.0 \text{ cm}^2 = 6.0 \times 10^{-4} \text{ m}^2$

Hot end $\theta_1 = 100^\circ\text{C}$, cold end $\theta_2 = 20^\circ\text{C}$

(a) Compute the thermal resistance of each slab using $R = L/(kA)$.

Brass slab:

$$R_1 = \frac{L_1}{k_1 A} = \frac{0.10 \text{ m}}{(109 \text{ Wm}^{-1}\text{K}^{-1})(6.0 \times 10^{-4} \text{ m}^2)} = 1.529 \text{ KW}^{-1}$$

Copper slab:

$$R_2 = \frac{L_2}{k_2 A} = \frac{0.10\text{m}}{(380\text{Wm}^{-1}\text{K}^{-1})(6.0 \times 10^{-4} \text{m}^2)} = 0.439\text{KW}^{-1}$$

Combine using the series rule:

$$R_{\text{total}} = R_1 + R_2 = 1.529\text{KW}^{-1} + 0.439\text{KW}^{-1} = 1.968\text{KW}^{-1}$$

Apply thermal Ohm's law:

$$\frac{Q}{t} = \frac{\theta_1 - \theta_2}{R_{\text{total}}} = \frac{80\text{K}}{1.968\text{KW}^{-1}} = 40.7\text{W}$$

The rate of heat flow along the composite is 40.7W.

(b) Interface temperature. The same heat current $Q/t = 40.7\text{W}$ flows through the brass slab (series means equal current). The temperature drop across just the brass slab is given by thermal Ohm's law applied to that slab alone:

$$\theta_1 - \theta = \frac{Q}{t} \times R_1 = (40.7\text{W})(1.529\text{KW}^{-1}) = 62.2\text{K}$$

So the interface temperature is:

$$\theta = \theta_1 - 62.2\text{K} = 100^\circ\text{C} - 62.2^\circ\text{C} = 37.8^\circ\text{C}$$

The interface temperature is 37.8°C.

Making Sense of the Answer: The interface temperature has settled much closer to the cold end (20°C) than to the hot end (100°C). This is exactly what we expected: the brass slab has 3.5 times the thermal resistance of the copper slab (1.529 versus 0.439 KW⁻¹), so in a voltage-divider sense the brass slab “drops” 62.2K out of the 80K total while the copper slab drops only 17.8K. The interface therefore lies close to the temperature of the low-resistance side, which is the cold end.

Think Like a Physicist: In any series composite, before doing any calculation, identify which slab has the higher thermal resistance. The interface temperature will lie close to the temperature of the lower-resistance side. This single rule of thumb catches arithmetic errors instantly: if your interface temperature lands on the wrong side of the midpoint, you have made a slip somewhere in your calculation.

REAL Example 7

A classroom at Usagara Secondary School has a brick exterior wall of total area 15m², thickness 150mm ($k = 0.70\text{Wm}^{-1}\text{K}^{-1}$). Set into this wall is a single-pane glass louvre window of area 1.8m², thickness 4.0mm ($k = 1.0\text{Wm}^{-1}\text{K}^{-1}$). On a hot Tanga afternoon, the outer surfaces of both the wall and the glass reach 40°C, while the inside surfaces are kept at 30°C by a ceiling fan. (a) Calculate the rate of heat gain through the glass window. (b) Calculate the rate of heat gain through the brick wall (the area outside the window). (c) Calculate the total rate of heat gain through the exterior wall including the window. (d) The window is only 11% of the total exterior-wall area; what percentage of the total heat gain does it contribute? Comment on what this tells us about building design in hot climates.

Solution

The window and the brick wall share the same inside and outside temperatures, so they sit in parallel as two independent heat-flow paths. We compute the thermal resistance of each path, then apply thermal Ohm's law to each path, then add the resulting heat currents.

Given:

Outside temperature $\theta_1 = 40^\circ\text{C}$, inside temperature $\theta_2 = 30^\circ\text{C}$, so $\Delta\theta = 10\text{K}$

Window: $A_w = 1.8\text{m}^2$, $L_w = 4.0\text{mm} = 4.0 \times 10^{-3} \text{m}$, $k_w = 1.0\text{Wm}^{-1}\text{K}^{-1}$

Brick wall (excluding window): $A_b = 15\text{m}^2$, $L_b = 150\text{mm} = 0.150\text{m}$, $k_b = 0.70\text{Wm}^{-1}\text{K}^{-1}$

(a) Heat gain through the glass window. Thermal resistance of the window pane:

$$R_w = \frac{L_w}{k_w A_w} = \frac{4.0 \times 10^{-3} \text{m}}{(1.0\text{Wm}^{-1}\text{K}^{-1})(1.8\text{m}^2)} = 2.22 \times 10^{-3} \text{KW}^{-1}$$

Applying thermal Ohm's law to the window:

$$\left(\frac{Q}{t}\right)_w = \frac{\Delta\theta}{R_w} = \frac{10K}{2.22 \times 10^{-3} \text{ KW}^{-1}} = 4500W$$

(b) Heat gain through the brick wall. Thermal resistance of the brick:

$$R_b = \frac{L_b}{k_b A_b} = \frac{0.150m}{(0.70Wm^{-1}K^{-1})(15m^2)} = 1.43 \times 10^{-2} \text{ KW}^{-1}$$

Again, applying thermal Ohm's law to the brick wall:

$$\left(\frac{Q}{t}\right)_b = \frac{\Delta\theta}{R_b} = \frac{10K}{1.43 \times 10^{-2} \text{ KW}^{-1}} = 700W$$

(c) Total heat gain. In parallel, the heat currents add:

$$\left(\frac{Q}{t}\right)_{\text{total}} = \left(\frac{Q}{t}\right)_w + \left(\frac{Q}{t}\right)_b = 4500W + 700W = 5200W$$

The total rate of heat gain into the classroom through this exterior wall is about 5.2kW.

(d) Fraction of the total heat gain contributed by the window:

$$\frac{4500W}{5200W} = 0.87 = 87\%$$

The glass window, despite occupying only 11% of the exterior-wall area, contributes 87% of the total heat gain. The brick wall, although 8.3 times larger in area than the window, contributes only 13% of the total heat load entering through the exterior. Windows dominate the thermal performance of the wall.

Making Sense of the Answer: The thermal-resistance numbers tell the whole story in one line: $R_b/R_w = 1.43 \times 10^{-2} / 2.22 \times 10^{-3} = 6.4$. The brick wall has about 6.4 times the thermal resistance of the window, so under the same 10K temperature difference it carries only 1/6.4 as much heat. This factor of 6.4 is the ratio of heat currents, and it is what pushes 87% of the total heat through the window even though the window covers only 11% of the wall area. Two effects combine to make the window so much more conductive than the wall it sits in: glass has a higher thermal conductivity than brick (1.0 versus 0.70), and the glass pane is 37 times thinner than the brick wall (4mm versus 150mm). The thinness matters much more than the material property, which is why even a large window is a drastically bigger heat path than the entire wall around it.

Think Like a Physicist: This is why traditional architecture in hot climates uses small windows, deep-set openings, wide overhangs, and heavy shutters, and why African vernacular buildings made of mud brick and thatch have always been cooler inside than modern glass-front imports. The physics is on the side of tradition. When architectural styles designed for cold climates (where large windows gather welcome passive solar heat) are imported into tropical cities, those same windows become unwanted heat-gain channels that work against the building's comfort. Understanding thermal resistance turns what would otherwise be cultural preference into quantified engineering: once the R-value of each element is known, the heat-flow rates through a wall can be predicted, and the wall can be designed to match the climate it stands in.

HOT Example 8

Two slabs of the same material ($k = 0.50Wm^{-1}K^{-1}$) are placed in parallel between two surfaces held at 80°C and 20°C. Each slab has cross-sectional area 0.20m². The first slab is 3.0cm thick; the second is 5.0cm thick. (a) Calculate the heat-flow rate through each slab. (b) State which slab carries more heat and explain why. (c) Find the total heat-flow rate through the parallel arrangement. (d) If the same two slabs were instead placed in series (with the same outer-face temperatures), find the new total heat-flow rate.

Solution

Given:

Conductivity (both slabs) $k = 0.50Wm^{-1}K^{-1}$

Cross-section (each slab) $A = 0.20m^2$

Thicknesses $L_1 = 3.0cm = 0.030m$, $L_2 = 5.0cm = 0.050m$

Hot face $\theta_1 = 80^\circ\text{C}$, cold face $\theta_2 = 20^\circ\text{C}$, so $\Delta\theta = 60K$

(a) Slab 1 (3.0cm thick):

$$R_1 = \frac{L_1}{kA} = \frac{0.030\text{m}}{(0.50\text{Wm}^{-1}\text{K}^{-1})(0.20\text{m}^2)} = 0.30\text{KW}^{-1}$$

Slab 2 (5.0cm thick):

$$R_2 = \frac{L_2}{kA} = \frac{0.050\text{m}}{(0.50\text{Wm}^{-1}\text{K}^{-1})(0.20\text{m}^2)} = 0.50\text{KW}^{-1}$$

In parallel, both slabs experience the same $\Delta\theta = 60\text{K}$:

$$\left(\frac{Q}{t}\right)_1 = \frac{\Delta\theta}{R_1} = \frac{60\text{K}}{0.30\text{KW}^{-1}} = \mathbf{200\text{W}} \text{ (for first slab)}$$

$$\left(\frac{Q}{t}\right)_2 = \frac{\Delta\theta}{R_2} = \frac{60\text{K}}{0.50\text{KW}^{-1}} = \mathbf{120\text{W}} \text{ (for second slab)}$$

(b) First slab carries more heat (200W versus 120W). The two slabs are made of the same material and have the same cross-section and the same temperature difference, so the only thing that differentiates them is thermal resistance. The thinner slab has a lower R, and by thermal Ohm's law a lower R carries a larger Q/t under the same $\Delta\theta$. Heat, like current, prefers the path of least resistance.

(c) Total heat-flow rate through the parallel arrangement. The resistances combine by the reciprocal-sum rule:

$$\frac{1}{R_{\text{total}}} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{0.30\text{KW}^{-1}} + \frac{1}{0.50\text{KW}^{-1}} = 5.33\text{W K}^{-1}$$

$$R_{\text{total}} = \frac{1}{5.33\text{W K}^{-1}} = 0.1875\text{KW}^{-1}$$

Applying thermal Ohm's law to the composite:

$$\left(\frac{Q}{t}\right)_{\text{total}} = \frac{\Delta\theta}{R_{\text{total}}} = \frac{60\text{K}}{0.1875\text{KW}^{-1}} = 320\text{W}$$

This matches the sum $200\text{W} + 120\text{W} = 320\text{W}$ obtained from part (a), as it must.

(d) For series:

$$R_{\text{total}} = R_1 + R_2 = 0.30\text{KW}^{-1} + 0.50\text{KW}^{-1} = 0.80\text{KW}^{-1}$$

Applying thermal Ohm's law:

$$\left(\frac{Q}{t}\right)_{\text{total}} = \frac{\Delta\theta}{R_{\text{total}}} = \frac{60\text{K}}{0.80\text{KW}^{-1}} = 75\text{W}$$

The series arrangement carries 75W, compared with 320W in parallel. The parallel arrangement carries more than four times as much heat as the series arrangement of the same two slabs.

Making Sense of the Answer: The ratio total thermal resistance in series to that in parallel $= 0.80 / 0.1875 = 4.27$ tells the whole story in a single number. For the same $\Delta\theta$, the heat current is inversely proportional to resistance, so rearranging the same two slabs from parallel to series increases the total resistance by a factor of 4.27 and decreases the heat current by the same factor. In parallel, the heat has two roads to choose from, and the roads share the load; in series, the heat is forced down one single road, with the full temperature drop spread over a longer combined path. **The parallel arrangement is always the better conductor, and the series arrangement is always the better insulator.** This is why insulation works by being thick (effectively a series arrangement of insulating layers) and why heat sinks work by being broad and finned (effectively a parallel arrangement of conducting paths).

Think Like a Physicist: When facing a composite-conductor problem, the very first move is to identify the arrangement. **Are the slabs sharing the same temperature difference but separate areas?** That is parallel: add the reciprocals of the resistances. **Are the slabs sharing the same heat flow but stacked end-to-end?** That is series: add the resistances directly. Once the arrangement is identified, the whole problem becomes an Ohm's-law calculation with thermal labels.

With composite conductors now in our toolkit, we are ready for one more geometry. So far every conductor has been flat, with parallel faces and constant cross-sectional area. But many real conductors are not flat: a hot-water pipe, a steam main, a refrigerator wall curved around a corner, a metal sphere holding hot tea. For these we need radial conduction, where the cross-sectional area changes with position. That is the subject of the next subsection.

Radial Conduction

All the geometries we have considered so far share one thing in common: the cross-sectional area available for heat flow is the same at every point along the conducting path. A slab has a constant rectangular cross-section. A lagged bar has a constant cross-section along its length. A composite wall has a constant cross-section at every depth through the wall. In each case, the area A sits comfortably in Fourier's law as a single constant that can be pulled outside any algebra we need to do.

Real heat-transfer situations do not always cooperate. A hot-water pipe carries water at 80°C and sits in a room at 20°C . Heat flows outward from the inner surface of the pipe wall to the outer surface, but the area available for that heat to cross is not constant: at the inner surface the area is small (the circumference is short), while at the outer surface the area is larger (the circumference is longer). Every molecular shell of pipe wall through which the heat must pass has a slightly bigger surface area than the shell just inside it. The same is true of a spherical metal flask of hot tea, a steam-main lagging jacket, a cooking pot whose walls curve around a corner, and an electrical cable sheathed in insulation.

In these situations, the area A in Fourier's law depends on the radius r at which we evaluate it. We can no longer treat A as a constant; we must treat it as a function of position. The result is a Fourier equation that is a differential equation, and solving it requires a small amount of calculus. The reward is a clean formula for the heat-flow rate through any curved conductor, and a natural generalisation of the thermal-resistance concept we have already built.

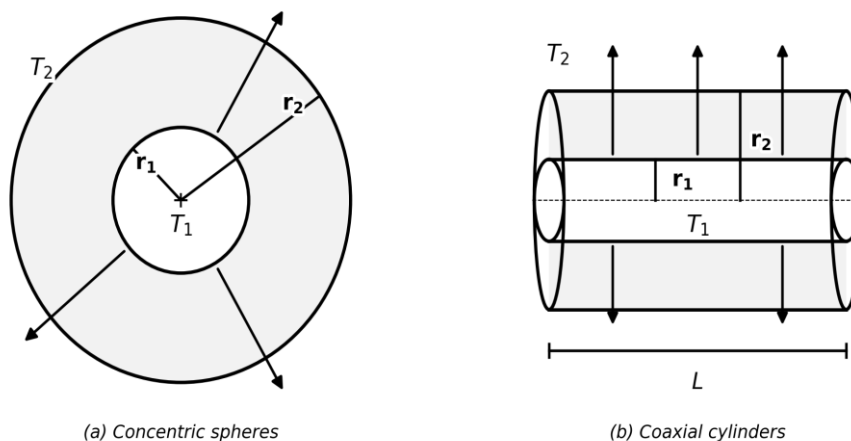


Figure: Two radial-conduction geometries. (a) Concentric spheres: the hot inner sphere is at temperature T_1 , the cold outer sphere is at T_2 , and heat flows radially outward through every spherical shell of radius r between r_1 and r_2 . (b) Coaxial cylinders of length L : the hot inner cylinder is at T_1 , the cold outer cylinder is at T_2 , and heat flows radially outward through every cylindrical shell of radius r between r_1 and r_2 .

Concentric Spheres

Consider two concentric spheres, as shown in the figure above. The inner sphere has radius r_1 and is held at the higher temperature T_1 ; the outer sphere has radius r_2 and is held at the lower temperature T_2 . Between them is a solid shell of some material of thermal conductivity k (for instance, a spherical flask of tea surrounded by a layer of lagging). In a steady state, heat flows radially outward through every concentric shell of radius r lying between r_1 and r_2 .

Consider one such shell, of inner radius r and thickness dr . This thin shell plays the role of a “slab” in Fourier's law, but with a few differences. Its area is the surface area of a sphere of radius r :

$$A(r) = 4\pi r^2$$

Its thickness is dr , and the temperature drop across it is dT (a small negative quantity, because temperature decreases as we move outward). Applying Fourier's law to this thin shell:

$$\frac{Q}{t} = -k(4\pi r^2) \frac{dT}{dr}$$

The minus sign confirms that heat flows in the direction of decreasing temperature. Because the arrangement is in a steady state, the heat-flow rate Q/t is the same through every concentric shell: all the heat that enters the inner surface of the lagging must emerge through its outer surface, with no accumulation anywhere in between. We can therefore treat Q/t as a constant during the integration and separate the variables:

$$\frac{Q}{t} \frac{dr}{r^2} = -4\pi k dT$$

Integrating from the inner surface ($r = r_1$, $T = T_1$) to the outer surface ($r = r_2$, $T = T_2$):

$$\frac{Q}{t} \int_{r_1}^{r_2} \frac{dr}{r^2} = -4\pi k \int_{T_1}^{T_2} dT$$

The integral on the left is $\int r^{-2} dr = -1/r$, and the integral on the right is straightforward:

$$\frac{Q}{t} \times \left(\frac{1}{r_1} - \frac{1}{r_2} \right) = 4\pi k (T_1 - T_2)$$

Solving for the heat-flow rate:

$$\frac{Q}{t} = \frac{4\pi k (T_1 - T_2)}{\frac{1}{r_1} - \frac{1}{r_2}}$$

This is the working formula for steady-state radial conduction through a spherical shell.

Thermal resistance of a spherical shell

Writing the result in Ohm's-law form, $Q/t = (T_1 - T_2)/R$, the thermal resistance R of the spherical shell is read off as:

$$R_{\text{sphere}} = \frac{\frac{1}{r_1} - \frac{1}{r_2}}{4\pi k}$$

Check the units: $1/r_1 - 1/r_2$ has units m^{-1} , $4\pi k$ has units $Wm^{-1}K^{-1}$, and their ratio comes out to KW^{-1} as it must. A spherical shell with small inner radius r_1 has large thermal resistance (if r_1 is small, $1/r_1$ is large), which is why *a tiny hot core loses heat slowly into a large outer volume*: the heat is constrained to pass through a small inner surface before it can fan out into the expanding outer shells.

Coaxial Cylinders

Now consider two coaxial cylinders of length L , as shown in the figure on the previous page. The inner cylinder has radius r_1 and is held at the higher temperature T_1 ; the outer cylinder has radius r_2 and is held at T_2 . Between them is a cylindrical shell of material of thermal conductivity k (for instance, the lagging around a hot-water pipe). In a steady state, heat flows radially outward through every cylindrical shell of radius r between r_1 and r_2 .

Consider a thin cylindrical shell of inner radius r and thickness dr , sharing the same axis and the same length L as the outer cylinder. Its lateral surface area (the curved surface through which heat flows radially outward) is:

$$A(r) = 2\pi rL$$

Applying Fourier's law to this thin shell, with thickness dr and temperature drop dT across it:

$$\frac{Q}{t} = -k(2\pi rL) \frac{dT}{dr}$$

Separating variables, noting that Q/t is constant throughout the steady state:

$$\frac{Q}{t} \frac{dr}{r} = -2\pi kL dT$$

Integrating from $r = r_1$ (where $T = T_1$) to $r = r_2$ (where $T = T_2$):

$$\frac{Q}{t} \int_{r_1}^{r_2} \frac{dr}{r} = -2\pi kL \int_{T_1}^{T_2} dT$$

The integral on the left is $\int (1/r) dr = \ln(r)$, so:

$$\frac{Q}{t} \ln\left(\frac{r_2}{r_1}\right) = 2\pi kL (T_1 - T_2)$$

Solving for the heat-flow rate:

$$\frac{Q}{t} = \frac{2\pi kL (T_1 - T_2)}{\ln\left(\frac{r_2}{r_1}\right)}$$

This is the working formula for steady-state radial conduction through a cylindrical shell of length L . The natural logarithm in the denominator is the signature of cylindrical geometry: it arises because the area grows linearly with radius ($A \propto r$), so the reciprocal area (which appears in the integral) integrates to a logarithm rather than an algebraic expression.

Thermal resistance of a cylindrical shell

Writing the result in Ohm's-law form, $Q/t = (T_1 - T_2)/R$, the thermal resistance of the cylindrical shell is:

$$R_{\text{cyl}} = \frac{\ln\left(\frac{r_2}{r_1}\right)}{2\pi kL}$$

Notice that R scales inversely with the length L of the pipe, as we would expect: a longer pipe has more surface area, more cylindrical shells in parallel along its length, and therefore lower resistance to radial heat flow. The logarithm in the numerator means that doubling the outer radius does not double the resistance; it only adds a small amount ($\ln 2 \approx 0.69$) to whatever logarithm was already there. Thick lagging is subject to diminishing returns.

Summary of the Three Geometries

We have now derived Fourier's law for three different geometries: the uniform slab (from Section 3.1.1), the spherical shell, and the cylindrical shell. The three formulas are collected below for reference.

Geometry	Thermal resistance R	Heat-flow rate Q/t
Uniform slab (thickness L , area A)	$\frac{L}{kA}$	$\frac{kA(T_1 - T_2)}{L}$
Spherical shell (radii r_1, r_2)	$\frac{\frac{1}{r_1} - \frac{1}{r_2}}{4\pi k}$	$\frac{Q}{t} = \frac{4\pi k (T_1 - T_2)}{\frac{1}{r_1} - \frac{1}{r_2}}$
Cylindrical shell (length L , radii r_1, r_2)	$\frac{\ln\left(\frac{r_2}{r_1}\right)}{2\pi kL}$	$\frac{2\pi kL (T_1 - T_2)}{\ln\left(\frac{r_2}{r_1}\right)}$

The essential point is that in each case, the heat-flow rate can be written as (temperature difference) divided by (thermal resistance). The thermal resistance depends only on the geometry and the thermal conductivity of the material; it does not depend on the temperature difference. The Ohm's-law analogy survives the change of geometry unharmed.

Two curves, two calculations. Bring your calculator and pen.

BINDER Example 9

A hot-water pipe of inner radius 2.0cm and outer radius 2.5cm carries water at 80°C. The outer surface of the pipe is at 30°C. The pipe is made of copper ($k = 380 \text{ W m}^{-1} \text{ K}^{-1}$) and has length 5.0m. Calculate the rate at which heat is conducted through the pipe wall.

Solution

This is the radial conduction through a cylindrical shell problem.

Thermal resistance of the cylindrical shell.

$$R_{\text{cyl}} = \frac{\ln\left(\frac{r_2}{r_1}\right)}{2\pi kL} = \frac{\ln(0.025/0.020)}{2\pi (380\text{Wm}^{-1}\text{K}^{-1})(5.0\text{m})} = 1.87 \times 10^{-5} \text{KW}^{-1}$$

Applying thermal Ohm's law:

$$\frac{Q}{t} = \frac{\Delta T}{R_{\text{cyl}}} = \frac{50\text{K}}{1.87 \times 10^{-5} \text{KW}^{-1}} = 2.68 \times 10^6 \text{W}$$

The rate of heat conduction through the copper pipe wall is $2.68 \times 10^6 \text{W}$ (about 2.7MW, or 2700Kw).

Making Sense of the Answer: A staggering answer: over two megawatts of heat flow through a 5-metre length of copper pipe. The number is large because copper is an excellent conductor ($k = 380$, one of the highest in the table) and the pipe wall is very thin (5mm of copper between hot water and cool outer surface). In practice a bare copper pipe would never maintain a 50K temperature difference across its wall for long: the outer surface would reach the water temperature within seconds, and the real heat loss would then be determined by convection from the outer surface to the room air, not by conduction through the copper wall. This is why hot-water pipes are always wrapped in lagging: the lagging is the actual bottleneck.

Think Like a Physicist: When computing radial conduction through a thin-walled metal conductor, expect a huge answer. The interesting physics is almost never the metal wall itself but what surrounds it. A good metal with a thin wall offers almost no resistance to heat flow, so its thermal resistance is tiny and the heat-flow rate it would carry under any realistic temperature difference is enormous. In any real installation, the temperature difference across the metal wall collapses to almost nothing within seconds, and the actual heat loss is then fixed by whatever sits outside the metal: lagging, paint, convection to the surrounding air. Whenever a problem gives you a metal-wall radial calculation and nothing else, the number you compute is a ceiling on the heat flow, not the heat flow a real pipe would actually deliver.

HOT Example 10

An *alkasusu* seller at **Forodhani Gardens** in **Zanzibar** keeps his *alkasusu* hot in a spherical metal flask of inner radius 10cm and outer radius 11cm, made of steel ($k = 50\text{Wm}^{-1}\text{K}^{-1}$). *Alkasusu* inside is at 85°C ; the outer surface of the flask is at 25°C . (a) Calculate the rate of heat loss through the flask wall. (b) The seller wraps the flask in a 2.0cm-thick layer of felt ($k = 0.040\text{Wm}^{-1}\text{K}^{-1}$), so the outer radius of the flask-plus-felt system becomes 13cm. Assuming the tea and the outer felt surface are still at 85°C and 25°C respectively, calculate the new rate of heat loss. (c) Comment on the ratio of the two heat-loss rates, and on which layer dominates the thermal resistance of the insulated system.

Solution

This is the radial conduction through concentric spheres problem. Part (a) uses a single spherical shell; part (b) uses two spherical shells in series.

(a) Rate of heat loss through the bare steel flask.

Thermal resistance of the steel shell:

$$R_{\text{steel}} = \frac{\frac{1}{r_1} - \frac{1}{r_2}}{4\pi k_{\text{steel}}} = \frac{\frac{1}{0.10\text{m}} - \frac{1}{0.11\text{m}}}{4\pi (50\text{Wm}^{-1}\text{K}^{-1})} = \frac{0.909\text{m}^{-1}}{628.3\text{Wm}^{-1}\text{K}^{-1}} = 1.45 \times 10^{-3} \text{KW}^{-1}$$

Rate of heat loss through the bare flask:

$$\left(\frac{Q}{t}\right)_{\text{bare}} = \frac{\Delta T}{R_{\text{steel}}} = \frac{60\text{K}}{1.45 \times 10^{-3} \text{KW}^{-1}} = 41400\text{W}$$

The bare flask would lose heat at 41400W (about 41kW).

(b) Rate of heat loss through the flask + felt lagging. The two spherical shells are in series (the same heat crosses both in turn), so the total thermal resistance is the sum of the individual resistances.

Thermal resistance of the felt shell:

$$R_{\text{felt}} = \frac{1}{4\pi k_{\text{felt}}} \left(\frac{1}{r_2} - \frac{1}{r_3} \right) = \frac{1}{4\pi (0.040 \text{ W m}^{-1} \text{ K}^{-1})} \left(\frac{1}{0.11 \text{ m}} - \frac{1}{0.13 \text{ m}} \right) = \frac{1.399 \text{ m}^{-1}}{0.5027 \text{ W m}^{-1} \text{ K}^{-1}} = 2.783 \text{ KW}^{-1}$$

Total thermal resistance of the two shells in series:

$$R_{\text{total}} = R_{\text{steel}} + R_{\text{felt}} = 1.45 \times 10^{-3} \text{ KW}^{-1} + 2.783 \text{ KW}^{-1} = 2.784 \text{ KW}^{-1}$$

Rate of heat loss through the lagged flask:

$$\left(\frac{Q}{t} \right)_{\text{lagged}} = \frac{\Delta T}{R_{\text{total}}} = \frac{60 \text{ K}}{2.784 \text{ KW}^{-1}} = 21.6 \text{ W}$$

The lagged flask loses heat at only 22W.

(c) Ratio of the two heat-loss rates:

$$\frac{\left(\frac{Q}{t} \right)_{\text{bare}}}{\left(\frac{Q}{t} \right)_{\text{lagged}}} = \frac{41400 \text{ W}}{21.6 \text{ W}} \approx 1900$$

The bare flask loses heat roughly 1900 times faster than the lagged flask.

Of the total thermal resistance of the lagged system, 2.783 out of 2.784 KW⁻¹ (99.95%) comes from the felt alone. The steel flask wall, despite being made of a mediocre conductor (not as good as copper, but still a metal), contributes only 0.05% of the total resistance. Essentially all the insulation is provided by the felt.

Making Sense of the Answer: The factor of 1900 is what makes tea flasks worth buying. Two centimetres of cheap felt converts a vessel that would cool in minutes into one that stays warm for hours. The physics behind this is the enormous conductivity ratio between steel and felt: $k_{\text{steel}}/k_{\text{felt}} = 50/0.040 = 1250$. Steel is 1250 times better at conducting heat than felt, and in series arrangement the resistance is dominated by whichever layer has the higher R . For two shells of comparable thickness, the one with the lower k wins the resistance battle by exactly the ratio of conductivities, which is why the felt controls everything even though it is only twice as thick as the steel.

Think Like a Physicist: This example illustrates a universal principle of heat-transfer engineering: whenever you have a layered system, the total resistance is dominated by the layer with the highest R . Adding more metal to the steel flask would barely help. Adding more felt to the lagging would help enormously. Real-world insulation design is therefore the art of identifying the weakest link in the series (the one with lowest R) and improving that one, rather than spending effort on layers that already dominate. The physics decides where your money should go.

With radial conduction now added to our toolkit, we have completed the study of heat conduction in its three main geometries: slab, cylinder, and sphere. But conduction is only one of the three heat-transfer mechanisms introduced at the start of the chapter. The next section takes up the second: convection, in which a moving fluid carries heat bodily from place to place, and in which the mechanism of transport is no longer the passing of molecular vibrations from atom to atom but the bulk displacement of a heated material itself.

THERMAL CONVECTION

When Kipanga grabbed the steel spoon in the kitchen and felt the handle grow hot, he was feeling conduction: heat passing atom by atom through a solid that did not itself move. Now leave the kitchen for a moment and look up. The warmth you feel rising from a *sufuria* boiling on a stove, the breeze that pushes in from the ocean during a hot afternoon, the plume of warm air above a parked minibus in midday sun, even the rolling motion of water as it heats inside a kettle: all of these are examples of a different heat-transfer mechanism, one in which the hot material itself physically carries the energy from one place to another. We call it **convection**. By definition:

Thermal convection is the process by which heat is transferred through a fluid (liquid or gas) by the actual movement of the fluid material itself from a region of higher temperature to a region of lower temperature.

Convection is fundamentally different from conduction in one important way. In **conduction**, the material stays where it is and energy migrates through it as a wave of molecular jostling. In **convection**, the material itself goes on a journey, and the energy rides along like a passenger. You could blindfold a cup of water, put a lump of ice in one corner and a drop of heat from a match in the other, and track the water molecules

as they moved: in conduction, almost none of them would move far from where they started. In convection, you would see them flowing around the cup in bulk, mixing hotter fluid with cooler fluid until the whole body came to a uniform temperature.

This distinction has an immediate consequence: *convection can only happen in materials whose molecules are free to move in bulk*. **Liquids** and **gases** qualify; solids do not. A metal bar can conduct heat but **cannot** convect it, *because its atoms are locked to their lattice sites*. A pot of water can both conduct and convect, but convection dominates quickly once the water starts moving. The air in a room above a radiator can barely conduct heat at all (air is one of the worst conductors in the k-table provided earlier), yet the room still warms throughout, because convection takes over the job.

Two Types of Convection

Convection currents can be set going by either of two different driving mechanisms, and it is important to keep them straight because the laws that govern the heat flow differ between them.

1) Natural convection

In **natural convection**, *the fluid moves because of density differences caused by uneven heating*. A pocket of fluid that has been warmed expands, becomes less dense than its surroundings, and is pushed upward by the buoyancy force: the same Archimedes effect that makes a cork float on water. At the top it eventually cools, becomes denser again, and sinks back down, displacing other warmed fluid as it falls. The result is a closed convection current, driven entirely by gravity acting on density gradients that the heating itself has created.

Every domestic example of natural heating works this way. A *sufuria* of water on a charcoal *jiko* shows convection currents in miniature, with hot water rising from the hot bottom of the pan and cooler water falling to take its place from the top. A room heated by a radiator or a fireplace does the same thing on a larger scale, as shown in the figure below.

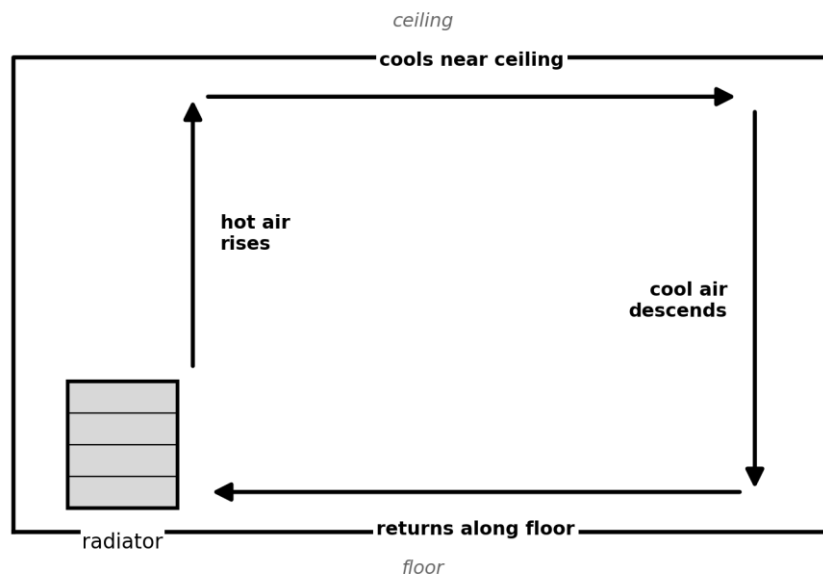


Figure: Natural convection in a room heated by a radiator. Hot air rises from the radiator, travels across the ceiling (cooling as it goes by losing heat to the ceiling and walls), descends along the far wall, and returns along the floor to the radiator. The loop is driven entirely by the density difference between warm and cool air.

On a much larger scale, the same principle drives the sea-land breeze cycle familiar to anyone who has lived along the Tanzanian coast. During a hot afternoon, the land heats faster than the ocean (because land has a much smaller specific heat capacity), warming the air above it. That warm air rises; cooler air over the ocean flows in at ground level to replace it; the result is the steady afternoon breeze from sea to land that keeps Dar es Salaam merely hot rather than intolerable. At night the effect reverses: the land cools faster than the ocean, the ocean now holds the warmer air, and the breeze blows from land out to sea. Trade winds, monsoons, and thunderstorm clouds are all driven by the same natural-convection mechanism operating on progressively larger scales.

2) Forced convection

In **forced convection**, the fluid is driven by an external agent that imposes motion on it regardless of whether there is any temperature-induced density gradient. A fan heater blows air past a hot element. A car radiator pumps coolant through metal channels past a hot engine block. The wind blows across a sweating face. A convection oven uses a built-in fan to circulate hot air past the food, cooking it more evenly than a still oven would. In every case an external force has taken over the job that gravity was doing in natural convection, and the fluid moves even where buoyancy would not carry it.

Forced convection is generally more effective at transferring heat than natural convection, because it is not limited by whatever modest density differences thermal expansion can create. A fan heater warms a room much faster than a radiator of the same power, because the fan forces air past the element at a speed far greater than natural buoyancy alone could achieve. This is why sweating cools the body more effectively in a breeze than in still air, and why hospitals use forced-air ventilation rather than trusting to natural circulation.

Factors Affecting the Rate of Convection

The rate at which a body loses heat by convection depends on three things. Any problem about convection can be approached by checking each of these three factors in turn.

Factor	What it means	Tanzanian example
Temperature difference between body and fluid	A larger excess temperature drives a steeper density gradient in natural convection and a larger heat flux in either mode.	A hot <i>sufuria</i> cools faster when the kitchen is cold than when it is already warm.
Properties of the fluid (viscosity, density, specific heat)	A low-viscosity, low-density, high-specific-heat fluid carries heat away more effectively per unit volume of flow.	A pot of tea cools faster in wind than in still water of the same temperature because air is less viscous and flows more freely.
Whether the convection is natural or forced	Forced convection is nearly always faster for a given temperature difference, because the fluid speed is not limited by buoyancy alone.	A person standing by an open minibus window on a moving <i>daladala</i> cools faster than a person sitting in a parked, windless minibus.

Dulong-Petit Law

When a hot body cools freely in still air (natural convection only) and the excess temperature above the surroundings is large, the rate of heat loss is found experimentally not to be proportional to the excess temperature itself but to a fractional power of it. The law was established by the French physicists Pierre Dulong and Alexis Petit in the early nineteenth century, and it states:

Under natural convection, the rate of heat loss from a hot body is directly proportional to the 5/4 power of the excess temperature of the body over its surroundings.

Mathematically:

$$\frac{dQ}{dt} \propto (\theta - \theta_s)^{5/4}$$

where θ is the temperature of the body and θ_s is the temperature of the surroundings. The fractional exponent 5/4 is a slightly steeper-than-linear dependence: doubling the excess temperature more than doubles the rate of heat loss (it multiplies it by $2^{5/4} \approx 2.38$).

The 5/4 power arises because natural convection depends on the strength of the convection current that the hot body creates, and that current is itself driven by the density difference between the warmed fluid near the body and the cool fluid far from it. A hotter body produces a larger density difference, hence a faster convection current, hence a more effective heat transfer. The precise exponent of 5/4 is an empirical result fitted to careful experiments; the important qualitative fact is that natural convection is a faster-than-linear process.

Dulong-Petit applies to the **natural-convection** regime with **large** excess temperatures (typically above about 50°C). It does **not** apply for **smaller excess temperatures**, and for the whole of the **forced-convection** regime at any excess temperature.

Newton's Law of Cooling

When a hot body cools either under **forced convection** (at **any** excess temperature) or under **natural convection** with a **small** excess (typically below about 30°C), the rate of heat loss is found to be directly proportional to the excess temperature itself. This was first stated by Isaac Newton in 1701 and has become one of the most-used laws in heat-transfer physics. The law states:

Under forced convection (at any excess temperature) or natural convection with a small excess temperature (below about 30°C), the rate of heat loss from a hot body is directly proportional to the excess temperature of the body over its surroundings.

Mathematically:

$$\frac{dQ}{dt} \propto (\theta - \theta_s)$$

The constant of proportionality depends on the body and its surroundings (its surface area, the fluid in contact with it, and whether the convection is natural or forced). Writing this constant as k , Newton's law becomes:

$$\frac{dQ}{dt} = k(\theta - \theta_s)$$

The full derivation of the corresponding temperature-versus-time expression, and the integration that predicts how long a hot cup of tea takes to cool, will be covered after the next section, which is devoted entirely to Newton's law and its consequences. For the purposes of this section it is enough to state the law, recognise its regime of validity, and understand why it sits alongside Dulong-Petit as one of two different laws for two different regimes of convection.

When does each law apply?

Knowing which of the two laws to use in a given situation is itself a piece of physics. The quick guide below captures the essentials.

Situation	Which law applies?
Forced convection, any excess temperature	Newton's law (linear in $\Delta\theta$)
Natural convection, small excess (roughly < 30°C)	Newton's law (linear in $\Delta\theta$)
Natural convection, large excess (roughly > 50°C)	Dulong-Petit law (proportional to $\Delta\theta^{5/4}$)
Natural convection, moderate excess (30°C to 50°C)	Transitional; neither law is exact

The thresholds given above are approximate and depend on the fluid involved, the geometry of the hot body, and the ambient conditions. In an A-level problem, assume Newton's law unless the excess temperature is very large and the convection is explicitly described as natural (no fan, no wind, no pump). If the problem mentions a fan or a moving fluid, assume Newton's law at any excess temperature.

Five examples. Three to warm up the concepts, two to drill the laws. Pens at the ready.

BINDER Example 11

A domestic iron radiator is installed on one wall of a room and heats the air around it. No fan is running. (a) Trace the complete convection-current loop that the radiator creates in the room, explaining what drives each segment of the loop. (b) Why does the upper part of the room warm up before the lower part? (c) How would the introduction of a small electric fan change the heating pattern?

Solution

(a) The radiator surface is hot and heats the air that touches it. The heated air expands, becomes less dense than the cooler air around it, and is pushed upward by buoyancy. This rising stream of hot air travels upward directly above the radiator until it reaches the ceiling. Along the way and at the ceiling, the air loses heat to the ceiling surface and the cooler surrounding air, becoming denser again. The now-cooler air moves along the ceiling away from the radiator, toward the far wall of the room. At the far wall, the air has cooled enough to become denser than the air below it; it descends down the far wall. Along the floor, this cool air is drawn back toward the radiator (partly by being displaced by more cool air coming down, partly by being pulled into the low-pressure region below the rising plume). When it reaches the radiator, it is heated again and the cycle repeats. The full loop is driven entirely by gravity acting on the density differences that the heating itself has created.

(b) The ceiling (upper part) warms before the floor (lower part) because the hot air rises first. Before the return flow along the floor has had time to bring cool air back to the radiator, heat has already collected at the ceiling. Also, the warm air is continuously replenished at the top of the room and continuously removed at the bottom, so the ceiling receives a steady delivery of fresh hot air while the floor receives a steady delivery of descending cool air.

(c) An electric fan converts the situation from natural convection to forced convection. The fan moves air at a speed much greater than buoyancy alone could produce, breaking up the orderly vertical temperature layering. Instead of the ceiling warming first and the floor last, the entire room is stirred into a roughly uniform temperature.

This is why ceiling fans are installed in Tanzanian homes: not to cool the air (they cannot change the air's temperature), but to mix it and to accelerate heat loss from the skin, which makes the body feel cooler.

Making Sense of the Answer: *The convection-loop trace is a chain of cause and effect, and each segment is driven by the result of the previous one. Hot air rises because it was heated. It cools because it travels along cooler surfaces. It descends because it became denser. It returns because it was drawn into the low-pressure zone left behind by the next rising plume. Every step is a consequence of the previous step. Any disruption to the chain (blocking the ceiling airflow with a curtain rail, obstructing the floor return with furniture) slows the whole loop.*

Think Like a Physicist: *When asked to trace a convection loop, do not simply say “hot air rises” and stop there. The full loop has four segments, and each segment has its own physical cause. A complete answer names all four: (1) the rising plume (driven by buoyancy), (2) the sideways flow near the ceiling (driven by displacement), (3) the descending flow along the cool wall (driven by increased density), and (4) the return flow along the floor (driven by continuity of flow). A one-segment answer is a one-segment understanding.*

REAL Example 12

During a hot afternoon in Mtwara, a breeze blows steadily from the ocean inland. At night the breeze reverses and blows from the land out to sea. (a) Explain both observations using the concepts of specific heat capacity and natural convection. (b) On a completely overcast day, when the sun is blocked and the temperatures of land and ocean stay roughly equal, there is essentially no breeze at all from either direction. Explain this observation.

Solution

(a) During the day, the sun heats both the land and the ocean. But land has a much smaller specific heat capacity than water (roughly $800\text{Jkg}^{-1}\text{K}^{-1}$ for soil and rock versus $4200\text{Jkg}^{-1}\text{K}^{-1}$ for water). Under the same solar input, the land temperature rises rapidly while the ocean temperature rises barely at all. By midday, the air over the hot land has been heated well above the temperature of the air over the cool ocean. The hot land-air is less dense; it rises. To replace it, cooler air flows in at ground level from the direction where the cool air is found, which is from the ocean. This is the sea breeze that makes Mtwara afternoons bearable.

At night, the sun no longer supplies heat, and both the land and the ocean cool by radiating heat into the sky. Again land has the smaller specific heat capacity, so it loses its daytime heat much faster than the ocean, which stays warm because it has so much stored thermal energy. By late night, the ocean is now warmer than the land. The air over the ocean is now the less-dense air; it rises. Cool air flows in at ground level from the now-cooler direction, which is from the land. This is the land breeze, blowing out to sea through the night and the early morning hours, which is why fishing dhows traditionally leave before dawn.

(b) On a completely overcast day, the sunlight is blocked by clouds and neither the land nor the ocean heats up significantly. With both surfaces at roughly the same temperature, the air above them is at roughly the

same temperature everywhere, with no temperature difference to drive a convection current. Without a temperature difference there is no density gradient; without a density gradient there is no buoyant force to push air upward over one region; without such a force there is no flow in at ground level to replace the rising air. The atmosphere over land and ocean simply sits there motionlessly, and the breeze that would normally be driven by convection is absent. This is why overcast coastal days often feel strangely still and unpleasant, even at ocean-side locations that are normally windy.

Making Sense of the Answer: *The sea-land breeze system is a natural convection current on a continental scale, and every feature of it is explained by the same physics that governs a *sufuria* on a charcoal jiko. A temperature difference creates a density gradient; the density gradient drives a flow; the flow is completed into a closed loop by the continuity of the atmosphere. The system is larger than a kitchen but the physics is the same. This is also why “onshore flow” (sea breeze) is used by meteorologists as a reliable indicator that the land has been warming faster than the ocean that day, which itself is a standard marker of clear-sky insolation.*

Think Like a Physicist: *Whenever a real-world atmospheric phenomenon is described as a “wind” or “breeze”, ask yourself what temperature difference is driving it. Almost every steady local wind on Earth (sea breezes, valley winds, monsoons, trade winds) has a temperature difference at its heart, and that temperature difference is almost always caused by differential heating of surfaces with different specific heat capacities. Identifying the temperature difference is the first step to understanding any wind system.*

BINDER Example 13

State the three factors that affect the rate of thermal convection from a hot body, and give a real-world Tanzanian example of each.

Solution

Factor 1: the temperature difference between the hot body and the surrounding fluid. A larger temperature difference drives a larger heat-flow rate by convection.

Example: a freshly boiled kettle (100°C) sitting in a 25°C kitchen loses heat much faster than the same kettle once it has cooled to 40°C. The first minutes of cooling feel urgently hot, while the last half hour feels barely warm; this is the temperature-difference factor at work.

Factor 2: the properties of the fluid in contact with the body, including its viscosity, density, and specific heat capacity.

Example: a hot *sufuria* placed outdoors on a windy day cools faster than the same *sufuria* placed in a bathtub of still water at the same air temperature, even though water has a much higher specific heat capacity than air. The reason is that air, with its low viscosity and low density, can be set moving far more easily than water; the convection current over the *sufuria* is much stronger in air than in still water, and the stronger current removes heat faster. Moving fluids beat still fluids, and the fluid’s willingness to move is the determining factor.

Factor 3: whether the convection is natural or forced. Forced convection is generally much faster than natural convection.

Example: a bus conductor standing at an open window of a moving bus cools faster than a passenger sitting in the same bus with the windows closed. In the first case, the external motion of the bus creates a forced-convection flow of air across the conductor’s skin, quickly sweeping away the warm boundary layer. In the second case, only natural convection is at work, and the boundary layer of warm air tends to cling to the skin, slowing heat loss.

Making Sense of the Answer: *The three factors are not independent of each other; in practice they combine. A large temperature difference in a moving fluid under forced convection produces the most rapid heat loss; a small temperature difference in a stagnant fluid under natural convection produces the slowest. Most everyday cooling situations involve some combination of all three, and estimating the cooling rate requires checking all three factors together.*

Think Like a Physicist: *A useful way to remember the three factors is as a triangle: (1) the temperature difference is the driving force, (2) the fluid is the medium through which the heat must be carried, and (3) whether the flow is natural or forced is the delivery mechanism. Change any one of the three and the cooling rate changes; but always check all three before making a prediction. An answer that mentions only the temperature difference and ignores the fluid or the convection mode is an incomplete answer.*

BINDER Example 14

A hot plate at 200°C cools freely in still air at 25°C. Under these conditions the initial rate of heat loss is measured as 500W. Assuming that the cooling continues to obey the appropriate physical law throughout, estimate the rate of heat loss when the plate has cooled to 100°C.

Solution

The plate is cooling **freely** in **still air** (natural convection) and the initial excess temperature is 175K (200°C – 25°C), well above 50°C, so the **Dulong-Petit law** governs the cooling.

Interpreting the data:

Surrounding temperature $\theta_s = 25^\circ\text{C}$

Initial plate temperature $\theta_1 = 200^\circ\text{C}$, so initial excess $\Delta\theta_1 = 200^\circ\text{C} - 25^\circ\text{C} = 175^\circ\text{C} = 175\text{K}$

Later plate temperature $\theta_2 = 100^\circ\text{C}$, so later excess $\Delta\theta_2 = 100^\circ\text{C} - 25^\circ\text{C} = 75^\circ\text{C} = 75\text{K}$

Initial rate $\left(\frac{dQ}{dt}\right)_1 = 500\text{W}$

Dulong-Petit law at two different excess temperatures:

$$\left(\frac{dQ}{dt}\right)_1 \propto (\Delta\theta_1)^{5/4}$$

$$\left(\frac{dQ}{dt}\right)_2 \propto (\Delta\theta_2)^{5/4}$$

Dividing one by the other cancels the proportionality constant:

$$\frac{\left(\frac{dQ}{dt}\right)_2}{\left(\frac{dQ}{dt}\right)_1} = \left(\frac{\Delta\theta_2}{\Delta\theta_1}\right)^{5/4}$$

Substituting the numerical values:

$$\frac{\left(\frac{dQ}{dt}\right)_2}{500\text{W}} = \left(\frac{75\text{K}}{175\text{K}}\right)^{5/4} = (0.4286)^{5/4}$$

$$\left(\frac{dQ}{dt}\right)_2 = 500\text{W} \times (0.4286)^{5/4} = 173\text{W}$$

The rate of heat loss when the plate has cooled to 100°C is about 173W.

Making Sense of the Answer: The excess temperature has fallen by a factor of $175/75 = 2.33$. Under a simple linear (Newton) law, the rate of heat loss would also fall by that same factor, to $500/2.33 = 215\text{W}$. Under Dulong-Petit, with the steeper $5/4$ exponent, the rate falls by a slightly larger factor of $2.33^{5/4} = 2.89$, to 173W. The difference of 215 vs 173W is a factor of about 1.24, which is the ratio of $2.33^{1/4} \approx 1.24$. This factor is small, but it is systematic: Dulong-Petit always predicts a somewhat faster cooling than Newton does, because the $5/4$ exponent amplifies every excess temperature non-linearly.

Think Like a Physicist: When two physical laws both describe the same phenomenon in overlapping regimes, it is worth computing the same problem under both laws and comparing the answers. The ratio of the two predictions tells you how much the choice of law matters in the specific situation at hand. For large excess temperatures the two laws diverge measurably; for small excess temperatures they converge, which is exactly why Newton's law works for small excesses even though the true underlying physics is closer to Dulong-Petit.

BINDER Example 15

Three hot bodies are cooling in a laboratory at 20°C. Body A is a cup of tea at 90°C sitting still on a table. Body B is the same cup of tea placed directly in front of a running desk fan. Body C is a bowl of soup at 35°C sitting still on the same table. For each body, decide whether its cooling is governed primarily by Newton's law of cooling or by the Dulong-Petit law, and justify your answer. No numerical calculation is needed.

Solution

Regime selection rests on two questions: is the convection natural or forced, and is the excess temperature large or small?

Body A: cup of tea at 90°C , no fan, excess temperature $= 90 - 20 = 70\text{K}$. This is natural convection (no external fluid driver) with a large excess temperature (well above the 50K threshold beyond which Dulong-Petit is the better description). Therefore, Dulong-Petit applies.

Body B: same cup of tea at 90°C , but with a fan blowing across it. The fan is an external agent driving the air past the tea, so the convection is forced regardless of the temperature difference. Under forced convection, Newton's law applies at any excess temperature. Therefore, Newton applies.

Body C: bowl of soup at 35°C , no fan, excess temperature $= 35 - 20 = 15\text{K}$. This is natural convection with a small excess temperature (well below the 30K threshold above which Dulong-Petit starts to take over). Therefore, Newton applies.

Making Sense of the Answer: *Of the three bodies, two obey Newton and only one obeys Dulong-Petit, which is a fair reflection of everyday life: most cooling situations involve either some forced airflow (a fan, a breeze, a breath) or a small excess temperature, and in both cases Newton's simpler linear law is the right tool. Dulong-Petit shines specifically for hot objects cooling in still air, which is less common in practice than it might seem. A laboratory setup with a hot plate on a bench and no air movement is one of the few situations where the $5/4$ power genuinely describes the cooling better than a linear law would.*

Think Like a Physicist: *The regime-selection question is often the most important physical decision in a convection problem, and it should be made before any calculation begins. A student who plugs numbers into the wrong law gets a wrong answer no matter how carefully the arithmetic is done. Checking the two questions (forced or natural? large or small excess?) takes only a moment, and it settles which equation to use for the rest of the problem. Asking the right question first is half of solving it right.*

Thermal convection has now introduced the second of the three heat-transfer mechanisms, and has given us two laws (Dulong-Petit and Newton's) to describe the rate of heat loss by moving fluids. The detailed integration of Newton's law, which lets us predict how the temperature of a cooling body changes with time, is the subject of another section. Before that, however, we turn to the third mechanism: radiation, in which heat travels as electromagnetic waves and requires no medium at all. The next section takes up thermal radiation, from the glow of a charcoal *jiko* to the light of the Sun itself.

THERMAL RADIATION

Conduction transfers heat through a material by molecular jostling. Convection transfers heat by bodily moving the material itself. Both mechanisms require a material to be present. But the warmth of the Sun reaches us across 150 million kilometres of near-perfect vacuum, where there is no material at all. A third heat-transfer mechanism must exist, one that needs no medium and can carry energy through empty space. We call it **thermal radiation**.

Thermal radiation is energy carried by electromagnetic waves, predominantly in the infrared part of the spectrum. Every body whose temperature is above absolute zero emits thermal radiation continuously, regardless of its surroundings: a hot charcoal *jiko* radiates toward your face, but your face also radiates back toward the *jiko*. A warm Dar es Salaam pavement radiates toward the sky, and the sky radiates back toward the pavement. The only difference between a hot body and a cold one is how much radiation each emits per second. In this section we will learn exactly how much, and how the radiation depends on the body's temperature, its surface properties, and the temperature of its surroundings.

Blackbody Radiation and Emissivity

Some surfaces are better emitters and absorbers of radiation than others. A dull, dark, rough surface absorbs nearly all the radiation that falls on it and re-emits it efficiently; a shiny, polished, silvery surface reflects nearly all incoming radiation and emits very little of its own. To set a baseline for comparison, physicists define an idealised surface called a **blackbody**.

A **blackbody** is a body that absorbs all radiation incident upon it, regardless of wavelength or angle of incidence. Because it absorbs everything, it reflects nothing; and because it absorbs so efficiently, it also re-emits more efficiently than any other body at the same temperature. A **blackbody** is therefore both a perfect absorber and a perfect emitter.

No real surface is a perfect blackbody, but some come close. Soot, lampblack, and the interior of a furnace with a tiny observation hole are all excellent approximations. At the other extreme, a polished silver mirror is a very poor absorber and emitter.

To quantify how close a real surface comes to being a blackbody, we define a number called **emissivity**, given the symbol ϵ (the Greek letter epsilon). The **emissivity of a surface** is the ratio of the power it actually radiates to the power a perfect blackbody of the same area and temperature would radiate. A **blackbody** has $\epsilon = 1$; a **perfect reflector** has $\epsilon = 0$. Real surfaces fall between these two extremes.

Surface	Approximate emissivity ϵ
Perfect blackbody	1.00
Lampblack, soot	0.95
Dark rough paint	0.90
Human skin	0.80
Red brick, concrete	0.90
Oxidised copper	0.65
Polished copper	0.04
Polished aluminium	0.05
Polished silver	0.02

A useful rule of thumb emerges from this table: *good absorbers are good emitters, and good reflectors are poor emitters.* This is not a coincidence; it is a deep consequence of the microscopic physics of how electromagnetic waves interact with matter. *A surface that is efficient at absorbing a given wavelength of radiation must, by the same mechanism, be efficient at emitting that wavelength.* This principle, known as **Kirchhoff's law of thermal radiation**, is the reason why the inside of a thermos flask is silvered: the low emissivity of the silver surface reduces radiative heat loss, and the same low absorptivity reduces radiative heat gain. The silver acts as a radiation barrier in both directions.

Stefan-Boltzmann Law

The total power radiated (amount of radiation per unit time) by a body depends on its **temperature**, its **surface area**, and its **emissivity**. The quantitative relationship was discovered experimentally by Josef Stefan in 1879 and derived theoretically by Ludwig Boltzmann in 1884. The result is one of the most important equations in thermal physics.

Stefan-Boltzmann law states that:

The total power radiated by a body is proportional to the fourth power of its absolute temperature, to its surface area, and to its emissivity.

Mathematically:

$$P = \epsilon \sigma A T^4$$

where P is the radiated power in watts, ϵ is the emissivity (dimensionless, between 0 and 1), σ is the Stefan-Boltzmann constant $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$, A is the surface area in m^2 , and T is the absolute temperature in kelvin. The temperature must be in kelvin because the law involves an absolute power; the Celsius scale would give a meaningless result.

The fourth-power dependence is extremely steep. Doubling the temperature increases the radiated power by a factor of 16. Tripling the temperature increases it by a factor of 81. This is why a small increase in temperature produces a dramatically visible increase in brightness: a charcoal jiko at 800K glows dull red, but at 1600K (only twice the temperature) it glows blindingly white, radiating 16 times as much energy.

Net radiative power

A body does not radiate into nothing. It sits in an environment (surroundings, walls, air) that is itself at some temperature T_s and is itself radiating back toward the body. The net power lost by the body is the difference between what it emits and what it absorbs:

$$P_{\text{net}} = \epsilon \sigma A (T^4 - T_s^4)$$

If the body is hotter than its surroundings ($T > T_s$), the net power is positive and the body loses energy by radiation. If the body is cooler ($T < T_s$), the net power is negative and the body gains energy from its surroundings. At thermal equilibrium ($T = T_s$), the net power is zero: the body emits and absorbs at exactly the same rate.

Wien's Displacement Law

The Stefan-Boltzmann law tells us how much total power a body radiates, but it does not tell us how that radiation is distributed across different wavelengths. A glowing charcoal *jiko* looks reddish; the Sun looks yellowish-white; a welding arc looks bluish-white. As the temperature rises, the dominant colour shifts from red through yellow to blue-white. This shift is quantified by **Wien's displacement law**. It states that:

The wavelength at which a blackbody emits maximum spectral power is inversely proportional to its absolute temperature.

Mathematically:

$$\lambda_{\text{max}} = \frac{\text{constant } b}{T} \text{ or } \lambda_{\text{max}} T = b$$

where λ_{max} is the peak wavelength in metres, T is the absolute temperature in kelvin, and b is Wien's displacement constant, $b = 2.90 \times 10^{-3} \text{ m K}$. The product $\lambda_{\text{max}} T$ is the same for every blackbody regardless of temperature: a hotter body shifts its peak to shorter wavelength so that the product stays constant.

The displacement law is beautifully illustrated by the spectral emission curves shown in the following figure.

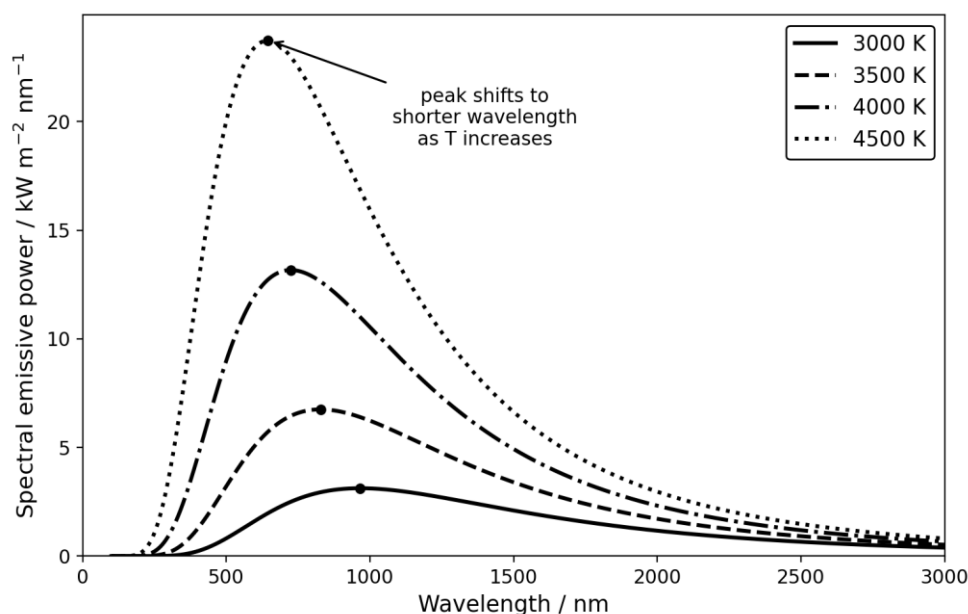


Figure: Spectral emission curves for blackbodies at four temperatures from 3000K to 4500K. As the temperature increases, the peak shifts to shorter wavelength (Wien's law) and the total area under the curve grows rapidly (Stefan-Boltzmann's law). The dots mark the peak wavelength at each temperature.

Key Observations from the Spectral Curves

Five features of the graph are worth studying carefully. Each one tells a different piece of the radiation story.

1. The peak shifts to shorter wavelength as temperature rises

At 3000K the peak sits deep in the infrared, beyond 900nm. At 4500K it has moved to about 640nm, near the red end of the visible spectrum. This leftward shift is Wien's displacement law in action: the product of peak wavelength and temperature stays constant, so a higher temperature forces the peak to a shorter wavelength. This is why a heating body changes colour from invisible (infrared) through dull red, orange, yellow, and eventually to white and blue-white as its temperature rises.

2. The peak grows taller at higher temperatures

The 4500K curve towers over the 3000K curve. This is not a separate effect from the wavelength shift; it is a physical consequence of it. All electromagnetic radiation travels at the speed of light, so *shorter wavelength means higher frequency*. *The energy carried by each photon is proportional to its frequency*. When the peak moves to shorter wavelength, it moves to a part of the spectrum where each photon is individually more energetic. A hotter body therefore radiates more power at its peak for two reasons working together: it emits more photons per second **and** each photon near the new peak carries more energy than the photons near the old peak. The taller peak is a direct consequence of the shift to shorter wavelength.

3. The total area under each curve grows rapidly with temperature

The area under the spectral curve represents the total power radiated per unit surface area. This area grows as the fourth power of the absolute temperature, which is Stefan-Boltzmann's law visible on the graph. Doubling the temperature increases the area by a factor of 16. The difference between the 3000K and 4500K curves is a ratio of $(4500/3000)^4 = (1.5)^4 = 5.06$, meaning the 4500K body radiates about five times as much total power per unit area as the 3000K body.

4. The peak height scales as the fifth power of temperature

This follows from observations 2 and 3 acting together. Stefan-Boltzmann says the total area grows as the fourth power of T . Wien says the peak shifts to shorter wavelength, which compresses the curve horizontally into a narrower band. More area squeezed into a narrower band forces the curve to grow taller. The combined effect is that the peak height is proportional to the fifth power of T : doubling the temperature makes the peak 32 times taller. This is why the height difference between curves on the graph is even more dramatic than the area difference.

5. The colour sequence of a heated body

As a body is heated continuously from room temperature, its dominant emission progresses through the following sequence: invisible (infrared only) at low temperatures, then dull red, bright red, orange, yellow, white (a mixture of all visible wavelengths), and finally blue-white at extremely high temperatures. The Sun's surface at about 5800K emits a spectrum that peaks in the visible range, which is why sunlight looks approximately white. A tungsten light-bulb filament at about 2800K peaks in the infrared, which is why incandescent bulbs are poor light sources: most of their radiated energy comes out as invisible heat rather than visible light.

The ideas have finished cooking; now let us sit down and enjoy them through three worked examples.

BINDER Example 16

A patient lies in an examination room at 16°C. The patient's skin temperature is 34°C, their body surface area is 1.6m², and the emissivity of human skin is 0.80. (a) Calculate the total power radiated by the patient. (b) Calculate the power the patient absorbs from the surroundings. (c) Find the net rate of radiative heat loss.

Solution

In this problem, Stefan-Boltzmann law applied twice: once for emission, once for absorption.

Given:

Skin temperature $T = 34^\circ\text{C} = 307\text{K}$, surroundings $T_s = 16^\circ\text{C} = 289\text{K}$

Surface area $A = 1.6\text{m}^2$, emissivity $\varepsilon = 0.80$

Stefan-Boltzmann constant $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2}\text{K}^{-4}$

(a) Total power radiated by the patient:

$$P_{\text{emit}} = \varepsilon \sigma A T^4 = (0.80)(5.67 \times 10^{-8} \text{ W m}^{-2}\text{K}^{-4})(1.6\text{m}^2)(307\text{K})^4 = 645\text{W}$$

The total power radiated by the patient is 645W.

(b) Power absorbed from the surroundings:

$$P_{\text{abs}} = \epsilon \sigma A T_s^4 = 0.80 \times 5.67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4} \times 1.6\text{m}^2 \times (289\text{K})^4 = 506\text{W}$$

The power the patient absorbs from the surroundings is 506W.

(c) Net rate of radiative heat loss:

$$P_{\text{net}} = P_{\text{emit}} - P_{\text{abs}} = 645\text{W} - 506\text{W} = 139\text{W}$$

The net rate of radiative heat loss is 139W.

The patient loses about 139W of heat by radiation alone.

Making Sense of the Answer: 139 watts is a significant fraction of the body's resting metabolic rate (roughly 80 to 100W for an average adult at rest). The fact that radiative heat loss alone exceeds the resting metabolic rate means the body must rely on other sources of internal heat (muscular activity, food digestion, shivering) to maintain its temperature in a 16°C room. This is why patients in cold examination rooms feel cold very quickly even without wind: radiation alone is pulling heat away faster than the body at rest can generate it.

Think Like a Physicist: In any Stefan-Boltzmann problem, it is better to compute both the emitted and absorbed powers separately before subtracting. It is tempting to jump directly to the net formula, but computing each term individually gives you a sanity check: the emitted power should always exceed the absorbed power for a body hotter than its surroundings, and the absorbed power should never be negative (it represents real energy arriving from the environment). A negative absorbed power means a sign error in your temperature conversion.

BINDER Example 17

A furnace radiates maximum energy at a wavelength of 1200Å. A star emits its maximum radiation at 4800Å.

(a) Calculate the surface temperature of each body. (b) Find the ratio of the power radiated per unit area by the furnace to that by the star.

Solution

Note: 1Å = 10⁻¹⁰m.

(a) Surface temperatures from Wien's displacement law:

For the furnace, $\lambda_{\text{max}} = 1200\text{Å} = 1200 \times 10^{-10} \text{ m} = 1.20 \times 10^{-7} \text{ m}$:

$$T_f = \frac{b}{\lambda_{\text{max}}} = \frac{2.90 \times 10^{-3} \text{ mK}}{1.20 \times 10^{-7} \text{ m}} = 24167\text{K}$$

For the star, $\lambda_{\text{max}} = 4800\text{Å} = 4.80 \times 10^{-7} \text{ m}$:

$$T_s = \frac{2.90 \times 10^{-3} \text{ mK}}{4.80 \times 10^{-7} \text{ m}} = 6042\text{K}$$

The furnace is at 24167K and the star is at 6042K (comparable to the Sun's surface temperature of 5800K).

(b) From Stefan-Boltzmann law:

$$P = \epsilon \sigma A T^4$$

The power per unit area is:

$$\frac{P}{A} = \epsilon \sigma T^4$$

Both a furnace cavity and a stellar surface emit radiation that closely matches the Planck blackbody spectrum, so we take $\epsilon = 1$ for both. The ratio is:

$$\frac{\left(\frac{P}{A}\right)_f}{\left(\frac{P}{A}\right)_s} = \left(\frac{T_f}{T_s}\right)^4 = \left(\frac{24167\text{K}}{6042\text{K}}\right)^4 = (4.0)^4 = 256$$

The furnace radiates 256 times as much power per unit area as the star.

Making Sense of the Answer: The furnace is only 4 times hotter than the star, but it radiates 256 times as much power per unit area. The fourth-power law amplifies every temperature difference dramatically. This

is why industrial furnaces and welding arcs consume such extraordinary amounts of energy: the power they radiate scales as the fourth power of their temperature, and at several thousand kelvin the losses become prodigious.

Think Like a Physicist: A useful way to estimate Stefan-Boltzmann ratios quickly: if the temperature ratio is n , the power ratio is n to the fourth. For $n = 2$, the ratio is 16; for $n = 3$, it is 81; for $n = 4$, it is 256; for $n = 10$, it is 10000. These round number let you predict the answer to any comparative radiation problem before touching a calculator.

BINDER Example 18

A piece of metal loses 255Js^{-1} by radiation when its temperature is 1200K and the surroundings are at 300K . Find the rate of heat loss by radiation when the metal cools to 600K , with the same surroundings.

Solution

Interpreting the data:

At state 1: $T_1 = 1200\text{K}$, $T_s = 300\text{K}$, $P_1 = 255\text{Js}^{-1}$

At state 2: $T_2 = 600\text{K}$, $T_s = 300\text{K}$

Net Stefan-Boltzmann equation at each state:

$$P_1 = \epsilon\sigma A(T_1^4 - T_s^4)$$

$$P_2 = \epsilon\sigma A(T_2^4 - T_s^4)$$

Dividing cancels $\epsilon\sigma A$:

$$\frac{P_2}{P_1} = \frac{T_2^4 - T_s^4}{T_1^4 - T_s^4}$$

Substituting (the K^4 units cancel in numerator and denominator):

$$\frac{P_2}{255\text{Js}^{-1}} = \frac{(600\text{K})^4 - (300\text{K})^4}{(1200\text{K})^4 - (300\text{K})^4}$$

Solving for P_2 :

$$P_2 = 15.0\text{Js}^{-1}$$

The rate of heat loss at 600K is 15.0Js^{-1} .

Making Sense of the Answer: The metal has halved in temperature (1200K to 600K), but the net radiative loss has fallen from 255W to 15W , a factor of 17. This is far steeper than the factor of 16 you would expect from the fourth-power law applied to the body temperature alone, because the net formula also includes the surroundings term. At 1200K , the surroundings (300K) are cold compared to the body and barely matter. At 600K , the surroundings are only a factor of 2 away, and their contribution to the absorbed power is now significant: the body's emission has fallen dramatically (by a factor of 16), but its absorption from the surroundings has stayed the same (the surroundings are still at 300K). The net loss shrinks even faster than the gross emission.

Think Like a Physicist: In any comparative radiation problem, always compute the full net expression rather than approximating by dropping the surroundings term. The surroundings term may be negligible when the body is very hot, but it becomes dominant as the body approaches the surroundings temperature. Dropping it early is a trap that produces the right-looking answer (a clean power of 4 ratio) at high temperatures but the wrong answer at moderate temperatures.

Stefan-Boltzmann and Wien together describe the total power and the spectral distribution of blackbody radiation. But radiation is a two-way street: every body emits and every body absorbs, simultaneously and continuously. The next subsection takes up this two-way picture, called Prevost's theory of heat exchange, and shows how it explains phenomena from premature-baby incubators to thermos flasks. After that we will derive the solar constant, predict Earth's equilibrium temperature, and show that Newton's law of cooling is a special limiting case of Stefan's law.

Prevost's Theory of Heat Exchange

It is tempting to think that a hot body radiates and a cold body absorbs. **Prevost's theory**, stated by Pierre Prevost in 1791, corrects this picture in an important way.

According to **Prevost's theory of heat exchange**: *Every body continuously emits thermal radiation and simultaneously absorbs thermal radiation from its surroundings, regardless of whether the body is hotter than, cooler than, or at the same temperature as its surroundings.*

In other words, the theory suggests that: *Emission and absorption are not alternatives; they happen together, at all times, for every body above absolute zero.* A body does not “stop emitting” when it reaches thermal equilibrium. It simply reaches a state in which its rate of emission equals its rate of absorption, so the net exchange is zero and its temperature stays constant.

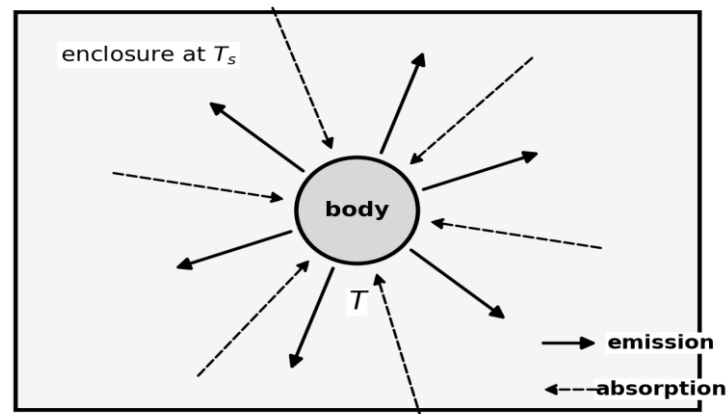


Figure: Prevost's theory of heat exchange. A body at temperature T sits inside an enclosure at temperature T_s . Solid arrows represent emission (outward from the body); dashed arrows represent absorption (inward from the enclosure walls). Both processes occur simultaneously and continuously, regardless of which is hotter.

Consequences of Prevost's theory

1. At thermal equilibrium ($T = T_s$), the body still radiates and still absorbs. The two rates are equal, so there is no net heat transfer and the temperature does not change. Equilibrium is a dynamic balance, not a cessation of activity.
2. A premature baby in a hospital incubator may be kept in warm air at, say, 35°C , yet still lose dangerous amounts of heat by radiation if the walls of the incubator are cold (for instance, if the incubator is near a cold window). The baby's skin radiates toward the cold walls, and the cold walls radiate back less than the baby emits. The net radiative loss can be large enough to cause hypothermia, even though the air temperature is perfectly adequate. Modern neonatal incubators solve this problem by using double-walled enclosures with warm inner walls, so that the wall radiation arriving at the baby's skin is close to what the baby emits.
3. A vacuum (thermos) flask reduces radiative heat transfer by silvering the inner and outer walls. The silver surface has very low emissivity ($\epsilon \approx 0.02$), so the hot liquid inside emits very little radiation outward, and the cold environment outside sends very little radiation inward. Both directions are suppressed by the same low emissivity, exactly as Prevost's theory predicts: reducing emission necessarily reduces absorption by the same factor.

The Solar Constant and Earth's Equilibrium Temperature

The Sun is the most important radiating body in our daily lives. Everything from the temperature of the atmosphere to the direction of the sea breeze depends on how much solar radiation reaches Earth's surface. In this subsection we derive the solar constant from Stefan-Boltzmann's law and use it to predict what Earth's temperature would be without an atmosphere.

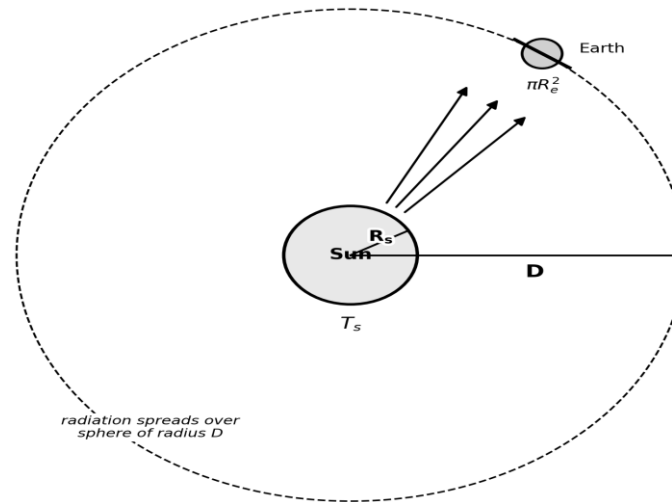


Figure: Sun-Earth geometry for the solar-constant derivation. The Sun (radius R_s , surface temperature T_s) radiates uniformly in all directions. At Earth's orbital distance D , the radiation has spread over a sphere of area $4\pi D^2$. The Earth intercepts a cross-sectional area πR_e^2 of this radiation.

Deriving the solar constant

Treat the Sun as a blackbody of radius R_s and surface temperature T_s . By Stefan-Boltzmann, the total power radiated by the Sun is:

$$P_{\text{Sun}} = \sigma(4\pi R_s^2)T_s^4$$

This radiation spreads uniformly in all directions. At a distance D from the Sun (the Earth's orbital radius), the same total power is distributed over a sphere of surface area $4\pi D^2$. The power per unit area at distance D is therefore:

$$\varphi_s = \frac{P_{\text{Sun}}}{4\pi D^2} = \frac{\sigma(4\pi R_s^2)T_s^4}{4\pi D^2}$$

The factors of 4π cancel, leaving:

$$\varphi_s = \sigma \left(\frac{R_s}{D} \right)^2 T_s^4$$

This quantity φ_s is the **solar constant**: the power per unit area of solar radiation arriving at the top of Earth's atmosphere, measured perpendicular to the Sun's rays. Its unit is Wm^{-2} .

Earth's equilibrium temperature

The Earth intercepts solar radiation over a circular cross-section of area πR_e^2 (the shadow the Earth would cast). The total power received by Earth is therefore:

$$P_{\text{received}} = \varphi_s \pi R_e^2$$

The Earth re-radiates this energy back into space from its entire surface area $4\pi R_e^2$ (treating the Earth as a blackbody sphere). At equilibrium, the power received equals the power emitted:

$$\varphi_s \pi R_e^2 = \sigma(4\pi R_e^2)T_e^4$$

The πR_e^2 cancels from both sides:

$$\varphi_s = 4\sigma T_e^4$$

Substituting the expression for φ_s :

$$\sigma \left(\frac{R_s}{D} \right)^2 T_s^4 = 4\sigma T_e^4$$

Cancelling σ and solving for T_e :

$$T_e^4 = \frac{R_s^2 T_s^4}{4D^2}$$

$$T_e = T_s \sqrt{\left(\frac{R_s}{2D}\right)}$$

This is the equilibrium temperature of a planet at orbital distance D from a star of radius R_s and surface temperature T_s , assuming the planet is a blackbody with no atmosphere.

Numerical computation

Given:

Sun's radius $R_s = 6.96 \times 10^8 \text{ m}$

Earth's orbital distance $D = 1.50 \times 10^{11} \text{ m}$

Sun's surface temperature $T_s = 5800\text{K}$

$$T_e = 5800\text{K} \times \sqrt{\frac{6.96 \times 10^8 \text{ m}}{2 \times 1.50 \times 10^{11} \text{ m}}} = 279\text{K}$$

Earth's equilibrium temperature without an atmosphere would be about 279K, which is 6°C. The observed mean surface temperature of Earth is about 288K (15°C). The difference of 9K is the **greenhouse effect**: the atmosphere traps outgoing infrared radiation and warms the surface above what bare radiative balance would predict. Without the greenhouse effect, Earth would be a frozen planet barely above the melting point of ice.

Understanding the role of reflection: albedo

The derivation above assumed that Earth absorbs every photon of sunlight that strikes it. In reality, a significant fraction of the incoming solar radiation is reflected straight back into space before it can be absorbed. Clouds reflect sunlight from their white upper surfaces. Ice caps and snow-covered land reflect it from theirs. Even the ocean reflects a portion at low angles. The fraction of incoming radiation that a planet reflects is called its **albedo**, given the symbol α (the Greek letter alpha). An albedo of $\alpha = 0$ means the planet absorbs everything (a perfect blackbody); an albedo of $\alpha = 1$ means the planet reflects everything (a perfect mirror). Earth's average albedo is approximately $\alpha \approx 0.30$, meaning about 30% of the sunlight arriving at the top of the atmosphere is reflected and only 70% is absorbed.

When the albedo is taken into account, the power actually absorbed by Earth is not $\varphi_s \pi R_e^2$ but $(1 - \alpha)\varphi_s \pi R_e^2$. The radiative equilibrium equation becomes:

$$(1 - \alpha)\varphi_s \pi R_e^2 = \sigma(4\pi R_e^2)T_e^4$$

Solving for T_e gives a corrected equilibrium temperature that is lower than the bare-blackbody value. With $\alpha = 0.30$ and $\varphi_s = 1380\text{Wm}^{-2}$, the corrected equilibrium temperature comes out to approximately 255K (−18°C), which is 33°C below the observed mean surface temperature of 288K (15°C). This 33°C difference is the true greenhouse lift: it is the full warming that Earth's atmosphere provides by trapping outgoing infrared radiation. The idealized calculation without albedo (279K, giving a 9K greenhouse lift) understated the effect because it assumed the planet absorbed more sunlight than it actually does.

Newton's Law of Cooling as a Limiting Case of Stefan's Law

Earlier in this chapter, Newton's law of cooling was stated as an empirical observation: the rate of heat loss is proportional to the excess temperature ($\theta - \theta_s$). We can now show that Newton's law follows directly from Stefan-Boltzmann's law when the excess temperature is small compared to the absolute temperature of the surroundings.

Starting with the net radiative power loss from Stefan-Boltzmann:

$$P_{\text{net}} = \varepsilon\sigma A(T^4 - T_s^4)$$

Factoring the difference of fourth powers.

Recall the algebraic identity: $a^4 - b^4 = (a - b)(a + b)(a^2 + b^2)$

Applying this with $a = T$ and $b = T_s$:

$$T^4 - T_s^4 = (T - T_s)(T + T_s)(T^2 + T_s^2)$$

Now applying the condition that the excess temperature is small: $(T - T_s) \ll T_s$.

This means $T \approx T_s$, so we can approximate the remaining factors:

$$\begin{aligned}(T + T_s) &\approx 2T_s \\ (T^2 + T_s^2) &\approx 2T_s^2\end{aligned}$$

Substituting these approximations into the factored expression:

$$T^4 - T_s^4 \approx (T - T_s) \times 2T_s \times 2T_s^2 = 4T_s^3(T - T_s)$$

The net radiative power loss therefore becomes:

$$P_{\text{net}} \approx \varepsilon \sigma A \times 4T_s^3(T - T_s) = 4\varepsilon \sigma AT_s^3(T - T_s)$$

This is exactly Newton's law of cooling. The rate of heat loss is directly proportional to the excess temperature, with a proportionality constant:

$$k = 4\varepsilon \sigma AT_s^3$$

So Newton's constant k is not truly a constant: it depends on the emissivity, the surface area, and the cube of the surroundings temperature. But for a given body cooling in a given room, all of these quantities are fixed, and k behaves as a constant throughout the cooling process.

This derivation reveals that Newton's law is a small-excess-temperature approximation to Stefan-Boltzmann's law, valid whenever $(T - T_s) \ll T_s$, which is the condition stated in the convection section.

With this solid theoretical foundation in place, we are now ready to see these ideas come alive through four more examples.

BINDER Example 19

Describe what happens to a blackbody as it is constantly heated, starting from room temperature. In your answer, describe the colour progression, the shift of the peak wavelength, and state which quantity increases most rapidly with temperature.

Solution

At room temperature (about 300K), the blackbody emits radiation entirely in the infrared, far beyond the range of human vision. The body appears dark to the eye, and its peak wavelength, given by Wien's law $\lambda_{\text{max}} = b/T$, is at about $2.90 \times 10^{-3} \text{ mK} / 300\text{K} \approx 9700\text{nm}$, deep in the infrared.

As the temperature rises, the peak wavelength shifts to shorter values. At about 800K, the peak enters the far-red end of the visible spectrum and the body begins to glow dull red. Continued heating brings the colour progression through bright red, orange, and yellow. At about 5800K (the temperature of the Sun's surface), the peak is in the middle of the visible spectrum and the body appears approximately white, because significant radiation is emitted at all visible wavelengths simultaneously. Above about 10000K, the peak moves into the blue-violet range and the body appears blue-white.

Throughout this progression, two quantities are increasing:

- (1) The total radiated power per unit area, which scales as T^4 (Stefan-Boltzmann).
- (2) The peak spectral emissive power, which scales as T^5 .

Of the two, the peak spectral emissive power increases most rapidly, because T^5 grows faster than T^4 . Doubling the temperature increases the total radiated power by a factor of 16 but increases the peak height by a factor of 32.

Making Sense of the Answer: *The colour progression is a direct visual demonstration of Wien's law: as the temperature rises, the peak moves to shorter wavelength, and the eye perceives whichever visible wavelength happens to be closest to the peak. The fact that the body eventually appears white (not a single colour) reflects the breadth of the blackbody curve: at high temperatures the curve is broad enough to span the entire visible spectrum, and the eye perceives the mixture as white.*

Think Like a Physicist: *When an exam question asks "what happens to a blackbody as it is heated?", three things must appear in the answer: (1) the peak shifts to shorter wavelength (Wien), (2) the total radiated power grows as the fourth power of absolute temperature (Stefan-Boltzmann), and (3) the peak height grows as the fifth power. State all three explicitly, because each earns a separate mark. Stating only the colour change without connecting it to the underlying laws is a description, not physics.*

REAL Example 20

A premature baby in Mbulu district hospital lies in an incubator. The baby's skin temperature is 36°C , skin surface area is 0.12m^2 , and skin emissivity is 0.80. The air inside the incubator is maintained at a comfortable 35°C . However, the incubator has a thin plastic wall facing a window, and the inner surface of this wall is only 18°C . (a) Calculate the net rate of radiative heat loss from the baby toward the cold wall. (b) A senior nurse drapes a warm cloth over the plastic wall, raising its inner surface temperature to 33°C . Calculate the new net rate of radiative heat loss. (c) Using Prevost's theory, explain why warm air alone was not enough to keep the baby safe.

Solution

This is a Prevost problem made quantitative. The baby emits radiation and absorbs radiation from the enclosure walls simultaneously. When the wall is cold, the absorption is small and the net loss is large, even though the air is warm.

Given:

Baby's skin: $T = 36^{\circ}\text{C} = 309\text{K}$, $A = 0.12\text{m}^2$, $\varepsilon = 0.80$

Cold wall: $T_s = 18^{\circ}\text{C} = 291\text{K}$

Warm wall (after blanket): $T_s = 33^{\circ}\text{C} = 306\text{K}$

Stefan-Boltzmann constant $\sigma = 5.67 \times 10^{-8} \text{Wm}^{-2}\text{K}^{-4}$

(a) Net radiative heat loss toward the cold wall at 291K:

$$P_{\text{net}} = \varepsilon\sigma A(T^4 - T_s^4) = 0.80 \times 5.67 \times 10^{-8} \text{Wm}^{-2}\text{K}^{-4} \times 0.12\text{m}^2 \times ((309\text{K})^4 - (291\text{K})^4) = 10.6\text{W}$$

The baby loses about 10.6W by radiation toward the cold wall.

(b) Net radiative heat loss after the blanket raises the wall to 306K:

$$P_{\text{net}} = 0.80 \times 5.67 \times 10^{-8} \text{Wm}^{-2}\text{K}^{-4} \times 0.12\text{m}^2 \times ((309\text{K})^4 - (306\text{K})^4) = 1.96\text{W}$$

After covering the wall, the baby loses only 1.96W by radiation.

(c) Prevost's theory says the baby emits and absorbs radiation simultaneously and continuously. The air temperature controls convective heat transfer, but it does not control radiative heat transfer: radiation passes straight through the air between the baby's skin and the enclosure walls without being absorbed by it. What matters for radiation is the temperature of the surfaces that the baby is exposed to. When one of those surfaces is a cold window wall at 18°C , the wall sends back far less radiation than the baby emits toward it. The result is a large net radiative loss (10.6W) that the warm air cannot compensate. The air may be at 35°C , but the baby's skin is exchanging radiation with an 18°C wall, not with the air. This is why warm air alone is not enough: radiation is a separate heat-transfer channel that bypasses the air entirely and responds only to the temperatures of the solid surfaces surrounding the baby.

Making Sense of the Answer: Covering the wall reduced the net radiative loss from 10.6W to about 2.0W, a factor of more than 5. For a tiny premature baby with a resting metabolic rate of only a few watts, this difference can be the margin between stable body temperature and dangerous hypothermia. The physics behind this is pure Prevost: raise the wall temperature and the wall sends more radiation back to the baby, increasing the absorbed power while the emitted power stays the same. The net loss shrinks because the absorption has risen, not because the emission has fallen.

Think Like a Physicist: Whenever a radiation problem involves a body inside an enclosure, check the temperature of the walls, not the temperature of the air. Air is nearly transparent to infrared radiation, so it plays almost no role in radiative heat exchange. The surfaces the body "sees" are what determine the absorbed power. A cold wall, a cold window, or a cold ceiling can drain a body by radiation even when the air is warm. This is the principle behind radiant heating panels, double-walled incubators, and the reason hospitals keep premature babies away from windows.

HOT Example 21

The solar constant at Earth's orbit is measured to be $\varphi_s = 1370\text{Wm}^{-2}$. Earth orbits the Sun at a distance $D = 1.50 \times 10^{11} \text{m}$, and the Sun's radius is $R_s = 6.96 \times 10^8 \text{m}$. (a) Using the solar constant formula, calculate the Sun's surface temperature. (b) The Sun's peak emission wavelength is measured to be about 500nm. Using Wien's law, independently calculate the Sun's surface temperature. (c) Comment on the agreement between the two methods and what it tells us about the Sun as a radiator.

Solution

(a) From the solar constant formula, rearranging for the surface temperature:

$$\varphi_s = \sigma \left(\frac{R_s}{D} \right)^2 T_s^4$$

Solving for T_s :

$$T_s^4 = \frac{\varphi_s}{\sigma \left(\frac{R_s}{D} \right)^2}$$

$$T_s = \sqrt[4]{\frac{\varphi_s}{\sigma \left(\frac{R_s}{D} \right)^2}} = \sqrt[4]{\frac{1370 \text{ Wm}^{-2}}{5.67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4} \left(\frac{6.96 \times 10^8 \text{ m}}{1.50 \times 10^{11} \text{ m}} \right)^2}} = 5800 \text{ K}$$

(b) From Wien's displacement law:

$$T_s = \frac{b}{\lambda_{\max}} = \frac{2.90 \times 10^{-3} \text{ mK}}{500 \times 10^{-9} \text{ m}} = 5800 \text{ K}$$

(c) The two methods give exactly the same value 5800K. The agreement confirms that the Sun's radiation closely follows the Planck blackbody spectrum: the same temperature that explains the total radiated power also explains the wavelength at which the radiation peaks. If the Sun were not a good blackbody, these two temperatures would disagree, and the size of the disagreement would measure how far the Sun's spectrum departs from the Planck curve.

Making Sense of the Answer: This cross-check is how astrophysicists validate stellar temperature measurements. A single temperature that simultaneously satisfies both Stefan-Boltzmann (total power) and Wien (spectral shape) is strong evidence that the object is radiating as a blackbody. For the Sun, the match is excellent. For other stars with thick atmospheres or unusual compositions, the two temperatures sometimes disagree, signalling that the simple blackbody model needs refinement for that particular star.

Think Like a Physicist: Whenever two independent methods give the same answer, the result is more trustworthy than either method alone. In experimental physics this is called a consistency check: if two different physical laws, applied to two different measurements, converge on the same number, then either both are correct or both are wrong in exactly the same way, which is vastly less likely. Building consistency checks into your problem-solving routine is one of the most powerful habits a physicist can develop.

HOT Example 22

A blackbody sphere of radius 5.0cm sits in a room at $T_s = 300\text{K}$. (a) Using the expression for Newton's cooling constant derived from Stefan's law, calculate the value of k for this sphere. (b) The sphere is heated to 320K and allowed to cool. Verify that the small-excess condition is satisfied. (c) Using Newton's law with your calculated k , estimate the initial net rate of heat loss. (d) Calculate the exact net rate of heat loss using the full Stefan-Boltzmann formula and comment on the accuracy of Newton's approximation.

Solution

Given:

Blackbody sphere: $\varepsilon = 1$, radius $r = 5.0\text{cm} = 0.050\text{m}$, so $A = 4\pi r^2 = 4\pi(0.050\text{m})^2 = 0.0314\text{m}^2$

Surroundings: $T_s = 300\text{K}$. Heated temperature: $T = 320\text{K}$

(a) Newton's cooling constant from Stefan's law:

$$k = 4\varepsilon\sigma AT_s^3 = 4 \times 1 \times 5.67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4} \times 0.0314\text{m}^2 \times (300\text{K})^3 = 0.192\text{WK}^{-1}$$

(b) The excess temperature is $T - T_s = 320\text{K} - 300\text{K} = 20\text{K}$. For Newton's law (as a limiting case of Stefan's law) to be valid, we need $(T - T_s) \ll T_s$. The ratio is $20\text{K} / 300\text{K} = 0.067$, which is about 7%. This is comfortably small (well below the rule-of-thumb threshold of about 10%), so the small-excess condition is satisfied and Newton's law should be a good approximation.

(c) Newton's law estimate of the initial net rate of heat loss:

$$P_{\text{Newton}} = k(T - T_s) = 0.192\text{WK}^{-1} \times 20\text{K} = 3.84\text{W}$$

(d) Exact Stefan-Boltzmann net formula:

$$P_{\text{exact}} = \varepsilon\sigma A(T^4 - T_s^4) = 5.67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4} \times 0.0314\text{m}^2 \times ((320\text{K})^4 - (300\text{K})^4) = 4.25\text{W}$$

Comparing the two results: Newton's law gives 3.84W, the exact formula gives 4.25W. The Newton estimate is about 10% below the exact value. This is a reasonable approximation for an excess of 20K out of 300K. The error arises because the approximation $T \approx T_s$ is not perfect: the true fourth-power curve is slightly steeper than the linear tangent that Newton's law represents. For smaller excesses the agreement would be even closer; for larger excesses it would deteriorate.

Making Sense of the Answer: The 10% discrepancy is itself a useful number. It tells the student that Newton's law is a practical tool for small excesses (a 7% excess gives a 10% error in the heat-loss rate), but not a perfect one. The exact Stefan-Boltzmann formula is always more accurate, and the price of using Newton's law is this systematic undershoot.

Think Like a Physicist: This problem demonstrates the physicist's habit of checking an approximation against the exact answer. Deriving an approximation is step one; computing the error it introduces is step two. An approximation without an error estimate is a guess, not physics. Whenever you use Newton's law of cooling, you should be able to state the condition under which it is valid and estimate how large the error is for the specific excess temperature in the problem.

With thermal radiation now complete, including its three major laws (Stefan-Boltzmann, Wien, and Newton's law as a limiting case) and the theory of heat exchange (Prevost), we have covered all three heat-transfer mechanisms: conduction, convection, and radiation. The next section returns to Newton's law of cooling and takes up its mathematical consequences: the exponential cooling curve, the integration of the cooling equation, the half-cooling time, and the experimental methods used to verify the law.

NEWTON'S LAW OF COOLING

In the convection section we stated Newton's law of cooling as an empirical observation: the rate of heat loss from a hot body is proportional to the excess temperature of the body over its surroundings. In the radiation section we showed that this law is a small-excess-temperature approximation to Stefan-Boltzmann's law. This section takes Newton's law and does the mathematics: we derive the differential equation, integrate it to find how the temperature changes with time, extract the half-cooling time, and develop the two-interval method that is the most heavily tested numerical technique in heat-related examination questions.

From Heat Loss to Temperature Change

Newton's law of cooling, in its heat-loss form, says:

$$\frac{dQ}{dt} = k(\theta - \theta_s)$$

where $\frac{dQ}{dt}$ is the rate at which the body loses heat (in watts), k is a positive constant that depends on the body's surface properties and surroundings, θ is the body's temperature, and θ_s is the surroundings temperature. But in practice we do not measure heat loss directly; we measure temperature. We need to convert this law into one that tells us how fast the temperature falls.

When a body of mass m and specific heat capacity c loses a small amount of heat dQ , its temperature drops by $d\theta$, where:

$$dQ = -mc d\theta$$

The minus sign is present because dQ is positive (heat leaves the body) while $d\theta$ is negative (the temperature falls). Dividing both sides by dt :

$$\frac{dQ}{dt} = -mc \frac{d\theta}{dt}$$

Setting this equal to Newton's law:

$$-mc \frac{d\theta}{dt} = k(\theta - \theta_s)$$

Dividing both sides by $-mc$:

$$\frac{d\theta}{dt} = -\frac{k}{mc}(\theta - \theta_s)$$

The quantity $\frac{k}{mc}$ is a positive constant for a given body cooling in a given environment. We give it the symbol λ (lambda):

$$\lambda = \frac{k}{mc}$$

The cooling equation becomes:

$$\frac{d\theta}{dt} = -\lambda(\theta - \theta_s)$$

The result is **Newton's law of cooling in temperature form** which can be stated as: *under forced convection (at any excess temperature) or natural convection with a small excess temperature (below about 30°C), the rate of fall of temperature of a body is proportional to the excess temperature of the body over its surroundings.*

The constant λ has units of inverse time (s^{-1} or min^{-1}) and is called the **cooling constant**. A large λ means rapid cooling; a small λ means slow cooling.

Integrating the Cooling Equation

The cooling equation is a separable first-order differential equation. To solve it, separate the variables and integrate.

$$\frac{d\theta}{\theta - \theta_s} = -\lambda dt$$

Integrate both sides. At $t = 0$ the body's temperature is θ_i (the initial temperature). At time t the temperature is θ :

$$\begin{aligned}\int_{\theta_i}^{\theta} \frac{d\theta}{\theta - \theta_s} &= -\lambda \int_0^t dt \\ [\ln(\theta - \theta_s)]_{\theta_i}^{\theta} &= -\lambda [t]_0^t \\ \ln(\theta - \theta_s) - \ln(\theta_i - \theta_s) &= -\lambda t \\ \ln\left(\frac{\theta - \theta_s}{\theta_i - \theta_s}\right) &= -\lambda t\end{aligned}$$

This is the logarithmic form of Newton's law of cooling.

Taking the exponential of both sides:

$$\frac{\theta - \theta_s}{\theta_i - \theta_s} = e^{-\lambda t}$$

Rearranging to make the temperature the subject:

$$\theta = \theta_s + (\theta_i - \theta_s)e^{-\lambda t}$$

This is the **integrated form of Newton's law of cooling**. It says that *the temperature of a cooling body decays exponentially from its initial value θ_i toward the surroundings temperature θ_s , with a time constant $1/\lambda$* . The excess temperature $(\theta - \theta_s)$ decays by the same fraction in each equal time interval, which is the hallmark of exponential decay.

The Cooling Curve

The graph of temperature against time is called the cooling curve. Its shape follows directly from the integrated form.

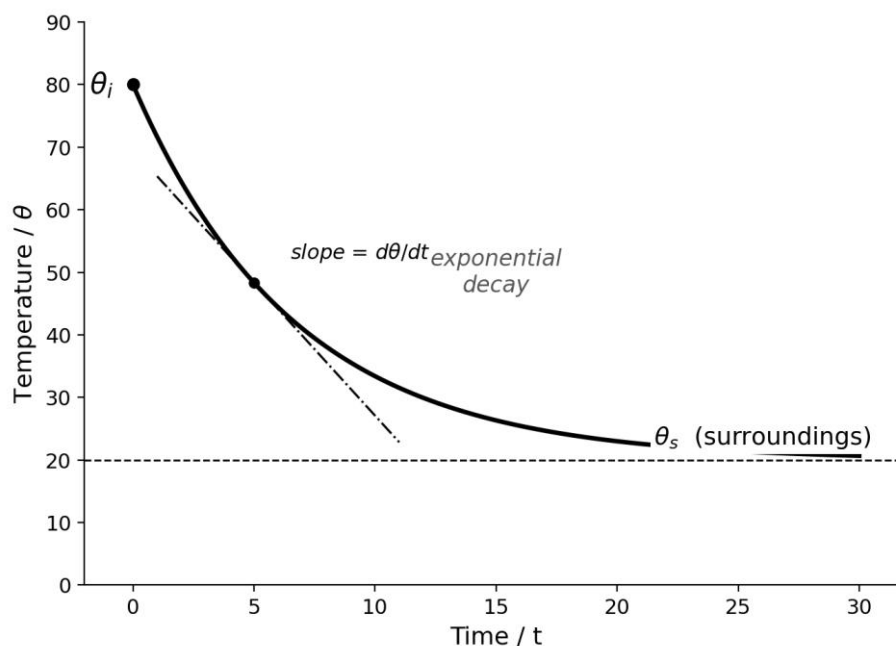


Figure: The cooling curve: temperature θ plotted against time t . The curve starts at the initial temperature θ_i and decays exponentially toward the surroundings temperature θ_s , which appears as a horizontal asymptote. The tangent at any point gives the instantaneous rate of cooling $d\theta/dt$ at that moment.

Key features of the cooling curve

1. The curve is steepest at the start (when the excess temperature is greatest) and becomes progressively flatter as the body approaches the surroundings temperature. This is because the rate of cooling is proportional to the excess: a large excess drives fast cooling, and fast cooling quickly reduces the excess, which slows the cooling further.
2. The curve has a horizontal asymptote at θ_s . Mathematically, the exponential factor $e^{-\lambda t}$ approaches zero as $t \rightarrow \infty$, so $\theta \rightarrow \theta_s$. But the exponential never reaches zero: the body gets closer and closer to θ_s without ever arriving. In practice, the body reaches a temperature indistinguishable from θ_s after a few multiples of $1/\lambda$.
3. The tangent to the curve at any point gives the instantaneous rate of cooling $\frac{d\theta}{dt}$ at that moment. The tangent is steep at the start and flat near the asymptote.

The Rate-of-Cooling Graph

If we plot the rate of cooling $-\frac{d\theta}{dt}$ against the excess temperature $(\theta - \theta_s)$, Newton's law predicts a straight line through the origin with gradient λ .

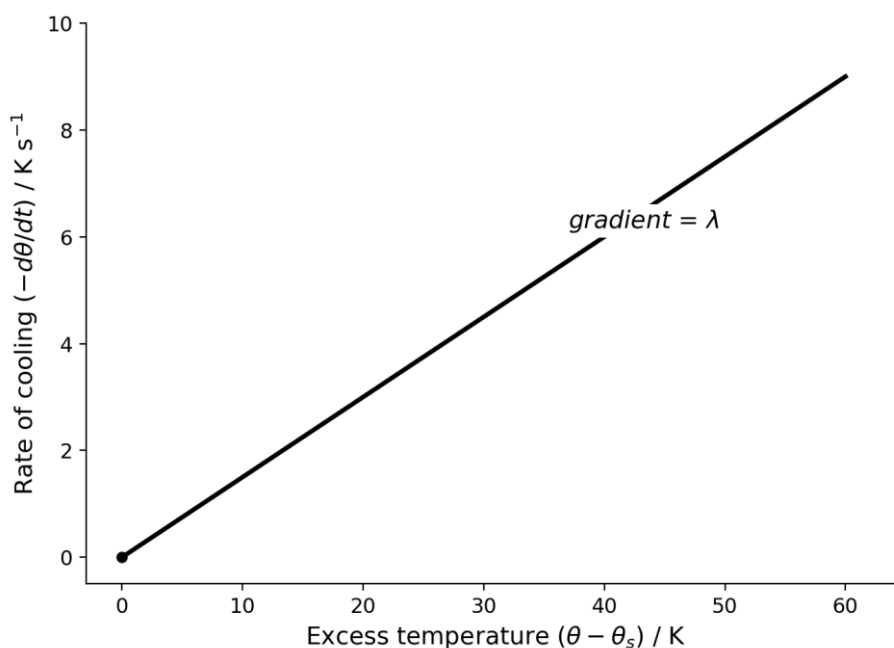


Figure: The rate-of-cooling graph. A plot of the rate of temperature fall $(-d\theta/dt)$ against the excess temperature $(\theta - \theta_s)$ gives a straight line through the origin. The gradient of this line is the cooling constant λ .

This graph is used experimentally to verify that a cooling body obeys Newton's law. If the data points fall on a straight line through the origin, the law is confirmed; if they curve, the cooling departs from the Newton regime (perhaps the excess is too large, and Dulong-Petit applies instead). The gradient of the line gives the numerical value of the cooling constant.

Half-Cooling Time

Just as radioactive decay has a half-life, Newton's cooling has a **half-cooling time**: the time required for the excess temperature to fall to half its initial value.

At time $t_{1/2}$, the excess temperature has fallen to half its initial value:

$$\theta - \theta_s = \frac{1}{2}(\theta_i - \theta_s)$$

Substituting into the logarithmic form:

$$\ln\left(\frac{\frac{1}{2}(\theta_i - \theta_s)}{\theta_i - \theta_s}\right) = -\lambda t_{1/2}$$

$$\ln\left(\frac{1}{2}\right) = -\lambda t_{1/2}$$

$$-\ln 2 = -\lambda t_{1/2}$$

$$t_{1/2} = \frac{\ln 2}{\lambda} = \frac{0.693}{\lambda}$$

The half-cooling time depends only on λ , not on the initial temperature or the surroundings temperature. This is exactly analogous to the half-life of radioactive decay ($t_{1/2} = \ln 2/\lambda$), and for the same reason: exponential decay always takes the same time to halve, regardless of where you start measuring.

The Two-Interval Method

The most common type of Newton's-law problem in the examinations gives three temperatures (or two temperature drops) at known time intervals and asks the student to find a fourth (or third) temperature or a cooling time. The systematic way to solve such problems is the two-interval method.

The key idea is this: the logarithmic form of Newton's law applies to any two moments in time, not just from the start. If the body cools from temperature θ_1 (higher temperature) to temperature θ_2 (lower temperature) in time interval t , then:

$$\ln\left(\frac{\theta_1 - \theta_s}{\theta_2 - \theta_s}\right) = \lambda t$$

If we know two such intervals (or one interval and a relationship between the two), we can write two equations and solve for the unknowns.

The equal-time shortcut

If both intervals have the **same** duration t , the right-hand sides of the two logarithmic equations are both λt , and we can set them equal or divide them. This produces a particularly clean ratio that eliminates λ entirely:

$$\frac{\theta_1 - \theta_s}{\theta_2 - \theta_s} = \frac{\theta_2 - \theta_s}{\theta_3 - \theta_s}$$

where $\theta_1, \theta_2, \theta_3$ are the temperatures at the start of the first interval, the end of the first interval (= start of the second), and the end of the second interval. This ratio says that the excess temperatures at three equally-spaced times form a geometric progression. It is the single most useful equation in Newton's-law problems.

In practical terms: if you know any three of the four quantities (the three temperatures and the surroundings temperature), you can find the fourth.

For now, let us pause the theory and bring the ideas down to earth with six carefully chosen worked examples.

A note on the regime of validity. *The problems that follow may use Newton's law at excess temperatures of 50K and above. Strictly speaking, the convection section showed that Newton's law is a small-excess approximation, valid when the excess is below about 30°C under natural convection. However, NECTA examination questions routinely apply Newton's law at much larger excesses without comment, and students are expected to use it unless the question specifically invokes Dulong-Petit. In practice, the two-interval method and the exponential decay formula work mathematically at any excess; the physical accuracy decreases at large excesses, but the examination expects the method regardless. When in doubt, use Newton's law unless the question tells you otherwise.*

BINDER Example 23

A body cools from 60°C to 40°C in 7 minutes. The surroundings are at 10°C. Find the temperature of the body after the next 7 minutes.

Solution

Given:

$$\theta_1 = 60^\circ\text{C}, \theta_2 = 40^\circ\text{C}, \theta_s = 10^\circ\text{C}, \theta_3 = ?$$

The excess temperatures at equally-spaced times (of 7 minutes) form a geometric progression:

$$\frac{\theta_1 - \theta_s}{\theta_2 - \theta_s} = \frac{\theta_2 - \theta_s}{\theta_3 - \theta_s}$$

Substituting the excess temperatures:

$$\frac{60^{\circ}\text{C} - 10^{\circ}\text{C}}{40^{\circ}\text{C} - 10^{\circ}\text{C}} = \frac{40^{\circ}\text{C} - 10^{\circ}\text{C}}{\theta_3 - 10^{\circ}\text{C}}$$

Solving for θ_3 :

$$\theta_3 = 28^{\circ}\text{C}$$

After the next 7 minutes, the body will be at 28°C .

Making Sense of the Answer: The body cooled by 20°C in the first 7 minutes (from 60 to 40) but only by 12°C in the next 7 minutes (from 40 to 28). The rate of cooling has slowed because the excess temperature has shrunk: at the start the excess was 50°C , driving rapid cooling; after 7 minutes it was only 30°C , driving slower cooling; after 14 minutes it will be only 18°C . This is the signature of exponential decay: the temperature drop in each successive equal interval is smaller than the previous one.

Think Like a Physicist: In a two-interval problem with **equal time intervals**, always write the geometric-progression equation first. It eliminates λ and gives a direct relationship between the three temperatures and the surroundings temperature.

BINDER Example 24

A body cools from 80°C to 50°C in 5 minutes. The surroundings are at 20°C . How long does it take to cool from 60°C to 30°C ?

Solution

This problem gives two cooling intervals with different start and end temperatures but the same λ . Use the logarithmic form on each interval.

Interval 1: 80°C to 50°C in 5 minutes.

$$\ln\left(\frac{80^{\circ}\text{C} - 20^{\circ}\text{C}}{50^{\circ}\text{C} - 20^{\circ}\text{C}}\right) = \lambda \times 5\text{min}$$

$$\lambda = \frac{\ln 2}{5\text{min}} = 0.1386 \text{ min}^{-1}$$

Interval 2: 60°C to 30°C in time t .

$$\ln\left(\frac{60^{\circ}\text{C} - 20^{\circ}\text{C}}{30^{\circ}\text{C} - 20^{\circ}\text{C}}\right) = 0.1386 \text{ min}^{-1} \times t$$

$$t = 10.0 \text{ min}$$

It takes 10.0 minutes to cool from 60°C to 30°C .

Making Sense of the Answer: The first interval (80 to 50, a 30°C drop) takes 5 minutes. The second interval (60 to 30, also a 30°C drop) takes 10 minutes, twice as long. A student who assumed equal drops take equal time would have answered 5 minutes and been wrong. The second interval takes longer because the body starts at a lower excess temperature (40°C versus 60°C) and ends at an even lower one (10°C versus 30°C), so the cooling is slower throughout. Equal temperature drops do not take equal time; exponential decay guarantees that each successive equal drop takes longer than the last.

Think Like a Physicist: The trap in this problem is the temptation to say “same drop, same time”. This is wrong for exponential cooling and right only for linear cooling. The cure is to always use the logarithmic equation, never a shortcut based on temperature drops alone.

REAL Example 25

A body is discovered in a hotel room in **Moshi**. The room temperature is a steady 20°C . A police doctor measures the body temperature at 10:00 PM and finds it to be 30°C . One hour later, at 11:00 PM, the body temperature has fallen to 27°C . Normal body temperature is 37°C . Assuming Newton’s law of cooling applies, estimate the time of death.

Solution

The strategy: use the two measured temperatures (30°C and 27°C , one hour apart) to find the cooling constant λ . Then work backwards from the known starting temperature of 37°C to find how long ago the body was at that temperature.

Finding λ from the two readings:

At 10:00 PM: $\theta = 30^\circ\text{C}$, excess $= \theta_1 - \theta_s = 30^\circ\text{C} - 20^\circ\text{C} = 10^\circ\text{C}$

At 11:00 PM: $\theta = 27^\circ\text{C}$, excess $= \theta_2 - \theta_s = 27^\circ\text{C} - 20^\circ\text{C} = 7^\circ\text{C}$

$$\ln\left(\frac{10}{7}\right) = \lambda \times 1\text{hr}$$

$$\lambda = 0.3567 \text{ hr}^{-1}$$

Finding the time t from the moment of death to 10:00 PM:

At the moment of death, the excess was:

$$\theta_i - \theta_s = 37^\circ\text{C} - 20^\circ\text{C} = 17^\circ\text{C}$$

At 10:00 PM, the excess was 10°C . The time elapsed from death to 10:00 PM is found from:

$$\ln\left(\frac{17}{10}\right) = \lambda t = 0.3567 \text{ hr}^{-1} \times t$$

$$t = 1.49 \text{ hr} \approx 1 \text{ hr } 29 \text{ min}$$

The body had been cooling for about 1 hour 29 minutes before the 10:00 PM reading.

Hence the estimated time of death is approximately **8:31 PM**.

Making Sense of the Answer: The method relies on two assumptions: that Newton's law applies (the excess is small enough for the linear regime) and that the room temperature stayed constant. At 10:00 PM the excess was 10°C , comfortably within Newton's regime. At the moment of death the excess was 17°C , still reasonable. In real forensic work, pathologists use more sophisticated models that account for clothing, body mass, and ambient conditions, but the basic principle is exactly this: measure two temperatures at known times, compute the cooling constant, and extrapolate backwards.

Think Like a Physicist: Working Newton's law backwards in time is the same logarithmic equation used forwards; the only difference is that the answer gives a time before the measurement rather than after it. Any problem that gives you two temperature readings at known times and asks "when was the body at a different temperature?" is a forensic-cooling problem in disguise, whether the object is a human body, a cup of coffee, or a piece of molten metal.

REAL Example 26

Mr. Akilikubwa presents the following student's working on the board. The question was: "A body cools from 84°C to 60°C in 8 minutes. Surroundings are at 12°C . Find the temperature after the next 8 minutes." The student wrote: "The body dropped 24°C in 8 minutes. By Newton's law, it will drop another 24°C in the next 8 minutes. Temperature $= 60 - 24 = 36^\circ\text{C}$." Mr. Akilikubwa turns to the class: "This student has made a fundamental error."

Where exactly is the mistake? Correct the working and find the right answer.

Solution

The student's mistake: assuming that equal temperature drops occur in equal time intervals. This is true only for linear cooling, which does not describe real cooling. Newton's law says the rate of cooling is proportional to the excess temperature, not constant. As the body cools, the excess shrinks, the driving force weakens, and each successive equal time interval produces a smaller drop than the previous one.

Correct working using the geometric-progression method:

Excess temperatures:

$$\theta_1 - \theta_s = 84^\circ\text{C} - 12^\circ\text{C} = 72^\circ\text{C}$$

$$\theta_2 - \theta_s = 60^\circ\text{C} - 12^\circ\text{C} = 48^\circ\text{C}$$

Geometric progression:

$$\frac{72^\circ\text{C}}{48^\circ\text{C}} = \frac{48^\circ\text{C}}{\theta_3 - 12^\circ\text{C}}$$

$$\theta_3 = 44^\circ\text{C}$$

The correct answer is 44°C , not 36°C .

The student's error gave a temperature that is 8°C too low.

Making Sense of the Answer: The student predicted a 24°C drop; the correct drop is only 16°C (from 60 to 44). In the first 8 minutes the body dropped 24°C (from 84 to 60) because the excess was large (72°C at the start). In the next 8 minutes the excess is smaller (48°C at the start), so the drop must be smaller. The ratio of consecutive drops ($16/24 = 0.667$) equals the ratio of consecutive excess temperatures ($48/72 = 0.667$). This constant ratio is the geometric-progression property: it is the fingerprint of exponential decay, and the surest sign that a student has understood Newton's law rather than guessed at it.

Think Like a Physicist: This error-correction format is a powerful study technique. Understanding why a wrong method gives the wrong answer teaches more than simply applying the right method mechanically. The student who can explain why equal drops do not take equal time has understood Newton's law at a deeper level than the student who merely plugs numbers into the formula. In an exam, if you are unsure of your method, check whether your answer implies equal drops in equal times. If it does, you have almost certainly made this error.

HOT Example 27

A body cools from 70°C to 50°C in 10 minutes, and from 50°C to 35°C in the next 10 minutes. Find the temperature of the surroundings.

Solution

Two equal time intervals of 10 minutes. The geometric-progression equation applies, but this time the unknown is θ_s rather than a temperature or a time.

The excess temperatures at three equally-spaced times form a geometric progression:

$$\frac{\theta_1 - \theta_s}{\theta_2 - \theta_s} = \frac{\theta_2 - \theta_s}{\theta_3 - \theta_s}$$

Cross-multiplying:

$$(\theta_1 - \theta_s)(\theta_3 - \theta_s) = (\theta_2 - \theta_s)^2$$

Substituting $\theta_1 = 70^\circ\text{C}$, $\theta_2 = 50^\circ\text{C}$, $\theta_3 = 35^\circ\text{C}$:

$$\begin{aligned} (70 - \theta_s)(35 - \theta_s) &= (50 - \theta_s)^2 \\ 2450 - 105\theta_s + \theta_s^2 &= 2500 - 100\theta_s + \theta_s^2 \\ -50 &= 5\theta_s \\ \theta_s &= -10^\circ\text{C} \end{aligned}$$

The surroundings temperature is -10°C .

Making Sense of the Answer: A negative surroundings temperature may look surprising for a body that starts at 70°C , but it is physically sensible: the body is cooling in a cold environment (perhaps outdoors on a chilly night or in a freezer room). The key test is that all three excess temperatures must be positive, and they are: $70 - (-10) = 80^\circ\text{C}$, $50 - (-10) = 60^\circ\text{C}$, $35 - (-10) = 45^\circ\text{C}$. Check the geometric progression: $80/60 = 1.333$, and $60/45 = 1.333$. They match. The answer is correct.

Think Like a Physicist: When the surroundings temperature is the unknown, the geometric-progression equation produces **a linear equation in θ_s** (because the quadratic terms cancel). This is a reliable algebraic pattern: learn it, and you can solve any "find the surroundings" problem in under a minute. The cancellation of the squared terms is not a coincidence; it is a consequence of the geometric-progression structure, and it happens **every time**.

HOT Example 28

A hot object has a half-cooling time of 6 minutes in a room at 25°C . The object starts at 105°C . (a) Find the temperature after 6 minutes, after 12 minutes, and after 18 minutes. (b) How long does it take for the object to reach 30°C ?

Solution

The half-cooling time halves the **excess temperature**, not the temperature itself. The initial excess is:

$$\theta_i - \theta_s = 105^\circ\text{C} - 25^\circ\text{C} = 80^\circ\text{C}$$

(a) After each half-cooling time of 6 minutes, the excess halves:

$$\text{After 6 min: excess} = \frac{80^\circ\text{C}}{2} = 40^\circ\text{C}, \text{ so } \theta = 25^\circ\text{C} + 40^\circ\text{C} = 65^\circ\text{C}$$

$$\text{After 12 min: excess} = \frac{40^\circ\text{C}}{2} = 20^\circ\text{C}, \text{ so } \theta = 25^\circ\text{C} + 20^\circ\text{C} = 45^\circ\text{C}$$

$$\text{After 18 min: excess} = \frac{20^\circ\text{C}}{2} = 10^\circ\text{C}, \text{ so } \theta = 25^\circ\text{C} + 10^\circ\text{C} = 35^\circ\text{C}$$

(b) At 30°C , the excess is $30^\circ\text{C} - 25^\circ\text{C} = 5^\circ\text{C}$. The excess must fall from 80°C to 5°C , which is a factor of $\frac{80}{5} = 16$. Since $2^4 = 16$, this requires 4 half-cooling times:

$$t = 4 \times 6 \text{ min} = \mathbf{24 \text{ min}}$$

Alternatively, using $\lambda = \frac{\ln 2}{6 \text{ min}} = 0.1155 \text{ min}^{-1}$:

$$t = \frac{\ln\left(\frac{80}{5}\right)}{0.1155 \text{ min}^{-1}} = \frac{\ln 16}{0.1155} \text{ min} = \frac{2.773}{0.1155} \text{ min} = 24.0 \text{ min}$$

The object reaches 30°C after 24 minutes.

Making Sense of the Answer: A common mistake is to halve the temperature instead of the excess. A student who halves 105 gets 52.5, then 26.25, and concludes the body reaches 25°C after only two half-cooling times (12 minutes). The correct sequence (65, 45, 35, 30) shows the body is still at 35°C after 18 minutes and does not reach 30°C until 24 minutes. The error comes from forgetting that **Newton's law describes the decay of the excess over the surroundings, not the decay of the temperature itself.**

Think Like a Physicist: Whenever a problem gives you the half-cooling time, convert it to the cooling constant immediately: $\lambda = \frac{\ln 2}{t_{1/2}}$. From there, every calculation uses the standard logarithmic equation. The half-cooling time is a convenient input format, not a substitute for the logarithmic method. For problems where the answer is an exact number of half-cooling times (as in part (a)), halving the excess is faster. For problems where the answer is not a whole number of half-cooling times (which is the more common case in exams), **the logarithmic equation is the only reliable method.**

Newton's law of cooling is now fully in our hands: stated, derived, integrated, graphed, and practised on the types of problem that examiners favour. The next section takes up the experimental determination of thermal conductivity, using Searle's apparatus for good conductors and Lee's disc for poor conductors, both of which rely on the steady-state and cooling-curve techniques we have built across this chapter.

EXPERIMENTAL DETERMINATION OF THERMAL CONDUCTIVITY

Everything we have derived so far in this chapter depends on knowing the thermal conductivity of the material in question. But how is that number measured in the first place? This section describes the two standard laboratory methods: **Searle's apparatus** for good conductors and **Lee's disc** for poor conductors. Both methods rest on the same principle but use different experimental strategies to suit the very different thermal behaviours of metals and insulators.

The Underlying Principle

In the steady state, Fourier's law gives the rate of heat flow through a uniform sample:

$$\frac{dQ}{dt} = \frac{kA(\theta_1 - \theta_2)}{x}$$

where k is the thermal conductivity, A is the cross-sectional area, θ_1 and θ_2 are the temperatures of the hot and cold faces, and x is the thickness of the sample. To measure k , we must independently measure all other quantities in this equation: the rate of heat flow $\frac{dQ}{dt}$, the area A , the thickness x , and the two surface temperatures. The two methods differ only in how they measure $\frac{dQ}{dt}$.

Searle's Apparatus (Good Conductors)

Searle's method is designed for materials with high thermal conductivity, such as copper, brass, aluminium, and iron. The sample is a long metal bar, typically 30cm to 50cm, with a large cross-sectional area.

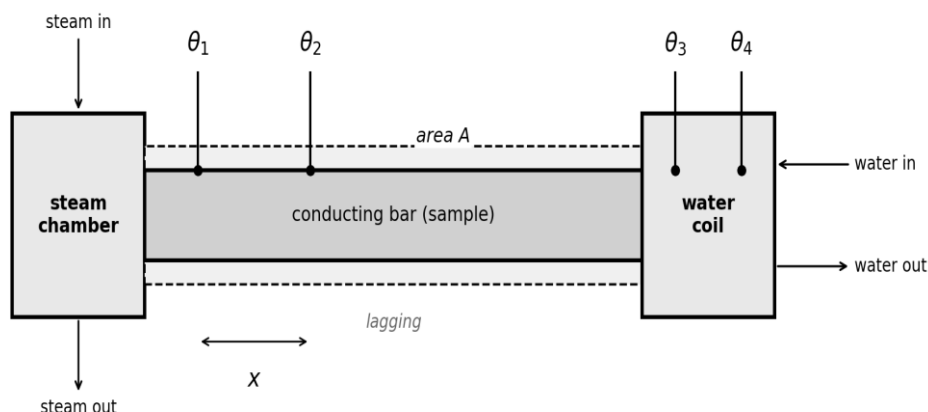


Figure: Searle's apparatus. Steam enters the chamber at the left, maintaining the hot end of the bar at a steady temperature. The bar is lagged to prevent heat loss from its sides. Two thermometers (θ_1 , θ_2) measure the temperature at two points along the bar, separated by a known distance x . At the cold end, water flows through a coil, entering at θ_3 and leaving at θ_4 .

How it works

Steam is passed continuously through the chamber at the hot end of the bar, maintaining a steady temperature. At the cold end, cold water flows through a copper coil at a known, measured rate. The bar is completely lagged along its length to prevent heat escaping from its sides. Under these conditions, all the heat that enters the bar at the hot end must exit through the cold end into the water. When the four thermometer readings become constant, the apparatus has reached its steady state.

At steady state, the rate of heat conducted through the bar equals the rate of heat absorbed by the water:

$$\frac{kA(\theta_1 - \theta_2)}{x} = \frac{m}{t} c(\theta_4 - \theta_3)$$

where $\frac{m}{t}$ is the mass of water flowing through the coil per unit time (measured by collecting and weighing water over a timed interval) and c is the specific heat capacity of water. The only unknown is k , so rearranging:

$$k = \frac{\frac{m}{t} c(\theta_4 - \theta_3) x}{A(\theta_1 - \theta_2)}$$

Why the bar must be lagged

Without lagging, heat would escape from the sides of the bar into the surrounding air, and not all of the heat entering at the hot end would reach the cold end. The equation above assumes that the rate of heat flow is the same at every cross-section along the bar. If heat leaks from the sides, this assumption fails, and the measured value of k will be too low.

Why a long bar is needed for good conductors

A good conductor carries heat so efficiently that even a small temperature difference drives a large heat flow. If the bar were short, the temperature difference ($\theta_1 - \theta_2$) between the two measurement points would be too small to read accurately on a mercury thermometer. A long bar gives a larger separation between the thermometers and a larger, more measurable temperature difference.

Lee's Disc (Poor Conductors)

Lee's disc method is designed for materials with low thermal conductivity, such as cardboard, glass, rubber, cork, and wood. The sample is a thin disc, typically a few millimetres thick.

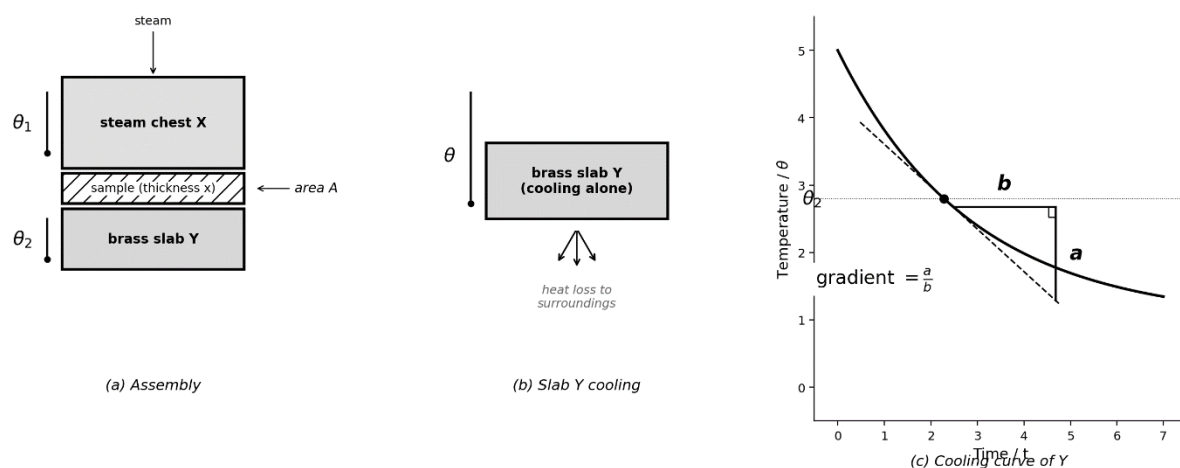


Figure: Lee's disc method. (a) The assembly: the thin sample disc is sandwiched between the steam chest X (at θ_1) and the brass slab Y (at θ_2). (b) The brass slab Y is removed and allowed to cool alone, with a thermometer recording its temperature. (c) The cooling curve of Y : the tangent at the temperature θ_2 gives the gradient a/b , which is the rate of temperature fall at that temperature.

How it works

The experiment has two stages.

Stage 1 (steady state): Steam is passed through the chest X , heating the top face of the sample disc. Heat conducts through the disc and into the brass slab Y below. The assembly is left until the thermometer readings θ_1 (on X) and θ_2 (on Y) become constant. At this point the steady state has been reached: the rate of heat conducted through the sample equals the rate at which slab Y loses heat to its surroundings.

Stage 2 (cooling curve): The sample disc and the steam chest are removed. The brass slab Y , now at temperature θ_2 , is left to cool alone in the air. Its temperature is recorded at regular intervals and plotted as a cooling curve (θ vs t). A tangent is drawn to the curve at the point where $\theta = \theta_2$. The gradient of this tangent, measured as $\frac{a}{b}$ from the tangent triangle, gives the rate of temperature fall of slab Y at the steady-state temperature.

Deriving the working equation

At steady state, the rate of heat conducted through the sample equals the rate at which slab Y loses heat. The rate of heat loss by slab Y at temperature θ_2 is:

$$\frac{dQ}{dt} = mc \frac{a}{b}$$

where m is the mass of slab Y , c is its specific heat capacity, and $\frac{a}{b}$ is the magnitude of the gradient of the cooling curve at $\theta = \theta_2$. Setting this equal to the Fourier conduction rate through the disc:

$$\frac{kA(\theta_1 - \theta_2)}{x} = mc \frac{a}{b}$$

Solving for k :

$$k = \frac{mc \frac{a}{b} x}{A(\theta_1 - \theta_2)}$$

Why a thin disc is needed for poor conductors

A poor conductor resists heat flow strongly. If the sample were thick, the cold face (slab Y) would take a very long time to warm up, and the apparatus might never reach a true steady state within a practical laboratory session. A thin disc ensures that the heat flow rate is large enough to bring slab Y to a measurable temperature within a reasonable time.

Why Different Methods for Different Conductors

The choice of method is not arbitrary; it is dictated by the physics of the material under test.

Good conductors need a **long** sample. A short bar of copper would have such a tiny temperature difference between its ends that the measurement would be swamped by thermometer uncertainty. A long bar gives a measurable ($\theta_1 - \theta_2$). The water-coil calorimeter at the cold end measures the heat flow rate directly and accurately.

Poor conductors need a **thin** sample. A thick slab of cardboard would never reach steady state in a reasonable time. A thin disc reaches steady state quickly. The cooling-curve method replaces the water-coil calorimeter because the heat flow rate through a poor conductor is too small to heat water measurably; instead, it is measured indirectly through the rate of temperature fall of the brass slab.

Attempting to use Searle's method on cardboard would fail because the heat flow would be too small to warm the water. Attempting to use Lee's disc on a 1-metre iron bar would fail because the bar would never reach steady state in a thin-disc arrangement. Each method is tailored to its material.

With the ideas now simmering nicely, let us serve them properly through four worked examples and enjoy the flavour of physics in action.

BINDER Example 29

State the principle on which Searle's apparatus is based. Identify the role of each of the four thermometers and explain why the conducting bar must be lagged.

Solution

Principle:

In the steady state, the rate of heat conducted through the bar equals the rate of heat absorbed by water flowing through the coil at the cold end. Fourier's law gives the conduction rate; the water-flow calorimetry equation gives the absorption rate. Equating the two allows the thermal conductivity k to be calculated.

Role of thermometers:

- Thermometers θ_1 and θ_2 measure the temperature at two points along the bar, separated by a known distance x . Their readings give the temperature gradient along the bar, which appears in Fourier's law as $\frac{(\theta_1 - \theta_2)}{x}$.
- Thermometers θ_3 and θ_4 measure the inlet and outlet temperatures of the cooling water. Their difference ($\theta_4 - \theta_3$) gives the temperature rise of the water, which, combined with the water flow rate and specific heat capacity, gives the rate of heat absorption.

Explanation on lagging bar:

The bar must be lagged to prevent heat loss from its sides. Without lagging, some of the heat entering at the hot end would escape to the surrounding air before reaching the cold end, and the water at the cold end would absorb less heat than the bar conducted. The calculated value of k would then be too low, because the equation assumes all conducted heat reaches the water.

Making Sense of the Answer: *Lagging is not a minor detail; it is the single most important experimental precaution in Searle's method. Without it, the fundamental equality (heat in = heat out) on which the entire method rests does not hold, and the experiment gives a systematically wrong answer.*

Think Like a Physicist: *In any experimental-method question, identify the equation being used and then explain each measurement in terms of a quantity in that equation. This maps every piece of apparatus to a physics symbol, which is exactly what the examiner wants to see. A description of the apparatus without linking it to the equation is a technician's account, not a physicist's.*

HOT Example 30

A copper rod 2.0m long and 1.0cm in radius is maintained at 250°C at one end. In the steady state, heat flows across each cross-section at a rate of 2.1Js⁻¹. The thermal conductivity of copper is $k = 380\text{Wm}^{-1}\text{K}^{-1}$. Find the temperature of the other end.

Solution

Fourier's law, rearranged to find the cold-end temperature.

Given:

Length $x = 2.0\text{m}$, radius $r = 1.0\text{cm} = 0.010\text{m}$

$$\text{Area } A = \pi r^2 = \pi(0.010\text{m})^2 = 3.14 \times 10^{-4} \text{ m}^2$$

$$\text{Hot-end temperature } \theta_1 = 250^\circ\text{C}, \text{ heat-flow rate } \frac{dQ}{dt} = 2.1\text{J s}^{-1} = 2.1\text{W}$$

From Fourier's law:

$$\frac{dQ}{dt} = \frac{kA(\theta_1 - \theta_2)}{x}$$

Rearranging:

$$\theta_1 - \theta_2 = \frac{\frac{dQ}{dt} \times x}{kA} = \frac{2.1\text{W} \times 2.0\text{m}}{380\text{W m}^{-1}\text{K}^{-1} \times 3.14 \times 10^{-4} \text{ m}^2} = 35.2\text{K} = 35.2^\circ\text{C}$$

$$\theta_2 = 250^\circ\text{C} - 35.2^\circ\text{C} = 214.8^\circ\text{C} \approx 215^\circ\text{C}$$

The cold end of the rod is at about 215°C .

Making Sense of the Answer: A 2-metre copper rod carrying only 2.1 watts has a temperature drop of just 35°C across its entire length. This tiny gradient is the signature of a good conductor: copper transmits heat so readily that even a modest heat flow requires almost no temperature difference to drive it. This is also why Searle's method uses a long bar: a short bar of copper would give an even smaller temperature difference, making the measurement impossibly imprecise.

Think Like a Physicist: When a problem asks for the temperature at the other end of a conductor, rearrange Fourier's law for the temperature difference first, then subtract from the known temperature. Never rearrange for the unknown temperature directly; the algebra is cleaner when you isolate the difference.

HOT Example 31

A brass boiler has a base of area 0.15m^2 and thickness 1.0cm . The boiler boils water at a rate of 6.0kg per minute. The thermal conductivity of brass is $k = 109\text{W m}^{-1}\text{K}^{-1}$ and the specific latent heat of vapourisation of water is $L_v = 2256\text{kJ kg}^{-1}$. Find the temperature of the surface in contact with the flame.

Solution

The boiler base conducts heat from the flame to the water. The water is boiling, so its temperature is 100°C . The rate of heat flow equals the rate at which the water absorbs latent heat to vapourise.

Given:

$$\text{Area } A = 0.15\text{m}^2, \text{ thickness } x = 1.0\text{cm} = 0.010\text{m}$$

$$\text{Water boiled: } \frac{m}{t} = 6.0\text{kg min}^{-1} = 0.10\text{kgs}^{-1}$$

$$\text{Latent heat } L_v = 2256\text{kJ kg}^{-1} = 2.256 \times 10^6 \text{ J kg}^{-1}$$

Rate of heat flow through the base:

$$\frac{dQ}{dt} = \frac{m}{t} \times L_v = 0.10\text{kgs}^{-1} \times 2.256 \times 10^6 \text{ J kg}^{-1} = 2.256 \times 10^5 \text{ W}$$

From Fourier's law, the temperature difference across the base:

$$\theta_1 - \theta_2 = \frac{\frac{dQ}{dt} \times x}{kA} = \frac{2.256 \times 10^5 \text{ W} \times 0.010\text{m}}{109\text{W m}^{-1}\text{K}^{-1} \times 0.15\text{m}^2} = 138\text{K} = 138^\circ\text{C}$$

$$\theta_1 = 100^\circ\text{C} + 138^\circ\text{C} = 238^\circ\text{C}$$

The flame-contact surface of the boiler base is at about 238°C .

Making Sense of the Answer: The flame side of the boiler is 138°C hotter than the water side. This large temperature difference across just 1cm of brass reflects the enormous rate of heat flow: 225 kilowatts passing through 0.15m^2 of brass is about 1500kW m^{-2} , which is comparable to an industrial furnace. The reason the base does not melt is that brass has a melting point above 900°C , well above the 238°C flame-contact temperature. If the water were to boil dry, however, the heat flow would stop being absorbed by vapourisation, the base temperature would rise rapidly, and the base could warp or melt.

Think Like a Physicist: Boiling-rate problems are Fourier problems with a twist: the rate of heat flow is not given directly but must be computed from the mass of water boiled per second and the latent heat of vapourisation. Convert the boiling rate to watts first ($dQ/dt = dm/dt \times L_v$), then use Fourier's law in the normal way. The water-side temperature is always 100°C (at standard pressure).

HOT Example 32

(a) Derive the working equation for the Lee's disc method of measuring the thermal conductivity of a poor conductor. (b) Explain why Searle's method cannot be used to measure the thermal conductivity of cardboard, and why Lee's disc method cannot be used for a 1-metre iron bar.

Solution

(a) In the steady state, the rate of heat conducted through the thin sample disc equals the rate at which the brass slab Y loses heat to its surroundings. By Fourier's law, the conduction rate through the disc of thickness x , area A , with face temperatures θ_1 (hot, from the steam chest) and θ_2 (cold, on slab Y) is:

$$\frac{dQ}{dt}_{\text{conduction}} = \frac{kA(\theta_1 - \theta_2)}{x}$$

The rate of heat loss by slab Y at temperature θ_2 is determined from the cooling curve. When slab Y cools alone, its temperature falls at a rate $\frac{d\theta}{dt}$ that can be read from the tangent to the cooling curve at $\theta = \theta_2$. The magnitude of this gradient is $\frac{a}{b}$ (from the tangent triangle). The rate of heat loss is therefore:

$$\frac{dQ}{dt}_{\text{loss}} = mc \frac{a}{b}$$

where m is the mass and c is the specific heat capacity of the brass slab. Equating the two rates:

$$\frac{kA(\theta_1 - \theta_2)}{x} = mc \frac{a}{b}$$

Solving for k :

$$k = \frac{mc \frac{a}{b} x}{A(\theta_1 - \theta_2)}$$

(b) Two reasons, one for each direction:

Searle's method cannot be used for cardboard because cardboard has very low thermal conductivity. The rate of heat flow through the sample would be so small that the cooling water at the cold end would barely warm up: the temperature rise ($\theta_4 - \theta_3$) would be unmeasurably small. Without a measurable heat-flow rate, k cannot be calculated. Additionally, a long bar of cardboard would take an impractically long time to reach steady state.

Lee's disc method cannot be used for a 1-metre iron bar because the method requires a thin disc sample. A 1-metre bar is far too thick: the cold face would take an extremely long time to reach a steady temperature, and the cooling-curve stage assumes that the slab Y is the only significant thermal mass below the sample, which is not true when the sample itself is a massive iron bar. Furthermore, iron is a good conductor: it would easily drive a large heat flow into slab Y, overwhelming the cooling-curve measurement. Searle's method with a water-coil calorimeter is the correct approach for iron.

Making Sense of the Answer: The two methods are complementary, not interchangeable. Each is optimised for a specific range of thermal conductivities. The dividing line is roughly around $k = 1 \text{ Wm}^{-1}\text{K}^{-1}$: materials above this value (metals, some dense ceramics) suit Searle; materials below it (wood, cardboard, rubber, cork, glass wool) suit Lee's disc. Knowing which method to use for which material is itself a piece of physics knowledge that the examiner tests.

Think Like a Physicist: When asked "why can't method X be used for material Y?", always answer in terms of a specific quantity in the working equation that becomes unmeasurably small or impractically large. For Searle on cardboard: the water temperature rise becomes unmeasurable. For Lee on iron: the sample is too thick to reach steady state. Vague answers like "it doesn't work" score no marks; specific answers tied to the equation score full marks.

With the experimental methods for measuring thermal conductivity now complete, we have covered the full arc of Chapter 3: the three heat-transfer mechanisms (conduction, convection, radiation), their quantitative laws, Newton's law of cooling, and the laboratory techniques that put numbers on the thermal properties of materials. The final section brings everything together by examining real-world applications that combine all three mechanisms.

APPLICATIONS OF HEAT TRANSFER

No real-world object loses or gains heat by just one mechanism. A cup of tea on a table cools simultaneously by conduction through the table surface, by convection as warm air rises from its surface, and by radiation from its sides toward the cooler walls of the room. The engineering of everyday thermal devices is the art of identifying which mechanism dominates and either amplifying it (if you want heat to flow) or suppressing it (if you want heat to stay). This section applies the physics of the entire chapter to a set of familiar devices and natural phenomena.

1. The vacuum flask

The vacuum flask (thermos flask) is the most elegant thermal device in everyday life. It suppresses all three heat-transfer mechanisms simultaneously, using a different physical strategy for each.

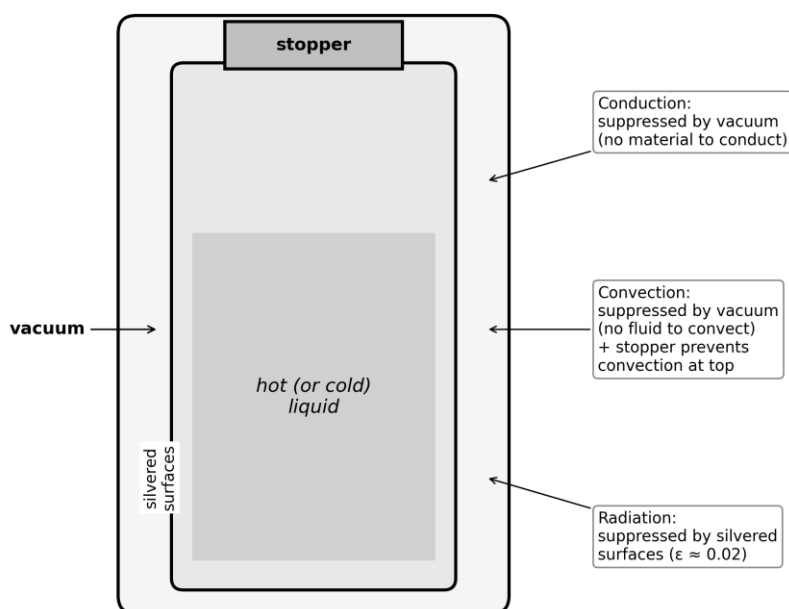


Figure: Cut-away view of a vacuum flask. The vacuum between the double walls eliminates conduction and convection. The silvered inner surfaces reduce radiation to a minimum. The stopper prevents convection at the top.

Conduction is suppressed by the vacuum between the double walls. With no material in the gap, there are no molecules to jostle their neighbours, and heat cannot conduct across. The only conduction paths remaining are through the glass walls themselves (which are thin) and through the stopper at the top.

Convection is suppressed by the same vacuum. With no air or other fluid in the gap, there is no medium to carry heat by bulk flow. The stopper at the top prevents warm air from rising out of the flask and being replaced by cooler air from the room.

Radiation is suppressed by silvering the inner surfaces of the double walls. Polished silver has an emissivity of only about 0.02, so the hot inner surface emits very little radiation outward, and the cold outer surface emits very little radiation inward. By Prevost's theory, low emissivity reduces both emission and absorption by the same factor, suppressing the net radiative exchange in both directions.

The result is a container that can keep a hot drink warm for many hours, or keep a cold drink cold for equally long. The flask does not "know" whether it is keeping heat in or keeping heat out; it simply reduces the rate of heat transfer in both directions by the same set of mechanisms.

2. The blanket

A blanket does not generate heat. It works by reducing the rate at which heat escapes from the body beneath it. The primary mechanism is **trapping air**: the loose weave of a blanket creates pockets of still air between the fibres, and air is one of the poorest thermal conductors. These trapped-air pockets *suppress both conduction* (air conducts poorly) and *convection* (the air pockets are too small for convection currents to develop). The fabric itself also *interrupts the direct radiation* path between the warm body and the cold room.

The same physics explains why a blanket can keep ice from melting. Wrapping ice in a blanket does not cool the ice; it reduces the rate at which heat from the warm room reaches the ice. The blanket suppresses the same three mechanisms in the same way, regardless of the direction of heat flow. This is a direct consequence of Fourier's law and Prevost's theory: anything that reduces the heat-transfer rate in one direction reduces it equally in the other.

3. Birds and animals in cold weather

When the temperature drops, birds fluff up their feathers and many animals curl into a ball. Both behaviours reduce heat loss, but by different physical strategies.

Fluffing feathers traps a thicker layer of still air close to the bird's body. Air is a poor conductor, so a thicker air layer increases the effective insulation thickness, raising the thermal resistance between the bird's warm skin and the cold surroundings. The trapped air also cannot convect because the feathers break it into small, isolated pockets. The bird has, in effect, put on a thicker blanket.

Curling into a ball reduces the surface area exposed to the cold environment. A sphere has the smallest surface-area-to-volume ratio of any shape, so by curling, an animal minimises the total area through which heat can escape by convection and radiation. It also reduces the area exposed to wind, suppressing forced convection. The physics is Stefan-Boltzmann (radiated power is proportional to surface area) and Newton's law of cooling (convective heat loss is proportional to surface area through the constant k , which depends on A).

4. The domestic “radiator”

The device commonly called a “radiator” in homes and offices is misnamed. It heats a room primarily by natural convection, not by radiation. Hot water flows through the metal fins of the radiator, heating the air in contact with them. The warm air rises (natural convection), travels across the ceiling, cools, descends along the far wall, and returns along the floor to the radiator, exactly as described in the convection section of this chapter. Radiation from the radiator surfaces does contribute, but it is a minor fraction of the total heat output because the radiator surface temperature (typically 50 to 70°C) is low enough that radiation is small compared to the vigorous convection current.

The name “radiator” is a historical holdover from the early days of central heating, when the distinction between convection and radiation was not yet well understood. In modern engineering, the same device is sometimes called a “convector”, which is more accurate.

5. The greenhouse effect

The glass of a greenhouse is transparent to short-wavelength visible radiation from the Sun but opaque to the longer-wavelength infrared radiation emitted by the warm soil and plants inside. Sunlight passes through the glass and warms the interior. The warmed interior radiates infrared back toward the glass, but the glass absorbs it and re-radiates some of it back inward instead of letting it escape. The interior reaches a higher equilibrium temperature than it would without the glass, because the outgoing radiation is partially blocked.

Earth's atmosphere behaves in a similar way. Greenhouse gases (mainly water vapour and carbon dioxide) are transparent to incoming short-wavelength solar radiation but absorb outgoing long-wavelength infrared radiation from the warm surface. The absorbed radiation is re-emitted in all directions, including back toward the surface, raising the equilibrium surface temperature. As derived in the radiation section, Earth's equilibrium temperature without an atmosphere, after accounting for the reflection of sunlight by clouds and ice (the albedo), would be about 255K (−18°C). The observed mean surface temperature is about 288K (15°C). The 33K greenhouse lift is what makes the planet habitable: without it, the entire surface would be permanently below the freezing point of water.

The important word in this description is equilibrium. The greenhouse effect does not cause an unlimited rise in temperature. It shifts the equilibrium upward: the surface warms until the rate of infrared emission

from the top of the atmosphere equals the rate of incoming solar absorption. At that new, higher equilibrium, the planet is in radiative balance again. Increasing the concentration of greenhouse gases shifts the equilibrium to a higher temperature, but the system always finds a new balance. The albedo adds a second layer of complexity: changes in cloud cover or ice extent alter how much sunlight is reflected, which in turn shifts the equilibrium temperature. The interplay between greenhouse warming and albedo reflection is what determines our planet's actual climate.

6. The solar water heater

A solar water heater uses a flat absorber panel, typically painted black (high absorptivity, high emissivity in the infrared), mounted on a roof. Sunlight heats the panel, which in turn heats water flowing through pipes bonded to its surface. The heated water rises naturally to a storage tank mounted above the panel (natural convection, thermosiphon effect), and cooler water from the bottom of the tank flows down to the panel to replace it. No pump is needed: the circulation is driven entirely by the density difference between the hot water rising from the panel and the cooler water descending from the tank.

Three mechanisms work together: radiation from the Sun heats the panel; conduction through the panel material transfers heat to the water; and natural convection circulates the water between the panel and the tank. The system is a self-sustaining heat engine that operates whenever the Sun is shining, converting solar radiation into stored thermal energy with no moving parts and no external power.

7. Why tile feels colder than wood

Ceramic tile and wooden floorboards in the same room are at the same temperature. Yet tile feels cold to bare feet while wood feels warm. The explanation is conduction: tile has a much higher thermal conductivity than wood (roughly $k \approx 1.0 \text{ W m}^{-1} \text{ K}^{-1}$ for tile versus $k \approx 0.15 \text{ W m}^{-1} \text{ K}^{-1}$ for wood). So the tile conducts heat away from the foot about seven times faster than wood does. The nerves in the foot respond to the rate of heat loss, not to the floor temperature. A high rate of heat loss triggers a strong "cold" sensation, even though the floor itself is at room temperature.

This is a pure conduction effect. The temperature is the same; the thermal conductivity is different. The lesson: the human body does not measure temperature; it measures heat flux.

Physics earns its keep when it explains the world around you. Let us put these ideas to work.

REAL Example 33

(a) Explain why birds swell their feathers in cold weather. State which mechanism of heat transfer is primarily reduced and what physical change achieves the reduction. (b) Explain why many animals curl into a ball when cold. State which mechanisms of heat transfer are reduced and what physical change achieves the reduction.

Solution

(a) When a bird swells its feathers, the feathers stand away from the body, trapping a thicker layer of air between the feathers and the skin. Air has very low thermal conductivity, so a thicker air layer increases the thermal resistance between the bird's warm body and the cold surroundings. The mechanism primarily reduced is conduction through the insulating air layer. The trapped air is also broken into small pockets that are too narrow for convection currents to develop, so convection through the feather layer is suppressed as well. The physical change is an increase in the effective insulation thickness.

(b) When an animal curls into a ball, it reduces the surface area exposed to the cold surroundings. A sphere has the smallest surface-area-to-volume ratio of any shape. The mechanisms reduced are convection (less surface area for warm air to rise from) and radiation (less surface area to emit radiation, since radiated power is proportional to the emitting area). Curling also reduces exposure to wind, suppressing forced convection. The physical change is a reduction in the effective surface area.

Making Sense of the Answer: Both behaviours achieve the same goal (reducing heat loss) by different physical means: the bird increases insulation thickness (raises thermal resistance); the animal decreases emitting area (lowers the surface factor in both Stefan-Boltzmann and Newton's law). Neither behaviour generates heat; both work by reducing the rate at which existing body heat escapes.

Think Like a Physicist: When answering an application question, always name three things: (1) which mechanism of heat transfer is affected (conduction, convection, radiation, or a combination), (2) what physical quantity in the relevant equation is changed (thickness, area, conductivity, emissivity), and (3) in

which direction the change goes (increased insulation, decreased area, lowered emissivity). This three-part structure maps directly to the marking scheme and guarantees full marks.

HOT Example 34

A room is heated by a domestic iron radiator through which hot water flows at a rate of 0.12kg s^{-1} . The temperature drop of the water between the inlet and the outlet of the radiator is 6.0K . The radiator has a total surface area of 1.5m^2 and its iron walls have a thickness of 2.0mm . The thermal conductivity of iron is $k = 80\text{W m}^{-1}\text{K}^{-1}$ and the specific heat capacity of water is $c = 4200\text{J kg}^{-1}\text{K}^{-1}$. (a) Calculate the rate at which heat is supplied to the room. (b) Calculate the mean temperature difference across the iron walls of the radiator.

Solution

(a) The rate of heat supplied equals the rate at which the water loses thermal energy as it flows through the radiator:

$$\begin{aligned}\frac{dQ}{dt} &= \frac{m}{t} c \Delta\theta \\ &= 0.12\text{kg s}^{-1} \times 4200\text{J kg}^{-1}\text{K}^{-1} \times 6.0\text{K} = 3024\text{W} \approx 3.0\text{kW}\end{aligned}$$

The radiator supplies about 3.0kW of heat to the room.

(b) This heat must conduct through the iron walls of the radiator to reach the room air. By Fourier's law:

$$\frac{dQ}{dt} = \frac{kA\Delta\theta}{x}$$

Rearranging for the temperature difference across the walls:

$$\Delta\theta = \frac{\frac{dQ}{dt} \times x}{kA} = \frac{3024\text{W} \times 2.0 \times 10^{-3}\text{m}}{80\text{W m}^{-1}\text{K}^{-1} \times 1.5\text{m}^2} = 0.050\text{K}$$

The mean temperature difference across the iron walls is only about 0.050K , or 50 millikelvin.

Making Sense of the Answer: A temperature difference of just 0.05K across the iron walls confirms what we observed in the radial-conduction section: metal walls are essentially transparent to heat. The thermal resistance of 2mm of iron is so tiny that it barely impedes the 3kW heat flow. The real bottleneck for heating the room is the convective heat transfer from the outer surface of the radiator to the room air, not the conduction through the metal. This is why radiator design focuses on maximising surface area (more fins) rather than reducing wall thickness.

Think Like a Physicist: This problem combines two different heat-transfer equations (water calorimetry for part (a), Fourier conduction for part (b)) on the same device. The key insight is that the heat-flow rate is the same in both equations: what the water loses is what the walls must conduct. Linking two equations through a shared quantity is a common exam technique, and it always starts with identifying what quantity is the same in both.

This chapter has taken the three heat-transfer mechanisms from their definitions to their mathematical laws, from their laboratory measurements to their everyday applications. Conduction, convection, and radiation are not three separate subjects; they are three faces of the same physical question: how does thermal energy move from one place to another? In every real situation, all three act together, and the physicist's job is to identify which one dominates and why. The tools built in this chapter, from Fourier's law to Stefan-Boltzmann's law to Newton's law of cooling, are the instruments for answering that question in any context the examination or the world can present. The miscellaneous worked examples that follow put all of these instruments on the table at once.

MISCELLANEOUS WORKED EXAMPLES ON HEAT TRANSFER**Example 35**

(a) A brick wall separating a warm room from the cold outdoors feels cold to the touch even though the room air is at 22°C. Explain which heat-transfer mechanism is responsible for the wall surface being below room temperature and why the wall surface cannot be at the same temperature as the room air.

(b) A brick wall of thickness 20cm, thermal conductivity $k = 0.72 \text{ Wm}^{-1}\text{K}^{-1}$, and area 10 m^2 separates a room at 22°C from outside air at 5°C. The inner surface has a convective film resistance of $0.12 \text{ m}^2\text{KW}^{-1}$ and the outer surface has a convective film resistance of $0.06 \text{ m}^2\text{KW}^{-1}$. (i) Calculate the total thermal resistance of the wall including both surface films. (ii) Calculate the rate of heat loss through the wall.

Solution

(a) The room air transfers heat to the wall surface by convection. This convective boundary layer has a finite thermal resistance (called the surface film resistance), meaning there must be a temperature difference between the room air and the wall surface to drive the heat flow. The wall surface is therefore always slightly cooler than the room air. The larger the rate of heat flow through the wall, the larger this temperature drop across the film. The wall cannot be at room temperature because that would imply zero temperature difference across the film, which would mean zero heat flow, contradicting the fact that heat is continuously flowing through the wall to the colder outdoors.

(b) Three thermal resistances in series: inner film, brick wall, outer film.

$$\text{Brick wall insulance: } \frac{x}{k} = \frac{0.20 \text{ m}}{0.72 \text{ Wm}^{-1}\text{K}^{-1}} = 0.278 \text{ m}^2\text{KW}^{-1}$$

$$\text{Total insulance in series: } (0.12 + 0.278 + 0.06) \text{ m}^2\text{KW}^{-1} = 0.458 \text{ m}^2\text{KW}^{-1}$$

Total resistance for area $A = 10 \text{ m}^2$:

$$R = \frac{r}{A} = \frac{0.458 \text{ m}^2\text{KW}^{-1}}{10 \text{ m}^2} = 0.0458 \text{ KW}^{-1}$$

(ii) Rate of heat loss:

$$\frac{dQ}{dt} = \frac{\Delta T}{R} = \frac{22^\circ\text{C} - 5^\circ\text{C}}{0.0458 \text{ KW}^{-1}} = \frac{17 \text{ K}}{0.0458 \text{ KW}^{-1}} = 371 \text{ W}$$

Example 36

(a) A student claims a vacuum flask “keeps things hot.” Another student says it “keeps things cold.” Explain how both statements are correct using the physics of the three heat-transfer mechanisms.

(b) Tea at 90°C in a thermos flask is placed in a room at 20°C. After 3 hours the tea has cooled to 85°C. (i) Find the cooling constant λ . (ii) How long does it take for the tea to reach 50°C? (iii) What is the half-cooling time?

Solution

(a) The vacuum flask suppresses all three heat-transfer mechanisms: the vacuum eliminates conduction and convection across the gap, the silvered surfaces reduce radiation, and the stopper prevents convection at the opening. These mechanisms operate identically regardless of the direction of heat flow. When a hot liquid is inside, the flask slows heat escaping outward. When a cold liquid is inside, the same barriers slow heat entering from the warm surroundings. The flask does not know whether it is keeping heat in or keeping heat out; it simply reduces the rate of heat transfer in both directions by the same set of mechanisms.

(b)(i) Using the logarithmic form of Newton’s law:

$$\lambda = \frac{\ln\left(\frac{90^\circ\text{C} - 20^\circ\text{C}}{85^\circ\text{C} - 20^\circ\text{C}}\right)}{3 \text{ hr}} = 0.0247 \text{ hr}^{-1}$$

(ii) Time to reach 50°C:

$$t = \frac{\ln\left(\frac{90^\circ\text{C} - 20^\circ\text{C}}{50^\circ\text{C} - 20^\circ\text{C}}\right)}{\lambda} = \frac{\ln\left(\frac{70 \text{ K}}{30 \text{ K}}\right)}{0.0247 \text{ hr}^{-1}} = 34.3 \text{ hr}$$

(iii) Half-cooling time:

$$t_{1/2} = \frac{\ln 2}{\lambda} = \frac{0.693}{0.0247\text{hr}^{-1}} = 28.1\text{hr}$$

Example 37

(a) A blacksmith heats a piece of iron continuously from room temperature. Describe and explain the colour changes the blacksmith observes, linking each change to the behaviour of the spectral emission curve.

(b) A tungsten filament operates at 2800K. (i) Find the peak emission wavelength. (ii) The visible spectrum spans approximately 400nm to 700nm. Is the peak within this range? What does this tell you about the efficiency of incandescent bulbs? (iii) If the filament temperature is increased to 3500K, by what factor does the total radiated power per unit area increase?

Solution

(a) At room temperature the iron emits only in the infrared and appears dark. As it heats, the peak of its spectral emission curve shifts to shorter wavelengths (Wien's law). At about 800K the peak enters the far-red edge of the visible spectrum and the iron glows dull red. Further heating shifts the peak through the visible range: bright red, orange, then yellow. At each stage the peak also grows taller (as the fifth power of temperature), so the iron becomes visibly brighter. At very high temperatures the peak moves into the blue region, but because the curve is broad, significant radiation is emitted at all visible wavelengths simultaneously, and the iron appears white rather than blue.

(b)(i) Wien's law:

$$\lambda_{\max} = \frac{b}{T} = \frac{2.90 \times 10^{-3} \text{ mK}}{2800\text{K}} = 1036\text{nm}$$

(ii) The peak is at 1036nm, which is in the near-infrared, well outside the visible range (400-700nm). Most of the filament's radiated energy comes out as invisible heat rather than visible light. This makes incandescent bulbs very inefficient as light sources: only a small fraction of the total radiated power falls within the visible spectrum.

(iii) Stefan-Boltzmann power ratio:

$$\frac{P_2}{P_1} = \left(\frac{3500\text{K}}{2800\text{K}}\right)^4 = (1.25)^4 = 2.44$$

The total radiated power per unit area increases by a factor of about 2.4.

Example 38

(a) A student performing Searle's experiment removes the lagging from the conducting bar to "see what happens." Explain how this error affects the calculated value of thermal conductivity and in which direction (too high or too low).

(b) In a Searle's apparatus experiment on a brass bar, the following data are recorded: distance between thermometers $x = 25\text{cm}$, bar diameter $d = 3.0\text{cm}$, bar thermometer readings $\theta_1 = 97^\circ\text{C}$ and $\theta_2 = 84^\circ\text{C}$, water flow rate $\frac{m}{t} = 2.4\text{gmin}^{-1}$, water inlet temperature $\theta_3 = 18^\circ\text{C}$, water outlet temperature $\theta_4 = 31^\circ\text{C}$. Calculate the thermal conductivity of brass.

Solution

(a) Without lagging, heat escapes from the sides of the bar into the surrounding air. Less heat reaches the water coil at the cold end than actually entered the bar at the hot end. The water warms less than it should, so the measured heat-flow rate (from the water temperature rise) is lower than the true conduction rate through the bar. Since the calculation assumes all conducted heat reaches the water, the equation underestimates the true heat flow, and the calculated value of k comes out too low.

(b) Converting units:

$$x = 0.25\text{m}, A = \pi(0.015\text{m})^2 = 7.07 \times 10^{-4} \text{ m}^2$$

$$\frac{m}{t} = 2.4\text{gmin}^{-1} = 4.0 \times 10^{-5} \text{ kgs}^{-1}$$

From Searle's working equation:

$$k = \frac{\frac{m}{t} c(\theta_4 - \theta_3) x}{A(\theta_1 - \theta_2)}$$

$$= \frac{4.0 \times 10^{-5} \text{ kgs}^{-1} \times 4200 \text{ Jkg}^{-1}\text{K}^{-1} \times 13\text{K} \times 0.25\text{m}}{7.07 \times 10^{-4} \text{ m}^2 \times 13\text{K}} = 59 \text{ Wm}^{-1}\text{K}^{-1}$$

The thermal conductivity of brass is approximately $59 \text{ Wm}^{-1}\text{K}^{-1}$.

Example 39

(a) In the Lee's disc method, the brass slab Y must be cooled alone after the steady state has been reached. Explain why the cooling curve cannot be obtained while the steam chest is still in contact with the sample.

(b) A Lee's disc experiment on a cardboard disc gives: disc thickness $x = 3.2\text{mm}$, disc diameter $d = 10.0\text{cm}$, steam chest temperature $\theta_1 = 99^\circ\text{C}$, brass slab temperature $\theta_2 = 54^\circ\text{C}$, mass of brass slab $m = 0.80\text{kg}$, specific heat of brass $c = 380 \text{ Jkg}^{-1}\text{K}^{-1}$, gradient of cooling curve at θ_2 is 0.32 Kmin^{-1} . Calculate the thermal conductivity of cardboard.

Solution

(a) While the steam chest is in contact, heat is continuously flowing into slab Y through the sample, maintaining Y at its steady-state temperature. The slab is not cooling; it is in thermal equilibrium between the heat arriving from the sample and the heat it loses to the air. To measure the rate at which Y loses heat at temperature θ_2 , we need Y to be cooling freely with no heat input. Removing the steam chest and sample ensures that the only process occurring is heat loss from Y to the surroundings, and the gradient of the resulting cooling curve at θ_2 gives exactly the rate of heat loss that was, during the steady state, being supplied by conduction through the sample.

(b) Converting units:

$$x = 3.2 \times 10^{-3} \text{ m}, A = \pi(0.050\text{m})^2 = 7.854 \times 10^{-3} \text{ m}^2$$

$$\frac{a}{b} = 0.32 \text{ Kmin}^{-1} = 5.33 \times 10^{-3} \text{ Ks}^{-1}$$

From Lee's disc working equation:

$$k = \frac{mc \frac{a}{b} x}{A(\theta_1 - \theta_2)} = \frac{0.80\text{kg} \times 380 \text{ Jkg}^{-1}\text{K}^{-1} \times 5.33 \times 10^{-3} \text{ Ks}^{-1} \times 3.2 \times 10^{-3} \text{ m}}{7.854 \times 10^{-3} \text{ m}^2 \times 45\text{K}} = 0.015 \text{ Wm}^{-1}\text{K}^{-1}$$

The thermal conductivity of cardboard is approximately $0.015 \text{ Wm}^{-1}\text{K}^{-1}$.

Example 40

(a) A nurse argues: "The patient loses heat by radiation only to objects colder than themselves." Evaluate this claim using Prevost's theory of heat exchange.

(b) A metal body of surface area 0.050m^2 and emissivity 0.60 sits in surroundings at 300K. (i) Calculate the net rate of radiative heat loss when the body is at 600K, at 400K, and at 310K. (ii) At which of these temperatures does the surroundings' contribution to the absorbed power exceed 10% of the body's emitted power? (iii) Below what excess temperature would Newton's law be a reasonable approximation?

Solution

(a) The claim is wrong. Prevost's theory states that every body emits and absorbs radiation continuously and simultaneously, regardless of the temperatures of the surrounding objects. The patient emits radiation toward all objects in the room, hot or cold, and simultaneously absorbs radiation from all of them. The net heat loss is the difference between emitted and absorbed power. A patient in a room full of objects at the same temperature still emits radiation; the reason there is no net loss at equilibrium is that the patient absorbs exactly as much as they emit, not that they stop emitting. The nurse's error is confusing net transfer with emission alone.

(b)(i) Using the net Stefan-Boltzmann formula at each temperature:

$$\text{At } 600\text{K: } P_{\text{net}} = 0.60 \times 5.67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4} \times 0.050\text{m}^2 \times ((600\text{K})^4 - (300\text{K})^4) = 207\text{W}$$

$$\text{At } 400\text{K: } P_{\text{net}} = 1.701 \times 10^{-9} \times (2.56 \times 10^{10} - 8.1 \times 10^9) \text{ W} = 29.7\text{W}$$

$$\text{At } 310\text{K: } P_{\text{net}} = 1.701 \times 10^{-9} \times (9.24 \times 10^9 - 8.1 \times 10^9) \text{ W} = 1.94\text{W}$$

(ii) At 600K, the emitted power is $1.701 \times 10^{-9} \times 1.296 \times 10^{11} = 220\text{W}$ and the absorbed power is $1.701 \times 10^{-9} \times 8.1 \times 10^9 = 13.8\text{W}$. The ratio is $13.8/220 = 6.3\%$, below 10%. At 400K, emitted = 43.5W, absorbed = 13.8W, ratio = 31.7%, well above 10%. At 310K, emitted = 15.7W, absorbed = 13.8W, ratio = 87.8%. The surroundings term exceeds 10% at 400K and 310K.

(iii) Newton's law approximation requires the excess to be small compared to the absolute surroundings temperature. With surroundings at 300K, the threshold is roughly 10% of 300K = 30K. So Newton's law is reasonable for body temperatures below about 330K (excess below about 30K).

Example 41

(a) In rural Tanzania, water stored in a porous clay pot outdoors can become noticeably cooler than the surrounding air temperature, while water stored in a sealed plastic jerry can stays at air temperature. Explain why the clay pot cools the water below the air temperature, identifying the mechanism responsible.

(b) A blackbody metal sphere of surface area 0.030m^2 , mass 2.0kg , and specific heat capacity $c = 400\text{Jkg}^{-1}\text{K}^{-1}$ is at 350K in surroundings at 300K . (i) Calculate the net rate of radiative heat loss. (ii) Hence calculate the initial rate of fall of temperature of the sphere.

Solution

(a) The clay pot is porous, so water slowly seeps through the walls to the outer surface and evaporates. Evaporation requires latent heat, which is drawn from the water and the pot itself, cooling them. This heat extraction continues as long as the outer surface remains moist and the surrounding air is not saturated with water vapour. The water temperature can fall below the air temperature because the evaporative cooling removes heat faster than convection and radiation from the warm air can replace it. The system reaches a dynamic equilibrium where the rate of evaporative cooling equals the rate of heat gain from the surroundings, at a temperature below the air temperature. A sealed plastic jerry can has no evaporation from its surface, so the only heat-transfer mechanisms are conduction, convection, and radiation between the water and the surroundings, which drive the water toward air temperature, not below it.

(b)(i) Net radiative heat loss:

$$P_{\text{net}} = \varepsilon \sigma A (T^4 - T_s^4) = 1 \times 5.67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4} \times 0.030\text{m}^2 \times ((350\text{K})^4 - (300\text{K})^4) = 11.7\text{W}$$

(ii) Rate of temperature fall:

$$\text{From } P_{\text{net}} = mc \frac{d\theta}{dt};$$

$$\frac{d\theta}{dt} = \frac{P_{\text{net}}}{mc} = \frac{11.7\text{W}}{2.0\text{kg} \times 400\text{Jkg}^{-1}\text{K}^{-1}} = 0.0146\text{Ks}^{-1}$$

Example 42

(a) Two identical cups of tea start at 80°C . One is placed in a room at 10°C , the other in a room at 20°C . The cups have the same mass, surface area, and emissivity. Do they have the same half-cooling time? Explain your reasoning.

(b) A blackbody sphere of radius 5.0cm , mass 4.1kg , and specific heat capacity $c = 450\text{Jkg}^{-1}\text{K}^{-1}$ sits in a room at 300K . (i) Find the cooling constant λ for this sphere. (ii) Hence find the half-cooling time.

Solution

(a) No, they do not have the same half-cooling time. The Newton cooling constant k depends on the cube of the surroundings temperature: $k = 4\varepsilon\sigma AT_s^3$. Since the cooling constant $\lambda = k/(mc)$ depends on T_s^3 , the cup in the 10°C room (283K) has a smaller λ than the cup in the 20°C room (293K). The half-cooling time $t_{1/2} = \ln 2 / \lambda$ is therefore longer in the colder room. This may seem counterintuitive (a colder room should cool things faster), but it reflects the radiation physics: the room at 20°C radiates more heat back to the cup, but the cup also emits more net power because the Stefan-Boltzmann net formula involves T_s^3 in the linearised Newton's constant. The dominant effect is that the radiative coupling between the cup and the surroundings is stronger at higher surroundings temperatures, giving a larger λ and therefore a shorter half-cooling time.

(b)(i) Newton's law of cooling is a small-excess approximation to Stefan-Boltzmann. The **heat-loss** constant for a body cooling by radiation is:

$$k = 4\varepsilon\sigma AT_s^3$$

Surface area of the sphere:

$$A = 4\pi(0.050\text{m})^2 = 0.0314\text{m}^2$$

Substituting:

$$k = 4 \times 1 \times 5.67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4} \times 0.0314\text{m}^2 \times (300\text{K})^3 = 0.192\text{WK}^{-1}$$

The cooling constant:

$$\lambda = \frac{k}{mc} = \frac{0.192\text{WK}^{-1}}{4.1\text{kg} \times 450\text{Jkg}^{-1}\text{K}^{-1}} = 1.04 \times 10^{-4} \text{ s}^{-1}$$

(ii) The half-cooling time:

$$t_{1/2} = \frac{\ln 2}{\lambda} = \frac{0.693}{1.04 \times 10^{-4} \text{ s}^{-1}} = 6660\text{s} = 111\text{min} \approx 1.85\text{hr}$$

Example 43

(a) A body at 90°C cools freely in still air at 20°C. Explain why the rate of temperature fall is not constant even though the surroundings temperature remains at 20°C throughout.

(b) A block of ice of mass 200g at 0°C is placed in a room at 25°C. The rate of heat absorption obeys Newton's law of cooling with $k = 1.2\text{WK}^{-1}$. The specific latent heat of fusion of ice is $L_f = 334\text{kJkg}^{-1}$.

(i) Calculate the rate of heat absorption by the ice. (ii) Calculate the time taken for the ice to melt completely. (iii) Explain why the rate of heat absorption remains constant during melting.

Solution

(a) Newton's law of cooling says the rate of temperature fall is proportional to the excess temperature over the surroundings. At the start, the excess is 70°C, so the rate of cooling is large. As the body cools, the excess shrinks, so the driving force for cooling weakens and the rate of temperature fall decreases. Each degree of cooling removes some of the driving force that caused the cooling in the first place. This self-slowness is the hallmark of exponential decay: the rate depends on the current state, not on the initial state.

(b)(i) The surroundings (25°C) are hotter than the ice (0°C), so heat flows from the surroundings to the ice. The excess of the surroundings over the ice is:

$$\frac{dQ}{dt} = k(\theta_s - \theta) = 1.2\text{WK}^{-1} \times (25^\circ\text{C} - 0^\circ\text{C}) = 1.2\text{WK}^{-1} \times 25\text{K} = 30\text{W}$$

(ii) Heat needed to melt the ice:

$$Q = mL_f = 0.200\text{kg} \times 334 \times 10^3 \text{ Jkg}^{-1} = 66800\text{J}$$

$$t = \frac{Q}{\frac{dQ}{dt}} = \frac{66800\text{J}}{30\text{W}} = 2227\text{s} = 37.1\text{min}$$

(iii) During melting, the ice stays at 0°C because the absorbed heat goes into breaking intermolecular forces (latent heat) rather than raising the temperature. Since the ice temperature does not change, the excess temperature remains constant at 25°C, and Newton's law of cooling gives a constant rate of heat absorption. This is unlike normal cooling, where the temperature changes and the rate decays exponentially. During a phase change, Newton's law produces a constant heat flow, not an exponential one.

Example 44

(a) Provide a mathematical explanation for why the body's temperature θ never actually reaches the surrounding temperature θ_s on the Newton cooling curve.

(b) A double-glazed window consists of two glass panes, each 4.0mm thick ($k_{\text{glass}} = 0.80\text{Wm}^{-1}\text{K}^{-1}$), separated by a 10mm air gap ($k_{\text{air}} = 0.025\text{Wm}^{-1}\text{K}^{-1}$). The window area is 1.5m². Inside temperature 22°C, outside temperature 5°C. (i) Calculate the thermal resistance of each of the three layers. (ii) Calculate the

total thermal resistance and the rate of heat loss. (iii) Calculate the rate of heat loss through a single glass pane of thickness 8.0mm (same area) and compare.

Solution

(a) The integrated form of Newton's law is $\theta = \theta_s + (\theta_i - \theta_s)e^{-\lambda t}$. For the body to reach θ_s , we would need $(\theta_i - \theta_s)e^{-\lambda t} = 0$. Since $(\theta_i - \theta_s)$ is a positive constant and $e^{-\lambda t}$ is always positive for any finite t , their product can never be zero. The exponential function decays toward zero but never reaches it. Therefore the body approaches θ_s ever more closely but never arrives: the surroundings temperature is an asymptote, not a destination.

(b)(i) Thermal resistance of each layer ($R = x/(kA)$):

$$\text{Each glass pane: } R_{\text{glass}} = \frac{4.0 \times 10^{-3} \text{ m}}{0.80 \text{ Wm}^{-1}\text{K}^{-1} \times 1.5 \text{ m}^2} = 3.33 \times 10^{-3} \text{ KW}^{-1}$$

$$\text{Air gap: } R_{\text{air}} = \frac{10 \times 10^{-3} \text{ m}}{0.025 \text{ Wm}^{-1}\text{K}^{-1} \times 1.5 \text{ m}^2} = 0.267 \text{ KW}^{-1}$$

(ii) Total resistance and heat loss:

$$R_{\text{total}} = (3.33 \times 10^{-3} + 0.267 + 3.33 \times 10^{-3}) \text{ KW}^{-1} = 0.274 \text{ KW}^{-1}$$

$$\frac{dQ}{dt} = \frac{\Delta T}{R} = \frac{(22 - 5) \text{ K}}{0.274 \text{ KW}^{-1}} = 62 \text{ W}$$

(iii) Single 8mm glass pane:

$$R_{\text{single}} = \frac{8.0 \times 10^{-3} \text{ m}}{0.80 \text{ Wm}^{-1}\text{K}^{-1} \times 1.5 \text{ m}^2} = 6.67 \times 10^{-3} \text{ KW}^{-1}$$

$$\frac{dQ}{dt} = \frac{\Delta T}{R} = \frac{(22 - 5) \text{ K}}{6.67 \times 10^{-3} \text{ KW}^{-1}} = 2550 \text{ W}$$

Have you noticed? The double-glazed window loses 62W; the single pane loses 2550W. The double-glazing reduces heat loss by a factor of about 41. The air gap, which is only 10mm thick, provides 97.6% of the total thermal resistance.

Example 45

(a) A solar water heater panel is painted black rather than white. Explain why, discussing what happens to both the absorptivity and the emissivity of the surface, and why the increased emissivity is an acceptable trade-off.

(b) A flat-plate solar water heater has an absorber area of 2.0 m^2 with absorptivity 0.90. The solar intensity on a clear day is 800 Wm^{-2} . The tank loses heat to the surroundings (air at 25°C) by Newton's law with $k = 8.0 \text{ WK}^{-1}$. (i) Calculate the power absorbed by the panel. (ii) Find the equilibrium temperature of the water. (iii) On a cloudy day the solar intensity drops to 300 Wm^{-2} . Find the new equilibrium temperature.

Solution

(a) A black surface has high absorptivity (close to 1), meaning it absorbs most of the incoming solar radiation and converts it to thermal energy. A white surface reflects most of the sunlight and absorbs very little. Painting the panel black maximises the solar energy captured. However, by Kirchhoff's law, a good absorber is also a good emitter: the black panel will also radiate more infrared energy than a white one would. This increased emissivity means the panel loses more heat by radiation at high temperatures. The trade-off is acceptable because the dominant heat loss from a solar panel at typical operating temperatures ($40\text{--}70^\circ\text{C}$) is convection to the surrounding air, not radiation. The gain from absorbing more sunlight far outweighs the modest increase in radiative loss.

(b)(i) Power absorbed:

$$P_{\text{abs}} = 0.90 \times 800 \text{ Wm}^{-2} \times 2.0 \text{ m}^2 = 1440 \text{ W}$$

(ii) At equilibrium, power absorbed = power lost:

$$P_{\text{abs}} = k(\theta - \theta_s)$$

$$1440 \text{ W} = 8.0 \text{ WK}^{-1} \times (\theta - 25^\circ\text{C})$$

$$\theta - 25^{\circ}\text{C} = \frac{1440\text{W}}{8.0\text{WK}^{-1}} = 180\text{K}$$

$$\theta = 25^{\circ}\text{C} + 180^{\circ}\text{C} = 205^{\circ}\text{C}$$

(iii) Cloudy day:

$$P_{\text{abs}} = 0.90 \times 300\text{Wm}^{-2} \times 2.0\text{m}^2 = 540\text{W}$$

$$\theta = 25^{\circ}\text{C} + \frac{540\text{W}}{8.0\text{WK}^{-1}} = 25^{\circ}\text{C} + 67.5^{\circ}\text{C} = 92.5^{\circ}\text{C}$$

Eleven examples have stretched every tool built in this chapter across problems that no single section could have posed. The Digging Deeper questions that follow push further still.

DIGGING DEEPER EXERCISE 3

EXERCISE 3A: BINDER QUESTIONS

Question 1

A metal spoon and a wooden *mwiko* are both resting in a pot of boiling water. A cook grabs the metal spoon handle for half a second and burns her fingers, then grabs the wooden *mwiko* handle and feels no pain.

- (a) Both handles were in contact with the same boiling water. Explain why the metal handle is painful even for a brief contact while the wooden handle is comfortable.
- (b) A student says: “*The metal spoon is hotter than the wooden mwiko.*” Is this correct? Explain.

Question 2

Air has a lower thermal conductivity ($k \approx 0.026 \text{ W m}^{-1} \text{ K}^{-1}$) than felt ($k \approx 0.04 \text{ W m}^{-1} \text{ K}^{-1}$). Yet wrapping a hot pipe in felt reduces heat loss far more effectively than leaving an open air gap of the same thickness.

- (a) Explain why a material with higher thermal conductivity can be a better insulator than one with lower thermal conductivity.
- (b) Identify the heat-transfer mechanism that the felt suppresses but the open air gap does not.
- (c) Explain why fibre-based insulation materials (felt, glass wool, feathers) all work by the same physical principle.

Question 3

A wall consists of three layers of different materials in series. In the steady state, the temperature drop across each layer is different.

- (a) Using the thermal Ohm’s law analogy, explain why the temperature drop across each layer is proportional to that layer’s thermal resistance.
- (b) Which layer has the largest temperature drop: the one with the highest thermal conductivity, the lowest thermal conductivity, or the greatest thickness? Explain.
- (c) A builder argues: “*The thickest layer always provides the most insulation.*” Explain why this claim is not necessarily correct.

Question 4

A copper bar has one end in steam at 100°C and the other end in melting ice at 0°C .

- (a) If the bar is perfectly lagged along its entire length, sketch the temperature profile along the bar and explain why it has this shape.
- (b) The lagging is now removed. Sketch the new temperature profile on the same axes. Explain why the profile curves and in which direction it curves.
- (c) In the unlagged case, is the temperature at the midpoint of the bar above or below 50°C ? Explain.

Question 5

Newton’s law of cooling applies under forced convection at any excess temperature, or under natural convection when the excess temperature is below about 30°C .

- (a) A beaker of water at 65°C cools in a room at 28°C under natural convection. Identify the temperature at which the cooling enters the Newton regime.
- (b) During the early part of the cooling (when the excess exceeds 30°C), does the actual cooling rate become faster or slower than Newton’s law would predict? Explain.

Question 6

A body at 1200K loses heat by radiation in surroundings at 300K . A student says: “*If I cool the body to 600K , the net radiative heat loss will be exactly one-sixteenth of the original, because Stefan-Boltzmann’s law involves T^4 .*” Explain the error in the student’s reasoning.

Question 7

A vacuum flask minimises heat transfer by all three mechanisms.

- (a) A small hole is accidentally made in the stopper. Which mechanism of heat transfer does this most immediately increase? Explain.
- (b) Over many years the silvered coating on the inner walls deteriorates. Explain how this affects the rate of heat transfer and which mechanism is involved.

- (c) A student argues: “*This flask keeps my coffee hot, so it cannot keep cold juice cold.*” Explain the error in this reasoning.

Question 8

Two identical metal blocks are placed in the same room at 20°C. Block A starts at 80°C and block B starts at 60°C.

- Which block cools faster initially? Explain.
- Which block takes longer to come within 2°C of room temperature? Explain.
- Do the two blocks have the same half-cooling time? Explain whether your answers to (a) and (b) are contradictory.

Question 9

A premature baby in a hospital incubator can lose dangerous amounts of heat by radiation even when the air inside the incubator is at the correct temperature of 36°C.

- Using Prevost’s theory, explain how the baby can experience a net radiative heat loss even when the air temperature is correct.
- Identify the specific physical feature of the incubator that causes this problem. What temperature would the walls need to be at to eliminate the net radiative loss?
- Modern incubators use double-walled enclosures with warm water circulating between the walls. Explain how this design solves the radiation problem.

Question 10

On a cold night, a layer of ice forms on a pond. As the ice thickens over many hours, its rate of growth becomes progressively slower, even though the air temperature above and the water temperature below remain constant.

- Identify the heat-transfer mechanism that carries heat from the water below through the ice to the cold air above.
- Explain why the rate of heat flow through the ice decreases as the ice becomes thicker, even though the temperature difference across it stays the same.
- Explain why the ice therefore acts as a self-limiting insulator: the thicker it grows, the harder it becomes for it to grow further.

Question 11

A body is cooling in a room at 25°C according to Newton’s law. After some time, the body (still well above room temperature) is suddenly moved to a second room at 10°C.

- Describe what happens to the rate of temperature fall at the instant of the transfer. Does it increase, decrease, or stay the same? Explain.
- Sketch the complete cooling curve from the start in the 25°C room through the transfer and into the 10°C room, marking the transfer point and both asymptotes clearly.
- Does the body have the same cooling constant λ in both rooms? Explain.

Question 12

The International Space Station orbits Earth in the near-vacuum of space. Its electronic equipment generates significant heat that must be removed to prevent overheating.

- Explain why natural convection cannot cool the equipment inside the station, even though the cabin is filled with air at normal atmospheric pressure.
- Fans circulate air over hot components. What type of convection is this, and why does it work in microgravity when natural convection does not?
- On the outside of the station, large radiator panels radiate waste heat into space. Explain why conduction and convection are both useless for removing heat from the exterior, and why radiation is the only viable mechanism.

EXERCISE 3B: REAL QUESTIONS**Question 13**

Kipanga visits his aunt’s new house in Dodoma. Walking barefoot from the wooden bedroom floor onto the marble bathroom floor, he exclaims: “*The bathroom is freezing! Someone left the window open!*” His aunt laughs: “*The window is shut. Both floors are at the same temperature.*”

- (a) Explain to Kipanga why the marble floor feels colder than the wooden floor even though a thermometer placed on each surface would read the same temperature.
- (b) Kipanga puts on rubber sandals and steps onto the marble again. He says it feels much less cold. Explain why, identifying the physical property of rubber that is responsible.

Question 14

A cook in a Tanzanian kitchen heats a clay pot on a charcoal *jiko*. The pot has a wooden handle.

- (a) Explain why the clay body of the pot becomes extremely hot while the wooden handle stays cool enough to grip, even though both are attached to the same pot.
- (b) The cook wraps a damp cloth around the handle as an extra precaution. Using two different heat-transfer mechanisms, explain why the damp cloth keeps her hand cooler than a dry cloth would.

Question 15

A blanket keeps a sleeping person warm on a cold night. The same blanket, wrapped around a block of ice, slows the melting of the ice on a hot day.

- (a) Does the blanket generate any heat? Explain.
- (b) Identify all three heat-transfer mechanisms that the blanket reduces and explain how it reduces each.
- (c) Explain why the blanket works in both directions (keeping heat in and keeping heat out) using a single physical principle.

Question 16

A solar water heater installed on a rooftop in Mwanza uses a flat black absorber panel connected to a storage tank mounted above it. There is no electric pump.

- (a) Explain why the absorber panel is painted black rather than left as shiny aluminium.
- (b) Water circulates automatically between the panel and the tank without any pump. Identify the mechanism that drives this circulation and explain how it works.
- (c) The heater is more effective on a clear sunny day than on an overcast day at the same air temperature. Explain why.

Question 17

A student writes in an essay: *“The greenhouse effect permanently traps heat inside Earth’s atmosphere, preventing any heat from escaping to space. This means Earth’s temperature will rise indefinitely.”*

- (a) Identify the two errors in the student’s reasoning.
- (b) Explain the correct equilibrium mechanism that determines Earth’s surface temperature.
- (c) Predict what would happen to Earth’s surface temperature if all greenhouse gases were suddenly removed from the atmosphere.

Question 18

On a cold night in Arusha, a cup of hot tea at 80°C and a glass of cold water at 10°C are left on the same table in a room at 20°C.

- (a) Which object loses heat faster initially? Explain.
- (b) Which object gains heat initially? Explain.
- (c) Will both objects eventually reach the same temperature? If so, what temperature, and which one reaches it first? Explain.

Question 19

In Dar es Salaam, daytime temperatures regularly exceed 30°C and houses with corrugated iron roofs can become unbearably hot inside.

- (a) Propose one modification to the building that would reduce heat entering by conduction, and explain the physics behind your proposal.
- (b) Propose one modification that would reduce heat entering by convection, and explain the physics.
- (c) Propose one modification that would reduce heat entering by radiation, and explain the physics.

Question 20

A mother in Iringa puts a pot of *ugali* on a charcoal *jiko* to keep it warm for a latecomer. She then wraps the pot in several layers of thick newspaper.

- (a) Identify the two mechanisms of heat transfer that the newspaper wrap primarily reduces. Explain how it reduces each.

- (b) After 30 minutes the *ugali* is still hot. After 2 hours it has cooled noticeably. Using Newton's law of cooling, explain why the rate of cooling is fast at first and slows down over time, even though the newspaper stays the same.
- (c) Would wrapping the pot in aluminium foil instead of newspaper keep the *ugali* hot for longer? Explain which mechanism each material is better at suppressing.

Question 21

Kipute pours equal amounts of boiling tea into a metal mug and a ceramic mug. After five minutes she touches both mugs. The metal mug feels cooler than the ceramic mug. She says: "*The metal mug cooled the tea faster because metal has higher thermal conductivity.*"

- (a) Is Kipute correct that the tea in the metal mug cooled faster? Explain.
- (b) Kipute's reasoning uses thermal conductivity. Explain why thermal conductivity alone does not fully explain the different cooling rates. What other property of the mug material matters?
- (c) If both mugs were perfectly lagged on the outside so that no heat could escape, would the tea in both mugs be at the same temperature after five minutes? Explain.

Question 22

A builder adds a 5cm layer of insulating foam to the inside surface of a concrete roof. Before the foam was added, the occupants complained that the house was hot during the day and cold at night.

- (a) Explain how the foam reduces the rate of heat flow through the roof during the day, identifying the physical property of the foam that makes it effective.
- (b) The builder argues that the greatest temperature drop in the composite roof-plus-foam system occurs across the foam layer, not across the much thicker concrete. Explain why this is correct.
- (c) At night, when the outside is cooler than the inside, the same foam layer now keeps the house warmer. Explain why the insulation works in both directions.

EXERCISE 3C: HOT QUESTIONS

Useful constants: Stefan-Boltzmann constant $\sigma = 5.67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4}$; Wien's displacement constant $b = 2.90 \times 10^{-3} \text{ mK}$.

Question 23

A composite lagged rod consists of a copper section (length 0.50m, $k = 380 \text{ Wm}^{-1}\text{K}^{-1}$) joined to a section of unknown material (length 0.50m). Both sections have cross-sectional area $2.0 \times 10^{-4} \text{ m}^2$. The hot end of the copper is at 200°C , the cold end of the unknown material is at 20°C , and the interface temperature is 170°C . Find the thermal conductivity of the unknown material.

Question 24

A vertical soil column consists of a sand layer (53cm, $k = 0.35 \text{ Wm}^{-1}\text{K}^{-1}$) on top of a loam layer (29cm, $k = 0.75 \text{ Wm}^{-1}\text{K}^{-1}$). The cross-sectional area is 1.0 m^2 . The top surface is at 37°C and the bottom at 14°C . Find the heat flux through the column and the temperature at the sand-loam interface. After light rain, the conductivities change to $k_{\text{wet sand}} = 1.8 \text{ Wm}^{-1}\text{K}^{-1}$ and $k_{\text{wet loam}} = 2.5 \text{ Wm}^{-1}\text{K}^{-1}$. By what factor does the heat flow increase?

Question 25

A body cools from temperature θ_1 to θ_2 in time t_1 , and from θ_2 to θ_3 in time t_2 , where the temperature drops are equal: $\theta_1 - \theta_2 = \theta_2 - \theta_3$. The surroundings are at θ_s . Using Newton's law of cooling, show algebraically that $t_2 > t_1$.

Question 26

A small blackbody sphere is heated to 500K and suspended in a vacuum at $T_s = 300\text{K}$. The sphere has mass 0.50kg, specific heat capacity $c = 400 \text{ Jkg}^{-1}\text{K}^{-1}$, and surface area $A = 3.0 \times 10^{-3} \text{ m}^2$.

- (a) Write the exact differential equation for the rate of temperature change of the sphere, using Stefan-Boltzmann's law directly.
- (b) Explain why this cooling is not exponential.
- (c) Calculate the initial rate of temperature fall.

Question 27

A patient lies in a hospital room at 16°C . The patient's skin is at 34°C , total skin area 1.6 m^2 , and emissivity 0.80.

- (a) Find the net rate of radiative heat loss.
- (b) A blanket with emissivity 0.05 is provided. Estimate the new net radiative heat loss from the blanket's outer surface, assuming the blanket surface reaches 22°C.

Question 28

A blackbody copper sphere (radius 30mm, density of copper $\rho_{\text{Cu}} = 8900\text{kgm}^{-3}$, specific heat capacity $c = 400\text{Jkg}^{-1}\text{K}^{-1}$) is cooled to -73°C and placed in an enclosure at 27°C . The initial rate of temperature rise is measured as 0.012Ks^{-1} . Use this data to calculate the Stefan-Boltzmann constant σ .

Question 29

A tungsten wire of length 100cm and perimeter 0.20cm operates at 2000K with emissivity 1.0. Find the rate of radiation from the wire, the wavelength of maximum emission, and the percentage of electrical energy radiated if the wire is connected to a 60W supply.

Question 30

A furnace radiates with maximum intensity at a wavelength of $1200\text{\AA}$. A star emits its maximum intensity at $4800\text{\AA}$.

- (a) Find the temperature of each.
- (b) Calculate the ratio of the power radiated per unit area by the furnace to that by the star.
- (c) At what wavelength does the human body (37°C) radiate maximum intensity?

Question 31

An ice layer 10cm thick has formed on a pond. The air above is at -10°C and the water below is at 0°C . The thermal conductivity of ice is $k = 2.1\text{Wm}^{-1}\text{K}^{-1}$, the density of ice is $\rho_{\text{ice}} = 900\text{kgm}^{-3}$, and the latent heat of fusion is $L_f = 3.34 \times 10^5\text{Jkg}^{-1}$.

- (a) Show that the time for the ice to grow from thickness x_1 to x_2 is given by $t = \frac{\rho_{\text{ice}} L_f}{2k\Delta T} (x_2^2 - x_1^2)$.
- (b) Hence compare the time for the ice to grow from 10cm to 11cm with the time for it to grow from 1cm to 2cm.

Question 32

A hot-water pipe of inner radius 2.0cm and outer radius 3.0cm (copper, $k = 380\text{Wm}^{-1}\text{K}^{-1}$) and length 5.0m is lagged with a 4.0cm-thick layer of asbestos ($k = 0.08\text{Wm}^{-1}\text{K}^{-1}$). Water inside is at 80°C and the outer lagging surface is at 22°C . Find the rate of heat loss through the lagged pipe and the temperature at the copper-asbestos boundary.

Question 33

Earth's albedo is approximately $\alpha = 0.30$. The solar constant is $\varphi_s = 1380\text{Wm}^{-2}$.

- (a) Modify the radiative equilibrium derivation to include the albedo, and show that the equilibrium temperature becomes lower than the bare-blackbody value.
- (b) The polar ice caps have higher albedo than the oceans they cover. If global warming melts the ice caps, explain what happens to the overall albedo and what this implies for the equilibrium temperature.
- (c) Is the effect described in part (b) a stabilising (negative) or a destabilising (positive) feedback? Explain.

Chapter 4

KINETIC THEORY OF GASES

INTRODUCTION

In the two weeks since Kipanga burned his hand on a stainless steel spoon in the school kitchen, he had turned one small injury into a long career. The red stripe on his palm had faded within three days, but the story had grown more impressive every time he told it. The water had been hotter. The spoon had been heavier. A rumour had started, somewhere in Form Four, that the heat had briefly stopped his wristwatch, and Kipanga, who did not own a wristwatch, had not confirmed the rumour, but he had not denied it either. He had also begun performing, at every morning break, a slow-motion reenactment of the grab-and-yelp, complete with sound effects that improved week by week, in front of a small audience of Form Two boys who found him genuinely heroic. He had been trying, without much success, to make “Mkono wa Moto” stick as his new nickname. Kipute, who had been there during the actual quarter-second event, had listened in silence for two weeks, and her silence was running out.

She found him in the courtyard that Friday morning, in the middle of the reenactment, frozen in an exaggerated pain-face in front of three rapt Form Two boys. She waited for the sound effect to finish. She waited for the applause. She waited for Kipanga to bow, which he did. Then she spoke with the calm of a person who had been preparing her words for fourteen days.

“Kipanga. A challenge. Make this bicycle pump noticeably hot in less than two minutes. No fire. No sunlight. Nothing rubbing against anything but air.”

“I accept.”

“I have not told you the prize.”

“I accept it anyway. Mkono wa Moto does not need details.”

*“If you succeed, I will call you **Professor Kipanga** for the whole of tomorrow. Every sentence, every corridor, every canteen. No rolling of the eyes. Full seriousness. For one day. But unless you can also explain, in physics, why the pump got hot, the title is only temporary, and I can take it back at any moment.”*

“That is not how bets work.”

“This is not a bet, Kipanga. This is a chapter.”

A pump was borrowed from the gardener. A flat-tyred bicycle appeared from behind the water tank. The three Form Two boys, now promoted to witnesses, arranged themselves in a respectful semicircle. Kipanga pumped. The first twenty strokes were a performance. The next forty were the first real physical work of his morning. By the sixtieth, Kipanga had forgotten the crowd, the bet, and possibly his own name. He looked less like a man demonstrating physics and more like a man being slowly defeated by a piece of workshop equipment. He stopped. He rested his palm flat on the pump barrel. He snatched it away.

“Aa! It is hot.”

One of the Form Two boys announced the time with the seriousness of a chief judge at the Olympics. One minute, fifty-one seconds. Kipanga lifted both arms above his head in triumph and, forgetting himself completely, let out a loud shout of “MKONO WA MOTO!” which the three Form Two boys, finally, repeated back in unison. It was the first time the nickname had actually worked in fourteen days. Kipanga was having an excellent morning.

Kipute walked forward four slow, careful steps, stopped in front of him, and bowed. The bow was small, slow, and completely serious.

“Congratulations, Professor Kipanga. Now, Professor Kipanga, would Professor Kipanga be kind enough to explain to the rest of us, in Professor Kipanga’s own words, exactly how Professor Kipanga made the pump hot? Or is Professor Kipanga’s knowledge on this matter a mystery even to Professor Kipanga himself?”

The three Form Two boys looked at each other. The word “Professor” had just appeared seven times in a single paragraph, each time with complete seriousness, as if at a state ceremony. One of them sat down on the concrete. Kipanga’s raised arms came slowly down to his sides.

“The pump is hot because I pumped it.”

“Professor Kipanga pumped it,” corrected Kipute. “Even Form One could tell us that. But no fuel was added. No match was struck. The sun was not on it. And Professor Kipanga’s own rules forbid friction, because the inside of the barrel is hot too, and nothing inside rubs against anything except air. So where, Professor Kipanga, did the heat come from?”

Kipanga opened his mouth. He closed it. He opened it a second time, slower, and closed it again with the weary finality of a man who has looked into his own mind and found that, on this particular question, his own mind is not at home.

At that moment Mr. Akilikubwa walked out of the staff room with his morning tea. He had been watching from the window for the last three minutes and had heard the word Professor six times already, which he thought, as a teaching tool, was almost perfect.

“Good morning, Professor Kipanga. A good piece of experimental physics. But a title given to you by Kipute is not a gift. It is a position, and a position comes with duties. The first and only duty of a physics Professor is to explain. So where, Professor, did the heat come from?”

“I do not know, sir.”

Mr. Akilikubwa nodded, picked up the pump by the handle, and began to walk toward the physics laboratory. *“Bring the bicycle. We are going to make the same thing happen again, slowly enough to understand. Professor Kipanga, you may walk beside me, as a Professor should. Mkono wa Moto, I am sorry to say, will have to wait.”*

Chapter 2 introduced temperature, and taught us to trust a thermometer over our fingers. Chapter 3 asked how heat moves, and gave us conduction, convection, and radiation. Between them, those two chapters have told us what heat does and where it goes. They have not once told us what temperature actually is. This is the chapter where physics finally answers that question.

Welcome to Chapter 4. Your own bicycle pump has been lying to you all your life about what is actually going on inside it. It is time to meet the molecules that tell the truth.

THE MOLECULAR PICTURE

Every Tanzanian student arriving in Form Five already knows what a gas is made of. From O-Level you have met the three states of matter, the bouncing of molecules in Brownian motion, the spreading of a smell across a room by diffusion, and the empirical gas laws that link pressure, volume, and temperature. What you have not yet been asked to do is to connect those molecular pictures to the numerical laws. Boyle’s law and Charles’ law were presented to you as experimental facts: pressure falls as volume rises, volume rises as temperature rises. Nobody told you why. Nobody asked what the molecules themselves were doing to produce those relationships. That question is the whole of Chapter 4.

The scale of what we are about to describe is worth pausing over. A single mole of gas contains roughly 6.0×10^{23} molecules, the Avogadro number, which means one breath of air carries more molecules than there are grains of sand on every beach on Earth. Those molecules are not close-packed. At atmospheric pressure the average spacing between neighbours is about ten times the diameter of a molecule, so the gas is mostly empty space with tiny fast-moving particles ricocheting through it. Each molecule travels in a straight line until it strikes another molecule or a wall, and each such collision lasts an extraordinarily short time compared with the time between collisions. The bicycle pump Kipanga was hammering was not a container of some continuous smooth substance. It was a container of about 10^{22} air molecules, each one moving at something like the speed of a rifle bullet, colliding with the barrel walls billions of times per second, and pushing back on his piston the only way they knew how.

The task of this chapter is to turn that picture into arithmetic. From six simple assumptions about how those molecules behave, we will derive the pressure a gas exerts on the walls of its container. From that pressure formula combined with the ideal gas equation, we will discover something genuinely surprising: the average kinetic energy of a gas molecule depends on one thing and one thing only, the Kelvin temperature. That single result will finally tell us what temperature is. It will also explain why Kipanga’s pump got hot, why a pressure cooker works, why hydrogen leaks out of Earth’s atmosphere while oxygen stays, and why a bicycle tyre left in the Dar es Salaam sun reads a higher pressure in the afternoon than it did in the morning. Six assumptions. One derivation. A great deal of physics. The next section starts with the assumptions.

ASSUMPTIONS OF THE KINETIC THEORY OF GASES

We had our molecular picture in the previous section. Now physics demands arithmetic. To derive a formula for the pressure a gas exerts on its container, we cannot work with an arbitrary cloud of complicated molecules doing whatever they please. We need a simplified model, one that keeps the essential physics (many small particles, random motion, collisions with walls) and throws away every complication the real gas would otherwise introduce (molecules of different shapes, molecules that stick to each other, collisions that lose energy, molecules with measurable size). The simplified model is called the ideal gas, and it is defined by six assumptions.

The six assumptions

Assumption 1: *A gas consists of a very large number of identical molecules moving in random directions with a range of different speeds.*

Assumption 2: *The volume occupied by the molecules themselves is negligible compared with the volume of the container.*

Assumption 3: *Collisions between the molecules and the walls of the container, and between molecules themselves, are perfectly elastic.*

Assumption 4: *Intermolecular forces are negligible except during the brief instant of collision.*

Assumption 5: *Between collisions, each molecule moves in a straight line at constant velocity, in accordance with Newton's first law.*

Assumption 6: *The time spent in a collision is negligible compared with the time between collisions.*

Stating the six assumptions correctly gives the basic framework. But understanding each one, knowing why it is there, and knowing when it breaks down, provides a much deeper grasp of the topic. We now take each of the six in turn and ask two questions of it: why is this simplification made, and under what conditions does it fail? A later section will return to these failures and use them to explain the difference between an ideal gas and a real gas. What follows here is the foundation of that comparison.

1. A very large number of identical molecules in random motion

Why this assumption is made. The arithmetic that produces the pressure formula relies on averages. An average is only meaningful when the number of values being averaged is **extremely large**, otherwise the result would fluctuate from one instant to the next and no single pressure could be assigned to the gas. One mole contains roughly 6.0×10^{23} molecules, which makes the statistical averaging exact for all practical purposes. The requirement that the molecules be **identical** means every one has the same mass m and behaves by the same rules, so the mass can be pulled out of the averaging as a single constant. **Randomness** of direction is what gives the gas its isotropy: no direction is preferred, which allows the derivation in the next section to split the total kinetic energy equally among the three Cartesian directions, a step we will depend on heavily.

Where it breaks down. This assumption fails only in systems where the number of particles is small enough for statistical fluctuations to matter. Nanoscale devices, or the thinnest upper regions of the atmosphere where only a handful of molecules cross a region per second, can show such fluctuations. In any gas you will meet at advanced-level, this assumption is essentially exact.

2. Molecular volume negligible compared with container volume

Why this assumption is made. The derivation that follows treats each molecule as a mathematical point, so that its position is a single location and it can be anywhere in the box with equal probability. This gives the gas access to the full volume V of the container. If the molecules had appreciable size, the available volume would be less than V , because each molecule would exclude a small region around itself from being occupied by another molecule. That complication would ruin the clean derivation and force us to carry an awkward correction through every equation.

Where it breaks down. At very high pressure, molecules are forced close together, and the fraction of the total volume occupied by the molecules themselves is no longer negligible. Inside a commercial oxygen cylinder at 200 atmospheres, molecules are packed tightly enough that the available volume is noticeably less than V , and the measured pressure is higher than $PV = nRT$ would predict. The van der Waals correction,

which you may meet in chemistry, introduces a replacement volume $V - nb$ to account for this exclusion, where b is a constant fitted to each real gas.

3. Perfectly elastic collisions

Why this assumption is made. Elastic collisions conserve total kinetic energy. If collisions were inelastic, if molecules lost a little energy each time they struck the wall or each other, the gas would slowly cool to zero temperature without any heat being removed, which is plainly not what happens. A gas in thermal equilibrium, left alone in a sealed insulated container, keeps the same temperature forever. The only way the molecular model can reproduce that stubborn fact is to require every collision to conserve kinetic energy. Without this assumption, the pressure formula we will soon derive would decay with time, and so would the temperature.

Where it breaks down. No gas violates this in the conditions an A-level student will meet. A very precise treatment of real gases accounts for inelastic molecule-on-molecule collisions where some kinetic energy is briefly transferred into rotational or vibrational excitation of the colliding molecules. In polyatomic gases this energy exchange becomes important when we discuss degrees of freedom later in this chapter. For the pressure derivation itself, elastic collisions at the walls is all we need.

4. Negligible intermolecular forces except during collision

Why this assumption is made. If two molecules attracted each other from a distance, their motion between collisions would not be a straight line, they would curve toward each other. The pressure derivation would then need to track those curving paths, which is not possible in closed form. By assuming no forces act between molecules except during the instant of direct contact, we guarantee that every molecule travels in a straight line between collisions, which is what makes assumption 5 consistent and what keeps the derivation tractable.

Where it breaks down. At low temperatures, molecules move slowly enough that the small attractive forces between them, the same van der Waals forces responsible for the volume correction above, become non-negligible. The molecules tend to linger near each other in brief clusters. This is exactly what allows real gases to be liquefied at low enough temperatures: an ideal gas, by assumption, cannot be liquefied at any pressure because it has no attractive forces to pull the molecules together. A real gas can. This is why the last section of this chapter will state that ideal behaviour requires both low density and high temperature.

5. Straight-line motion between collisions

Why this assumption is made. This is a direct consequence of assumption 4 plus Newton's first law. A molecule with no force acting on it moves in a straight line at constant velocity. The derivation in the next section tracks a single molecule between two consecutive collisions with the same wall, and that tracking requires the molecule to travel a known distance (a round trip of $2L$ in a cubic box of side L) in a known time ($2L/u$ where u is its x-component of velocity). A molecule that was curving between collisions would travel further than $2L$ between hits, and the time between hits would not be given by our simple formula.

Where it breaks down. Same physics as assumption 4, because it is the same physics. Attractive forces curve the paths slightly at close range; at low temperature and high density those curves matter.

6. Collision time negligible compared with time between collisions

Why this assumption is made. During the brief instant of a wall-collision, the molecule is in contact with the wall and is experiencing a large force, it is not behaving as a free particle. During the much longer time between collisions, it is a free particle obeying Newton's first law. The derivation treats collisions as instantaneous events that reverse the x-component of velocity and leave everything else unchanged. This simplification is valid provided the molecule spends an overwhelmingly larger fraction of its time travelling freely than it does in contact with the wall. In air at room temperature, a typical molecule travels for something like 10^{-9} s between collisions and each collision itself lasts only about 10^{-13} s, so the ratio is ten thousand to one in favour of free motion. The assumption is excellent for all ordinary conditions.

Where it breaks down. At pressures so high that molecules collide almost continuously (liquids, dense gases, the interior of a giant planet), this assumption fails. The molecules are effectively always in contact, and the free-flight picture no longer applies. This is one of several reasons why liquids are not described by $PV=nRT$.

These six assumptions together define what physicists call an **ideal gas**. An ideal gas is not a real substance. It is a deliberately idealised model in which molecules are vanishingly small, never interact except on

contact, and never lose energy. Real gases behave like ideal gases only when three conditions are met together: **low pressure** (so molecular volume is negligible compared to container volume), **high temperature** (so molecules move quickly enough that attractive forces between them do not matter), and **low density** (so molecules collide rarely enough that the time-between-collisions assumption is comfortable). The last section of this chapter will return to these conditions and examine what happens when they are not met. For now, we take the six assumptions as given and put them to work: in the next section we will use them to derive the pressure of an ideal gas from first principles.

Before we leave this section, a small test. The Professor who won his title in the introduction is about to lose it to a planet. Listen carefully.

REAL Example 1

Kipute approaches Mr. Akilikubwa in the staff room during break, with a physics textbook open at a picture of the planet Jupiter.

Kipute: “Sir, this chapter says Jupiter has a gaseous atmosphere going all the way down to its core, and that the pressure near the centre is millions of atmospheres. Is it still a gas at those pressures?”

Mr. Akilikubwa: “Technically, yes. The material is gaseous in the sense that its molecules are not locked into a crystal. But it behaves nothing like the air in this room. Tell me which of our six assumptions would fail first if we tried to apply $PV=nRT$ to the centre of Jupiter.”

Kipute (thinking): “At millions of atmospheres, molecules are crushed very close together. The spacing between them must be tiny. So the volume occupied by the molecules themselves is no longer small compared with the volume they are confined to. That is assumption 2.”

Mr. Akilikubwa: “Correct. What else?”

Kipute: “When molecules are that close, they must be in contact nearly all the time. The attractive forces between them would be significant. That is assumption 4.”

Mr. Akilikubwa: “Also correct. And the consequence?”

Kipute: “Assumption 5 then fails automatically, because if forces are acting between collisions, the molecules do not move in straight lines. And if molecules are essentially always in contact, assumption 6 also fails, because there is no free-flight time between collisions.”

Mr. Akilikubwa: “So how many of our six fail at high density?”

Kipute: “Four. Two, four, five, and six.”

Mr. Akilikubwa: “Exactly four. Only assumptions 1 and 3 survive. The number of molecules is still large, and collisions are still elastic to the same degree they always are. This is why gas giants are not described by $PV=nRT$ and why the equation we are about to derive is useful only in the limit of low density and high temperature, which, thankfully for the NECTA examiners and for this chapter, describes every gas you will ever be asked to analyse at Form Five level.”

Making Sense of the Answer: The point of this dialogue is not Jupiter. The point is that the six assumptions are not independent of one another. When one fails at high density, several others fail at the same time, because they describe different aspects of the same underlying physics, namely that molecules must be widely spaced and weakly interacting for the ideal gas model to apply. Remember this pattern. It will return in the final section of this chapter when we ask under what conditions a real gas behaves ideally, and we will discover that the answer is essentially an inversion of the Jupiter case: low pressure, high temperature, low density.

Think Like a Physicist: When an exam question asks which of the six assumptions fails under some extreme condition, do not evaluate all six one by one. Find the single physical quantity that the condition is pushing, high density, low temperature, small particle count, and then ask which assumptions are built on that quantity being extreme. The assumptions cluster. High density always breaks 2, 4, 5, and 6 together, because all four of them silently assume molecules are widely spaced and independent between contacts. Low temperature always breaks 4 and 5 together, because both depend on attractive forces being weak. Learn the clusters, not the individual failures, and you will answer these questions in half the time.

PRESSURE OF AN IDEAL GAS

The six assumptions in the previous section give us everything we need to start calculating. We are about to derive the most important result in the kinetic theory of gases, a formula that tells us, from first principles and starting from nothing more than Newton's laws, what pressure a gas of many molecules will exert on the walls of its container. The derivation will follow the career of a single molecule through one round trip inside a cubical box, use Newton's second law to turn the molecule's wall-collisions into a force, extend the result to all the molecules in the container, and close with a symmetry argument that connects the motion along one axis to the motion in all three. Every stage will rest on one or more of the six assumptions. By the end, you will understand not just the final formula but what each assumption is doing along the way to keep it honest.

Consider a cubical container of side L containing a single gas molecule of mass m . The molecule is moving in some arbitrary direction with some speed, but for the moment we will focus only on the component of its motion along the x -axis, which we call u . The figure below shows the geometry. The molecule bounces back and forth between the two walls perpendicular to the x -axis, the wall at $x = 0$ on the left and the wall at $x = L$ on the right, and we are going to calculate how much force it exerts on one of those walls, say the right one.

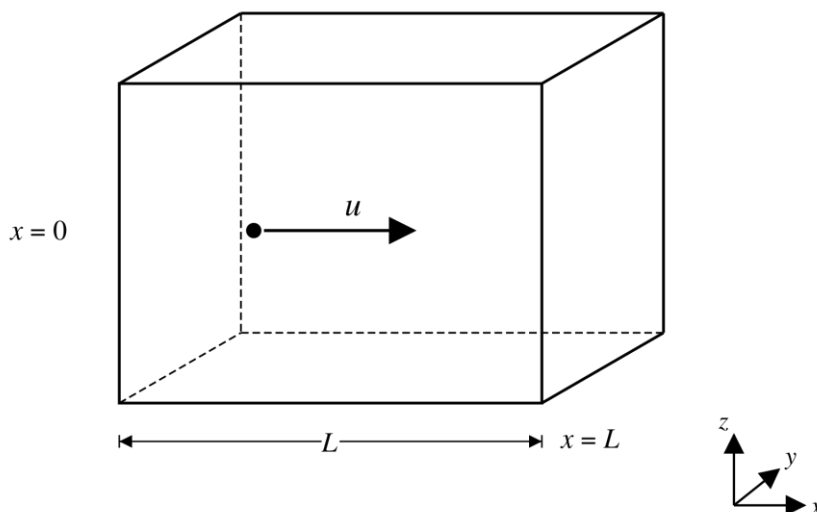


Figure: A single gas molecule of mass m inside a cubical container of side L , moving with x -component of velocity u . The molecule bounces elastically between the walls at $x = 0$ and $x = L$.

Determining the momentum change in a single collision

Consider one bounce at the wall $x = L$. Just before the bounce, the molecule's x -component of momentum is $+mu$. Just after the bounce, because collisions with the walls are perfectly elastic (assumption 3), the x -component of velocity has reversed to $-u$, and so the momentum is now $-mu$. The y - and z -components of the molecule's momentum are unchanged by this collision, because the wall acts only along the x -direction. The change in the x -component of momentum is therefore:

$$\Delta(mu) = (-mu) - (+mu) = -2mu$$

The magnitude of this change is $2mu$. By Newton's third law, the wall has received an equal and opposite impulse of $+2mu$ on every such bounce. Every time the molecule strikes the wall, it delivers a little kick of size $2mu$ to it. The question we must answer next is: *how often does this kick arrive?*

Determining the time between collisions

Between successive bounces the molecule travels in a straight line at constant velocity (assumptions 4 and 5). Starting from the wall at $x = L$ it moves off to the opposite wall at $x = 0$, a distance L , reverses there (another elastic bounce), and returns to the wall at $x = L$, another distance L . The total round-trip distance

between one hit on the right-hand wall and the next hit on the same wall is therefore $2L$. At constant x-speed u , the time this round trip takes is:

$$t = \frac{2L}{u}$$

This is the time between successive kicks delivered by our single molecule to the wall at $x = L$.

Determining the average force on the wall

We now have everything we need for Newton's second law. In its most general form, Newton's second law states that force is the rate of change of momentum. Our molecule delivers momentum of size $2mu$ to the wall every time $2L/u$, so the average force the molecule exerts on the wall is:

$$F = \frac{2mu}{\frac{2L}{u}} = \frac{mu^2}{L}$$

This is the average force that one molecule exerts on one wall, expressed purely in terms of the molecule's mass, its x-speed, and the size of the container.

A word of care about what this “average” really means. The molecule is not pressing on the right-hand wall continuously. It strikes the wall once per round trip, delivers its impulse in an instant, then spends the next $2L/u$ seconds flying away toward the opposite wall, returning, and preparing to strike again. Between consecutive hits, its contribution to the force on the right-hand wall is exactly zero. The formula $F = mu^2/L$ gives the time-averaged force over one complete round trip, smoothing this on-off delivery into a single equivalent steady force. It is this round-trip-average, and not an instantaneous force, that we will add up when we sum over all molecules in the next step. Every molecule contributes the same kind of round-trip-averaged force to every wall it can reach, and the bookkeeping works out cleanly precisely because each molecule spends equal time heading toward, and equal time heading away from, any given wall.

Determining the total force on the wall from all N molecules

A real gas does not contain a single molecule. It contains an enormous number, N of them (assumption 1). Each molecule has its own x-component of velocity: call them u_1 , u_2 , and so on up to u_N . Because the molecules do not interact except during instantaneous collisions (assumption 4), and because those collision times are negligible (assumption 6), each molecule contributes its own round-trip-averaged force to the right-hand wall independently. The total average force on the wall is simply the sum of all the individual contributions:

$$F_{\text{total}} = \frac{m(u_1^2 + u_2^2 + \dots + u_N^2)}{L}$$

This sum of squares cries out for averaging. We define the **mean square** of the x-components as *the average of the individual squared velocities*:

$$\bar{u}^2 = \frac{u_1^2 + u_2^2 + \dots + u_N^2}{N}$$

Multiplying both sides by N gives the original sum of squares as $N\bar{u}^2$, and substituting this into the force expression above gives the total force on the wall in its final compact form:

$$F_{\text{total}} = \frac{Nm\bar{u}^2}{L}$$

Determining the pressure on the wall

The right-hand wall is a square face with area L^2 . Pressure is the force per unit area of wall, averaged over time and over the whole surface. Dividing the total force above by the wall area, and recognising that $L^3 = V$ is the volume of the container:

$$P = \frac{F_{\text{total}}}{L^2} = \frac{Nm\bar{u}^2}{L^3} = \frac{Nm\bar{u}^2}{V}$$

Notice that the division by wall area only becomes physically legitimate now, at the end of the sum. A single molecule does not spread its force over L^2 of wall. It strikes a single point and moves on. Pressure is a

collective quantity: it has physical meaning only once we have added the contributions from many molecules whose collision points are scattered randomly across the whole wall. Assumption 1, random motion, is what guarantees that the scattering is uniform and that the total force really is distributed over the full L^2 .

At this stage the formula is exact. But it is not yet useful, because it has been written in terms of the mean square of the x-components of velocity alone, and the x-direction is not special. Anyone setting up the container with the axes labelled differently would have produced the same formula with the mean of $\bar{u}^2$ replaced by $\bar{v}^2$ or $\bar{w}^2$. The molecules move in three dimensions, and what we really want is a formula in terms of the full three-dimensional motion. The next step makes that connection.

Invoking isotropy

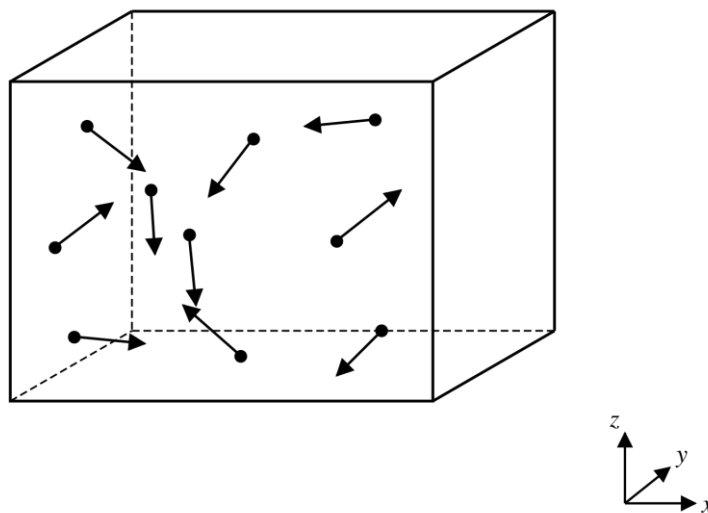


Figure: The same cubical container, with ten representative molecules drawn out of the enormous number present. Their velocity arrows point in every direction. No direction is preferred. This randomness is the engine of the isotropy argument.

For any individual molecule, let u , v , w be its three velocity components along the x-, y-, and z-axes, and let c be the magnitude of its resultant velocity. By the ordinary Pythagorean rule in three dimensions:

$$c^2 = u^2 + v^2 + w^2$$

This equation is true for every single molecule in the container. Taking the average of both sides over all N molecules (*the average of a sum is the sum of the averages*):

$$\bar{c}^2 = \bar{u}^2 + \bar{v}^2 + \bar{w}^2$$

Now we invoke the full force of assumption 1. The molecules are moving in completely random directions. No direction of space is preferred, neither the x-axis nor the y-axis nor the z-axis has any special claim on the motion. The average of any quantity built out of the x-component of velocity must therefore be identical to the average of the corresponding y- and z-quantities:

$$\bar{u}^2 = \bar{v}^2 = \bar{w}^2$$

Picture it. Inside the cube, the molecules are flying simultaneously in every imaginable direction: towards the ceiling, towards the floor, towards the left wall, towards the right wall, towards every diagonal in between. No direction gets a special invitation. On average, exactly one-third of each molecule's kinetic energy is associated with its motion along the x-axis, exactly one-third with its motion along the y-axis, and exactly one-third with its motion along the z-axis. This is not an approximation, not an estimate, not a rough rule of thumb. It is an exact consequence of the randomness we assumed at the very beginning. Randomness gives us isotropy, and isotropy gives us the factor of one-third that is about to appear in the final formula.

With the three mean squares equal, the sum on the previous line reduces immediately to $\bar{c}^2 = 3\bar{u}^2$, or rearranged:

$$\bar{u}^2 = \frac{\bar{c}^2}{3}$$

The central formula

Substituting this into the pressure formula $P = Nm\bar{u}^2/V$ gives the central result of this section:

$$P = \frac{Nm\bar{c}^2}{3V}$$

or, multiplying both sides by V :

$$PV = \frac{1}{3}Nm\bar{c}^2$$

This is the pressure of an ideal gas expressed in terms of its molecular properties. The equation is called the **kinetic equation of gases**, and everything else we will do with the molecular picture in this chapter flows from it.

The same result can be rewritten in a form that is usually more convenient in problems. The total mass of gas in the container is Nm (the number of molecules times the mass of each one), and the volume of the container is V , so the density of the gas is:

$$\rho = \frac{Nm}{V} \text{ or } Nm = \rho V$$

Substituting $Nm = \rho V$ into the pressure formula:

$$P = \frac{\rho\bar{c}^2}{3}$$

It is customary to write the mean square speed $\bar{c}^2$ as v_{rms}^2 , where v_{rms} is called the **root mean square speed** of the molecules. In that notation:

$$P = \frac{1}{3}\rho v_{\text{rms}}^2$$

which rearranges to give v_{rms} directly from the pressure and density of the gas:

$$v_{\text{rms}} = \sqrt{\frac{3P}{\rho}}$$

These three formulas are all the same physics rewritten in different clothing. Which one you reach for in a given problem is a matter of which quantities you have been given. If you know the number of molecules and the mass of each, use the first. If you know the density and the pressure, use the third to find the rms speed in one line. Nobody will ask you to reproduce the whole derivation on paper in an examination, but every formula in this summary is examinable, and most problems in this chapter will use at least one of them. Six examples follow, each exercising a different muscle.

BINDER Example 2

A sealed cubical container of volume 5.0L holds 3.0×10^{24} molecules of gas, each of mass 3.3×10^{-26} kg. The root mean square speed of the molecules is 400m/s. Calculate the pressure of the gas inside the container.

Solution

Three of the molecular properties are given to us directly: N , m , and v_{rms} . The volume is given in litres and must be converted to cubic metres before substitution. The formula of choice is therefore:

$$P = \frac{Nm v_{\text{rms}}^2}{3V}$$

Converting the volume:

$$V = 5.0 \text{ L} = 5.0 \times 10^{-3} \text{ m}^3$$

Substituting the numerical values:

$$P = \frac{(3.0 \times 10^{24})(3.3 \times 10^{-26} \text{ kg})(400 \text{ m/s})^2}{3 \times 5.0 \times 10^{-3} \text{ m}^3} = 1.056 \times 10^6 \text{ Pa} \approx 1.06 \text{ MPa}$$

The pressure inside the container is $1.056 \times 10^6 \text{ Pa}$ (or about 1.06MPa, roughly ten times atmospheric pressure).

Making Sense of the Answer: Ten atmospheres is the kind of pressure you would find in a fully inflated bicycle tyre, or in a small compressed-gas canister. It is comfortably within the range of pressures you would expect from a sealed container holding something like five moles of gas (3.0×10^{24} molecules is about five moles) in only five litres. The molecules are more tightly packed than they would be at atmospheric pressure, so they hit the walls more often per second, so the pressure is higher. Everything adds up.

Think Like a Physicist: The two most common slips in a calculation of this kind are forgetting that $1\text{L} = 10^{-3} \text{ m}^3$ (a factor of a thousand), and forgetting that v_{rms} appears in the formula squared, not to the first power. Both errors are silent until the final answer, where they produce results that are wrong by factors of thousands or by large factors of 400. Write the unit after every number as you substitute, and watch them multiply and cancel. If the units at the end come out as Pa, you have almost certainly done the calculation right.

BINDER Example 3

At standard temperature and pressure the density of nitrogen gas is 1.25 kg/m^3 . Calculate the root mean square speed of the nitrogen molecules.

Solution

The third summary formula was built precisely for this case: the pressure and density are given, and the rms speed is the unknown:

$$v_{\text{rms}} = \sqrt{\frac{3P}{\rho}}$$

At standard temperature and pressure: $T=273\text{K}$, $P=101325\text{Pa}$

Substituting the numerical values:

$$v_{\text{rms}} = \sqrt{\frac{3 \times 101325\text{Pa}}{1.25\text{kg/m}^3}} = \sqrt{243180\text{m}^2/\text{s}^2} = 493.1\text{m/s}$$

The root mean square speed of the nitrogen molecules at STP is about 493m/s.

Making Sense of the Answer: 493m/s is roughly 1800km/h, about 1.4 times the speed of sound in air (which is close to 340m/s at this temperature). That ordering is universal: the rms speed of a gas always exceeds the speed of sound in that same gas, because a sound wave is a coordinated, cooperative vibration of many molecules moving together, while v_{rms} measures the uncoordinated individual speeds in all directions. A sound wave can only travel as fast as the molecules can physically carry the disturbance, and they cannot carry it faster than they themselves are moving. The factor of roughly 1.4 between v_{rms} and the speed of sound is not accidental, it turns out to equal $\sqrt{3/\gamma}$ where γ is a ratio we will meet in a later chapter. For now, note only that the two speeds are of the same order of magnitude, with v_{rms} always the larger of the two.

Think Like a Physicist: The unit reduction $\text{Pa}/(\text{kg/m}^3) \rightarrow \text{m}^2/\text{s}^2$ is worth learning by heart. It is the reason the formula $v_{\text{rms}} = \sqrt{3P/\rho}$ returns a speed in m/s when the pressure is in pascals and the density is in kilograms per cubic metre, with no extra conversion factors. Every time you meet a formula whose units work out cleanly on the first try, take a moment to notice. SI was designed that way for exactly this convenience.

BINDER Example 4

A balloon contains helium gas at a pressure of $1.2 \times 10^5 \text{ Pa}$, and the helium has a density of 0.18 kg/m^3 .

(a) Calculate the root mean square speed of the helium atoms.

(b) The balloon is gently squeezed at constant temperature to half its original volume. Find the new pressure inside the balloon, and explain the reasoning in terms of what stays the same and what changes at the molecular level.

Solution

(a) Using $v_{\text{rms}} = \sqrt{(3P/\rho)}$:

$$v_{\text{rms}} = \sqrt{\frac{3 \times 1.2 \times 10^5 \text{ Pa}}{0.18 \text{ kg/m}^3}} = 1414 \text{ m/s}$$

The rms speed of helium atoms at this pressure and density is about 1414m/s (which is roughly 1.41km/s).

(b) At constant temperature, the average kinetic energy per molecule does not change (this is a result we will prove formally in a later section of this chapter, but for now take it as given). Since the mass of each helium atom is fixed, the rms speed must also stay the same: the molecules do not speed up or slow down when the balloon is squeezed slowly at constant temperature.

What does change is the density. The same mass of helium now occupies only half the volume, so the density has doubled. From the formula $P = (1/3)\rho v_{\text{rms}}^2$, with the v_{rms}^2 unchanged and the density doubled, the pressure must also double:

$$P_{\text{new}} = 2 \times 1.2 \times 10^5 \text{ Pa} = 2.4 \times 10^5 \text{ Pa}$$

The new pressure is $2.4 \times 10^5 \text{ Pa}$, exactly twice the original.

This is Boyle's law, which you met at O-level as an empirical fact, now derived from the molecular picture: *at constant temperature, halving the volume doubles the pressure because it doubles the density without changing the molecular speeds.*

Making Sense of the Answer: *The helium rms speed (about 1414m/s) came out much larger than the nitrogen rms speed in the previous example (493m/s at nearly the same pressure). This is not because the helium is somehow hotter or more energetic; the two gases in these two problems are at similar conditions. Helium moves faster because its atoms are lighter: a helium atom is about seven times lighter than a nitrogen molecule, so for the same average kinetic energy per molecule, the lighter helium atom must move faster by a factor of $\sqrt{7} \approx 2.6$, which is close to the ratio $1414/493 \approx 2.87$ that we actually observe. The small discrepancy comes from slightly different temperatures in the two problems. You will see this mass-dependent speed pattern over and over in this chapter: lighter molecules move faster at the same temperature. It is the reason hydrogen leaks out of Earth's atmosphere but oxygen does not.*

Think Like a Physicist: *Whenever a gas-law problem says "at constant temperature", do not reach for the calculator first. Ask: what stays the same (temperature $\rightarrow$ kinetic energy per molecule $\rightarrow v_{\text{rms}}$), what changes (volume $\rightarrow$ density), and what does the pressure formula then tell you. Three-quarters of Boyle's-law problems are solved in the head before you pick up a pen. The remaining quarter is just checking that the arithmetic matches what the physical reasoning has already told you must happen.*

REAL Example 5

A deflating balloon with a small pinhole emits a steady hiss and pushes a steady jet of gas out of the hole for several seconds, yet the kinetic theory states that the pressure of the gas inside the balloon is produced by discrete molecular collisions with the inner wall, each one a sharp, brief, isolated impulse at a single point. Explain, in terms of the molecular picture, why the pressure driving the gas through the hole is steady rather than erratic, naming the two averaging processes responsible.

Solution

The pressure driving the gas through the hole is not the force of a single collision but the time-averaged and area-averaged force delivered to the balloon wall surrounding the hole. Two averaging processes operate together.

- 1) **Averaging in time:** many millions of molecular collisions arrive at every patch of wall during each millisecond, so the running sum of those collisions over any reasonable measurement interval is smooth.
- 2) **Averaging in area:** the wall is struck by molecules across its whole surface, and by the randomness (of assumption 1) the collision points are distributed uniformly, so the force per unit area is the same everywhere.

So the steady pressure driving the steady jet is the doubly averaged result of enormously many microscopic events, and the hiss is steady for the same reason: the listener's ear averages thousands of tiny pressure fluctuations per second into a single perceived continuous sound.

Making Sense of the Answer: Pressure, temperature, density, and every other macroscopic thermal quantity in this book behaves like a smooth continuous variable even though the molecular reality beneath it is grainy and discrete. The smoothness is never the bare physics; it is the physics plus the averaging properties of whatever instrument, or organ, or thin wall of rubber, is sampling the molecular events.

Think Like a Physicist: Every time a macroscopic measurement appears smooth and steady, ask what microscopic process is being averaged to produce the smoothness and over what timescale and length scale the averaging is happening. A faster instrument, or one with a smaller sensitive face, would begin to see the fluctuations that a slower instrument hides. What counts as “smooth” is as much a property of the measurement apparatus as it is of the underlying physics.

HOT Example 6

A bicycle pump barrel is sealed at the far end and open at the near end, where a piston slides freely. A pressure gauge is fitted to the sealed end. With the piston fully withdrawn, the sealed air column occupies the full length L of the barrel and the gauge reads $2.0 \times 10^5 \text{ Pa}$. **Kipanga** pushes the piston in **gently** and **slowly**, and holds it at a position where the sealed air column has shrunk to one third of its original length. Throughout the compression, the air inside the barrel remains at steady room temperature.

- Predict the new gauge reading.
- Show by direct analysis that the per-molecule force has tripled compared with its value when the piston was fully withdrawn.
- Kipanga, in the chapter introduction, pumped his bicycle tyre vigorously for nearly two minutes and observed that the pump barrel became noticeably hot. If Kipanga had compressed this same pump barrel quickly rather than gently and slowly, would the gauge reading at the one-third-length position be equal to the answer of part (a), larger than it, or smaller? Identify which assumption of the present problem would be violated, and explain qualitatively why.

Solution

(a) At constant temperature, the average kinetic energy per molecule does not change, so the mean square speed $\bar{c}^2$ is unchanged. The pump is sealed, so the number of molecules N is unchanged. What changes is the volume of the sealed air column: the cross-sectional area A of the barrel is fixed, so the volume drops from AL to $A\left(\frac{L}{3}\right) = \frac{AL}{3} = \frac{V}{3}$. From the kinetic equation $P = \frac{Nm\bar{c}^2}{3V}$, with N , m , and $\bar{c}^2$ all unchanged and V replaced by one-third its original value, the pressure must triple:

$$P_{\text{new}} = \frac{Nm\bar{c}^2}{3 \times \frac{V}{3}} = 3 \times \frac{Nm\bar{c}^2}{3V} = 3P$$

$$P_{\text{new}} = 3 \times 2.0 \times 10^5 \text{ Pa} = 6.0 \times 10^5 \text{ Pa}$$

The gauge reading at the one-third-length position is $6.0 \times 10^5 \text{ Pa}$, three times the original, which is the molecular derivation of the same result Boyle's law would give directly.

(b) A single molecule with x -component of velocity u exerts a round-trip-averaged force on the sealed end equal to:

$$F = \frac{mu^2}{L'}$$

where L' is the length of the sealed air column, the distance the molecule must travel between successive hits on the sealed end. Initially, with the piston fully withdrawn, $L' = L$, so the initial per-molecule force is:

$$F_1 = \frac{mu^2}{L}$$

After compression, $L' = L/3$, so the new per-molecule force is:

$$F_2 = \frac{mu^2}{L/3} = \frac{3mu^2}{L} = 3F_1$$

Hence the per-molecule force has tripled.

(c) If Kipanga compresses the barrel quickly rather than gently, the gauge reading at the one-third-length position will be **larger** than the $6.0 \times 10^5 \text{ Pa}$ predicted by part (a).

The assumption being violated is that the temperature of the gas inside the barrel stays constant during the compression. In the idealised slow compression of parts (a) and (b), heat has time to flow out through the barrel walls into the surrounding air, keeping the gas at room temperature throughout. In a fast compression, there is no time for that heat flow. The mechanical work done by Kipanga's hand on the piston is pushed into the gas and cannot escape fast enough, so the temperature of the gas rises during the compression. A higher temperature means a larger mean square speed, and therefore, by the kinetic equation $P = Nm\bar{c}^2/(3V)$ with $\bar{c}^2$ now increased as well as V decreased, the pressure rises by more than the Boyle's-law factor of three. (This is exactly the physics that heated the pump barrel in the chapter introduction and briefly made Kipanga the proud owner of the title *Mkono wa Moto*. A slow, patient compression converts work into pressure alone. A fast, vigorous compression converts work into pressure and heat together.)

Making Sense of the Answer: The two parts of this example have shown the same physics from two different angles. Part (a) used the finished formula $P = Nm\bar{c}^2/(3V)$ and simply noted that shrinking V by a factor of three triples P . Part (b) went all the way back to derivation of the kinetic equation of gases and showed, molecule by molecule, exactly why that factor of three appears. The reason is the round-trip distance: shortening it makes each molecule hit the sealed end more often per second, not harder per hit. It is instructive that nothing about the force of a single collision changed. The molecules are still arriving at the same $2mu$ per kick they always delivered. What changed was the rate of arrivals, because the molecules now have a shorter commute between hits.

Think Like a Physicist: Parts (a) and (b) did not compute the same quantity. Part (a) found a gauge reading; part (b) found a per-molecule force. What they share is a scaling factor: both the collective pressure and the single-molecule force triple when the column shortens to one third. This is the useful habit to carry forward. Whenever a macroscopic result surprises you, ask what it predicts at the single-molecule level and check the two agree on ratios, not on numbers. Ratios are the currency of consistency between the microscopic picture and the macroscopic one.

HOT Example 7

Nitrogen gas at standard temperature and pressure (273K, $1.013 \times 10^5 \text{ Pa}$) has density 1.25 kg/m^3 and a root mean square speed of approximately 493m/s. Consider a flat region of container wall of area 1.00 m^2 in contact with this gas.

- Calculate the total average force delivered to this one square metre of wall by the gas.
- The typical x-component of molecular velocity, averaged across only those molecules currently heading toward the wall, is approximately $v_{\text{rms}}/\sqrt{3}$. Use this to estimate the momentum kick delivered to the wall by one average molecule in a single collision, given that the molecular mass of N_2 is $4.65 \times 10^{-26} \text{ kg}$.
- Estimate the number of molecular collisions striking the one-square-metre region of wall each second. Comment on the size of your answer and what it tells you about why the pressure appears steady.

Solution

- Pressure is force per unit area, so the total average force on a one-square-metre region of wall is simply:

$$F = P \times A = 1.013 \times 10^5 \text{ Pa} \times 1.00 \text{ m}^2 = 1.013 \times 10^5 \text{ N}$$

The one square metre of wall receives, on average, 1.013×10^5 newtons of force from the gas at every instant, which is roughly the weight of ten tonnes pressing on it. (This is actually the same force on both sides, which is why the wall does not collapse inward; atmospheric pressure presses on the outside equally hard.)

- Using the given $v_{\text{rms}} = 493 \text{ m/s}$:

$$u_{\text{avg}} = \frac{493 \text{ m/s}}{\sqrt{3}} = 284.6335 \text{ m/s}$$

The momentum change delivered to the wall in one collision is $2mu$, so the average momentum kick per collision is:

$$\Delta p = 2mu_{\text{avg}} = 2 \times 4.65 \times 10^{-26} \text{ kg} \times 284.6335 \text{ m/s} = 2.65 \times 10^{-23} \text{ kgm/s}$$

Each molecule arriving at the wall delivers about 2.65×10^{-23} kgm/s of momentum in a single collision (an extraordinarily small number).

(c) The total force on the wall is the total rate of momentum delivery: force equals momentum per second. So, if n is the number of collisions per second on the one-square-metre region, and each collision delivers on average Δp of momentum, then:

$$F = n \times \Delta p$$

Rearranging and substituting:

$$n = \frac{F}{\Delta p} = \frac{1.013 \times 10^5 \text{ N}}{2.65 \times 10^{-23} \text{ kgm/s}} \approx 3.8 \times 10^{27} \text{ collisions per second}$$

Comment: the number of collisions per square metre per second is of order 10^{27} , which is extremely large. Each collision delivers a very small momentum, but their combined rate is so large that the total force per unit area is smooth and constant, which explains why the pressure of the gas appears steady even though it is produced by discrete molecular events.

Making Sense of the Answer: *The number 3.8×10^{27} per square metre per second is not a unit to be memorised; it is a picture to be kept in mind whenever the word “pressure” appears in a physics problem. What we casually call the pressure of a gas is, at the molecular level, a storm. It is four octillion separate collision events pounding every square metre of container wall every second, each one too small and too brief to be noticed individually, their combined effect smooth and steady only because there are so enormously many of them. The REAL Example 5 about the balloon, explained why this storm looks steady to a pressure gauge. This calculation has shown how intense the storm actually is. Together the two examples close out this section with a single message: pressure is not what it looks like. It is what emerges when you average.*

Think Like a Physicist: *Reversing a formula is one of the most useful and least taught skills in physics. The pressure formula $P = F/A$ is usually reached for when pressure is unknown; this example reached for it when the collision rate was unknown. Every formula in physics can be rearranged to answer a different question than the one it was originally introduced to answer. When you meet a new formula, make a habit of asking: if I knew every quantity in this equation except one, which quantities could be the unknown, and what physical question would each unknown correspond to? You will find that most formulas secretly answer half a dozen different questions depending on which variable you are solving for, and the richer your repertoire of reversals, the more problems you will find you already know how to solve.*

We have now derived the central formula (called kinetic equation of gases) of this chapter and tested it against seven different kinds of problem. The pressure of an ideal gas is set by its density and the mean square speed of its molecules. What we have not yet done, and what part (c) of HOT Example 6 had to borrow in advance, is show that the average kinetic energy per molecule depends only on the Kelvin temperature. For that we need to bring in the ideal gas equation and compare it with what we have just derived. That is the work of the next two sections.

THE IDEAL GAS EQUATION

The kinetic equation we have just derived is a beautiful piece of physics. It tells us how the pressure of a gas arises from the rain of molecular collisions on the container walls, and it gives us a formula for that pressure in terms of the number of molecules, their mass, and their mean square speed. But it has one limitation that makes it awkward for practical problems: it does not contain the temperature of the gas anywhere.

What we need is a second equation. An equation that speaks the language of pressure, volume, temperature, and the amount of gas you can measure out of a bottle. This second equation must be consistent with the kinetic equation; the two of them, placed side by side, must describe the same gas and produce answers that agree with each other. And when we put them side by side, something quietly spectacular will happen, but we will keep that surprise for the end of this section.

The second equation is called the **ideal gas equation**, and it comes not from the molecular picture but from three empirical laws that you met at O-Level. Each law describes how one pair of gas variables behaves when the others are held constant. We remind ourselves of all three, combine them into one, and give the resulting constant of proportionality a name. That is the whole plan for this section.

Three empirical gas laws

Boyle's law states that for a fixed mass of gas held at constant temperature, the volume is inversely proportional to the pressure:

$$V \propto \frac{1}{P} \quad (\text{constant } n \text{ and } T)$$

Charles' law states that for a fixed mass of gas held at constant pressure, the volume is directly proportional to the absolute temperature:

$$V \propto T \quad (\text{constant } n \text{ and } P)$$

Avogadro's law states that for a gas held at constant pressure and constant temperature, the volume is directly proportional to the amount of gas:

$$V \propto n \quad (\text{constant } P \text{ and } T)$$

Combining the three into one

Each law tells us how V depends on one variable with the others held fixed. If we now allow all three variables to change together, the combined dependence must be the product of the three separate proportionalities:

$$V \propto \frac{nT}{P}$$

To turn the proportionality into an equation we introduce a constant of proportionality, conventionally written as R , called the **universal molar gas constant** or simply the **gas constant**. Rearranging:

$$PV = nRT$$

This is the **ideal gas equation**. It is also called the **ideal gas law** and the **equation of state** of an ideal gas, and the three names are used interchangeably. The word "state" here is a technical word: it means the condition of the gas as specified by its pressure, volume, and temperature together. An equation that links all of these in one relation deserves to be called the *equation of its state*.

The value of the constant R has been measured to great precision:

$$R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$$

The same value holds for every gas that behaves ideally. It is called the **universal** molar gas constant for precisely that reason, and the universality is a deep fact: it is a statistical consequence of the assumptions we laid down at the start of this chapter, and it would not be true if different gases had different molecular forces or different collision behaviours. Nitrogen, oxygen, helium, argon, and every other ideal gas obey $PV = nRT$ with the same R , which is why one equation serves them all.

For problems that compare two states of the same fixed amount of gas, it is often more convenient to write the equation in ratio form. Dividing $PV = nRT$ by nT gives $\frac{PV}{nT} = R$, which is the same constant for every state. Therefore, comparing state 1 with state 2:

$$\frac{P_1 V_1}{n_1 T_1} = \frac{P_2 V_2}{n_2 T_2}$$

If the amount of gas does not change between the two states (the gas is sealed), $n_1 = n_2$ and they cancel, leaving the **combined gas law**:

$$\frac{P_1 V_1}{T_1} = \frac{P_2 V_2}{T_2}$$

The same equation in molecular units

The ideal gas equation as written above uses the number of moles n , which is a chemist's unit. For a physicist interested in the molecular picture, it is often more natural to use the total number of molecules N instead. The two are connected by Avogadro's constant:

$$N = n N_A \text{ or } n = \frac{N}{N_A}$$

Substituting $n = \frac{N}{N_A}$ into $PV = nRT$:

$$PV = \frac{N}{N_A} RT = NkT$$

where:

$$k = \frac{R}{N_A} = \frac{8.314 \text{ J mol}^{-1} \text{ K}^{-1}}{6.022 \times 10^{23} \text{ mol}^{-1}} = 1.381 \times 10^{-23} \text{ J K}^{-1}$$

The constant is called the **Boltzmann constant**, named after the Austrian physicist Ludwig Boltzmann who laid most of the statistical foundations of the kinetic theory in the late nineteenth century. k is the gas constant expressed per molecule rather than per mole. Wherever a piece of physics involves the thermal energy of a single particle (atom, molecule, electron, photon), k is the constant that appears. It shows up in semiconductor physics, in the Boltzmann distribution of atomic energies, in the thermal noise voltage of a resistor, and in many other places far beyond gases. It is one of the most universal constants in physics, and it was born in this chapter, as a conversion factor between two different ways of counting the amount of gas.

The critical distinction between n and N

We now have two forms of the same equation:

$$PV = nRT \quad (\text{macroscopic form})$$

$$PV = NkT \quad (\text{microscopic form})$$

They describe the same gas and they give the same answer. The only difference is what unit of gas-amount they use. n counts **moles**; N counts **molecules**. A factor of the Avogadro constant (about 10^{23}) separates them, and this is where students confuse. Substituting n into a formula that wants N , or the other way round, gives an answer wrong by twenty-three orders of magnitude. A number like that is visibly absurd (a litre of gas does not contain 0.04 molecules, nor does it contain 2×10^{22} moles), so the error is usually caught, but not always, and the correction costs time in an examination.

Temperature must always be in Kelvin

Both forms of the equation require the **absolute** temperature, which means the temperature measured in kelvin, not in degrees Celsius or degrees Fahrenheit. The conversion is the one you already know:

$$T(\text{K}) = T(^{\circ}\text{C}) + 273$$

This is the most common silent error in the whole chapter. A temperature of 27°C is not 27 K; it is 300 K. A temperature of -3°C is not a negative temperature at all in the ideal gas equation; it is 270 K. Substituting Celsius where Kelvin is wanted gives answers wrong by factors of ten or more, and unlike the n -versus- N error, this one is often silent: the answer is numerically plausible, just badly wrong. The only protection is habit. Every time you pick up a temperature, convert it to kelvin **first**, before you substitute, and write the kelvin value in your working so the examiner can see you did the conversion. If the question gives temperature in kelvin already (for example, “300 K”), write “ $T = 300\text{K}$ ” again at the start of your working so you see that it is kelvin and not something else.

The quiet surprise: comparing the two equations

We now have two equations for the same quantity PV . On one hand, from the kinetic theory:

$$PV = \frac{1}{3} Nm\bar{c}^2$$

On the other hand, from the ideal gas equation in molecular form:

$$PV = NkT$$

Both are true for the same gas in the same state. They must agree. Setting the right-hand sides equal to each other:

$$\frac{1}{3} Nm\bar{c}^2 = NkT$$

The factor of N cancels from both sides, and multiplying by $3/2$:

$$\frac{1}{2} m \bar{c}^2 = \frac{3}{2} kT$$

Look carefully at what this equation says. The left-hand side is the **average translational kinetic energy** of a **single** gas molecule (remember, $\bar{c}^2$ is the **mean** of the squares of the individual molecular speeds, so $\frac{1}{2} m \bar{c}^2$ is the **mean** kinetic energy of a molecule). The right-hand side is $\frac{3}{2} kT$, a quantity that depends on nothing but the temperature. It does not depend on the mass of the molecule, on the type of gas, on the pressure, or on the volume. Only on **T**.

This is the single most important result of the entire chapter. Temperature, on the molecular scale, is the average translational kinetic energy of the molecules, up to the constant factor $\frac{3}{2}k$. The thermometer you built in Chapter 2 was not measuring a mysterious thing called “hotness.” *It was measuring, very indirectly, the average kinetic energy of the molecules hitting it.* **Hotness**, in the language of molecules, *means molecular speed*.

What we have just discovered is that the kinetic theory (which knows nothing about temperature) plus the ideal gas equation (which knows nothing about molecules) together *imply* that **temperature is directly related to molecular kinetic energy**. Neither piece says this on its own. Only the two, placed side by side, do. This is a derived result, and it is the payoff of the entire chapter so far. The next section unpacks the result in full and shows what it means for different kinds of gas, for different temperatures, and for the kinds of molecules you meet in everyday life.

Four examples follow.

BINDER Example 8

A sealed container of volume 0.050 m^3 holds 2.0 mol of nitrogen gas (molar mass 28 g/mol) at a temperature of 300 K . Calculate: (a) the pressure of the gas, (b) the root mean square speed of the nitrogen molecules, and (c) verify numerically that the kinetic equation $PV = \frac{1}{3} Nm \bar{c}^2$ gives the same PV as the ideal gas equation.

Solution

(a) From $PV = nRT$:

$$P = \frac{nRT}{V} = \frac{(2.0 \text{ mol})(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(300 \text{ K})}{0.050 \text{ m}^3} = 9.98 \times 10^4 \text{ Pa}$$

(b) The density of the gas (total mass divided by volume) uses the total mass nM :

$$\rho = \frac{nM}{V}$$

Substituting ρ to root mean square speed formula $v_{\text{rms}} = \sqrt{\frac{3P}{\rho}}$:

$$v_{\text{rms}} = \sqrt{\frac{3P}{\frac{nM}{V}}} = \sqrt{\frac{3PV}{nM}}$$

Substituting the numerical values:

$$v_{\text{rms}} = \sqrt{\frac{3 \times 9.98 \times 10^4 \text{ Pa} \times 0.050 \text{ m}^3}{2.0 \text{ mol} \times 28 \times 10^{-3} \text{ kg/mol}}} = 517 \text{ m/s}$$

(c) For the verification step, let us compute $\frac{1}{3} Nm \bar{c}^2$ using $Nm = nM$ (the total mass of gas) and $\bar{c}^2 = v_{\text{rms}}^2$:

$$\frac{1}{3} Nm \bar{c}^2 = \frac{1}{3} (2.0 \times 28 \times 10^{-3} \text{ kg})(517 \text{ m/s})^2 = 4990 \text{ J}$$

Comparing with PV from the ideal gas equation:

$$PV = (9.98 \times 10^4 \text{ Pa})(0.050 \text{ m}^3) = 4990 \text{ J}$$

The two expressions for PV agree, confirming that the kinetic equation and the ideal gas equation describe the same physical gas.

Making Sense of the Answer: The pressure came out almost exactly one atmosphere, which is sensible for two moles of gas in fifty litres at room temperature (this is close to the molar volume at standard conditions). The rms speed of 517 m/s matches the canonical value for nitrogen at 300 K quoted in physics textbooks. The two independent calculations of PV agreeing numerically is not a coincidence; it is a check that the entire derivation of the previous section was arithmetically correct, and it is the kind of consistency check a careful student performs whenever they suspect a slip.

Think Like a Physicist: Whenever a chapter gives you two different-looking formulas for the same quantity, compute both and check they agree. This takes sixty seconds and catches arithmetic or conceptual errors immediately. The habit pays off enormously in long multi-part problems, where a silent error at step two can propagate through to the final answer without you noticing.

BINDER Example 9

Two sealed containers, each of volume 0.040 m^3 , are held at the same temperature of 300 K. Container A holds 0.50 mol of helium (molar mass 4.0 g/mol). Container B holds 0.50 mol of argon (molar mass 40 g/mol). For each container, calculate (a) the pressure, (b) the root mean square speed of the atoms. Comment on which quantities are the same in the two containers and which are not, and why.

Solution

(a) Pressure from $PV = nRT$. Both containers have the same n , V , and T , so both have the same pressure:

$$P = \frac{nRT}{V} = \frac{(0.50 \text{ mol})(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(300 \text{ K})}{0.040 \text{ m}^3} = 3.12 \times 10^4 \text{ Pa}$$

The pressure is the same in both containers, approximately $3.12 \times 10^4 \text{ Pa}$.

(b) rms speed from $\mathbf{v_{rms}} = \sqrt{\frac{3PV}{nM}} = \sqrt{\frac{3nRT}{nM}} = \sqrt{\frac{3RT}{M}}$. Helium ($M = 4.0 \times 10^{-3} \text{ kg/mol}$):

$$v_{\text{rms,He}} = \sqrt{\frac{3 \times (8.314 \text{ J mol}^{-1} \text{ K}^{-1}) \times (300 \text{ K})}{4.0 \times 10^{-3} \text{ kg/mol}}} \approx 1368 \text{ m/s}$$

Argon ($M = 40 \times 10^{-3} \text{ kg/mol}$):

$$v_{\text{rms,Ar}} = \sqrt{\frac{3 \times (8.314 \text{ J mol}^{-1} \text{ K}^{-1}) \times (300 \text{ K})}{40 \times 10^{-3} \text{ kg/mol}}} \approx 433 \text{ m/s}$$

Equal and unequal quantities in the two containers:

- The **pressures** are *equal* because pressure depends only on n , V , T (from the ideal gas equation), and these are the same for both containers.
- The **average kinetic energies per atom** are *equal*, because average KE depends only on temperature ($\frac{3}{2}kT$), and the temperatures are identical.
- The **root mean square speeds** are *unequal*, because the atoms of helium are ten times lighter than the atoms of argon, and since $\frac{1}{2}mv_{\text{rms}}^2 = \frac{3}{2}kT$, a smaller mass forces a larger v_{rms} . Numerically, $\frac{v_{\text{rms,He}}}{v_{\text{rms,Ar}}} = \sqrt{\frac{40}{4}} = 3.16$, matching the ratio $\frac{1368}{433} = 3.16$ we computed.

Making Sense of the Answer: The headline result is that two gases at the same temperature have the same average kinetic energy per molecule, but different molecular speeds, with the lighter gas moving faster. This is why helium balloons leak faster than air-filled balloons (the lighter helium atoms pass through microscopic pores in the rubber more quickly), and why hydrogen has escaped Earth's atmosphere over geological time while heavier oxygen has not.

Think Like a Physicist: When comparing two gases at the same temperature, the mass dependence of v_{rms} is the one thing that distinguishes them. Pressure depends only on n , V , T ; kinetic energy per molecule depends only on T ; the only species-dependent quantity is the rms speed, through the molar mass. If a

question asks what is different between two gases at equal temperature, the answer almost always involves v_{rms} or some consequence of it (diffusion rate, effusion rate, speed of sound, escape from gravity).

REAL Example 10

At the summit of Mount Kilimanjaro (altitude 5895m), the atmospheric pressure is approximately 48kPa, compared with 101kPa at sea level. The air temperature at the summit averages about 263K, compared with 301K at sea level in the surrounding Tanzanian lowlands. Mountaineers at the summit report significant difficulty breathing. Using the ideal gas equation in its molecular form, determine whether this difficulty is caused primarily by the lower number of air molecules per unit volume, or by the slower speeds of those molecules at the colder temperature. Justify your conclusion with an appropriate calculation.

Solution

The ideal gas equation in molecular form is $PV = NkT$, which rearranges to give the number density (molecules per unit volume):

$$\frac{N}{V} = \frac{P}{kT}$$

The ratio of number densities at the summit and at sea level:

$$\frac{\left(\frac{N}{V}\right)_{\text{summit}}}{\left(\frac{N}{V}\right)_{\text{sea}}} = \frac{\frac{P_{\text{summit}}}{kT_{\text{summit}}}}{\frac{P_{\text{sea}}}{kT_{\text{sea}}}} = \frac{P_{\text{summit}}}{P_{\text{sea}}} \times \frac{T_{\text{sea}}}{T_{\text{summit}}} = \frac{48\text{kPa}}{101\text{kPa}} \times \frac{301\text{K}}{263\text{K}} \approx 0.54$$

So the number density at the summit is about 54% of the sea-level value; there are 46% fewer air molecules per cubic metre. The temperature ratio alone ($263/301 = 0.87$) would only have reduced the number density by 13% if pressure had stayed the same. Most of the number-density drop comes from the lower pressure, not from the lower temperature.

For molecular speeds, $v_{rms} \propto \sqrt{T}$, so the ratio of rms speeds is:

$$\frac{v_{rms,\text{summit}}}{v_{rms,\text{sea}}} = \sqrt{\frac{T_{\text{summit}}}{T_{\text{sea}}}} = \sqrt{\frac{263\text{K}}{301\text{K}}} \approx 0.935$$

Air molecules at the summit move about 6.5% more slowly than at sea level, a modest change.

The difficulty of breathing at altitude is therefore primarily due to the lower **number density** of air molecules (46% fewer oxygen molecules available per breath) (and only minimally due to the slower motion of those molecules). A climber at the summit taking a normal breath draws about half as many oxygen molecules into the lungs as at sea level. The 6.5% slower molecular speeds are negligible in comparison.

Making Sense of the Answer: *The physics of altitude sickness is essentially a counting problem. Mountaineers struggle to breathe not because air molecules are cold and sluggish, but because there are far fewer of them per litre of lung volume. Supplemental oxygen tanks, used above 7000 m, work by restoring the number density of oxygen molecules reaching the lungs, not by warming them. The ideal gas equation in its molecular form, $PV = NkT$, turns what might seem a question about cold mountain air into a simple ratio of numbers of molecules per cubic metre.*

Think Like a Physicist: *Whenever pressure drops and temperature also drops, ask which is dropping faster. The combined gas law tells you that PV/T is constant, so a drop in P with a drop in T can partially cancel, leaving the number density almost unchanged, or it can compound, leaving the number density dramatically lower, depending on the relative sizes. The only way to tell which is by computing the ratio, as this example has done. Intuition is unreliable when two effects act in opposite directions on the same quantity.*

HOT Example 11

A sealed rigid gas cylinder in a storage yard holds gas at 20 °C and 12MPa. Kipute, walking past the yard one Saturday afternoon, notices that the cylinder has been standing in direct sunlight for several hours. She checks the gauge: the pressure has risen to 14MPa. The cylinder is rigid, so the volume of gas is unchanged, and it is sealed, so the amount of gas is unchanged. Find (a) the new temperature of the gas; (b) the average translational kinetic energy of a gas molecule at the original temperature of 20 °C; (c) the percentage increase in the root mean square speed of the molecules between the two states.

Solution

Convert the initial temperature to kelvin: $T_1 = 20 + 273 = 293 \text{ K}$.

(a) At constant V and constant n , $PV = nRT$ reduces to $\frac{P}{T} = \text{constant}$:

$$\frac{P_1}{T_1} = \frac{P_2}{T_2}$$

$$T_2 = T_1 \times \frac{P_2}{P_1} = 293 \text{ K} \times \frac{14}{12} = 341.8 \text{ K} \approx 342 \text{ K} = 69^\circ \text{C}$$

The new temperature is about 69°C .

(b) The average translational kinetic energy **per molecule** follows the molecular meaning of temperature established in this section, namely $\text{KE} = \frac{3}{2} kT$:

$$\text{KE} = \frac{3}{2} \times (1.38 \times 10^{-23} \text{ J K}^{-1}) \times (293 \text{ K}) = 6.07 \times 10^{-21} \text{ J}$$

Each molecule carries, on average, about 6×10^{-21} joules of kinetic energy.

This is a tiny amount individually; the energy of an entire mole of gas is obtained by multiplying by the Avogadro constant, giving roughly 3600J per mole.

(c) Since $v_{\text{rms}}^2 \propto T$, the rms speed satisfies $v_{\text{rms}} \propto \sqrt{T}$. The ratio of rms speeds is:

$$\frac{v_{\text{rms},2}}{v_{\text{rms},1}} = \sqrt{\frac{T_2}{T_1}} = \sqrt{\frac{342 \text{ K}}{293 \text{ K}}} = 1.080$$

The rms speed increases by a factor of 1.080, a percentage increase of 8.0%.

For comparison, the pressure has increased by the factor $14/12 = 1.167$, a fractional increase of 16.7%, exactly matching the temperature ratio. Pressure scales linearly with Kelvin temperature; molecular speed scales only with its square root.

Making Sense of the Answer: *The cylinder has heated by 49K, and its pressure has jumped by 2MPa. This is why gas cylinders in Tanzania are required by safety regulations to be stored in shaded areas; a cylinder that was safe at 12MPa in the morning can exceed its pressure rating by noon if left in direct tropical sunlight. Kipute, noticing the problem, is being a thoughtful physicist. If the cylinder had been full to its rated pressure to begin with, the temperature rise could have caused a rupture, which is a serious industrial hazard.*

Think Like a Physicist: *Three quantities in this example scale with Kelvin temperature in three different ways. Pressure at constant V is linear in T ; average kinetic energy per molecule is also linear in T ; rms speed is linear in the square root of T . Whenever a problem asks you to compare two thermal states, stop and identify which of these three scalings applies to each quantity it asks about. The answer then writes itself as a ratio.*

Part (b) of the last example leaned on a result, $\text{KE} = \frac{3}{2} kT$, that we met only briefly at the end of the consistency argument. That result is rich enough to deserve a section of its own: it connects temperature to molecular speeds, to the speeds of different gases at the same temperature, to escape velocities and planetary atmospheres, and to the eventual definition of internal energy for an ideal gas. The next section is devoted to unpacking it.

TEMPERATURE AND MOLECULAR KINETIC ENERGY

Chapter 2 asked what temperature is and gave an answer that was, on reflection, a little evasive. Temperature was defined operationally, as whatever a thermometer measures, with the zero point of the Celsius scale pegged to the freezing of water and the zero point of the Kelvin scale pegged to the absolute zero of thermal motion. That was enough to do thermometry, but it did not say what temperature actually is. A thermometer reading of 300K was a number on a scale; it was not an explanation. This section removes the evasion. We have now assembled enough machinery to say what temperature really is on the molecular scale, and the

answer, when it comes, is disarmingly simple. **Temperature** is *molecular translational kinetic energy, multiplied by a constant*. That is the whole of it.

The result we are about to unpack was quietly stated at the end of the previous section, when we equated the kinetic equation with the ideal gas equation. We now take it seriously, develop its immediate consequences, and use it to answer a question that has been hanging over the chapter since the first paragraph: why does Kipanga's pump get hot when he compresses the air, and what does "hot" actually mean down at the level of the molecules themselves?

The key result

From the two equations for PV:

$$\frac{1}{3} N m \bar{c}^2 = N k T$$

Cancelling N and multiplying by 3/2:

$$\frac{1}{2} m \bar{c}^2 = \frac{3}{2} k T$$

The left side is the average translational kinetic energy of a single molecule of the gas, which we will denote KE. The right side is three-halves of the Boltzmann constant times the Kelvin temperature. So:

$$KE = \frac{3}{2} k T \quad \left(\text{or } T = \frac{2}{3k} \times KE = \text{constant} \times KE \right)$$

This is the central result of the chapter. Read it slowly, because it says more than it looks like it says.

What this equation is saying

The average translational kinetic energy per molecule of an ideal gas depends on one thing and one thing only: the **Kelvin** temperature. It does not depend on the mass of the molecule. It does not depend on the chemical identity of the gas. It does not depend on the pressure. It does not depend on the volume. It does not depend on whether the gas is in a balloon, a cylinder, or a tyre. Two different gases at the **same** temperature, whether they are helium and argon, or nitrogen and xenon, or hydrogen and mercury vapour, have **exactly** the same average translational kinetic energy per molecule. There is no small correction. There is no qualifying clause. Temperature is the measure of translational kinetic energy per molecule, and nothing more.

This is a surprising claim when you first meet it, because it seems to conflict with everyday intuition. Heavier molecules should surely carry more energy at a given temperature, you might think. They don't. What does change with molecular mass is the **speed** of the molecules: *lighter molecules move faster, heavier molecules move more slowly, in just the right proportion to keep the kinetic energy $\frac{1}{2} m \bar{c}^2$ the same for both*. A hydrogen molecule and an oxygen molecule at room temperature are both carrying about 6×10^{-21} J of kinetic energy on average. The hydrogen molecule, being sixteen times lighter, moves four times faster to make up the difference. This is the numerical fact we will confirm with the second worked example of this section.

Three formulas for the root mean square speed

Rearranging the central result for the rms speed of a single molecule ($\frac{1}{2} m \bar{c}^2 = \frac{3}{2} k T$):

$$v_{\text{rms}} = \sqrt{\bar{c}^2} = \sqrt{\frac{3kT}{m}}$$

This is the first form, and it expresses the rms speed in terms of the per-molecule mass m and the Boltzmann constant k. For numerical calculations it is sometimes more convenient to use the molar mass M instead of the per-molecule mass, using $m = \frac{M}{N_A}$ and $k = \frac{R}{N_A}$:

$$v_{\text{rms}} = \sqrt{\frac{3 \times \frac{R}{N_A} \times T}{\frac{M}{N_A}}} = \sqrt{\frac{3RT}{M}}$$

This second form is the most practical for calculations, because molar masses are tabulated in every physics and chemistry reference while per-molecule masses require a division by N_A that adds work without insight. Finally, from the pressure derivation we already have the density form:

$$v_{\text{rms}} = \sqrt{\frac{3P}{\rho}}$$

Three equivalent formulas, each matched to a different question. Use the first when you know the per-molecule mass (rare). Use the second when you know the molar mass and the temperature (the usual NECTA case). Use the third when you know the pressure and the density (the usual lab case).

Total kinetic energy of a gas sample

Average kinetic energy per molecule is $\frac{3}{2}kT$. The total translational kinetic energy of a gas sample containing N molecules is obtained by multiplying:

$$KE_{\text{total}} = N \times \frac{3}{2}kT = \frac{3}{2}NkT = \frac{3}{2}nRT$$

where the last equality uses $NkT = nRT$. For one mole of gas at room temperature ($T = 298\text{K}$), the total translational kinetic energy is $\frac{3}{2}(1)(8.314)(298) \approx 3716\text{J}$, or about 3.7kJ . This is the same kind of energy as the kinetic energy of a one-kilogram mass moving at $\sqrt{2 \times 3716} \approx 86\text{m/s}$, approximately the speed of a cricket ball hurled by a fast bowler. A litre of air at room temperature carries, in its molecular motion, roughly the kinetic energy of a hard-thrown cricket ball. This is why compressing a gas heats it noticeably: that molecular energy is genuinely substantial, and mechanical work done on the gas adds to it.

With this central result now firmly in hand, let us bring it down to earth through a few carefully chosen worked examples.

BINDER Example 12

Calculate the average translational kinetic energy of a gas molecule at (a) 0°C and (b) 100°C . Find the ratio of the two energies and explain why it is not 2, despite the temperature having apparently doubled between freezing and boiling water.

Solution

Converting to kelvin: $T_1 = 273\text{K}$, $T_2 = 373\text{K}$.

(a) At 0°C :

$$KE_1 = \frac{3}{2}kT = \frac{3}{2} \times (1.381 \times 10^{-23} \text{J K}^{-1}) \times (273\text{K}) = 5.65 \times 10^{-21}\text{J}$$

(b) At 100°C :

$$KE_2 = \frac{3}{2} \times (1.381 \times 10^{-23} \text{J K}^{-1}) \times (373\text{K}) = 7.73 \times 10^{-21}\text{J}$$

Ratio:

$$\frac{KE_2}{KE_1} = \frac{373}{273} = 1.37$$

The kinetic energy at 100°C is 1.37 times the kinetic energy at 0°C , not 2 times, because the Celsius temperature of 100°C is not twice 0°C on the absolute scale. On the Kelvin scale, which is the one the kinetic theory uses, 0°C is 273K and 100°C is 373K , and 373 is only 1.37 times 273 . The Celsius scale, because its zero point is arbitrary, gives misleading ratios. This is why kinetic theory always uses Kelvin.

Making Sense of the Answer: The 37% increase in kinetic energy between freezing water and boiling water is modest; it is not the factor of two that a naive reading of “0 to 100” suggests. This is why heating a pot of water from 0°C to 100°C takes far less energy than heating it from 0K to 273K . Most of the thermal energy of a substance at everyday temperatures was already there at the freezing point.

Think Like a Physicist: Ratios of Celsius temperatures are physically meaningless. Ratios of Kelvin temperatures are the real thing. Whenever a question asks by what factor some thermal quantity has changed, first rewrite the temperatures in kelvin, then take the ratio.

BINDER Example 13

Hydrogen gas (molar mass 2.0 g/mol) and oxygen gas (molar mass 32 g/mol) are both held at a temperature of 300 K. Calculate the root mean square speed of each gas. Then calculate the average translational kinetic energy per molecule for each gas, and verify numerically that the two values are equal.

Solution

rms speeds from $v_{\text{rms}} = \sqrt{\frac{3RT}{M}}$. For hydrogen ($M = 2.0 \times 10^{-3} \text{ kg/mol}$):

$$v_{\text{rms}, \text{H}_2} = \sqrt{\frac{3 \times (8.314 \text{ J mol}^{-1} \text{ K}^{-1}) \times (300 \text{ K})}{2.0 \times 10^{-3} \text{ kg/mol}}} = 1934 \text{ m/s}$$

For oxygen ($M = 32 \times 10^{-3} \text{ kg/mol}$):

$$v_{\text{rms}, \text{O}_2} = \sqrt{\frac{3 \times (8.314 \text{ J mol}^{-1} \text{ K}^{-1}) \times (300 \text{ K})}{32 \times 10^{-3} \text{ kg/mol}}} = 484 \text{ m/s}$$

Hydrogen molecules move four times faster than oxygen molecules (ratio $1934/484 \approx 4.00$, matching $\sqrt{\frac{32}{2}}$ since oxygen is sixteen times heavier per molecule).

Kinetic energy per molecule uses $\frac{1}{2} m v_{\text{rms}}^2$.

Per-molecule masses:

$$m_{\text{H}_2} = \frac{M}{N_A} = \frac{2.0 \times 10^{-3} \text{ kg mol}^{-1}}{6.022 \times 10^{23} \text{ mol}^{-1}} = 3.32 \times 10^{-27} \text{ kg}, m_{\text{O}_2} = \frac{32 \times 10^{-3} \text{ kg mol}^{-1}}{6.022 \times 10^{23} \text{ mol}^{-1}} = 5.31 \times 10^{-26} \text{ kg}$$

$$\text{KE}_{\text{H}_2} = \frac{1}{2} \times (3.32 \times 10^{-27} \text{ kg}) \times (1934 \text{ m/s})^2 = 6.21 \times 10^{-21} \text{ J}$$

$$\text{KE}_{\text{O}_2} = \frac{1}{2} \times (5.31 \times 10^{-26} \text{ kg}) \times (484 \text{ m/s})^2 = 6.22 \times 10^{-21} \text{ J}$$

The two kinetic energies are equal to within rounding, as the theory predicts.

Understand temperature confirmation: Both can be verified independently against the temperature formula:

$$\text{KE} = \frac{3}{2} kT = \frac{3}{2} \times (1.381 \times 10^{-23} \text{ J K}^{-1}) \times (300 \text{ K}) = 6.21 \times 10^{-21} \text{ J}$$

All three numbers agree. The critical insight is confirmed: *at a given temperature, different gases have different molecular speeds, but the same kinetic energy per molecule.*

Making Sense of the Answer: Hydrogen molecules at room temperature move at nearly two kilometres per second, which is comparable to a rifle bullet at the moment it leaves the muzzle. Oxygen molecules, being sixteen times heavier, move at only a quarter of that speed, but with exactly the same kinetic energy per molecule. The speed difference is what matters for diffusion rates and for escape from gravity; the kinetic energy equality is what matters for thermal equilibrium. Both will come back in the next example.

Think Like a Physicist: When two gases are in thermal equilibrium, ask what is equal at the molecular level and what is not. Temperature is equal, therefore kinetic energy per molecule is equal. Masses are not equal, therefore speeds are not equal. Number of molecules is set by partial pressures, not by temperature. Identifying these three separately untangles most kinetic-theory problems on sight.

REAL Example 14

Earth's atmosphere contains significant amounts of oxygen and nitrogen but only trace amounts of hydrogen and helium, even though hydrogen and helium are far more abundant in the Universe as a whole. The escape velocity from Earth's gravity is approximately 11.2 km/s. Using the result of the previous example, explain why hydrogen can escape Earth's atmosphere over geological time while oxygen cannot.

Solution

The root mean square speed of hydrogen molecules at room temperature is about 1.9 km/s (from the previous example), which is approximately 17% of Earth's escape velocity of 11.2 km/s. The corresponding figure for oxygen is 0.48 km/s, only about 4% of escape velocity. Although neither rms speed exceeds escape velocity, the rms is not the full story: molecular speeds in a gas follow a distribution, with some molecules moving slower than the rms speed and some moving faster. In the high-speed tail of the distribution, a small fraction of molecules always exceeds any given speed, including the escape velocity. For hydrogen, whose rms is already 17% of escape, this fraction is non-negligible: over geological time (billions of years), enough hydrogen molecules in the high-altitude tail of the distribution exceed escape velocity, travel upward, and leave the atmosphere permanently. For oxygen, whose rms is only 4% of escape, the fraction of the distribution exceeding escape velocity is astronomically smaller, and the rate of escape is effectively zero on any timescale relevant to a planet.

It is worth understanding that: The same reasoning explains why the Moon has almost no atmosphere (its escape velocity is only 2.4 km/s, which even oxygen can exceed over geological time), why Jupiter retains even hydrogen (its escape velocity is 59.5 km/s, far above any molecular speed at its cold outer temperatures), and why the top of Mount Kilimanjaro, although thinner in number density, has the same chemical composition as the lowlands: 5895m is not nearly high enough for the escape-velocity-plus-distribution argument to matter, because molecules escape only from the very top of the atmosphere, far above any mountain.

Making Sense of the Answer: *The chemical composition of a planet's atmosphere is set by a competition between thermal kinetic energy and gravitational binding energy. Lighter molecules, which move faster at the same temperature, lose this competition over time. Earth has retained enough oxygen and nitrogen to support life largely because these molecules are heavy enough to stay gravitationally bound. Had Earth been smaller, or warmer, or both, we would not be here to have this conversation.*

Think Like a Physicist: *When an average quantity is smaller than some threshold, the tail of the distribution often matters more than the average itself. Averages describe typical molecules; tails describe rare events that can be the only thing that matters over long timescales. Atmospheric escape is a rare-event phenomenon, and so is radioactive decay, and so is thermal fluctuation in electronic noise, and so are many other phenomena in physics where the average is large but the tail is what the universe notices.*

HOT Example 15

At what temperature must nitrogen gas (molar mass 28g/mol) be heated in order for its root mean square molecular speed to equal that of hydrogen gas at 300K? Comment briefly on the size of your answer.

Solution

Using:

$$v_{\text{rms}} = \sqrt{\frac{3RT}{M}}$$

The condition we are asked to satisfy is equality of root mean square speeds between the two gases at their respective temperatures:

$$v_{\text{rms},\text{N}_2} \text{ at } T = v_{\text{rms},\text{H}_2} \text{ at } 300\text{K}$$

Squaring both sides (so we can use $v_{\text{rms}}^2 = \frac{3RT}{M}$ directly) and writing out each side:

$$\frac{3R \times T}{M_{\text{N}_2}} = \frac{3R \times 300\text{K}}{M_{\text{H}_2}}$$

The factor 3R cancels from both sides, leaving a clean ratio equation. Solving for T:

$$T = \frac{M_{\text{N}_2}}{M_{\text{H}_2}} \times 300\text{K} = \frac{28\text{g mol}^{-1}}{2.0\text{g mol}^{-1}} \times 300\text{K} = 14 \times 300\text{K} = 4200\text{K}$$

The temperature required is about fourteen times room temperature and comparable to the surface temperature of the Sun (which is about 5800K). This shows that heavy molecules would need extremely high temperatures to match the speed of light molecules at ordinary temperatures, which is why light gases such as hydrogen escape from planetary atmospheres while heavier ones are retained.

Making Sense of the Answer: *the temperature ratio $4200/300 = 14$ is exactly the molar mass ratio $28/2 = 14$, as it must be from $v_{\text{rms}} \propto T/M$. This is the same scaling law that governs atmospheric escape, stellar composition, and the speed of sound in different gases. A single quadratic proportionality, $\bar{c}^2 \propto T/M$, carries an enormous amount of astrophysics inside it.*

Think Like a Physicist: *whenever two molecular speeds are equal, the product MT must be equal. This “equal-speed” condition often appears disguised in problems about diffusion, effusion, or spectral broadening. Recognising that MT is the invariant, not M or T separately, turns a one-line calculation into a sentence.*

Having enjoyed the worked examples, we now realise that something has been quietly left unsaid. Every result in this section has been about translational motion. A molecule that only translates, like a helium atom, is fully described by the story we have just told. But many gases are made of molecules that can also rotate, like the dumbbell-shaped nitrogen and oxygen that fill the air around you, and some can even vibrate internally, like the bent-shape water vapour and the linear carbon dioxide. These additional ways of storing energy do not show up in the pressure formula, but they do show up in the total energy a gas carries at a given temperature. The next subtopic makes the correction, and in doing so introduces an integer that will quietly control a great deal of physics still to come.

DEGREES OF FREEDOM AND INTERNAL ENERGY

Every result so far in this chapter has been about translational motion. The kinetic equation counted the momentum transferred to a wall when a molecule bounced off it, and that bouncing is translational motion. The central result $\text{KE} = \frac{3}{2} kT$ is the average **translational** kinetic energy per molecule, not the total energy. For a **monatomic** gas like helium or argon, whose atoms are essentially spherical point-masses with nothing inside them to wobble or spin, this is the whole story. A helium atom at 300 K carries $\frac{3}{2} kT$ of energy and nothing more.

But most real gases are not monatomic. Nitrogen and oxygen, the two gases that make up 99% of the air you are breathing, are **diatomic**: each molecule is a pair of atoms held together by a chemical bond, something like a tiny dumbbell. A dumbbell can translate (the whole molecule moves through space), but it can also **rotate** end over end, and the energy stored in that rotation does not show up in the pressure formula. Carbon dioxide, water vapour, methane, and ozone are **polyatomic**: three or more atoms in a single molecule, with even more ways to move internally. This section extends the kinetic theory to account for those extra motions, introduces the concept of **degrees of freedom** as a bookkeeping tool, and arrives at a compact formula for the total internal energy of any ideal gas in terms of a single integer.

What a degree of freedom is

A **degree of freedom** of a molecule is an independent way in which the molecule can store kinetic energy. For a point-like atom moving through three-dimensional space, there are three such ways: motion along the x-axis, motion along the y-axis, and motion along the z-axis. Each of these is independent, in the sense that the energy associated with motion along one axis does not appear in any other. The total translational kinetic energy of the atom is a sum of three independent contributions, one per axis.

Translational motion gives three degrees of freedom. Rotational motion adds more, provided the molecule can actually rotate in a distinguishable way. A single spherical atom has no distinguishable axis of rotation (a perfect sphere looks the same when rotated about any axis through its centre), so a monatomic molecule has zero rotational degrees of freedom. A **diatomic molecule**, shaped like a dumbbell, has **two perpendicular axes about which rotation is distinguishable** (both perpendicular to the bond axis); rotation about the bond axis itself stores negligible energy because the atoms are essentially points on the bond, so only two rotational degrees of freedom contribute. A **polyatomic molecule** has **three rotational degrees of freedom**, because its non-linear arrangement of atoms distinguishes three perpendicular axes of rotation.

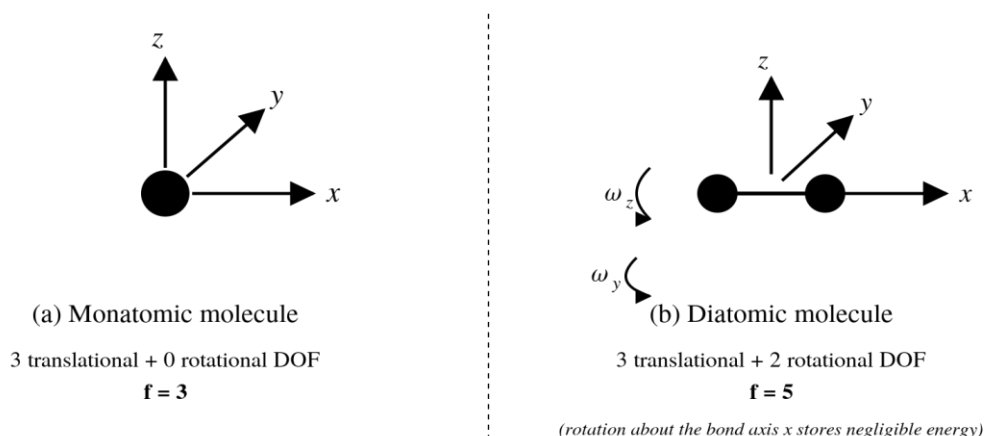


Figure: Degrees of freedom of a monatomic molecule (left) and a diatomic molecule (right). The monatomic molecule has only three translational degrees of freedom. The diatomic molecule has three translational and two rotational degrees of freedom (about axes perpendicular to the bond).

The equipartition theorem, stated informally

Having identified the degrees of freedom, the question is how much energy each one stores. The answer is a deep result of statistical mechanics, called the **equipartition theorem**, which we state here without proof:

In thermal equilibrium, each degree of freedom of a molecule stores, on average, $\frac{1}{2}kT$ of energy per molecule.

This statement is consistent with what we already know. The three translational degrees of freedom give an average translational kinetic energy of $3 \times \frac{1}{2}kT = \frac{3}{2}kT$ per molecule, exactly the result from the previous section. The equipartition theorem generalises this: whatever other degrees of freedom a molecule has, each one also contributes $\frac{1}{2}kT$. A full proof of the theorem requires the Boltzmann distribution of energies in a gas and the equipartition of energy among independent quadratic terms in the Hamiltonian, and lies beyond this chapter; the theorem is revisited in Chapter 5 in connection with heat capacities. For now, we take it as given and use it.

Total kinetic energy per molecule

For a molecule with a total number of degrees of freedom $f = f_{\text{trans}} + f_{\text{rot}}$, the total average kinetic energy per molecule is:

$$KE_{\text{tot}} = \frac{f}{2}kT$$

Filling in the three gas types:

- Monatomic gases (helium, argon, neon) have $f = 3 + 0 = 3$, so $KE_{\text{tot}} = \frac{3}{2}kT$.
- Diatomic gases (hydrogen, nitrogen, oxygen) have $f = 3 + 2 = 5$, so $KE_{\text{tot}} = \frac{5}{2}kT$.
- Polyatomic gases (carbon dioxide, water vapour, ozone) have $f = 3 + 3 = 6$, so $KE_{\text{tot}} = 3kT$.

Internal energy of a gas sample

The **internal energy** U of a gas sample is the total kinetic energy of all its molecules. For N molecules each carrying $\frac{f}{2}kT$:

$$U = N \times \frac{f}{2}kT = \frac{f}{2}NkT = \frac{f}{2}nRT$$

Three compact results for the three gas types follow immediately. For a monatomic gas ($f = 3$):

$$U = \frac{3}{2}nRT \quad (\text{monatomic})$$

For a diatomic gas ($f = 5$):

$$U = \frac{5}{2}nRT \quad (\text{diatomic})$$

For a polyatomic gas ($f = 6$):

$$U = 3nRT \quad (\text{polyatomic})$$

Three formulas, controlled by a single integer f . That integer is carrying enormous predictive power, as we will see at the end of this section and in Chapter 5.

Summary of internal energy for the three gas types

Gas type	Examples	Degrees of freedom f	Internal energy per mole
Monatomic	He, Ne, Ar	3 (3 trans + 0 rot)	$U = (3/2) RT$
Diatomic	H ₂ , N ₂ , O ₂	5 (3 trans + 2 rot)	$U = (5/2) RT$
Polyatomic	CO ₂ , H ₂ O, O ₃	6 (3 trans + 3 rot)	$U = 3 RT$

With these ideas now neatly in place, let us anchor them through a few worked examples before they drift away.

BINDER Example 16

Calculate the total internal energy of 3.0 mol of helium gas (monatomic) at a temperature of 27 °C.

Solution

$T = 27 + 273 = 300\text{K}$. Helium is monatomic, so $f = 3$ and the internal energy is:

$$U = \frac{3}{2} nRT = \frac{3}{2} \times (3.0\text{mol}) \times (8.314\text{Jmol}^{-1} \text{K}^{-1}) \times (300\text{K}) = 11224\text{J}$$

The total internal energy of 3.0 mol of helium at 300 K is approximately $1.12 \times 10^4 \text{ J}$ (or about 11.2 kJ).

Making Sense of the Answer: Eleven kilojoules is the energy that would be released if the three moles of helium gas were to be cooled all the way to absolute zero. It is comparable to the chemical energy in about 0.3 grams of sugar, which is a small but non-trivial quantity. Gas at room temperature is carrying real energy, distributed across an astronomical number of molecules in kinetic form.

Think Like a Physicist: The internal energy of a monatomic gas depends only on n , T , and R , with no dependence on the gas identity or molar mass. Helium, neon, argon, and xenon in equal numbers of moles at the same temperature all have exactly the same internal energy. This is the cleanest possible gas.

BINDER Example 17

Calculate the total internal energy of 2.0 mol of nitrogen gas (diatomic) at 27 °C. By what factor is this larger than the internal energy of 2.0 mol of helium gas at the same temperature?

Solution

Convert temperature: $T = 300\text{K}$. Nitrogen is diatomic, so $f = 5$ and:

$$U_{\text{N}_2} = \frac{5}{2} nRT = \frac{5}{2} \times (2.0\text{mol}) \times (8.314\text{Jmol}^{-1} \text{K}^{-1}) \times (300\text{K}) = 12471\text{J}$$

For 2.0 mol of helium at the same temperature:

$$U_{\text{He}} = \frac{3}{2} nRT = \frac{3}{2} \times (2.0\text{mol}) \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) \times (300\text{K}) = 7483\text{J}$$

Ratio of internal energies:

$$\frac{U_{\text{N}_2}}{U_{\text{He}}} = \frac{12471}{7483} = \frac{5}{3} \approx 1.67$$

The nitrogen has 5/3 times the internal energy of the helium at the same temperature, purely because each nitrogen molecule has two additional rotational degrees of freedom that helium atoms do not have. The ratio 5/3 is exactly the ratio of the degrees of freedom.

Making Sense of the Answer: Two gases at the same temperature and in the same amount have different internal energies if they have different molecular structures. A diatomic gas stores 67% more energy per mole than a monatomic gas, because its molecules have extra rotational motion that the monatomic atoms cannot participate in. This is why heating a given volume of nitrogen from 0 °C to 100 °C requires more heat than heating the same volume of helium over the same temperature range. The rotational motion is a storage reservoir that the heat has to fill along with the translational one.

Think Like a Physicist: The ratio $U_{\text{diatomic}}/U_{\text{monatomic}}$ at equal n and T equals the ratio of degrees of freedom, 5/3. The absolute values require a full substitution; the ratio requires no numerical work at all. When two gas samples differ only in molecular structure at matching n and T , reach for the ratio first.

REAL Example 18

Account for the following:

- It takes longer to heat a room full of humid air than dry air to the same temperature rise.
- A helium-filled balloon released indoors warms up faster than an air-filled balloon of the same size as the room temperature changes.
- Monatomic gases such as argon are used as inert fillers in insulating double-glazed windows rather than air, despite being more expensive.

Solution

- Dry air is mostly diatomic (nitrogen and oxygen, $f = 5$), while water vapour is polyatomic ($f = 6$) and stores more internal energy per mole per kelvin than the diatomic molecules. Humid air therefore requires more heat than dry air to undergo the same temperature rise, so heating a humid room takes longer.
- Helium is monatomic ($f = 3$) while air is effectively diatomic ($f = 5$), so helium needs less heat per mole to raise its temperature by a given amount. For the same heat flow through the balloon wall, the helium-filled balloon therefore reaches the room's temperature faster than the air-filled balloon.
- Argon atoms are monatomic ($f = 3$) and store less thermal energy per collision than the diatomic nitrogen and oxygen molecules of air ($f = 5$). In addition, argon atoms are heavier than the average air molecule, so at the same temperature they move more slowly and transport heat across the gap less rapidly. Both effects lower the thermal conductivity of argon below that of air, making argon a better insulator despite its greater cost.

Making Sense of the Answer: The three situations look superficially different, but they are all controlled by the same integer f . Humid air stores more energy per kelvin because water vapour has more degrees of freedom. Helium balloons warm up faster because helium has fewer degrees of freedom. Argon is a better insulator because its monatomic structure carries less heat per molecule. The physical question in each case is how many independent ways a molecule can absorb, store, or transport thermal energy, and f is the answer.

Think Like a Physicist: Whenever a real-world thermal question involves “this substance heats faster than that one” or “this gas insulates better than that one”, ask first which molecules have more degrees of freedom and which are heavier. The answer is usually a combination of the two, and reasoning about both factors separately is clearer than reaching for a table of heat capacities. Molecular structure, through f and m , sits underneath most everyday thermal phenomena even when the phenomena themselves look unrelated.

HOT Example 19

A rigid sealed container of volume 0.040 m^3 holds a mixture of 1.0 mol of helium gas and 1.0 mol of nitrogen gas, both at a temperature of 300 K . Each gas behaves ideally and neither reacts with the other. Calculate:

- the total internal energy of the mixture,
- the total pressure exerted on the walls of the container.

Solution

(a) The total internal energy is the sum of the internal energies of the two component gases, because the two gases do not interact. Helium is monatomic ($f = 3$); nitrogen is diatomic ($f = 5$):

$$U_{\text{tot}} = U_{\text{He}} + U_{\text{N}_2} = \frac{3}{2} n_{\text{He}} RT + \frac{5}{2} n_{\text{N}_2} RT$$

Substituting numerical values:

$$U_{\text{tot}} = \frac{1}{2} \times (3 \times 1.0 \text{ mol} + 5 \times 1.0 \text{ mol}) \times (8.314 \text{ J mol}^{-1} \text{ K}^{-1}) \times (300 \text{ K}) = 9977 \text{ J}$$

The total internal energy of the mixture is approximately 9977 J .

(b) For non-reacting helium-nitrogen gaseous mixture:

$$P_{\text{tot}} = \frac{(n_{\text{He}} + n_{\text{N}_2}) RT}{V}$$

$$P_{\text{tot}} = \frac{(2.0 \text{ mol}) \times (8.314 \text{ J mol}^{-1} \text{ K}^{-1}) \times (300 \text{ K})}{0.040 \text{ m}^3} = 1.25 \times 10^5 \text{ Pa}$$

The total pressure on the walls of the container is approximately $1.25 \times 10^5 \text{ Pa}$.

Making Sense of the Answer: The total pressure of the mixture is the same as it would be for 2.0 mol of a single ideal gas in the same container at the same temperature. Pressure depends on n , V , and T alone; it does not care whether the molecules are all the same kind or a mixture of different kinds. Internal energy, by contrast, does care: because helium and nitrogen have different numbers of degrees of freedom, they carry different amounts of internal energy per mole even at the same temperature. Pressure is blind to molecular structure; internal energy is not.

Think Like a Physicist: For any gas mixture at uniform temperature, each component obeys $PV = nRT$ independently, with its own partial pressure contributed to the total. This is Dalton's law of partial pressures, and it turns multi-component gas problems into separate single-component problems that are added at the end. The algebra never gets harder as more gases are added; it only gets longer.

Everything in this section has depended on a single integer, the number of degrees of freedom f . Monatomic gases have $f = 3$; diatomic gases have $f = 5$; polyatomic gases have $f = 6$. The internal energy of any ideal gas is then $(f/2)nRT$, and that is the end of the story at this level. But the same integer f will reappear in Chapter 5 in the formulas for the heat capacity at constant volume and the heat capacity at constant pressure of an ideal gas. The humble integer f , counting the ways a molecule can store kinetic energy, turns out to control also how much heat a gas absorbs when warmed, and how fast sound travels through it, and how adiabatic compression changes its temperature. Carrying f correctly through this section is not just a bookkeeping habit; it is preparation for a chapter's worth of physics that is yet to come.

With the ideal gas story now fully told, we find ourselves at a natural moment to ask a harder question. We have spent five sections assuming that molecules have no volume of their own, that they do not attract each other, and that all collisions are perfectly elastic. These assumptions produced an equation of state so clean and predictive that it is hard to remember they were assumptions at all. But every real gas departs from this idealisation under the right conditions, and the departures are neither small nor exotic. The next subtopic asks where the ideal gas model breaks down, why it breaks down, and what the real physics looks like beyond the approximation.

REAL GASES AND THE LIMITS OF THE IDEAL GAS MODEL

Everything we have done so far has rested on the six assumptions of the kinetic theory. Those assumptions gave us a theory clean enough to derive, predictive enough to use, and simple enough to remember. But no real gas obeys them perfectly. A real gas is made of molecules that do occupy a small volume of their own, that do attract each other weakly even when they are not colliding, and that do, at sufficiently low temperatures, stop moving in straight lines and instead form a liquid. The ideal gas equation is a beautiful, but it is an approximation, and this section looks honestly at where the approximation breaks down and why.

Which assumptions fail first

When a real gas is pushed away from ideal behaviour, it is not all six assumptions that fail at once. Two of them fail first, and the other four remain adequate until very extreme conditions are reached.

The first to fail is **assumption 2**, that the volume of the molecules themselves is negligible compared with the volume of the container. At ordinary pressures this is excellent: one mole of gas at atmospheric pressure and room temperature occupies about 25 litres, while the actual molecules inside it occupy only a few cubic centimetres. But at **high pressure** *the container shrinks around the gas while the molecules themselves do not shrink, and the fraction of the container occupied by the molecules becomes significant*. A gas compressed to one hundred atmospheres has molecules occupying perhaps a few percent of the container volume, and at one thousand atmospheres the fraction is larger still. The molecules begin to get in each other's way, and the equation $PV = nRT$ no longer holds to the precision the assumption claimed.

The second to fail is **assumption 4**, that intermolecular forces are negligible except during the instant of a collision. This assumption fails at **low temperature**. At room temperature, molecules are moving fast enough that the weak attractive forces between them (van der Waals forces, named after the physicist who first wrote them into an equation of state) are too feeble to deflect a molecule noticeably from its straight-line path. But at **low temperatures** *the molecules are moving slowly, and the attractive forces have time to act*. Molecules begin to stick to each other briefly, to cluster, and at sufficiently low temperatures to condense into a liquid altogether. The gas stops behaving as a collection of independent projectiles and starts behaving as a collection of mutually attracting particles.

The other four assumptions (large number of molecules, elastic wall collisions, straight-line motion between collisions, negligible collision time) remain adequate well into the regime where (2) and (4) have failed, and they are not usually what limits the ideal gas model in practice.

Conditions for near-ideal behaviour

The two failure modes just described point directly at the conditions under which a real gas behaves most like an ideal gas. The ideal gas approximation is best at **low pressure**, because *low pressure means a large container volume for a given amount of gas, so the molecules occupy a negligible fraction of that volume*. It is best at **high temperature**, because *at high temperature the molecules move fast and the weak intermolecular attractions have no time to act on them*. It is best at **low density**, because *low density combines both advantages: molecules are far apart (so their own volume is negligible) and typically moving fast (so attractions do not matter)*.

These three conditions are equivalent statements of the same physical idea, because low pressure at a given temperature implies low density, and high temperature at a given density implies faster-moving molecules less affected by attractions. Students often remember the compact phrase: *real gases behave most ideally at low pressure and high temperature*. That is enough for the examination.

The van der Waals equation

The Dutch physicist Johannes van der Waals proposed in 1873 a modified equation of state that corrects for the two effects identified above. Schematically, his equation takes the form:

$$(P + \text{correction for attractions}) \times (V - \text{correction for molecular volume}) = nRT$$

The first correction raises the effective pressure to account for attractive forces between molecules; the second reduces the available volume to account for the space the molecules themselves take up. At low pressure and high temperature, both corrections become small and the equation reduces back to $PV = nRT$. The detailed form of the corrections, and their derivation, lie beyond the A-Level syllabus and are not required for the examination. What matters here is the conceptual point: the van der Waals equation exists,

it works better than $PV = nRT$ for real gases away from the ideal regime, and it does so by restoring the two assumptions the ideal gas model simplified away.

An ideal gas cannot be liquefied

One further consequence of the failure of assumption (4) deserves its own paragraph. A real gas can be liquefied by cooling it sufficiently and, if necessary, compressing it; every industrial gas sold in Tanzania has a liquid phase, produced in this way for **storage** and **transport**. An ideal gas, by contrast, can never be liquefied. This follows directly from assumption (4): if intermolecular forces are truly zero, as the ideal gas model requires, then there is nothing to hold the molecules together into a liquid. No matter how cold or how compressed the ideal gas becomes, its molecules continue to bounce off each other elastically and move in straight lines between collisions. They never stick. The ability of real gases to condense into liquids is itself the clearest evidence that intermolecular forces are not zero, and therefore that no real gas is strictly ideal.

With the honest limits of the ideal gas model now acknowledged, we can return to the many situations where the model is perfectly adequate and see what it can do. The kinetic theory of gases is not just an elegant derivation; it is also a working tool used in weather forecasting, in industrial gas handling, in the design of insulating windows, and in the physiology of breathing. The final subtopic of this chapter gathers these applications together and shows kinetic theory at work in the world you live in.

APPLICATIONS OF THE KINETIC THEORY OF GASES

The kinetic theory of gases is not confined to physics examinations and laboratory demonstrations. It is a working tool, used every day by meteorologists explaining the weather, by doctors treating patients with respiratory problems, by engineers designing gas cylinders and window insulation, and by astronomers working out the chemistry of stellar atmospheres. Every phenomenon in which gas molecules do something visible or measurable can be traced back, eventually, to the handful of ideas developed in this chapter: molecules move rapidly and randomly, their speeds depend on temperature and mass, and the pressure they exert is the collective drumming of their collisions. This section takes six applications, each drawn from a different corner of life, and shows the kinetic theory at work in the world we live in.

1. Why atmospheric pressure falls with altitude

The atmosphere is held to the Earth by gravity. Each layer of air must support the weight of all the air above it, and since air at sea level has the greatest amount of atmosphere above it, sea-level pressure is the highest. As you climb a mountain, the amount of atmosphere above you decreases, and so does the pressure. At the top of Mount Kilimanjaro (5895m) the atmospheric pressure is about half its sea-level value; at the cruising altitude of a passenger jet (about 11000m) it is about one-quarter; at the altitude of the International Space Station there is essentially no atmosphere at all.

The molecular picture makes this quantitative. At any altitude, the local pressure equals the weight per unit area of the air column above. This weight is set by the total number of molecules in the column, which falls roughly exponentially with height because *faster-moving molecules can reach greater altitudes against gravity while slower ones cannot*. The result is that the number density of air molecules, N/V , falls with altitude in parallel with the pressure. A climber at the summit of Kilimanjaro draws about 54% as many air molecules into each breath as a climber at sea level. This is the number-density consequence of a pressure change, and it is why mountaineers at altitude struggle to breathe even though the chemical composition of the air (approximately 78% nitrogen and 21% oxygen) is identical at the summit and at sea level.

2. Gas cylinders and their safety

Domestic liquid petroleum gas (LPG) cylinders, widely used for cooking in Tanzanian homes, typically hold gas at pressures of several hundred kilopascals at room temperature. Industrial gas cylinders (oxygen, acetylene, nitrogen, argon) commonly hold gas at pressures of 10 to 20 megapascals, approximately 100 to 200 atmospheres. The walls of these cylinders must withstand the pressure for years without failure; the engineering rests on the physics developed in this chapter.

A sealed cylinder is a rigid container. Its volume does not change, and the amount of gas inside is fixed. From the ideal gas equation with V and n constant, the pressure is proportional to the absolute temperature: if the temperature rises, so does the pressure, in direct proportion on the Kelvin scale. A cylinder left in direct tropical sunlight, as in the HOT example of an earlier section, can see its temperature rise from 20°C (293K) to around 70°C (343K), a 17% increase on the Kelvin scale. Its internal pressure rises by the same 17% fraction. A cylinder rated to hold gas safely at 12MPa in the morning can therefore see its internal pressure rise to over 14MPa by midday, pushing it close to or beyond its pressure rating.

This is the physical reason Tanzanian safety regulations require LPG and industrial gas cylinders to be stored in shaded, well-ventilated areas. The prohibition is not arbitrary bureaucracy; it is the direct application of the Kelvin-temperature scaling of gas pressure inside a fixed volume. Molecules inside a heated cylinder are moving faster and hitting the walls harder, and if the walls cannot bear the increased battering, they can rupture. A ruptured high-pressure cylinder releases its stored compressed gas explosively and has been responsible for industrial accidents across the world. The kinetic theory explains both why such accidents happen and how to prevent them.

3. Argon in double-glazed windows

Modern energy-efficient windows use two panes of glass separated by a thin sealed gap, typically a centimetre or two wide. The gap reduces heat loss through the window because still air is a poor thermal conductor. Filling the gap with argon instead of air reduces the heat loss further, and argon-filled windows are now standard in high-efficiency construction despite being more expensive than air-filled ones.

Two pieces of kinetic-theory physics explain why argon insulates better than air. Argon is a monatomic gas with only three degrees of freedom per atom, compared with five for the diatomic molecules nitrogen and oxygen that dominate air. Argon atoms therefore carry less thermal energy per collision than air molecules, which reduces the rate at which heat is transported across the gap. Additionally, argon atoms are heavier than the average air molecule (40g/mol versus about 29g/mol for air), so at the same temperature they move more slowly (v_{rms} proportional to $1/\sqrt{M}$), further slowing the transport of heat. Both effects act in the same direction, making argon a better thermal insulator than air by about 30%. The extra cost of argon is recovered over the lifetime of the window through reduced heating and cooling bills.

4. Diffusion of gases: Graham's law

When two gases are brought into contact, they spontaneously intermix until the mixture is uniform. This process, called diffusion, is driven entirely by molecular motion: molecules of each gas wander randomly through the other, and after enough time the concentrations become uniform. The rate at which a gas diffuses depends on how fast its molecules are moving, which in turn depends on their molar mass. Since $v_{\text{rms}} = \sqrt{3RT/M}$, lighter gases diffuse faster than heavier ones at the same temperature, and the ratio of diffusion rates is inversely proportional to the square root of the ratio of molar masses:

$$\frac{\text{rate}_1}{\text{rate}_2} = \sqrt{\frac{M_2}{M_1}}$$

This is Graham's law of diffusion, first stated empirically by the Scottish chemist Thomas Graham in 1848, long before the kinetic theory gave it a molecular explanation. It tells us, for example, that hydrogen (molar mass 2) diffuses through a porous membrane four times faster than oxygen (molar mass 32), because $\sqrt{32/2} = \sqrt{16} = 4$. Industrial applications include the enrichment of uranium for nuclear reactors, which relies on the very small difference in diffusion rates between the two uranium isotopes U-235 and U-238, separated through thousands of stages of gaseous diffusion. A small molecular-mass difference, amplified through many stages, produces the enriched product.

5. The speed of sound in gases

Sound is a pressure wave in a gas, propagated by successive collisions between molecules. A disturbance at one point compresses the local gas slightly, which pushes neighbouring molecules, which in turn push their neighbours, and so on, carrying the compression outward at a finite speed. It is plausible (and correct) that the speed of sound in a gas is close to the average speed of the molecules themselves, since the molecules are the carriers of the disturbance. A full derivation, accounting for adiabatic compression effects not yet treated in this book, gives:

$$v_{\text{sound}} = \sqrt{\frac{\gamma RT}{M}}$$

where γ is a number close to 1.4 for diatomic gases like air. Comparing this with the rms speed of the molecules, $v_{\text{rms}} = \sqrt{3RT/M}$, we see that the speed of sound is only slightly smaller than v_{rms} , differing by the factor $\sqrt{\gamma/3} \approx 0.68$. This agrees with experiment: in air at room temperature, the speed of sound is about 343m/s, while v_{rms} for air molecules is about 500m/s.

Two consequences follow.

- First, sound travels faster in lighter gases: the speed of sound in helium is nearly three times that in air, because helium molecules are lighter ($M = 4$ instead of 29) and therefore moving faster at the same temperature. This is why inhaling helium from a balloon makes your voice temporarily higher-pitched: the vocal cords vibrate at the same rate, but the sound waves in the helium-filled vocal tract travel faster, so the resonances shift to higher frequencies.
- Second, sound travels faster in warmer gases: the speed of sound in air rises by about 0.6m/s for each kelvin of temperature increase. This is why distant sounds carry further on cool evenings than on hot afternoons: the cooler air refracts sound downward, bending it along the ground rather than letting it escape upward.

6. Breathing and gas exchange

Breathing moves a mixture of gases (nitrogen, oxygen, carbon dioxide, water vapour, and trace gases such as argon) into and out of the lungs. Each of these gases, while mixed with the others, exerts its own partial pressure, contributing independently to the total pressure of the mixture. This is Dalton's law of partial pressures, met in the mixture example of the previous section. In the lungs, oxygen diffuses from the air into the blood, and carbon dioxide diffuses from the blood into the air, each following its own partial-pressure gradient. At sea level, the partial pressure of oxygen in inhaled air is about 21kPa; in the blood returning from the body, it is about 5kPa; the 16kPa difference is what drives oxygen into the blood.

At high altitude, total atmospheric pressure is lower, so the partial pressure of oxygen is also lower even though the percentage of oxygen in air has not changed. A climber at the summit of Kilimanjaro inhales air whose oxygen partial pressure is only about 10kPa. The partial-pressure gradient between lung air and returning blood is roughly halved, and the rate at which oxygen enters the blood is reduced correspondingly. This is the molecular reason for altitude sickness, and it is why climbers at extreme altitudes (above 7000m) typically use supplemental oxygen, which restores the partial-pressure gradient by raising the oxygen concentration in the inhaled gas.

The same physics operates in reverse for scuba divers. At 30m depth below the sea surface, the pressure on a diver is four atmospheres (one from the atmosphere plus three from the water column). Every gas the diver breathes has its partial pressure multiplied by four. Nitrogen, which is inert at sea-level partial pressures, becomes physiologically active at high partial pressure and can cause "nitrogen narcosis," a condition resembling drunkenness. Deep technical divers breathe gas mixtures (such as helium-oxygen) specifically chosen to reduce the partial pressures of problem gases, designing the breathing gas on the basis of Dalton's law.

With these six applications now in place, let us enjoy the ideas one last time through a pair of carefully chosen worked examples.

REAL Example 20

Explain for the following:

- A climber at the summit of Mount Kilimanjaro struggles to breathe even though the chemical composition of the air is the same as at sea level.
- Inhaling helium from a balloon temporarily raises the pitch of a person's voice.
- Hydrogen gas diffuses through a porous membrane four times faster than oxygen gas at the same temperature.

Solution

- At the summit of Kilimanjaro, the total atmospheric pressure is about half its sea-level value, so the partial pressure of oxygen in the inhaled air is also about half. The partial-pressure gradient that drives oxygen from the lungs into the blood is reduced, lowering the rate at which oxygen enters the bloodstream even though the oxygen fraction in the air (about 21%) has not changed.
- The speed of sound in a gas is proportional to $1/\sqrt{M}$, so it is higher in helium ($M = 4\text{g/mol}$) than in air ($M \approx 29\text{g/mol}$). When the vocal tract is filled with helium, sound waves travel faster through it, and the resonant frequencies of the tract shift upward. The vocal cords vibrate at their normal rate, but the filtering action of the tract emphasises higher frequencies, making the voice sound higher-pitched.
- Graham's law states that the rate of diffusion of a gas is proportional to $1/\sqrt{M}$. The ratio of diffusion rates for hydrogen ($M = 2$) to oxygen ($M = 32$) is $\sqrt{(32/2)} = \sqrt{16} = 4$. Hydrogen molecules are sixteen times lighter than oxygen molecules, so at the same temperature they move four times faster on average, and they diffuse four times faster through the same membrane.

Making Sense of the Answer: *three apparently different phenomena, one at altitude, one in a party balloon, and one in an industrial separation plant, all turn out to be applications of the same family of ideas:*

molecular speeds depend on temperature and mass, and physical processes that depend on molecular speed scale in predictable ways. This is the hallmark of kinetic theory: one set of principles, applied across a wide range of scales from the lung to the atmosphere to the uranium enrichment plant.

Think Like a Physicist: whenever an application question involves how fast a gas does something (diffuses, mixes, escapes, conducts sound), the first question to ask is how its molecular speed depends on the variables given. If the temperature is fixed, the answer almost always involves $1/\sqrt{M}$. If the gas is specified but the temperature varies, the answer involves $\sqrt{T}$. Recognising which scaling applies, and recognising it quickly, turns most application questions into a single ratio calculation.

HOT Example 21

An industrial argon gas cylinder of internal volume 0.050m^3 is filled with argon (molar mass 40g/mol) at 20°C to a pressure of 15MPa . The cylinder is rated to withstand pressures up to 18MPa . The cylinder is accidentally left standing in direct sunlight, where its temperature rises to 55°C .

- Calculate the new pressure inside the cylinder.
- State whether the cylinder is still within its safe pressure rating. Explain.
- Calculate the root mean square speed of the argon atoms at the higher temperature.

Solution

$$T_1 = 20 + 273 = 293\text{K}; T_2 = 55 + 273 = 328\text{K}.$$

- (a) The cylinder is rigid (V constant) and sealed (n constant), so from $PV = nRT$ the pressure is proportional to temperature on the Kelvin scale:

$$P_2 = P_1 \times \frac{T_2}{T_1} = 15\text{MPa} \times \frac{328}{293} = 16.8\text{MPa}$$

The new pressure inside the cylinder is approximately 16.8MPa .

- (b) The cylinder is rated to 18MPa . The new pressure of 16.8MPa is below the rating, so the cylinder is still within its safe operating range, though with a reduced safety margin of only about 1.2MPa . Continued heating of the cylinder could push the pressure past the rating and into the failure regime, so the cylinder should be moved to a shaded area promptly as a precaution.

- (c) The root mean square speed of argon atoms at 328K uses $v_{\text{rms}} = \sqrt{\frac{3RT}{M}}$:

$$v_{\text{rms}} = \sqrt{\frac{3 \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) \times (328\text{K})}{40 \times 10^{-3}\text{kg/mol}}} = 452\text{m/s}$$

The argon atoms have a root mean square speed of approximately 452m/s at 55°C .

Making Sense of the Answer: the pressure rise of about 12% for a temperature rise of 35K is consistent with the direct Kelvin proportionality of the ideal gas equation at constant volume. The cylinder survives on this particular afternoon, but only with a small safety margin, and a cylinder rated closer to its initial operating pressure could easily be pushed past its limit by the same amount of solar heating. Cylinder safety ratings include temperature allowances for exactly this reason, and the recommended storage temperature is well below the conditions reachable in direct tropical sunlight.

Think Like a Physicist: sealed-cylinder problems at different temperatures are Kelvin-scale ratio problems. Convert both temperatures to kelvin, take the ratio, and apply it to the pressure. The molar mass, molecular speed, and other molecular properties do not enter the pressure calculation at all (they would only enter if the question asked about those properties); they are needed only when the question asks about speeds or kinetic energies separately. Identifying which quantities are needed for which part of the question saves time in a multi-part examination question.

This chapter has taken the kinetic theory of gases from six assumptions to pressure, to temperature, to internal energy, and finally to six applications drawn from the world around you. Molecular motion is the common thread; the equations that follow from assuming many rapidly-moving particles bouncing elastically off each other turn out to describe the air in your lungs, the gas in your cooking cylinder, the insulation in your windows, the sound of your voice, and the weather above your head. The miscellaneous worked examples that follow bring all of these ideas together at once.

MISCELLANEOUS WORKED EXAMPLES ON THE KINETIC THEORY OF GASES**Example 22**

- (a) A gas in a sealed container exerts a steady pressure on the walls, even though the total momentum of all the molecules inside is zero. Explain how this is possible.
- (b) A container of volume $5.0 \times 10^{-3} \text{ m}^3$ holds 3.0×10^{24} molecules of a gas. Each molecule has a mass of $3.3 \times 10^{-26} \text{ kg}$, and the mean square speed of the molecules is $1.6 \times 10^5 \text{ m}^2 \text{ s}^{-2}$. Calculate the pressure exerted by the gas on the walls of the container.

Solution

- (a) Pressure arises from the change of momentum of individual molecules when they collide with the walls, not from any net momentum carried by the gas as a whole. A molecule approaching a wall reverses its component of momentum perpendicular to that wall, and the wall experiences an equal and opposite impulse. Because molecules are moving in all directions, the total vector momentum of the gas remains zero, but the rate of momentum change at each wall is non-zero and equal in magnitude on all walls. Pressure is the time-averaged force per unit area from this continuous rain of momentum reversals, so it does not require any net drift of the gas.
- (b) Applying the kinetic equation $P = \frac{Nm}{3V} \bar{c}^2$:

$$P = \frac{(3.0 \times 10^{24}) \times (3.3 \times 10^{-26} \text{ kg}) \times (1.6 \times 10^5 \text{ m}^2 \text{ s}^{-2})}{3 \times (5.0 \times 10^{-3} \text{ m}^3)} = 1.06 \times 10^6 \text{ Pa}$$

The pressure on the walls of the container is approximately $1.06 \times 10^6 \text{ Pa}$.

Example 23

- (a) Real gases obey the ideal gas equation very closely under some conditions but deviate from it under others. Explain the conditions under which a real gas departs most noticeably from ideal behaviour and identify which assumption of the kinetic theory fails first under those conditions.
- (b) The molecules in a sample of gas are observed, and their speeds are recorded as follows:
3 molecules at 2.2m/s, 7 molecules at 6.1m/s, 5 molecules at 7.8m/s, 2 molecules at 4.0m/s, and 1 molecule at 2.5m/s. Calculate the root mean square speed of the sample.

Solution

- (a) A real gas deviates most strongly from the ideal gas equation at high pressure and at low temperature.
- At high pressure, the volume of the molecules themselves is no longer negligible compared with the volume of the container, and the assumption of negligible molecular volume fails first.
 - At low temperature, the weak attractive forces between molecules have time to act on the slowly-moving molecules, and the assumption of negligible intermolecular forces fails.
- (b) The root mean square speed is obtained by taking the average of the squares of the speeds and then taking the square root. The mean of the squares is:

$$\bar{c}^2 = \frac{3 \times (2.2 \text{ m/s})^2 + 7 \times (6.1 \text{ m/s})^2 + 5 \times (7.8 \text{ m/s})^2 + 2 \times (4.0 \text{ m/s})^2 + 1 \times (2.5 \text{ m/s})^2}{3 + 7 + 5 + 2 + 1}$$

$$\bar{c}^2 = \frac{(14.52 + 260.47 + 304.20 + 32.00 + 6.25) \text{ m}^2 \text{ s}^{-2}}{18} = 34.30 \text{ m}^2 \text{ s}^{-2}$$

$$v_{\text{rms}} = \sqrt{\bar{c}^2} = \sqrt{34.30 \text{ m}^2 \text{ s}^{-2}} = 5.86 \text{ m/s}$$

The root mean square speed of the molecules in the sample is approximately 5.86m/s.

Example 24

- (a) Two sealed containers hold different gases at the same temperature. It is observed that although the average kinetic energy per molecule is the same in both containers, the average molecular speeds are different. Explain this observation.
- (b) 2.0 moles of oxygen gas are contained at a temperature of 300K. Calculate the total translational kinetic energy of the gas sample, and comment on the magnitude of your answer by comparing it with the kinetic energy of a cricket ball of mass 0.16kg bowled at 36m/s.

Solution

- (a) The average kinetic energy per molecule of an ideal gas depends only on the absolute temperature, according to the relation $KE = (3/2)kT$. Since the two gases are at the same temperature, they have the same average kinetic energy per molecule. However, the molecules of the two gases have different masses, so for the same kinetic energy $\frac{1}{2}m\bar{c}^2$ they must have different mean square speeds. The lighter gas has the higher rms speed and the heavier gas the lower rms speed, such that the product $m\bar{c}^2$ is the same for both.
- (b) Total translational KE of the gas:

$$KE_{\text{total}} = \frac{3}{2} nRT = \frac{3}{2} \times (2.0\text{mol}) \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) \times (300\text{K}) = 7483\text{J}$$

Kinetic energy of the cricket ball:

$$KE_{\text{ball}} = \frac{1}{2} mv^2 = \frac{1}{2} \times (0.16\text{kg}) \times (36\text{m/s})^2 = 104\text{J}$$

The total translational KE of 2.0mol of oxygen at 300K is approximately 7483J, which is about seventy-two times the kinetic energy of the cricket ball. This shows that a modest quantity of gas at room temperature carries substantial molecular kinetic energy, distributed across an astronomical number of molecules.

Example 25

- (a) Explain what is meant by a degree of freedom of a gas molecule, and justify why a nitrogen molecule has five accessible degrees of freedom at room temperature while a helium atom has only three.
- (b) 2.0 moles of helium gas and 2.0 moles of nitrogen gas are held in separate sealed rigid containers, both initially at 300K. Each sample is heated at constant volume and its temperature rises by 40K. Calculate the heat supplied to each sample, and compare the two values.

Solution

- (a) A degree of freedom is an independent way in which a molecule can store kinetic energy. A helium atom, being essentially a point mass with no distinguishable axis of rotation, can store kinetic energy only in its translational motion along three perpendicular directions, giving three degrees of freedom. A nitrogen molecule, being a dumbbell shape, has the same three translational degrees of freedom but can also rotate about two axes perpendicular to its bond, adding two further degrees of freedom. Rotation about the bond axis itself stores negligible energy because the atoms lie essentially on the axis and the moment of inertia about it is very small, so it is not counted. Nitrogen therefore has $f = 5$ accessible degrees of freedom at room temperature.
- (b) At constant volume the heat supplied equals the increase in internal energy of the gas:

$$Q = \Delta U = \frac{f}{2} \times nR\Delta T$$

For helium ($f = 3$):

$$Q_{\text{He}} = \frac{3}{2} \times (2.0\text{mol}) \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) \times (40\text{K}) = 998\text{J}$$

For nitrogen ($f = 5$):

$$Q_{\text{N}_2} = \frac{5}{2} \times (2.0\text{mol}) \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) \times (40\text{K}) = 1663\text{J}$$

Nitrogen requires approximately 1663J and helium approximately 998J to produce the same 40K temperature rise. Nitrogen needs about 1.67 times more heat because its molecules have two additional rotational degrees of freedom that must also be filled with energy as the gas warms; the ratio is exactly 5/3, the ratio of the degrees of freedom.

Example 26

- (a) A student measures the temperature of a cold room as 10°C and the temperature of a warm room as 30°C . The student claims that the warm room is therefore three times as hot as the cold room. Explain why this claim is incorrect and state the correct ratio.
- (b) Estimate the total number of air molecules in a classroom of dimensions $12\text{m} \times 8\text{m} \times 3.5\text{m}$ at a temperature of 25°C and a pressure of $1.0 \times 10^5 \text{Pa}$. Use $k = 1.381 \times 10^{-23} \text{J K}^{-1}$.

Solution

- (a) The Celsius scale places its zero at the freezing point of water, which is not the absolute zero of molecular motion. Ratios of temperatures measured on the Celsius scale are therefore not ratios of any physically meaningful quantity. The kinetic energy per molecule is proportional to the Kelvin temperature, not to the Celsius temperature. Converting: $10^{\circ}\text{C} = 283\text{K}$ and $30^{\circ}\text{C} = 303\text{K}$. The correct ratio of hotness (measured by average molecular kinetic energy) is $303/283 \approx 1.07$, so the warm room is only about 7% hotter than the cold room, not three times as hot.
- (b) Volume of the classroom:

$$V = 12\text{m} \times 8\text{m} \times 3.5\text{m} = 336\text{m}^3$$

Temperature in kelvin: $T = 25 + 273 = 298\text{K}$. Applying $PV = NkT$ to find N :

$$N = \frac{PV}{kT} = \frac{(1.0 \times 10^5 \text{Pa}) \times (336\text{m}^3)}{(1.381 \times 10^{-23} \text{J K}^{-1}) \times (298\text{K})} = 8.2 \times 10^{27}$$

The classroom contains approximately 8.2×10^{27} air molecules.

Example 27

- (a) Explain the molecular origin of the latent heat of fusion of a solid, using ideas from the kinetic theory of gases.
- (b) A sealed cylinder contains a fixed mass of oxygen gas initially at a temperature of 20°C and a pressure of $2.0 \times 10^5 \text{Pa}$. The gas is compressed to half its original volume and simultaneously heated to a new temperature of 100°C . Calculate the new pressure of the gas, the new root mean square speed of the oxygen molecules, and the new average translational kinetic energy per molecule. (Take the molar mass of oxygen to be 32g/mol , $k = 1.381 \times 10^{-23} \text{J K}^{-1}$.)

Solution

- (a) In a solid, molecules are held in fixed positions by strong attractive intermolecular forces and vibrate about those fixed positions with a small kinetic energy. When heat is supplied at the melting point, that heat does not raise the kinetic energy of the molecules (the temperature does not change) but instead does work against the intermolecular forces, increasing the potential energy of the molecular arrangement. When enough heat has been supplied to weaken the forces that hold the solid together, the molecules become free to move past each other and the solid has melted into a liquid. The latent heat of fusion is the energy required per unit mass to perform this separation without any accompanying rise in temperature.
- (b) Convert temperatures: $T_1 = 20 + 273 = 293\text{K}$; $T_2 = 100 + 273 = 373\text{K}$.

Applying the combined gas law $\frac{P_1 V_1}{T_1} = \frac{P_2 V_2}{T_2}$ with $V_2 = \frac{1}{2} V_1$:

$$P_2 = P_1 \times \frac{V_1}{V_2} \times \frac{T_2}{T_1} = (2.0 \times 10^5 \text{Pa}) \times 2 \times \frac{373\text{K}}{293\text{K}} = 5.09 \times 10^5 \text{Pa}$$

The new pressure of the gas is $5.09 \times 10^5 \text{Pa}$.

New rms speed from $v_{\text{rms}} = \sqrt{\frac{3RT}{M}}$, at $T_2 = 373\text{K}$:

$$v_{\text{rms}} = \sqrt{\frac{3 \times (8.314 \text{J mol}^{-1} \text{K}^{-1}) \times (373\text{K})}{32 \times 10^{-3} \text{kg/mol}}} = 539\text{m/s}$$

The root mean square speed of the oxygen molecules at 373K is 539m/s .

Average translational KE per molecule at $T_2 = 373\text{K}$:

$$\text{KE} = \frac{3}{2} kT = \frac{3}{2} \times (1.381 \times 10^{-23} \text{J K}^{-1}) \times (373\text{K}) = 7.73 \times 10^{-21} \text{J}$$

The average translational kinetic energy per oxygen molecule at the new temperature is $7.73 \times 10^{-21} \text{J}$.

Example 28

- (a) Two non-reacting ideal gases are mixed in a single container. Account for the fact that the total pressure exerted by the mixture on the walls of the container equals the sum of the pressures that each gas would exert alone in the same container at the same temperature.
- (b) A sealed rigid container of volume 10.0 litres contains a mixture of 0.40 mol of helium gas and 0.60 mol of nitrogen gas at a temperature of 27°C. Calculate the total pressure of the mixture in the container and the total internal energy of the mixture. (Take helium to be monatomic and nitrogen to be diatomic.)

Solution

- (a) In an ideal gas, intermolecular forces are negligible except during collisions, so the molecules of each species move independently of the molecules of the other species. Each species therefore collides with the walls of the container at a rate determined only by its own number density, molecular mass, and the temperature, unaffected by the presence of the other species. The pressure exerted by each species is the pressure it would exert if it alone occupied the container at the same temperature, which is called its partial pressure. The total force per unit area on the wall is the sum of the forces contributed by each species individually, so the total pressure equals the sum of the partial pressures. This statement is called Dalton's law of partial pressures.
- (b) Convert: $V = 10.0 \text{ litres} = 10.0 \times 10^{-3} \text{ m}^3 = 1.0 \times 10^{-2} \text{ m}^3$; $T = 27 + 273 = 300 \text{ K}$.
 Total moles: $n = n_{\text{He}} + n_{\text{N}_2} = 0.40 + 0.60 = 1.0 \text{ mol}$.
 Total pressure from $PV = nRT$:

$$P_{\text{tot}} = \frac{nRT}{V} = \frac{(1.0 \text{ mol}) \times (8.314 \text{ J mol}^{-1} \text{ K}^{-1}) \times (300 \text{ K})}{1.0 \times 10^{-2} \text{ m}^3} = 2.49 \times 10^5 \text{ Pa}$$

The total pressure of the gas mixture is $2.49 \times 10^5 \text{ Pa}$.

Total internal energy = sum of internal energies of each component.

Helium has $f = 3$ and nitrogen has $f = 5$:

$$U_{\text{tot}} = \frac{3}{2} n_{\text{He}} RT + \frac{5}{2} n_{\text{N}_2} RT$$

$$U_{\text{tot}} = \frac{RT}{2} (3n_{\text{He}} + 5n_{\text{N}_2})$$

$$U_{\text{tot}} = \frac{(8.314 \text{ J mol}^{-1} \text{ K}^{-1}) \times (300 \text{ K})}{2} \times (3 \times 0.40 \text{ mol} + 5 \times 0.60 \text{ mol}) = 5236 \text{ J}$$

The total internal energy of the gas mixture is 5236 J.

Example 29

- (a) The kinetic theory of gases uses the root mean square speed (v_{rms}) of the molecules rather than their simple arithmetic mean speed. Explain the physical reason for this choice.
- (b) A sample of gas contains 10 molecules with the following speeds: three molecules at 300 m/s, four molecules at 500 m/s, two molecules at 700 m/s, and one molecule at 1000 m/s. Calculate the simple arithmetic mean speed of the molecules, and separately the root mean square speed. Comment on why the two values differ.

Solution

- (a) The pressure of a gas on the walls of its container depends on the rate at which molecules transfer momentum to the walls, and the momentum transfer per collision is proportional to the molecular speed. However, the kinetic energy transferred to or from each molecule is proportional to the square of its speed. The quantities that determine pressure and temperature are therefore controlled by the average of the squares of the molecular speeds, not by the average of the speeds themselves. The root mean square speed, being the square root of the mean square, is the speed that enters the pressure and temperature equations correctly and has the right dimensions. The simple mean speed, while conceptually simpler, does not appear in the fundamental equations of the kinetic theory.
- (b) Simple arithmetic mean:

$$v_{\text{mean}} = \frac{3 \times (300 \text{ m/s}) + 4 \times (500 \text{ m/s}) + 2 \times (700 \text{ m/s}) + 1 \times (1000 \text{ m/s})}{10} = 530 \text{ m/s}$$

Root mean square speed:

$$\bar{c}^2 = \frac{3 \times (300\text{m/s})^2 + 4 \times (500\text{m/s})^2 + 2 \times (700\text{m/s})^2 + 1 \times (1000\text{m/s})^2}{10} = 3.25 \times 10^5 \text{ m}^2\text{s}^{-2}$$

$$v_{\text{rms}} = \sqrt{\bar{c}^2} = \sqrt{3.25 \times 10^5 \text{ m}^2\text{s}^{-2}} = 570\text{m/s}$$

The simple arithmetic mean is 530m/s, while the root mean square speed is 570m/s. The rms is larger because squaring gives extra weight to the faster molecules, and the subsequent square root does not cancel this weighting.

Note the following fact: *Whenever the distribution of speeds has any spread, the rms speed exceeds the simple mean speed; the two are equal only if all molecules move at the same speed.*

Example 30

- Explain the molecular mechanism by which the pressure of a gas in a sealed rigid container increases when the temperature of the gas is raised.
- A rigid gas cylinder is rated to withstand internal pressures up to $1.5 \times 10^6 \text{ Pa}$. The cylinder is filled with nitrogen gas at a temperature of 20°C to a pressure of $1.0 \times 10^6 \text{ Pa}$. Determine the maximum temperature to which the gas can be heated before the pressure reaches the cylinder's rating, and calculate the percentage increase in the root mean square speed of the nitrogen molecules between the original temperature and this maximum.

Solution

- At a higher temperature, the molecules of the gas have a larger mean square speed and therefore move faster. Each molecule striking a wall delivers a larger momentum change per collision (because of its greater speed) and also strikes the walls more often per second (because it crosses the container more rapidly between collisions). Both effects multiply together: the rate of momentum transfer to the walls, and hence the pressure, rises in direct proportion to the absolute temperature at constant volume.
- At constant V and constant n , $PV = nRT$ gives $\frac{P}{T} = \text{constant}$.

$$T_1 = 20 + 273 = 293\text{K}:$$

$$T_{\text{max}} = T_1 \times \frac{P_{\text{max}}}{P_1} = (293\text{K}) \times \frac{1.5 \times 10^6 \text{ Pa}}{1.0 \times 10^6 \text{ Pa}} = 440\text{K} = 167^\circ\text{C}$$

The maximum safe temperature for the gas is 167°C .

Since $v_{\text{rms}} \propto \sqrt{T}$:

$$\frac{v_{\text{rmsmax}}}{v_{\text{rms}_1}} = \sqrt{\frac{T_{\text{max}}}{T_1}} = \sqrt{\frac{440\text{K}}{293\text{K}}} = 1.225$$

The rms speed increases by a factor of 1.225, a percentage increase of 22.5%.

Example 31

- Two sealed rigid containers hold different gases at the same temperature. The same quantity of heat is supplied to each container at constant volume. One gas is monatomic and the other is diatomic. Explain which gas undergoes the larger temperature rise and why.
- Two separate rigid containers, one holding 3.0 moles of helium gas and the other holding 3.0 moles of nitrogen gas, are initially at a temperature of 27°C . A quantity of heat equal to 5000J is supplied to each container at constant volume. Calculate the final temperature of each gas, and the ratio of the temperature rises.

Solution

- The heat supplied at constant volume goes entirely into increasing the internal energy of the gas. A diatomic molecule has five degrees of freedom and stores energy in translational motion and in two modes of rotation, while a monatomic molecule has only three degrees of freedom and stores energy in translation alone. For a given amount of heat, the diatomic gas must distribute that energy across five channels instead of three, so the energy per degree of freedom (and hence the temperature rise) is smaller in the diatomic gas. The monatomic gas therefore undergoes the larger temperature rise.

- At constant volume, $Q = \Delta U = \frac{f}{2} \times nR\Delta T$, so $\Delta T = \frac{2Q}{fnR}$.

For helium ($f = 3$):

$$\Delta T_{\text{He}} = \frac{2 \times (5000\text{J})}{3 \times (3.0\text{mol}) \times (8.314\text{Jmol}^{-1}\text{K}^{-1})} = 133.7\text{K}$$

For nitrogen ($f = 5$):

$$\Delta T_{\text{N}_2} = \frac{2 \times (5000\text{J})}{5 \times (3.0\text{mol}) \times (8.314\text{Jmol}^{-1}\text{K}^{-1})} = 80.2\text{K}$$

Final temperatures:

$$T_{\text{fHe}} = (27 + 273)\text{K} + 133.7\text{K} = 433.7\text{K} = 160.7^\circ\text{C}$$

$$T_{\text{fN}_2} = (27 + 273)\text{K} + 80.2\text{K} = 380.2\text{K} = 107.2^\circ\text{C}$$

Ratio of temperature rises:

$$\frac{\Delta T_{\text{He}}}{\Delta T_{\text{N}_2}} = \frac{133.7\text{K}}{80.2\text{K}} = \frac{5}{3}$$

The helium reaches 160.7°C and the nitrogen reaches 107.2°C .

The ratio of temperature rise of helium to that of nitrogen is 5:3.

Example 32

- (a) Two ideal gases with different molar masses are at the same temperature. When placed in thermal contact, no net transfer of heat occurs between them. Explain why this is so, even though the rms speeds of the molecules in the two gases are not equal.
- (b) For one mole of an ideal gas at a temperature of 400K , calculate: (i) the average translational kinetic energy of a single molecule, (ii) the total translational kinetic energy of the gas sample, and (iii) the root mean square speed of nitrogen molecules at this temperature (molar mass of nitrogen = 28g/mol).

Solution

- (a) Thermal equilibrium between two gases is a statement about the average kinetic energy per molecule, not about the average molecular speed. At the same temperature, molecules of both gases carry the same average translational kinetic energy, given by $(3/2)kT$. When molecules of the two gases interact at the boundary between them, collisions between a fast-moving light molecule and a slower-moving heavy molecule on average transfer no net energy in either direction, because the average kinetic energies are already equal. No heat flows between the two gases because no energy imbalance exists at the molecular level, even though the molecular speeds differ.

- (b)(i) Average translational kinetic energy per molecule:

$$\text{KE} = \frac{3}{2} kT = \frac{3}{2} \times (1.381 \times 10^{-23}\text{JK}^{-1}) \times (400\text{K}) = 8.29 \times 10^{-21}\text{J}$$

The average translational kinetic energy per molecule is $8.29 \times 10^{-21}\text{J}$.

- (ii) Total translational KE of one mole:

$$\text{KE}_{\text{total}} = \frac{3}{2} nRT = \frac{3}{2} \times (1\text{mol}) \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) \times (400\text{K}) = 4988\text{J}$$

The total translational kinetic energy of one mole at 400K is 4988J .

- (iii) Rms speed of nitrogen molecules:

$$v_{\text{rms}} = \sqrt{\frac{3RT}{M}} = \sqrt{\frac{3 \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) \times (400\text{K})}{28 \times 10^{-3}\text{kg/mol}}} = 597\text{m/s}$$

The root mean square speed of nitrogen molecules at 400K is 597m/s .

Example 33

- (a) Explain why the rotational kinetic energy of gas molecules contributes to the internal energy of a diatomic gas but does not contribute to the pressure it exerts on the walls of its container.

- (b) 2.0 moles of nitrogen gas (diatomic) are held at a temperature of 27°C. Calculate: (i) the total translational kinetic energy of the gas sample, (ii) the rotational kinetic energy per molecule, (iii) the total rotational kinetic energy of the sample, (iv) the total internal energy of the sample.

Solution

- (a) Pressure arises from the change of momentum of molecules when they collide with the walls of the container, and only the translational motion of the centre of mass of a molecule contributes to this momentum change. A rotating diatomic molecule carries angular momentum about its own centre of mass, but rotation does not change the position of the centre of mass and therefore does not change the molecule's translational momentum. The rotational kinetic energy is internal to the molecule and does not influence the momentum transferred to a wall. Internal energy, however, is the total kinetic energy of all the molecules in all their modes of motion, so it counts translational and rotational kinetic energies together. Rotational KE adds to internal energy but not to pressure.

(b) $T = 27 + 273 = 300\text{K}$.

- (i) Total translational KE = $(3/2)nRT$:

$$KE_{\text{trans}} = \frac{3}{2} \times (2.0\text{mol}) \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) \times (300\text{K}) = 7483\text{J}$$

The total translational kinetic energy is 7483J.

- (ii) For a diatomic molecule, rotational degrees of freedom $f_{\text{rot}} = 2$.

By equipartition each rotational degree of freedom stores $(1/2)kT$ of energy per molecule, so the rotational KE per molecule is:

$$KE_{\text{rot},1} = f_{\text{rot}} \times \frac{1}{2} kT = 2 \times \frac{1}{2} \times (1.381 \times 10^{-23}\text{J K}^{-1}) \times (300\text{K}) = 4.14 \times 10^{-21}\text{J}$$

The rotational kinetic energy per molecule is $4.14 \times 10^{-21}\text{J}$.

- (iii) Total rotational KE of the sample = $(f_{\text{rot}}/2)nRT$:

$$KE_{\text{rot,total}} = \frac{2}{2} \times nRT = nRT = (2.0\text{mol}) \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) \times (300\text{K}) = 4988\text{J}$$

The total rotational kinetic energy of the sample is 4988J.

- (iv) Total internal energy = translational KE + rotational KE = $(f/2)nRT$ with $f = 5$:

$$U = \frac{5}{2} nRT = \frac{5}{2} \times (2.0\text{mol}) \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) \times (300\text{K}) = 12471\text{J}$$

The total internal energy of the sample is 12 471J, which is indeed the sum of the translational (7483J) and rotational (4988J) contributions.

- (v) Ratio of rotational to translational KE:

$$\frac{KE_{\text{rot,total}}}{KE_{\text{trans}}} = \frac{4988\text{J}}{7483\text{J}} = \frac{2}{3}$$

The ratio of rotational to translational kinetic energy is 2:3.

Understand: This fraction equals exactly the ratio of rotational to translational degrees of freedom (2 versus 3), as the equipartition theorem requires.

Example 34

- (a) The speed at which a sound wave travels through a gas is close to, but not exactly equal to, the root mean square speed of the molecules of the gas. Explain why the two speeds are closely related but not identical.
- (b) A sample of an ideal gas undergoes a change in which its volume is doubled and its pressure is raised to four times its original value. If the root mean square speed of its molecules before the change was 250m/s, calculate the root mean square speed after the change.

Solution

- (a) A sound wave is propagated through a gas by successive collisions between molecules: a local compression is passed from molecule to molecule across the gas, and the speed at which this handoff

occurs cannot exceed the speed at which the molecules themselves are moving. The speed of sound is therefore approximately equal to the rms speed of the molecules. The two are not exactly equal because a sound wave involves compression, which is an adiabatic process, and the full analysis gives

$v_{\text{sound}} = \sqrt{\frac{\gamma RT}{M}}$, whereas $v_{\text{rms}} = \sqrt{\frac{3RT}{M}}$. The two formulas differ by the factor $\sqrt{\gamma/3}$, which for a diatomic gas with $\gamma = 1.4$ gives a sound speed about 68% of the rms speed; for a monatomic gas with $\gamma = \frac{5}{3}$ the ratio is $\sqrt{\frac{5}{9}} = 0.75$. The two speeds therefore track each other closely but are distinct.

- (b) From the ideal gas equation $PV = NkT$, if V is doubled and P is quadrupled:

$$\frac{T_2}{T_1} = \frac{P_2 V_2}{P_1 V_1} = \frac{4P \times 2V}{P \times V} = 8$$

The new temperature is therefore 8 times the original temperature.

Since $v_{\text{rms}} \propto \sqrt{T}$:

$$\frac{v_{\text{rms}_2}}{v_{\text{rms}_1}} = \sqrt{\frac{T_2}{T_1}} = \sqrt{8} = 2.83$$

$$v_{\text{rms}_2} = 2.83 \times (250 \text{ m/s}) = 707 \text{ m/s}$$

The new root mean square speed of the molecules is 707 m/s.

Example 35

- (a) An ideal gas, strictly defined, cannot be liquefied however much it is cooled or compressed. Explain why the liquefaction of real gases is not in conflict with the kinetic theory but is actually evidence about one of its assumptions.
- (b) A sealed flask contains a fixed mass of neon gas initially at a temperature of 20°C and a pressure of $1.2 \times 10^5 \text{ Pa}$. The flask and its contents are cooled slowly to -50°C . Calculate the percentage change in the pressure of the gas and the percentage change in the root mean square speed of the neon atoms, and comment on whether the total internal energy of the sample has changed.

Solution

- (a) Liquefaction requires molecules to stick together, and molecules can stick together only if there are attractive forces between them. The kinetic theory of an ideal gas assumes these intermolecular forces to be zero, so an ideal gas cannot liquefy by definition. Every real gas, however, liquefies when cooled sufficiently, because real molecules do attract each other weakly through van der Waals forces. The existence of liquefaction in real gases is therefore direct evidence that the assumption of zero intermolecular forces, while an excellent approximation at ordinary conditions, is not strictly true, and it fixes the regime (low temperature, high density) where that assumption begins to fail.

- (b) $T_1 = 20 + 273 = 293 \text{ K}$; $T_2 = -50 + 273 = 223 \text{ K}$.

At constant V and constant n , $P \propto T$:

$$\frac{P_2}{P_1} = \frac{T_2}{T_1} = \frac{223}{293} = 0.761$$

Fractional change in pressure:

$$\frac{P_2 - P_1}{P_1} = 0.761 - 1 = -0.239$$

The pressure decreases by a percentage of 23.9%.

Fractional change in rms speed, using $v_{\text{rms}} \propto \sqrt{T}$:

$$\frac{v_{\text{rms}_2}}{v_{\text{rms}_1}} = \sqrt{\frac{T_2}{T_1}} = \sqrt{0.761} = 0.872$$

$$\frac{v_{\text{rms}_2} - v_{\text{rms}_1}}{v_{\text{rms}_1}} = 0.872 - 1 = -0.128$$

The rms speed decreases by a percentage of 12.8%.

The total internal energy of the sample has decreased. Neon is monatomic ($f = 3$), so the internal energy is $U = (3/2)nRT$, which is directly proportional to the Kelvin temperature. Since T has fallen from 293K to 223K, the internal energy has fallen by the same fraction of 0.239 (or 23.9%). The energy removed has been transferred as heat to the cooling medium surrounding the flask; it has not disappeared, but it is no longer in the gas.

Fourteen miscellaneous worked examples have put molecules, pressure, temperature, and degrees of freedom to work on problems that blend the chapter's ideas together. Having enjoyed the full combination of ideas, it is time to sharpen our thinking; the Digging Deeper Exercise is ready in the next page.

DIGGING DEEPER EXERCISE 4

EXERCISE 4A: BINDER QUESTIONS

Question 1

A student claims that the pressure of a gas in a container increases when the gas is compressed because “*the molecules touch each other more often.*” Identify two errors in the student’s reasoning and state what really causes the pressure increase.

Question 2

- (a) Explain why the kinetic theory assumes that molecules in an ideal gas move in straight lines between collisions rather than in curved paths.
- (b) Explain why the kinetic theory assumes that the time spent in a collision is negligible compared with the time between collisions.

Question 3

A cubic container and a cylindrical container hold equal amounts of the same gas at the same temperature and pressure. Compare the average rate at which molecules strike a square centimetre of wall in the cube with the average rate at which molecules strike a square centimetre of the curved wall of the cylinder, and explain your reasoning.

Question 4

- (a) A student states that doubling the Celsius temperature of a gas must double the average kinetic energy per molecule. Explain why this claim is wrong.
- (b) Show by algebra that no pair of Celsius temperatures exists for which the student’s rule in (a) gives the correct ratio of 2 for the kinetic energies. Then describe the limit in which the student’s rule becomes approximately correct, and explain briefly why the approximation improves in that limit.

Question 5

One mole of an ideal gas is held in container A, and two moles of the same ideal gas are held in container B of the same volume and at the same temperature.

- (a) Identify which of the following quantities is the same in both containers: the pressure, the total internal energy, the average kinetic energy per molecule, and the root mean square speed of the molecules.
- (b) Identify which of the above quantities is twice as large in container B as in container A.
- (c) Using your answers to (a) and (b), explain briefly the distinction between quantities that are intensive and quantities that are extensive.

Question 6

A student computes the simple arithmetic mean of the speeds of a sample of ten molecules and calls the result the root mean square speed. Explain the student’s error and identify the one circumstance in which the student’s result would happen to equal the true root mean square speed.

Question 7

- (a) Explain why a diatomic molecule such as nitrogen has two rotational degrees of freedom rather than three, even though rotation about three perpendicular axes is geometrically possible.
- (b) Explain why the corresponding argument in (a) does not apply to a polyatomic molecule such as water vapour, which is treated as having three rotational degrees of freedom.

Question 8

A student argues that compressing an ideal gas (reducing its volume while keeping its temperature constant) must increase the gas’s internal energy. Identify the error in the student’s claim by appealing to the result, established in this chapter, that the internal energy of an ideal gas depends only on its absolute temperature.

Question 9

A sealed container holds a real gas. The container is compressed to one-tenth of its original volume at constant temperature.

- (a) State whether the pressure of the real gas rises by exactly a factor of ten, by more than ten, or by less than ten, and explain which assumption of the kinetic theory is responsible for your answer.

- (b) The compressed gas is then cooled to a very low temperature. Describe the new phenomenon that may occur for the real gas but has no analogue in the ideal gas model, and explain briefly why the ideal gas model cannot accommodate it.

Question 10

The absolute temperature of a sample of an ideal gas is increased from 300K to 1200K. Calculate the factor by which the root mean square speed of its molecules changes, and identify the scaling law you have used.

EXERCISE 4B: REAL QUESTIONS**Question 11**

Kipute notices that a gas cylinder left in direct afternoon sun in Dar es Salaam feels noticeably warmer to the touch than one kept in the shade. The shop owner says this is the reason the spare cylinders are always stored in a shaded store-room.

- (a) Explain, using the molecular picture, why the pressure inside the warm cylinder has increased even though no gas has been added.
- (b) Explain why Tanzanian safety regulations require cylinders to be stored in shaded, well-ventilated places rather than in closed rooms.

Question 12

At Horombo Huts on Mount Kilimanjaro (altitude 3720m), a mountain climber feels mildly breathless. By the time he reaches Uhuru Peak (altitude 5895m) he is gasping between every few words. His guide assures him that the air composition has not changed: it is still about 78% nitrogen and 21% oxygen. Explain to the mountain climber what physical quantity has in fact changed with altitude, and how this accounts for the increasing difficulty of breathing.

Question 13

A mother cooks beans in a sealed pressure cooker. The cooker's safety valve is set to release steam if the internal pressure exceeds twice atmospheric pressure.

- (a) Explain, using the kinetic theory, why the pressure inside the sealed cooker rises as the contents are heated from below.
- (b) Explain why the pressure reaches a stable value rather than rising indefinitely, even if the flame remains lit.

Question 14

A diver descends to a depth of 20 metres below the surface of the sea off Zanzibar, where the total ambient pressure is about three atmospheres. The diver is breathing ordinary compressed air.

- (a) Explain why the partial pressure of nitrogen in the diver's lungs at this depth is approximately three times its value at sea level.
- (b) Nitrogen is chemically inert at sea level but is known to induce symptoms resembling drunkenness ("nitrogen narcosis") at high partial pressures. Explain, using molecular collision rates, why a chemically inert gas can become physiologically active once its partial pressure is raised sufficiently.

Question 15

Ally claims that a helium balloon rises because "*helium atoms are lighter than air atoms.*" **Bakari** objects that Ally is wrong because air is a mixture of different gases rather than a single kind of atom. Resolve the disagreement using the kinetic theory, and explain the sense in which Ally's statement captures the correct physics despite the careless wording.

Question 16

An architect designing an upmarket office in Arusha specifies double-glazed windows filled with argon gas rather than air, despite the higher cost of the argon-filled version.

- (a) Explain, using kinetic theory, two distinct molecular properties of argon that make it a better insulator than air.
- (b) Explain whether replacing argon with krypton (molar mass 84 g/mol, monatomic) would improve the insulation further, identifying which of the two properties changes in the favourable direction.

Question 17

Kipute pumps up her bicycle tyre on a hot Dodoma afternoon. She notices that the metal barrel of the pump becomes warm during vigorous pumping, even though she has not heated it with any external flame. Explain what happens to the air trapped in the pump during each downstroke of the handle, and account for the warmth of the pump body in molecular terms.

Question 18

Air in a shaded well in a rural village is at 24°C, while air on a Dar es Salaam street at midday is at 36°C. A student claims that “*the air molecules on the street move much faster than those in the well.*” Evaluate this claim by computing the ratio of root mean square speeds in the two locations, and comment on whether the adverb “much” is appropriate.

EXERCISE 4C: HOT QUESTIONS

Useful constants: Boltzmann constant $k = 1.381 \times 10^{-23} \text{ J K}^{-1}$; Avogadro constant $N_A = 6.022 \times 10^{23} \text{ mol}^{-1}$; universal gas constant $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$; standard atmospheric pressure $= 1.013 \times 10^5 \text{ Pa}$.

Question 19

A steel cylinder of internal volume 0.040 m^3 contains nitrogen gas at 17°C and a pressure of $8.0 \times 10^6 \text{ Pa}$. A fault develops and the gas temperature rises to 127°C. Calculate the quantity of heat supplied to the gas during this temperature rise.

Question 20

Starting from the kinetic equation $PV = \frac{1}{3}Nm\bar{c}^2$ and the equation of state $PV = NkT$, derive the relation between the root mean square speed of the molecules of an ideal gas and its Kelvin temperature. Hence prove that for any two ideal gases A and B at the same temperature, the ratio of their rms speeds equals the square root of the inverse ratio of their molar masses, and explain briefly the physical reason why this ratio does not depend on temperature.

Question 21

A rigid cylinder holds 4.0 moles of neon gas at 27°C. It absorbs 2000J of heat at constant volume. A second rigid cylinder, identical in shape and size, holds 4.0 moles of oxygen gas also at 27°C and also absorbs 2000J at constant volume. Which cylinder ends up at the higher pressure, and by what factor?

Question 22

A pure gas is observed to diffuse through a porous membrane at a rate equal to one divided by 1.41 times the rate of hydrogen at the same temperature. Identify the gas and state briefly why your identification is unique among common gases at A-Level. (The molar mass of hydrogen is 2.0g/mol.)

Question 23

A rigid sealed container of volume $2.0 \times 10^{-3} \text{ m}^3$ holds helium gas at 127°C and $5.0 \times 10^5 \text{ Pa}$. The container is placed in a room at 27°C and the helium inside cools according to Newton’s law with cooling constant $\lambda = 2.0 \times 10^{-3} \text{ s}^{-1}$. Find the pressure inside the container 10 minutes after it is placed in the room.

Question 24

A sealed vessel of volume 25 litres contains 1.5 moles of helium, 2.0 moles of nitrogen, and 0.5 moles of water vapour, all at 127°C. A dessicant is then introduced that completely absorbs the water vapour, removing it from the gas phase. Calculate the percentage change in the total pressure of the mixture and the percentage change in the total internal energy of the mixture, and explain why the two percentages are not equal.

Question 25

An ideal monatomic gas expands from volume V_0 to volume $2V_0$ along a process in which the pressure and volume are related by $P = \alpha V$, where α is a positive constant. The initial pressure is P_0 and the initial absolute temperature is T_0 . Find the factor by which the root mean square speed of the molecules has changed by the end of the expansion, and find the final internal energy of the gas as a multiple of its initial internal energy.

Chapter 5

THERMODYNAMICS

INTRODUCTION

The aluminium kettle had been sitting on Mr. Akilikubwa's desk since before the first bell, plugged into the wall with the quiet confidence of a demonstration that has been planned for days. It was a small kettle; the kind a college student might keep in a hostel room for boiling water for instant noodles at eleven o'clock at night. But this one had a modification. Balanced on top of the lid, perfectly centred, was a 200-gram iron weight, the squat flat kind that sits on a kitchen scale when a *mama ntilie* is measuring out rice. It was not decorating the kettle. It was sitting there with purpose, the way things on Mr. Akilikubwa's demonstration bench always had purpose, even when that purpose was not yet visible.

Kipanga noticed the kettle the moment he walked in. He noticed it the way a man notices a parked car with its hazard lights on, not as a crisis exactly, but as something that will probably become relevant to him personally before the day is over.

He took his seat. He kept one eye on the kettle.

Mr. Akilikubwa opened the register, clicked the kettle on, and began calling names without looking up. The kettle was apparently not part of the lesson yet. It was simply a kettle on a desk, minding its own business, heating up.

Kipanga: *"Sir, the kettle is on."*

Mr. Akilikubwa: *"Yes, Kipanga."*

Kipanga: *"Should the kettle be on?"*

Mr. Akilikubwa: *"That is an interesting question and the answer is yes."*

Three minutes passed. The kettle began to hiss, a sound that started low and private and slowly became impossible to ignore. Kipanga, who had been taking notes from the previous lesson, put her pen down. The class, by some collective instinct, turned to face the front.

Then the lid moved.

It did not fly off. It did not explode. It simply lifted, just slightly, perhaps two centimetres, tilted forward at a polite angle, and hovered there, trembling, as a thick tongue of steam curled out from underneath it. The 200-gram iron weight, which had been sitting perfectly flat, was now visibly higher than it had been when the lesson started. The steam was holding it up. Quietly, steadily, and without any drama, the steam inside that small aluminium kettle was pushing a 200-gram piece of iron upward against gravity.

The class stared.

Mr. Akilikubwa looked up from the register. He regarded the floating lid with the mild satisfaction of a man who has seen this before and is still, after all these years, quietly pleased that it works.

Mr. Akilikubwa: *"Kipanga."*

Kipanga: *"Sir."*

Mr. Akilikubwa: *"Please come and press the lid back down."*

Kipanga blinked. This seemed reasonable. The lid was misbehaving. He was being asked to correct it. He was a helpful person. He stood up, walked to the desk, extended his right index finger, and pressed down on the lid with the casual confidence of someone who has never thought carefully about what is underneath it.

For a long and instructive moment, Kipanga's finger and the steam inside the kettle conducted a quiet negotiation.

The steam won!

The lid pushed back. Not violently, not dangerously, but with a steady, warm, entirely undeniable upward force that Kipanga had not budgeted for. He pressed harder. The lid pushed back harder. He was, at this moment, a 200-gram iron weight with a university scholarship. He was losing to water vapour.

He stepped back. His finger was pink and slightly damp.

Kipanga: *"It pushed me."*

Mr. Akilikubwa: *"Yes."*

Kipanga: *"The steam pushed my finger. Upward."*

Mr. Akilikubwa: “Yes.”

Kipanga: “That is my hand.”

Mr. Akilikubwa: “Also yes. Please sit down before you have any further conversations with the kettle.”

Kipanga sat down. He looked at his finger. He looked at the kettle. The lid was still hovering. The iron weight was still in the air. He had the expression of a man who has just been bested in arm-wrestling by a cup of boiling water and is not quite sure which part of his life philosophy this should prompt him to revise.

Kipute had not moved from her seat. She had not volunteered, had not rushed forward, had not touched anything. She had simply watched the entire exchange with the calm attention of someone assembling evidence.

Kipute: “Kipanga, where did the energy come from? To push your finger. And to keep that weight in the air.”

Kipanga: “From the steam.”

Kipute: “And where did the steam get it?”

Kipanga: “From the hot water.”

Kipute: “And where did the hot water get it?”

Kipanga looked at the power cord running from the kettle to the wall socket. Then he looked at his finger again. Then he looked at the iron weight, still hovering above the lid.

Kipanga: “From the electricity.”

Kipute: “So electricity went in through that cable, and a 200-gram weight went up into the air. Something converted one into the other. What was the something?”

Kipanga opened his mouth. He closed it. He was not used to being asked questions whose answers he did not immediately have. He was especially not used to Kipute being the one asking them.

At that moment Mr. Akilikubwa switched the kettle off. The hissing faded. The lid sank back with a soft aluminium clunk. The iron weight settled. The steam dissipated. The classroom returned to normal, except for the question still hanging in the air, which is exactly where Mr. Akilikubwa intended it to remain.

Mr. Akilikubwa: “Kipute has asked the question that this entire chapter exists to answer. Electricity went in. A weight went up. Something in the middle did the converting. That something was a thermodynamic process. And the framework that lets us track every joule, every step of the way, from the wall socket to the iron weight and everywhere in between, is called the **first law of thermodynamics**.”

He pulled the power cord from the wall and set the kettle aside.

Mr. Akilikubwa: “Chapter 4 told you what temperature is at the molecular level. It told you that the internal energy of a gas is the total kinetic energy of its molecules, and that the temperature of the gas is simply a measure of how fast those molecules are moving. Beautiful. Correct. But Chapter 4 had one quiet limitation that I never pointed out, because you were not ready for it. In Chapter 4, the gas sat inside its container and did its own molecular business. The walls never moved. The gas never pushed anything. No piston slid. No lid lifted. No finger was embarrassed in front of thirty witnesses.”

Kipanga stared at the kettle with fresh resentment.

Mr. Akilikubwa: “This chapter takes the walls off the nail. Here, the gas is allowed to push on pistons, on lids, on the atmosphere, on whatever is in its way. And when it pushes something through a distance, it does work. Mechanical work. The same work you calculated in Form Four when you pushed a box along a floor. The same joules. The same equation. The difference is that here, the gas is doing the pushing, and the energy to do that pushing came from somewhere, from heat, or from compression, or from both at once. The first law is simply the rule that the accounts must balance. Every joule that goes into a gas must either raise its internal energy or come out again as work. Not some of the joules. Not most of the joules. Every single one.”

The class was quiet in the way that classes are quiet when something has genuinely landed.

Kipute wrote something in her notebook. Kipanga was still looking at his finger, but his expression had changed. It was no longer the expression of a man who had been beaten by a kettle. It was the expression of a man who was beginning to understand that the kettle had been trying to teach him something.

Chapters 2, 3, and 4 have been building toward this point from three different directions. Chapter 2 established that heat is energy in transit between bodies at different temperatures, gave us the Zeroth Law to make the word temperature logically meaningful, and developed the practical tools of specific heat capacity and latent heat that let us calculate how many joules it takes to boil a kettle, melt a block of ice, or cool a hospital ward. Chapter 3 traced the three routes heat can travel, through conduction, convection, and radiation, and showed that the rate of flow matters as much as the total amount. Chapter 4 then reached inside the gas itself, connected temperature to molecular kinetic energy, and delivered the single most important result so far: the internal energy of an ideal gas is $U = (f/2)nRT$, a number written entirely in terms of how many molecules are present and how fast they are moving.

In all three of those chapters, the gas was a passive participant. It sat inside its container, accumulated energy, increased its temperature, and was grateful. Nothing in those chapters required the gas to push on anything.

This chapter removes that restriction.

When the walls of a container are free to move, the gas can do mechanical work on its surroundings. The steam in Mr. Akilikubwa's kettle is pushing the lid upward. A gas in a car engine is pushing a piston downward. The hot gas in a turbine is pushing a blade sideways at hundreds of metres per second. In all of these, the same principle applies: the gas exchanges energy with its surroundings by two channels simultaneously. The first channel is heat, which is energy flowing across a temperature difference, as in Chapters 2 and 3. The second channel is work, which is energy transferred by a force acting through a distance, exactly as in mechanics. The first law of thermodynamics is the single equation that keeps the accounts for both channels at once: every joule entering the gas either increases its internal energy or leaves as work done on the surroundings. There is no third option, no hidden drawer, no energy that simply disappears. The books balance.

This chapter unfolds in three parts.

Welcome to Chapter 5. Kipanga's finger has already had its first lesson. The rest of the chapter will explain what it felt.

THE LANGUAGE OF THERMODYNAMICS

The kettle is back on the bench and the weight is sitting still, but Kipute's question has not gone anywhere. Electricity went in, a weight went up, and something in the middle did the converting. Before we can write an equation for what that something was, we need a small set of words. Thermodynamics is a subject where the vocabulary carries its own weight. Using "heat" when one means "temperature," or "system" when one means "surroundings," will send a student sideways for pages at a time. This section is the dictionary. There are no equations here and no worked examples. What there is, is a short list of terms that every subsequent section will assume has been understood.

Some of the vocabulary is already in place. Chapter 2 distinguished temperature from heat, defined thermal equilibrium, and gave us the **Zeroth Law** that makes the word temperature meaningful in the first place. Chapter 4 introduced the **internal energy** U of a gas and tied it to the temperature and to the molecular degrees of freedom. Those definitions are assumed from this point forward, and the reader is invited to turn back to the earlier chapters if any of that language has gone cold. What is new, and what this section adds, is the vocabulary we need to talk about a gas that is exchanging energy with the world outside it, which is to say, a gas that is doing something to Kipanga's finger.

System, surroundings, and boundary

When a physicist wants to think about some piece of the world, she starts by drawing an imaginary line around it. Everything inside the line is called the **system**. Everything outside the line is called the **surroundings**. The line itself is called the **boundary**.

Consider the kettle. If we are interested in what the steam did to the lid, the natural choice is to call the steam (together with the water underneath it, if we are being thorough) the system. The boundary is then the inner surface of the kettle, together with the underside of the lid. The surroundings are everything else in the story, namely the kettle walls themselves, the 200-gram weight, Kipanga's finger, the hot plate below, and the rest of the classroom. This choice is a decision, not a discovery. We could equally well call the kettle-and-its-contents the system, in which case the boundary becomes the outer surface of the kettle and the hot plate is

surroundings. Neither choice is wrong. They are two different questions about the same physical object, and which one we prefer depends on what we are trying to calculate.

Throughout this chapter, the system will almost always be some quantity of gas inside some container, and the surroundings will be everything outside that container. The boundary will sometimes be rigid, like the walls of a sealed tank, and sometimes movable, like the upper face of a piston or the bottom of a hovering kettle lid. But the principle is general. The first thing one does in any thermodynamics problem is decide what the system is, and once that choice is made, one stays with it. Changing one's mind halfway through a question and calling the piston the system when the gas was the system a moment ago is how sign errors are born.

Open, closed, and isolated systems

Three things can, in principle, cross the boundary of a system.

- 1) Matter can cross, in the form of molecules moving from one side to the other.
- 2) Energy can cross as heat, which is the mode of transfer we met in Chapter 2.
- 3) Energy can also cross as work, which is the mode that appears whenever a force is applied through a distance.

Which of these three the boundary lets through determines what kind of system we are dealing with.

An **open system** allows all three to cross. A pot of *ugali* on a charcoal *jiko* with its lid off is open. Steam escapes into the air (matter leaves), the flame delivers heat through the base of the pot (heat enters), and a *mwiko* stirring the *ugali* does work on the mixture (work enters). Mr. Akilikubwa's kettle during the demonstration was also open, briefly, whenever the lid lifted and the steam escaped around its edges.

A **closed system** allows heat and work to cross, but not matter. The quantity of stuff inside is fixed. A sealed saucepan with a snug lid is closed. No steam escapes. Heat still flows in from the flame, and pressing on the lid still does work on the air trapped inside. The kettle, when its lid was resting flat and no steam was getting past it, was behaving as a closed system.

An **isolated system** allows nothing at all to cross. No matter, no heat, no work. An isolated system is hermetically sealed, perfectly rigid, and perfectly insulated from everything around it. The iconic example is the thermos flask of hot tea that sits on a desk and refuses to cool for hours. Its double-walled vacuum blocks conduction and convection, its silvered inner surface blocks radiation, and its screw-top blocks mass flow. No real flask is perfectly isolated, and the tea will eventually cool if given enough time, but for the duration of a physics problem the flask is close enough to isolated that we treat it as isolated.

Throughout this chapter we work almost entirely with closed systems. The first law of thermodynamics, which is the business of the next section, is a statement about closed systems. A gas sealed in a rigid tank is closed. A gas sealed beneath a movable piston is also closed, because the piston moves but the molecules stay inside. The kettle with its lid hovering was at the edge of closed and open, which is why the demonstration was interesting but also why a quantitative analysis would need us to be careful about the steam leaking out. For the next several sections we will work with containers that do not leak.

Thermodynamic state and state variables

Once we have decided what the system is, we need to describe its condition. The physical condition of a gas can be specified completely by a small number of measurable quantities: its **pressure P**, its **volume V**, its **temperature T**, and the **amount of substance n**, which we measure in moles. These four quantities are called **state variables**. They have this name because *a complete specification of all four amounts to a complete description of the gas*. Two samples of gas with the same P, V, T, and n are thermodynamically identical, whether one is in Bukoba and the other in Mbeya. Nothing else need be said about them. Their future behaviour, if handed over to another physicist, will be indistinguishable.

The set of values (P, V, T, n) at a particular moment is called the **thermodynamic state** of the system. The state is a snapshot, a list of numbers true at one instant. At the moment Mr. Akilikubwa switched on the hot plate, the air inside the kettle was in one state. By the time the lid lifted, it was in a different state: the temperature had risen, the pressure had built up, and some of the water had joined the gas as steam, so the amount of substance present in the gaseous part of the system had increased. Every step of the demonstration was a change from one state to another.

The Zeroth Law of Chapter 2 is quietly doing its work every time we write the letter T in that list. Without the Zeroth Law, the temperature of the gas would not be a single well-defined number for the whole body;

it would be some tangle of pair-by-pair relationships, and the state description would fall apart before we got started. It does not fall apart, because the Zeroth Law holds.

For an ideal gas, the four state variables are not independent. They are tied together by the equation of state $PV = nRT$ that we derived in Chapter 4. Knowing any three of them determines the fourth. Specify P , V , and n , and T is forced. Specify P , T , and n , and V follows. The state space of an ideal gas is therefore really three-dimensional, even though we talk about four variables. In practice this is convenient, because most problems hold one of the four variables fixed while the others change, and the language of “constant pressure” or “constant temperature” is more natural than insisting on three degrees of freedom and solving the constraint each time.

Equilibrium and quasi-static processes

The state variables P , V , T have meaningful values only when the gas has had time to settle down, so that its molecules have redistributed themselves evenly, any temperature gradients have died out, and the pressure has equalised from wall to wall. The notion of thermal equilibrium between two separate bodies was defined in Chapter 2. Here we extend it to a single body in equilibrium with itself. A gas is in **thermodynamic equilibrium** when its state variables are well-defined, uniform throughout the body of gas, and unchanging with time in the absence of any external disturbance. If we have left a tank of gas alone for long enough, its pressure is the same at the top as at the bottom, its temperature is the same in the middle as at the walls, and if no one switches on a flame or pulls out a piston, all the state variables will still be the same an hour from now. That is equilibrium.

Thermodynamics is a subject about equilibrium states joined by processes. The gas starts in some equilibrium state A and finishes in some other equilibrium state B . The transition is, strictly speaking, complicated. During the rapid lifting of Mr. Akilikubwa’s kettle lid, the pressure was not uniform across the gas inside the kettle; pressure waves were bouncing from wall to wall, and the temperature near the hot plate was different from the temperature at the top. For calculations, this is inconvenient. A gas whose state is not well defined cannot be assigned a point on a PV diagram, and a process whose intermediate stages cannot be drawn cannot have its work easily computed.

The standard way around this is to assume the process happens slowly enough that the gas passes through a continuous sequence of near-equilibrium states. A process that satisfies this condition is called **quasi-static**. Quasi-static is an idealisation, not a description of any real process, but it is an idealisation good enough to get the right answer to within a few percent in most A-Level problems, which is the accuracy we are aiming for. When a question describes the gas as “slowly compressed” or “gradually heated,” it is instructing us to treat the process as quasi-static. When a question says the tyre “suddenly bursts” or the gas is “released rapidly,” it is flagging precisely the opposite, namely a non-quasi-static process that we handle differently. We will meet that different handling later in the chapter, when we reach the adiabatic process.

Thermodynamic process and the path that joins the states

A **thermodynamic process** is any change of state. The gas starts in state A and ends in state B , and the process is the sequence of intermediate states that takes it from one to the other. The demonstration with the kettle was a thermodynamic process, as was the cooling of the tea in the thermos flask, and as is every compression of the air in a bicycle pump, every expansion of the gases in a car engine, and every breath the reader has ever taken.

Here is the central idea that makes this subject interesting, and it is the idea that will be formalised in the section on work done by a gas. The same initial state A and the same final state B can be connected by many different processes, by many different paths through state space. Different paths do different things along the way. One path might heat the gas first and compress it afterwards. Another might compress it first and heat it afterwards. A third might do both at once. The final state is the same either way, but how much heat flowed in along the way, how much work was done, and in which direction each was pointing, all of these depend on the path. In thermodynamic language, **heat and work** are *path-dependent* quantities. The **state** itself is *not*. A snapshot does not care how you arrived at it. The bookkeeping does.

This is the deepest reason why the four processes we are about to meet give four different answers for the work done, even when they start and end in identical states. It is also the reason we draw PV diagrams, which are the tool for making path-dependence visible. For now, hold on to the shape of the idea. State is where you are; process is how you got there; and the “how” affects the balance sheet even when the “where” is the same.

The four workhorses

Thermodynamic processes come in many varieties, but four of them are worth particular study. They are simple, they are common, and between them they cover almost all the physics an A-level student needs. What distinguishes the four is which of the state variables is held constant during the process. The names, in the tradition of the subject, are Greek.

The four processes are collected in the preview table below. The heat Q and work W columns are left empty for now, because filling them in is the business of the sections that follow. Treat the table as a roadmap. Put a bookmark at this page. After the isochoric section is read, come back and fill in the isochoric row; after the isobaric section is read, fill in the isobaric row; and so on. By the end of the chapter the table will be complete, and the complete version will become the single most important reference in this chapter.

Process	What is constant	PV diagram shape	Heat Q	Work W
Isochoric	Volume V	Vertical straight line	to come	zero (no volume change)
Isobaric	Pressure P	Horizontal straight line	to come	to come
Isothermal	Temperature T	Gentle hyperbola, $PV = \text{constant}$	to come	to come
Adiabatic	Heat exchange ($Q = 0$)	Steep hyperbola, $PV^\gamma = \text{constant}$	zero (no heat flows)	to come

The Greek roots explain the names. **Isochoric** combines **iso** (*equal*) with **choros** (*space, volume*), so the gas is held at constant volume. **Isobaric** combines **iso** with **baros** (*pressure*), so the gas is held at constant pressure. **Isothermal** combines **iso** with **thermos** (*heat, temperature*), so the gas is held at constant temperature. And **adiabatic** comes from **adiabatos**, meaning “*not passing through*”. It is a name about the heat, which in an adiabatic process does not pass through the walls of the container. Physicists have been borrowing words from Greek for thousands of years, and the habit is now too strong to change.

Each of these four processes is met in turn as the chapter unfolds. For each we draw the PV diagram, derive the work done, find how much heat flows, and compute the change in internal energy. By the time the synthesis section is reached, the reference table above will be complete in every cell. Before any of that can happen, however, we need the law at the heart of this subject, the law that answers Kipute’s question about what converted the electricity into a hovering iron weight. That is the business of the next section.

THE FIRST LAW OF THERMODYNAMICS

At the end of the introduction, Kipute asked a question. Electricity had flowed in through a cable, a 200-gram iron weight had gone up two centimetres into the air, and something in the middle had done the converting. The previous section gave the something a name, a **thermodynamic process**, and put six words on the desk to talk about it: system, surroundings, boundary, state, equilibrium, path. Naming a thing is a good first step. It is not the same as understanding it. The rule that tracks every joule of Kipute’s electricity, from the hot plate through the water and the steam and out again into the gravitational potential energy of a kitchen-scale weight, is called the **first law of thermodynamics**, and it is the single most important sentence in this chapter.

Return to the kettle for a moment. The hot plate is on, the lid is resting flat, and the steam inside is quietly building pressure. Three things are happening at once, and the first law is nothing more than the observation that these three things always balance. Heat is flowing into the kettle from the hot plate, because the hot plate is hotter than the water. The water and the steam inside the kettle are getting hotter, which means their internal energy is rising, because internal energy is just the total kinetic energy of the molecules, and hotter molecules move faster. And if the lid is free to move, some of that incoming energy is going back out again as mechanical work done on the lid, and on whatever is resting on the lid. Three things. Three arrows. One kettle. The books balance.

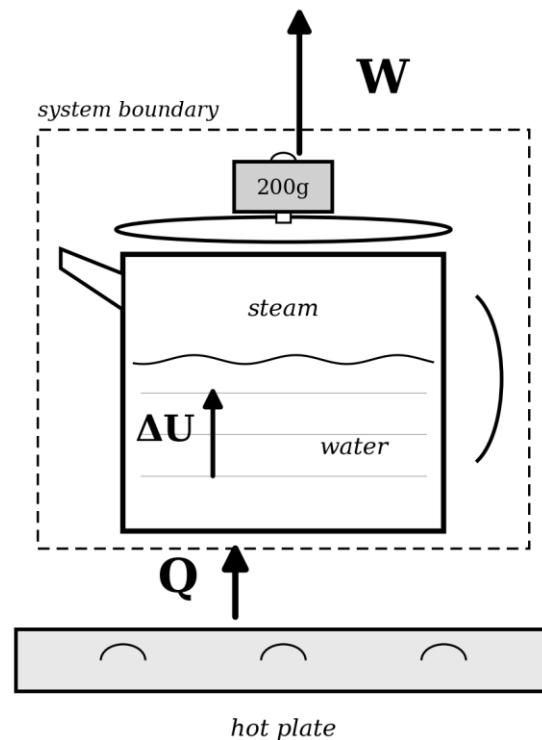


Figure: The three energy flows at work in Mr. Akilikubwa's kettle. Heat Q enters from the hot plate below. Work W leaves through the rising lid (which is lifting the 200g iron weight against gravity). The change in internal energy ΔU , shown as a vertical arrow inside the body of water, tracks whatever is neither entering nor leaving. The first law says these three arrows balance to the last joule.

The formal statement is short, and it is worth reading twice. *For any closed system undergoing a thermodynamic process, the heat Q supplied to the system is equal to the increase ΔU in its internal energy plus the work W done by the gas on its surroundings:*

$$Q = \Delta U + W$$

Every joule of heat that goes into the gas either stays inside the gas as added internal energy, or leaves again as mechanical work done on the surroundings, or splits itself between the two in some ratio. There is no third option. There is no hidden drawer in the wall of the kettle where a joule can disappear and reappear later. If a joule went in as heat, it is still there in one of the two forms, or it has already left again as work. That is the whole of the law, and that is why, at the instant Mr. Akilikubwa switched on the hot plate, the outcome of the entire demonstration was already decided. The electricity was going to go in. Some of it was going to stay inside the kettle as hotter water and hotter steam. The rest was going to come back out again as the gravitational potential energy of a rising iron weight. The exact split depended on the details, but the total was going to balance. Kipute's question, ten minutes later, was really a question about the split.

The law is often also written in the rearranged form:

$$\Delta U = Q - W,$$

which says the same thing with the terms shuffled around: *the change in internal energy equals the heat supplied minus the work done by the gas*. Both forms are correct. A reader should be comfortable moving between the two without fuss. What must not change, whichever form is in use, is the meaning of the three symbols. Q is always *the heat supplied to the gas by its surroundings*. W is always *the work done by the gas on its surroundings*. ΔU is always *the change in the internal energy of the gas itself*. Reverse any of those three definitions, and the signs start to go the wrong way.

Sign convention

A sign convention is the most important piece of mental hygiene in this chapter. More students lose marks on NECTA thermodynamics questions through sign errors than through any other single mistake, and the only defence is to decide on the convention at the start and never deviate from it. The one we adopt is the standard one, locked into the form of the first law written above, and every worked example in this chapter will rehearse it explicitly before quoting a final answer.

Symbol	Positive means	Negative means	Unit
Q	heat enters the system	heat leaves the system	J
W	gas expands; gas does work on surroundings	gas is compressed; surroundings do work on gas	J
ΔU	internal energy rises; temperature rises	internal energy falls; temperature falls	J

A habit worth installing immediately: when the numbers at the end of a calculation come out with a sign that contradicts the physics of the situation, such as an answer that says the gas cooled while it was being compressed, or that the gas released heat while it was being flamed from below, the mistake is almost always in a sign, not in the formula. Verifying the sign of each quantity against physical expectation, before the final answer is stated, is not a habit to be learned in a panic the night before an exam. It is a habit to be installed right now, in the first worked example of the section, and rehearsed in every one that follows.

The law on which the first law rests

The first law of thermodynamics is not really a new law. It is the **law of conservation of energy**, applied to a system that exchanges energy with its surroundings through two particular channels: **heat** and **work**. Conservation of energy is itself one of the oldest and most secure principles in physics. It was known in its mechanical form in the eighteenth century, when it was noticed that the kinetic energy of a falling body equals the loss of its gravitational potential energy, no matter the shape of the path it falls along. What the first law adds to that older statement is the recognition that heat, which was once thought to be a separate substance called caloric that flowed invisibly from hot things to cold things, is in fact just another form of energy, and can be traded for mechanical work in a strict ratio that never changes.

The ratio is worth remembering, because it is the experimental foundation of the subject. Every 4.186 joules of mechanical work, fully converted into heat, raises the temperature of one gram of water by one degree Celsius. The figure was established by James Prescott Joule in the 1840s, in a series of patient and beautifully designed experiments that used a falling weight to turn a paddle inside an insulated tank of water. The falling weight did a measurable amount of mechanical work. The paddle stirred the water. The friction between the paddle and the water converted that mechanical work into heat. Joule measured the temperature rise of the water, weighed the weight, measured the height it fell through, and computed the ratio of mechanical work to heat produced. He repeated the experiment with different weights, different heights, different volumes of water, and even with different liquids, and he always obtained the same number. His answer has not been improved by more than a small correction in the hundred and eighty years since. The unit of energy we use today is named after him, and the first law is in some sense simply the promotion of his observation from a single experimental ratio into a universal principle: *mechanical work and heat are two forms of the same thing, and the books of energy, written in joules, must balance whenever the two are exchanged.*

Evidence for the validity of the first law

The first law has never been caught breaking. Every serious attempt to construct a counterexample has failed. The evidence for its validity comes from four independent directions, each of which would be compelling on its own, and which together leave the law on an experimental foundation as solid as anything in physics.

1. Joule's paddle-wheel experiment

Repeated over the last century and a half with ever more refined apparatus, Joule's experiment always gives the same mechanical-to-heat conversion ratio. The ratio is universal: it does not depend on the material being

heated, on the mechanism by which the mechanical work is done, or on the scale of the apparatus. This universality is exactly what one would expect from a conservation law, and exactly what one would not expect from anything else.

2. Electrical heating

When a current passes through a resistor, electrical energy is converted into heat at a rate predicted exactly by the power dissipated in the resistor. Electric kettles, of the kind sitting on Mr. Akilikubwa's desk, work on this principle, and they convert their input electrical energy into heat with an efficiency close to one hundred percent. If any joule of the input were going missing, the mismatch would show up in a careful measurement of electrical energy supplied versus thermal energy released. No such mismatch has ever been found. Billions of kettles have been boiled by this physics. None has cheated.

3. The absence of any perpetual motion machine of the first kind

A **perpetual motion machine** of the first kind *is a machine that produces more energy than it consumes, or produces energy from no source at all*. Such a machine would solve every energy problem the world has ever had, and its inventor would become the wealthiest human in history. Thousands have been proposed, from medieval overbalanced wheels to modern patents involving hidden magnets and claims of overunity generation. None has ever worked. Patent offices around the world now routinely decline to examine such proposals, because three centuries of experience have made clear that the first law will defeat them all.

4. The balance sheet of every real engine

Every engine ever operated, from the steam engines of eighteenth-century England to the gas turbines at Ubungu power station in Dar es Salaam, has consumed at least as much energy in fuel as it has produced in useful work. A careful audit of any real engine, accounting for the fuel energy burned, the waste heat rejected, and the useful work extracted, always finds that the books balance. In every case the amount of fuel energy consumed is at least as large as the useful work produced. In fact it is considerably larger, because most of the fuel energy escapes as waste heat rather than becoming useful work, but the reason that this should be so is not a matter of the first law. It is a matter of the second, which is waiting at the end of this chapter.

Limitations of the first law

For all its power and all its evidence, the first law is not the whole story. It has limitations, and being clear about them prepares the ground for the second law at the end of the chapter. Three of these limitations are worth naming explicitly, and every A-level student should be able to articulate them.

1. The first law does not predict the direction of a process

Consider a cup of hot tea on a desk in a cool classroom. Over the next hour the tea will cool, and the classroom air will warm very slightly, until the two are at the same temperature. This is the familiar direction. Now imagine running the film backwards: the cooled tea spontaneously heating itself up, while the classroom air gives up a tiny fraction of its thermal energy, until the cup of tea is once again steaming. No joule is missing in the backwards film. Every joule that left the tea in the forward process is simply flowing back into it. The first law is perfectly satisfied by the backwards film. But the backwards film never happens. Nobody has ever seen a cup of tea warm itself by extracting heat from a cooler room, and nobody ever will. The first law does not forbid this process, and that silence is its first limitation.

2. The first law does not distinguish between high-quality and low-quality forms of energy

From the first law's point of view, a joule of mechanical work and a joule of heat are interchangeable; both are just a joule. But in practice, they are not interchangeable. A joule of mechanical work can always be completely converted into a joule of heat, as happens every time someone rubs their hands together on a cold Iringa morning, or every time a brake pad dissipates kinetic energy on a long descent from the Usambara escarpment. The reverse is not true. A joule of heat can never be completely converted back into a joule of mechanical work, no matter how ingenious the machinery. Some fraction is always lost, rejected to a cooler reservoir. Heat is a lower-quality form of energy than work, because work freely becomes heat while heat does not freely become work. The first law, by counting both in joules, erases this distinction, and to that extent it under-describes the physics.

3. The first law sets no ceiling on the efficiency of energy conversion

Because the first law does not distinguish between heat and work by quality, it cannot predict how efficiently one can be converted into the other. A heat engine supplied with a thousand joules of heat might, for all the

first law knows, deliver anywhere from zero to a thousand joules of useful work. Engineers designing real engines need to know the ceiling in advance, and the first law cannot tell them where it is.

All three limitations share a common theme. They concern what the first law does not predict, not what it predicts wrongly. Within its proper domain, which is the tracking of energy between the three accounts of heat, work, and internal energy, the first law has an unbroken record of two centuries. For the questions that lie outside that domain, such as which direction a process will spontaneously run in, or how efficient an engine can possibly be, a deeper principle is needed.

Enough theory. Four worked examples now follow, each short, each rehearsing the sign convention, and each chosen so that a different species of sign error gets its audition. We begin with the tamest.

BINDER Example 1

A gas absorbs 500J of heat, and during the process its internal energy increases by 200J. Find the work done by the gas, and describe in physical terms what the gas is doing.

Solution

Apply the first law in its standard form, with every sign decided by the physics of the situation before any numbers are substituted in. The gas absorbs heat, so heat enters the system, which in our sign convention means Q is positive: $Q = +500\text{J}$. The internal energy of the gas increases, which means the gas becomes hotter, consistent with heat going in and the temperature rising, so ΔU is positive: $\Delta U = +200\text{J}$. The work done by the gas is what is being asked for, and the first law supplies it directly:

$$W = Q - \Delta U = 500\text{J} - 200\text{J} = 300\text{J}$$

The sign of W is positive, which means the gas has done 300J of work on its surroundings.

Making Sense of the Answer: *Of the 500J that entered the gas as heat, 200J stayed inside it as raised internal energy, in the form of hotter, faster-moving molecules, and 300J left again as mechanical work done on the surroundings. The two destinations add up to 500J exactly, as the first law requires. In physical terms, the gas is expanding against its surroundings while being heated from below. One might picture a gas in a cylinder with a movable piston, warmed by a flame, pushing the piston upward while its temperature also rises. Every joule is accounted for.*

Think Like a Physicist: *The first law has told us how the books balance, but it has said nothing about the mechanism. Is the gas lifting a piston? Pushing back the atmosphere? Turning a turbine blade? The equation is silent on the question. It guarantees only that the total is 300J of work and 200J of retained heat, whatever the arrangement happens to be. This is both the power of the law, since it is universal and applies equally to any closed system, and its deliberate limitation: the law does not care about the particular piston. A physicist who wants the particular piston must bring in a separate piece of physics, which in this chapter will be the PV diagram of the next section.*

BINDER Example 2

A gas is compressed quickly by an external force, which does 300J of work on the gas. During the compression, 120J of heat leaks out of the gas into the surroundings. Find the change in internal energy of the gas.

Solution

This example is designed to rehearse every sign in the convention, and it rewards readers who take each symbol in turn, verifying the sign against the physical situation, before touching the equation at all.

The external force does work on the gas, which means the gas is being compressed, not expanding. In our sign convention, which defines W as the work done by the gas on its surroundings, the work done by the gas is the negative of the work done on the gas. So $W = -300\text{J}$. The minus sign is doing real work here. A reader who plugs $+300\text{J}$ into the formula will obtain a final answer with the wrong sign, and will report that the gas cooled down when in fact it has warmed up. Heat leaks out of the gas to the surroundings, which in the convention means Q is negative, so $Q = -120\text{J}$. Now, and only now, apply the first law in its rearranged form:

$$\Delta U = Q - W = (-120\text{J}) - (-300\text{J}) = -120\text{J} + 300\text{J} = +180\text{J}.$$

Making Sense of the Answer: *The change in internal energy is positive. The gas has become hotter, which is exactly what one expects when a gas is rapidly compressed. The compression added 300J of work to the internal budget, and the heat leakage removed 120J, leaving a net gain of 180J. This is the physics of a*

bicycle pump barrel getting warm as air is rapidly pushed through it, which is precisely what Kipanga's pump was doing in the opening pages of chapter 4. It is also the physics of a diesel engine cylinder heating up fast enough to ignite fuel without a spark, which is one of the applications we will meet later.

Think Like a Physicist: The two minus signs that appeared in this problem are not a quirk of algebra; they are the physics. Each minus sign is a separate decision about which direction the energy is flowing. One minus sign records that work is being done on the gas rather than by it. The other records that heat is leaving the gas rather than entering it. A student who treats signs as a mechanical recipe, plugging numbers in hopefully and checking at the end, will fail on this class of problem. A student who treats each sign as answering the question 'which way?' will succeed. The habit to build is not 'put a minus sign in', but 'decide, before I write anything down, whether this quantity is going in or coming out'. Build that habit here, on a simple numerical problem, because it is the habit that carries you through the adiabatic process in a later section.

REAL Example 3

On a cold Iringa morning, a student waiting outside the assembly hall rubs her palms together vigorously. Within a few seconds her palms become noticeably warm, even though no flame, no stove, and no electrical heater is anywhere near her.

- Using the first law of thermodynamics, explain the physical mechanism.
- A few seconds after she stops rubbing, her palms begin to cool back down to the temperature of the morning air. What form does the first law take during that cooling phase?

Solution

(a) The first law of thermodynamics applies to any closed system, not only to a gas inside a container. For this example, take the two palms together as the system. The surroundings are the rest of the student's body, the cold Iringa air, the inside of her school uniform, and everything else beyond the outer surface of her skin.

While she is rubbing, the muscles of her arms and shoulders are doing mechanical work on her palms through the friction between them. Friction converts mechanical work into heat inside the material where the friction occurs, which in this case is the skin of the palms themselves. In the language of the first law, the palms are receiving energy as work rather than as heat. The system is a solid, not a gas, so there is no significant expansion, which means no energy is leaving again as expansion work on the surroundings. All of the incoming work stays in the palms, raising their internal energy. Hotter molecules in the palms mean a higher temperature, and the palms feel warm. The balance during the rubbing phase is therefore

$$\Delta U = -W \quad (\text{palms gain internal energy from work done on them}),$$

with the heat term Q approximately zero over the few seconds of rubbing, because the palms have not had time to lose any appreciable heat to the cold air around them.

(b) When she stops rubbing, the external work stops. The palms are now warmer than the surrounding air, so heat begins to flow outward from palms to air, exactly as chapter 2 described: heat flows spontaneously from the hotter body to the cooler one. In the first law, the work term is now zero, and the heat term is negative because heat is leaving the system. The balance becomes

$$\Delta U = Q \quad (\text{palms lose internal energy as heat escapes to the air}),$$

with Q negative. The internal energy of the palms decreases, which is the physical sensation of her palms cooling back down to the morning temperature.

Making Sense of the Answer: The same first law, written in exactly the same form, describes both halves of the situation. What changes between the two halves is not the law itself but which of the two channels, work or heat, is active. During the rubbing, the work channel is carrying energy in and the heat channel is quiet. After the rubbing, the work channel is quiet and the heat channel is carrying energy out. The student's palms behave as a thermodynamic system exactly like a kettle, an engine cylinder, or the air inside a sealed tyre. The same accounting runs in all four cases.

Think Like a Physicist: Notice what this example has done for us. It has taken the first law, which was stated originally for a gas, and applied it to a pair of human palms. The palms are solid tissue, not gas; there is no piston, no container, no pressure; and yet the first law applies without modification. This is the lesson. The first law is not a law about gases. It is a law about closed systems. Any closed system will do: a gas, a liquid, a solid, a pair of hands, a rock warming in the sun, the earth itself radiating to space. The skill in applying the first law to an unfamiliar situation is never mathematical. It is always conceptual, and it

always comes down to three decisions. What is the system. What counts as heat. What counts as work. Once those three decisions are made, the arithmetic writes itself.

HOT Example 4

One gram of water at 100°C is vapourised at atmospheric pressure into 1671cm³ of steam at the same temperature. The specific latent heat of vapourisation of water at 100°C is 2.26×10⁶Jkg⁻¹, and atmospheric pressure is 1.013×10⁵Pa. Calculate:

- the heat supplied to the water during the vapourisation,
- the external work done by the expanding water against the atmosphere, and
- the increase in the internal energy of the water as it becomes steam.

Solution

This example is a direct bridge between chapter 2 and chapter 5, and the bridge is worth making explicit as we go. The heat supplied comes from chapter 2's latent-heat formula, because the process is a phase change at constant temperature rather than a simple heating of a fixed-phase sample. The external work is a new chapter 5 quantity, computed from the volume change against a constant external pressure. The internal energy change follows directly from the first law, once the other two are in hand. We take the three parts in order.

(a) The water is vapourising at constant temperature, so there is no temperature rise at all, and the heating formulae that involve a temperature change do not apply. What applies instead is chapter 2's latent-heat formula, $Q = mL$, where m is the mass of water vapourising and L is the specific latent heat of vapourisation. Substituting:

$$Q = (1.0 \times 10^{-3} \text{ kg})(2.26 \times 10^6 \text{ J kg}^{-1}) = 2.26 \times 10^3 \text{ J} = 2260 \text{ J}$$

The sign is positive, because heat is being supplied to the water from the flame below.

(b) The water begins as a liquid with a volume of about 1cm³ (with $\rho_{\text{water}} = 1 \text{ g cm}^{-3}$, 1g of liquid water has a volume of 1cm³ at this temperature) and ends as steam with a volume of 1671cm³. The change in volume is therefore

$$\Delta V = 1671 \text{ cm}^3 - 1 \text{ cm}^3 = 1670 \text{ cm}^3 = 1.670 \times 10^{-3} \text{ m}^3$$

Throughout the expansion, the system is pushing against the surrounding atmosphere at a constant pressure of 1.013×10⁵Pa, so the work done by the expanding water on the atmosphere is:

$$W = P\Delta V = (1.013 \times 10^5 \text{ Pa})(1.670 \times 10^{-3} \text{ m}^3) = 169 \text{ J}$$

The sign is positive, because the system is expanding and therefore doing work on its surroundings.

(c) Change in internal energy. With Q and W in hand, the first law supplies ΔU directly:

$$\Delta U = Q - W = 2260 \text{ J} - 169 \text{ J} = 2091 \text{ J}$$

Making Sense of the Answer: Three numbers now sit side by side, and the relationship between them is the whole point of the example. Of the 2260J supplied to vapourise one gram of water, only 169J went into mechanical work against the atmosphere. That is about seven percent of the total. The remaining 2091J, which is roughly ninety-three percent, stayed inside the water as an increase in internal energy, even though the temperature did not rise by a single degree. The vapour molecules are at exactly the same temperature as the liquid water they came from, but they are no longer held close to each other by the intermolecular forces that existed in the liquid. Separating those molecules from each other, against their own mutual attraction, takes energy, and that energy is now locked inside the vapour as an increase in its internal energy. This is the physical content of latent heat. Most of the energy supplied during boiling goes into breaking molecular bonds, and only a small fraction goes into pushing back the atmosphere.

Think Like a Physicist: Notice what the first law has just done. It has stitched together three quantities that at first sight belong to three different chapters. The heat supplied Q came from chapter 2's latent-heat formula. The work done W came from chapter 5's picture of expansion against atmospheric pressure. And the increase in internal energy ΔU reached back to chapter 4's molecular account of what internal energy actually is, namely the kinetic energy of molecular translation, supplemented by rotational and vibrational modes, but now also including the energy stored in separating molecules against the intermolecular forces that were holding them together. The first law does not care which chapter supplied each of its three terms. It cares only that the three terms balance, which they do, to the last joule.

Four worked examples now lie behind us, and the first law has emerged from each of them with its books still balancing and its dignity intact. What has not yet been cracked open is the work term itself. W has been a single number so far, pulled from either algebra (Example 1) or a product $P\Delta V$ at fixed pressure (Example 4). Both cases are special; in a general thermodynamic process, the pressure declines to hold still while the gas expands, and W must be computed by an integral rather than a single multiplication. The integral, graphed, turns into an area; and the diagram on which that area lives is the most useful picture in this chapter. That picture is called the PV diagram, and the next section introduces it.

WORK DONE BY A GAS

The previous section ended with a promise. The work term W was going to open up, and the tool for opening it was going to be a diagram. That is now overdue. Every worked example so far has treated W as a single number, pulled either from the first law by arithmetic or from a product $P\Delta V$ at fixed pressure. Both of those are special cases. The general rule, for any process at all, is that the work done by a gas is not a number but an integral, and the integral has a picture, and the picture carries the physics. That picture is called the **PV** diagram, and the next few pages live on it.

Before the diagram, we need the integral. Before the integral, we need the piece of physics the integral integrates. That piece, in turn, needs only one picture to see clearly, and the picture is this one.

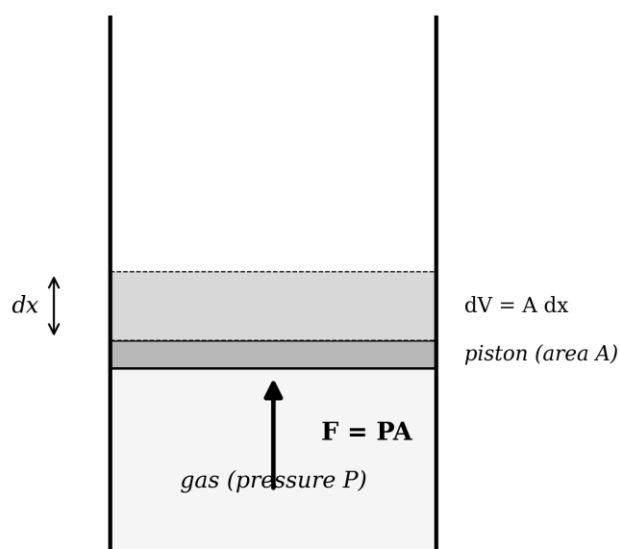


Figure: A gas trapped inside a cylinder with a movable piston of cross-sectional area A . The gas exerts an upward force $F = PA$ on the piston, where P is the gas pressure. When the piston slides outward by a small distance dx , the volume swept out (shown shaded) is $dV = A dx$, and the small quantity of work done by the gas on the piston is $dW = F dx = P dV$. This tiny identity is the foundation of everything that follows.

From force on a piston to work done by a gas

Picture the cylinder in the figure. A gas at pressure P fills the lower half; a piston of cross-sectional area A closes the top. The gas molecules, in the course of their chaotic motion from Chapter 4, are constantly colliding with the underside of the piston. Each collision transfers a tiny amount of momentum upward. Averaged over all molecules and all collisions, these transfers add up to a steady net force on the piston, directed upward, and equal, by the very definition of pressure, to the pressure multiplied by the area on which it acts:

$$F = PA$$

This is not new physics. It is the pressure formula from Chapter 4, applied to one particular wall of the container. What is new is what we do with it next.

Let the piston slide outward by a small distance dx . Nothing mysterious happens in this motion; the gas pushes, the piston moves, work is done by the pusher on the pushed. The small amount of work done by the gas on the piston in this small motion is the force multiplied by the distance:

$$dW = Fdx = PAdx$$

Now comes the observation that turns this piece of mechanics into a piece of thermodynamics. The volume swept out by the piston as it moves through dx is Adx , and this is exactly the change in the volume of the gas:

$$dV = Adx$$

Substituting $Adx = dV$ into the expression for dW gives the result of this section in a single line:

$$dW = PdV$$

The small quantity of work done by a gas, when its volume changes by a small amount dV , is the pressure multiplied by that volume change. The identity holds for a piston, for the hovering lid of a kettle, for the expanding gases in a car engine, and for any other situation in which a gas expands against some movable boundary. The boundary does not have to be a flat piston; the formula is a local statement about an infinitesimal volume change, and any shape of boundary can be broken down into many small flat pieces. For a finite change in volume, from some initial V_1 to some final V_2 , the total work done is the sum of many such small contributions. In the language of calculus, this sum is an integral:

$$W = \int_{V_1}^{V_2} PdV$$

This is the central formula of the section. It says that the **work done** by a gas during any process, whether slow or fast, whether the pressure stays fixed or changes wildly, whether the gas is ideal or real, *is the integral of the pressure with respect to the volume, taken along the path of the process*. A student who carries this formula into the rest of the chapter can derive everything else.

The integral is the area under the curve

An integral of the form $\int PdV$ has an immediate graphical meaning that is worth pausing on, because it is the reason the PV diagram is worth drawing in the first place. If we plot P on the vertical axis against V on the horizontal axis, then during any thermodynamic process the state of the gas traces out a curve on this diagram. The quantity PdV , for a thin vertical strip of width dV under the curve, is exactly the area of that strip: height P , width dV , area $P dV$. The integral, which sums all these strips from V_1 to V_2 , is therefore the total area enclosed between the curve, the V -axis, and the two vertical lines $V = V_1$ and $V = V_2$. In a single phrase:

$$W = \text{area under the PV curve}$$

This is how the PV diagram does physics. It makes the work done by the gas visible as an area on a page. Two processes whose curves enclose the same area do the same amount of work; two processes whose curves enclose different areas do different amounts of work; and the shape of the curve matters as much as its endpoints. **Foundations of Physics** taught us that the area under a force-displacement graph equals the work done, which is exactly the same story as the one told here, translated from mechanical units into thermodynamic ones.

Checking the sign against the first-law convention

The sign convention from the previous section required W positive for expansion and negative for compression. The integral gets this right automatically, provided the limits are written in the order the process actually runs. When the gas expands, the final volume V_2 exceeds the initial volume V_1 , the integration runs from a smaller number to a larger one, and since the pressure is positive throughout, the integral is positive. $W > 0$: the gas has done work on its surroundings. When the gas is compressed, V_2 is less than V_1 , the integration runs from a larger number to a smaller one, the integral comes out negative, and $W < 0$: the surroundings have done work on the gas. When the volume does not change at all, $V_2 = V_1$, the integral vanishes, and $W = 0$. That last case is the isochoric process, which gets its own subsection soon.

The convention locked at the start of the previous section survives intact. No new rules are needed; the integral and the sign convention already agree.

Work is path-dependent

Here is the observation that makes this subject interesting, and that the rest of the chapter will build on. The integral $\int PdV$ depends on the shape of the curve between V_1 and V_2 , and not merely on the endpoints. Two processes connecting the same initial state $A = (V_1, P_1)$ to the same final state $B = (V_2, P_2)$ can trace out very

different curves, enclose very different areas, and therefore deliver very different amounts of work. This is true even though both processes start at the same place and end at the same place. Work is a **path function**, not a **state function**. Its value cannot be recovered from the endpoints alone; the whole history of the process matters.

The figure below makes the point graphically. Three different processes carry the gas from the same point A in the upper left to the same point B in the lower right, each along a differently shaped curve, and each enclosing a visibly different area beneath itself. The three areas are the three values of W .

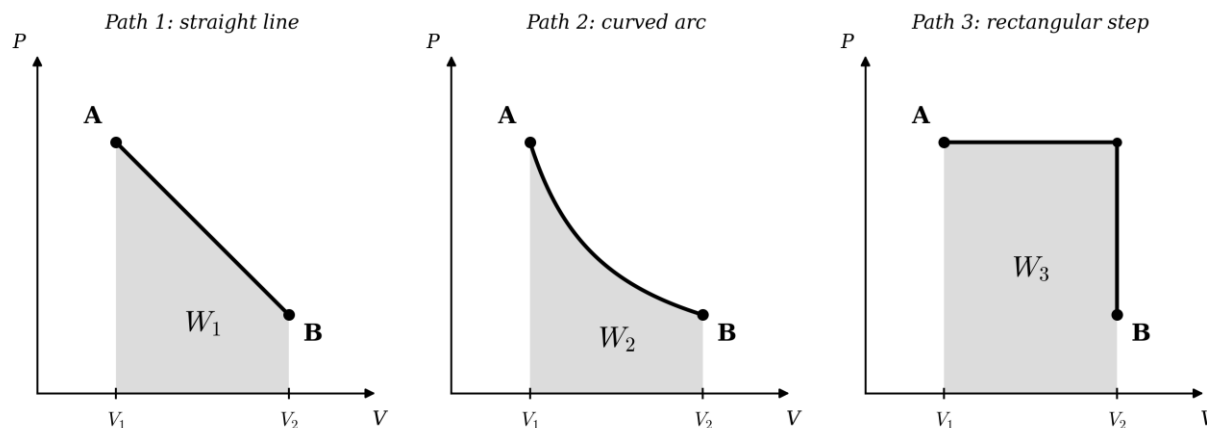


Figure: Three thermodynamic processes connecting the same initial state A to the same final state B, along three different paths. The shaded area beneath each path is the work W done by the gas during that process. The three areas are visibly unequal: W_1 (straight line), W_2 (curved arc, lying below the line), and W_3 (rectangular step, lying above the line). Same endpoints, three different answers. Work is path-dependent.

The curved arc of Path 2 lies everywhere below the straight line of Path 1, so the area under Path 2 is smaller than the area under Path 1, and $W_2 < W_1$. The rectangular step of Path 3 goes first along the top of the box at constant high pressure, then straight down at constant volume; the area under this step is the full rectangle of width $V_2 - V_1$ and height P_1 , which is everywhere above the straight line, so $W_3 > W_1$. All three processes carry the gas from the same A to the same B, yet all three do different amounts of work against the surroundings.

State functions and path functions

The path-dependence of W is worth contrasting with another quantity that behaves very differently. The state variables P , V , T , and n are, as their name suggests, functions of state: each has a definite value at every equilibrium state, and the change in any of them between two states depends only on the two states, not on how the gas got from one to the other. If the gas ends at $V = V_2$, the volume is V_2 regardless of whether the gas reached V_2 by a straight line, a curved arc, or a rectangular step. The same is true of the pressure, the temperature, and the amount of substance. It is also true of the internal energy U , which Chapter 4 showed depends only on the temperature for an ideal gas, and therefore only on the state, not the path. The change ΔU between two states is the same no matter which process connected them.

Work W does not behave this way. Neither does heat Q . Both are path functions, not state functions. There is no such thing as "the work of state B" or "the heat of state B"; there is only the work done and the heat supplied during a particular process, and both depend on the path that process took through state space. ***This is why the symbols ΔW and ΔQ are wrong on principle and should never be written:*** there is no W of state A and no W of state B to take the difference of. The correct symbols are simply W and Q , each referring to the full accumulation over the path. The first law, $Q = \Delta U + W$, has a delta **only** on U , precisely because only U is a state function of the three.

This distinction is not a nicety. It is the reason the four thermodynamic processes about to be met, namely isochoric, isobaric, isothermal, and adiabatic, give four genuinely different answers for the work done, even when their endpoints are arranged to coincide. Each process traces a different path. Each path encloses a different area. Each area is the work done. The physics is in the curves.

This is enough theory for one section. The formula $dW = PdV$, the integral $W = \int PdV$, the area-under-curve reading, and the path-dependence that results are the four pieces that every subsequent calculation uses. Four worked examples now put them to work on increasingly serious geometry: a constant-pressure slab (the

rectangle under a horizontal line), a rectangular step (two perpendicular legs), a comparison between two different paths connecting the same endpoints, and a Tanzanian football at the Miono school field.

BINDER Example 5

A gas in a cylinder with a loose, frictionless piston is heated slowly so that its pressure remains constant at $2.0 \times 10^5 \text{ Pa}$ throughout the process. Its volume expands from 3.0L to 8.0L. Calculate the work done by the gas during the expansion.

Solution

The pressure is constant throughout the process, so it comes out of the integral as a common factor, leaving only the integration over the volume change:

$$W = \int_{V_1}^{V_2} P dV = P \int_{V_1}^{V_2} dV = P(V_2 - V_1) = P\Delta V$$

Before substituting, the volumes must be converted from litres to cubic metres (the SI unit), because the pressure has been given in pascals, and the product $P\Delta V$ will only come out in joules if every factor carries SI units:

$$V_1 = 3.0\text{L} = 3.0 \times 10^{-3}\text{m}^3, \quad V_2 = 8.0\text{L} = 8.0 \times 10^{-3}\text{m}^3$$

The volume change is therefore:

$$\Delta V = V_2 - V_1 = 8.0 \times 10^{-3}\text{m}^3 - 3.0 \times 10^{-3}\text{m}^3 = 5.0 \times 10^{-3}\text{m}^3$$

Substituting into the formula for the work:

$$W = P\Delta V = (2.0 \times 10^5 \text{ Pa})(5.0 \times 10^{-3} \text{ m}^3) = 1.0 \times 10^3 \text{ J} = 1000 \text{ J}$$

The work done by the gas is 1000J.

Making Sense of the Answer: Geometrically, the PV diagram for this process is a horizontal line at $P = 2.0 \times 10^5 \text{ Pa}$, running from $V = 3.0\text{L}$ to $V = 8.0\text{L}$. The work done by the gas is the area of the rectangle beneath this line: height P , width ΔV , area $P\Delta V$. This is the simplest possible PV-diagram calculation, because the shape is a rectangle and the area is a single multiplication. All 1000J of work flowed from the gas to whatever mechanism was holding the piston in place, whether a stone on top, the atmosphere, a spring, or a thoughtful lab technician's hand.

Think Like a Physicist: Constant-pressure processes are the gentle giants of thermodynamics. Their PV curve is a horizontal line, their area is a rectangle, and their work is a single multiplication. Whenever a problem says "**the pressure remains constant**" or "**isobaric**", reach for $W = P\Delta V$ without hesitation. And always, always convert the volume to cubic metres before multiplying by a pressure in pascals. Leaving V in litres gives an answer in $\text{Pa}\cdot\text{L}$, which is smaller than a joule by a factor of 1000.

BINDER Example 6

A gas starts at state ($V_1 = 2.0\text{L}$, $P_1 = 1.0 \times 10^5 \text{ Pa}$) and undergoes a two-step process on the PV diagram. In step one, it is heated at constant volume until its pressure has risen to $P_2 = 3.0 \times 10^5 \text{ Pa}$. In step two, at this new constant pressure, its volume expands to $V_3 = 5.0\text{L}$. Calculate the total work done by the gas during the two-step process, and identify which step contributed what.

Solution

The two steps are separate integrals on the PV diagram. Step one is a vertical segment (constant volume), and step two is a horizontal segment (constant pressure). Each is computed separately and the two results are added.

Step one (isochoric heating):

The volume is constant at $V = 2.0\text{L}$ throughout the step, which means $dV = 0$ at every point along it. Therefore:

$$W_1 = \int P dV = 0$$

No volume change means no work done by the gas, no matter how much the pressure happens to rise during the step. This is a slightly counter-intuitive fact for students meeting it for the first time, so it is worth reading twice. The gas is being heated vigorously; its molecules are hitting the walls faster and harder; the pressure is tripling. And yet the gas is doing no work, because nothing is moving. Work, in the mechanical sense, requires a force to be applied through a distance, and here the distance is zero.

Step two (isobaric expansion):

The pressure is constant at $P = 3.0 \times 10^5 \text{ Pa}$ while the volume expands from 2.0L to 5.0L. Converting to SI units:

$$V_2 = 2.0\text{L} = 2.0 \times 10^{-3}\text{m}^3, \quad V_3 = 5.0\text{L} = 5.0 \times 10^{-3}\text{m}^3, \quad \Delta V = 3.0 \times 10^{-3}\text{m}^3.$$

The work done during step two is:

$$W_2 = P\Delta V = (3.0 \times 10^5 \text{ Pa})(3.0 \times 10^{-3} \text{ m}^3) = 900\text{J}$$

The total work done by the gas during the two-step process:

$$W = W_1 + W_2 = 0 + 900\text{J} = 900\text{J}$$

Making Sense of the Answer: On the PV diagram, the two-step path looks like an L lying on its side: a vertical leg going up (isochoric heating, no width, no area) joined at the top to a horizontal leg going right (isobaric expansion, a genuine rectangle). Only the horizontal leg contributes to the work, and it contributes its full rectangular area of 900J. The vertical leg, despite involving a dramatic tripling of the pressure, contributes nothing at all, because its width on the V-axis is zero. All of the work comes from the second step.

Think Like a Physicist: This example encodes two facts that apply to every multi-leg problem. First, vertical segments on a PV diagram do zero work, always, regardless of how the pressure changes along them. Second, horizontal segments do $P\Delta V$ of work, equal to the rectangular area beneath them. For any staircase path made of purely horizontal and purely vertical pieces, the total work is simply the sum of the horizontal rectangles. This observation is a fast way to estimate the work for any rectangular-step process, and it foreshadows the next few sections, where the isobaric process (all horizontal) and the isochoric process (all vertical) appear as two of the four thermodynamic workhorses.

REAL Example 7

Kipanga is inflating a football with a hand pump at the Miono school field. On his first attempt, in no hurry, he pumps slowly and pauses between strokes. On his second attempt, running late for a match against the neighbouring school, he pumps quickly and does not pause. Both times, the football ends at the same final pressure. Kipanga notices that the pump barrel gets noticeably hotter after the fast pumping than after the slow pumping. Explain why, using the first law of thermodynamics.

Solution

Take the air inside the pump barrel as the system, and consider what happens during one compression stroke. The handle pushes down; the air compresses; the pressure inside the barrel rises until the one-way valve opens and the air moves into the football. Whether Kipanga pumps slowly or quickly, the air starts at atmospheric pressure and ends at football pressure. The two endpoints agree.

What differs between the two strokes is the amount of heat the air exchanges with the barrel walls during the stroke itself. Compressing a gas always tries to heat it up (molecules get squeezed into less space, they hit the walls harder, the temperature rises). In a slow stroke, there is time for that extra thermal energy to leak out to the barrel walls and from there to the surrounding air, so the air inside the barrel stays close to room temperature throughout. In a fast stroke, there is no time for the heat to escape; it stays trapped inside the air, so the air gets genuinely hot by the end of the stroke.

Now apply the first law, $Q = \Delta U + W$, to the air in the barrel for one stroke. In the slow stroke, ΔU is small (the temperature hardly rises), Q is negative (heat leaves the air), and W is modest. In the fast stroke, ΔU is large (the air finishes much hotter), Q is nearly zero (no time to lose heat), and W is larger. By the first law, the extra work done by Kipanga's arm in the fast stroke has nowhere to go except into ΔU , so the air finishes hotter. That hot air, once it passes through the valve, dumps its heat onto the valve and the upper barrel walls. Stroke after stroke, those walls accumulate heat faster than they can shed it to the surroundings. The barrel gets warm. The slow stroke deposits very little heat per stroke, and the barrel has plenty of time between strokes to cool, so it stays near ambient.

Making Sense of the Answer: Both pumpings end with the same air in the same football at the same final pressure. The endpoints agree. But the amount of work Kipanga's arm has done on the air, and the amount of heat the air has shed to the barrel, differ between the two paths. This is path-dependence of W and Q in its most everyday form: a football pump whose temperature reports on how fast you were pumping.

Think Like a Physicist: The barrel's temperature is a readout, not a cause. Kipanga's arm was the real source of the heat. The first law is what converts his muscular work, via the intermediate step of compressed

air, into the warmth of the barrel walls. When a thermodynamic system behaves differently at different speeds, the first question is always whether the path is different, and the second is how ΔU , Q , and W are rearranging themselves along that path. The endpoints being the same does not settle anything; the path is where the physics lives.

HOT Example 8

A gas is taken from state A, where $P_A = 3.0 \times 10^5 \text{ Pa}$ and $V_A = 1.0 \times 10^{-3} \text{ m}^3$, to state B, where $P_B = 1.0 \times 10^5 \text{ Pa}$ and $V_B = 3.0 \times 10^{-3} \text{ m}^3$, along two different paths on the PV diagram. Path 1 is a straight line directly from A to B. Path 2 consists of an isobaric expansion at $P = 3.0 \times 10^5 \text{ Pa}$ from V_A to V_B , followed by an isochoric drop at $V = V_B$ from P_A down to P_B . Compute the work done by the gas along each path.

Solution

The initial and final states A and B are fixed. Only the shape of the path between them differs. We compute each path separately by finding the area beneath its PV curve. The two paths are shown in the figure below; the coordinates of A and B are given in the display units of the axes, and the shaded region in each panel is the area whose value we are about to compute.

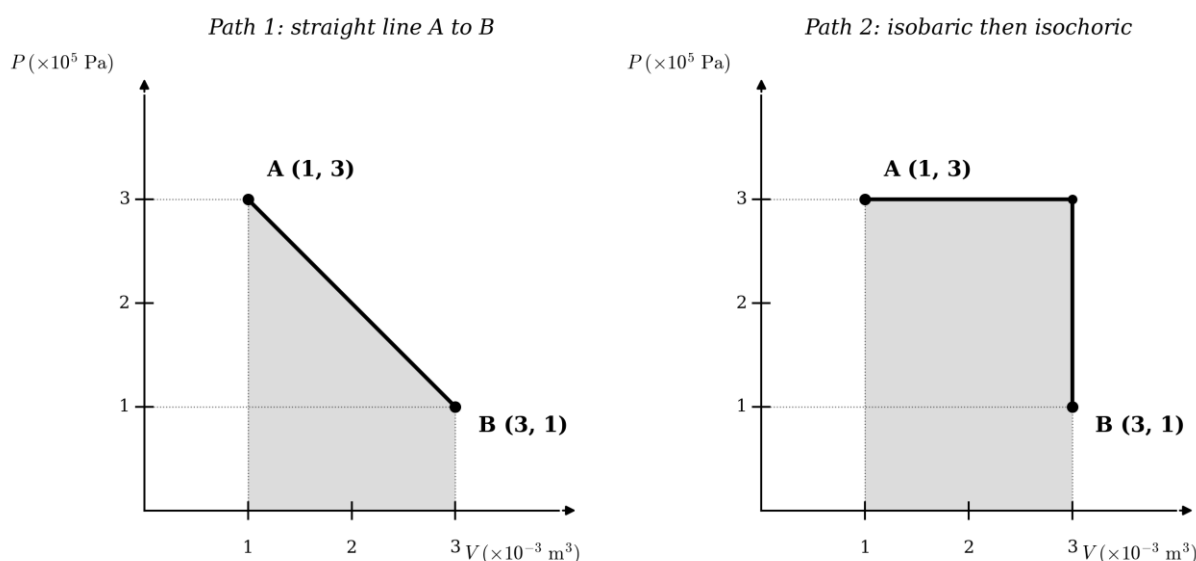


Figure: The same initial state A and final state B, connected by two different paths. The coordinate pair in brackets next to each point gives (V, P) in the display units of the axes. On the left, a straight line from A to B encloses a trapezoidal area beneath it. On the right, an isobaric expansion at the upper pressure (horizontal leg) followed by an isochoric pressure drop at V_B (vertical leg) encloses the full rectangle beneath the horizontal leg. The rectangle is visibly larger than the trapezoid, and the calculation that follows turns that visual difference into a precise number of joules.

Path 1: the straight line

This encloses a trapezoidal area, so:

$$W_1 = \frac{P_A - P_B}{2} \times (V_B - V_A) + P_B(V_B - V_A) = \frac{P_A + P_B}{2} \times (V_B - V_A)$$

Substituting:

$$W_1 = \frac{3.0 + 1.0}{2} \times 10^5 \text{ Pa} \times (3.0 - 1.0) \times 10^{-3} \text{ m}^3 = 400 \text{ J}$$

The work done along Path 1 is 400J.

Path 2: isobaric expansion followed by isochoric drop

This encloses a rectangular area, so:

$$W_2 = P_A(V_B - V_A) = (3.0 \times 10^5 \text{ Pa})(3.0 - 1.0) \times 10^{-3} \text{ m}^3 = 600 \text{ J}$$

The work done along Path 2 is 600J.

Making Sense of the Answer: Same A, same B, two different values for W . The difference is real and fully quantitative: Path 2 does 200 joules more work than Path 1. Why? Path 2 stays at the high pressure P_A throughout the volume expansion and then drops the pressure only after the volume has finished changing. It captures the full high-pressure rectangle as its work. Path 1 lets the pressure drop gradually during the

expansion, so its average pressure during the volume change is smaller, and the trapezoid beneath it encloses less area than the rectangle. The difference between the two areas is the difference between the two values of W . The physics is visible on the diagram; the diagram agrees with the arithmetic; and the arithmetic agrees with the first law, as we are about to verify.

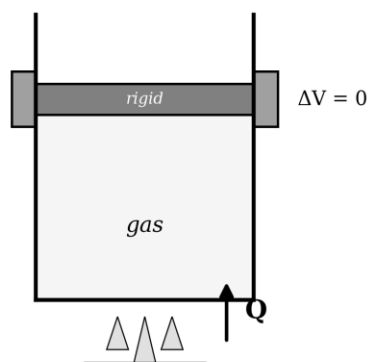
Think Like a Physicist: Here is the bonus consequence of this example, and it is the reason this problem is classified as HOT. Because the endpoints A and B are identical in the two processes, the change in internal energy ΔU must also be identical in both. U is a state function; it cares only about the endpoints. Write the first law for each path: $Q_1 = \Delta U + W_1$ and $Q_2 = \Delta U + W_2$. Subtract the two equations: $Q_2 - Q_1 = W_2 - W_1 = 200\text{J}$. The second path requires 200 joules more heat than the first path, because it does 200 joules more work, and the internal energy is the same in both cases. This is how Q and W share the path-dependence burden. Whatever extra work is done by the gas must be fed by extra heat from the surroundings, because ΔU is already pinned down by the endpoints.

Work is now in hand. The next section takes up the other half of the first law, the heat Q , and in particular how much heat is needed to raise the temperature of a gas by a given amount.

MOLAR HEAT CAPACITIES OF AN IDEAL GAS

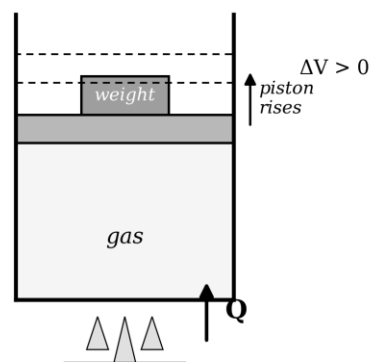
The previous section opened the work term W . This section opens the heat term Q . The question to be answered, in a single sentence, is this: *how much heat is required to raise the temperature of a gas by a given amount?* Chapter 2 answered the corresponding question for solids and liquids using a single specific heat capacity, because solids and liquids hardly expand when heated, and so the distinction between heating at constant volume and heating at constant pressure made no practical difference for them. Gases are different. A gas heated at constant volume needs less heat than the same gas heated at constant pressure, for the same temperature rise, because in the second case some of the heat is being spent on expansion work against whatever is holding the pressure fixed. Two different setups, two different answers, two different heat capacities. The picture that makes this vivid is the following.

Heating at constant V



Constant volume:
piston cannot move

Heating at constant P



Constant pressure:
weight holds P fixed

Figure: Two identical samples of the same gas, in two identical cylinders, are heated from below by identical flames delivering the same quantity of heat Q . On the left, the piston is clamped rigid, so the gas cannot expand: the volume is held constant, and all of the incoming heat goes into raising the internal energy of the gas, which raises its temperature. On the right, the piston is free to slide, with a fixed weight on top keeping the pressure constant, so as the gas heats up it expands, lifts the weight, and does work against gravity. Some of the incoming heat becomes internal energy; the rest leaves as work. For the same Q , the right-hand gas ends at a lower temperature than the left-hand gas. Or, equivalently, to reach the same temperature rise, the right-hand gas requires more heat than the left-hand one. The difference is exactly the expansion work.

The physics the figure encodes is the punchline of the whole section: a gas has not one but two heat capacities, one for each of the two arrangements. The left arrangement gives the **molar heat capacity at constant volume**, written C_v . The right arrangement gives the **molar heat capacity at constant pressure**, written C_p . Both have units of $\text{J mol}^{-1} \text{K}^{-1}$, and both refer to the heat needed per mole of gas, per degree of temperature rise. The rest of the section derives each one from the molecular picture of Chapter 4, connects them through the first law, and lands on the famous relation between them: **Mayer's equation**.

Heat capacity at constant volume

Consider the left-hand cylinder in the figure. The piston is rigid, so the gas cannot expand, meaning $\Delta V = 0$ during the heating. By the result from the previous section, $W = 0$ for a process at constant volume. The first law then simplifies:

$$Q = \Delta U + W = \Delta U + 0 = \Delta U$$

Every joule of heat that flows into the gas goes into raising its internal energy. This is a very clean setup, because the heat Q is literally the change in U , nothing lost to anything else.

Now invoke the result from Chapter 4 that expressed the internal energy of an ideal gas in terms of its temperature. For a monatomic ideal gas with n moles and absolute temperature T , the **equipartition theorem** gave:

$$U = \frac{3}{2} \times nRT$$

The change in internal energy for a small temperature change ΔT is therefore:

$$\Delta U = \frac{3}{2} \times nR\Delta T$$

Combining this with $Q = \Delta U$:

$$Q = \frac{3}{2} \times nR\Delta T \quad (\text{monatomic, constant } V).$$

The **molar heat capacity at constant volume** is defined as *the heat required per mole per degree temperature rise at constant volume*:

$$C_v = \frac{Q}{n\Delta T} \quad \text{at constant } V$$

Substituting the monatomic expression for Q :

$$C_v = \frac{\frac{3}{2} \times nR\Delta T}{n\Delta T}$$

$$C_v = \left(\frac{3}{2}\right)R \quad (\text{monatomic ideal gas}).$$

For a diatomic gas at moderate temperatures, Chapter 4 gave five active degrees of freedom (three translational plus two rotational), so $U = (5/2)nRT$ and the molar heat capacity at constant volume is:

$$C_v = \left(\frac{5}{2}\right)R \quad (\text{diatomic ideal gas}).$$

In general, for a gas with f active degrees of freedom per molecule, equipartition gives $U = (f/2)nRT$ and therefore:

$$C_v = \left(\frac{f}{2}\right)R = f \times \frac{R}{2}$$

The molar heat capacity at constant volume is, at this level of theory, a direct measurement of *how many degrees of freedom each molecule of the gas has, multiplied by the universal constant $R/2$* . This is one of the most beautiful connections in the whole of classical physics: a macroscopic, thermometer-and-stopwatch quantity like C_v counts the microscopic degrees of freedom of the molecules.

Numerically, with $R = 8.314 \text{ J mol}^{-1} \text{K}^{-1}$, the predictions are $C_v = 12.5 \text{ J mol}^{-1} \text{K}^{-1}$ for a monatomic gas and $C_v = 20.8 \text{ J mol}^{-1} \text{K}^{-1}$ for a diatomic gas at room temperature. The measured values for helium, neon, and

argon (monatomic) cluster tightly around 12.5. The measured values for nitrogen, oxygen, hydrogen, and carbon monoxide (diatomic, at room temperature) cluster tightly around 20.8. The theory is not approximately right; it is exactly right to within experimental precision.

Heat capacity at constant pressure

Now consider the right-hand cylinder in the figure. The piston is free to slide, and the weight on top keeps the pressure constant at whatever value the weight's mass and the piston's area deliver. Same gas, same temperature rise ΔT , different arrangement. The first law still holds, but this time W is not zero; the gas does expansion work against the weight and the atmosphere:

$$Q = \Delta U + W \quad (\text{constant } P)$$

The internal energy change ΔU depends only on the temperature change (for an ideal gas, U depends only on T , as Chapter 4 established), so ΔU is the same as in the constant-volume case for the same ΔT :

$$\Delta U = nC_v\Delta T$$

What about the work? Because the pressure is constant throughout, the work done by the gas is simply $W = P\Delta V$, from the previous section. For an ideal gas at constant pressure, the ideal-gas law $PV = nRT$, differentiated while holding P fixed, gives $P\Delta V = nR\Delta T$. So:

$$W = P\Delta V = nR\Delta T \quad (\text{constant } P)$$

Substituting both pieces into the first law:

$$Q = nC_v\Delta T + nR\Delta T = n(C_v + R)\Delta T$$

The **molar heat capacity at constant pressure** is defined, in parallel with the constant-volume version, as *the heat required per mole per degree at constant pressure*:

$$C_p = \frac{Q}{n\Delta T} \quad \text{at constant } P$$

Substituting the expression for Q :

$$C_p = \frac{n(C_v + R)\Delta T}{n\Delta T}$$

$$C_p = C_v + R \quad \text{or} \quad C_p - C_v = R$$

This is **Mayer's equation**, named for the German physician and physicist Julius Robert von Mayer who derived it in the 1840s, at about the same time Joule was pounding water with his paddle-wheel across the Channel. The equation is short, universal, and carries the whole payoff of the picture the section opened with. *The heat capacity at constant pressure exceeds the heat capacity at constant volume by exactly R , the universal gas constant*, no matter what gas is involved and no matter how many degrees of freedom its molecules have. The universal gas constant is precisely the expansion-work penalty per mole per kelvin for heating at constant pressure instead of constant volume.

Numerically, the predictions are:

$$C_p = \frac{3}{2}R + R = \frac{5}{2}R = 20.8 \text{ Jmol}^{-1} \text{ K}^{-1} \text{ for a monatomic gas and}$$

$$C_p = \frac{5}{2}R + R = \frac{7}{2}R = 29.1 \text{ Jmol}^{-1} \text{ K}^{-1} \text{ for a diatomic gas at room temperature.}$$

Again the experimental values agree.

The adiabatic index

The ratio of the two molar heat capacities is denoted by the Greek letter gamma and called the **adiabatic index** of the gas:

$$\gamma = \frac{C_p}{C_v}$$

For a monatomic ideal gas, $\gamma = (5/2)R / (3/2)R = 5/3 \approx 1.67$. For a diatomic ideal gas at room temperature, $\gamma = (7/2)R / (5/2)R = 7/5 = 1.40$. For polyatomic gases such as water vapour or carbon dioxide, γ is typically smaller still, around 1.33, because more degrees of freedom per molecule means a larger C_v and therefore a

ratio closer to 1. The value $\gamma = 1$ (which **no** real gas ever reaches) would mean the two heat capacities are equal, which by Mayer's equation would require $R = 0$, which the universe has declined to arrange.

The ratio γ is not an idle number. It will reappear shortly as the exponent that distinguishes an adiabatic process from an isothermal one. An isothermal process obeys $PV = \text{constant}$, and an adiabatic process obeys $PV^\gamma = \text{constant}$, where the γ is exactly this ratio. The difference between slow compression and fast compression in the hand pump of the previous section is mediated by γ , and so γ deserves its careful introduction here, before the processes that use it.

The table below summarises the predictions of the theory for the three main categories of ideal gas. It is worth learning.

Gas type	f	C_v	C_p	γ	Examples
monatomic	3	$(3/2)R$	$(5/2)R$	5/3	He, Ar
diatomic	5	$(5/2)R$	$(7/2)R$	7/5	N_2 , O_2
polyatomic	6 or more	$3R$	$4R$	4/3	H_2O , CO_2

Two rearrangements of Mayer's equation and the adiabatic index are worth deriving explicitly. Starting from:

$$\gamma = \frac{C_p}{C_v} \quad \text{and} \quad C_p = C_v + R,$$

substitute the second equation into the first to eliminate C_p :

$$\gamma = \frac{(C_v + R)}{C_v} = 1 + \frac{R}{C_v},$$

which rearranges to the first useful form:

$$C_v = \frac{R}{\gamma - 1}$$

Multiplying both sides by γ and using $C_p = \gamma C_v$ gives the second:

$$C_p = \frac{\gamma R}{\gamma - 1}$$

Both forms are handy when a problem gives γ and asks for either heat capacity, which is exactly the direction many examination problems take. They will also be the bridge by which the adiabatic equation $PV^\gamma = \text{constant}$ gets connected to the first law in the adiabatic section.

Specific heat capacities of gases

Everything derived so far in this section has been **molar**, meaning "per mole". Molar heat capacities are natural for gases because the equipartition theorem from Chapter 4 applies per molecule, and moles count molecules; so molar heat capacities come out in clean multiples of $R/2$. But many practical problems, give **specific** heat capacities instead, that is heat per kilogram per kelvin, with units of $J kg^{-1} K^{-1}$. Chapter 2 used specific heat capacities throughout, for solids and liquids. This subsection bridges the two conventions for gases, and a student must be comfortable switching between them.

The specific heat capacity at constant volume, written c_v (lower-case c, to distinguish it from the molar C_v), is defined as *the heat needed per kilogram per degree temperature rise at constant volume*. The relationship to the molar capacity follows from noting that one mole of a gas has a mass equal to its molar mass M (in $kg mol^{-1}$). Thus:

$$m = \frac{n}{M}, \text{ and } c_v = \frac{C_v}{M}$$

Analogously:

$$c_p = \frac{C_p}{M}$$

Dividing Mayer's equation by the molar mass on both sides:

$$c_p - c_v = \frac{R}{M}$$

This is Mayer's equation in specific form. It says that *the specific heat capacity at constant pressure exceeds the specific heat capacity at constant volume by R/M , which is the universal gas constant per unit mass.*

A second form follows from the ideal-gas law.

From $PV = nRT$ with $n = \frac{m}{M}$, so $PV = \frac{mRT}{M}$, which rearranges to:

$$\frac{R}{M} = \frac{PV}{mT} = \frac{P}{\left(\frac{m}{V}\right)T} = \frac{P}{\rho T}$$

Substituting this back into the specific form of Mayer's equation gives:

$$c_p - c_v = \frac{P}{\rho T}$$

This form is especially convenient when the problem gives the gas's density, pressure, and temperature, because it bypasses the molar mass entirely. Both forms say the same thing as Mayer's equation said in its molar form; all three are Mayer's equation, expressed in different units.

The adiabatic index γ is independent of which convention is used, because the molar mass M appears in both the numerator and the denominator of the ratio and cancels out:

$$\gamma = \frac{C_p}{C_v} = \frac{c_p}{c_v}$$

This universality is why γ is the quantity most worth remembering: it has the same numerical value regardless of whether the data is expressed per mole or per kilogram.

A small rule for deciding which form of Mayer's equation to reach for. If the problem gives data in moles, use the molar form $C_p - C_v = R$. If the problem gives data in kilograms or via density, use whichever of the two specific forms is more direct: R/M when M is handed to you, and $P/(\rho T)$ when density, pressure, and temperature are given. Choosing the form that matches the given data is not stylistic; it avoids converting every number through molar mass, and it avoids the arithmetic error that conversions invite.

That is all the theory. Six worked examples follow.

BINDER Example 9

Two moles of helium are heated from 27°C to 127°C . Calculate the heat required (a) if the heating is done at constant volume, and (b) if the heating is done at constant pressure. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

The amount of substance is $n = 2\text{ mol}$. The temperature change is $\Delta T = 127^\circ\text{C} - 27^\circ\text{C} = 100\text{K}$ (the size of a degree is the same on the Celsius and the kelvin scales, so no conversion is needed). Helium is monatomic, so $C_v = (3/2)R$ and $C_p = (5/2)R$.

(a) Constant volume:

$$Q = nC_v\Delta T = n \times (3/2)R \times \Delta T = (2\text{mol})(3/2)(8.314\text{Jmol}^{-1}\text{K}^{-1})(100\text{K}) = 2494\text{J}$$

(b) Constant pressure:

$$Q = nC_p\Delta T = n \times (5/2)R \times \Delta T = (2\text{mol})(5/2)(8.314\text{Jmol}^{-1}\text{K}^{-1})(100\text{K}) = 4157\text{J}$$

Making Sense of the Answer: Same gas, same amount, same temperature rise, two very different heat requirements. The constant-pressure case needs 1663J more heat than the constant-volume case, and Mayer's equation tells us where that extra heat went: into expansion work against the atmosphere, at a rate

of R per mole per kelvin, totalling $nR\Delta T = (2\text{mol})(8.314\text{Jmol}^{-1}\text{K}^{-1})(100\text{K}) = 1663\text{J}$. The arithmetic closes the loop. The extra heat in the constant-pressure case is exactly the expansion work done.

Think Like a Physicist: A common examination trap is to quote a single "heat capacity of helium" and expect it to apply regardless of conditions. There is no such thing for a gas. Every calculation involving a gas must first decide whether the process is at constant volume, at constant pressure, or neither; and only after that decision can the correct heat capacity be selected. "Heat capacity at constant volume" and "heat capacity at constant pressure" are two different numbers, separated by exactly R , and choosing between them is a physics decision, not a formula look-up.

BINDER Example 10

Compute the molar heat capacity at constant volume C_v , the molar heat capacity at constant pressure C_p , and the adiabatic index γ for (a) argon and (b) nitrogen. Express each value both in $\text{Jmol}^{-1}\text{K}^{-1}$ and as a multiple of R . Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

(a) Argon is monatomic; $f = 3$, so the formulae give:

$$C_v = \left(\frac{3}{2}\right)R = \frac{3}{2} \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) = 12.47\text{Jmol}^{-1}\text{K}^{-1}$$

$$C_p = C_v + R = \left(\frac{3}{2}\right)R + R = \left(\frac{5}{2}\right)R = \frac{5}{2} \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) = 20.79\text{Jmol}^{-1}\text{K}^{-1}$$

$$\gamma = \frac{C_p}{C_v} = \frac{\left(\frac{5}{2}\right)R}{\left(\frac{3}{2}\right)R} = \frac{5}{3} = 1.67$$

(b) Nitrogen is diatomic; $f = 5$, so:

$$C_v = \left(\frac{5}{2}\right)R = \frac{5}{2} \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) = 20.79\text{Jmol}^{-1}\text{K}^{-1}$$

$$C_p = C_v + R = \left(\frac{5}{2}\right)R + R = \left(\frac{7}{2}\right)R = \frac{7}{2} \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) = 29.10\text{Jmol}^{-1}\text{K}^{-1}$$

$$\gamma = \frac{C_p}{C_v} = \frac{\left(\frac{7}{2}\right)R}{\left(\frac{5}{2}\right)R} = \frac{7}{5} = 1.40$$

Making Sense of the Answer: A curious coincidence falls out: the constant-volume heat capacity of nitrogen, $20.79\text{Jmol}^{-1}\text{K}^{-1}$, is numerically equal to the constant-pressure heat capacity of argon. This is not accident. Going from argon ($f = 3$) to nitrogen ($f = 5$) adds two degrees of freedom worth of C_v , which is $(2/2)R = R$. And going from constant V to constant P also adds exactly R , by Mayer's equation. So $(5/2)R$ for nitrogen at constant V equals $(3/2)R + R = (5/2)R$ for argon at constant P . The gas constant R is everywhere in thermodynamics, and here it cancels two completely different physical effects against each other.

Think Like a Physicist: Memorising these six numbers is unnecessary if Mayer's equation and the $(f/2)R$ rule are understood. Argon's C_v is $(3/2)R$; everything else in the list can be generated from that plus the two rules. C_p is $C_v + R$. Nitrogen's C_v is C_v of argon plus the extra $(2/2)R$ from the two rotational modes. The adiabatic index is just the ratio. A student who has the two rules on the desk does not need to carry the table in memory; the table carries itself.

BINDER Example 11

A gas has density 1.775kgm^{-3} at 27°C and 10^5Pa , and its specific heat capacity at constant pressure is $c_p = 0.846\text{kJ kg}^{-1}\text{K}^{-1}$. Find the specific heat capacity at constant volume c_v and the adiabatic index γ of the gas.

Solution

The data is given per kilogram (a density and a specific heat capacity), so we reach for the specific form of Mayer's equation rather than the molar form.

$$c_p - c_v = \frac{P}{\rho T} \text{ or } c_v = c_p - \frac{P}{\rho T}$$

Substituting:

$$c_v = 846 \text{ J kg}^{-1} \text{ K}^{-1} - \left(\frac{10^5 \text{ Pa}}{1.775 \text{ kg m}^{-3} \times 300 \text{ K}} \right) = 658 \text{ J kg}^{-1} \text{ K}^{-1} = 0.658 \text{ kJ kg}^{-1} \text{ K}^{-1}$$

The adiabatic index is the ratio:

$$\gamma = \frac{c_p}{c_v} = \frac{0.846 \text{ kJ kg}^{-1} \text{ K}^{-1}}{0.658 \text{ kJ kg}^{-1} \text{ K}^{-1}} = 1.29$$

Making Sense of the Answer: A γ of about 1.3 is characteristic of a polyatomic gas with $f = 6$ active degrees of freedom ($\gamma = 4/3 \approx 1.33$ exactly). Carbon dioxide, sulphur dioxide, and water vapour all sit in this neighbourhood. The small departure of 1.29 from the theoretical 1.33 is a real effect (no gas is ever perfectly ideal) and tells the physicist that the molecules are starting to interact with one another at this density and pressure. For examination purposes, 1.29 is close enough to $4/3$, to classify the gas as polyatomic.

Think Like a Physicist: The first decision in every heat-capacity problem is not which equation to apply, but which form of the equation matches the given data. Molar data calls for the molar form. Kilogram data calls for the specific form. Density-pressure-temperature data calls for the $P/(\rho T)$ form. A student who learns to read the data before reaching for an equation will not waste time converting perfectly good numbers through a molar mass they did not need. The equations are all saying the same physics; the skill is matching form to data.

REAL Example 12

A small kiosk in Mbeya measures roughly $4\text{m} \times 3\text{m} \times 2.5\text{m}$. On a cold morning the air inside is at 10°C ; the owner wants to warm it to 22°C before opening for business, using a charcoal *jiko*. Take the density of air at 10°C and atmospheric pressure to be 1.25 kg m^{-3} , and its specific heat capacity at constant pressure to be $c_p = 1005 \text{ J kg}^{-1} \text{ K}^{-1}$. Calculate the heat the *jiko* must deliver to the air. Why is the constant-pressure specific heat capacity the right one to use, even though the kiosk is not a piston-and-cylinder arrangement?

Solution

The last question is where the physics is. A *kiosk* is not a sealed box. Air leaks out under the door and around the windows as it warms, so the air inside never becomes pressurised above atmospheric. Throughout the heating, the pressure inside the *kiosk* stays equal to the pressure outside, which is atmospheric. The air is being heated at constant pressure, and the correct heat capacity is therefore c_p . A piston is not needed; a leaky boundary does the same job.

With the physics settled, the arithmetic is short. The volume of the *kiosk*:

$$V = 4\text{m} \times 3\text{m} \times 2.5\text{m} = 30\text{m}^3$$

The mass of air it contains:

$$m = \rho V = (1.25 \text{ kg m}^{-3})(30\text{m}^3) = 37.5\text{kg}$$

The temperature rise:

$$\Delta T = 22^\circ\text{C} - 10^\circ\text{C} = 12\text{K}$$

The heat required at constant pressure:

$$Q = mc_p \Delta T = (37.5\text{kg})(1005 \text{ J kg}^{-1} \text{ K}^{-1})(12\text{K}) = 4.52 \times 10^5 \text{ J} (\approx 450\text{kJ}).$$

Making Sense of the Answer: 450kJ is the energy released by burning roughly 15 grams of charcoal (charcoal releases about 30MJ per kilogram when it burns). So the *jiko* needs to deliver 15 grams' worth of its fuel energy specifically to the air in the room. In practice the owner burns much more than 15 grams, because most of the *jiko*'s heat escapes through its own flue, warms the pot sitting on top, or leaks into the walls and the floor rather than going into the air itself. The 450kJ is a lower bound, not an operational estimate. But it captures the fundamental physics that controls the morning's charcoal budget.

Think Like a Physicist: The *kiosk* obeys constant pressure for the same reason an open cup of tea does and for the same reason the lungs do during a slow breath: any boundary that lets gas cross freely

automatically imposes constant pressure on the gas inside, because excess pressure just pushes more gas out until the inside and outside agree. A piston is one way to keep pressure constant, but it is not the only way, or even the common way. Most constant-pressure processes in everyday life are constant-pressure because of leaky boundaries, not because of moving pistons. c_p , not c_v , is therefore the default heat capacity for heating a gas in an ordinary room or an ordinary pot. c_v applies only when the container is deliberately sealed, which, outside of a physics laboratory, is the exception rather than the rule.

BINDER Example 13

The heat required to raise the temperature of 3.00mol of a polyatomic gas from 320K to 370K at constant pressure is 4.99kJ. Find (a) its molar heat capacity at constant pressure C_p , (b) its molar heat capacity at constant volume C_v , (c) its adiabatic index γ , and (d) the heat needed to raise 4.00mol of the same gas from 300K to 400K at constant volume. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

(a) The molar heat capacity at constant pressure:

$$C_p = \frac{Q}{n\Delta T} = \frac{4990\text{J}}{3.00\text{mol} \times 50\text{K}} = 33.3\text{Jmol}^{-1}\text{K}^{-1}$$

This is very close to $4R = 33.26\text{Jmol}^{-1}\text{K}^{-1}$, confirming the gas is polyatomic ($f = 6$).

(b) Apply Mayer's equation:

$$C_v = C_p - R = 33.3\text{Jmol}^{-1}\text{K}^{-1} - 8.314\text{Jmol}^{-1}\text{K}^{-1} = 25.0\text{Jmol}^{-1}\text{K}^{-1}$$

This is very close to $3R = 24.94\text{Jmol}^{-1}\text{K}^{-1}$, again consistent with $f = 6$.

(c) The adiabatic index:

$$\gamma = \frac{C_p}{C_v} = \frac{33.3\text{Jmol}^{-1}\text{K}^{-1}}{25.0\text{Jmol}^{-1}\text{K}^{-1}} = 1.33$$

This matches the polyatomic prediction $\gamma = 4/3$ exactly.

(d) Now apply the calculated C_v to the new setup: $n = 4.00\text{mol}$, $\Delta T = 100\text{K}$, constant volume. Since $\Delta V = 0$, all heat goes into internal energy:

$$Q = nC_v\Delta T = (4.00\text{mol})(25.0\text{Jmol}^{-1}\text{K}^{-1})(100\text{K}) = 1.00 \times 10^4\text{J} = 10.0\text{kJ}$$

Making Sense of the Answer: Four numbers fall out consistently: $C_p = 4R$, $C_v = 3R$, $\gamma = 4/3$, and a final Q of 10.0kJ for the second setup. The first three match the polyatomic predictions to within the arithmetic's precision. The fourth applies the calculated C_v to a new question, which is what heat capacities are for

Think Like a Physicist: The flow of logic in this example is the flow of logic in every heat-capacity problem: measure whichever heat capacity the experimental setup gives you (here, C_p , because the setup was at constant pressure); use Mayer's equation to convert to the other (C_v); use the ratio to find γ ; then apply either value to whatever new question arrives. Thermodynamics is economical with experimental data in a way that most of physics is not. Mayer's equation is the reason.

HOT Example 14

A container holds 0.5 moles of nitrogen gas initially at 300K and $1.0 \times 10^5\text{Pa}$. The gas is heated so that its temperature rises by 200K. Calculate the heat supplied, the work done by the gas, and the change in internal energy, in each of the following three cases, and comment on the pattern:

- The heating is done at constant volume.
- The heating is done at constant pressure.
- The heating is done at constant temperature.

Solution

Nitrogen is diatomic gas. For a diatomic ideal gas at room temperature, $C_v = (5/2)R$ and $C_p = (7/2)R$. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$ and $n = 0.5\text{mol}$ throughout.

(a) Constant volume, $\Delta T = 200\text{K}$. Since $\Delta V = 0$, no expansion work is done, and all the heat goes into internal energy:

$$W = 0$$

$$\Delta U = nC_v\Delta T = (0.5\text{mol})(5/2)(8.314\text{Jmol}^{-1}\text{K}^{-1})(200\text{K}) = 2079\text{J}$$

$$Q = \Delta U + W = 2079\text{J} + 0 = 2079\text{J}$$

(b) Constant pressure, $\Delta T = 200\text{K}$. Now the gas does expansion work, and the heat required is larger by a factor of $C_p / C_v = \gamma = 7/5 = 1.4$:

$$Q = nC_p\Delta T = (0.5\text{mol})(7/2)(8.314\text{Jmol}^{-1}\text{K}^{-1})(200\text{K}) = 2910\text{J}$$

The internal energy change is the same as in part (a), because ΔU depends only on ΔT for an ideal gas:

$$\Delta U = 2079\text{J}$$

The work done is the difference:

$$W = Q - \Delta U = 2910\text{J} - 2079\text{J} = 831\text{J}$$

As a consistency check, W should also equal $nR\Delta T$ for an ideal gas at constant pressure:

$$nR\Delta T = (0.5\text{mol})(8.314\text{Jmol}^{-1}\text{K}^{-1})(200\text{K}) = 831\text{J} \checkmark$$

(c) Constant temperature, $\Delta T = 0$. The internal energy of an ideal gas depends only on T , so if T does not change, neither does U :

$$\Delta U = 0$$

The first law then gives $Q = W$: whatever heat is supplied must emerge as work, and vice versa. Since $\Delta T = 0$, the formula $Q = nC_v\Delta T$ or $Q = nC_p\Delta T$ would return zero, but these formulae assume constant V or constant P , neither of which applies here. At constant temperature, both V and P must change together (along the hyperbola $PV = nRT$). So Q and W are not zero; they are equal to each other, and their value depends on how much the gas expanded or was compressed.

The calculation cannot be completed without additional information, and this is the point of including part (c): to show that the constant-volume and constant-pressure formulae are not general-purpose, and that the isothermal process needs its own treatment, coming in a later section.

Making Sense of the Answer: *The three cases form a pattern worth keeping in mind. At constant volume, all of Q goes into ΔU ; W is zero. At constant pressure, some of Q goes into ΔU and the rest into W ; the split is determined by Mayer's equation. At constant temperature, none of Q goes into ΔU ; all of it goes into W . These are three of the four cornerstone processes of thermodynamics, and each one pins down a different one of the three thermodynamic variables (V , P , T) while the other two are allowed to adjust.*

Think Like a Physicist: *A gas can be heated along many different paths, and the heat required changes with the path even when the temperature change is the same. The gas does not care what label a physicist attaches to the process; it just responds to whichever of its state variables are being held fixed. The skill is to identify which variable is being held fixed in a given experimental setup, and then select the formula that corresponds to that constraint. At constant V , use C_v ; at constant P , use C_p ; at constant T , neither formula applies and the isothermal tools must be brought in.*

The two heat capacities now sit on the desk, connected by Mayer's equation, together with their ratio γ that will matter greatly for the adiabatic process. What remains is to put these tools to work on the four cornerstone processes of thermodynamics, taken one at a time. The first of these, the simplest by far, pins the volume down and lets everything else adjust. It is called the isochoric process, and it is the subject of the next section.

THE ISOCHORIC PROCESS (CONSTANT VOLUME)

Four thermodynamic processes now arrive, each with its own section. Each is defined by pinning one of the state variables down and letting the others adjust. The simplest of the four, and the first to meet, pins the volume. It is called the **isochoric process**, from the Greek **isos** (*equal*) and **choros** (*space*), and the picture that introduces it is probably already sitting on the reader's kitchen shelf.

Imagine a pressure cooker on a *jiko* in a kitchen. The lid is screwed firmly down; the safety valve is closed; the walls of the cooker are thick and stiff. Inside, a gas, perhaps the steam and air above a pot of beans being cooked. Below, the flame. As the flame heats the gas inside, something a little surprising happens, or fails to happen. The walls of the cooker do not move. The volume of the gas inside does not change. Every molecule of the gas is now bouncing between walls that stay exactly where they were, at a distance it knows

intimately. All that changes is how hard the molecules bounce. They bounce harder, because they are hotter. The pressure inside rises, but the volume does not.

This is the whole of the isochoric process, in one sentence. The gas inside is heated, but nothing moves. The picture on a PV diagram is therefore a single vertical line.

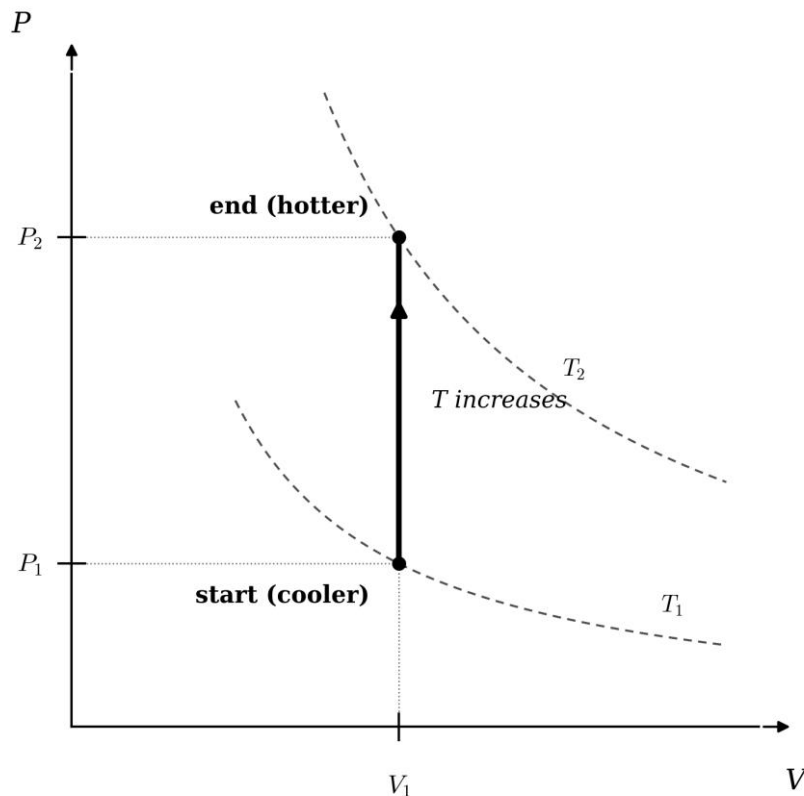


Figure: The isochoric process on a PV diagram. The volume is held fixed at V_1 throughout the process, so the state of the gas moves only in the vertical direction, from the starting pressure P_1 (cooler) to the ending pressure P_2 (hotter). The two dashed curves are isotherms, representing the hyperbola $PV = nRT$ at two different temperatures: T_1 through the lower endpoint and T_2 through the upper endpoint, with $T_2 > T_1$. The isochoric process carries the gas straight up from one isotherm to the other. No left or right motion is possible when the volume is fixed.

The first law of thermodynamics, applied at constant volume

The defining feature of the isochoric process is $\Delta V = 0$. From the result of an earlier section, work done by a gas at constant volume is

$$W = \int PdV = 0 \quad (\text{since } dV = 0 \text{ at every point along the path}).$$

No volume change means no expansion against anything, and therefore no work. With $W = 0$, the first law simplifies dramatically:

$$Q = \Delta U + W = \Delta U + 0 = \Delta U$$

Every joule of heat supplied to the gas, in an isochoric process, goes into raising its internal energy. There is no other destination available to the heat, because the only two destinations are internal energy and work, and the work channel is shut. This is the isochoric process's chief feature, and the reason it is the simplest of the four processes: it has one channel where the others have two.

Using the Chapter 4 expression $U = (f/2)nRT$, the change in internal energy for a temperature change ΔT is:

$$\Delta U = \left(\frac{f}{2}\right)nR\Delta T = nC_V\Delta T$$

So the isochoric heat supply takes the convenient form:

$$Q = \Delta U = nC_V\Delta T \quad (\text{isochoric})$$

This is the first key result of the section, and it is what the student reaches for any time an isochoric calculation needs a heat.

How pressure and temperature relate during isochoric heating

A second result follows immediately from the ideal-gas law. For a fixed amount n of an ideal gas, $PV = nRT$ at any moment during the process. If V is held fixed throughout, then n , R , and V are all constants, and the equation rearranges to:

$$P = \left(\frac{nR}{V}\right) \times T \text{ or } P = \text{constant} \times T,$$

which says P is proportional to T during the process. Equivalently, the ratio of P to T stays the same at every moment:

$$\frac{P}{T} = \text{constant} \quad (\text{isochoric})$$

Alternatively, comparing the initial and final states of the process:

$$\frac{P_1}{T_1} = \frac{P_2}{T_2}$$

This is the isochoric form of the gas law, often called **Gay-Lussac's law** after the French chemist Joseph Louis Gay-Lussac, who published it about 220 years ago. It is a direct consequence of the ideal-gas law applied at fixed volume, and it is the relationship that any isochoric problem involving pressures and temperatures together will use.

The isochoric fingerprint

Three results together identify the isochoric process and make every calculation involving it a one-step exercise in substitution. They are worth memorising as a block:

W = 0. No work, because no volume change.

Q = ΔU = nC_vΔT. All heat goes into internal energy.

P/T = constant. Pressure and absolute temperature rise together in fixed proportion.

Any problem involving a **rigid container**, a **sealed bottle**, a **pressure cooker** with the valve **shut**, a **gas cylinder** being heated in the sun, or any other situation in which a gas is heated or cooled without changing its volume can be read off these three results directly. The worked examples in this section walk through four such situations in sequence.

Nothing moves, everything rises, four examples now follow.

BINDER Example 15

A rigid container holds 0.32g of oxygen gas. The oxygen is heated from 25°C to 35°C. Given that the molar heat capacity at constant volume of oxygen is $C_v(\text{O}_2) = 20\text{Jmol}^{-1}\text{K}^{-1}$, and the molar mass of oxygen is $M(\text{O}_2) = 32\text{gmol}^{-1}$, find the heat supplied to the oxygen.

Solution

Convert the mass of oxygen to moles:

$$n = \frac{m}{M} = \frac{0.32\text{g}}{32\text{gmol}^{-1}} = 0.010\text{mol}$$

The temperature change is:

$$\Delta T = 35^\circ\text{C} - 25^\circ\text{C} = 10\text{K}$$

Since the container is rigid ($\Delta V = 0$), the isochoric heat formula applies:

$$Q = nC_v\Delta T = (0.010\text{mol})(20\text{Jmol}^{-1}\text{K}^{-1})(10\text{K}) = 2.0\text{J}$$

Making Sense of the Answer: Two joules is a tiny amount of heat, which should not be a surprise: 0.32 grams is a tiny amount of oxygen (about the mass of a grain of rice), and ten degrees is a modest temperature rise.

Think Like a Physicist: The problem specified a "rigid container" deliberately. Without that phrase, the student would not know whether to reach for C_v or C_p , and the answer would be different by a factor of 7/5 for a diatomic gas like oxygen. Rigid means the volume is fixed, which means the process is isochoric, which means C_v is correct. Every word in a thermodynamics problem statement is doing work; "**rigid**",

"insulated", "open", "sealed", and "at constant pressure" are the code words that identify which process is being described.

BINDER Example 16

An insulated rigid container holds 2mol of helium gas. A heater inside the container supplies 60J of heat to the helium. Find the rise in temperature of the gas. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

Helium is monatomic, so $C_v = (3/2)R = 12.47\text{Jmol}^{-1}\text{K}^{-1}$. The container is rigid, so the process is isochoric, and C_v is the correct heat capacity. Apply the isochoric heat formula $Q = nC_v\Delta T$ and solve for ΔT :

$$\Delta T = \frac{Q}{nC_v} = \frac{60\text{J}}{2\text{mol} \times 12.47\text{Jmol}^{-1}\text{K}^{-1}} = 2.4\text{K}$$

The temperature rise is 2.4K.

Making Sense of the Answer: The temperature rise is small, about two-and-a-half degrees, because sixty joules is not much heat and two moles of helium is a fair-sized sample. The phrase "insulated container" in the problem statement was not a tripwire; it simply confirmed that all the heat supplied by the heater went into the gas rather than leaking out through the walls. If the container had not been insulated, most of the sixty joules would have leaked away as heat to the surroundings, and the gas inside would barely have warmed at all. The insulation ensures the heat accounting closes cleanly: all 60J supplied, all 60J ending up in ΔU of the gas.

Think Like a Physicist: This is the inverse of the previous example: instead of finding Q given ΔT , find ΔT given Q . The same formula handles both directions. Whenever a heater is embedded in an insulated rigid container, the student should recognise it as the cleanest possible isochoric experiment: every joule delivered by the heater is a joule gained by the gas, and the temperature rise follows from C_v alone. Real laboratory measurements of C_v are made in exactly this kind of apparatus, called a bomb calorimeter, and the formula above is the equation used at the end of the measurement.

REAL Example 17

A pressure cooker on a jiko contains air and steam at 20°C and 1.0atm when it is first sealed. The safety valve is set to open if the internal pressure reaches 2.5atm. Assuming the gas inside is an ideal gas, and taking the cooker's volume to be constant while the valve is still closed, find the temperature of the gas at the moment the safety valve opens.

Solution

While the valve is closed, the volume of the gas inside the cooker is fixed. The process is therefore isochoric, and the relation between pressure and absolute temperature is Gay-Lussac's law:

$$\frac{P_1}{T_1} = \frac{P_2}{T_2}$$

Temperatures must be in kelvin for this to work, because the ideal-gas law requires absolute temperature.

$$T_1 = 20^\circ\text{C} + 273 = 293\text{K}$$

The pressures can be left in atmospheres, because they appear as a ratio and the units cancel. Rearrange Gay-Lussac's law for T_2 :

$$T_2 = T_1 \times \frac{P_2}{P_1} = 293\text{K} \times \frac{2.5\text{atm}}{1.0\text{atm}} = 732.5\text{K} = 459.5^\circ\text{C}$$

The temperature of the gas is 459.5°C.

Making Sense of the Answer: Four hundred and sixty degrees Celsius is hot. A cook heating a well-sealed pressure cooker on a jiko for long enough would indeed push the internal temperature into this range, which is precisely why pressure cookers have safety valves set to open well before such temperatures are reached. The student should be aware that real pressure cookers do not reach 459°C: long before that temperature is approached, either the food finishes cooking and the cook turns off the flame, or the safety valve opens and releases steam. The calculation shows what would happen in the worst case of a jammed valve, and is therefore a calculation safety engineers actually perform.

Think Like a Physicist: The isochoric step is the first half of the pressure cooker's story. The second half, which starts the instant the valve opens, is no longer isochoric: gas leaves through the valve, the volume of contained gas decreases, and the pressure stays pinned at 2.5atm by the valve mechanism. That second phase is isobaric, and it is the subject of the next section. Real devices rarely stay in a single thermodynamic

process from start to finish; they typically walk through a sequence of processes, and the skilled student learns to recognise where one process ends and the next begins. The moment the valve opens in this problem is exactly such a transition.

HOT Example 18

A rigid sealed pressure cooker contains 2.0mol of air at 20°C and 1.0atm. A heating element inside the cooker supplies energy at a steady rate of 500Js⁻¹. The safety valve is set to open when the internal pressure reaches 2.5atm. Treating the air as a diatomic ideal gas, find:

- the time elapsed between the heater switching on and the moment the safety valve opens,
- the total heat supplied by the heater up to that moment, and
- the total change in internal energy of the gas up to that moment.
- Comment on the relationship between the answers to (b) and (c) in the light of the first law of thermodynamics.

Solution

Since air is diatomic, the molar heat capacity at constant volume is:

$$C_v = \left(\frac{5}{2}\right)R = \frac{5}{2} \times (8.314\text{Jmol}^{-1}\text{K}^{-1}) = 20.79\text{Jmol}^{-1}\text{K}^{-1}$$

Initial temperature in kelvin:

$$T_1 = 20^\circ\text{C} + 273 = 293\text{K}.$$

Final temperature from Gay-Lussac's law:

$$T_2 = T_1 \times \frac{P_2}{P_1} = 293\text{K} \times \frac{2.5\text{atm}}{1.0\text{atm}} = 732.5\text{K}$$

Temperature change:

$$\Delta T = T_2 - T_1 = 732.5\text{K} - 293\text{K} = 439.5\text{K}$$

The heat required to produce this temperature change in the isochoric cooker:

$$Q = nC_v\Delta T = (2.0\text{mol})(5/2)(8.314\text{Jmol}^{-1}\text{K}^{-1})(439.5\text{K}) = 18300\text{J} = 18.3\text{kJ}$$

- (a) Time for the heater to deliver this much energy at 500Js⁻¹:

$$t = \frac{Q}{\text{rate}} = \frac{18300\text{J}}{500\text{Js}^{-1}} = 37\text{s}$$

- (b) The total heat supplied by the heater by that moment is Q itself:

$$Q = 18.3\text{kJ}$$

- (c) The total change in internal energy is what the first law delivers after the work term has been evaluated. At constant volume, $W = 0$, so:

$$\Delta U = Q - W = 18.3\text{kJ} - 0 = 18.3\text{kJ}$$

- (d) The two answers are numerically equal, both 18.3kJ. This is because the cooker is rigid, the gas cannot expand, so the work channel in the first law is closed: $W = 0$. With no work to account for, the first law $Q = \Delta U + W$ collapses to $Q = \Delta U$, and every joule of heat delivered by the heater ends up, with nothing lost to anything else, as an increase in the internal energy of the gas. This equality between Q and ΔU is the defining feature of every isochoric process.

Making Sense of the Answer: The three numerical answers tell a consistent story. A 500-watt heater running for 37 seconds delivers 18.3kJ of heat, and that heat raises 2.0 moles of air from 293K to 732.5K in a rigid cooker whose volume holds fixed throughout. The pressure rises from 1.0 to 2.5 atmospheres in exact proportion to the temperature, as Gay-Lussac's law requires for an isochoric process. Every number checks against every other number; nothing is left over, and nothing is missing.

Think Like a Physicist: Rate problems appear throughout thermodynamics whenever the question asks "how long?" rather than "how much?" The recipe is always the same: compute the total energy required for the temperature change using the appropriate heat capacity, then divide by the supply rate to get the time. Heat capacity does the physics; division does the timing.

That is the first of the four thermodynamic processes in hand. The volume was held fixed; the work channel was closed; Q and ΔU were equal. The next process in the sequence keeps the gas's behaviour a little livelier: the volume is allowed to change, but the pressure is now the variable that is held fixed. This is the isobaric process, and it is the subject of the next section.

THE ISOBARIC PROCESS (CONSTANT PRESSURE)

In the previous section the volume was held fixed. Now the pressure is held fixed instead. The second of the four thermodynamic processes is called the **isobaric process**, from the Greek *isos* (equal) and **baros** (weight, pressure), because the word originally referred to a gas heated with a fixed weight pressing on its piston. The weight imposes a fixed pressure on the gas below it; the pressure stays at that fixed value throughout the heating; the volume, now free to change, adjusts upwards to accommodate the expanding gas.

The proper imagination of this process is exactly the one the word describes: a vertical cylinder, capped above by a frictionless piston, loaded on top with a stone of exactly the right mass to keep the pressure inside the cylinder at atmospheric. A flame below heats the gas. The gas expands; the piston rises; the stone goes up with it. Nothing about the pressure changes: the stone is still the same stone, the atmosphere above is still the same atmosphere, and the gas inside is still at atmospheric pressure plus the small extra pressure the stone contributes. The volume changes; the temperature changes; the pressure does not.

An even more familiar picture you can form in your mind, and one already met earlier in this chapter, is an ordinary room. The Mbeya kiosk of an earlier worked example was heated at constant pressure precisely because its boundary was leaky: air could exit under the door and around the windows as it warmed and expanded. The kiosk never built up any excess pressure because the excess air simply left. A classroom, a kitchen, a bedroom on a cold morning in the southern highlands, all behave as leaky isobaric systems whenever the air inside is warmed. The piston-and-stone arrangement and the leaky room are two manifestations of the same physics: pressure pinned, volume free, and the first law of thermodynamics now managing two channels at once.

The picture on a PV diagram is a horizontal line, and the area beneath it is the work done.

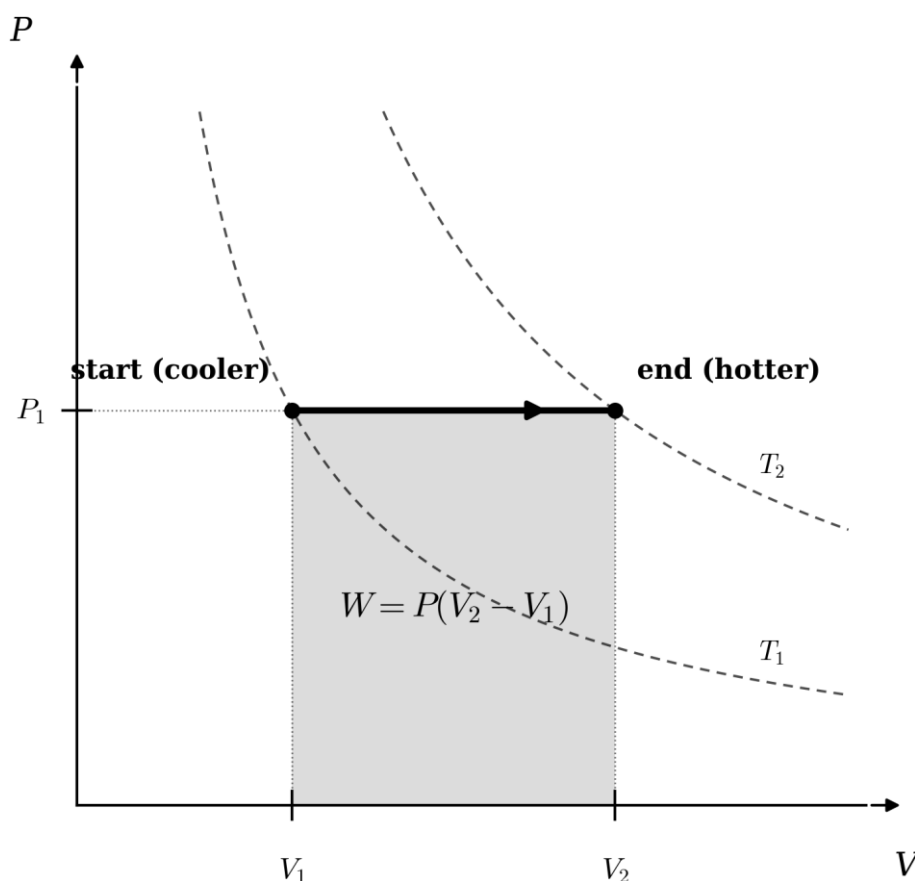


Figure: The isobaric process on a PV diagram. The pressure is held fixed at P_1 throughout, so the state of the gas moves horizontally from the starting volume V_1 (cooler, on the lower isotherm T_1) to the final volume V_2 (hotter, on the upper isotherm T_2). The shaded rectangle below the horizontal line represents the work done by the gas during the expansion: $W = P(V_2 - V_1)$. The dashed curves are isotherms passing through

the two endpoints, showing that as the gas moves to the right along the horizontal line, it is moving onto progressively hotter isotherms.

The first law at constant pressure

The defining feature of the isobaric process is that the pressure is constant. From the result of an earlier section, work done by a gas at constant pressure is a single multiplication rather than an integral:

$$W = P(V_2 - V_1) = P\Delta V$$

This is the most direct isobaric result. But there is a more useful form, available because the gas is ideal. Start from the ideal-gas law at both states:

$$PV_1 = nRT_1, \quad PV_2 = nRT_2$$

Subtract the first equation from the second. Since P , n , R are all constant, they come out as common factors:

$$P(V_2 - V_1) = nR(T_2 - T_1) = nR\Delta T$$

So the isobaric work done has the clean alternative form:

$$W = P\Delta V = nR\Delta T \quad (\text{isobaric})$$

The right-hand form is often more convenient than the left, because it uses the temperature change directly and bypasses the volumes. Whichever form is faster in a given problem, it helps to know that both are available.

With the work term in hand, the first law gives the heat and the change in internal energy in parallel. The heat required at constant pressure is

$$Q = nC_p\Delta T \quad (\text{isobaric}),$$

which was already the defining relation for C_p in the molar heat capacities section. The change in internal energy, for any process of an ideal gas, depends only on the temperature change:

$$\Delta U = nC_v\Delta T \quad (\text{any process of an ideal gas}).$$

This last formula is worth pausing on, because it looks surprising in an isobaric context. The gas **did** expand; the volume **did** change; and yet ΔU still uses C_v , the heat capacity at constant volume. The reason is that the internal energy of an ideal gas depends only on temperature, as the molecular derivation in Chapter 4 established. The C_v in $\Delta U = nC_v\Delta T$ is not saying anything about constant volume; it is using the universal proportionality between U and T for an ideal gas. The proportionality constant happens to equal C_v , but that is a property of the gas, not a condition on the process.

With Q , ΔU , and W all expressed in terms of ΔT , the first law $Q = \Delta U + W$ is simply Mayer's equation in disguise:

$$nC_p\Delta T = nC_v\Delta T + nR\Delta T,$$

which divides through by $n\Delta T$ to give $C_p - C_v = R$, the Mayer relation already familiar. This is a useful self-check: any isobaric calculation that tracks Q , ΔU , and W separately must satisfy this relation at the end.

How volume and temperature relate during isobaric heating

The ideal-gas law $PV = nRT$, applied at fixed pressure, has the same useful rearrangement that Gay-Lussac's law had at fixed volume. With P , n , R all constant, the equation reads:

$$V = \frac{nR}{P} \times T \text{ or } V = \text{constant} \times T,$$

which says V is proportional to T during the process. Equivalently:

$$\frac{V}{T} = \text{constant} \quad (\text{isobaric}),$$

or, comparing initial and final states,

$$\frac{V_1}{T_1} = \frac{V_2}{T_2}$$

This is known as **Charles's law**, after the French physicist Jacques Charles, who established it about 240 years ago through careful measurements on gases in open tubes. The law predates Gay-Lussac's by some twenty years, and historically it was Charles's observations that planted the idea of an absolute zero of temperature: extrapolating the V-versus-T line backwards, one finds that V would reach zero at around -273°C , which is no accident.

The isobaric fingerprint

Four results together identify the isobaric process. Any problem whose setup includes the words "**constant pressure**", "**isobaric**", "**open pot**", "**leaky room**", or a **piston loaded with a constant weight** can be solved by reaching for these directly:

$W = P\Delta V = nR\Delta T$. Work done by the gas, equal to the rectangular area under the horizontal process line.

$Q = nC_p\Delta T$. Heat supplied, using the constant-pressure molar heat capacity.

$\Delta U = nC_v\Delta T$. Change in internal energy, using the constant-volume molar heat capacity even though this is a constant-pressure process, because U depends only on T for an ideal gas.

$V/T = \text{constant}$. Volume and absolute temperature rise together in fixed proportion (Charles's law).

The piston rises, the stone rises, four examples now follow.

BINDER Example 19

$2 \times 10^{-2} \text{ kg}$ of nitrogen gas is heated at constant pressure (1.0 atm) so that its temperature rises by 45°C . Given $M(\text{N}_2) = 28 \times 10^{-3} \text{ kg mol}^{-1}$ and $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$, find the heat supplied.

Solution

The amount of substance in moles:

$$n = \frac{m}{M} = \frac{2 \times 10^{-2} \text{ kg}}{28 \times 10^{-3} \text{ kg mol}^{-1}} = 0.714 \text{ mol.}$$

Nitrogen is diatomic, so its molar heat capacity at constant pressure is:

$$C_p = \frac{7}{2}R = \frac{7}{2} \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1} = 29.10 \text{ J mol}^{-1} \text{ K}^{-1}$$

The temperature rise is $\Delta T = 45 \text{ K}$ (the size of a kelvin equals the size of a degree Celsius). Applying the isobaric heat formula:

$$Q = nC_p\Delta T = (0.714 \text{ mol})(29.10 \text{ J mol}^{-1} \text{ K}^{-1})(45 \text{ K}) = 935 \text{ J}$$

Making Sense of the Answer: *Nine hundred and thirty-five joules is a moderate amount of heat, equivalent to running a 100-watt bulb for about nine seconds, or burning roughly 0.03 grams of charcoal. Twenty grams of nitrogen gas warming by forty-five degrees is a small but not trivial heating job, and the 935 J figure is appropriately sized. Had the process been carried out at constant volume instead, with $C_v = (5/2)R = 20.79 \text{ J mol}^{-1} \text{ K}^{-1}$ in place of C_p , the heat needed would have been about 668 J; some 267 J less, exactly equal to $nR\Delta T$, which is the expansion work done by the gas in the constant-pressure case.*

Think Like a Physicist: *The phrase "heated at constant pressure" in an examination problem is a code word for isobaric, which in turn is a code word for "reach for C_p ". This code-word recognition is what separates a confident thermodynamics student from a confused one. Whenever "constant pressure", "atmospheric pressure", "open container", or "unloaded piston" appears in a problem, the same mental switch must flip: isobaric, so use C_p .*

BINDER Example 20

A gas at 27°C and 10^5 Pa has an initial volume of 0.04 m^3 . It is heated at constant pressure until its volume becomes 0.05 m^3 . The sample contains 48 g of gas with molar mass 28 g mol^{-1} . Find (a) the work done by the gas during the expansion, and (b) the final temperature of the gas.

Solution

(a) Work done by the gas at constant pressure:

$$W = P\Delta V = P(V_2 - V_1) = (10^5 \text{ Pa})(0.05 \text{ m}^3 - 0.04 \text{ m}^3) = 1000 \text{ J} = 1.0 \text{ kJ}$$

(b) For the final temperature, apply Charles's law:

$$\frac{V_1}{T_1} = \frac{V_2}{T_2}$$

Rearranging for T_2 :

$$T_2 = T_1 \times \frac{V_2}{V_1}$$

With $T_1 = 27^\circ\text{C} + 273 = 300\text{K}$:

$$T_2 = 300 \text{ K} \times \frac{0.05 \text{ m}^3}{0.04 \text{ m}^3} = 300 \text{ K} \times 1.25 = 375 \text{ K} = 102^\circ\text{C}$$

Making Sense of the Answer: A twenty-five-percent increase in volume produces a twenty-five-percent increase in absolute temperature, exactly as Charles's law requires for an isobaric process. The gas has gone from just below room temperature to just above the boiling point of water. Note that the problem provided more data than the two questions actually needed: the mass 48g and the molar mass 28gmol^{-1} were not used in either part. This is common in physics problems and deliberate: real experimental setups give the physicist more measurements than any single calculation requires, and part of the skill is recognising which numbers matter for the question at hand. The moles n could have been computed from the given data and used in the alternative form $W = nR\Delta T$; both routes give the same answer.

Think Like a Physicist: The two formulae for isobaric work, $W = P\Delta V$ and $W = nR\Delta T$, are both in hand from the derivation above. A careful student should notice that either form would work for part (a) of this problem. The first needs P and ΔV (both given directly). The second needs n (computed from m and M) and ΔT (which is what part (b) is asking for, so using this form would have created a circular dependency). The first form wins on efficiency. Whenever both forms of a formula are available, checking which data the problem hands you saves time: pick the form that takes raw inputs, not the form that needs a downstream result.

REAL Example 21

A classroom with a radiator running for 10 minutes feels noticeably warmer at the end than at the start. The air in the room is at atmospheric pressure throughout, because the door, the windows, and the gaps around them are not airtight. Use the isobaric model to explain what happens to the volume of air as its temperature rises, where any "extra" air ends up, and why a well-sealed classroom would feel warmer than a leaky one for the same heat supplied by the radiator.

Solution

Take the air inside the classroom as the thermodynamic system. Because the door and windows are not airtight, any rise in internal pressure inside the room would immediately push air out through the gaps, so the pressure inside the room stays matched to atmospheric pressure throughout the heating. The process is therefore isobaric, and the relevant law is Charles's law:

$$\frac{V}{T} = \text{constant}$$

As the air's absolute temperature rises, its volume must rise in fixed proportion. But the classroom's physical dimensions, namely its walls, its floor, and its ceiling, do not move. The air cannot expand into extra space inside the room, because no extra space appears. Instead, air at the edges is pushed out through the leaks.

The air that remains inside is not the same air that was there at the start. It has been warmed by the radiator, some of its molecules have left through the leaks, and its density is correspondingly lower. Every cubic metre of classroom air now contains fewer molecules than it did at the start, because the same room volume now contains the same total number of molecules minus however many have leaked out. This is what warm air means: air at the same pressure but lower density than cold air.

A well-sealed classroom heats faster than a leaky one because all the radiator's heat stays in the room rather than walking out through the gaps. In a leaky room, each air molecule that escapes carries away a share of the internal energy the radiator has been supplying, so the energy that actually remains to warm the air inside is only a fraction of what the radiator delivered. In a sealed room, no molecules escape, and every joule the

radiator delivers to the air stays as internal energy of that same air. For the same radiator running for the same time, the sealed room reaches a higher final temperature.

Making Sense of the Answer: Three physical effects combine in the right way. Charles's law says volume rises with temperature at constant pressure. Fixed room dimensions say the volume of the enclosure cannot change. Leaky boundaries reconcile the two by letting the "extra" air exit through the gaps. The result is that a room at constant pressure and fixed geometry responds to heating by expelling air, not by expanding. This is not a defect of the isobaric model; it is the isobaric model applied to a real enclosed space whose boundaries are porous.

Think Like a Physicist: The principle that a leaky enclosure behaves isobarically is one of the most useful pieces of thermodynamic intuition a student can carry out of this chapter. It applies to a heated room, a heated lecture hall, a heated daladala interior with the windows open, a cooking pot with a loose lid, and the alveoli of the human lungs during a slow breath. A tightly sealed enclosure, by contrast, behaves isochorically: volume is fixed, no gas exits, and the pressure rises with temperature. The difference between the two extremes (and every case in between) is nothing more subtle than whether the boundary lets gas pass. Tight boundaries trap gas and raise pressure; leaky boundaries release gas and hold pressure.

HOT Example 22

A cylinder contains 0.250 mol of carbon dioxide at 17°C and 1.00 atm. The gas is heated at constant pressure until its temperature reaches 127°C. Find (a) the work done by the gas, (b) the change in internal energy of the gas, (c) the heat supplied to the gas. (d) How much work would have been done if the constant pressure had been 0.50 atm instead, with the same initial temperature and same final temperature? $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Solution

$$T_1 = 17^\circ\text{C} + 273 = 290\text{K}, \quad T_2 = 127^\circ\text{C} + 273 = 400\text{K}, \quad \Delta T = T_2 - T_1 = 400\text{K} - 290\text{K} = 110\text{K}$$

Since carbon dioxide is polyatomic, $\gamma = 1.33$

Calculating the two molar heat capacities:

$$C_v = \frac{R}{\gamma - 1} = \frac{8.314 \text{ J mol}^{-1} \text{ K}^{-1}}{1.33 - 1} = 25.2 \text{ J mol}^{-1} \text{ K}^{-1}$$

$$C_p = C_v + R = 25.2 \text{ J mol}^{-1} \text{ K}^{-1} + 8.314 \text{ J mol}^{-1} \text{ K}^{-1} = 33.5 \text{ J mol}^{-1} \text{ K}^{-1}$$

(a) Work done by the gas at constant pressure:

$$W = nR\Delta T = (0.250 \text{ mol})(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(110 \text{ K}) = 229 \text{ J}$$

(b) Change in internal energy of the gas (depends only on ΔT for an ideal gas):

$$\Delta U = nC_v\Delta T = (0.250 \text{ mol})(25.2 \text{ J mol}^{-1} \text{ K}^{-1})(110 \text{ K}) = 693 \text{ J}$$

(c) Heat supplied, from the first law:

$$Q = \Delta U + W = 693 \text{ J} + 229 \text{ J} = 922 \text{ J}$$

As a consistency check, compute Q directly using C_p :

$$Q = nC_p\Delta T = (0.250 \text{ mol})(33.5 \text{ J mol}^{-1} \text{ K}^{-1})(110 \text{ K}) = 921 \text{ J} \checkmark$$

The two routes agree to within rounding, as they must for any isobaric process of an ideal gas.

(d) If the constant pressure had been 0.50 atm instead, with the same T_1 and the same T_2 , the work done would be:

$$W' = nR\Delta T = (0.250 \text{ mol})(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(110 \text{ K}) = 229 \text{ J}$$

The work is the same. This is worth pausing on. At 0.50 atm, the gas occupies twice the volume it occupied at 1.00 atm (by Boyle's law applied to the initial state), and expanding at half the pressure through double the volume gives exactly the same $P\Delta V$. More formally, $W = P\Delta V$, but ΔV itself depends on P through the ideal-gas law: for the same ΔT , $\Delta V = nR\Delta T/P$ scales inversely with P , and the P in the numerator of $P\Delta V$ cancels the P in the denominator of ΔV . The result is:

$$W = P \times \frac{nR\Delta T}{P} = nR\Delta T,$$

which contains no pressure at all. The isobaric work done on a given sample during a given temperature change is independent of the pressure at which that heating happens.

Making Sense of the Answer: *Three numerical answers and one surprise. The 229J of work, the 693J of ΔU , and the 922J of heat all fit together through the first law, and the two independent routes to Q agree. The surprise is that the work is the same regardless of the pressure, which contradicts an intuition students often bring in from the PV-diagram picture: half the pressure, the reader thinks, must give half the work. But the picture has to adjust along with the numbers. At 0.50atm the initial and final volumes are both twice what they were at 1.00atm, so the rectangle on the PV diagram is shorter in height and longer in width in exactly the right proportions to preserve its area. The rectangle's shape changes; its area does not. This is why reasoning from a formula sometimes delivers insights that a picture can miss.*

Think Like a Physicist: *The independence of isobaric work from pressure is an example of a more general principle: thermodynamic formulae are often more compact than their derivations, and they often reveal cancellations that are not obvious from the underlying physics. $W = nR\Delta T$ carries no pressure on its face, even though $W = P\Delta V$ does. The P drops out because ΔV itself depends on P . A student who internalises $W = nR\Delta T$ as the "main" isobaric formula, with $W = P\Delta V$ as its longer sibling, will move faster through calculations where only n and ΔT are at hand. More generally: when two forms of the same formula exist, the shorter one is usually the one that has absorbed cancellations the longer one still shows. Preferring the shorter form is a small but real time-saver on examination questions that must be answered against the clock.*

Two of the four thermodynamic processes now sit on the desk. The first held the volume fixed and shut the work channel. The second held the pressure fixed and opened both channels, with all the heat required at a rate set by C_p rather than C_v . The third process holds a different variable fixed, namely the temperature itself, and it unlocks a subtle point about how the first law manages heat and work when the internal energy of an ideal gas is not allowed to change. That process is called the isothermal process, and it is the subject of the next section.

THE ISOTHERMAL PROCESS (CONSTANT TEMPERATURE)

The previous section held the pressure fixed and let the volume breathe. This section holds the third and subtlest of the three primary state variables: the **temperature**. The process in which a gas's temperature remains unchanged throughout is called the **isothermal process**, from the Greek **isos** (*equal*) and **thermos** (*heat, temperature*). Holding the temperature fixed has a consequence the student may find surprising: the internal energy of the gas does not change either. U for an ideal gas depends only on T , and if T stays put, so does U . The first law then delivers a clean and slightly unusual result, $Q = W$, meaning every joule of heat flowing into the gas leaves again as work done by the gas on its surroundings. The gas is a pure energy conduit.

Imagine a gas sits inside a thin-walled cylinder. The cylinder is not an insulator; it is a conductor of heat. Surrounding the cylinder is a large reservoir of water, the kind one might find in a slow-moving river, or a bucket with a heating element carefully keeping its temperature fixed. A piston fits into the top of the cylinder. When the piston is pushed down very slowly, or pulled up very slowly, the gas inside does work against the piston, or has work done upon it. At every moment during this slow movement, whatever temperature change would tend to occur inside the gas is immediately undone by heat flowing into or out of the gas through the thin conducting wall. The reservoir is so large that no measurable heating or cooling of the reservoir itself occurs, no matter how long the experiment runs. The gas simply stays at the reservoir's temperature, and the heat bookkeeping records every joule.

Two conditions must hold for this description to apply. The cylinder must conduct heat freely, through thin walls, in constant contact with a large heat reservoir. And the process must be **slow** enough that heat has time to flow in or out at every instant along the way. A **quasi-static process**, the physicist calls it. A rapid compression in a thick-walled cylinder is the opposite of isothermal, and is the subject of the next section.

The picture on the PV diagram is a hyperbola.

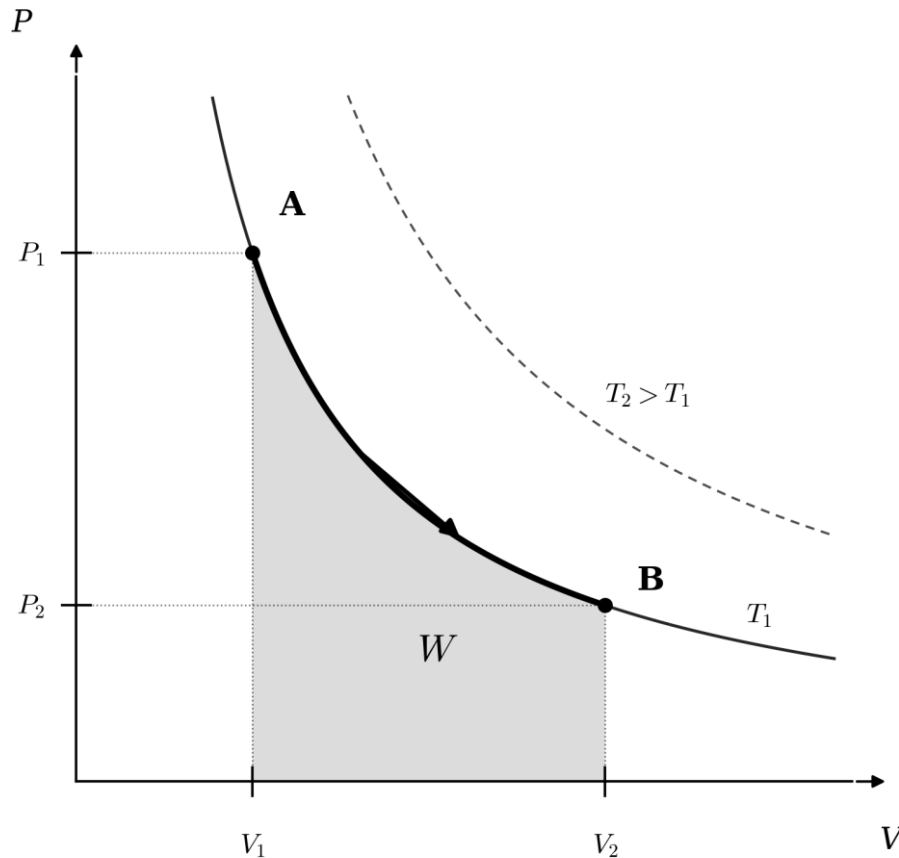


Figure: The isothermal process on a PV diagram. The solid curve traces the hyperbola $PV = \text{constant}$ at temperature T_1 , carrying the gas from state A at (V_1, P_1) down-right to state B at (V_2, P_2) during an expansion. The shaded region beneath the curve between V_1 and V_2 represents the work done by the gas, W . A second isotherm, drawn dashed at a higher temperature $T_2 > T_1$, sits further from the origin: higher- T isotherms always lie further out than lower- T ones, because $PV = nRT$ grows with T . The process shown is an expansion; the same curve, traced in the opposite direction from B to A, is an isothermal compression.

The first law at constant temperature

For an ideal gas, the internal energy U depends only on the temperature T , as the molecular derivation in Chapter 4 established. If T is held fixed throughout a process, then

$$\Delta U = 0 \quad (\text{isothermal, ideal gas})$$

The first law $Q = \Delta U + W$ simplifies at once:

$$Q = W \quad (\text{isothermal, ideal gas})$$

Every joule of heat that enters the gas leaves again as work done by the gas; every joule of work done on the gas must leave again as heat rejected to the surroundings. The gas's own energy store does not grow or shrink. This is a remarkable property: the gas acts as a perfect intermediary between heat and work, converting one to the other with no net change in its own state of excitement.

Boyle's law and the hyperbola

The second consequence of fixed temperature comes from the ideal-gas law. With T , n , and R all constant, $PV = nRT$ becomes:

$$PV = \text{constant} \quad (\text{isothermal}),$$

or, comparing initial and final states,

$$P_1 V_1 = P_2 V_2$$

This is **Boyle's law**, named for the Anglo-Irish natural philosopher Robert Boyle, who published it about 360 years ago, when thermodynamics as a discipline was still 150 years away. Boyle measured the pressure and volume of a fixed amount of air at room temperature, confirmed that their product stayed constant as

long as the temperature did, and handed physics a relationship that holds for every isothermal process of every ideal gas ever since.

On a PV diagram, $PV = \text{constant}$ traces a hyperbola. Different temperatures give different hyperbolas, each called an **isotherm** for its particular temperature. Higher- T isotherms lie further from the origin, because for the same V , a higher T means a higher P , and for the same P , a higher T means a larger V . The dashed curve in the figure above, drawn at the higher temperature T_2 , is one such isotherm; the solid curve, drawn at T_1 , is another.

Work done during an isothermal process

The work done by the gas as its volume changes from V_1 to V_2 is the integral of PdV , just as in any process. What distinguishes the isothermal case is that P is not constant; it varies with V along the hyperbola, following $P = nRT/V$. Substituting into the work integral:

$$W = \int PdV = \int \frac{nRT}{V} dV$$

Since n , R , and T are all constants throughout the process, they come out of the integral:

$$W = nRT \int_{V_1}^{V_2} \frac{dV}{V} = nRT [\ln V]_{V_1}^{V_2} = nRT \ln \frac{V_2}{V_1}$$

Using Boyle's law $P_1V_1 = P_2V_2$ to exchange the volume ratio for a pressure ratio $\left(\frac{V_2}{V_1} = \frac{P_1}{P_2}\right)$ gives the equally useful alternative form:

$$W = nRT \ln \frac{V_2}{V_1} = nRT \ln \frac{P_1}{P_2}$$

The two forms are mathematically identical. The first uses the volume ratio V_2/V_1 ; the second uses the pressure ratio P_1/P_2 (note the inversion; the subscripts swap, because when V goes up, P goes down). Either can be used, depending on which data the problem gives.

Since $Q = W$ for an isothermal process of an ideal gas, the same formula delivers the heat:

$$Q = W = nRT \ln \frac{V_2}{V_1}$$

What the sign of W tells us

Two cases arise, and the sign of W distinguishes them cleanly. If the gas expands, $V_2 > V_1$, so $V_2/V_1 > 1$, and $\ln(V_2/V_1) > 0$. Therefore, $W > 0$: positive work, done by the gas on its surroundings. The gas pushed back the piston; the surroundings absorbed the energy. And $Q > 0$ too: heat flowed in from the reservoir, exactly replacing what the gas spent as work.

If the gas is compressed, $V_2 < V_1$, so $V_2/V_1 < 1$, and $\ln(V_2/V_1) < 0$. Therefore, $W < 0$: negative work, as calculated. *What does a negative work mean physically?* It means the gas did not do positive work on its surroundings; rather, the surroundings did positive work on the gas; they pushed the piston down against the gas's resistance. The sign convention established in an earlier section defines W as work done by the gas on its surroundings. Negative W means the flow went the other way. And $Q < 0$ in lockstep: heat flowed out of the gas into the reservoir, carrying away exactly the work the surroundings had put in. The gas remained at its original temperature throughout, while acting as a pipe connecting the compressor's arm to the river's water.

Examiners like to ask, after a calculation of a negative W for an isothermal compression, "State the significance of the negative sign". The answer is not "the gas lost energy"; the gas did not lose energy, $\Delta U = 0$ throughout. The answer is "**the surroundings did work on the gas**" or equivalently, "the gas was compressed, not expanded". This distinction is worth owning.

The isothermal fingerprint

Four results identify every isothermal process of an ideal gas. You should be able to state them without looking at notes:

PV = constant. Boyle's law holds. The PV diagram traces a hyperbola.

$\Delta U = 0$. Internal energy depends only on T , and T does not change.

$Q = W$. Every joule of heat in leaves as work out; every joule of work in leaves as heat out.

$W = nRT \ln(V_2/V_1) = nRT \ln(P_1/P_2)$. Work done by the gas, equal to the shaded area under the hyperbola on the PV diagram.

The next **condition for applicability** is equally important: *the process must be slow, and the boundary must conduct heat freely*. Any process that violates either condition is not isothermal, and the formula above does not apply to it. The next section will formalize the opposite extreme, in which heat does not flow at all during the process.

The hyperbola is drawn, the logarithm is loaded, four examples now follow.

BINDER Example 23

One mole of an ideal gas at 320K is compressed isothermally from a volume of 8.0L to a volume of 4.0L. Find the work done by the gas and state the significance of the sign of your answer. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Solution

The process is isothermal, so:

$$W = nRT \ln \frac{V_2}{V_1} = (1.0 \text{ mol})(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(320 \text{ K}) \ln \frac{4.0 \text{ L}}{8.0 \text{ L}} = -1844 \text{ J} = -1.84 \text{ kJ}$$

Significance of the negative sign. The work done by the gas is negative, which means the gas did **not** do positive work on its surroundings. Instead, the surroundings did 1.84kJ of positive work on the gas in order to compress it from 8.0L down to 4.0L. The gas was on the receiving end of the energy transfer, not the giving end.

Making Sense of the Answer: *The magnitude, about two kilojoules, is the energy cost of squeezing a mole of gas at room temperature to half its volume. In practice, this is the work a person must supply through a piston handle. Note that the gas's internal energy did not rise in response, because the process was isothermal; the heat reservoir absorbed the 1.84kJ of work as heat, and the temperature inside the cylinder never budged. If the reservoir had not been there, and the gas were compressed fast instead of slowly, the gas would have heated up and the calculation would have needed the adiabatic formulae of the next section.*

Think Like a Physicist: *Three things to own about the logarithm. First, $\ln(0.5)$ is negative, about -0.693 ; students sometimes forget this and return a positive answer. Second, the two conventions $\ln(V_2/V_1)$ for the isothermal work and the sign rule "positive W means work done by the gas" together guarantee that compression ($V_2 < V_1$) produces a negative W , automatically, without any extra sign juggling. Third, the formula's dependence on n and T is linear: compressing two moles gives twice the work; doing it at twice the temperature gives twice the work; halving the volume always gives $-nRT \times \ln(2) = -0.693nRT$ of work, whatever n and T happen to be. These scalings are worth internalizing.*

BINDER Example 24

Sixteen moles of an ideal gas at 320K are compressed isothermally from a volume of 18.0L to a volume of 4.0L. Find the work done by the gas and explain the sign. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Solution

The process is isothermal. Again we apply $W = nRT \ln(V_2/V_1)$:

$$W = (16 \text{ mol})(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(320 \text{ K}) \ln \frac{4.0 \text{ L}}{18.0 \text{ L}} = -64020 \text{ J} = -64.0 \text{ kJ}$$

The sign is negative. The surroundings did 64kJ of work on the gas in order to compress it from 18L down to 4L. At the same time, 64kJ of heat flowed out of the gas into the surrounding reservoir, because $Q = W$ for an isothermal process. The gas was a short-term stopover for the energy; the surroundings' work handle was pushing energy in at one side and the reservoir was receiving it as heat on the other.

Making Sense of the Answer: *Sixteen times the moles of the previous example, plus a larger volume ratio (4.5 rather than 2), gives a much larger work: 64kJ instead of 1.84kJ. Both scalings check out. More gas means more molecules to push against each other's pressure during compression; a larger compression ratio means the work integral extends over a larger range. The linearity in n is exact; the dependence on V_2/V_1 is logarithmic, so squeezing a gas down to a quarter of its volume does not cost four times.*

Think Like a Physicist: *A quick way to spot a sign error in an isothermal problem: if the gas was compressed, W must be negative, and Q must be negative too (heat leaving the gas). If the gas expanded,*

both must be positive (heat entering, work done by the gas on surroundings). Any answer where the process direction and the sign of W disagree is a calculation error, either in the ratio V_2/V_1 or in the logarithm's sign. Check before moving on.

REAL Example 25

A car tyre suddenly bursts on the Morogoro road. Kipute, who is standing nearby, notices that the air rushing out of the hole feels noticeably cold, not warm. Puzzled, she asks Mr. Akilikubwa about it later in the laboratory.

Kipute: "Mr. Akilikubwa, when the tyre burst, the escaping air was cold. I thought compressed air should be warmer than the surroundings, not colder. Why did it feel cold?"

Mr. Akilikubwa: "A good question. The air inside the tyre was at about 2 atmospheres of pressure and at roughly the temperature of the road surface, which on a hot afternoon in Morogoro might be 40°C . When the tyre burst, the air expanded very quickly to atmospheric pressure. That expansion did work against the atmosphere outside, pushing air out of the way to make room for itself. **Where did that work-energy come from?** From inside the gas itself, because the burst was so fast that no heat from the tyre walls or the surrounding air had time to flow into the escaping gas during the expansion. With no heat coming in, and work going out, the gas's internal energy fell; and for an ideal gas, falling internal energy means falling temperature. The air cooled as it escaped, by many tens of degrees. That cold rush is what you felt."

Kipute: "So the compressed air wasn't warmer, it was cooler by the time it reached me."

Mr. Akilikubwa: "Yes. Two temperatures at play: the tyre-wall temperature, which never changed in the brief instant, and the air-inside-the-burst temperature, which dropped sharply during the expansion because there was no time for heat to flow. That quick-expansion process, where the gas cannot exchange heat with its surroundings, has a name. It is called an **adiabatic** process, and it is the subject of our next lesson. The first law still applies: $Q = \Delta U + W$, but with $Q = 0$ because no heat flowed."

Kipute: "Interesting. Now, what if the tyre had been deflated slowly? Say, through a pinhole so small that the air took hours to escape. Would it still have cooled?"

Mr. Akilikubwa: "Ah, this is where the isothermal picture we have been studying comes in. If the deflation is slow enough; very slow, hours-slow, then during the expansion, heat has time to flow from the tyre wall, from the road, from the surrounding air, into the escaping gas at every moment. The gas draws in just enough heat to replace the energy it is spending as work against the atmosphere. Its temperature therefore stays roughly equal to the temperature of its surroundings throughout the expansion. No dramatic cooling; no dramatic warming. The gas emerges at the ambient temperature, having quietly traded heat from the road for work done against the air. That is the isothermal case, and the formula we have just derived, $W = nRT \ln(V_2/V_1)$, computes the work exactly."

Kipute: "So fast expansion cools the gas; slow expansion keeps the temperature constant. The speed of the process picks which one."

Mr. Akilikubwa: "Precisely. And the physics that separates them is not the gas itself; it is the time available for heat to flow through the walls of whatever contains the gas. Thick walls, fast process, little time: adiabatic. Thin walls, slow process, plenty of time: isothermal. Every gas expansion in the world sits somewhere between these two idealizations, and the physicist's first job is to decide which one it is closer to."

Making Sense of the Answer: The same physical act: air escaping a tyre; has two completely different outcomes depending on the speed. Fast: adiabatic, gas cools sharply. Slow: isothermal, gas stays at ambient temperature. The two limits are the two thermodynamic idealizations between which real processes live. The burst tyre was adiabatic (bang, no time for anything); a slow pinhole leak would have been isothermal (hours, all the time in the world).

Think Like a Physicist: Real-world gas expansions and compressions are never exactly isothermal or exactly adiabatic, but almost all of them are close to one or the other. The classification is fast: if the time for the process is much shorter than the time for heat to flow through the boundary, the process is effectively adiabatic. If the time for the process is much longer, it is effectively isothermal. The ratio of the process time to the heat-flow time determines which idealization applies. Students who learn to make this call quickly are students who get thermodynamics right on the first try, because they have already chosen the correct formula before touching the calculator.

HOT Example 26

Two moles of an ideal gas are compressed isothermally from 10L to 5L. Perform the calculation at two different temperatures: (a) $T = 300\text{K}$, and (b) $T = 600\text{K}$. For each case, find the work done on the gas and the heat released by the gas. (c) Comment on the ratio of the two works and the physical reason behind the scaling. (d) For the hotter case, by what factor does the ratio V_2/V_1 need to change for the work to match the colder case's answer? Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

(a) At $T = 300\text{K}$:

$$W_a = nRT \ln \frac{V_2}{V_1} = (2.0 \text{ mol})(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(300 \text{ K}) \times \ln \frac{5\text{L}}{10\text{L}} = -3458 \text{ J} \approx -3.46 \text{ kJ}$$

The work is -3.46kJ , so the surroundings do 3.46kJ of positive work on the gas. Since $Q = W$ for an isothermal process:

$$Q_a = -3.46 \text{ kJ},$$

meaning 3.46kJ of heat leaves the gas into the surroundings during the compression.

(b) At $T = 600\text{K}$, with the same n and same volume ratio:

$$W_b = (2.0 \text{ mol})(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(600 \text{ K}) \times \ln \frac{5\text{L}}{10\text{L}} = -6916 \text{ J} \approx -6.92 \text{ kJ}$$

And the heat released:

$$Q_b = -6.92 \text{ kJ}$$

(c) The ratio of the two works:

$$\frac{W_b}{W_a} = \frac{-6.92 \text{ kJ}}{-3.46 \text{ kJ}} = 2.00.$$

Comment: The work at 600K is exactly twice the work at 300K .

Physical reason: The isothermal work formula $W = nRT \ln(V_2/V_1)$ is linear in T . Doubling T doubles W for fixed n and fixed volume ratio. Compressing hot gas costs twice as much work as compressing the same amount of the same cold gas through the same volume ratio, because at the higher temperature the molecules are banging harder (with higher KE) against the piston and the piston must push harder to displace them.

(d) For the hotter case at 600K , we ask: *what volume ratio would make the work match the colder case at 300K , namely -3.46kJ ?* Set up the equation:

$$(2.0 \text{ mol})(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(600 \text{ K}) \times \ln(\text{ratio}) = -3458 \text{ J},$$

which rearranges to:

$$\begin{aligned} \ln(\text{ratio}) &= \frac{-3458 \text{ J}}{9977 \text{ J}} = -0.3466, \\ \text{ratio} &= e^{-0.3466} \\ \frac{V_2}{V_1} &= e^{-0.3466} = 0.707 \end{aligned}$$

So at 600K , the gas need only be compressed to about 71% of its original volume (not to 50%, as in the colder case) to lose the same 3.46kJ of work.

Making Sense of the Answer: Parts (a), (b), and (c) together show the linear-in- T scaling of isothermal work in action: doubling the temperature doubles the work at the same volume ratio. Part (d) inverts the question: given a target work, how much compression does a hotter gas need? Answer: less.

Think Like a Physicist: The isothermal formula is linear in n , linear in T , and logarithmic in the volume ratio. Three different scalings, three different meanings. A doubling of n doubles the work (more molecules, each contributing the same per-molecule expansion energy). A doubling of T doubles the work (same molecules, each pushing harder). A doubling of the volume ratio adds $\ln(2) \times nRT$ to the work (logarithmic, so ratios add in log-space rather than multiplying). Students who carry these three scalings in their head can estimate any isothermal problem's answer to within a factor of two before they reach for a calculator, which is exactly the speed an examination demands.

Three of the four thermodynamic processes now sit on the desk. The volume was held fixed; the pressure was held fixed; the temperature was held fixed. The fourth and last process holds something less obvious: the heat exchange itself is held at zero. In such a process the gas is thermally sealed from its surroundings, and any change of volume must be paid for from the gas's own internal energy, with consequences for the

temperature that the isothermal case kept carefully disguised. That process is the adiabatic process, and it is the subject of the next section, where the dialogue of REAL Example 25 will receive its formal treatment.

THE ADIABATIC PROCESS (NO HEAT EXCHANGE)

The previous section held the temperature constant by allowing heat to flow freely through a thin conducting wall, in a process slow enough for that flow to keep up at every instant. The fourth and last of the basic thermodynamic processes goes to the opposite extreme. Instead of holding T constant, it holds the heat exchange Q at zero. Either the gas is wrapped in walls so **insulating** that no heat can flow through them in the available time, **or** the process is so **fast** that heat has no time to flow even through ordinary walls. Either way, $Q = 0$, and the gas exchanges energy with the outside world purely as mechanical work. The process is called the **adiabatic process**, from the Greek **a** (not) and **dia-bainein** (*to pass through*), in honour of the heat that does not pass through.

Imagine a cylinder with thermos-flask walls and a piston at the top. The walls keep heat out as well as a vacuum-jacketed flask keeps a hot drink hot, which is to say, very nearly perfectly. When the gas inside is compressed by pushing the piston down, the work done on the gas has nowhere to go except into raising its internal energy, which raises its temperature. When the gas is allowed to expand by letting the piston rise, the work the gas does on the piston comes out of its internal energy, which falls, taking the temperature with it. The bicycle pump that warmed up under fast pumping in Chapter 4, and the burst tyre whose escaping air felt cold in the dialogue of an earlier section, are both examples of this process happening in the wild. Fast processes do not give heat time to leak through walls; they are adiabatic by virtue of speed alone, even when the walls are not particularly insulating.

On a PV diagram, the adiabat is a curve that resembles the isotherm but is steeper. The figure below shows the two together, drawn from a common starting point so the comparison is direct.

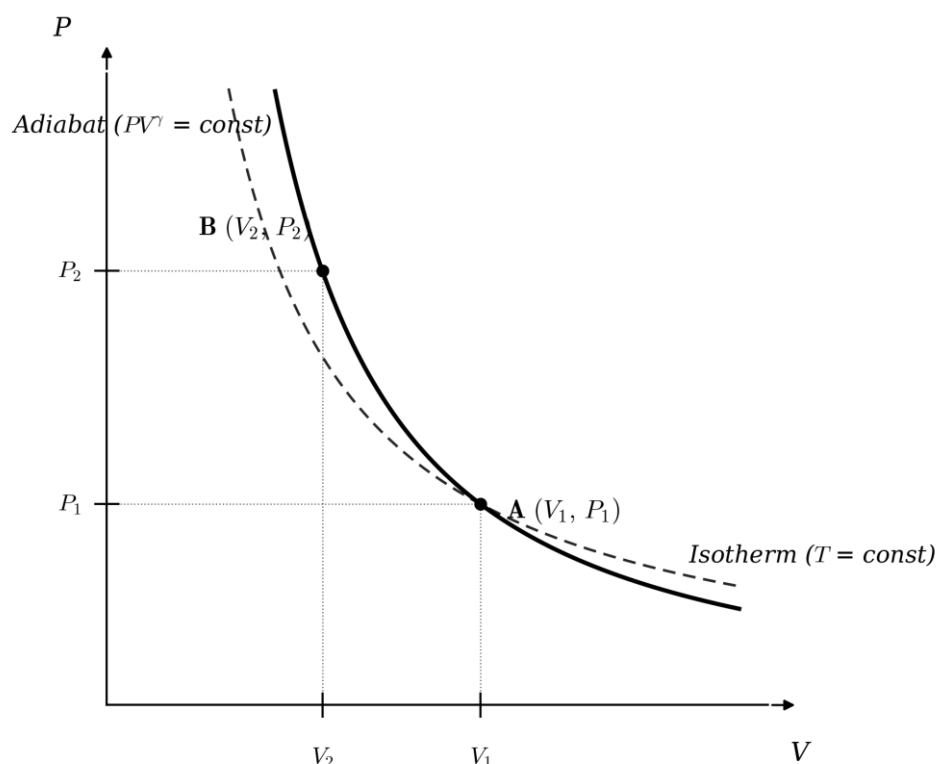


Figure: An adiabatic process and an isothermal process drawn from the same initial point A. The adiabat (solid) follows $PV^\gamma = \text{constant}$; the isotherm (dashed) follows $PV = \text{constant}$. The adiabat is everywhere steeper than the isotherm: when the gas is compressed adiabatically from A to B, both V and T decrease less than they would along the isotherm, so the pressure rises faster per unit of volume change. The physical reason is the first law: an isothermal compression sheds heat to the surroundings as it goes, keeping T fixed; an adiabatic compression keeps that heat in, so T rises during the process and the rising T contributes its own pressure rise on top of the geometric one from compression. The adiabat is therefore steeper than the isotherm by exactly the factor γ , the adiabatic index.

The first law with no heat exchange

With $Q = 0$, the first law $Q = \Delta U + W$ rearranges immediately:

$$\Delta U = -W \quad (\text{adiabatic})$$

This is the central physics of the adiabatic process, and it deserves a moment to absorb. In every other process so far, work was a side effect of moving energy between the heat reservoir and the gas. In an adiabatic process, work and internal energy are directly tied: every joule of work the gas does on its surroundings is a joule taken out of its internal energy, and every joule of work done on the gas by its surroundings is a joule added to its internal energy. There is no third party, no heat reservoir, no thermometer-fixing river. The gas's energy store is the only buffer.

Since the internal energy of an ideal gas depends only on temperature, $\Delta U = nC_v\Delta T$, the equation rearranges to a temperature statement:

$$nC_v\Delta T = -W,$$

which says: *when work is done by the gas ($W > 0$), $\Delta T < 0$ and the gas cools; when work is done on the gas ($W < 0$), $\Delta T > 0$ and the gas warms.* The cold air rushing from the burst tyre, the warm bicycle-pump barrel, and the warm hands rubbed in fast friction all share this single mechanism.

Derivation of the adiabatic equation

The first law in differential form, with $Q = 0$, gives a relation between an infinitesimal volume change and an infinitesimal temperature change. The work done by the gas during a small expansion dV is PdV , and the change in internal energy during a small temperature change dT is nC_vdT . Setting $\Delta U = -W$ in their differential forms:

$$nC_vdT = -PdV \quad (1)$$

This is one equation in two infinitesimal unknowns dT and dV . A second equation comes from the ideal-gas law $PV = nRT$. Differentiating both sides (the product rule on the left, only T varying on the right with n and R constant):

$$PdV + VdP = nRdT \quad (2)$$

Now eliminate dT between (1) and (2). From (1):

$$dT = -\frac{PdV}{nC_v}$$

Substituting into (2):

$$PdV + VdP = nR \times \left(-\frac{PdV}{nC_v}\right) = -\frac{R}{C_v}PdV$$

Move the right-hand-side term to the left and combine the PdV terms:

$$PdV + \frac{R}{C_v}PdV + VdP = 0,$$

$$\left(1 + \frac{R}{C_v}\right)PdV + VdP = 0$$

The factor in parentheses simplifies neatly. Using Mayer's equation:

$C_p - C_v = R$, the ratio $\frac{R}{C_v} = \frac{C_p - C_v}{C_v} = \frac{C_p}{C_v} - 1 = \gamma - 1$, so $1 + \frac{R}{C_v} = 1 + \gamma - 1 = \gamma$. The equation becomes:

$$\gamma PdV + VdP = 0 \quad (3)$$

Pause for a moment to look at what has just happened. The temperature has been eliminated. Equation (3) involves only P , V , and the constant γ ; no T at all. This is genuinely surprising. We started from a first-law equation that featured the temperature change dT , and we ended with a relationship purely between P , V , and dP , dV . The temperature has been hidden inside γ , where it is no longer a variable but a fixed property of the gas. The relationship that survives is between pressure and volume only. Students often find this elimination step disorienting on first reading; it is worth re-reading until the surprise has worn off, because

the next step turns equation (3) into the famous adiabatic equation commonly examined in advanced thermodynamics problems.

Divide equation (3) through by PV (this assumes the gas has non-zero pressure and volume, which is always true for a real gas in a real situation):

$$\gamma \frac{dV}{V} + \frac{dP}{P} = 0$$

Each term is now exact: dV/V is the differential of $\ln V$, and dP/P is the differential of $\ln P$. Integrating both sides:

$$\int \gamma \frac{dV}{V} + \int \frac{dP}{P} = \int 0$$

$$\gamma \ln V + \ln P = \text{constant or } \ln V^\gamma + \ln P = \text{constant}$$

Combining logarithms:

$$\ln(PV^\gamma) = \text{constant},$$

which means the argument of the logarithm is itself a constant:

$$PV^\gamma = \text{constant} \quad (\text{adiabatic})$$

This is the **adiabatic equation**, and it is the central result of this section. *For any quasi-static adiabatic process of an ideal gas, the product P times V to the power γ stays constant from the start of the process to the end.* Or, comparing initial and final states using subscripts on each variable:

$$P_1 V_1^\gamma = P_2 V_2^\gamma \quad (4)$$

Two companion forms

The adiabatic equation (4) connects pressure and volume. Two further relations connect temperature with each of them, derived by substituting the ideal-gas law into (4).

First, eliminate P using $P = nRT/V$. Equation (4) becomes:

$$\frac{nRT_1}{V_1} V_1^\gamma = \frac{nRT_2}{V_2} V_2^\gamma$$

The nR cancels, and the V^γ/V combines into V to the power $(\gamma-1)$:

$$T_1 V_1^{\gamma-1} = T_2 V_2^{\gamma-1}$$

which is often written compactly as:

$$TV^{\gamma-1} = \text{constant} \quad (\text{adiabatic})$$

This is the form of choice when V is what is changing during the process and T is what is being asked for.

Second, eliminate V using $V = nRT/P$. Equation (4) becomes:

$$P_1 \left(\frac{nRT_1}{P_1} \right)^\gamma = P_2 \left(\frac{nRT_2}{P_2} \right)^\gamma$$

The constants nR cancel between the two states, and after simplification:

$$\frac{T_1^\gamma}{P_1^{\gamma-1}} = \frac{T_2^\gamma}{P_2^{\gamma-1}}$$

This rearranges to a particularly useful form for problems where pressure changes adiabatically:

$$\left(\frac{T_2}{T_1} \right)^\gamma = \left(\frac{P_2}{P_1} \right)^{\gamma-1}$$

$$\frac{T_2}{T_1} = \left(\frac{P_2}{P_1} \right)^{\frac{\gamma-1}{\gamma}}$$

Writing it in this fraction-and-exponent form lets the student read off the temperature ratio directly from the pressure ratio. This is the form the worked examples below will reach for most often.

Three forms of the adiabatic work, all equivalent

The work done by the gas during an adiabatic process can be written in three equivalent forms. Each is convenient in a different context, and each can be derived from a different starting point. The derivations are short but worth seeing in full.

Form 1: from the work integral

Start from the general definition of work done by a gas during a quasi-static process:

$$W = \int_{V_1}^{V_2} P dV$$

During an adiabatic process, the pressure depends on volume through $PV^\gamma = \text{constant}$. Setting that constant equal to its value at the initial state, $P_1 V_1$ raised to γ :

$$P V^\gamma = P_1 V_1^\gamma,$$

which rearranges to express the pressure during the process in terms of V :

$$P = \frac{P_1 V_1^\gamma}{V^\gamma} = P_1 V_1^\gamma V^{-\gamma}$$

The factor $P_1 V_1^\gamma$ with subscript 1 is constant during the process; only V varies. So:

$$W = \int_{V_1}^{V_2} P dV = P_1 V_1^\gamma \int_{V_1}^{V_2} V^{-\gamma} dV = P_1 V_1^\gamma \left[\frac{V^{1-\gamma}}{1-\gamma} \right]_{V_1}^{V_2} = P_1 V_1^\gamma \times \frac{(V_2^{1-\gamma} - V_1^{1-\gamma})}{1-\gamma}$$

Multiply through and simplify each term:

$$W = \frac{P_1 V_1^\gamma V_2^{1-\gamma} - P_1 V_1^\gamma V_1^{1-\gamma}}{1-\gamma} = \frac{P_1 V_1^\gamma V_2^{1-\gamma} - P_1 V_1}{1-\gamma}$$

The term $P_1 V_1^\gamma V_2^{1-\gamma}$ can be simplified further by using adiabatic relation: $P_1 V_1^\gamma = P_2 V_2^\gamma$

$$W = \frac{P_2 V_2^\gamma V_2^{1-\gamma} - P_1 V_1}{1-\gamma} = \frac{P_2 V_2 - P_1 V_1}{1-\gamma}$$

Multiplying numerator and denominator by -1 to put the formula into its conventional form (positive denominator):

$$W = \frac{P_1 V_1 - P_2 V_2}{\gamma - 1} \quad (\text{form 1})$$

This is the adiabatic work in pressure-volume form. It can be applied directly whenever the initial and final pressures and volumes are known.

Form 2: from form 1 via the ideal-gas law

The ideal-gas law applied at each end of the process gives:

$$P_1 V_1 = nRT_1, \quad P_2 V_2 = nRT_2$$

Substitute these into form 1:

$$W = \frac{nRT_1 - nRT_2}{\gamma - 1} = \frac{nR(T_1 - T_2)}{\gamma - 1} \quad (\text{form 2})$$

This is the temperature form. It is the form to use when initial and final temperatures and the moles are at hand, but volumes or pressures are not.

Form 3: from the first law directly

The first law of thermodynamics with $Q = 0$ gives $\Delta U = -W$, which rearranges to $W = -\Delta U$. The internal energy of an ideal gas depends only on temperature:

$$\Delta U = nC_v(T_2 - T_1)$$

So:

$$W = -nC_v(T_2 - T_1) = nC_v(T_1 - T_2) \quad (\text{form 3})$$

This is the internal-energy form. It is the fastest way to compute W when ΔU has already been computed for some other reason, and it provides an independent route to the answer that can verify forms 1 or 2.

How the three forms reconcile

Forms 1 and 2 differ only by the substitution $PV = nRT$, applied at each state. They are mathematically identical for any ideal gas.

Forms 2 and 3 differ only by Mayer's equation. From the rearrangements derived in the molar heat capacities section, $C_v = R/(\gamma - 1)$. Substituting into form 3:

$$W = nC_v(T_1 - T_2) = n \times \frac{R}{\gamma - 1} \times (T_1 - T_2) = \frac{nR(T_1 - T_2)}{\gamma - 1},$$

which is exactly form 2. So forms 2 and 3 are also identical for any ideal gas.

All three forms are therefore one formula in three different costumes. Use whichever costume matches the given data: form 1 for pressures and volumes, form 2 for moles and temperatures, form 3 for an internal-energy cross-check. A well-set adiabatic problem will let you compute the answer with one form and verify it with another, which is the surest defence against arithmetic errors involving γ .

Why the adiabat is steeper than the isotherm

Both curves pass through A on the figure above. From A, the isotherm follows $PV = \text{constant}$, which has slope $dP/dV = -P/V$ along the curve. The adiabat follows $PV^\gamma = \text{constant}$, which has slope $dP/dV = -\gamma P/V$. The adiabat's slope is steeper than the isotherm's by exactly the factor γ at every point. For $\gamma = 1.4$ (diatomic), the adiabat falls 1.4 times as fast per unit of V ; for $\gamma = 5/3$ (monatomic), it falls $5/3$ times as fast.

The physical reason for the steepness difference is the temperature. Isothermal compression keeps T fixed by allowing heat to leave. Adiabatic compression has no such heat outlet, so the temperature rises during the compression. A higher T means a higher P at any given V (because $PV = nRT$), so the adiabat reaches higher pressures for the same volume reduction than the isotherm does. Conversely, isothermal expansion absorbs heat to keep T fixed; adiabatic expansion has no heat to absorb, so T falls, and the adiabat reaches lower pressures for the same volume increase. Both effects make the adiabat steeper.

The adiabatic fingerprint

Five results identify every adiabatic process of an ideal gas. Each one will be reached for in the worked examples:

$Q = 0$. The defining condition. No heat exchanged with the surroundings.

$PV^\gamma = \text{constant}$. The pressure-volume relation. The PV diagram traces an adiabat, steeper than an isotherm by the factor γ .

$TV^{\gamma-1} = \text{constant}$. Temperature-volume relation. The form to use when V is known and T is sought.

$\frac{T_2}{T_1} = \left(\frac{P_2}{P_1}\right)^{\frac{\gamma-1}{\gamma}}$. Temperature-pressure relation. The form to use when P is known and T is sought.

$\Delta U = -W = nC_v(T_1 - T_2)$. The first law with $Q = 0$. Three equivalent forms of W are available, all reconcilable.

The walls insulate, the gas heats and cools by its own arithmetic, six examples now follow.

BINDER Example 27

0.15mol of a monatomic ideal gas, initially at 250kPa and 320K, expands adiabatically until its pressure drops to 100kPa. Find the final temperature of the gas and the work done by the gas during the expansion. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

Monatomic means $\gamma = 5/3$.

Final temperature, from the temperature-pressure form of the adiabatic equation:

$$\frac{T_2}{T_1} = \left(\frac{P_2}{P_1}\right)^{\frac{(\gamma-1)}{\gamma}}$$

With $\frac{(\gamma-1)}{\gamma} = \frac{5/3-1}{5/3} = 0.4$:

$$T_2 = (320\text{K}) \times \left(\frac{100\text{kPa}}{250\text{kPa}}\right)^{0.4} = 222\text{K}$$

So the final temperature is 222K.

The gas has cooled by nearly one hundred degrees during its adiabatic expansion.

Work done by the gas, using the temperature form (form 2):

$$W = \frac{nR(T_1 - T_2)}{\gamma - 1} = \frac{(0.15\text{ mol})(8.314\text{ J mol}^{-1}\text{ K}^{-1})(320\text{K} - 222\text{K})}{5/3 - 1} = 183\text{J}$$

Making Sense of the Answer: A roughly hundred-kelvin cooling and 183J of work for a 0.15mol sample is a quietly impressive amount of physics from a small quantity of gas. The work done came out of the gas's own internal energy, so we can verify the result with form 3: $\Delta U = nC_v\Delta T = (0.15\text{mol})(3/2)(8.314\text{ J mol}^{-1}\text{ K}^{-1})(222\text{ K} - 320\text{ K}) \approx -183\text{J}$, and $W = -\Delta U = +183\text{J}$, agreeing exactly. Two completely different formulae, two routes through the algebra, one identical answer. That is the kind of internal consistency every adiabatic problem should carry.

Think Like a Physicist: The exponent $(\gamma-1)/\gamma$ is small for any real gas: 2/5 for monatomic, 2/7 for diatomic, 1/4 for polyatomic. This means the temperature ratio T_2/T_1 never changes by very much for moderate pressure ratios, even though the pressure ratio itself can be large. In this problem, the pressure ratio was 0.4 but the temperature ratio came out to 0.69, which is a much milder change. This is why adiabatic cooling, while real, is rarely catastrophic for gases that change pressure by factors of two or three. Order-of-magnitude pressure changes are needed to produce order-of-magnitude temperature changes adiabatically.

BINDER Example 28

A motorcar tyre is at 4.0atm and 27°C. The tyre suddenly bursts, and the air inside expands to atmospheric pressure (1.0atm). Taking $\gamma = 1.4$ for the air, calculate the temperature of the escaping air immediately after the burst.

Solution

The expansion is **sudden**, fast enough that no heat has time to flow into the air through the surrounding tyre or atmosphere during the brief expansion event. So the process is adiabatic, and the temperature-pressure form of the adiabatic equation applies.

$$T_1 = (27 + 273)\text{K} = 300\text{K}$$

Apply the adiabatic temperature-pressure relation:

$$\frac{T_2}{T_1} = \left(\frac{P_2}{P_1}\right)^{\frac{(\gamma-1)}{\gamma}}$$

With $\frac{(\gamma-1)}{\gamma} = \frac{1.4-1}{1.4} = 0.286$:

$$T_2 = (300\text{K}) \times \left(\frac{1.0\text{atm}}{4.0\text{atm}}\right)^{0.286} = 202\text{K} = -71^\circ\text{C}.$$

The air emerging from the burst tyre is therefore at about -71°C (almost a hundred degrees below the road surface temperature).

Making Sense of the Answer: Negative seventy-one degrees Celsius sounds dramatic and is dramatic. But the cooling is brief: the cold air mixes within seconds with the surrounding atmospheric air, which is around 300K, and the temperature equilibrates rapidly. The -71°C is the temperature of the air for the very short interval immediately after the burst, before that mixing has had a chance to happen. It is what a thermometer pointed straight at the burst hole, in the first fraction of a second, would read. After a few seconds the cold

air is gone and the residual air at the burst site is back to ambient. Adiabatic cooling is a transient effect, sharp but short.

Think Like a Physicist: This is the canonical exam-style adiabatic question: given the initial state and the final pressure, find the final temperature. The adiabatic temperature-pressure formula is the workhorse, and the only arithmetic challenge is the non-integer exponent. For $\gamma = 1.4$ (most exam problems), the exponent is $2/7 \approx 0.286$. For $\gamma = 5/3$ (monatomic), it is $2/5 = 0.4$. For $\gamma = 4/3$ (polyatomic), it is $1/4 = 0.25$. The rest of the calculation is direct substitution.

BINDER Example 29

Air at atmospheric pressure ($1.0 \times 10^5 \text{ Pa}$) is compressed adiabatically to half its initial volume. Taking $\gamma = 1.4$ for the air, find the final pressure.

Solution

From the adiabatic pressure-volume relation:

$$P_1 V_1^\gamma = P_2 V_2^\gamma,$$

rearranging for P_2 :

$$P_2 = P_1 \times \left(\frac{V_1}{V_2}\right)^\gamma$$

With $V_2 = V_1/2$, the volume ratio $V_1/V_2 = 2$:

$$P_2 = (1.0 \times 10^5 \text{ Pa}) \times (2)^{1.4} = 2.64 \times 10^5 \text{ Pa}$$

So the final pressure is about $2.64 \times 10^5 \text{ Pa}$.

Making Sense of the Answer: Halving the volume isothermally would have doubled the pressure, by Boyle's law. Halving it adiabatically produces a 2.64-fold increase, more than doubling. The extra factor of 1.32 comes from the temperature rise that accompanies adiabatic compression: the gas heats up, and that temperature rise contributes its own pressure increase on top of the geometric one from compression alone. The factor by which the adiabatic pressure exceeds the isothermal pressure for the same volume change is exactly $\left(\frac{V_1}{V_2}\right)^{\gamma-1} = 2^{0.4} = 1.32$, which matches: $2.64 / 2.0 = 1.32$.

Think Like a Physicist: The adiabatic compression ratio $2^\gamma = 2.64$ sits between two simpler benchmarks. Boyle's law for the same volume change would give $2^1 = 2$ (pure compression, no temperature effect). Real gases have γ between 1 and 5/3, so the adiabatic compression ratio for halved volume sits between 2.0 and $2^{1.67} \approx 3.17$. Most real gases come out at 2.6 ± 0.3 , which is exactly where this answer sits.

REAL Example 30

A scuba diver in Lake Tanganyika is at a depth of 30m, where the water pressure is approximately 4.0atm. She releases an air bubble from her regulator. The bubble, initially at the local water temperature of 22°C, rises rapidly to the lake's surface, where the pressure is just 1.0atm. Taking $\gamma = 1.4$ for the air inside the bubble, find the temperature of the air inside the bubble when it arrives at the surface, and explain in physical terms why the bubble is colder than the water it has risen through.

Solution

The word "rapidly" is the physical cue. The bubble's centimetre-scale interior with its thin gas-water boundary cannot conduct heat fast enough to keep up with a fast rise of several seconds, so during the rise **no** significant heat flows from the surrounding water into the air inside the bubble. With $Q \approx 0$, the bubble's expansion is adiabatic.

$$T_1 = (22 + 273) \text{ K} = 295 \text{ K}$$

With $(\gamma-1)/\gamma = 2/7 \approx 0.286$:

$$T_2 = (295 \text{ K}) \times \left(\frac{1.0}{4.0}\right)^{0.286} = 198 \text{ K} = -75^\circ\text{C}.$$

So the air inside the bubble at the moment it breaks the surface is at about -75°C .

Physical explanation: As the bubble rises, the water pressure pushing inward on it decreases, so the bubble expands. The expanding air pushes back the water around it, doing work on the water; work that comes out of the air's internal energy because the rise is fast and there is no time for heat to flow into the bubble from the surrounding water through the bubble's thin gas-water interface. Falling internal energy means falling temperature. The faster the bubble rises, the closer this picture is to perfectly adiabatic. A slow rise would allow the warm water to keep the bubble at ambient temperature throughout (an isothermal expansion). The bubble's actual behaviour is somewhere between, but for a typical fast-rising bubble it is close enough to adiabatic that the calculation above gives a useful estimate.

Making Sense of the Answer: -75°C is sharp cooling for a bubble rising through 22°C water. In practice, the bubble does not stay at -75°C for long: as it breaks the surface, the air inside mixes with the atmosphere and warms back to ambient within seconds. The -75°C is the air's transient state during the rise itself, and is what would be measured by a tiny thermometer riding inside the bubble. The bubble's wall, and the water immediately surrounding the bubble, would also cool slightly through contact with the cold air, but this transient effect is too small to be felt by the diver.

Think Like a Physicist: The adiabatic model applies whenever the timescale of the volume change is short compared with the timescale of heat conduction through the boundary. For a centimetre-scale bubble rising metres per second through water, the rise takes a few seconds; heat conduction across the bubble's tiny diameter takes considerably longer. The adiabatic approximation is therefore reasonable. For a slow geological process such as magma rising through kilometres of crust over millions of years, the expansion is effectively isothermal, even though the volumes are vastly larger and the pressure changes far more dramatic. Speed and length scales together decide which idealization applies; the underlying physics does not change with size, only with the ratio of process time to heat-flow time.

HOT Example 31

An ideal gas ($n = 2.0\text{mol}$, $\gamma = 1.4$) is initially at 300K , $1.0 \times 10^5\text{Pa}$, and occupies a volume of 0.050m^3 . The gas is compressed adiabatically until its volume is reduced to 0.020m^3 . Find: (a) the final temperature, (b) the final pressure. (c) Verify the two answers are consistent with the ideal-gas law. (d) Calculate the work done on the gas during the compression. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

(a) Final temperature from $TV^{\gamma-1} = \text{constant}$:

$$T_2 = T_1 \times \left(\frac{V_1}{V_2}\right)^{\gamma-1}$$

With $\gamma - 1 = 0.4$:

$$T_2 = (300\text{K}) \times \left(\frac{0.050\text{m}^3}{0.020\text{m}^3}\right)^{0.4} = 433\text{K}$$

(b) Final pressure from $PV^{\gamma} = \text{constant}$:

$$P_2 = P_1 \times \left(\frac{V_1}{V_2}\right)^{\gamma} = (1.0 \times 10^5 \text{ Pa}) \times (2.5)^{1.4} = 3.6 \times 10^5 \text{ Pa}$$

(c) At state 2:

$$P_2 V_2 = (3.609 \times 10^5 \text{ Pa})(0.020 \text{ m}^3) = 7218\text{J},$$

$$nRT_2 = (2.0 \text{ mol})(8.314 \text{ J mol}^{-1}\text{K}^{-1})(433 \text{ K}) = 7200\text{J} \approx 7218\text{J} \quad (P_2 V_2)$$

These agree to within rounding ($P_2 V_2 \approx nRT_2$), as the ideal-gas law requires. The adiabatic answers are internally consistent.

(d) Work done **by** the gas, from form 2:

$$W_{\text{gas}} = \frac{nR(T_1 - T_2)}{\gamma - 1} = \frac{(2.0 \text{ mol})(8.314 \text{ J mol}^{-1}\text{K}^{-1})(300 \text{ K} - 433 \text{ K})}{0.4} = -5530\text{J}$$

So the work done by the gas is -5530J , meaning the work done on the gas by the surroundings is $+5530\text{J}$:

$$W_{\text{on gas}} = +5530 \text{ J} = 5.5\text{kJ}$$

Verification using the internal-energy form. The change in internal energy is:

$$\begin{aligned}\Delta U &= nC_v\Delta T = (2.0 \text{ mol}) \times \frac{R}{\gamma - 1} \times (T_2 - T_1) \\ &= (2.0 \text{ mol}) \times \frac{8.314 \text{ J mol}^{-1} \text{ K}^{-1}}{0.4} \times (433\text{K} - 300\text{K}) = +5530 \text{ J}\end{aligned}$$

Since $Q = 0$, ΔU equals the work done on the gas, and both come out to 5.5kJ. Two completely different formulae deliver the same answer, as the first law of thermodynamics requires for any adiabatic process.

Making Sense of the Answer: Five and a half kilojoules of work compresses 2 moles of gas from 50L to 20L, raising its temperature from room temperature to 160°C in the process. All 5.5kJ of that work ends up stored as internal energy in the gas, because $Q = 0$ means no energy leaks out as heat. This is exactly why a tightly compressed gas cylinder, freshly filled, feels warm to the touch, the compressor's work has gone into the gas. The two-route consistency check (PV-and-V-form work versus internal-energy form) confirms the answer in two independent ways, which is the safest possible check for a calculation involving exponents and γ .

Think Like a Physicist: The four parts of this problem map directly onto the four ways to think about an adiabatic process. Part (a) used the temperature-volume relation; part (b) used the pressure-volume relation; part (c) used the ideal-gas law as a consistency check between the two; part (d) used the work formula and confirmed it via the internal-energy form. Each part is a different face of the same underlying physics. Being able to switch fluidly between these four perspectives, and using one to check another, is the mark of a student who has mastered adiabatic processes rather than just memorised a formula. Every adiabatic question can be checked this way.

HOT Example 32

One mole of oxygen gas ($\gamma = 1.4$) is initially at 2.0atm and 320K, occupying some volume V . The gas is first compressed isothermally until its volume halves. Then, starting from the end state of the first compression, the gas expands adiabatically and reversibly back to its original volume V . Find:

- The pressure of the gas at the end of the isothermal compression.
- The work done by the gas during the isothermal compression.
- The pressure of the gas at the end of the adiabatic expansion.
- The temperature of the gas at the end of the adiabatic expansion.
- Comment on whether the gas has returned to its initial state, and explain in physical terms.

Solution

Three states are involved in this problem, and naming them clearly at the start prevents confusion later. **State 1** is the initial state. **State 2** is the end of the isothermal compression. **State 3** is the end of the adiabatic expansion.

State 1: $P_1 = 2.0\text{atm}$, $T_1 = 320\text{K}$, volume = V (a quantity left as a symbol because no numerical V is given).

- (a) Isothermal compression from V to $V/2$. Boyle's law applies because the temperature is held constant:

$$\begin{aligned}P_1V_1 &= P_2V_2, \\ P_2 &= P_1 \times \frac{V_1}{V_2} = (2.0 \text{ atm}) \times \frac{V}{V/2} = (2.0\text{atm}) \times 2 = 4.0\text{atm}\end{aligned}$$

The pressure of the gas at the end of the isothermal compression is 4.0atm.

- (b) Work done by the gas during the isothermal compression:

$$W_{12} = nRT \ln \frac{V_2}{V_1} = (1.0 \text{ mol})(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(320 \text{ K}) \times \ln \frac{V/2}{V} = -1844 \text{ J} \approx -1.84 \text{ kJ}$$

The work done by the gas is -1844 J (or -1.84 kJ)

The negative sign means the surroundings did 1.84kJ of work on the gas during the compression. The gas itself did -1.84 kJ of work on the surroundings.

(c) Adiabatic expansion from $V_2 = V/2$ back to $V_3 = V$. The starting state of this expansion is **State 2**:

$P_2 = 4.0\text{atm}$, $V_2 = V/2$, $T_2 = 320\text{K}$. Using the adiabatic pressure-volume relation:

$$P_2 V_2^\gamma = P_3 V_3^\gamma,$$

$$P_3 = P_2 \times \left(\frac{V_2}{V_3}\right)^\gamma = (4.0\text{ atm}) \times (V/2 \div V)^{1.4} = (4.0\text{ atm}) \times (0.5)^{1.4} = 1.52\text{atm}$$

The pressure of the gas at the end of the adiabatic expansion is 1.52atm.

(d) Temperature at State 3, using the adiabatic temperature-volume relation:

$$T_2 V_2^{\gamma-1} = T_3 V_3^{\gamma-1},$$

$$T_3 = T_2 \times \left(\frac{V_2}{V_3}\right)^{\gamma-1} = (320\text{ K}) \times (0.5)^{0.4} = 243\text{K}$$

The temperature of the gas at the end of the adiabatic expansion is 243K.

(e) Comment on the final state: The gas has returned to its original volume V , but it has **not** returned to its original state. Three numbers tell the story:

Volume: $V_3 = V = V_1$. **Equal** to the initial value.

Pressure: $P_3 = 1.52\text{atm}$, less than $P_1 = 2.0\text{atm}$. **Lower** than initial.

Temperature: $T_3 = 243\text{K}$, less than $T_1 = 320\text{K}$. **Lower** than initial.

Why has the gas not returned to its original state, despite returning to its original volume? Because the two processes were genuinely different and did not cancel out. The isothermal compression took heat out of the gas ($Q = W = -1.84\text{kJ}$) into the surroundings while leaving the temperature unchanged. The adiabatic expansion that followed exchanged no heat at all with the surroundings, but lost work-energy from inside the gas, so the temperature fell. The gas ended at a state with lower T , lower P , and the original V . To return the gas exactly to State 1 from State 3, an additional process step would be required: heating the gas at constant volume from 243K back to 320K, which would also restore the pressure (because $PV = nRT$ and V is fixed). Different paths between the same volumes can leave the gas at very different states.

Making Sense of the Answer: *The volume returned to its starting value, but the gas's other state variables did not. Of the 1.84kJ of work the surroundings put into the gas during the isothermal compression, the gas immediately dumped 1.84kJ as heat to the heat reservoir; nothing was stored. Then in the adiabatic expansion, the gas spent some of its own internal energy doing work against the surroundings, and the temperature fell as a result. The two processes have completely different energy bookkeeping: the isothermal one trades work in for heat out and conserves U ; the adiabatic one trades U for work and conserves Q . Doing them in sequence does not undo the first with the second. State functions (V) returned to their starting values; path functions (Q and W along each step) did not, and the cumulative path leaves the gas elsewhere on the PV diagram.*

Think Like a Physicist: *The PV diagram is the most efficient way to see why this problem ends where it does. The isothermal step traces a hyperbola from State 1 to State 2 (going leftward, up and to the left, since volume halves and pressure doubles). The adiabatic step traces a steeper curve from State 2 back to volume V , but because the adiabat is steeper than the isotherm, it does not return to the same starting pressure when it reaches the original volume. It lands lower, at State 3 directly below State 1 on the diagram. The gap between State 1 and State 3, both at volume V , is exactly the difference between an isotherm and an adiabat through State 2. Examination questions often test this kind of cyclic-but-not-quite reasoning, because it forces the student to recognise that path-dependence applies to compound processes too: not just W and Q for one process, but also the entire net thermodynamic state at the end of a sequence.*

Four thermodynamic processes are now in hand: V fixed, P fixed, T fixed, and Q fixed. Each one pinned down a different state variable and let the others adjust; each one carries its own work formula and its own characteristic curve on the PV diagram. The next section assembles all four into a single mastery picture, with a master summary table that the student should learn to recognise and apply, and a synthesis figure showing all four processes radiating from a common starting point. The combined-process example just delivered is one species of multi-step problem; the next section systematises the technique for handling them.

MASTERING THE PV DIAGRAM

The four cornerstone thermodynamic processes are now in hand. Each one was met as a separate problem in its own section, with its own picture, its own derivation, and its own worked examples. This closing section pulls the four together. The aim is twofold. First, to assemble what has been learned into a single mastery picture and a single summary table that can be revisited at any moment to verify a half-remembered formula. Second, to graduate from individual processes to the multi-process problems that examiners actually set, in which the gas walks through two or three processes in sequence and the student must account for the total work and total heat across all of them.

Imagine a gas at some particular state, with a particular pressure, volume, and temperature. From that single starting point, the gas can be carried in any of four canonical directions, each fixing one of the variables. The picture below shows all four processes radiating from a common starting state A on the PV diagram. This is the master diagram of the chapter, worth keeping near to hand whenever a thermodynamics problem is being read.

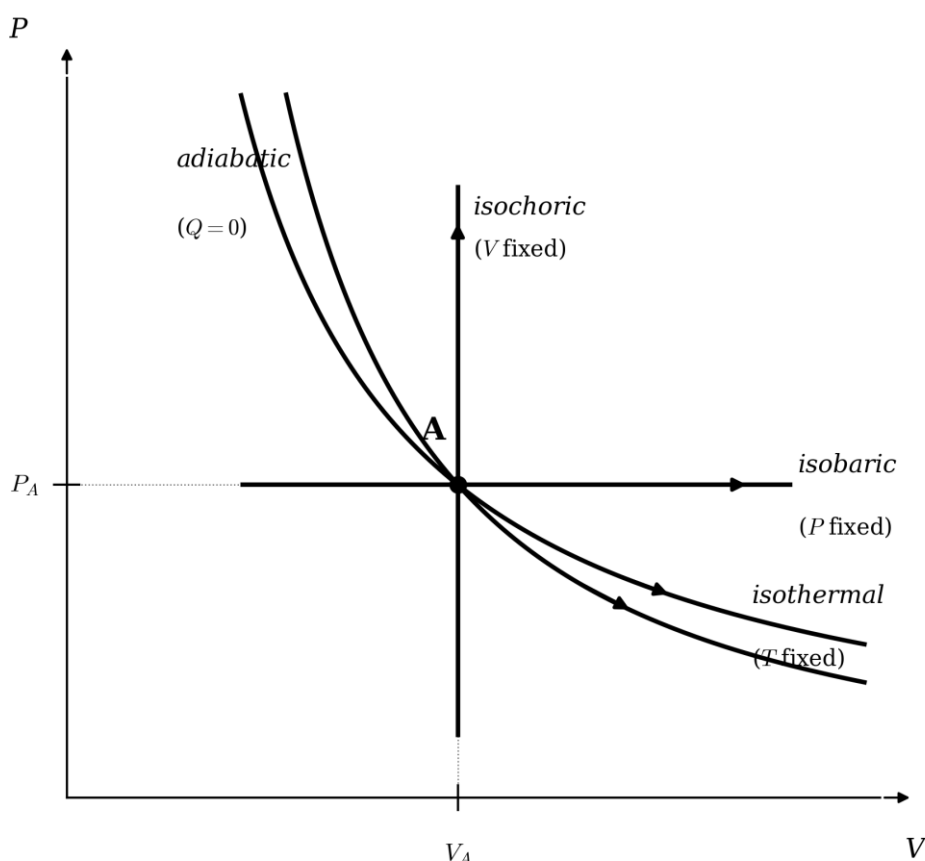


Figure: The four cornerstone thermodynamic processes drawn from a common starting point A on a PV diagram. The vertical line is the isochoric process (V fixed, only P and T change). The horizontal line is the isobaric process (P fixed, only V and T change). The gentler hyperbola is the isothermal process (T fixed, $PV = \text{constant}$). The steeper hyperbola is the adiabatic process ($Q = 0$, $PV^\gamma = \text{constant}$). At point A all four processes pass through the same state; from A onward they diverge. Knowing which way the gas's state moves on this diagram tells the student immediately which formula to reach for.

The master process summary table

The four formulae for Q , W , and ΔU across the four processes can be assembled into a single comparison table. Every entry is derivable from what was established in the four preceding sections; the table simply puts them side by side so the patterns become visible at a glance.

Process	Constant	Process equation	Q	W	ΔU
Isochoric	V	$\frac{P}{T} = \text{const}$	$nC_v\Delta T$	0	$nC_v\Delta T$
Isobaric	P	$\frac{V}{T} = \text{const}$	$nC_p\Delta T$	$P\Delta V = nR\Delta T$	$nC_v\Delta T$
Isothermal	T	$PV = \text{const}$	$= W$	$nRT \ln \frac{V_2}{V_1}$	0
Adiabatic	$Q = 0$	$PV^\gamma = \text{const}$	0	$\frac{P_1 V_1 - P_2 V_2}{\gamma - 1}$	$= -W$

Three patterns deserve highlighting.

- Two of the four processes have one zero entry each, but in different columns. Isochoric has $W = 0$, because no volume change means no work. Adiabatic has $Q = 0$, by definition. The remaining quantity in each case is computed from the first law: for the isochoric process, all the heat goes into internal energy, so $Q = \Delta U$; for the adiabatic process, all the work comes out of internal energy, so $W = -\Delta U$.
- Although the ΔU column shows different expressions for each process, the formula $\Delta U = nC_v\Delta T$ applies to every process of an ideal gas, because the internal energy depends only on temperature. The table simply chooses the most illuminating form for each row: $nC_v\Delta T$ for isochoric and isobaric, where the formula is the natural way to compute ΔU ; zero for isothermal, which is just $nC_v\Delta T$ with $\Delta T = 0$; and $-W$ for adiabatic, which is the first law with $Q = 0$. Whether the volume or pressure changes is irrelevant for ΔU ; only ΔT matters, and ΔT might be zero or might be derivable from W .
- The entries in the Q column simplify dramatically for two of the four. Isothermal Q equals W , and adiabatic Q is zero by definition. Each of these simplifications has been derived in its own section and should now be remembered as a feature of the table.

Closure of cycles

A particularly important application of the table is to thermodynamic cycles. A cycle is a sequence of processes that returns the gas to exactly the state at which it started. State variables P , V , T , and U are by definition unchanged at the end of a cycle, since the gas is back where it began. So:

$$\Delta U_{\text{cycle}} = 0,$$

regardless of how complicated the path was. The first law of dynamics for a complete cycle then gives:

$$Q_{\text{cycle}} = W_{\text{cycle}}$$

Net heat absorbed by the gas over the cycle equals net work done by the gas over the cycle. This is how every heat engine in the world operates: heat absorbed at one stage of the cycle, work delivered at another, and the gas returned to the start to repeat the procedure. The petrol engine in a car, the diesel engine in a daladala, and the gas turbine at Ubungo power station in Dar es Salaam all exploit exactly this closure. The numerical net work per cycle equals the area enclosed by the cycle on the PV diagram, traversed in its appropriate direction. A clockwise cycle encloses a positive area, meaning W is positive: the gas does net work on its surroundings, and the device is producing power. A counter-clockwise cycle encloses a negative area, meaning W is negative: the surroundings do net work on the gas, and the device is consuming power. Engines are clockwise; refrigerators and heat pumps are counter-clockwise. The direction of traversal alone tells us which.

Combined processes and the state-transfer rule

Many examination problems describe a sequence of two or more thermodynamic processes carried out in succession. A gas might be compressed isothermally and then heated isochorically, or expanded adiabatically and then cooled isobarically. The general principle for handling such combined-process problems is short and important enough to deserve its own statement.

The state-transfer rule: *In any sequence of thermodynamic processes, the final state of one process is the initial state of the next. The values of P , V , T , and n at the end of process 1 carry over unchanged as the starting values for process 2. The numbers do not reset; they propagate forward.*

This means a two-process problem is solved by treating the two processes in turn. Use the formulae appropriate to process 1 to find its final state. Take that final state as the initial state of process 2. Use the formulae appropriate to process 2 to find its final state. Sum any quantities that accumulate (work, heat) across the two processes; carry forward state variables (P , V , T , n) in their final values.

A common student error is to apply the initial conditions of process 1 to the start of process 2, effectively rewinding the gas to before process 1 took place. This must not be done. The state-transfer rule is the principle that prevents the error, and it must be invoked explicitly in every combined-process problem.

Real-world applications of the four processes

Real machines and natural processes rarely involve a single thermodynamic process in isolation. Most consist of sequences of processes in cycles, or approximations to single processes that can be analysed using the appropriate idealization. A few examples worth knowing.

1. Petrol engine (four-stroke)

The four strokes of a petrol engine map one-to-one onto the four cornerstone processes, in approximate form.

- **Intake stroke:** the piston descends and air-fuel mixture is drawn in at roughly atmospheric pressure (isobaric).
- **Compression stroke:** the piston rises rapidly, compressing the mixture before significant heat can leak through the cylinder walls (nearly adiabatic).
- **Combustion:** the spark plug fires and the fuel burns explosively while the piston is still near top dead centre, so the combustion happens at nearly constant volume (nearly isochoric).
- **Power stroke:** the hot high-pressure gas pushes the piston back down rapidly, expanding nearly adiabatically.
- **Exhaust stroke:** the burnt gas is pushed out at near-atmospheric pressure (isobaric).

All four cornerstone processes appear in a single cycle of the engine, in a sequence that has been refined and tuned by mechanical engineers for more than a century but whose underlying physics is exactly the four cases this chapter has built up.

2. Diesel engine

Differs from the petrol engine in one critical respect: there is no spark plug. The compression ratio is so high that the adiabatic compression of the air alone raises its temperature above the ignition point of diesel fuel. Fuel is then injected directly into the hot compressed air and ignites on contact, with no spark needed. The temperature rise during compression is computed using $TV^{\gamma-1} = \text{constant}$, which is a standard adiabatic calculation, and is the reason a diesel engine has its higher compression ratio than a petrol engine.

3. Refrigerator and air conditioner

These devices run a thermodynamic cycle in the opposite direction to an engine. A working fluid (typically a refrigerant gas) is compressed adiabatically by a motor-driven compressor, raising its temperature. The hot fluid then releases heat to the kitchen or outside air at near-constant pressure (isobaric). The fluid is then expanded back through a valve, cooling adiabatically. The cold fluid then absorbs heat from inside the refrigerator at near-constant pressure (isobaric). The cycle is counter-clockwise on the PV diagram, the work done on the gas is positive, and the device is a refrigerator (consuming electrical energy to move heat from the cold compartment to the warm room) rather than an engine (producing energy from a temperature difference).

4. Speed of sound in air

Sound waves are alternating compressions and rarefactions of air, with frequencies in the hundreds or thousands of hertz. At such high frequencies, each compression or rarefaction lasts only milliseconds, far too short for heat to flow significantly through the air during one cycle. Sound propagation in air is therefore an adiabatic, not isothermal, phenomenon, with the speed of sound given by the formula, $\sqrt{\gamma RT/M}$, where γ enters because the propagation is adiabatic. The fact that γ rather than 1 appears in this formula was first appreciated by Laplace, correcting an earlier calculation by Newton that had assumed isothermal propagation and produced a speed about fifteen percent too low.

5. Gas turbine power station

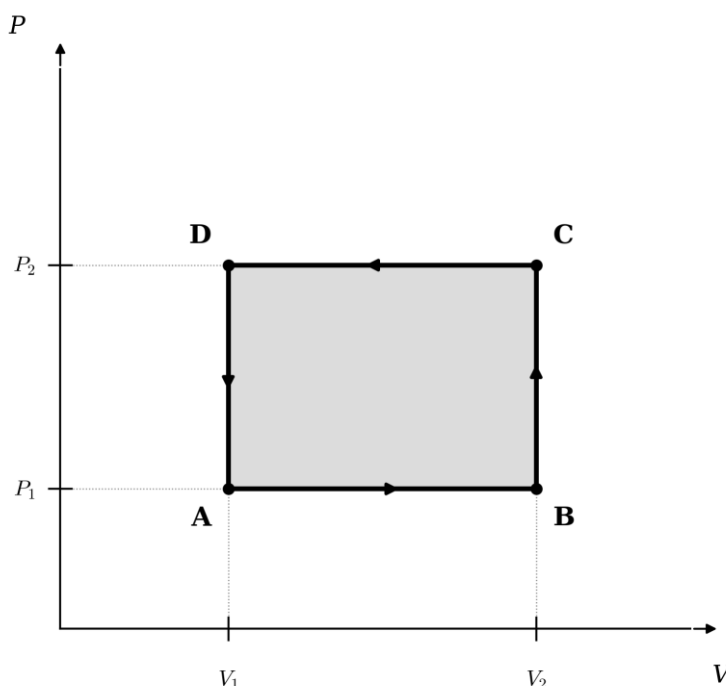
The gas turbines that supply Dar es Salaam with much of its electrical power, and the diesel and gas-fired backup units at Kidatu and elsewhere, exploit a Brayton-style cycle. Air is compressed in a multi-stage compressor (nearly adiabatic), heated in a combustion chamber by burning natural gas at constant pressure (isobaric), expanded through the turbine blades to drive the generator (nearly adiabatic), and finally exhausted to the atmosphere (isobaric). The cycle, drawn on a PV diagram, is a closed loop traversed clockwise, and its enclosed area is the net work delivered to the generator per cycle. The same physics that runs in a daladala engine runs, scaled up by orders of magnitude, in the turbine halls of every power station in the country.

The four processes have now been brought together, three examples now follow.

BINDER Example 32

A fixed amount of an ideal gas is taken around a rectangular cycle on a PV diagram, as shown in the figure below. The corners of the rectangle are at $A(V_1, P_1)$, $B(V_2, P_1)$, $C(V_2, P_2)$, and $D(V_1, P_2)$, with $V_1 = 1.0 \times 10^{-3} \text{m}^3$, $V_2 = 3.0 \times 10^{-3} \text{m}^3$, $P_1 = 1.0 \times 10^5 \text{Pa}$, and $P_2 = 2.0 \times 10^5 \text{Pa}$. The gas is taken around the cycle in the order $A \rightarrow B \rightarrow C \rightarrow D \rightarrow A$.

- Identify each leg of the cycle by the type of thermodynamic process it represents.
- Compute the net work done by the gas per cycle.
- State whether the net heat flow over the cycle is into or out of the gas.



Solution

- Identification of each leg:

$A \rightarrow B$. The gas moves from (V_1, P_1) to (V_2, P_1) , so the volume increases from V_1 to V_2 while the pressure stays fixed at P_1 . This is an **isobaric expansion** at low pressure.

$B \rightarrow C$. The gas moves from (V_2, P_1) to (V_2, P_2) , so the pressure increases from P_1 to P_2 while the volume stays fixed at V_2 . This is an **isochoric pressure rise** at large volume.

$C \rightarrow D$. The gas moves from (V_2, P_2) to (V_1, P_2) , so the volume decreases from V_2 to V_1 while the pressure stays fixed at P_2 . This is an **isobaric compression** at high pressure.

$D \rightarrow A$. The gas moves from (V_1, P_2) to (V_1, P_1) , so the pressure decreases from P_2 to P_1 while the volume stays fixed at V_1 . This is an **isochoric pressure drop** at small volume.

- The two isochoric legs ($B \rightarrow C$ and $D \rightarrow A$) contribute zero work, because the volume does not change during them. Only the two isobaric legs do work.

On A→B, the gas expands at the low pressure P_1 :

$$W_{A \rightarrow B} = P_1(V_2 - V_1) = (1.0 \times 10^5 \text{ Pa})(3.0 \times 10^{-3} \text{ m}^3 - 1.0 \times 10^{-3} \text{ m}^3) = +200 \text{ J}$$

On C→D, the gas is compressed at the high pressure P_2 :

$$W_{C \rightarrow D} = P_2(V_1 - V_2) = (2.0 \times 10^5 \text{ Pa})(1.0 \times 10^{-3} \text{ m}^3 - 3.0 \times 10^{-3} \text{ m}^3) = -400 \text{ J}$$

The net work done by the gas per cycle is the sum of the four legs:

$$\begin{aligned} W_{\text{cycle}} &= W_{A \rightarrow B} + W_{B \rightarrow C} + W_{C \rightarrow D} + W_{D \rightarrow A} \\ &= +200 \text{ J} + 0 \text{ J} + (-400 \text{ J}) + 0 \text{ J} = -200 \text{ J} \end{aligned}$$

So the gas does -200 J of net work per cycle.

The negative sign tells us that the surroundings have done 200 J of net work on the gas over the complete cycle, not the other way round.

Cross-checking with the enclosed-area method: the rectangle has width $V_2 - V_1 = 2.0 \times 10^{-3} \text{ m}^3$ and height $P_2 - P_1 = 1.0 \times 10^5 \text{ Pa}$, so its area is $(2.0 \times 10^{-3} \text{ m}^3)(1.0 \times 10^5 \text{ Pa}) = 200 \text{ J}$. The magnitude matches. The sign is supplied by the direction of traversal. Tracing the arrows on the figure, the path A→B→C→D→A goes rightward at the bottom, up on the right, leftward at the top, down on the left. This is counter-clockwise on a standard PV diagram. Counter-clockwise traversal makes the net work negative, which agrees with the leg-by-leg calculation: $W_{\text{cycle}} = -200 \text{ J}$.

(c) By cycle closure, $Q_{\text{cycle}} = W_{\text{cycle}} = -200 \text{ J}$. Negative net heat means heat flows out of the gas over the complete cycle.

The gas absorbs heat during some legs and rejects more heat during others; the net is outward. Since work flows in (the surroundings drive the gas around the cycle) and heat flows out (the gas dumps heat to its surroundings), this cycle is a refrigeration-style device. It uses external work to move heat from the gas to the environment.

Making Sense of the Answer: Two-hundred joules of net heat exit the gas per cycle, and the same magnitude of work is supplied to the gas per cycle. This is the energy bookkeeping of a refrigeration-style device. The four legs do not all contribute work, and not all in the same direction; the net work is what survives after the four are summed. The leg-by-leg sum and the enclosed-area calculation agree on both the magnitude (200 J) and the sign (negative, by the counter-clockwise traversal). The lesson: the sign of net work in a cycle problem is determined not by the shape of the cycle but by the direction in which the cycle is traversed. Direction matters as much as area.

Think Like a Physicist: Two methods are available for net work in a cycle. Method one: leg-by-leg calculation, summing the work done during each individual process and using the appropriate formula for that process. This method is foolproof but slow. Method two: enclosed-area calculation, which gives the magnitude in one step (just the area on the PV diagram), with the sign supplied by the direction of traversal (clockwise positive, counter-clockwise negative). The two methods must agree, and a careful student uses one to compute and the other to check. When they disagree, as in the discussion above, the disagreement points to an error in identifying the direction, which a careful re-examination of the arrows resolves.

REAL Example 33

Kipute is studying a textbook diagram showing the working fluid of an air conditioner used in a Dar es Salaam supermarket cold-storage room, traced as a closed cycle on a PV diagram. The cycle is drawn in the counter-clockwise direction. Most engine cycles encountered in the chapter are drawn clockwise.

- Explain what the direction of traversal of a cycle on a PV diagram tells us about the net work done by the gas over the complete cycle.
- Explain why a heat engine cycle must be clockwise on a PV diagram and a refrigerator cycle must be counter-clockwise.
- State, with reason, the direction of net heat flow over the air conditioner's cycle.

Solution

(a) The net work done by the gas over a complete cycle equals the area enclosed by the cycle on the PV diagram, signed according to the direction of traversal. A clockwise cycle encloses a positive area; the net work done by the gas is positive, meaning the gas delivers work to its surroundings on balance. A counter-

clockwise cycle encloses a negative area; the net work done by the gas is negative, meaning the surroundings do work on the gas on balance. The direction of traversal therefore determines the sign of the net work; the area determines its magnitude.

(b) A heat engine exists to produce work from a temperature difference. The gas inside the engine must, on balance, be doing work on the surroundings, since that work is the engine's useful output (turning a turbine blade or driving a piston). The net work done by the gas must therefore be positive, and from part (a) this requires the cycle to be traversed clockwise. A refrigerator or air conditioner does the opposite. It pumps heat from a cold compartment to a warmer surrounding, against the natural flow direction, by consuming work supplied from outside (typically electrical work driving a compressor). The work flows into the gas, so the net work done by the gas is negative, and the cycle must be traversed counter-clockwise. The direction of the loop on the PV diagram therefore identifies the device unambiguously: clockwise for an engine, counter-clockwise for a refrigerator or heat pump.

(c) For the air conditioner, the cycle is counter-clockwise, so the net work done by the gas is negative. By cycle closure (ΔU over a complete cycle is zero, so $Q_{\text{cycle}} = W_{\text{cycle}}$), the net heat flow over the cycle is also negative. Negative net heat means heat flows out of the gas on balance over the cycle, and into the surroundings. This is exactly what an air conditioner must do: it absorbs heat from inside the cold-storage room (where the gas is at low temperature during one part of the cycle), and rejects a larger amount of heat to the warmer outside air (where the gas is at high temperature during another part of the cycle). The difference between the heat rejected and the heat absorbed equals the work supplied by the electrical compressor, again by the first law applied over the cycle.

Making Sense of the Answer: *The supermarket air conditioner is correctly drawn counter-clockwise because that is what a refrigerator must look like on a PV diagram. The cycle's direction is its identity card: engines run clockwise, refrigerators run counter-clockwise. Both involve the same four cornerstone processes (in various combinations), the same first law, and the same kind of PV diagram, but with the loop traversed in opposite directions. The energy bookkeeping over the cycle differs only in sign, and the sign is read from the direction of the arrows.*

Think Like a Physicist: *Direction of traversal tells you the sign of the net work (clockwise positive, counter-clockwise negative). The sign of the net work tells you the device type (positive for engine, negative for refrigerator or heat pump). The sign of the net work also tells you the sign of the net heat flow, by cycle closure. Three pieces of physics, all read off a single arrow on a single diagram. When an examination question shows a PV cycle and asks for the device type or the direction of net heat flow, the answer is found by glancing at the direction of the arrows, not by computing anything. Reading direction first, numbers second, is the habit this section is most concerned to install.*

HOT Example 34

2500mol of air ($\gamma = 1.4$) is cooled at constant pressure of 1.0atm from 35°C to 26°C. Find (a) the molar heat capacity at constant volume of the gas (b) the change in internal energy of the gas, (c) the work done by the gas, and (d) the heat supplied to the gas. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

Note that the sample is cooled, not heated, so the temperature change is negative:

$$\Delta T = 26^\circ\text{C} - 35^\circ\text{C} = -9\text{K}$$

(a) The molar heat capacity at constant volume:

$$C_v = \frac{R}{\gamma - 1} = \frac{8.314\text{Jmol}^{-1}\text{K}^{-1}}{1.4 - 1} = \frac{8.314\text{Jmol}^{-1}\text{K}^{-1}}{0.4} = 20.79\text{Jmol}^{-1}\text{K}^{-1}.$$

(b) The change in internal energy depends only on temperature for an ideal gas:

$$\Delta U = nC_v\Delta T = (2500\text{mol})(20.79\text{Jmol}^{-1}\text{K}^{-1})(-9\text{K}) = -467775\text{J} \approx -468\text{kJ}$$

Negative ΔU means the internal energy of the gas has fallen, which makes sense: the gas was cooled, so its molecular kinetic energy fell.

(c) The work done by the gas at constant pressure, using $W = nR\Delta T$:

$$W = nR\Delta T = (2500\text{mol})(8.314\text{Jmol}^{-1}\text{K}^{-1})(-9\text{K}) = -187065\text{J} \approx -187\text{kJ}$$

Negative W means the gas did not do positive work on its surroundings; instead, the surroundings did positive work on the gas as it contracted. This too makes sense: a cooling gas at constant pressure shrinks, and the atmosphere outside pushes it inward.

(d) The heat supplied to the gas, from the first law:

$$Q = \Delta U + W = (-468\text{kJ}) + (-187\text{kJ}) = -655\text{kJ}$$

Negative Q means heat did not flow into the gas; it flowed out of the gas. The cooling process delivered 655kJ of heat from the gas to the surroundings.

Verification using the alternative isobaric heat formula:

$$Q = nC_p\Delta T, \quad \text{where} \quad C_p = C_v + R = 20.79\text{Jmol}^{-1}\text{K}^{-1} + 8.314\text{Jmol}^{-1}\text{K}^{-1} = 29.10\text{Jmol}^{-1}\text{K}^{-1}.$$

$$Q = (2500\text{mol})(29.10\text{Jmol}^{-1}\text{K}^{-1})(-9\text{K}) = -654750\text{J} \approx -655\text{kJ}$$

Both routes agree, as the first law requires.

Making Sense of the Answer: *Three negative numbers tell a single coherent story. The gas cooled (ΔT negative), so its internal energy fell (ΔU negative). The gas contracted as it cooled (because pressure was held constant, so V is proportional to T by Charles's law), so the work done by the gas was negative (W negative). And heat was extracted from the gas (Q negative). All three signs follow from the single physical fact: the gas was cooled. The energy bookkeeping closes: 468kJ of heat removed comes from the falling internal energy, and an additional 187kJ comes from the work done by the surroundings as they squeezed the contracting gas inward. Total heat extracted: 655kJ. The two-route check (using C_p directly versus going through C_v and the first law) confirms the arithmetic.*

Think Like a Physicist: $nC_p \Delta T$ is the fastest route to Q for an isobaric process; $nC_v \Delta T$ plus $P\Delta V$ is the longer route via the first law. Both work, both should agree, and a careful student uses both: the first to compute, the second to check. The relation between the two is Mayer's equation, which is the reason C_p exceeds C_v by exactly R , and is the reason any isobaric problem can be cross-checked with this kind of two-route method. The 2500-mole sample size in this problem is realistic for a small industrial process; the magnitude of the heat (655kJ) sets the scale of how much cooling capacity the surrounding refrigeration system must deliver per such cooling event. Industrial thermodynamics works in exactly these scales.

This closes the systematic study of the four cornerstone thermodynamic processes and their applications in cycles. The first law of thermodynamics has been the constant companion of every problem in this chapter: it has told us how heat, work, and internal energy must balance in every process, and it has imposed the cycle-closure condition that makes engines and refrigerators work the way they do. The first law accounts for energy. But energy accounting alone does not tell the whole story of thermodynamics. It does not, for instance, tell us why heat flows from hot bodies to cold bodies and not the other way; energy would be conserved either way. To understand the directionality of natural processes, a second law is needed, and the next section introduces it. The second law is what gives the difference between possible and impossible processes its precise mathematical meaning, and it is the foundation on which the concept of entropy is built.

THE SECOND LAW OF THERMODYNAMICS: A BRIEF INTRODUCTION

The first law of thermodynamics has accompanied every problem in this chapter so far, telling us how heat, work, and internal energy must balance in any process. The first law is a law of **conservation**: energy is neither created nor destroyed; it merely changes form. But energy conservation alone does not capture everything thermodynamics has to say. Two facts about the natural world are not predicted, and not even forbidden, by the first law. **The first** is that *some processes happen spontaneously while their reverses, despite being equally allowed by energy conservation, never do*. **The second** is that *heat-engine efficiency has a ceiling that depends only on the temperatures of the hot and cold reservoirs, and not on the engineering skill of the designer*. To capture these facts, a second law of thermodynamics is needed. This closing section introduces that law qualitatively and applies it to the kind of efficiency calculations that examiners frequently set.

Place a hot cup of tea and a cold cup of water side by side on the classroom desk. Wait. After some minutes the tea has cooled, the water has warmed, and eventually both are at the same room temperature. Wait longer; nothing further happens. Now ask whether the reverse process could ever occur: a tepid cup at room temperature spontaneously dividing itself into a hotter portion and a colder portion, with no external agent acting on it. Energy conservation does not forbid such a thing. The total energy of the two cups before and

after would be identical; only the distribution of that energy across the two cups would change. The first law would happily permit it. And yet no one in the history of physics has ever observed it. Hot bodies cool, cold bodies warm, and the equilibration runs in only one direction. The reverse, though energetically permitted, is forbidden by something else.

That something else is the **second law of thermodynamics**. It captures the directionality of natural processes, the fact that some directions in the space of physical possibilities are open and the reverse directions are closed. It is one of the most far-reaching statements in physics, and although the full mathematical machinery of the second law belongs to a more advanced course, the qualitative content of the law and its main quantitative consequence (the efficiency ceiling for heat engines) can be stated and used at this level.

Two classical statements of the second law

Two equivalent statements of the second law have come down through the history of thermodynamics. The first is due to Lord Kelvin and Max Planck, the second to Rudolf Clausius. Each captures the same underlying truth, but each starts from a different observation.

Kelvin-Planck statement: *No heat engine operating in a cycle can convert all of the absorbed heat into useful work. Some of the absorbed heat must always be rejected to a cooler body.*

This statement comes from observing real heat engines. The petrol engine in a car, the gas turbine in a power station, and any other heat-using device that runs in a cycle, all reject some of the heat they absorb to a cooler reservoir, in the form of warm exhaust, hot cooling water, or radiated heat. None of them converts all of the input heat into work. Engineers have tried, for two centuries, to design an engine that does. None has succeeded, and the second law tells us none ever will.

Clausius statement: *Heat cannot flow spontaneously from a cooler body to a hotter body. To move heat from cold to hot, work must be supplied from outside.*

This statement comes from observing the natural direction of heat flow. Heat always flows from hot to cold without external help; the reverse never occurs without external help. A refrigerator does indeed move heat from a cold compartment to a warmer kitchen, but only because an electric motor consumes work to drive the compressor. Unplug the refrigerator and the heat flow reverses, leaking from the warm kitchen back into the cold compartment, the compartment warming up to room temperature exactly as the second law predicts.

Equivalence: The Kelvin-Planck and Clausius statements look superficially different. One is about engines; the other is about heat flow. But they can be shown to be logically equivalent: any device that violates one of them could be combined with a normal device to violate the other. A full proof of this equivalence is beyond the scope of this book; for our purposes here, it is enough to know that the two statements are different verbal expressions of the same underlying physical principle. Both are correct, both are useful, and which to invoke in a given problem depends on whether the problem is about engines (use Kelvin-Planck) or about heat flow (use Clausius).

Carnot's theorem and the efficiency ceiling

The most useful quantitative consequence of the second law for examination purposes is **Carnot's theorem**, named for the French engineer Sadi Carnot, who established it about 200 years ago in a remarkable monograph written when he was just 28 years old. The theorem states a strict ceiling on the efficiency of any heat engine, expressed in terms of only **two quantities**: *the temperature of the hot reservoir from which the engine absorbs heat, and the temperature of the cold reservoir to which it rejects heat.*

We define the **efficiency η** of a heat engine in the natural way, as the ratio of useful work output per cycle to heat input per cycle:

$$\eta = \frac{\text{useful work output per cycle}}{\text{heat input per cycle}}$$

This efficiency lies between 0 (a useless engine that produces no work) and 1 (a perfect engine that converts all heat to work). The first law alone allows any value in this range. The second law restricts the range further. **Carnot's theorem** says:

No heat engine operating between a hot reservoir at absolute temperature T_h and a cold reservoir at absolute temperature T_c can have an efficiency greater than the Carnot efficiency:

$$\eta_{\max} = 1 - \frac{T_c}{T_h},$$

where the temperatures must be expressed in kelvin. This bound is independent of the working substance, the design of the engine, and any other detail. It depends only on the two reservoir temperatures.

The ceiling has immediate practical consequences. Consider a typical petrol engine, in which the combustion temperature might reach about 800K and the surroundings sit at about 320K. The Carnot efficiency for these temperatures is:

$$\eta_{\max} = 1 - \frac{320\text{K}}{800\text{K}} = 1 - 0.40 = 0.60 = 60\%$$

So the absolute ceiling on the efficiency of any engine running between these temperatures is 60%. No matter how cleverly the engine is designed, the second law forbids it from doing better. Real petrol engines achieve about 25% to 30% in practice, well below the ceiling, because of friction, incomplete combustion, heat loss through the cylinder walls, and other practical inefficiencies. But the second law tells us that even a perfectly designed engine, with no friction or losses, could not exceed 60% in this temperature range. The ceiling is not an engineering target; it is a law of nature.

This is also the long-promised explanation for why a working engine gets hot. If 100% efficiency were possible, an engine could absorb its fuel's chemical energy and deliver every joule as useful work, leaving no heat behind. Real engines must reject 70% to 75% of the absorbed heat to the surroundings, and that rejected heat is what warms the engine block, the radiator water, and the air around the exhaust. The mechanic at every petrol station in the country who shrugs at a hot engine and says "a hot engine is a working engine" is, knowingly or not, invoking the second law. The first law alone could not have told him so.

Two laws are now in hand, three examples now follow.

BINDER Example 35

Identify which form of the second law of thermodynamics is most directly invoked by each of the following statements:

- (a) A refrigerator cannot work without a power supply.
- (b) No heat engine can be 100% efficient.

Solution

(a) "A refrigerator cannot work without a power supply" is most directly an instance of the **Clausius form**. A refrigerator's job is to move heat from a colder compartment to a warmer kitchen, against the natural direction of heat flow. The Clausius form says exactly that this transfer requires work input from outside, which is the power supply mentioned in the statement.

(b) "No heat engine can be 100% efficient" is most directly an instance of the **Kelvin-Planck form**, which says exactly that some heat must always be rejected to a cooler reservoir, so the work output cannot equal the full heat input.

Making Sense of the Answer: *The two forms of the second law specialize to two different real-world problem types. Refrigerators, heat pumps, and any device that moves heat against the natural direction land under the Clausius form. Engines, turbines, and any device that converts heat to work land under the Kelvin-Planck form. The two forms are equivalent in the sense that any violation of one would imply a violation of the other, but in practice, choosing the form that matches the problem at hand makes the reasoning quicker.*

Think Like a Physicist: *Both classical statements of the second law are statements of impossibility. The first law is a positive statement (energy is conserved). The second law, by contrast, is a series of negative statements (this cannot happen, that cannot happen, no engine can be 100% efficient, no heat can flow spontaneously from cold to hot). Statements of impossibility are sometimes more powerful than statements of possibility, because every observed instance of the forbidden process not occurring is evidence for the law. After two centuries of careful observation, the second law has accumulated more evidence than almost any other law in physics, and despite the simplicity of its forms, no exception has ever been found.*

REAL Example 36

A Dar es Salaam taxi driver tells a passenger that a friend of his has designed an engine that recovers 95% of the fuel's energy as useful work. The engine operates between a combustion temperature of 700K and an ambient temperature of 305K.

- Use Carnot's theorem to determine whether such an engine is possible.
- Find the correct theoretical maximum efficiency for an engine operating between these two temperatures, and state in physical terms what the result implies about the friend's claim.

Solution

(a) Carnot's theorem sets the upper bound on the efficiency of any heat engine operating between two reservoir temperatures. Computing that bound for $T_h = 700\text{K}$ and $T_c = 305\text{K}$:

$$\eta_{\max} = 1 - \frac{T_c}{T_h} = 1 - \frac{305\text{K}}{700\text{K}} = 1 - 0.4357 \approx 0.564 = 56.4\%$$

So the maximum theoretical efficiency between these temperatures is about 56.4%. The friend's claim of 95% therefore exceeds this ceiling by a substantial margin. By Carnot's theorem, no heat engine operating between 700K and 305K can have an efficiency above 56.4%, no matter how cleverly it is designed. The friend's claim is therefore impossible.

(b) The correct theoretical maximum efficiency is 56.4%, as computed in part (a). In physical terms, this means that for every 100J of heat absorbed by the engine from its hot reservoir at 700K, at most 56.4J can emerge as useful work, and at least 43.6J must be rejected to the cold reservoir at 305K as waste heat. Any engine that claims to deliver more than 56.4J of work per 100J of absorbed heat between these two reservoirs is claiming to violate the second law of thermodynamics, which is one of the most well-tested laws in all of physics. The friend's claim is therefore impossible regardless of the design details.

Making Sense of the Answer: A 95% claim is not just slightly too high; it is too high by a factor that makes the device physically impossible, not merely difficult or inefficient. The second law is not an engineering challenge to be overcome with better machining or better fuel; it is a fundamental statement about what nature permits. A claim that exceeds the Carnot bound by even a percent or two would warrant scepticism; a claim that exceeds it by 38 percentage points warrants flat refusal. Real petrol-engine efficiencies in this temperature range run around 25% to 30%, well below the 56% Carnot ceiling, because of friction, incomplete combustion, and other practical inefficiencies. There is room for engineers to improve real engines toward the Carnot ceiling, but no room above it.

Think Like a Physicist: Whenever a problem presents an efficiency claim and two reservoir temperatures, the first move is to compute the Carnot bound. If the claim exceeds the bound, the device is impossible, full stop. If the claim falls below the bound, the device is in principle permitted, though it may still be difficult or impractical to build. The Carnot bound is the most economical possible filter for engine claims: one calculation, one comparison, and the verdict is in. Beware of impressive-sounding claims that do not respect the bound. Real engineering improvements work below the ceiling, never above it.

HOT Example 37

A thermal power station in Tanzania burns heavy fuel oil. The combustion temperature in the boiler is 900K, and the cooling tower rejects waste heat to the surroundings at an ambient temperature of 310K.

- Calculate the maximum theoretical efficiency of the power station.
- The actual electrical efficiency achieved by the station is 38%. Calculate the fraction of fuel energy lost as waste heat, and explain in physical terms where this waste energy actually goes.
- Does the operation of this power station violate the first law of thermodynamics? Does it violate the second law? Explain each answer briefly.

Solution

(a) The maximum theoretical efficiency, by Carnot's theorem with $T_h = 900\text{K}$ and $T_c = 310\text{K}$:

$$\eta_{\max} = 1 - \frac{T_c}{T_h} = 1 - \frac{310\text{K}}{900\text{K}} = 1 - 0.3444 \approx 0.656,$$

or about 65.6%.

This is the absolute ceiling on the efficiency of any heat engine, including this power station, that operates between these two temperatures.

(b) The actual efficiency is 38%, meaning that 38% of the fuel's combustion energy emerges as useful electrical work. The remaining 62% is lost as waste heat:

$$\text{fraction lost} = 1 - 0.38 = 0.62 = 62\%.$$

Physically, this 62% of fuel energy is rejected to the atmosphere through the cooling tower, where it appears as warm water vapour rising from the tower and warm air around the station. None of this energy has been destroyed; the first law forbids that. It has been **degraded** from useful, high-temperature thermal energy (concentrated in the 900K boiler steam) to less useful, low-temperature thermal energy (dispersed in the 310K cooling water and surrounding air). The amount of energy is conserved, but its quality, or its capacity to do further work, has fallen.

(c) **Whether the power station violates the first law:** The first law states that energy is conserved in every process. The power station takes in 100J of fuel energy per arbitrary unit input, delivers 38J as electrical work, and rejects 62J as waste heat. The total output, $38\text{J} + 62\text{J} = 100\text{J}$, equals the input. Energy is conserved exactly. The first law is not violated.

Whether the power station violates the second law: The second law, in the form of Carnot's theorem, sets the maximum possible efficiency between 900K and 310K at 65.6%. The actual efficiency, 38%, is below this ceiling. Therefore, the second law is not violated either. The power station operates within the bounds set by both laws, just well below the second-law ceiling because of practical inefficiencies (heat loss in pipes, friction in the turbine, incomplete combustion, and so on).

Making Sense of the Answer: *Three results emerge in tight agreement. The Carnot ceiling is 65.6%; the actual efficiency is 38%; the difference (27.6 percentage points) is the gap left by practical engineering, the room for improvement that engineers continually push against but cannot eliminate. The 62% of fuel energy not converted to electrical work is not destroyed; it is rejected as low-grade heat to the atmosphere. The cooling tower of every thermal power station in the country is the visible evidence of the second law: the plumes of warm vapour rising from the towers carry away the rejected heat that the second law insists must be discarded somewhere.*

Think Like a Physicist: *The first law is concerned with **quantities** of energy; the second law is concerned with the **quality**, or the capacity, of that energy to do useful work. The same hundred joules can be high-quality if concentrated at a high temperature, or low-quality if dispersed at a low temperature. A power station's job is to convert as much high-quality input energy to useful work as possible before it inevitably becomes low-quality waste heat. The Carnot bound expresses the unavoidable price the second law extracts from this conversion, in proportion to the temperature ratio of cold to hot reservoirs. The closer the cold reservoir is to absolute zero, or the hotter the hot reservoir, the smaller this price. Practical engines aim for both: the highest combustion temperature the materials can stand, and the coldest cooling water available.*

This is where the chapter, and our introduction to thermodynamics, closes. The first law has been the constant companion of every problem: it has accounted for the conservation of energy in every process, from the simplest heating of a gas at constant volume to the most complicated multi-stage cycle. The second law, introduced briefly here, has supplied the directionality and the efficiency ceiling that the first law alone cannot deliver. Together, the two laws form the foundation of all classical thermodynamics, and although a full mathematical treatment of the second law lies in more advanced courses, the verbal statements and the Carnot bound that have been presented here are sufficient to answer every NECTA examination question on the second law that has been set in recent years. The physics of heat, work, and energy is a vast subject, and what has been built up in this chapter is only the entrance. But it is an entrance that opens directly onto every engine, refrigerator, and power station in the country, and on every cup of tea that cools toward room temperature.

The chapter's ten sections have each held one idea at a time. The miscellaneous worked examples that follow no longer have that courtesy. They will combine ideas across sections, and sometimes across chapters, in exactly the way examiners prefer.

MISCELLANEOUS WORKED EXAMPLES ON THERMODYNAMICS

Example 38

- (a) Explain in molecular terms why a polyatomic ideal gas requires more heat than a monatomic ideal gas to undergo the same temperature rise at constant volume, even though both samples contain the same number of moles.
- (b) 0.50 mol of an ideal monatomic gas absorbs 200J of heat. Calculate the temperature rise of the gas if (i) the process takes place at constant volume, (ii) the process takes place at constant pressure. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

(a) A monatomic molecule, having no internal structure, can store kinetic energy only in three independent translational degrees of freedom. A polyatomic molecule, in addition to its three translational modes, can store kinetic energy in rotational modes about its molecular axes and in vibrational modes along its bonds. The equipartition theorem assigns an energy of $\frac{1}{2}kT$ to every accessible degree of freedom per molecule, so a polyatomic molecule must absorb energy not only into its translational motion but also into its rotational and vibrational modes before its temperature rises by the same amount. With more energy stores to fill per molecule, more heat is required for the same temperature rise.

(b)(i) At constant volume, all of the heat raises the internal energy of the gas, and for a monatomic gas $Q = nC_v\Delta T$ with $C_v = \frac{3}{2}R$.

$$\Delta T = \frac{Q}{nC_v} = \frac{200\text{J}}{0.50\text{mol} \times \frac{3}{2} \times 8.314\text{Jmol}^{-1}\text{K}^{-1}} = 32.1\text{K}$$

(ii) At constant pressure, $Q = nC_p\Delta T$, with $C_p = C_v + R = \frac{3}{2}R + R = \frac{5}{2}R$ for a monatomic gas.

$$\Delta T = \frac{Q}{nC_p} = \frac{200\text{J}}{0.50\text{mol} \times \frac{5}{2} \times 8.314\text{Jmol}^{-1}\text{K}^{-1}} = 19.2\text{K}$$

The temperature rise is approximately 32.1K at constant volume and 19.2K at constant pressure.

Example 39

- (a) Distinguish, by reference to the first law of thermodynamics, between the work done by an ideal gas during an isothermal expansion and the work done during an adiabatic expansion when both processes start from the same initial state and produce the same volume change. Account for which of the two processes delivers the larger amount of work, and explain why.
- (b) 2.0 mol of an ideal gas ($\gamma = 1.4$) initially at 350K occupies a volume of 0.020m³ and expands to a volume of 0.040m³. Calculate the work done by the gas if (i) the expansion is isothermal at 350K, (ii) the expansion is adiabatic. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

(a) In an isothermal expansion the temperature is held constant, so the internal energy of an ideal gas does not change, and the first law $\Delta U = Q - W$ reduces to $Q = W$; every joule of heat that flows in is delivered as work. In an adiabatic expansion no heat enters the gas, so $\Delta U = -W$, and the work is supplied entirely from the gas's own internal energy. As the gas does work adiabatically its internal energy and therefore its temperature fall, the pressure drops faster than the corresponding isotherm, and the area under the adiabatic PV-curve is smaller than the area under the isothermal curve over the same volume range. The isothermal expansion therefore delivers more work than the adiabatic expansion between the same initial state and the same final volume.

(b)(i) For the isothermal expansion at constant $T = 350\text{K}$, the work done by the gas is $W = nRT \ln V_2/V_1$.

$$W = nRT \ln \frac{V_2}{V_1} = 2.0\text{mol} \times 8.314\text{Jmol}^{-1}\text{K}^{-1} \times 350\text{K} \times \ln \frac{0.040\text{m}^3}{0.020\text{m}^3} = 4034\text{J}$$

(ii) For the adiabatic expansion, $TV^{\gamma-1} = \text{constant}$, gives the final temperature.

$$T_2 = T_1 \left(\frac{V_1}{V_2} \right)^{\gamma-1} = 350\text{K} \times \left(\frac{0.020}{0.040} \right)^{0.4} = 265.3\text{K}$$

Since $Q = 0$, the work done by the gas equals the loss of internal energy:

$$W = nC_v(T_1 - T_2) \text{ with } C_v = R/(\gamma - 1) = \frac{5}{2}R$$

$$W = nC_v(T_1 - T_2) = 2.0\text{mol} \times \frac{5}{2} \times 8.314\text{Jmol}^{-1}\text{K}^{-1} \times (350\text{K} - 265.3\text{K}) = 3520\text{J}$$

The work done is approximately 4034J in the isothermal expansion and 3520J in the adiabatic expansion, and the isothermal value is the larger of the two as predicted.

Example 40

- (a) When a small mass of water at 100°C is fully vapourised at atmospheric pressure, only a small fraction of the heat supplied appears as work done by the steam against the atmosphere, while the larger fraction is stored as an increase in internal energy of the water-steam system. Account for this partition of energy in physical terms.
- (b) 5.0g of water at 100°C is fully vapourised at atmospheric pressure ($1.013 \times 10^5\text{Pa}$). Calculate the work done by the steam as it forms, the increase in internal energy of the water-steam system, and the total heat supplied. Take the specific latent heat of vapourisation of water as $L_v = 2.26 \times 10^6\text{Jkg}^{-1}$, the density of liquid water as $\rho_l = 1000\text{kgm}^{-3}$, and the density of steam at 100°C and atmospheric pressure as $\rho_s = 0.598\text{kgm}^{-3}$.

Solution

(a) In the liquid state the water molecules sit close together, held by strong intermolecular (attractive) forces, so liquid water has a small specific volume. In the vapour state the molecules are far apart and the gas occupies a large specific volume. Vapourising water therefore requires energy on two distinct accounts. First, the molecules must be pulled apart against their mutual attraction, a process that increases the internal energy of the system without producing any external mechanical effect. Second, the new vapour pushes back the surrounding atmosphere as it expands into the space it now needs, doing mechanical work on the surroundings. Because the intermolecular attractions in liquid water are strong relative to the modest pressure of the atmosphere, the energy needed to overcome these attractions is much larger than the work needed to displace the atmosphere; the partition is therefore heavily weighted towards internal energy.

(b) Volume of the liquid: $V_l = m/\rho_l = 5.0 \times 10^{-3}\text{kg} / 1000\text{kgm}^{-3} = 5.0 \times 10^{-6}\text{m}^3$. Volume of the steam at $1.013 \times 10^5\text{Pa}$ and 100°C : $V_s = m/\rho_s = 5.0 \times 10^{-3}\text{kg} / 0.598\text{kgm}^{-3} = 8.36 \times 10^{-3}\text{m}^3$. The work done by the steam against the atmosphere as it forms is $W = P\Delta V$ at constant atmospheric pressure.

$$W = P(V_s - V_l) = 1.013 \times 10^5\text{Pa} \times (8.36 \times 10^{-3}\text{m}^3 - 5.0 \times 10^{-6}\text{m}^3) = 846\text{J}$$

The total heat supplied is $Q = mL_v$, by the definition of the specific latent heat of vapourisation at constant pressure.

$$Q = mL_v = 5.0 \times 10^{-3}\text{kg} \times 2.26 \times 10^6\text{Jkg}^{-1} = 11300\text{J}$$

From the first law of thermodynamics, the increase in internal energy is $\Delta U = Q - W$.

$$\Delta U = Q - W = 11300\text{J} - 846\text{J} = 10454\text{J}$$

The work done by the steam against the atmosphere is approximately 846J, the increase in internal energy is approximately 10454J, and the total heat supplied is approximately 11300J. Of the heat supplied, only about 7.5% emerges as external work; the remaining 92.5% is stored as internal energy.

Example 41

- (a) Justify, by reference to the equipartition of energy, why a polyatomic ideal gas with seven accessible degrees of freedom (three translational, three rotational, and one vibrational) has a value of γ closer to unity than that of a monatomic ideal gas. Determine the value of γ for such a polyatomic gas.
- (b) 3.0 mol of the polyatomic gas described above is contained at a temperature of 400K and undergoes an isothermal expansion in which its volume doubles. Calculate the work done by the gas during this expansion. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Solution

(a) By the equipartition theorem each accessible degree of freedom stores an average energy of $\frac{1}{2}kT$ per molecule, so the molar internal energy of an ideal gas with f accessible degrees of freedom is $U = \frac{f}{2}RT$, the molar heat capacity at constant volume is $C_v = \frac{f}{2}R$, and the molar heat capacity at constant pressure is $C_p = C_v + R = \frac{f+2}{2}R$. The ratio of principal heat capacities is therefore $\gamma = C_p/C_v = (f+2)/f$. For $f = 3$ (monatomic) this gives $\gamma = 5/3 \approx 1.67$. For $f = 7$ (polyatomic with three translational, three rotational, and one vibrational degree of freedom) this gives $\gamma = 9/7 \approx 1.29$, a value closer to unity. As f rises, the additional energy stores dilute the rise in temperature per joule of heat absorbed, the difference between C_p and C_v becomes a smaller fraction of either, and γ approaches one.

$$\gamma = \frac{f+2}{f} = \frac{7+2}{7} = \frac{9}{7} \approx 1.29$$

(b) For an isothermal expansion the work done by the gas is $W = nRT \ln V_2/V_1$, independent of γ .

$$W = nRT \ln \frac{V_2}{V_1} = 3.0 \text{ mol} \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1} \times 400 \text{ K} \times \ln 2 = 6915 \text{ J}$$

The value of γ for the polyatomic gas is $9/7 \approx 1.29$, and the work done in the isothermal expansion is approximately 6915J.

Example 42

- (a) Identify the three distinct contributions to the total heat input when a fixed mass of liquid water in an open container is heated from below room temperature, brought to its boiling point, partially vapourised, and the new steam is allowed to expand against atmospheric pressure as it forms.
- (b) 0.50kg of water in an open container is heated from 25°C and partially boiled away at atmospheric pressure ($1.013 \times 10^5 \text{ Pa}$) until 0.30kg of liquid water remains. Calculate: (i) the total heat supplied to the system, (ii) increase in internal energy; given the specific heat capacity of water $c = 4200 \text{ J kg}^{-1} \text{ K}^{-1}$, the specific latent heat of vaporisation $L_v = 2.26 \times 10^6 \text{ J kg}^{-1}$, and the density of steam at 100°C and 1 atmosphere $\rho_s = 0.598 \text{ kg m}^{-3}$. Neglect the heat capacity of the container.

Solution

(a) The **first** contribution is the *sensible heat* needed to raise the temperature of the whole liquid mass from 25°C to its boiling point of 100°C, given by $mc\Delta T$. The **second** contribution is the energy that increases the internal energy of the vapourised portion, *breaking the intermolecular attractions* in the liquid as molecules escape into the vapour phase. The **third** contribution is the **mechanical work** done by the new steam as it forms and pushes back the surrounding atmosphere by an amount equal to the volume of the new vapour. The standard latent heat of vapourisation, quoted at constant atmospheric pressure, packages the second and third contributions into a single coefficient, so the total heat supplied is conveniently written as the sum of the sensible heat and the latent heat term.

(b) Sensible heat to raise the liquid from 25°C to 100°C, with $m = 0.50 \text{ kg}$, $c = 4200 \text{ J kg}^{-1} \text{ K}^{-1}$ and $\Delta T = 75 \text{ K}$:

$$Q_1 = mc\Delta T = 0.50 \text{ kg} \times 4200 \text{ J kg}^{-1} \text{ K}^{-1} \times 75 \text{ K} = 157500 \text{ J}$$

Mass vapourised: $\Delta m = 0.50 \text{ kg} - 0.30 \text{ kg} = 0.20 \text{ kg}$. Latent heat associated with the vaporisation:

$$Q_2 = \Delta m \times L_v = 0.20 \text{ kg} \times 2.26 \times 10^6 \text{ J kg}^{-1} = 4.52 \times 10^5 \text{ J}$$

Within Q_2 , the work done by the new steam against the atmosphere is $W = P\Delta V$, where ΔV is the volume occupied by the steam after it forms, $V = \Delta m/\rho_s$.

$$W = P \times \frac{\Delta m}{\rho_s} = 1.013 \times 10^5 \text{ Pa} \times \frac{0.20 \text{ kg}}{0.598 \text{ kg m}^{-3}} = 33880 \text{ J}$$

The remaining portion of Q_2 , namely $Q_2 - W = 4.52 \times 10^5 \text{ J} - 3.39 \times 10^4 \text{ J} = 4.18 \times 10^5 \text{ J}$, is stored as the **internal energy increase** associated with breaking the intermolecular attractions during vapourisation. The **total heat supplied** is the sum of the sensible heat and the latent heat term:

$$Q_{\text{total}} = Q_1 + Q_2 = 1.575 \times 10^5 \text{ J} + 4.52 \times 10^5 \text{ J} = 6.10 \times 10^5 \text{ J}$$

The **total heat** supplied to the system is approximately $6.10 \times 10^5 \text{ J}$, of which, $1.575 \times 10^5 \text{ J}$ ($\approx 26\%$) raises the temperature of the liquid, $4.18 \times 10^5 \text{ J}$ ($\approx 68\%$) increases the **internal energy** of the new vapour, and $3.39 \times 10^4 \text{ J}$ ($\approx 6\%$) is delivered as work against the atmosphere.

Example 43

- (a) Explain why, when heat is supplied to a gas in a sealed rigid container of metal walls in contact with a hotter reservoir, the rate at which the internal energy of the gas rises depends on both the thermal conductivity of the container walls and the cross-sectional area available for conduction.
- (b) A sealed rigid steel container of internal volume 0.020 m^3 contains 0.80 mol of monatomic ideal gas, initially at 27°C . One end of the container, of cross-sectional area 0.010 m^2 and wall thickness 5.0 mm , is in steady contact with a steam reservoir at 100°C , while the rest of the container surface is perfectly insulated. The thermal conductivity of steel is $k = 50 \text{ W m}^{-1} \text{ K}^{-1}$. Calculate the initial rate at which the temperature of the gas rises with time. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Solution

(a) Heat reaches the gas only through the conducting wall, and Fourier's law fixes the rate at which energy crosses that wall as $dQ/dt = kA\Delta T/L$, where k is the thermal conductivity of the wall material, A is the cross-sectional area through which heat flows, ΔT is the temperature difference across the wall, and L is the wall thickness. Increasing k means each square metre of wall passes heat more readily; increasing A increases the total area available for heat passage; and both raise the rate at which energy enters the gas. Because the container is rigid the gas does no work, so by the first law every joule arriving in the gas is delivered to its internal energy, and the rate of temperature rise dT/dt is proportional to dQ/dt . The two material parameters of the wall therefore appear directly in the rate at which the gas heats up.

(b) Temperature difference across the wall: $\Delta T = 373 \text{ K} - 300 \text{ K} = 73 \text{ K}$. Initial rate of heat conduction into the gas:

$$\frac{dQ}{dt} = \frac{kA\Delta T}{L} = \frac{50 \text{ W m}^{-1} \text{ K}^{-1} \times 0.010 \text{ m}^2 \times 73 \text{ K}}{5.0 \times 10^{-3} \text{ m}} = 7300 \text{ W}$$

At constant volume the heat raises only the internal energy: $dQ = nC_v dT$, with $C_v = \frac{3}{2}R$ for the monatomic gas, hence:

$$\frac{dT}{dt} = \frac{1}{nC_v} \times \frac{dQ}{dt} = \frac{7300 \text{ W}}{0.80 \text{ mol} \times \frac{3}{2} \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1}} = 732 \text{ K s}^{-1}$$

The initial rate at which the temperature of the gas rises is approximately 732 K s^{-1} .

The rate is large because the small quantity of monatomic gas has a small heat capacity while the steel wall is thin and conducts well.

Example 44

- (a) For an ideal gas undergoing a quasi-static adiabatic process, derive the relation $\frac{dT}{T} + (\gamma - 1)\frac{dV}{V} = 0$.
- (b) 2.0 mol of an ideal gas ($\gamma = 1.4$) is initially at 320 K and 0.050 m^3 . Using the relation derived in part (a), calculate the temperature of the gas when its volume has decreased to 0.020 m^3 during a quasi-static adiabatic compression.

Solution

(a) In an adiabatic process $Q = 0$, so the first law $dU = dQ - dW$ for an infinitesimal step reduces to $dU = -dW$. For a quasi-static process the work done by the gas in expanding by dV against its own pressure is $dW = PdV$, and the internal energy of an ideal gas depends only on temperature, with $dU = nC_v dT$. Combining these gives $nC_v dT = -PdV$. Substituting $P = nRT/V$ from the ideal gas equation and dividing through by nT yields:

$$\frac{dT}{T} = -\frac{R}{C_v} \times \frac{dV}{V}$$

The combination R/C_v equals $(C_p - C_v)/C_v = \gamma - 1$, so $dT/T = -(\gamma - 1)dV/V$. Rearranging gives the required result:

$$\frac{dT}{T} + (\gamma - 1) \frac{dV}{V} = 0$$

(b) Integrating the differential relation between the initial state (T_1, V_1) and the final state (T_2, V_2):

$$\begin{aligned} \int_{T_1}^{T_2} \frac{dT}{T} &= -(\gamma - 1) \int_{V_1}^{V_2} \frac{dV}{V} \\ \ln \frac{T_2}{T_1} &= -(\gamma - 1) \ln \frac{V_2}{V_1} \\ \ln \frac{T_2}{T_1} &= \ln \left(\frac{V_2}{V_1} \right)^{-(\gamma-1)} = \ln \left(\frac{V_1}{V_2} \right)^{(\gamma-1)} \\ \frac{T_2}{T_1} &= \left(\frac{V_1}{V_2} \right)^{(\gamma-1)} \end{aligned}$$

Solving for T_2 with $T_1 = 320\text{K}$, $V_1 = 0.050\text{m}^3$, $V_2 = 0.020\text{m}^3$ and $\gamma = 1.4$:

$$T_2 = T_1 \left(\frac{V_1}{V_2} \right)^{\gamma-1} = 320\text{K} \times \left(\frac{0.050}{0.020} \right)^{0.4} = 462\text{K}$$

The temperature of the gas at the end of the adiabatic compression is approximately 462K.

Example 45

- (a) A refrigerator powered by an electrical supply is placed in a perfectly sealed and perfectly insulated room and is left running with its door open. Predict, with reason from the first law of thermodynamics, what happens to the average temperature of the room over time.
- (b) The refrigerator described in part (a) is rated at an electrical input of 150W. Estimate the rate at which the temperature of the air in the room rises if the room has internal dimensions $5.0\text{m} \times 4.0\text{m} \times 3.0\text{m}$ and contains air of density $\rho = 1.2\text{kgm}^{-3}$ and specific heat capacity at constant volume $c_v = 720\text{Jkg}^{-1}\text{K}^{-1}$. Neglect the heat capacity of the refrigerator and of the contents of the room.

Solution

(a) Treating the room and the refrigerator together as a single closed and insulated system, the only energy crossing the boundary of the system is the electrical energy supplied through the cable. The first law applied to the whole system therefore gives $\Delta U_{\text{system}} = W_{\text{electrical}}$, so the entire electrical input goes into the internal energy of the air in the room. The refrigerator does not destroy any energy by removing heat from its interior; it only relocates heat, taking it from inside the cabinet (which the open door now exposes to the room) and delivering it, together with the work it consumes, back into the same room. The average temperature of the room therefore rises steadily as the compressor runs.

(b) Volume of the room: $V = 5.0\text{m} \times 4.0\text{m} \times 3.0\text{m} = 60\text{m}^3$.

Mass of air: $m = \rho V = 1.2\text{kgm}^{-3} \times 60\text{m}^3 = 72\text{kg}$.

The rate at which the air absorbs heat at constant volume equals the electrical power input:

$$\begin{aligned} mc_v \frac{dT}{dt} &= P_{\text{elec}} \\ \frac{dT}{dt} &= \frac{P_{\text{elec}}}{mc_v} = \frac{150\text{W}}{72\text{kg} \times 720\text{Jkg}^{-1}\text{K}^{-1}} = 2.89 \times 10^{-3}\text{Ks}^{-1} \end{aligned}$$

The temperature of the room rises at approximately $2.89 \times 10^{-3}\text{Ks}^{-1}$.

This is about 0.17K every minute or 10K every hour, sufficient to be uncomfortable within an hour and dangerous within a working day.

Example 46

- (a) Show, with reasons, that the net heat absorbed by an ideal gas during one complete cycle of any closed thermodynamic process is necessarily equal to the net work done by the gas during that cycle.
- (b) 1.5 mol of an ideal gas ($\gamma = 1.4$) executes the following cycle. Starting at state A with $P = 1.0 \times 10^5\text{Pa}$, $V = 0.040\text{m}^3$, $T = 320\text{K}$, the gas undergoes (i) an isobaric expansion to $V = 0.080\text{m}^3$, (ii) an

isochoric cooling until the pressure has fallen to $0.50 \times 10^5 \text{ Pa}$, (iii) an isothermal compression back to state A. Calculate the net work done by the gas during the complete cycle. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Solution

(a) For any cycle the gas returns to its starting state, so every state variable, including the internal energy, returns to its initial value and $\Delta U_{\text{cycle}} = 0$. The first law for the complete cycle therefore reduces to $Q_{\text{net}} = W_{\text{net}}$: every joule of net heat absorbed during the cycle is delivered as net work, and the two are necessarily equal whatever the shape of the cycle, whatever the working substance, and whatever the sequence of processes. The equality is a direct consequence of internal energy being a state function.

(b) The three legs of the cycle. Leg (i) isobaric $A \rightarrow B$ at $P = 1.0 \times 10^5 \text{ Pa}$, expansion from $V = 0.040 \text{ m}^3$ to $V = 0.080 \text{ m}^3$:

$$W_1 = P\Delta V = 1.0 \times 10^5 \text{ Pa} \times (0.080 \text{ m}^3 - 0.040 \text{ m}^3) = 4000 \text{ J}$$

At B the temperature has risen to $T_B = T_A \times V_B/V_A = 320 \text{ K} \times 2 = 640 \text{ K}$.

Leg (ii) isochoric $B \rightarrow C$ at $V = 0.080 \text{ m}^3$, cooling until $P = 0.50 \times 10^5 \text{ Pa}$: no volume change, so $W_2 = 0$, and $T_C = T_B \times P_C/P_B = 640 \text{ K} \times 0.50 = 320 \text{ K}$.

The gas at C is therefore at the same temperature as at A, so leg (iii) is the isothermal compression at 320 K from $V_C = 0.080 \text{ m}^3$ to $V_A = 0.040 \text{ m}^3$:

$$W_3 = nRT \ln \frac{V_A}{V_C} = 1.5 \text{ mol} \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1} \times 320 \text{ K} \times \ln \frac{0.040}{0.080} = -2766 \text{ J}$$

The negative sign indicates that work is done on the gas during the compression. The net work for the complete cycle is the sum of the three:

$$W_{\text{net}} = W_1 + W_2 + W_3 = 4000 \text{ J} + 0 + (-2766 \text{ J}) = 1234 \text{ J}$$

The net work done by the gas during the complete cycle is approximately 1234 J (and by the cycle closure rule this is also the net heat absorbed during the cycle).

Example 47

- (a) In industrial gas-compression machinery a gas is often compressed in two adiabatic stages, with isobaric cooling between the stages, rather than in a single adiabatic stage. Identify the physical reason engineers prefer the two-stage arrangement, and account for it in terms of the work that must be supplied to the compressor.
- (b) 0.50 mol of an ideal gas ($\gamma = 1.4$) initially at $1.0 \times 10^5 \text{ Pa}$ and 300 K is compressed in two stages. Stage 1: adiabatic compression until the volume halves. Stage 2: isobaric cooling at the new pressure until the temperature returns to 300 K . Calculate the total work done on the gas during the two stages combined, and the heat rejected during the isobaric cooling stage. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Solution

(a) A single-stage adiabatic compression heats the gas substantially as it does work on it, so the gas is then at a high temperature and high pressure. The work needed to push it the rest of the way to a target final pressure is large because the pressure is already large. Cooling the gas back to its starting temperature between stages, at the elevated intermediate pressure, removes the temperature rise without changing the pressure, so the second adiabatic stage starts from a colder gas at the same pressure and ends at the target final pressure with a smaller total temperature rise and a smaller swept volume. The two areas under the two adiabats on the PV diagram together cover less area than the single adiabat to the same final pressure, so the integrated work supplied to the compressor is reduced. The intercooling step rejects to the surroundings the heat that the first-stage compression deposited in the gas, and that heat would otherwise have to be paid for as additional compressor work in the single-stage arrangement.

(b) Initial volume of the gas: $V_1 = \frac{nRT_1}{P_1} = \frac{0.50 \text{ mol} \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1} \times 300 \text{ K}}{1.0 \times 10^5 \text{ Pa}} = 1.247 \times 10^{-2} \text{ m}^3$.

Stage 1, adiabatic compression to $V_2 = \frac{V_1}{2} = 6.236 \times 10^{-3} \text{ m}^3$:

$$T_2 = T_1 \left(\frac{V_1}{V_2} \right)^{\gamma-1} = 300 \text{ K} \times 2^{0.4} = 396 \text{ K}$$

$$P_2 = P_1 \left(\frac{V_1}{V_2} \right)^\gamma = 1.0 \times 10^5 \text{ Pa} \times 2^{1.4} = 2.64 \times 10^5 \text{ Pa}$$

Work done on the gas in the adiabatic stage equals the increase in internal energy, with $C_v = \frac{R}{\gamma-1} = \frac{5}{2}R$:

$$W_{1,\text{on}} = nC_v(T_2 - T_1) = 0.50 \text{ mol} \times \frac{5}{2} \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1} \times (396 \text{ K} - 300 \text{ K}) = 996 \text{ J}$$

Stage 2 is isobaric cooling at P_2 from $T_2 = 396 \text{ K}$ back to 300 K .

The volume falls from $V_2 = 6.236 \times 10^{-3} \text{ m}^3$ to $V_3 = V_2 \times \frac{T_3}{T_2} = 6.236 \times 10^{-3} \times \frac{300}{396} = 4.726 \times 10^{-3} \text{ m}^3$.

Work done on the gas:

$$W_{2,\text{on}} = -P_2(V_3 - V_2) = -2.64 \times 10^5 \text{ Pa} \times (4.726 - 6.236) \times 10^{-3} \text{ m}^3 = 398 \text{ J}$$

Total work done on the gas in the two stages combined:

$$W_{\text{on,total}} = W_{1,\text{on}} + W_{2,\text{on}} = 996 \text{ J} + 398 \text{ J} = 1394 \text{ J}$$

Heat rejected during the isobaric cooling, $Q_{\text{rejected}} = nC_p(T_2 - T_3)$ with $C_p = \frac{7}{2}R$:

$$Q_{\text{rejected}} = nC_p(T_2 - T_3) = 0.50 \text{ mol} \times \frac{7}{2} \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1} \times (396 \text{ K} - 300 \text{ K}) = 1395 \text{ J}$$

The total work done on the gas during the two stages is approximately 1394 J and the heat rejected during the isobaric cooling stage is approximately 1395 J.

Example 48

- (a) A piston of mass m and cross-sectional area A rests on top of a fixed mass of gas in a vertical cylinder open to the atmosphere above and is free to slide vertically. By a force-balance argument on the piston, derive the pressure of the gas inside the cylinder in terms of the atmospheric pressure, the piston's weight and the area, and explain why this gas pressure is independent of the piston's vertical position, provided the piston is in mechanical equilibrium.
- (b) A vertical cylinder of internal cross-sectional area $A = 0.020 \text{ m}^2$ is closed below and contains 0.10 mol of an ideal gas. A piston of mass $m = 5.0 \text{ kg}$ rests on top of the gas, free to slide vertically. The atmospheric pressure is $P_{\text{atm}} = 1.013 \times 10^5 \text{ Pa}$ and gravity is $g = 9.81 \text{ ms}^{-2}$. The gas is heated and the piston rises by $\Delta h = 0.10 \text{ m}$. Calculate the work done by the gas during this process, identifying separately the work done against the atmosphere and the work done in raising the piston. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Solution

(a) A free piston in mechanical equilibrium has zero net vertical force on it. The downward forces are its weight mg and the atmospheric pressure P_{atm} acting over its top surface, while the upward force is the gas pressure acting over its bottom surface. Equating these on a piston of cross-sectional area A gives:

$$P_{\text{gas}}A = P_{\text{atm}}A + mg, \text{ hence } P_{\text{gas}} = P_{\text{atm}} + \frac{mg}{A}$$

The result depends only on the piston's weight, the atmospheric pressure and the area, none of which involves the piston's vertical position; the gas pressure is therefore the same whether the piston sits low in the cylinder or high in it.

(b) Pressure of the gas in the cylinder:

$$P_{\text{gas}} = P_{\text{atm}} + \frac{mg}{A} = 1.013 \times 10^5 \text{ Pa} + \frac{5.0 \text{ kg} \times 9.81 \text{ ms}^{-2}}{0.020 \text{ m}^2} = 1.038 \times 10^5 \text{ Pa}$$

Volume swept by the piston: $\Delta V = A\Delta h = 0.020 \text{ m}^2 \times 0.10 \text{ m} = 2.0 \times 10^{-3} \text{ m}^3$.

Total work done by the gas at constant gas pressure:

$$W_{\text{total}} = P_{\text{gas}}\Delta V = 1.038 \times 10^5 \text{ Pa} \times 2.0 \times 10^{-3} \text{ m}^3 = 207.5 \text{ J}$$

Work done by the gas against the atmosphere:

$$W_{\text{atm}} = P_{\text{atm}}\Delta V = 1.013 \times 10^5 \text{Pa} \times 2.0 \times 10^{-3} \text{m}^3 = 202.6 \text{J}$$

Work done by the gas in raising the piston (equal to the gain in gravitational potential energy of the piston):

$$W_{\text{piston}} = mg\Delta h = 5.0 \text{kg} \times 9.81 \text{ms}^{-2} \times 0.10 \text{m} = 4.91 \text{J}$$

The total work done by the gas is approximately 207.5J, of which approximately 202.6J pushes back the atmosphere and approximately 4.91J raises the piston, in the expected ratio $P_{\text{atm}}:mg/A$.

Example 49

- (a) The speed of sound in an ideal gas is given by $v = \sqrt{\gamma RT/M}$, where γ is the ratio of principal heat capacities. Account for the appearance of the factor γ , rather than the simpler factor 1, in this formula by reference to the time-scale on which compressions and rarefactions in a sound wave act on the gas.
- (b) Calculate the speed of sound in dry air at 27°C, treating air as a diatomic ideal gas with $\gamma = 1.4$ and effective molar mass $M = 29 \times 10^{-3} \text{kgmol}^{-1}$. Compare the result with the root mean square speed of air molecules at the same temperature, and comment on the ratio. Take $R = 8.314 \text{Jmol}^{-1} \text{K}^{-1}$.

Solution

(a) A sound wave is a sequence of compressions and rarefactions that pass through a gas at audio frequencies, sweeping any small region of gas through one pressure cycle in a time of order one millisecond. Heat conduction across the wavelength of an audible sound wave acts on a much longer time-scale, so essentially no heat flows in or out of any particular region during one cycle; the local compressions and rarefactions are therefore adiabatic rather than isothermal. The bulk modulus that controls the wave speed is consequently the adiabatic bulk modulus γP , not the isothermal value P , and in the formula $v = \sqrt{B/\rho}$ the factor γ appears. If the propagation were isothermal one would have $v = \sqrt{RT/M}$, the simpler formula with γ replaced by unity, but acoustic measurements unambiguously favour the larger adiabatic value, vindicating the kinetic-theory picture.

(b) Speed of sound at 300K:

$$v_{\text{sound}} = \sqrt{\frac{\gamma RT}{M}} = \sqrt{\frac{1.4 \times 8.314 \text{Jmol}^{-1} \text{K}^{-1} \times 300 \text{K}}{29 \times 10^{-3} \text{kgmol}^{-1}}} = 347 \text{ms}^{-1}$$

Root mean square speed at the same temperature:

$$v_{\text{rms}} = \sqrt{\frac{3RT}{M}} = \sqrt{\frac{3 \times 8.314 \text{Jmol}^{-1} \text{K}^{-1} \times 300 \text{K}}{29 \times 10^{-3} \text{kgmol}^{-1}}} = 508 \text{ms}^{-1}$$

Ratio of the two speeds:

$$\frac{v_{\text{sound}}}{v_{\text{rms}}} = \sqrt{\frac{\gamma}{3}} = \sqrt{\frac{1.4}{3}} = 0.683$$

The speed of sound in dry air at 27°C is approximately 347ms⁻¹, against an rms molecular speed of approximately 508ms⁻¹; the sound speed is about 68% of the rms molecular speed. The two are of the same order of magnitude because both are set by the same molecular kinetic energy scale, but they are not equal because the sound wave samples a typical molecular speed weighted by its ability to transfer momentum forward, not the bare rms speed.

Example 50

- (a) In a vigorous downstroke of a bicycle pump, the air trapped in the barrel is compressed rapidly. Justify why this compression can be modelled as adiabatic, and account for the physical reason the metal barrel of the pump becomes warm to the touch after a few strokes.
- (b) A cyclist pumps her bicycle tyre. Just before one downstroke, the air in the pump barrel occupies a volume of $4.0 \times 10^{-4} \text{m}^3$ at 27°C and atmospheric pressure ($1.013 \times 10^5 \text{Pa}$). After the downstroke, the air has been compressed to a volume of $1.0 \times 10^{-4} \text{m}^3$, and a thermometer in good thermal contact with the air reads 187°C immediately after. Treating the compression as adiabatic, calculate the value of γ for the air, and compare it with the value expected for a diatomic ideal gas.

Solution

(a) The downstroke takes a small fraction of a second, while heat conduction through the air to the metal barrel and out to the surroundings happens on a much longer time-scale, so during the stroke very little heat leaves the air. The compression is therefore approximately adiabatic, and the work done on the gas by the cyclist appears as an increase in internal energy, raising the temperature. After the stroke the now-hot air slowly transfers heat to the cooler metal barrel by conduction; since this transfer is repeated stroke after stroke, the barrel accumulates the deposited heat and becomes noticeably warm to the touch.

(b) For an adiabatic process $T_1 V_1^{\gamma-1} = T_2 V_2^{\gamma-1}$. Taking logarithms and rearranging gives:

$$\gamma - 1 = \frac{\log T_2/T_1}{\log V_1/V_2}$$

With $T_1 = 300\text{K}$, $T_2 = 460\text{K}$, $V_1 = 4.0 \times 10^{-4}\text{m}^3$ and $V_2 = 1.0 \times 10^{-4}\text{m}^3$:

$$\gamma - 1 = \frac{\log 460/300}{\log 4.0/1.0} = 0.308$$

$$\gamma = 1 + 0.308 = 1.31$$

The deduced value of γ is approximately 1.31, lower than the textbook value of 1.40 for an ideal diatomic gas. The discrepancy is consistent with the compression not being perfectly adiabatic; a small leak of heat from the gas to the metal barrel during the stroke pulls the apparent γ down towards the isothermal limit of unity.

Example 51

- (a) An ideal gas occupies one half of a thermally insulated container of total volume $1.0 \times 10^{-2}\text{m}^3$. The other half is initially evacuated. The partition between the two halves is suddenly removed, and the gas freely fills the whole container. Show, using the first law of thermodynamics, that the temperature of the gas does not change during this process, even though its volume doubles. Account for why the equation $PV^\gamma = \text{constant}$ does **not** apply here.
- (b) 0.20 mol of an ideal gas ($\gamma = 1.4$) is initially confined at 27°C and $1.0 \times 10^5\text{Pa}$ to one half of the $1.0 \times 10^{-2}\text{m}^3$ thermally insulated container described in part (a), with the other half evacuated. The partition is suddenly removed. Calculate (i) the final temperature of the gas, (ii) the final pressure of the gas, (iii) the change in internal energy of the gas.

Solution

(a) The container is thermally insulated, so no heat enters or leaves the gas, $Q = 0$. The other half of the container is evacuated, so the gas does no work on anything as it expands into the vacuum, $W = 0$. The first law $\Delta U = Q - W$ therefore gives $\Delta U = 0$. For an ideal gas the internal energy depends only on temperature, so $\Delta U = 0$ implies $\Delta T = 0$ and the temperature is unchanged at 300K . The relation $PV^\gamma = \text{constant}$ is derived under the assumption of a quasi-static adiabatic process, in which the gas is in equilibrium at every step and does work PdV at every step. Joule's free expansion is adiabatic but not quasi-static and does no work, so the derivation does not apply, and the gas does not follow an adiabat on the PV diagram; it simply jumps from its initial state to its final state along a horizontal isotherm.

(b)(i) Final temperature, by the argument of part (a):

$$T_f = T_i = 300\text{K}$$

(ii) For an ideal gas at constant temperature, PV is constant, so $P_f V_f = P_i V_i$, with $V_f = 2V_i$:

$$P_f = P_i \times \frac{V_i}{V_f} = 1.0 \times 10^5\text{Pa} \times \frac{V_i}{2V_i} = 5.0 \times 10^4\text{Pa}$$

(iii) Change in internal energy:

$$\Delta U = nC_v \Delta T = nC_v \times 0 = 0$$

The final temperature of the gas is 300K , the final pressure is $5.0 \times 10^4\text{Pa}$, and the change in internal energy is zero, all in agreement with the qualitative argument from the first law.

Example 52

- (a) A closed cycle on a PV diagram consists of two curves of one type joined by two curves of a different type, and the cycle is traversed clockwise. Identify what the two types of curve must be in order for the

cycle to operate as a Carnot engine, and justify your identification by reference to the requirement that all heat absorbed and rejected by a Carnot engine occurs at exactly two temperatures.

- (b) A Carnot engine operates between a hot reservoir at 500K and a cold reservoir at 300K. During one complete cycle, the engine absorbs 1500J of heat from the hot reservoir. Calculate (i) the efficiency of the engine, (ii) the work output per cycle, (iii) the heat rejected to the cold reservoir per cycle.

Solution

(a) A Carnot engine, by Carnot's construction, exchanges heat with the surroundings only at exactly two temperatures, the hot-reservoir temperature T_h and the cold-reservoir temperature T_c . Heat absorption from the hot reservoir at constant T_h occurs along an isothermal expansion at T_h ; heat rejection to the cold reservoir at constant T_c occurs along an isothermal compression at T_c . The two isotherms are joined by processes during which no heat is exchanged with anything, that is, by adiabatic legs: one expansion taking the gas from T_h down to T_c and one compression returning it from T_c up to T_h . The Carnot cycle therefore consists of two isothermal curves joined by two adiabatic curves, and these are the only curves consistent with heat exchange occurring at exactly two temperatures.

(b)(i) Efficiency of the Carnot engine:

$$\eta = 1 - \frac{T_c}{T_h} = 1 - \frac{300\text{K}}{500\text{K}} = 0.40$$

(ii) Work output per cycle:

$$W = \eta \times Q_h = 0.40 \times 1500\text{J} = 600\text{J}$$

(iii) Heat rejected per cycle, by the first law for the cycle:

$$Q_c = Q_h - W = 1500\text{J} - 600\text{J} = 900\text{J}$$

The efficiency of the Carnot engine is 40%, the work output per cycle is 600J, and the heat rejected to the cold reservoir per cycle is 900J.

Example 53

- (a) An ideal gas undergoes a process in which its temperature varies with volume according to $T = T_0 + \alpha V^2$, where T_0 and α are positive constants and V is the volume of the gas. Determine, in terms of T_0 , α and the universal gas constant R , the volume at which the pressure of one mole of the gas reaches its minimum value, and derive an expression for the corresponding minimum pressure.
- (b) For one mole of helium gas with $T_0 = 250\text{K}$ and $\alpha = 1.2 \times 10^6 \text{K m}^{-6}$, calculate (i) the volume at which the pressure of the gas is a minimum, (ii) the minimum pressure attained, (iii) the temperature of the gas at the minimum pressure. Take $R = 8.314 \text{J mol}^{-1} \text{K}^{-1}$.

Solution

(a) For one mole the ideal gas equation gives $PV = RT$, hence $P = RT/V$. Substituting $T = T_0 + \alpha V^2$ yields P as an explicit function of V :

$$P(V) = \frac{R(T_0 + \alpha V^2)}{V} = \frac{RT_0}{V} + R\alpha V$$

The first term decreases with V while the second increases with V , so $P(V)$ has a single minimum. Setting $dP/dV = 0$ to locate it:

$$\frac{dP}{dV} = -\frac{RT_0}{V^2} + R\alpha = 0$$

$$V = \sqrt{\frac{T_0}{\alpha}}$$

Substituting V back into $P(V)$ gives the minimum pressure:

$$P_{\min} = \frac{RT_0}{V} + R\alpha V = R\sqrt{\alpha T_0} + R\sqrt{\alpha T_0} = 2R\sqrt{\alpha T_0}$$

(b)(i) With $T_0 = 250\text{K}$ and $\alpha = 1.2 \times 10^6 \text{K m}^{-6}$:

$$V_{\min} = \sqrt{\frac{250\text{K}}{1.2 \times 10^6 \text{Km}^{-6}}} = \sqrt{2.083 \times 10^{-4} \text{m}^6} = 1.44 \times 10^{-2} \text{m}^3$$

(ii) Minimum pressure:

$$P_{\min} = 2R\sqrt{\alpha T_0} = 2 \times 8.314 \text{Jmol}^{-1} \text{K}^{-1} \times \sqrt{1.2 \times 10^6 \text{Km}^{-6} \times 250\text{K}} = 2.88 \times 10^5 \text{Pa}$$

(iii) At $P_{\min}$ the temperature is $T = T_0 + \alpha V^2 = T_0 + \alpha(T_0/\alpha) = 2T_0$

$$2T_0 = 2 \times 250\text{K} = 500\text{K}$$

The minimum pressure of the helium occurs at a volume of approximately $1.44 \times 10^{-2} \text{m}^3$, where the pressure is approximately $2.88 \times 10^5 \text{Pa}$ and the temperature is exactly 500K .

Sixteen miscellaneous worked examples lie behind you. The Digging Deeper questions that follow are now yours to attempt.

DIGGING DEEPER EXERCISE 5

EXERCISE 5A: BINDER QUESTIONS

Question 1

Two students write the first law of thermodynamics in different forms. One writes $\Delta U = Q - W$; the other writes $\Delta U = Q + W$. Each insists that the other is wrong. Resolve the disagreement by identifying what W means in each convention, and account for why the two equations are equivalent statements of the same physics.

Question 2

During an isothermal compression of an ideal gas, heat must leave the gas if its temperature is to remain constant. A student concludes from this fact that the internal energy of the gas must therefore decrease. Identify the error in the student's reasoning, and account for what becomes of the work that the surroundings do on the gas during the compression.

Question 3

Two different processes carry an ideal gas from the same initial state A to the same final state B along different paths on a PV diagram. Account for the fact that the change in internal energy ΔU is identical for both processes but the work done W and the heat exchanged Q can differ between them.

Question 4

Heat-capacity tables quote a single value of specific heat capacity for solids and liquids, but two distinct molar heat capacities (at constant volume and at constant pressure) for gases. A student concludes that the inequality $C_p > C_v$ is a feature peculiar to gases and does not exist for solids or liquids. Justify the student's observation by accounting for what changes, and what does not change, when energy is added to each phase of matter at constant pressure.

Question 5

Two ideal gases A and B are observed in separate experiments to have the same molar heat capacity at constant volume C_v . A student concludes from this single piece of evidence that the two gases must be the same substance, or at least be molecules of the same type. Evaluate this conclusion: justify what equal C_v values do, and do not, tell you about A and B at the molecular level, and identify what additional measurement would let you distinguish A from B.

Question 6

A welded steel pressure vessel sits in a Dar es Salaam yard. Overnight the temperature drops from 32°C to 22°C. State, with reason from the equation of state of an ideal gas, whether the pressure inside the vessel falls by exactly 10/305 of its initial value, by more than 10/305, or by less than 10/305, and identify the assumption that underlies your conclusion.

Question 7

Two identical balloons are filled, one with helium and the other with ordinary air, both at the same temperature and at atmospheric pressure. The helium balloon is found to be much lighter than the air balloon. Account, by reference to Avogadro's law, for why the lighter helium balloon nevertheless contains the same number of molecules as the heavier air balloon, and identify the molecular property responsible for the mass difference.

Question 8

A student claims that, since the temperature of an ideal gas does not change during an isothermal process, no heat can flow into or out of the gas. Identify the error in this reasoning, and account for what the heat flow accomplishes if it is not raising the temperature.

Question 9

A student claims that any process happening rapidly must be adiabatic in the sense that the relation $PV^\gamma = \text{constant}$ holds throughout. Evaluate this claim, distinguishing carefully between the condition $Q \approx 0$ and the further condition that the process is quasi-static, and identify which condition must be added to 'happens rapidly' for the relation $PV^\gamma = \text{constant}$ to apply.

Question 10

On a single PV diagram, sketch an isothermal compression and an adiabatic compression of the same ideal gas, both starting from the same initial state and reducing the volume by the same factor. Mark the two end-states clearly, indicate which of the two compressions reaches the higher final pressure, and account for the molecular reason behind the difference.

Question 11

A closed cycle on a PV diagram is traversed in the counter-clockwise sense. From this single piece of information alone, identify with reason: (a) the sign of the net work done by the gas over the complete cycle, (b) the sign of the net heat absorbed by the gas over the complete cycle, (c) whether the device is functioning as a heat engine or as a refrigerator.

Question 12

A heat pump installed to warm a Dar es Salaam house during cool July weather is observed to deliver, for each joule of electrical energy it consumes from the grid, more than one joule of heat to the interior of the house. Account, by reference to the first law of thermodynamics applied to the heat pump's working substance, for why this is not a violation of energy conservation.

Question 13

A student argues that the Kelvin-Planck and Clausius statements of the second law of thermodynamics cannot both be correct, because the first speaks about heat engines while the second speaks about the spontaneous flow of heat, and the two sound completely unrelated. Resolve the student's confusion by identifying the single underlying physical principle that both statements express, and account for why a violation of either statement would necessarily imply a violation of the other.

EXERCISE 5B: REAL QUESTIONS**Question 14**

A factory technician at an aluminium plant in Dar es Salaam observes that the compressed air emerging from the plant's industrial compressors feels noticeably hot to the touch. The technician reasons that *'energy is conserved, so the heat absorbed by the gas during the compression must equal the work done on it by the compressor.'* Identify the flaw in the technician's reasoning, and describe an everyday observation in industrial machinery that demonstrates the flaw directly.

Question 15

A boiler-plant operator in Mwanza, briefing a junior technician at the end of a shift, describes *'the gas doing 5000J of work'*, without specifying whether the work is done by the gas on the surroundings or by the surroundings on the gas. Account for why this ambiguity matters for first-law bookkeeping, and describe how a careful technician would record the same physics unambiguously.

Question 16

During a road accident, a sealed LPG cooking-gas cylinder is thrown into the wreckage of a burning vehicle. Account, by reference to the first law of thermodynamics, for why such cylinders are at risk of catastrophic explosion despite being designed for high-pressure operation. State the safety feature that manufacturers add to convert this risk from explosion to controlled release, and describe in one sentence how the feature works.

Question 17

A mother at home heats milk in an open saucepan and notices that the milk surface rises slightly as it warms, well before any boiling begins. Account for this rise in terms of an isobaric thermal expansion, and explain why the rise stops being noticeable once the milk reaches its boiling point.

Question 18

A bicycle pump becomes warm to the touch after vigorous use, even though no flame or external heater has been near it. A cyclist suggests that friction in the pump's seal is responsible for the warming. A physicist friend insists that this is at most a secondary effect and that the dominant cause is something else. Reconstruct the physicist's argument, and justify by appeal to direct observation why the friction explanation alone cannot account for the warming.

Question 19

Kipanga reads in a textbook that the indicator diagram of an internal combustion engine can be approximated by two adiabats and two isobars on a PV diagram, traversed clockwise. He asks **Mr. Akilikubwa** why all four cornerstone thermodynamic processes (isochoric, isobaric, isothermal, and adiabatic) are not used equally in the cycle, given that the chapter has spent so much effort on each. Provide the explanation Mr. Akilikubwa would give, framed in terms of the mechanical and thermal constraints of a real engine running at thousands of cycles per minute.

Question 20

A technician in Arusha claims to have built a 'self-powered fan' that runs on warm air alone, with no electrical input or fuel, by extracting energy from the heat of the surrounding atmosphere. Identify which form of the second law of thermodynamics this proposed device would violate, and account for why the device is impossible no matter how cleverly it is engineered.

Question 21

A government official in the energy ministry proposes building a coal-fired power station with a thermal efficiency of 80%, operating between a boiler at 500K and a cooling tower at 300K. Evaluate the proposal by reference to Carnot's theorem, and state with reason whether the project is feasible as proposed. If it is not, briefly suggest one direction in which the design parameters would have to move to make 80% theoretically achievable.

Question 22

An LPG cooking-gas cylinder used in a Dodoma household is moved from a shaded outdoor storage area at 25°C into a hot indoor kitchen at 40°C, with the valve closed throughout. The cylinder's safety pressure-release valve was calibrated to open when the internal pressure exceeds the rated limit, on the assumption that the cylinder is at 25°C. Account for the physical change inside the cylinder as it is brought indoors, and discuss whether the safety calibration is now conservative (over-cautious), exact, or risky.

Question 23

A factory in Tanga installs a refrigeration system to cool its server room. The factory engineer informs the manager that *'the refrigerator dumps more heat into the surroundings than it absorbs from the server room itself.'* A new accountant, suspicious of this claim, alleges that it violates the conservation of energy. Resolve the accountant's confusion by applying the first law of thermodynamics to the refrigeration cycle.

Question 24

A laboratory technician in Babati compresses 2.0mol of an ideal gas isothermally at 300K from a volume of 4.0L to a volume of 1.0L, and observes that approximately 6.9kJ of heat must be removed from the gas during the compression to keep its temperature constant. (a) State, with reason, whether the change in internal energy of the gas during the compression is positive, negative, or zero. (b) Account, by reference to the first law of thermodynamics, for why heat must be removed from the gas even though its temperature does not change. (c) Verify the magnitude of the heat removed by computing the work done on the gas. Take $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$.

Question 25

A pressure cooker boils food more rapidly than an open pan of water because food cooks faster at higher temperatures. Explain, by reference to the relationship between the saturation vapour pressure of water and its boiling temperature, why a sealed cooker enables higher cooking temperatures than an open pan, and account for the role of the cooker's safety valve in preventing the internal pressure (and hence the cooking temperature) from rising indefinitely.

EXERCISE 5C: HOT QUESTIONS

Useful constants: $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$, **atmospheric pressure** = $1.013 \times 10^5\text{Pa}$.

Question 26

An ideal monatomic gas is taken from state A at (P_0, V_0, T_0) to state B by two different paths. Path 1: isothermal expansion from A until the volume is $2V_0$, then isochoric heating until the pressure returns to P_0 . Path 2: isobaric expansion from A at pressure P_0 until the volume is $2V_0$. (a) Show that both paths terminate at the same final state B. (b) Verify that ΔU is the same along both paths. (c) Find the difference in heat absorbed between path 2 and path 1, expressed in terms of $P_0 V_0$.

Question 27

An ideal gas of molar mass 16g mol^{-1} is found to have $\gamma = 1.30$ at room temperature. Identify the most likely number of accessible degrees of freedom for the gas, name a real gas consistent with these properties, and justify the identification.

Question 28

An ideal gas at pressure P_0 , volume V_0 , and temperature T_0 undergoes a quasi-static adiabatic compression to volume $V_0/4$. The gas has $\gamma = 1.4$. Find the final pressure and the final temperature, both as multiples of the initial values, and find the work done on the gas as a multiple of $P_0 V_0$.

Question 29

The compression stroke of a four-stroke petrol engine takes about 1ms. Above a 5.0mm-thick aluminium piston of cross-section 0.005m^2 , 0.5g of gas with specific heat capacity $1000\text{J kg}^{-1}\text{K}^{-1}$ sits at a temperature 100K above the piston. Taking the thermal conductivity of aluminium as $230\text{W m}^{-1}\text{K}^{-1}$, estimate the time for heat to leak from the gas through the piston by conduction, and by comparing the two timescales justify whether the adiabatic approximation is appropriate for the compression stroke.

Question 30

A heat engine using an ideal monatomic gas operates on a rectangular cycle on a PV diagram with corners at $A(V_1, P_1)$, $B(V_1, P_2)$, $C(V_2, P_2)$, $D(V_2, P_1)$, traversed in the order $A \rightarrow B \rightarrow C \rightarrow D \rightarrow A$, with $V_1 = 1.0 \times 10^{-3}\text{m}^3$, $V_2 = 4.0 \times 10^{-3}\text{m}^3$, $P_1 = 1.0 \times 10^5\text{Pa}$ and $P_2 = 4.0 \times 10^5\text{Pa}$. Find the net work done per cycle, the heat absorbed per cycle, and the efficiency of the engine. Compare the efficiency with the Carnot ceiling between the two temperature extremes of the cycle.

Question 31

A diesel engine compresses air from atmospheric pressure and 27°C to one-sixteenth of its initial volume in a single quasi-static adiabatic compression stroke. The autoignition temperature of diesel fuel in compressed air is approximately 480°C . Verify that the compression alone raises the air's temperature above the autoignition temperature, and explain why this makes the diesel engine self-igniting without the spark plug needed in a petrol engine. Take $\gamma = 1.4$ for air.

Question 32

An ideal gas at 76cmHg and 27°C is contained in a cylinder fitted with a heavy frictionless piston. The piston is suddenly released and the gas expands adiabatically until its volume doubles. Taking $\gamma = 1.4$, find the final pressure of the gas, and the final temperature.

Question 33

A Carnot heat engine operates between two reservoirs. With the cold reservoir at 300K and the hot reservoir at 500K, the engine produces 200J of work per cycle. The cold reservoir is now cooled to 250K, with the hot reservoir unchanged, and the engine continues to absorb the same quantity of heat from the hot reservoir per cycle. Find the new work output per cycle, and the percentage increase in efficiency.

Question 34

An ideal gas with $\gamma = 1.4$, initially at $1.0 \times 10^5\text{Pa}$, 0.020m^3 , and 300K, undergoes a two-stage process. Stage 1: quasi-static adiabatic compression to one-quarter of its initial volume. Stage 2: isobaric cooling at the resulting pressure until the temperature returns to 300K. Find the total work done on the gas across both stages, the heat rejected during the cooling stage, and the net change in internal energy at the end.

Question 35

A vertical glass tube of uniform cross-section is closed at the bottom and open at the top. A column of mercury of length 0.20m is trapped above a column of air of length 0.40m at 27°C . Atmospheric pressure is 76cmHg. The tube is now slowly inverted so that it is closed at the top and open at the bottom, with the temperature held constant at 27°C . Find the new length of the trapped air column.

EXAMINATION QUESTIONS FOR PART ONE

Useful constants: Young's modulus of steel = $2.0 \times 10^{11} \text{ Pa}$, Young's modulus of copper = $1.1 \times 10^{11} \text{ Pa}$, Young's modulus of aluminium = $7.0 \times 10^{10} \text{ Pa}$, $c_{\text{water}} = 4200 \text{ J kg}^{-1} \text{ K}^{-1}$, $L_f(\text{ice}) = 3.36 \times 10^5 \text{ J kg}^{-1}$, $L_v(\text{water}) = 2.26 \times 10^6 \text{ J kg}^{-1}$, $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$, atmospheric pressure = $1.013 \times 10^5 \text{ Pa}$, $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$, Wien displacement constant $b = 2.898 \times 10^{-3} \text{ m K}$, $0^\circ \text{C} = 273 \text{ K}$, solar constant $S = 1370 \text{ W m}^{-2}$, $k_B = 1.381 \times 10^{-23} \text{ J K}^{-1}$, $N_A = 6.022 \times 10^{23} \text{ mol}^{-1}$

Question 1

- A student claims that the strain energy stored in a stretched wire is the kinetic energy of its atoms. Identify the error and account for the form in which the energy is actually stored.
- A daladala's rear leaf-spring suspension creaks under heavy passenger load but is silent when the vehicle is empty. Account for the difference, with reference to the elastic deformation of the leaf springs.
- A steel wire of length 2.0m and cross-sectional area $1.5 \times 10^{-6} \text{ m}^2$ is stretched by an axial force of 300N. Take Young's modulus of steel as $2.0 \times 10^{11} \text{ Pa}$. Calculate the extension of the wire and the elastic strain energy stored in it.

Question 2

- A student examining a stress-strain curve for a metal identifies the proportional limit and the elastic limit as the same point. Identify the error and account for the physical difference between the two limits.
- A bicycle pump used in a **Kigoma** shop, after months of daily use, develops a permanent slight bend in its piston rod even though no single push has snapped the rod. Account for the bend.
- A copper wire of yield stress $2.5 \times 10^8 \text{ Pa}$ and breaking stress $4.0 \times 10^8 \text{ Pa}$ has a circular cross-section of diameter 1.0mm. Calculate the maximum load the wire can support before plastic deformation begins, and the maximum load it can support before fracture.

Question 3

- Account for why the strain energy per unit volume of an elastic material is given by $\frac{1}{2} \times \text{stress} \times \text{strain}$, and not by $\text{stress} \times \text{strain}$ alone.
- A child in **Shinyanga** builds a slingshot using a thick rubber band. He notices that stretching the band to twice its initial extension launches the projectile to approximately four times the previous range, at the same angle. Account for this observation.
- A spring of force constant 200 N m^{-1} is stretched 0.30m from its natural length and then released, propelling a 0.20kg projectile horizontally from rest along a frictionless surface. Calculate the elastic potential energy stored in the spring and the launch velocity of the projectile.

Question 4

- For a fluid, distinguish between the isothermal bulk modulus and the adiabatic bulk modulus, and account for why the two values differ.
- The speed of sound in fresh water at room temperature is observed to be substantially higher than the speed of sound in air at the same temperature. Identify the physical property of water mainly responsible, and account for why this property is so much larger in water than in air.
- An ideal monatomic gas is at a pressure of $1.5 \times 10^5 \text{ Pa}$. Calculate (i) the isothermal bulk modulus of the gas, (ii) the adiabatic bulk modulus of the gas. Take $\gamma = 5/3$.

Question 5

- A bar of varying cross-section carries the same axial tensile force throughout its length. Account for why the strain in the bar is not the same everywhere, and identify where in the bar the strain is greatest.
- A glass tumbler dropped on a tiled floor in a kitchen typically shatters into many fragments, while a metal cup dropped from the same height suffers only a small dent. Account for this difference in behaviour.
- A steel rod of uniform cross-sectional area A_0 carries an axial tensile force F . A small circular hole drilled through the rod reduces its local cross-sectional area to $A_0/4$ at the hole. Treating the stress concentration factor at the edge of the hole as 3, calculate the factor by which the peak local stress at the hole exceeds the stress in the unaffected part of the rod.

Question 6

- (a) A cable made of n parallel wires of identical cross-section is loaded with a force F . A student claims that each wire carries the same load F because they share the load. Identify the error in the student's reasoning and state the correct sharing of the load among the n wires.
- (b) The lift cables of a passenger lift in a **Dar es Salaam** office building consist of multiple parallel steel wire ropes rather than a single thicker rope. Account for two engineering reasons (beyond the simple multiplication of cross-section) why a multi-rope arrangement is preferred.
- (c) A rigid horizontal beam is supported in equilibrium by three vertical wires of identical material and identical length, attached to the beam at equal spacing. The two outer wires each have cross-sectional area A ; the middle wire has cross-sectional area $2A$. A vertical load W is hung from the centre of the beam. Calculate the tension in each of the three wires.

Question 7

- (a) A small mass attached to the lower end of a vertical elastic wire and pulled downward from its equilibrium position oscillates after release. Identify the type of motion that results, and account for why the period of this motion depends on the elastic properties of the wire as well as the mass.
- (b) A bus driver in **Mtwara** notices a steady low-pitched hum from his bus's wheel-and-axle assembly at one particular cruising speed, but not at higher or lower speeds. Account for this resonance in terms of the elastic properties of the suspension.
- (c) A 2.0kg mass hangs in equilibrium from a vertical steel wire of length 1.5m and cross-sectional area $5.0 \times 10^{-7}\text{m}^2$. Take Young's modulus of steel as $2.0 \times 10^{11}\text{Pa}$. The mass is pulled a small distance below its equilibrium position and released. Calculate the period of the resulting vertical oscillation.

Question 8

- (a) A student plots stress against strain for a metal wire and finds that the curve becomes non-linear shortly after leaving the origin. Identify what physical limit has been passed, and account for why the slope of the straight portion alone gives the Young's modulus.
- (b) A factory in **Mwanza** tensile-tests every batch of incoming steel wire before installing it in a suspension bridge. Account for why a tensile test alone is sufficient to detect a sample of below-specification steel, but is not sufficient to confirm an above-specification sample as fully usable.
- (c) In a tensile test on a steel rod of original length 200mm and original diameter 10mm , the following data are recorded at the proportional limit: applied force 14kN ; extension 0.180mm . Calculate the Young's modulus of the steel.

Question 9

- (a) A student claims that a load suddenly released onto an elastic wire produces the same maximum extension as the same load placed onto the wire gradually. Identify the error and account for what happens to the wire in the moments after a sudden release.
- (b) A worker on a **Sumbawanga** construction site lowers a sack of cement onto a wire-supported hook by carefully releasing the sack at the level of the hook, rather than dropping it from above his shoulders. Account for why the dropping method is more likely to snap the supporting wire even though the cement weight is the same in both cases.
- (c) A 5.0kg load is released from rest at a height of 0.10m above the lower end of a vertical steel wire of original length 2.0m and cross-sectional area $1.0 \times 10^{-6}\text{m}^2$. Take Young's modulus of steel as $2.0 \times 10^{11}\text{Pa}$ and $g = 9.81\text{ms}^{-2}$. Calculate the maximum dynamic extension of the wire and the factor by which it exceeds the static extension produced by the same load.

Question 10

- (a) The fundamental frequency of a stretched wire fixed at both ends depends on its tension, length, and linear mass density. Account for why a wire of higher Young's modulus does not, by that fact alone, have a higher fundamental frequency than a similar wire of lower Young's modulus.

- (b) A guitarist tuning her instrument in **Zanzibar** finds that tightening the tuning peg of a string raises its pitch. Account for this behaviour by reference to the formula for the fundamental frequency of a stretched wire.
- (c) A steel wire of length 0.75m and linear mass density $8.0 \times 10^{-3} \text{kgm}^{-1}$ is stretched between two fixed pegs and is observed to vibrate at its fundamental frequency of 220Hz. Calculate the tension in the wire.

Question 11

- (a) A student claims that any physical property of a substance that varies with temperature can serve as the basis of a thermometer. Identify three additional requirements that the property must satisfy, and account for why each is needed.
- (b) A laboratory in **Mpanda** maintains a working thermometer for routine use and a reference thermometer kept locked in a cupboard, used only to recalibrate the working thermometer. Account for why the reference thermometer is kept specifically for this purpose rather than simply rotating two identical thermometers.
- (c) A platinum resistance thermometer has resistance 25.0Ω at the ice point and 35.5Ω at the steam point. In a constant-temperature bath, the resistance is 31.5Ω . Calculate the temperature of the bath on the platinum-resistance scale.

Question 12

- (a) Account for why the constant-volume gas thermometer is regarded as the near-ideal thermometer over a wide range of temperatures.
- (b) An industrial calibration laboratory in **Dar es Salaam** keeps a constant-volume gas thermometer as its primary reference standard, using it once or twice a year to recalibrate platinum resistance thermometers used for daily measurements. Account for why the gas thermometer, despite its bulk and inconvenience, is preferred as the primary reference.
- (c) A constant-volume gas thermometer has a pressure of $0.800 \times 10^5 \text{Pa}$ when its bulb is at the triple point of water (273.16K). When the bulb is placed in a heated bath, the pressure becomes $1.099 \times 10^5 \text{Pa}$. Calculate the temperature of the bath on the ideal-gas scale, in kelvin and in degrees Celsius.

Question 13

- (a) A student claims that since a platinum resistance thermometer gives the same reading as the ideal-gas scale at the ice point and at the steam point, the two scales must agree at every temperature in between. Identify the error in this claim, and account for the physical origin of any disagreement.
- (b) A power-station engineer uses a platinum resistance thermometer (rather than a mercury-in-glass thermometer) to monitor the temperature of superheated steam at about 400°C . Account for two reasons why the platinum thermometer is preferred for this application.
- (c) A platinum resistance thermometer has resistance 25.0Ω at the ice point and 35.6Ω at the steam point. In a heated bath, its resistance is 41.6Ω . A constant-volume gas thermometer placed in the same bath gives a temperature reading of 158.0°C . Calculate the temperature of the bath on the platinum-resistance scale, and the discrepancy between this value and the ideal-gas scale value.

Question 14

- (a) A thermocouple consists of two different metal wires joined at two junctions, one at a known reference temperature and the other at the unknown temperature to be measured. Account for the origin of the thermoelectric EMF generated between the two junctions, and identify why the cold junction is necessary.
- (b) A meat-processing factory in **Iringa** uses a thermocouple probe (rather than a glass thermometer) to measure the temperature inside cuts of meat as they leave the oven. Account for two practical advantages of the thermocouple in this application.
- (c) A copper-constantan thermocouple has its cold junction held in a bath of melting ice at 0°C . When the hot junction is held at the steam point of water (100°C), the EMF measured across the open ends of the thermocouple is 4.10mV. Treating the EMF as varying linearly with the temperature of the hot junction in the working range 0 to 200°C , calculate the temperature of an oil bath that produces an EMF of 7.40mV in the same thermocouple.

Question 15

- (a) Account for why the modern thermodynamic temperature scale is defined using the triple point of water as a single fixed reference point, rather than the ice point and the steam point used historically.
- (b) A standards laboratory in **Dar es Salaam** realises a triple-point cell of water as part of its temperature-reference chain. Identify what is special about a triple-point cell that makes its temperature more reproducible than a simple ice-water bath.
- (c) A constant-volume gas thermometer reads a pressure of 100kPa when its bulb is at the triple point of water (273.16K). The same thermometer, with its bulb placed in a vessel of liquid nitrogen at its boiling point, reads a pressure of 28.0kPa. Calculate the temperature of the boiling point of liquid nitrogen on the ideal-gas scale, in kelvin and in degrees Celsius.

Question 16

- (a) Account for why a thermistor (a semiconductor temperature sensor) has an electrical resistance that decreases with rising temperature, the opposite of a metal resistance thermometer, and account for why this property makes thermistors particularly suited to detecting small changes in temperature near room temperature.
- (b) A digital clinical thermometer used in a **Bagamoyo** health clinic has a thermistor at the tip of its probe rather than a platinum resistance element. Account for two practical reasons (beyond cost) why a thermistor is preferred in this application.
- (c) A thermistor has resistance 10.0k Ω at 25°C and 3.30k Ω at 50°C. The resistance varies with absolute temperature according to $R = R_0 e^{(B/T)}$, where R_0 and B are constants. Calculate the value of B and the resistance of the thermistor at body temperature 37°C.

Question 17

- (a) Account, by reference to the physical structure of a clinical mercury-in-glass thermometer, for why the reading remains at the maximum reached during measurement, even when the thermometer has been removed from the patient and cooled to room temperature.
- (b) A nurse at a **Chalinze** health clinic must shake the clinical thermometer vigorously after each patient's measurement before the next reading. Account for the physical reason this shaking is necessary.
- (c) The bulb of a clinical thermometer contains 0.20cm³ of mercury at 0°C. The cubic expansion coefficient of mercury is $\gamma_{\text{Hg}} = 1.82 \times 10^{-4} \text{K}^{-1}$. Take the cubic expansion of the glass as negligible. Calculate the change in the volume of mercury when the thermometer reading rises from room temperature 24°C to the patient's body temperature 39.5°C.

Question 18

- (a) Distinguish between the operating ranges and sensitivities of a mercury-in-glass thermometer and an alcohol-in-glass thermometer, and account for why each is preferred for its own range of applications.
- (b) A weather station outside **Mbeya** in the Southern Highlands, where night-time temperatures occasionally fall below -20°C in the cooler months, uses an alcohol-in-glass thermometer (rather than a mercury-in-glass thermometer) for its outdoor minimum-temperature record. Account for the choice.
- (c) A mercury-in-glass thermometer has a bulb containing 0.50cm³ of mercury at 0°C. It is graduated such that a 1.0mm rise of the mercury thread in its stem corresponds to a temperature rise of 1.0K. Take the cubic expansion coefficient of mercury as $\gamma_{\text{Hg}} = 1.82 \times 10^{-4} \text{K}^{-1}$ and assume the expansion of the glass is negligible. Calculate the cross-sectional area of the bore of the stem.

Question 19

- (a) A radiation thermometer (pyrometer) measures the temperature of a body without making physical contact with it. Account for the physical principle that allows this measurement, and identify the upper temperature limit above which contact thermometers (mercury-in-glass, platinum resistance) become impractical or impossible.
- (b) A steel-mill furnace operator at **Tanga** monitors the temperature of molten steel from outside the furnace using an optical pyrometer, while a poultry-incubator operator at a **Bagamoyo** farm uses a digital probe thermometer to monitor the temperature inside the incubator. Account for why each operator's choice of thermometer is appropriate to the application.

- (c) An optical pyrometer is calibrated by viewing a black body at a series of known temperatures, and is found to give an output reading proportional to the fourth power of the absolute temperature, in agreement with the Stefan-Boltzmann law. When pointed at a black body at 1000K, the pyrometer reads 1.00 (in arbitrary units). When pointed at a different glowing body, treated as a black body, the pyrometer reads 4.00. Calculate the temperature of this body.

Question 20

- (a) Distinguish between thermal conductivity and thermal conductance, and account for why the conductance of a wall containing a window of poorer thermal insulation than the wall itself depends not only on the conductivity of the two materials but also on the area of each.
- (b) An office building in **Lindi** has a south-facing wall containing both a brick portion and a single glass window. The architect adds a second window of the same area on the same wall. Account for whether the heat lost through the wall doubles, less than doubles, or more than doubles, identifying the dominant heat-loss path.
- (c) A wall of total area 10m^2 has a brick portion of area 8.5m^2 and thickness 0.10m , with thermal conductivity $1.0\text{Wm}^{-1}\text{K}^{-1}$. A glass window of area 1.5m^2 and thickness 5.0mm fills the rest, with thermal conductivity $0.80\text{Wm}^{-1}\text{K}^{-1}$. When the inside is at 25°C and the outside at 15°C , calculate the rate of heat conduction through (i) the brick portion, (ii) the glass window, and (iii) the wall as a whole.

Question 21

- (a) Account for why a cylindrical hot-water pipe loses heat at a rate per unit length that depends, for a given temperature difference between the pipe surface and the surroundings, on the natural logarithm of the ratio of the outer to the inner radius of any insulating layer wrapped around it, rather than on the layer thickness alone.
- (b) A solar water heater installed on a **Mwanza** rooftop carries hot water from the rooftop tank to a tap on the ground floor through a copper pipe wrapped in two distinct layers of lagging: an inner layer of mineral wool (low thermal conductivity) and an outer layer of foil-faced foam (slightly higher conductivity but also low). Account for why the engineer uses two layers in series rather than the same total thickness of a single material.
- (c) A copper hot-water pipe of inner radius 2.0cm carries water at 80°C through a building where the surrounding air is at 20°C . The pipe is lagged with two concentric layers: an inner layer of thickness 2.0cm and thermal conductivity $0.05\text{Wm}^{-1}\text{K}^{-1}$, surrounded by an outer layer of thickness 1.0cm and thermal conductivity $0.10\text{Wm}^{-1}\text{K}^{-1}$. Treating the pipe surface as held at 80°C and the outermost lagging surface as held at 20°C , calculate the rate of heat loss per unit length of the pipe.

Question 22

- (a) Newton's law of cooling, in the form rate of cooling proportional to temperature difference between the body and its surroundings, holds rigorously for forced convection but only as an approximation for natural convection at small temperature differences. Account for why this distinction arises.
- (b) A barbecue cook in **Bukoba** finds that a piece of grilled meat cools faster if she fans it with a banana leaf than if she leaves it on the table to cool in still air. Account for the difference, identifying which heat-transfer mechanism is enhanced by the fanning.
- (c) A small body cools from 80°C to 60°C in still air at 20°C in 5.0 minutes. When the same body is placed in a steady draught at the same air temperature, it cools from 80°C to 60°C in 1.5 minutes. Calculate the ratio of the cooling-law constant under forced convection to the constant under free convection.

Question 23

- (a) A blackbody is heated continuously from 1000K through 2000K to 5000K. Account for the systematic shift in the wavelength at which it radiates maximum energy, and account for the corresponding change in the colour at which the body appears to a human observer.
- (b) Astronomers in a Tanzanian observatory classify stars by colour: red giants appear cool red, our Sun appears yellow-white, and certain hot young stars appear bluish white. Account for this colour-temperature relationship by reference to Wien's displacement law.

- (c) Wien's displacement constant is $b = 2.898 \times 10^{-3} \text{ mK}$. A particular star emits its maximum radiation at a wavelength of 480nm. Calculate the surface temperature of the star, and compare it with the surface temperature of the Sun (which emits its maximum radiation at approximately 500nm).

Question 24

- (a) A grey body has the same emissivity at all wavelengths, but the emissivity is less than that of a perfect blackbody at the same temperature. Account, by reference to Kirchhoff's law of thermal radiation, for why a body that is a poor emitter at a particular wavelength must also be a poor absorber at the same wavelength.
- (b) The roof of a corrugated-iron house in a **Dar es Salaam** neighbourhood is painted matt black to reduce the temperature of the interior on a hot afternoon. Account, by reference to the difference in dominant radiation wavelengths during the day and at night, for whether the matt black paint actually achieves the desired cooling.
- (c) A grey-bodied object of emissivity 0.60 is held at a steady temperature of 400K in surroundings at 300K. Take the Stefan-Boltzmann constant as $\sigma = 5.67 \times 10^{-8} \text{ Wm}^{-2} \text{ K}^{-4}$. Calculate the net rate of radiation loss per unit area of the object's surface.

Question 25

- (a) A vacuum flask is designed to minimise heat transfer between its contents and the surroundings by all three mechanisms (conduction, convection, and radiation). Identify the design feature that minimises each mechanism, and account for why the flask cannot completely eliminate any of them.
- (b) A market trader in **Chakechake** fills a vacuum flask with hot tea at 90°C in the morning and finds that the tea is still pleasantly hot at 70°C when she drinks it eight hours later. Account for the tea's slow cooling, and identify the residual heat-loss mechanism that ultimately limits the flask's effectiveness.
- (c) A vacuum flask of internal surface area 0.010m² holds tea at 80°C in a room at 20°C. The internal walls of the flask are silvered, with effective emissivity 0.05. Take the Stefan-Boltzmann constant as $\sigma = 5.67 \times 10^{-8} \text{ Wm}^{-2} \text{ K}^{-4}$. Calculate the rate of heat loss by radiation through the silvered vacuum gap, and compare it with the approximate rate of heat loss from an unlagged open cup of the same area by free convection alone, taking the convection coefficient as $h = 10 \text{ Wm}^{-2} \text{ K}^{-1}$.

Question 26

- (a) A solar absorber panel is painted matt black on its sun-facing side to absorb the maximum fraction of incident solar radiation, but the same blackness causes it to radiate strongly in the infrared at its operating temperature. Account for why this radiative loss is unavoidable, and identify a coating strategy (a 'selective surface') that allows a high absorption of solar radiation without the corresponding high emission of thermal radiation.
- (b) A solar water heater installed on a rooftop has a flat absorber panel of area 2.0m². On a clear day at noon the panel receives a solar flux of 800Wm⁻² and reaches an operating temperature of 70°C while the surroundings are at 30°C. Account for why the panel does not continue to heat indefinitely, and identify the limiting factor on its steady-state temperature.
- (c) A solar absorber panel of area 2.0m² and emissivity 0.90 (assumed equal to its solar absorptivity) sits on a rooftop and receives solar radiation of flux 800Wm⁻². At its steady operating temperature of 70°C in surroundings at 30°C, calculate the rate at which the panel absorbs solar radiation, the rate at which it loses heat by infrared radiation to the surroundings, and the net useful heat-collection rate.

Question 27

- (a) Account, by reference to thermal radiation to the sky, for why frost can form on the surface of a tin roof or on the leaves of a plant on a clear night even when the air temperature one metre above the surface is several degrees above 0°C.
- (b) A farmer in **Mufindi** observes that on clear cold nights his crops in an open field develop frost on their leaves, while on cloudy nights with the same air temperature no frost appears. Account for the role of cloud cover in suppressing frost formation.
- (c) On a clear night, the effective sky temperature for thermal radiation is approximately 250K, while on an overcast night with low clouds it is approximately 280K. The ground surface is at 5°C (278K) on both nights and emits as a blackbody. Calculate the net rate of radiative heat loss per unit area of ground surface for each case, and account for why one case can produce frost while the other cannot.

Question 28

- (a) The Earth, treated as a perfect absorber of solar radiation and a perfect emitter of thermal radiation, would settle at a steady-state temperature determined by balancing the absorbed solar power with the thermal-infrared emission to space. Account for the geometric factor (involving the Earth's cross-section as seen from the Sun and its surface area for emission) that enters this balance.
- (b) Climate scientists at a Tanzanian research institute observe that the average measured temperature at the Earth's surface (around 288K) is substantially higher than the calculated equilibrium temperature for a bare planet of the same albedo. Account for this discrepancy by reference to the Earth's atmosphere.
- (c) Take the solar constant at the Earth's orbit as $S = 1370 \text{ W m}^{-2}$ and the Earth's average albedo (the fraction of solar radiation reflected back to space) as 0.30. Treating the Earth as a sphere that absorbs the un-reflected solar radiation and re-emits as a blackbody, calculate the equilibrium effective temperature of the Earth on the Stefan-Boltzmann scale.

Question 29

- (a) A typical horticultural greenhouse uses glass walls and roof to admit solar radiation while restricting the loss of heat at night. Account for the optical property of glass that allows it to act in this dual role, and identify the wavelength region in which glass is transparent and the region in which it is opaque.
- (b) A market gardener in **Iringa** builds a polythene-sheet greenhouse to extend the growing season for tomatoes. The polythene is transparent to visible light but, unlike glass, is also fairly transparent to thermal infrared. Account for why the polythene greenhouse is less effective at retaining heat than a glass one, and account for the limited improvement that polythene provides over no enclosure at all.
- (c) A greenhouse glass panel of area 4.0 m^2 is at a steady temperature of 20°C while the air outside is at 5°C . The glass behaves as a blackbody for thermal infrared radiation. Calculate the rate at which the panel loses heat to the cooler outdoor air by net infrared radiation, comparing this with the rate at which an equivalent perfectly transparent (zero-emissivity) film of the same area would lose. (Take the air's effective radiation temperature as 5°C .)

Question 30

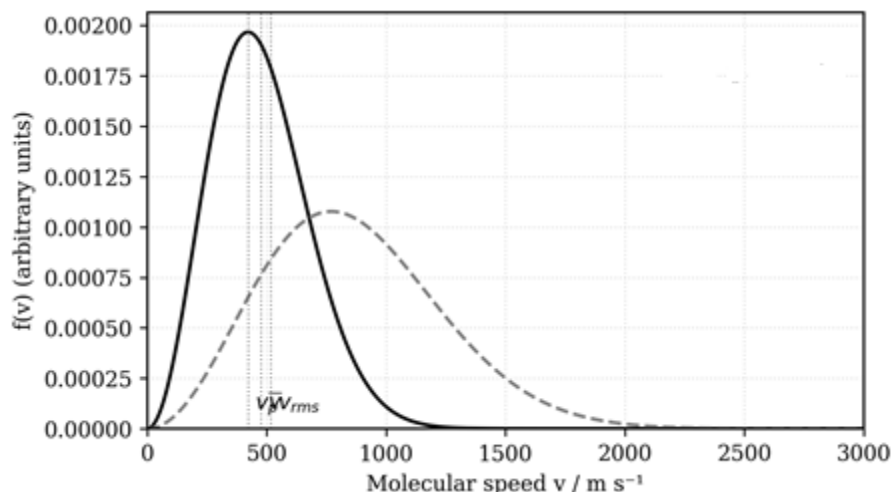
- (a) Account, by reference to the kinetic theory of an ideal gas, for why the root-mean-square speed of the molecules of a gas at a given temperature depends on the molar mass of the gas, but the average translational kinetic energy per molecule does not.
- (b) At a balloon-display event, two helium-filled balloons are released simultaneously with hydrogen-filled balloons. The hydrogen balloons are observed to deflate (lose gas through the rubber wall) noticeably faster than the helium balloons, even though both gases are at the same temperature and the rubber walls are identical. Account for this observation, identifying the molecular property responsible.
- (c) Calculate the root-mean-square speeds of hydrogen and nitrogen molecules at 300K, given molar masses $M_{\text{H}_2} = 2.0 \times 10^{-3} \text{ kg mol}^{-1}$ and $M_{\text{N}_2} = 28 \times 10^{-3} \text{ kg mol}^{-1}$. Find the ratio of the two speeds and comment briefly on its physical significance. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Question 31

- (a) Account for why the total translational kinetic energy of all the molecules in a sample of an ideal gas is far larger than the kinetic energy of a single sports ball, even though each individual molecule has only a tiny kinetic energy and even though the gas appears to be at rest.
- (b) A football placed on a windless day at the centre of a **Benjamin Mkapa** pitch appears stationary. A nearby kicked football flies through the air. State, with reason from the kinetic theory of gases, whether the stationary football's mass-and-air system contains more, less, or comparable translational kinetic energy than the kicked football itself.
- (c) Calculate the total translational kinetic energy of 2.0 mol of an ideal gas at 300K. Compare this with the kinetic energy of a 0.45kg football kicked at 25 ms^{-1} . Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Question 32

The figure below shows the Maxwell-Boltzmann speed distribution for molecules of a fixed-mass sample of nitrogen at two different temperatures.



- Identify the curve corresponding to the higher temperature, and account for the shift of the peak and the broadening of the distribution as the temperature rises.
- On a hot afternoon in Dodoma, a few molecules of perfume sprayed in one corner of a still room are smelt at the opposite corner only some seconds later, even though the mean speed of perfume molecules at 300K is comparable to the speed of a fast bicycle. Account, by reference to the Maxwell-Boltzmann distribution and to the molecules' frequent collisions, for the apparent slowness of perfume diffusion.
- From the figure, identify the most probable speed and estimate the rms speed of nitrogen molecules at 300K. Compare your estimate of the rms speed with the calculated value from $v_{\text{rms}} = \sqrt{3RT/M}$, taking $M_{\text{N}_2} = 28 \times 10^{-3} \text{kgmol}^{-1}$ and $R = 8.314 \text{Jmol}^{-1}\text{K}^{-1}$.

Question 33

- Distinguish between effusion and diffusion as processes by which gas molecules escape from one region to another, and account for why Graham's law (*rate of effusion inversely proportional to the square root of molar mass*) holds for effusion but is only approximately applicable to diffusion.
- A factory in Tanzania uses a long set of porous-membrane stages to enrich uranium for nuclear-power use, where the natural mixture of UF_6 molecules contains both $^{235}\text{UF}_6$ and $^{238}\text{UF}_6$ at very similar molar masses (349 and 352 g/mol respectively). Account for why a single membrane stage gives only a tiny enrichment of the lighter isotope, and why many stages in series are required.
- Two identical containers each have a small hole in the wall. One contains helium gas, the other oxygen gas, both at the same pressure and temperature. Calculate the ratio of the rate at which helium effuses through its hole to the rate at which oxygen effuses through the identical hole. Take $M_{\text{He}} = 4.0 \times 10^{-3} \text{kgmol}^{-1}$, $M_{\text{O}_2} = 32 \times 10^{-3} \text{kgmol}^{-1}$.

Question 34

- The speed of sound in an ideal gas is given by $v = \sqrt{\gamma RT/M}$. Account for why a person who inhales helium and then speaks has an audibly higher-pitched voice than usual, even though the vocal cords themselves vibrate at the same frequency as before.
- A safety officer at a Tanga industrial gas plant warns workers against inhaling helium directly from a cylinder for amusement, despite the high-pitched-voice effect being apparently harmless. Account for the physical danger of inhaling pure helium, identifying the property of helium that the warning addresses.
- Calculate the speed of sound in dry air and in helium, both at 27°C, taking $\gamma_{\text{air}} = 1.40$, $M_{\text{air}} = 29 \times 10^{-3} \text{kgmol}^{-1}$, $\gamma_{\text{He}} = 5/3$, $M_{\text{He}} = 4.0 \times 10^{-3} \text{kgmol}^{-1}$, and $R = 8.314 \text{Jmol}^{-1}\text{K}^{-1}$. Find the ratio of the two speeds.

Question 35

- A diatomic molecule at room temperature stores its thermal energy in five accessible degrees of freedom: three translational and two rotational. Account, by reference to the geometry of a dumbbell-shaped molecule, for why two and not three rotational modes are available.

- (b) A factory in Mwanza compares the cost of filling its argon balloons and its nitrogen balloons. The technician notes that for the same volume and temperature, the gases contain the same number of moles, but the nitrogen balloon has more total internal energy than the argon balloon. Account for this observation in terms of degrees of freedom.
- (c) Calculate the total internal energy of 1.0 mol of nitrogen (N_2) at 300K. Find the fraction of this internal energy that resides as translational kinetic energy and the fraction that resides as rotational kinetic energy. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Question 36

- (a) The molar heat capacity at constant volume of an ideal gas with f accessible degrees of freedom is $C_v = (f/2)R$, and at constant pressure is $C_p = ((f+2)/2)R$. Account, by reference to the first law of thermodynamics applied at constant pressure, for the additional R in C_p relative to C_v .
- (b) A welder in a workshop uses bottled argon gas as a shielding gas for arc welding because the arc stays cooler in argon than it would in nitrogen at the same heating power. Account for this observation in terms of the heat capacities of the two gases.
- (c) For a diatomic ideal gas with $f = 5$ accessible degrees of freedom, calculate C_v , C_p , and the ratio of principal heat capacities $\gamma = C_p/C_v$. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Question 37

- (a) The first law of thermodynamics states that the change in internal energy of a closed system equals the heat added to it minus the work done by it. Account for why, in any heating process at constant pressure, the heat supplied to an ideal gas always exceeds the change in its internal energy by exactly the work done by the gas during the expansion.
- (b) A welder in a workshop fills a flexible balloon with helium from a cylinder, then leaves it on a sunny bench. The sun warms the helium and the balloon expands. Account, by reference to the first law of thermodynamics, for why the heat absorbed by the helium from the Sun exceeds the rise in the gas's internal energy by an amount equal to the work the helium has done in pushing back the surrounding atmosphere.
- (c) 0.50 mol of an ideal monatomic gas at 300K is heated at the constant pressure $1.0 \times 10^5 \text{ Pa}$ until its temperature is 600K. Calculate the heat supplied, the work done by the gas, and the change in its internal energy. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Question 38

- (a) The change in internal energy of an ideal gas between two given states depends only on the two states, not on the path joining them. Account for why the work done by the gas in moving between the same two states is different along different paths.
- (b) A factory technician in Mwanza uses a piston compressor to raise a gas from atmospheric pressure to operating pressure. He notices that the compressor uses less work in cool weather (when the gas is allowed to remain near atmospheric temperature during compression) than in hot weather (when the gas heats during the compression and the pressure rises more rapidly). Account for this observation.
- (c) An ideal gas undergoes a transition from state A at (P_0, V_0) to state B at $(P_0/2, 2V_0)$ along two different paths. Path 1: isochoric drop from A to $(P_0/2, V_0)$, then isobaric expansion to B. Path 2: isobaric expansion from A to $(P_0, 2V_0)$, then isochoric drop to B. Calculate the work done by the gas along each path, expressed in terms of $P_0 V_0$, and identify which path delivers more work.

Question 39

- (a) The net work done by an ideal gas during one complete traversal of any closed cycle on a PV diagram equals the area enclosed by the cycle, with the sign of the area set by the direction of traversal. Account, by reference to the integral $W = \oint P dV$, for why this geometric area rule applies to cycles of every shape, not just rectangular ones.
- (b) A small two-stroke engine in a motorbike has its indicator diagram (a real-time PV record of the gas in the cylinder) recorded by a diagnostic technician. The diagram is roughly triangular rather than perfectly rectangular. Account for why the technician can still infer the engine's net work output per cycle from the shape of the diagram alone.

- (c) An ideal gas executes a triangular cycle on a PV diagram with corners at A($V = 1.0 \times 10^{-3} \text{ m}^3$, $P = 1.0 \times 10^5 \text{ Pa}$), B($V = 1.0 \times 10^{-3} \text{ m}^3$, $P = 2.0 \times 10^5 \text{ Pa}$), and C($V = 2.0 \times 10^{-3} \text{ m}^3$, $P = 1.0 \times 10^5 \text{ Pa}$), traversed in the order $A \rightarrow B \rightarrow C \rightarrow A$. The path $B \rightarrow C$ is a straight line on the PV diagram. Calculate the net work done by the gas per cycle.

Question 40

- (a) In an isothermal expansion of an ideal gas the temperature is held constant by allowing heat to flow in from the surroundings. Account, by reference to the first law and the property that internal energy of an ideal gas depends only on temperature, for why the heat absorbed during the expansion exactly equals the work done by the gas.
- (b) A natural-gas pipeline in a Tanzanian gas-distribution network carries compressed methane from a wellhead at 80 atmospheres into a distribution network at 5 atmospheres, with the gas held at approximately ground temperature by long contact with the soil. Account for why the heat absorbed by the gas from the soil equals the work the gas does in expanding, even though the gas does not visibly heat up during the expansion.
- (c) One mole of methane (treated as an ideal gas) at 290K is expanded isothermally from a pressure of 80 atmospheres to a pressure of 5 atmospheres. Calculate the work done by the gas during the expansion. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

Question 41

- (a) Carnot's theorem fixes the maximum efficiency of any heat engine operating between two reservoirs at temperatures T_{hot} and T_{cold} , in kelvin, as $\eta_{\text{Carnot}} = 1 - T_{\text{cold}}/T_{\text{hot}}$. Account for why this maximum cannot be approached except in the limit of a reversible (and therefore infinitely slow) cycle, and account for the consequences for real engines run at finite power.
- (b) A geothermal power station built near the volcanic Lake Natron in northern Tanzania extracts heat from underground steam at 150°C and rejects waste heat to the atmosphere at 27°C . Account, with reason from Carnot's theorem, for why the engineer designing the plant should expect a thermal efficiency of at most about 30%, and identify two factors that drive the realised efficiency below this Carnot ceiling.
- (c) A heat engine operates between an underground steam reservoir at 150°C and a cooling tower at 27°C . Calculate the Carnot efficiency for this engine, and the maximum work that could be extracted from each kilojoule of heat absorbed from the steam reservoir.

Question 42

- (a) In a calorimetry experiment, the measured final temperature is always slightly lower than the true value if the system is losing heat to the surroundings during the mixing process. Explain which mechanism is primarily responsible and why the error increases with longer experimental times.
- (b) A domestic refrigerator manufacturer in Mwanza advertises its product as having a coefficient of performance of 12 between an interior temperature of 4°C and a kitchen temperature of 28°C . State, with reason from Carnot's theorem applied in reverse, whether this claim is feasible.
- (c) A domestic refrigerator maintains its interior at 4°C in a kitchen at 28°C . Its compressor consumes 200W of electrical power. Calculate the maximum possible coefficient of performance for this refrigerator, and the maximum rate at which it could extract heat from its interior.

Question 43

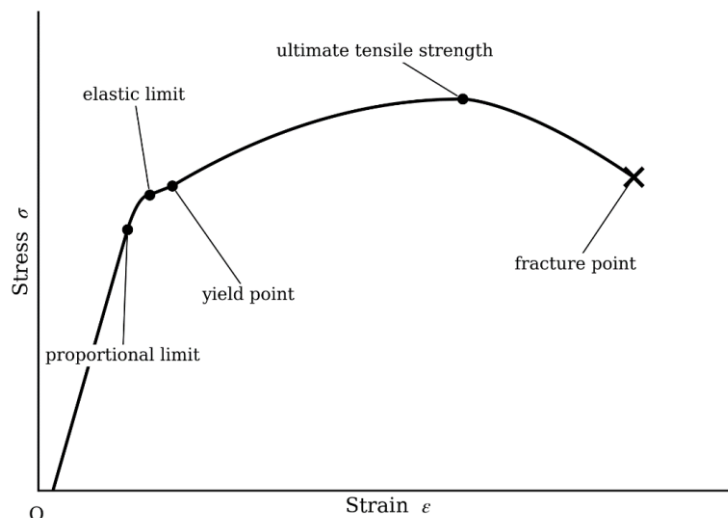
- (a) A corrugated iron roof in Dar es Salaam becomes extremely hot in midday sun but cools rapidly after sunset. Explain the rapid cooling, identifying which heat-transfer mechanisms dominate during the night-time cooling and why the roof cools faster than the concrete walls of the same building.
- (b) A gas-filled cylinder in an industrial laboratory is briefly connected through a valve to an adjacent evacuated cylinder. Both cylinders are well insulated. The gas rushes from the filled to the empty cylinder until pressure equalises. Account for why the final temperature of the gas (an ideal gas, well above its boiling point) is essentially the same as its initial temperature, even though the expansion is rapid and the system is thermally insulated.
- (c) 0.50 mol of an ideal gas at 27°C and $1.0 \times 10^5 \text{ Pa}$ fills one half of a thermally insulated $2.0 \times 10^{-3} \text{ m}^3$ container, with the other half evacuated. The partition is suddenly removed. Calculate the final temperature, the final pressure, and the change in internal energy of the gas. Take $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$.

ANSWERS TO DIGGING DEEPER EXERCISES

EXERCISE 1A

1. Hooke's law states that within the elastic regime, the extension of a wire is directly proportional to the applied force, so doubling the force doubles the extension. This proportionality holds because the interatomic restoring forces in the wire respond linearly to small displacements of the atoms from their equilibrium positions. Once the stress exceeds the **proportional limit**, the linearity breaks down because the interatomic forces are no longer well approximated as linear, and the extension begins to grow faster than the force.
2. The long thin wire extends more. The extension is given by $\Delta L = FL/(AE)$. For the same material (same E) and the same force (F), the extension is directly proportional to the original length L and inversely proportional to the cross-sectional area A . The long thin wire has a larger L and a smaller A , both of which act to increase the extension. The two geometric factors that determine the answer are therefore the original length and the cross-sectional area.
3. Young's modulus E describes resistance to longitudinal (tensile or compressive) deformation, and is the relevant property for a wire, rod, or column being stretched or compressed along its length. The shear modulus G describes resistance to angular (shape-changing) deformation at constant volume, and is the relevant property for a block being sheared sideways or a shaft being twisted. The bulk modulus B describes resistance to uniform volumetric (size-changing) deformation at constant shape, and is the relevant property for a fluid or solid subjected to a uniform hydrostatic pressure from all directions.
4. A fluid has no permanent intermolecular forces resisting tangential displacement of one layer relative to another. A static shear stress applied to a fluid causes its layers to flow continuously past each other, so the fluid never reaches a static equilibrium under shear and its shear modulus is zero. A fluid does, however, resist a uniform compression because squeezing the molecules closer together still meets a strong repulsive force, so its bulk modulus is finite. The essential physical difference between a solid and a fluid is therefore the presence in a solid of intermolecular forces that resist shape change, and their absence in a fluid.
5. The rubber band extends more, because rubber's Young's modulus is much smaller than steel's, and strain $\epsilon = \sigma/E$ is therefore larger in the rubber for the same applied stress. However, the steel wire stores more elastic potential energy. The energy stored is $U = \frac{1}{2} F \Delta L$, but the steel wire's much larger restoring force at any given small extension dominates over the rubber's larger total extension, and the steel ends up with more energy in the elastic regime that the question specifies. (At very large strains far beyond the steel's elastic limit, this comparison reverses, but within the small-force elastic regime steel wins.)
6. The total elastic potential energy stored in a stretched wire is the energy contained in the entire piece, measured in joules. The energy per unit volume is the energy stored in each cubic metre of the material, measured in joules per cubic metre. The total energy depends on the size of the wire, while the energy per unit volume depends only on the material and the stress applied. When comparing two different materials, the total energy is uninformative because it conflates the size of the sample with the property of the material. The energy per unit volume is the meaningful comparison because it isolates the material property from the geometry, and shows directly which material is intrinsically better at storing elastic energy.
7. Elastic fatigue is the gradual weakening of a material under many cycles of loading at stresses well below its breaking stress, caused by the slow growth of microscopic cracks at surface flaws with each loading cycle. It differs from ordinary overloading, which produces an immediate plastic deformation in a single event when the elastic limit is exceeded. Elastic fatigue accumulates damage cycle by cycle without any visible warning, while overloading damage is immediate and obvious. An everyday example is a spring balance in daily use at a market, whose zero reading drifts slowly over months and years because of elastic fatigue in its spring, even though no single load has ever come close to overloading it.

8.



- Between the origin and the proportional limit, the metal obeys Hooke's law and stress is directly proportional to strain.
- Between the proportional limit and the elastic limit, the metal still returns to its original length on unloading but the stress-strain relation is no longer linear.
- Between the elastic limit and the yield point, plastic flow begins and any unloading would leave a permanent set.
- Between the yield point and the ultimate tensile strength, the metal undergoes large plastic strain at gradually increasing stress as it work-hardens.
- Between the ultimate tensile strength and the fracture point, the metal forms a localised neck that thins rapidly until it breaks.

9. When a bimetallic strip is heated, both metals try to expand, but the metal with the larger coefficient of linear expansion tries to grow more than the metal with the smaller coefficient. The bond between them prevents either layer from sliding relative to the other, so the strip bends into an arc. The metal with the larger coefficient of expansion (brass, $\alpha \approx 19 \times 10^{-6} \text{K}^{-1}$) lies on the outer side of the curve, because the outside of a curve is longer than the inside, and the brass needs the extra length to accommodate its greater natural expansion. A practical device that depends on this bending is the mechanical thermostat, in which the bimetallic strip opens or closes an electrical contact as the temperature crosses a set point.

10. Searle's two-wire apparatus uses two identical wires hanging side by side from the same rigid support. One wire is the test wire, which carries the variable load whose effect on its extension is to be measured. The other is the reference wire, which is kept under a constant small load just heavy enough to keep it straight. A light frame attached to each wire connects them through a sensitive spirit level fitted with a micrometer screw gauge. As the test wire stretches under the applied load, the spirit level tips slightly and the micrometer is adjusted to recentre the bubble. The reading gives the extension of the test wire, and Young's modulus is then extracted from the slope of the load-extension graph.

The two advantages of the two-wire arrangement over a single-wire experiment are: **(1) thermal compensation**, because both wires hang in the same air at the same temperature, so any expansion or contraction caused by room temperature changes affects both equally and cancels out of the difference; **(2) compensation for any small movements of the support beam**, because both wires are suspended from the same beam and any settling, vibration, or unintended motion of the beam affects the tops of both wires equally and does not appear in the difference.

EXERCISE 1B

11. The spring is suffering from elastic fatigue. Each time Mama Amina hangs a load on the balance and then removes it, the spring is stretched and released within its elastic limit, but microscopic flaws on the surface of the wire grow a tiny amount with each cycle. After two years and many thousands of weighings, this damage has accumulated into a small permanent plastic deformation of the spring, which is why the unloaded length is now longer than when it was new and the pointer reads 50g instead of zero. Applying a correction of -50g to every reading is not a complete solution because the elastic fatigue continues to accumulate with every new use, so the zero error will keep growing. Worse, **fatigue can also reduce the spring constant** of the spring, so the same load no longer produces the same extension, meaning that the entire calibration has shifted and a single zero correction will not restore accuracy across the full range of the balance.

12. A long bridge in a coastal city experiences large temperature swings between cold nights and hot afternoons. If the steel cables and deck were rigidly fixed at both ends, the prevented thermal expansion would generate a thermal stress $\sigma = E\alpha\Delta T$ large enough

to crack the anchor blocks and fatigue the steel. Expansion joints allow the bridge to lengthen and shorten freely with temperature, so that the temperature change produces motion rather than stress.

13. The tyre acts as an elastic shock absorber between the road and the rigid metal frame of the car. Every small bump, pothole, or stone in the road would otherwise transmit a sharp impulsive force directly to the wheel rim and from there to the suspension and the passengers. The compliant rubber and air-filled structure of the tyre deforms elastically over a small distance when it meets a bump, spreading the impact force over a longer time and converting much of the impulsive shock into a gentle compression. A car with completely rigid wheels would transmit every roughness in the road surface, however small, as a sharp jolt to the chassis. The ride would be intolerably rough, the suspension would wear out quickly from the high-frequency loading, and the metal frame and bodywork would be subjected to repeated impulsive stresses that would fatigue them rapidly.

14. When the truck crosses the riveted bridge, the steel members bend slightly under the changing load, and at each riveted joint the plates slip a microscopic distance against one another before friction holds them again. Each microscopic slip releases a small burst of stored elastic energy as sound, and the accumulated bursts from many joints across the bridge produce the creaking. A welded bridge is silent because welded joints are continuous metal with no slipping surfaces, so there is no friction-driven release of elastic energy at the junctions.

15. When a horizontal beam bends under a downward load, the top of the beam is in compression and the bottom is in tension. Concrete is strong in compression but weak in tension, so a pure concrete beam would crack along the bottom. The steel bars are placed near the bottom because that is where the tensile stress is largest, and steel has a tensile strength which by far is higher than that of concrete; the steel therefore carries the tension while the concrete carries the compression.

16. The hard outer shell stops almost instantly on impact and would transmit the full deceleration directly to the rider's head if there were nothing soft inside it. The crushable foam liner deforms plastically over several centimetres during the impact, extending the stopping distance of the head and converting the kinetic energy of the head into the work of crushing the foam. Because the stopping distance is much longer, the peak deceleration on the brain is much smaller, and the impact is survivable.

17. The crack started at the corner because corners are points of high stress concentration: any uniform stress applied across a glass panel is amplified at sharp corners by a large factor (just as the corners of the de Havilland Comet windows became fatigue crack starters). The missing rubber spacers were responsible for producing the stress in the first place. Without rubber spacers to absorb the thermal expansion of the glass relative to the wooden frame, the glass developed a thermal stress $\sigma = E\alpha\Delta T$ each time the room warmed in the afternoon, and this stress was amplified at the corners until a crack opened.

18. A sudden jerk extends the stiff nylon line through a very small distance before it reaches its breaking point, which means the line absorbs the kinetic energy of the rod and the fish over a tiny stopping distance. From $U = \frac{1}{2}F\Delta L$, a small ΔL requires a very large peak force F to absorb the same energy, and that peak force exceeds the breaking strength even though the average force during a steady pull would not. A stretchy line would have spread the same energy over a longer extension and reduced the peak force enough to keep the line intact.

19. Strength and stiffness are independent properties of a material. Old bone has roughly the same Young's modulus as young bone (so it feels just as stiff under small loads) but a much smaller breaking stress, because the porous microstructure that develops with age provides easier paths for cracks to propagate. The grandmother's bone was therefore stiff enough to feel healthy right up to the moment of the fall, then failed at a stress that a young bone would have absorbed without difficulty.

EXERCISE 1C

20. Deceleration: $a = \frac{\Delta v}{\Delta t} = \frac{3.0}{0.40} = 7.5 \text{ m/s}^2$ (upward).

Net upward force on cage: $F_{\text{net}} = ma = (1500)(7.5) = 11250 \text{ N}$.

Tension in cable: $T = mg + ma = 1500(9.8 + 7.5) = 25950 \text{ N}$.

$\sigma = \frac{T}{A} = 25950 / (4.0 \times 10^{-4}) = 6.49 \times 10^7 \text{ Pa}$

21. 125°C

22. $8.03 \times 10^8 \text{ Pa}$

23. 17.7 m/s

24. $U_A : U_B = 3 : 5$

25. $2.58 \times 10^6 \text{ N}$

26. 25.09 mm .

27. 228°C

28. Maximum working stress: $\sigma_{\text{working}} = \frac{\sigma_{\text{yield}}}{n} = \frac{2.5 \times 10^8}{5} = 5.0 \times 10^7 \text{ Pa}$.

Area of one cable: $A = \frac{\pi(5.0 \times 10^{-3})^2}{4} = 1.963 \times 10^{-5} \text{ m}^2$.

Maximum force per cable: $F_{\text{max}} = \sigma_{\text{working}} A = (5.0 \times 10^7)(1.963 \times 10^{-5}) = 981.7 \text{ N}$.

Total weight allowed: $W_{\text{max}} = 6F_{\text{max}} = 5890 \text{ N}$.

$$m_{\max} = \frac{W_{\max}}{g} = \frac{5890}{9.8} = 601\text{kg}$$

29. 3.83kg

30. For a conical pendulum, the vertical component of the wire tension supports the weight: $T\cos 30^\circ = mg$

$$T = \frac{mg}{\cos 30^\circ} = \frac{(0.20)(9.8)}{0.866} = 2.263\text{N}$$

$$\text{Cross-sectional area: } A = \frac{\pi(0.50 \times 10^{-3})^2}{4} = 1.963 \times 10^{-7}\text{m}^2$$

$$\Delta L = \frac{TL}{AE_c} = \frac{(2.263)(0.80)}{(1.963 \times 10^{-7})(1.2 \times 10^{11})} = 7.69 \times 10^{-5}\text{m} \approx 0.077\text{mm}$$

31. Fractional density change is approximately equal to the negative of the fractional volume change:

$$\frac{\Delta \rho}{\rho_0} = \frac{(1003.0 - 1000.0)}{1000} = 3.0 \times 10^{-3}$$

From the bulk modulus, $\Delta P = -B\left(\frac{\Delta V}{V_0}\right) = B\left(\frac{\Delta \rho}{\rho_0}\right)$:

$$\Delta P = (2.2 \times 10^9)(3.0 \times 10^{-3}) = 6.6 \times 10^6\text{Pa}$$

The pressure at depth h is $\Delta P = \rho_0 gh$:

$$h = \frac{\Delta P}{\rho_0 g} = \frac{(6.6 \times 10^6)}{(1000)(9.8)} = 673\text{m}$$

32. Conservation of energy. As the ball falls a distance ΔL (the maximum extension of the wire), it loses gravitational PE equal to $mg\Delta L$ and gains elastic PE equal to $\frac{1}{2}k(\Delta L)^2$, where $k = AE/L$ is the spring constant of the wire:

$$mg\Delta L = \frac{1}{2} \frac{AE}{L} (\Delta L)^2$$

Solving for ΔL :

$$\Delta L = \frac{2mgL}{AE} = \frac{2(2.0)(9.8)(3.0)}{(1.0 \times 10^{-6})(2.0 \times 10^{11})} = 5.88 \times 10^{-4}\text{m} \approx 0.59\text{mm}$$

This is twice the static extension of the wire under the same weight, as expected for a sudden load applied to a Hookean wire.

33. $7.5 \times 10^{-5}\text{m}$ (or 0.075mm)

EXERCISE 2A

1. Kipanga's common sense is not common physics. The Zeroth Law guarantees that "being in thermal equilibrium" is a **transitive** relation: if A and B are in thermal equilibrium, and B and C are in thermal equilibrium, then A and C must also be. This is not a mathematical identity; it is an empirical law that has to be tested experimentally, because there is no a priori reason why thermal equilibrium should behave transitively. Without it, a thermometer placed in body B could not be used to compare body A with body C. The Zeroth Law is therefore the foundation on which every thermometer's logic rests.

2. The mistake is the claim that a body "contains" heat. Heat is not a quantity stored inside a body; it is energy transferred between two bodies because of a temperature difference between them. A body can have a temperature, and it can exchange heat with another body.

3. (a) A thermometric property is a physical property that changes **reproducibly and monotonically** with temperature and can therefore be used to measure temperature. On another hand, an ordinary physical property is any measurable feature of a substance, not all of which satisfy this criterion.

(b) Two examples of non-thermometric properties:

- The colour of a metal at room temperature, because it does not change with small temperature changes.
- The density of water between 0°C and 4°C , because it does not change monotonically with temperature (it increases then decreases), so a single density value corresponds to two temperatures.

4. Two disadvantages of the ice point:

- The ice point depends on the atmospheric pressure above it. A laboratory at high altitude with lower atmospheric pressure measures an ice point at a slightly different temperature from a laboratory at sea level.
- The ice point is affected by dissolved air in the water. Pure water with no dissolved air freezes at a slightly higher temperature than water saturated with dissolved air, and the difference is large enough to be significant at the precision required for a primary reference.

5. (a) Thermal equilibrium between the thermometer and the bath.

(b) Before thermal equilibrium is reached, the thermometer is at a different temperature from the bath, and its reading reflects only its own temperature, not the bath's. Only at thermal equilibrium does the thermometer's temperature equal the bath's, and only then does the reading correspond to the true temperature of the bath.

6. (a) Water has the largest value of specific heat capacity among common liquids. The numerical value is $4200 \text{ J kg}^{-1} \text{ K}^{-1}$.

(b) A given mass of oil absorbs less than half the heat per degree of temperature rise compared with the same mass of water. For the same cooling duty, a system using oil would need either more than twice the mass flow rate of oil, or allow the oil to heat up by more than twice as much, either of which would require larger pumps, larger pipes, and larger radiators. Water's larger specific heat capacity makes it the more efficient coolant, whenever it is available.

7. During melting, intermolecular forces holding the solid rigid structure must be partially loosened to allow the molecules to slide past one another as a liquid. Loosening the intermolecular forces requires energy, which is supplied as heat absorbed from the surroundings. During freezing, the reverse process occurs: the molecules lock back into rigid positions, and the energy previously used to loosen the intermolecular forces is released as heat to the surroundings. The temperature stays at 0°C throughout because the energy change concerns only the molecular potential energy (bond configuration), not the molecular kinetic energy (which sets the temperature).

8. (a) Three differences:

- Evaporation occurs only at the surface of a liquid; boiling occurs throughout the volume.
- Evaporation occurs at any temperature above the freezing point; boiling occurs only at a specific temperature (the boiling point) that depends on the applied pressure.
- Evaporation is a slow process; boiling is rapid and produces visible bubbles in the liquid.

(b) Both processes absorb the same latent heat of vapourisation per kilogram of liquid converted to vapour, because the molecular cost of separating a liquid molecule into the vapour phase is the same regardless of where the separation happens.

9. A thermometer placed in contact with a body extracts heat from the body to warm itself, and the amount extracted is determined by the thermometer's thermal capacity which is small for small bulb. If the thermometer's thermal capacity is very small compared with the body's, the heat extracted disturbs the body's temperature only **negligibly**, and the reading is close to the body's original temperature. If the thermometer's thermal capacity were comparable to the body's, the act of measurement would significantly cool the body, and the reading would correspond to the disturbed temperature rather than the original one. The measurement would then be meaningless. A small bulb minimises the mass of the thermometric substance and therefore the disturbance.

EXERCISE 2B

10. (a) Kipute is using temperature correctly. Temperature is a property of the air (or of a body), not a property of a person's response to it. Kipanga is inferring the air's temperature from his own sweating, but sweating also depends on what he ate, how much he is exerting himself, his hydration level, and his clothing. These personal factors vary from day to day, so sweating is not a reliable indicator of the air's temperature.

(b) A thermometer measures a physical property of a substance inside the thermometer itself such as the length of a mercury column, the resistance of a platinum coil, the EMF of a thermocouple junction, that changes reproducibly and only with temperature. The reading depends on the thermometer's physics and not on the person using it.

11. (a) Two reasons:

- The thermometer begins each measurement well below body temperature (the antiseptic bath is cool), so during the brief time in the mouth it absorbs heat rapidly but has not yet reached thermal equilibrium when she reads it.
- Shaking the thermometer immediately before insertion leaves no time for the mercury column to settle from its lowest position, and the reading reflects a transient state rather than the true equilibrium temperature.

(b) Two changes:

- Allow each thermometer to remain in the child's mouth long enough (typically two to three minutes) to reach thermal equilibrium before reading.
- Rinse and dry the thermometer between patients instead of cooling it in antiseptic, so it starts each measurement near room temperature rather than well below it.

12. The father's reasoning is wrong because the gas flame is not twice as hot as the charcoal flame; the two flame temperatures are broadly similar. What differs is the **rate at which each cooker delivers heat** to the pot. The gas flame is more concentrated and has a much higher heat transfer rate to the base of the pot, so more joules of heat per second reach the beans. The charcoal flame, being diffuse and partially shielded by ash, delivers heat more slowly. Since the heat required to cook the beans is the same in both cases, the faster delivery rate of the gas cooker means the same heat is delivered in a shorter time.

13. The water-filled bottle will keep Mama Juma warmer for longer. If both bottles lose heat to the surroundings at the same rate, the bottle that can release more heat per degree of its own temperature drop is the one that stays warm longer. The water's specific

heat capacity is $4200\text{J kg}^{-1}\text{K}^{-1}$, more than double the oil's $2000\text{J kg}^{-1}\text{K}^{-1}$, so a kilogram of water at 60°C stores more than twice the heat (above the room temperature) that a kilogram of oil at 60°C stores. With the same rate of heat loss, the water bottle's temperature falls more slowly than the oil bottle's, and Mama Juma stays warm for roughly twice as long.

14. (a) Water has a specific heat capacity of $4200\text{J kg}^{-1}\text{K}^{-1}$, higher than any other common liquid. A small mass of water flowing through the engine can absorb large quantities of heat for only a modest temperature rise. Used engine oil has a much smaller specific heat capacity, roughly $2000\text{J kg}^{-1}\text{K}^{-1}$, so for the same flow rate the oil would absorb less than half the heat per degree of temperature rise. The engine would heat up further before the oil could remove the heat, and overheating would continue.

(b) One other property: the boiling point (water boils at 100°C , which sets the upper operating temperature; a coolant that boils too low would vapourise in the engine).

Other acceptable properties: freezing point, corrosion resistance to the engine metals, viscosity, non-flammability.

15. (a) The technician's advice is not physically sensible. The three thermometers disagree because they use three different thermometric substances (mercury expansion, platinum resistance, thermocouple EMF), and each has its own slight non-linearity with temperature. Recent calibration does not remove this inherent disagreement; it only fixes the readings at the two calibration points.

(b) The student should record the platinum resistance thermometer's reading of 48.5°C . Among the three instruments, the platinum thermometer has the smallest departure from true thermodynamic temperature between the fixed points, and its reading can be taken as the most accurate of the three.

16. When the sambusas are dropped into the oil, heat flows from the hot oil to the cold sambusas, cooling the oil. The oil's thermal capacity is $C = mc$, fixed by the mass of oil in the pot. The cooling caused by the sambusas is $\Delta T = Q/C$, where Q is the heat the sambusas extract. When twenty sambusas are added at once, Q is four times larger than when five are added, so the oil's temperature drops four times further. A larger drop means the sambusas cook in cooler oil and therefore take longer. The vendor can avoid the problem by using a larger pot of oil (higher thermal capacity) so that the same number of sambusas produces a smaller temperature drop, or by frying the sambusas in smaller batches so that each batch encounters oil close to the intended cooking temperature.

17. When the soda bottle is placed in the iced water basin, heat flows from the soda (and from the water) to the ice. As it receives heat, the ice begins to melt, absorbing the latent heat of fusion $L_f = 3.34 \times 10^5\text{J kg}^{-1}$ without changing its own temperature. The ice therefore acts as an energy sink that keeps the water in the basin at 0°C until all of it has melted. In the 4°C basin, there is no latent-heat sink; every joule extracted from the soda raises the temperature of the surrounding water, reducing the temperature difference between the soda and the water and slowing the cooling rate. So the iced-water basin stays colder longer and cools the soda faster, just as Kipute claims.

Question 18

When the lid is off, water molecules at the surface of the stew can escape as vapour by evaporation, and each escaping molecule carries away the latent heat of vapourisation. This heat is lost to the air and does not go into warming the stew. A significant fraction of the flame's heat is consumed replacing the lost molecules' energy rather than raising the stew's temperature. With the lid on, the evaporated molecules accumulate in the air trapped above the stew and many of them condense back onto the surface, returning the heat. The lidded pot therefore reaches boiling sooner.

19. (a) The valve stays silent for the first ten minutes because the water inside the cooker is being warmed from room temperature toward its (elevated) boiling point, and no net steam is being produced yet. Once the water reaches its boiling point at the cooker's internal pressure, steam begins to accumulate, the internal pressure rises to the valve's set point, and the valve begins to hiss as it vents steam.

(b) The continued hissing means that the rate at which steam is being produced (by the flame's boiling of the water) exceeds the rate at which the valve can release it without the pressure rising further. The valve is therefore maintaining the pressure at its set value by matching its steam-release rate to the steam-production rate.

(c) If the flame is reduced but the valve continues to hiss, the cooker's internal temperature does **not** drop. The temperature is locked at the boiling point corresponding to the valve's set pressure, and so long as any steam is being produced and the valve is hissing, the pressure (and hence the temperature) stays at that set value. The reduction in flame strength reduces only the rate of steam production, and so the rate at which the valve hisses.

EXERCISE 2C

20. 0.47°C .

In a real waterfall, the actual temperature rise is smaller because: (1) some of the gravitational potential energy goes into kinetic energy of spray and air motion, which is dissipated to the surrounding atmosphere rather than into the falling water itself, and (2) some of the falling water evaporates on the way down, absorbing latent heat that would otherwise have raised the temperature.

21. 16.3°C

22. 21.2g , 100°C .

23. $a = 40\mu\text{V}/^\circ\text{C}$, $b = -0.040\mu\text{V}/^\circ\text{C}^2$, $1950\mu\text{V}$ or 1.95mV

24. The rate of growth of an ice layer of thickness x is $dx/dt = k\Delta T/(\rho L_f x)$, derived in the chapter from balancing the rate of heat conducted through the ice with the rate at which the released heat of fusion is produced. Rearranging and integrating:

$$x dx = \frac{k\Delta T}{\rho L_f} dt$$

Integrating from $x_1 = 0.060\text{m}$ at time 0 to $x_2 = 0.070\text{m}$ at time t :

$$\begin{aligned}\int_{x_1}^{x_2} x dx &= \frac{k\Delta T}{\rho L_f} \int_0^t dt \\ \frac{x_2^2 - x_1^2}{2} &= \frac{k\Delta T}{\rho L_f} t \\ t &= \frac{\rho L_f (x_2^2 - x_1^2)}{2k\Delta T}\end{aligned}$$

Substituting the numerical values, with $\Delta T = 8\text{K}$:

$$t = \frac{920\text{kg m}^{-3} \times 3.34 \times 10^5\text{J kg}^{-1} \times ((0.070\text{m})^2 - (0.060\text{m})^2)}{2 \times 2.2\text{W m}^{-1}\text{K}^{-1} \times 8\text{K}} = 1.135 \times 10^4\text{s} \approx 3.15\text{h}$$

The ice layer grows from 6.0cm to 7.0cm in about 3.15h.

25. 46.4°C .

26. (a) 53.33°C (b) 52.45°C (c) 1.7%

27. 692g

28. (a) $1.68 \times 10^5\text{J}$ (b) 1.4min or 84s

29. 0.80gh^{-1}

30. Specific heat capacity of the solid phase (segment 1, $t = 0$ to 60s, $\Delta T = 40\text{K}$):

$$\begin{aligned}Q_1 &= P \times t_1 = 50\text{W} \times 60\text{s} = 3000\text{J} \\ c_{\text{solid}} &= \frac{Q_1}{m \Delta T_1} = \frac{3000\text{J}}{0.100\text{kg} \times 40\text{K}} = 750\text{J kg}^{-1}\text{K}^{-1}\end{aligned}$$

Latent heat of fusion (segment 2, $t = 60\text{s}$ to 140s , $\Delta t_2 = 80\text{s}$):

$$\begin{aligned}Q_2 &= P \times t_2 = 50\text{W} \times 80\text{s} = 4000\text{J} \\ L_f &= \frac{Q_2}{m} = \frac{4000\text{J}}{0.100\text{kg}} = 4.0 \times 10^4\text{J kg}^{-1}\end{aligned}$$

Specific heat capacity of the liquid phase (segment 3, $t = 140\text{s}$ to 200s , $\Delta T = 70\text{K}$):

$$\begin{aligned}Q_3 &= P \times t_3 = 50\text{W} \times 60\text{s} = 3000\text{J} \\ c_{\text{liquid}} &= \frac{Q_3}{m \Delta T_3} = \frac{3000\text{J}}{0.100\text{kg} \times 70\text{K}} = 429\text{J kg}^{-1}\text{K}^{-1}\end{aligned}$$

31. Kinetic energy of the bullet just before impact:

$$\text{KE} = \frac{1}{2}mv^2 = \frac{1}{2} \times 0.005\text{kg} \times (350\text{m s}^{-1})^2 = 306.25\text{J}$$

Heat required to raise the bullet's temperature from 25°C to its melting point 327°C :

$$Q_{\text{warm}} = m c_{\text{Pb}}(327 - 25)\text{K} = 0.005\text{kg} \times 128\text{J kg}^{-1}\text{K}^{-1} \times 302\text{K} = 193.28\text{J}$$

Since $\text{KE} = 306.25\text{J} > Q_{\text{warm}} = 193.28\text{J}$, the bullet reaches its melting point and continues to absorb heat. The remaining energy goes into melting some of the lead at 327°C :

$$Q_{\text{remaining}} = \text{KE} - Q_{\text{warm}} = 306.25\text{J} - 193.28\text{J} = 112.97\text{J}$$

Mass of lead that melts:

$$m_{\text{melt}} = \frac{Q_{\text{remaining}}}{L_{f,\text{Pb}}} = \frac{112.97\text{J}}{2.5 \times 10^4\text{J kg}^{-1}} = 4.52 \times 10^{-3}\text{kg} = 4.52\text{g}$$

About 4.5g of the bullet's mass melts on impact.

32. 0.147g, 0.3%.

EXERCISE 3A

1. (a) Metal has much higher thermal conductivity than wood. In the brief contact, the metal conducts heat away from the skin rapidly, drawing enough thermal energy to cause a burn. Wood conducts heat so slowly that very little energy reaches the skin in half a second, so the temperature of the skin barely changes.

(b) Not correct. Both handles are at the same temperature (100°C , the temperature of boiling water). The student is confusing temperature (which is the same) with the rate of heat transfer (which depends on thermal conductivity). The metal feels hotter because it delivers heat to the hand faster, not because it is at a higher temperature.

2. (a) In an open air gap, although air itself is a poor conductor, the air is free to move. Warm air near the hot surface rises and is replaced by cooler air, setting up convection currents that carry heat across the gap far more effectively than conduction alone. Felt prevents this air movement.

(b) Convection. The felt traps air in small pockets between its fibres, preventing bulk air movement. The open gap allows free convection.

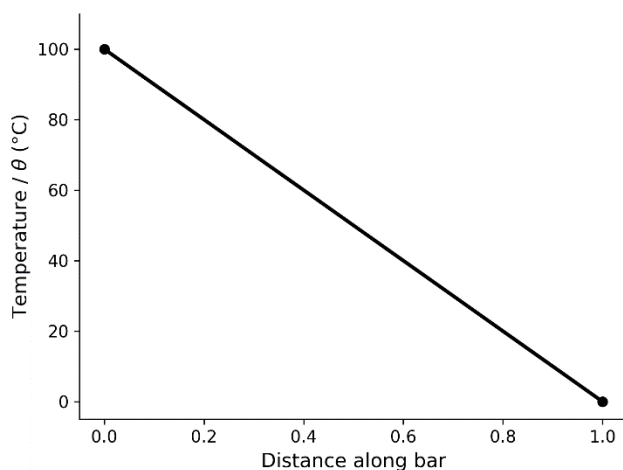
(c) All fibrous materials work by trapping air in small, isolated pockets. The fibres themselves conduct a small amount of heat, but their main function is to break the air into cells too small for convection currents to develop. The insulation comes from the still air, not from the fibre material itself.

3. (a) In the electrical analogy, temperature difference is analogous to voltage drop and heat flow rate is analogous to current. In a series circuit, the same current flows through every resistor, and the voltage drop across each is proportional to its resistance ($V = IR$). Similarly, the same heat flow rate passes through every layer, and the temperature drop across each layer is proportional to its thermal resistance ($\Delta\theta = \dot{Q} \times R$).

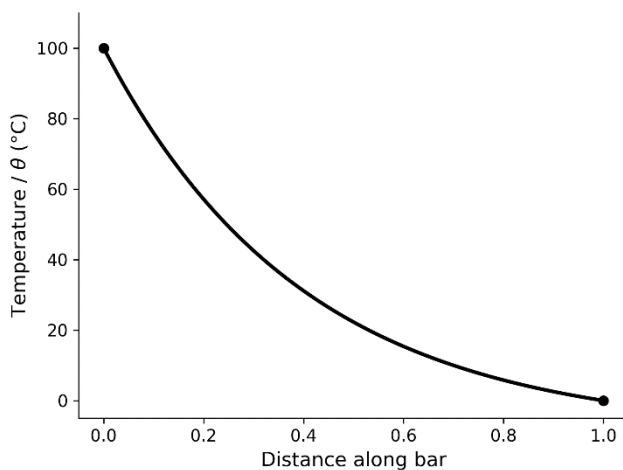
(b) The layer with the largest thermal resistance has the largest temperature drop. This is the layer with the combination of lowest thermal conductivity and greatest thickness, since $R = L/(kA)$ and A is constant for series arrangement. So it is not necessarily the thickest layer or the one with the lowest conductivity alone; it is the one with the highest L/k ratio.

(c) A thick layer of a good conductor (high k) can have a lower thermal resistance than a thin layer of a poor conductor (low k). Thickness matters, but only in combination with conductivity. The builder's claim ignores the role of k .

4. (a)



(b)



Explanation: When unlagged, heat leaks from the sides along the entire length. The heat flow rate decreases as you move from the hot end toward the cold end (because some heat has already escaped). A decreasing heat flow rate means a decreasing temperature gradient, so the profile curves.

Direction: It is concave upward (curves toward the hot end). This because the temperature drops steeply near the hot end (where the heat flow is large) and flattens near the cold end (where less heat remains).

(c) Below 50°C. The heat leaking from the sides throughout the bar means less heat reaches the midpoint than in the lagged case. With less heat flowing through the second half of the bar, the temperature at the midpoint is pulled downward.

5. (a) Newton's regime begins when the excess drops below about 30°C . Since the room is at 28°C , this occurs at about 58°C (excess = $58 - 28 = 30^\circ\text{C}$). Below 58°C , Newton's law applies; above 58°C , Dulong-Petit governs.

(b) Faster. The Dulong-Petit law gives a cooling rate that increases more rapidly than linearly with excess temperature (it follows a power law with index greater than 1). Newton's law, being linear in the excess, underestimates the cooling rate at large excess temperatures.

6. The student applied T^4 to the body's emission alone, ignoring the surroundings' contribution. The net heat loss is $\epsilon\sigma A(T^4 - T_s^4)$, not $\epsilon\sigma AT^4$. The ratio of net losses is not the same as the ratio of emitted powers.

7. (a) Convection. The hole allows warm air to rise out of the flask and be replaced by cooler air from outside, creating a convection current through the opening. This is the most immediate effect because convection can transport heat much faster than radiation or conduction through the stopper material.

(b) Radiation increases. The silvered surfaces had very low emissivity ($\epsilon \approx 0.02$), suppressing radiative exchange across the vacuum gap. As the coating deteriorates, the emissivity rises, and the surfaces emit and absorb more radiation. Since the vacuum cannot carry heat by conduction or convection, radiation is the only mechanism operating across the gap, and its increase directly raises the overall heat-transfer rate.

(c) The student's error is assuming the flask has a preferred direction. The flask suppresses heat transfer in both directions by the same set of mechanisms. When hot coffee is inside, it slows heat flowing outward. When cold juice is inside, the same vacuum, silvering, and stopper slow heat flowing inward from the warm room. The flask reduces the rate of heat transfer regardless of direction.

8. (a) Block A cools faster initially. It has a larger excess temperature ($80 - 20 = 60^\circ\text{C}$ versus $60 - 20 = 40^\circ\text{C}$), and Newton's law says the rate of temperature fall is proportional to the excess.

(b) Block A takes longer. It starts further from room temperature, so it has more excess to shed. Even though it cools faster at every moment (because it is always hotter than B at the corresponding elapsed time), the extra distance it must cover outweighs the faster speed.

(c) Yes, they have the same half-cooling time. The half-cooling time $t_{1/2} = \ln 2/\lambda$ depends only on λ , which is the same for both blocks (same mass, surface, material, and surroundings).

The answers to (a) and (b) are not contradictory: A cools faster in absolute terms (more degrees per minute) but not in relative terms (the same fraction of its excess decays in each half-cooling time). The exponential decay is the same; only the starting point differs.

9. (a) Prevost's theory: every body emits and absorbs radiation simultaneously. The baby's skin (at about 36°C) emits infrared radiation toward all surrounding surfaces. If the incubator walls are cooler than the baby's skin (for example, at 25°C), the walls emit less radiation back toward the baby than the baby emits toward them. The net radiative exchange is from the baby to the walls. The air temperature is irrelevant to radiation; radiation does not pass through air by convection but travels directly between surfaces.

(b) The problem is that the incubator walls are at a lower temperature than the baby. To eliminate the net radiative loss, the walls would need to be at the same temperature as the baby's skin (about 36°C). Then, by Prevost's theory, the baby would absorb exactly as much radiation from the walls as it emits, and the net radiative exchange would be zero.

(c) The double-walled design circulates warm water between the inner and outer walls, raising the inner wall temperature close to the baby's skin temperature. The baby now sees warm walls radiating strongly back toward it, reducing the net radiative loss to near zero. The air inside is also warm, so convective losses are minimal. Both mechanisms are addressed simultaneously.

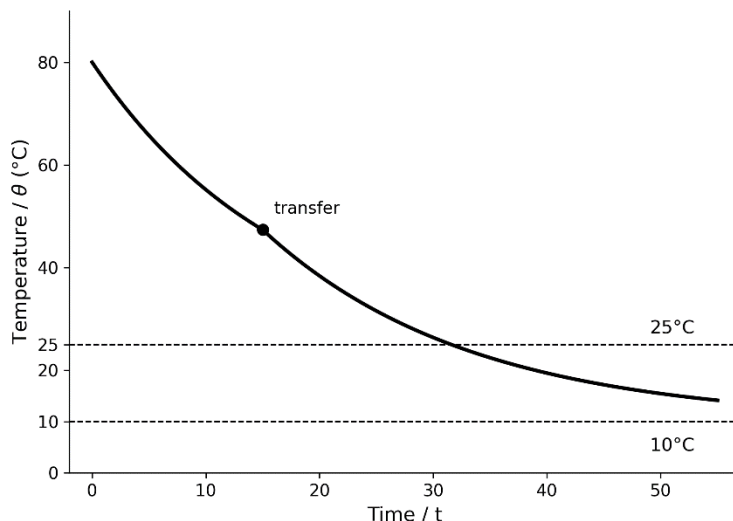
10. (a) Conduction. Heat flows from the liquid water at 0°C (just below the ice) upward through the solid ice to the cold air above. The water is the heat source; the air is the heat sink; the ice is the conducting path.

(b) Fourier's law: $\frac{dQ}{dt} = \frac{kA\Delta\theta}{x}$. The temperature difference $\Delta\theta$ and the area A stay the same, but the thickness x of the ice increases as more ice forms. A thicker ice layer means a longer conduction path, which increases the thermal resistance and decreases the heat flow rate.

(c) The heat that flows through the ice is the heat that freezes more water at the bottom surface. As the ice thickens, the heat flow rate drops (from (b)), so less water freezes per second, and the ice grows more slowly. The ice is insulating the water below from the cold air above. Each new layer of ice makes the insulation thicker, which slows the formation of the next layer. The process is self-limiting: the ice can never grow fast enough to overcome its own insulating effect.

11. (a) The rate of temperature fall increases instantly. In the 25°C room, the driving force was the excess over 25°C . In the 10°C room, the excess is suddenly larger (body temperature minus 10°C instead of minus 25°C), so Newton's law gives a larger rate of temperature fall at that instant.

(b) The cooling curve starts as a standard exponential decay toward 25°C. At the transfer point, the curve kinks downward: the slope becomes steeper and the curve now decays toward 10°C instead of 25°C.



(c) No. The cooling constant $\lambda = k/(mc)$ depends on k , which for radiative cooling is $k = 4\epsilon\sigma AT_s^3$. Since T_s is different in the two rooms (298K versus 283K), the value of k and therefore λ is different. The cooling constant in the 10°C room is slightly smaller than in the 25°C room.

12. (a) Natural convection requires buoyancy: warm air must be less dense than cooler air and must rise under the influence of gravity. In microgravity (free fall around Earth), there is no effective gravitational force to drive density-driven flow. Warm air stays where it is and does not rise, so natural convection currents cannot form. Hot spots develop around equipment with no natural mechanism to carry the heat away.

(b) Forced convection. The fans mechanically push air over the hot surfaces, replacing warm air with cooler air regardless of gravity. Forced convection does not rely on buoyancy; it relies on an external force (the fan blades) to move the fluid. This works in any gravitational environment, including microgravity.

(c) Outside the station, there is a near-vacuum: no material medium exists for conduction (which requires molecular contact) or convection (which requires a fluid). Radiation is the only mechanism that can transport energy through a vacuum, because electromagnetic waves do not need a medium. The radiator panels are large, dark surfaces designed to maximise radiative emission into the cold of space (effective temperature about 3K), and they are the station's only route for rejecting waste heat.

EXERCISE 3B

13. (a) Marble has much higher thermal conductivity than wood. When the foot touches marble, heat flows rapidly from the warm skin into the cold marble, producing a strong sensation of cold. Wood conducts heat so slowly that very little energy leaves the foot in the same contact time, so the foot stays warm. The sensation of cold is caused by the rate of heat loss from the skin, not by the floor temperature.

(b) Rubber has low thermal conductivity (similar to wood). The sandal acts as an insulating layer between the foot and the marble, reducing the rate of heat flow from the skin. The foot loses heat much more slowly and feels less cold.

14. (a) Clay has low thermal conductivity compared to the high temperature gradient across it, but the fire is in direct contact with the clay body, supplying heat continuously. The wooden handle is separated from the fire and connected to the clay body through a narrow junction. Wood has very low thermal conductivity, so heat from the clay body conducts very slowly along the handle. The handle stays cool because the rate of heat arriving through the wood is too small to raise its temperature significantly.

(b) **First mechanism: evaporation.** The damp cloth loses water by evaporation, which draws latent heat from the handle and the cloth, cooling both. **Second mechanism: conduction.** The water in the cloth has higher thermal capacity than dry cloth, so it absorbs more heat before its temperature rises, acting as a temporary heat sink that delays the warming of the handle.

15. (a) No. The blanket generates no heat. It is a passive insulator that reduces the rate of heat transfer between the object it covers and the surroundings.

(b) **Conduction:** the blanket material and the air trapped within it have low thermal conductivity, resisting heat flow through the blanket's thickness. **Convection:** the blanket traps air in small pockets between its fibres, preventing convection currents from

carrying heat across the layer. **Radiation:** the blanket's fabric interrupts the direct line of sight between the warm body (or ice) and the surroundings, blocking radiative exchange.

(c) Fourier's law is symmetric: the rate of heat flow depends on the temperature difference and the thermal resistance, regardless of which side is hotter. The blanket increases the thermal resistance between the object and the surroundings. This slows heat flow outward (keeping the person warm) and inward (keeping the ice cold) by the same mechanism. The blanket does not choose a direction; it resists heat flow in both.

16. (a) A black surface has high absorptivity (close to 1), so it absorbs most of the incoming solar radiation and converts it to thermal energy. A shiny aluminium surface reflects most of the sunlight and absorbs very little. Painting the panel black maximises the solar energy captured by the panel.

(b) Natural convection (thermosiphon effect). The sun heats the water in the panel, making it less dense. The warm water rises naturally through the pipes to the storage tank above. Cooler, denser water from the bottom of the tank flows down to the panel to replace it. This density-driven circulation continues as long as the panel is hotter than the tank, with no pump needed.

(c) On a clear day, the solar intensity reaching the panel is much higher than on an overcast day, because clouds reflect and scatter a large fraction of the incoming sunlight. The rate of energy absorption by the panel depends on the intensity of the radiation striking it, not on the air temperature. Even at the same air temperature, the panel receives less radiant energy on a cloudy day and therefore heats the water less.

17. (a) Error 1: the greenhouse effect does not prevent heat from escaping. It absorbs outgoing infrared and re-emits some of it back toward the surface, but the rest escapes to space. The atmosphere is not a sealed container. **Error 2:** the temperature does not rise indefinitely. The system reaches an equilibrium where the rate of outgoing radiation from the top of the atmosphere equals the rate of incoming solar absorption. At that equilibrium, the temperature stabilises.

(b) The surface warms until it radiates enough infrared energy that the total outgoing radiation from the top of the atmosphere balances the incoming solar absorption. Greenhouse gases shift this equilibrium to a higher surface temperature (because the surface must be warmer to drive enough radiation through the absorbing atmosphere), but the system always finds a new balance.

(c) Without greenhouse gases, the atmosphere would be transparent to all wavelengths. The surface would radiate directly to space with no atmospheric absorption or re-emission. The equilibrium temperature would drop to about 255K (-18°C), far below the freezing point of water. The planet's surface would be permanently frozen.

18. (a) The hot tea loses heat faster. Its excess temperature over the room ($80 - 20 = 60^{\circ}\text{C}$) is much larger than the cold water's deficit ($20 - 10 = 10^{\circ}\text{C}$). Newton's law of cooling says the rate of heat transfer is proportional to the temperature difference, so the tea's heat loss rate is about six times larger than the cold water's heat gain rate.

(b) The cold water gains heat. It is below room temperature, so heat flows from the warm room air into the cold water by convection and conduction, and from the warm surroundings by radiation.

(c) Both will eventually reach 20°C (room temperature). The cold water reaches it first because it has a much smaller temperature gap to close (only 10°C versus 60°C for the tea). Even though the tea cools faster in absolute terms (more watts of heat flow), it has a much longer journey to make.

19. (a) Conduction: add a layer of insulating material (such as foam board or ceiling board) beneath the iron roof. The insulation has very low thermal conductivity, creating a high thermal resistance between the hot roof and the interior, reducing the rate of heat conduction into the house.

(b) Convection: install a ventilated air gap (ridge vents at the top, eave vents at the bottom) between the roof and the ceiling. Hot air beneath the iron roof rises and escapes through the ridge vents, pulling cooler outside air in through the eaves. This removes heat by natural convection before it reaches the living space.

(c) Radiation: paint the outer surface of the iron roof with reflective white or silver paint, or install a reflective foil barrier beneath the roof. A reflective surface has low absorptivity, so it absorbs less solar radiation. A foil barrier beneath the roof has low emissivity, so it radiates less heat downward into the house.

20. (a) Conduction: newspaper has low thermal conductivity and the multiple layers trap air between them, increasing the total thermal resistance. **Convection:** the newspaper wrapping prevents warm air from rising freely from the pot surface and being replaced by cooler air, suppressing natural convection.

(b) Newton's law of cooling says the rate of temperature fall is proportional to the excess temperature over the surroundings. When the *ugali* is very hot, the excess is large and the rate of heat loss is high, so the temperature falls quickly. As the *ugali* cools, the excess shrinks, the driving force for heat loss weakens, and the rate of cooling slows. This self-slowness is the nature of exponential decay: the rate depends on the current excess, not on the initial temperature.

(c) Aluminium foil is a better reflector of infrared radiation (low emissivity), so it would reduce radiative heat loss more effectively than newspaper. However, aluminium is a good thermal conductor, so it would reduce conductive and convective heat loss less

effectively than newspaper (which traps insulating air layers). The best solution would be aluminium foil on the outside (to reflect radiation) with newspaper underneath (to trap air and resist conduction and convection).

21. (a) Yes, Kipute is correct that the tea in the metal mug cooled faster. Metal conducts heat rapidly from the tea through the mug wall to the outer surface, where it is lost to the air by convection and radiation. Ceramic conducts heat more slowly, so less heat reaches the outer surface per second.

(b) Thermal conductivity determines how fast heat passes through the wall, but it is not the only factor. The thermal capacity of the mug material also matters. A ceramic mug is typically thicker and has higher specific heat capacity than a thin metal mug, so the ceramic absorbs and stores more heat from the tea before its outer surface warms enough to lose heat rapidly. The metal mug, being thin with low thermal capacity, warms through almost instantly and begins losing heat immediately. The rate of cooling depends on both the conductivity (how fast heat reaches the surface) and the thermal mass of the mug (how much heat the mug itself absorbs along the way).

(c) Yes. If no heat can escape from either mug, both systems (tea + mug) would reach thermal equilibrium internally. The tea would cool slightly as it shares heat with the mug material, and the mug would warm. The final temperature in each case would depend on the relative thermal capacities of the tea and the mug, which are different for metal and ceramic. So the final temperatures would be slightly different, but both would be close to the initial tea temperature because the tea's thermal capacity is much larger than either mug's.

22. (a) Foam has very low thermal conductivity compared to concrete. Adding the foam layer places a high thermal resistance in series with the concrete. By Fourier's law, the rate of heat flow through the composite is determined by the total thermal resistance, which is now much larger than concrete alone. The foam reduces the rate of heat conduction through the roof dramatically.

(b) In a series arrangement, the temperature drop across each layer is proportional to its thermal resistance. Foam has much higher thermal resistance per unit thickness than concrete (because its conductivity is much lower). Even though the concrete is thicker, its low resistance means only a small temperature drop occurs across it. Most of the total temperature difference between outside and inside appears across the thin foam layer, because that is where the resistance is concentrated.

(c) The foam's thermal resistance works identically regardless of which side is hotter. During the day, it resists heat flowing inward from the hot roof to the cool interior. At night, it resists heat flowing outward from the warm interior to the cold roof. The resistance is a property of the material, not of the direction of heat flow. This is a direct consequence of Fourier's law being symmetric with respect to the sign of the temperature difference.

EXERCISE 3C

23. $76 \text{ W m}^{-1} \text{ K}^{-1}$

24. 12.1 W m^{-2} , 18.7°C , 4.6

25. From Newton's law of cooling, the time for each interval:

$$t_1 = \frac{1}{\lambda} \ln \left(\frac{\theta_1 - \theta_s}{\theta_2 - \theta_s} \right)$$

$$t_2 = \frac{1}{\lambda} \ln \left(\frac{\theta_2 - \theta_s}{\theta_3 - \theta_s} \right)$$

Since the drops are equal, $\theta_1 - \theta_2 = \theta_2 - \theta_3 = d$ (say). Then $\theta_3 = \theta_2 - d$ and $\theta_1 = \theta_2 + d$. The excess temperatures are:

$$\theta_1 - \theta_s = \theta_2 - \theta_s + d \text{ and } \theta_3 - \theta_s = \theta_2 - \theta_s - d$$

The argument of the logarithm in t_2 is:

$$\frac{\theta_2 - \theta_s}{\theta_2 - \theta_s - d}$$

and in t_1 :

$$\frac{\theta_2 - \theta_s + d}{\theta_2 - \theta_s}$$

Let $E = \theta_2 - \theta_s$ (the excess at the midpoint). Then:

$$\frac{t_2}{t_1} = \frac{\ln \left(\frac{E}{E-d} \right)}{\ln \left(\frac{E+d}{E} \right)}$$

Since $E > d$ (the body is above the surroundings throughout), we have $\frac{E}{E-d} > \frac{E+d}{E}$ because subtracting d from a smaller denominator produces a larger ratio than adding d to a larger numerator. Therefore $\ln \left(\frac{E}{E-d} \right) > \ln \left(\frac{E+d}{E} \right)$, giving $t_2 > t_1$.

26. (a) The exact equation from Stefan-Boltzmann:

$$mc \frac{dT}{dt} = -\epsilon \sigma A (T^4 - T_s^4)$$

$$\frac{dT}{dt} = -\frac{\epsilon \sigma A}{mc} (T^4 - T_s^4)$$

(b) The right-hand side contains T^4 , not $(T - T_s)$. The equation is not of the form $dT/dt = -\lambda(T - T_s)$, so the solution is not a simple exponential. The T^4 dependence makes the equation nonlinear: the rate of cooling changes disproportionately as T changes, unlike Newton's law where the rate is strictly proportional to the excess.

(c) 0.00925 K s^{-1}

27. (a) 139W (b) 2.7W

28. $6.54 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$

29. 1.81W, 1450nm, 3%

30. (a) Furnace: 24200K, Star: 6042K (b) $P_{\text{furnace}} : P_{\text{star}} = 257 : 1$ (c) 9350nm (or 9.35 μm).

31. (a) Rate of heat loss per unit area through ice of thickness x :

$$\frac{dQ}{dt} = \frac{kA\Delta T}{x}$$

This heat freezes a thin layer dx of water:

$$dQ = \rho_{\text{ice}} A dx \times L_f$$

So:

$$\begin{aligned} \rho_{\text{ice}} A L_f \frac{dx}{dt} &= \frac{kA\Delta T}{x} \\ x dx &= \frac{k\Delta T}{\rho_{\text{ice}} L_f} dt \end{aligned}$$

Integrating from x_1 to x_2 :

$$\begin{aligned} \int_{x_1}^{x_2} x dx &= \frac{k\Delta T}{\rho_{\text{ice}} L_f} \int_0^t dt \\ \frac{1}{2} (x_2^2 - x_1^2) &= \frac{k\Delta T}{\rho_{\text{ice}} L_f} t \\ t &= \frac{\rho_{\text{ice}} L_f}{2k\Delta T} (x_2^2 - x_1^2) \end{aligned}$$

(b) Time ratio:

$$\frac{t_{10 \rightarrow 11}}{t_{1 \rightarrow 2}} = \frac{(0.11)^2 - (0.10)^2}{(0.02)^2 - (0.01)^2} = 7.0$$

Growing from 10cm to 11cm takes 7 times longer than growing from 1cm to 2cm.

32. Two cylindrical shells in series. Radial thermal resistance of each:

$$R = \frac{\ln\left(\frac{r_2}{r_1}\right)}{2\pi kL}$$

Copper shell ($r_1 = 0.020\text{m}$, $r_2 = 0.030\text{m}$):

$$R_{\text{Cu}} = \frac{\ln(0.030/0.020)}{2\pi \times 380 \times 5.0} = \frac{0.4055}{11938} = 3.40 \times 10^{-5} \text{ KW}^{-1}$$

Asbestos shell ($r_1 = 0.030\text{m}$, $r_2 = 0.070\text{m}$):

$$\begin{aligned} R_{\text{asb}} &= \frac{\ln(0.070/0.030)}{2\pi \times 0.08 \times 5.0} = \frac{0.8473}{2.513} = 0.337 \text{ KW}^{-1} \\ R_{\text{total}} &= 3.40 \times 10^{-5} + 0.337 \approx 0.337 \text{ KW}^{-1} \end{aligned}$$

(Copper resistance is negligible.)

$$\frac{dQ}{dt} = \frac{58\text{K}}{0.337 \text{ KW}^{-1}} = 172\text{W}$$

Temperature at copper-asbestos boundary:

$$\theta_{\text{boundary}} = 80^\circ\text{C} - 172\text{W} \times 3.40 \times 10^{-5} \text{ KW}^{-1} = 80^\circ\text{C} - 0.006^\circ\text{C} \approx 80.0^\circ\text{C}$$

(The copper wall is essentially isothermal: the entire temperature drop occurs across the asbestos.)

33. (a) With albedo, the power absorbed by Earth is:

$$P_{\text{abs}} = (1 - \alpha) \varphi_s \pi R_e^2$$

The power radiated by Earth as a blackbody sphere:

$$P_{\text{rad}} = \sigma (4\pi R_e^2) T_e^4$$

At equilibrium:

$$\begin{aligned} (1 - \alpha) \varphi_s \pi R_e^2 &= \sigma \times 4\pi R_e^2 T_e^4 \\ T_e^4 &= \frac{(1 - \alpha) \varphi_s}{4\sigma} \end{aligned}$$

With $\alpha = 0$ (no reflection), this gives the bare-blackbody value. With $\alpha = 0.30$, the factor $(1 - \alpha) = 0.70$ reduces the numerator, so T_e^4 is smaller and therefore T_e is lower than the bare-blackbody value.

(b) Ice and snow have high albedo (0.8 to 0.9); ocean water has low albedo (0.06 to 0.10). If ice caps melt, they are replaced by dark ocean, which absorbs more sunlight. The overall planetary albedo decreases. With lower α , the factor $(1 - \alpha)$ increases, so more solar energy is absorbed, and the equilibrium temperature rises further.

(c) This is a destabilising (positive) feedback. A small initial warming melts some ice, which lowers the albedo, which causes more warming, which melts more ice, which lowers the albedo further. Each step amplifies the initial perturbation rather than counteracting it. This is called the ice-albedo feedback and is one of the most important positive feedbacks in climate physics.

EXERCISE 4A

1. Errors:

First error: The assumption that molecules touch each other at all; in the ideal gas model the molecules are taken as essentially point masses that interact only briefly at collisions.

Second error: Identifying the cause of pressure with collisions between molecules, whereas pressure is produced by collisions of the molecules with the walls of the container.

Correct reasoning:

Compressing the gas reduces the container volume, so each molecule travels a shorter distance between successive wall collisions and strikes the walls more frequently per second; the increased rate of momentum transfer to the walls is the true cause of the pressure increase.

2. (a) Intermolecular forces are assumed to be negligible except during the instant of a collision, so there is no force acting on a molecule in the intervals between collisions. With no force there is no acceleration, and by Newton's first law the molecule moves in a straight line at constant velocity.

(b) Assuming the collision time is negligible allows us to treat each molecule as moving freely between instantaneous collisions. The assumption simplifies the derivation of pressure by making it unnecessary to track the detailed mechanics of what happens during a collision and enables the momentum change at the wall to be treated as a sudden reversal.

3. The average rate at which molecules strike a square centimetre of wall is the same in both containers. Pressure is force per unit area on the wall, and the two containers have the same pressure; since all walls experience the same rate of momentum transfer per unit area, the rate at which molecules strike one square centimetre is the same on any wall, whether flat or curved, in either container.

4. (a) The average kinetic energy per molecule of an ideal gas is proportional to the Kelvin temperature, not the Celsius temperature, and the two scales have different zeros. Doubling the Celsius temperature does not double the Kelvin temperature, so the kinetic energy does not double.

(b) The student's rule requires two conditions simultaneously:

Doubling in Celsius ($\theta_2 = 2\theta_1$ in $^{\circ}\text{C}$) and

Doubling in Kelvin ($\theta_2 + 273 = 2(\theta_1 + 273)$ in K).

Substituting the first into the second gives $2\theta_1 + 273 = 2\theta_1 + 546$, which reduces to $273 = 546$, a contradiction. No pair of Celsius temperatures satisfies both conditions, so the student's rule never gives the correct ratio of kinetic energies exactly.

The rule becomes approximately correct in the limit of very high temperatures, where $\theta_1 \gg 273^{\circ}\text{C}$. In this limit the additive offset of 273 is negligible compared with the Celsius values themselves, so the Celsius and Kelvin temperatures are approximately equal and their ratios are approximately the same. At stellar temperatures (thousands of kelvin) the rule holds to good accuracy.

5. (a) The average kinetic energy per molecule and the root mean square speed are the same in both containers, because both depend only on the Kelvin temperature.

(b) The pressure and the total internal energy are both twice as large in container B, because both scale in proportion to the number of moles (and hence the number of molecules) at the same temperature.

(c) An intensive quantity, such as temperature, pressure, or kinetic energy per molecule, is independent of the amount of matter present. An extensive quantity, such as total internal energy or total number of molecules, scales in proportion to the amount of matter. Doubling the amount of gas doubles the extensive quantities and leaves the intensive quantities unchanged.

6. The student has confused two different statistical operations. The root mean square speed is obtained by first squaring each speed, then averaging the squares, then taking the square root; the simple arithmetic mean averages the speeds directly without squaring. Whenever the molecular speeds differ from one another, squaring gives extra weight to the larger speeds, and the rms speed exceeds the arithmetic mean. The two would agree only in the special case where every molecule in the sample moves at exactly the same speed.

7. (a) A diatomic molecule consists of two atoms connected by a bond, and we can take the line joining the atoms as the bond axis. Rotation about an axis perpendicular to the bond carries the atoms in circular paths of finite radius and stores energy. Rotation about the bond axis itself, however, rotates the atoms around their own centres and stores negligible energy because the moment of inertia about that axis is very small (the atoms lie essentially on the axis). Only two of the three possible rotational axes therefore store appreciable kinetic energy, giving two rotational degrees of freedom.

(b) A polyatomic molecule such as water vapour has its atoms arranged in a non-linear configuration, so no one axis passes through every atom. All three perpendicular rotational axes carry the atoms in circular paths of non-zero radius and have non-zero moments of inertia. All three rotations therefore store appreciable kinetic energy, giving three rotational degrees of freedom.

8. The student's error is to identify work done on the gas with a rise in internal energy. The internal energy of an ideal gas depends only on its absolute temperature, not on its volume or pressure. Since the temperature is held constant during the compression, the internal energy does not change; the work done on the gas does not accumulate as increased molecular kinetic energy, because molecular kinetic energy is fixed by the temperature alone.

9. (a) For a real gas, the pressure rises by more than a factor of ten because the volume available to the molecules shrinks faster than the container volume: the volume occupied by the molecules themselves, negligible at the original volume, becomes a significant fraction of the container at the smaller volume. The assumption of negligible molecular volume (assumption 2 of the kinetic theory) fails, so the real gas pressure exceeds what the ideal gas equation predicts.

(b) When the compressed gas is cooled to a very low temperature, the attractive intermolecular forces begin to dominate over the slower-moving molecules, and the gas may liquefy. The ideal gas model, which assumes zero intermolecular forces, cannot describe liquefaction because in the absence of attractive forces there is nothing to hold molecules together into a liquid phase.

10. The rms speed of an ideal gas is proportional to the square root of the Kelvin temperature:

$$v_{\text{rms}} \propto \sqrt{T}$$

The factor by which it changes is therefore:

$$\frac{v_{\text{rms}_2}}{v_{\text{rms}_1}} = \sqrt{\frac{T_2}{T_1}} = \sqrt{\frac{1200\text{K}}{300\text{K}}} = 2$$

The root mean square speed doubles.

EXERCISE 4B

11. (a) At the higher temperature the nitrogen or LPG molecules inside the cylinder move with a greater mean square speed. Each molecule delivers a larger momentum change when it strikes a wall, and each molecule strikes the walls more frequently per second. Both effects multiply the rate of momentum transfer to the walls, and the pressure, which is this rate per unit area, rises in direct proportion to the Kelvin temperature at fixed volume.

(b) A cylinder stored in a closed, hot room has no way to lose heat to the surroundings and its internal temperature can climb further; the pressure would rise with it, possibly exceeding the safety rating and causing the cylinder to rupture. A shaded, well-ventilated place keeps the cylinder near ambient temperature by both blocking solar radiation and allowing convective heat loss to circulating air, keeping the internal pressure safely below the rating.

12. What has changed with altitude is the total atmospheric pressure, which at the summit of Kilimanjaro is about half its sea-level value. Since the fractional composition of air has not changed, the partial pressure of oxygen is also about half its sea-level value. The rate at which oxygen molecules diffuse from the lung air into the blood is driven by the partial-pressure gradient between the two, and with that gradient halved, the rate of oxygen uptake is reduced correspondingly. The mountain climber must breathe harder and more frequently to move the same amount of oxygen into his blood, and at the summit the body's demand begins to outstrip what his breathing can deliver.

13. (a) As the contents of the cooker are heated, the water molecules in the vapour phase move with greater mean square speed and strike the walls of the cooker more frequently and with greater momentum change per collision. The pressure of the vapour, which is the rate of momentum transfer to the walls per unit area, rises in direct proportion to the Kelvin temperature of the vapour at fixed volume.

(b) The pressure reaches a stable value when the safety valve opens and begins releasing steam. The escape of molecules through the valve removes them from the container and prevents the number density (and hence the pressure) from rising further. The valve maintains a near-constant pressure by venting just enough steam to balance the rate at which more water vapourises from the heated contents.

14. (a) Dalton's law states that each gas in a mixture exerts the same partial pressure as it would if it alone occupied the container at the same temperature. At sea level the total pressure is one atmosphere and the partial pressure of nitrogen is about 78% of this, roughly 0.78 atmospheres. At 20 metres depth the diver breathes air compressed to three atmospheres; the fractional composition

has not changed, so each species' partial pressure is multiplied by three, giving a nitrogen partial pressure of about 2.34 atmospheres, approximately three times its surface value.

(b) At higher partial pressure, the number density of nitrogen molecules in the blood and tissues is correspondingly higher, so the rate at which nitrogen molecules collide with receptors, cell membranes, and other biological targets is also correspondingly higher. Even if each individual collision has negligible chemical effect, the cumulative rate of such collisions becomes large enough to disturb nerve transmission and produce the narcotic symptoms observed. Physiological activity of an inert gas depends on collision rates, not on chemistry.

15. Bakari is correct that air is a mixture of several gases, mainly nitrogen and oxygen, not a single kind of atom; Ally's wording is sloppy. But Ally's physical intuition is correct. At the same temperature, helium atoms (molar mass 4 g/mol) and the average air molecule (effective molar mass about 29 g/mol) carry the same average kinetic energy per molecule. A container filled with helium at a given temperature and pressure therefore holds the same number of molecules per unit volume as a container of air at the same temperature and pressure, but each helium atom is about seven times lighter than the average air molecule. The helium-filled balloon displaces a volume of air whose mass greatly exceeds the mass of the helium inside, and the buoyant force exceeds the weight of the balloon, causing it to rise. Ally's statement captures this mass-per-molecule comparison, even if the language of "atoms of air" is imprecise.

16. (a) Argon is monatomic, with only three translational degrees of freedom, while air consists mainly of diatomic nitrogen and oxygen molecules with five degrees of freedom (three translational plus two rotational). An argon atom therefore carries less thermal energy per collision than an air molecule and transports heat across the gap less effectively. In addition, argon atoms are heavier (molar mass 40 g/mol) than the average air molecule (29 g/mol), so at the same temperature the argon atoms move more slowly and arrive at the opposite pane less frequently. Both effects reduce the thermal conductivity of argon below that of air, making argon a better insulator.

(b) Krypton is also monatomic, so its number of degrees of freedom is the same as argon's (three). The insulation would not improve through that property. However, krypton atoms are much heavier than argon atoms (molar mass 84 g/mol compared with 40 g/mol), so krypton atoms move even more slowly at the same temperature and transport heat across the gap even less rapidly. The mass effect improves, the degrees-of-freedom effect does not; on balance, krypton provides better insulation than argon, which is why premium high-efficiency windows use it.

17. When Kipute pushes the handle down, the piston reduces the volume of air trapped in the pump cylinder rapidly. Because the compression is fast, the air has no time to lose heat to the surroundings through the walls; the compression is effectively an insulated one. The work done on the air during compression goes entirely into raising the kinetic energy of the air molecules, which means the temperature and the mean square speed of the molecules inside the pump have increased. These faster, hotter molecules then transfer their extra thermal energy by direct contact to the inner wall of the pump barrel, warming the metal. The metal feels warm not because of any external source of heat but because Kipute's muscular work has been converted, through the compression, into thermal energy of the gas and thence of the pump.

18. $T_{\text{well}} = 24 + 273 = 297\text{K}$; $T_{\text{street}} = 36 + 273 = 309\text{K}$.

Since $v_{\text{rms}} \propto \sqrt{T}$ for a given gas, the ratio of rms speeds is:

$$\frac{v_{\text{rms,street}}}{v_{\text{rms,well}}} = \sqrt{\frac{T_{\text{street}}}{T_{\text{well}}}} = \sqrt{\frac{309\text{K}}{297\text{K}}} = 1.020$$

The molecules on the hot street move only about 2% faster on average than those in the cool well. The word "much" is **inappropriate**: a 12°C difference in Celsius looks substantial in human terms, but on the Kelvin scale it represents only a 4% change in absolute temperature, and molecular speeds respond to the square root of that, giving barely 2%. The student has fallen into the familiar trap of treating Celsius differences as though they were fractional changes in the physically meaningful quantity, which is the Kelvin temperature.

EXERCISE 4C

19. Find the amount of nitrogen present from $PV = nRT$:

$$n = \frac{P_1 V}{RT_1} = \frac{(8.0 \times 10^6 \text{Pa}) \times (0.040 \text{m}^3)}{(8.314 \text{Jmol}^{-1}\text{K}^{-1}) \times (290\text{K})} = 132.7 \text{mol}$$

At constant volume the heat supplied equals the change in internal energy. Nitrogen is diatomic ($f = 5$), so $\Delta U = (f/2)nR\Delta T$:

$$Q = \frac{5}{2} \times (132.7 \text{mol}) \times (8.314 \text{Jmol}^{-1}\text{K}^{-1}) \times (110\text{K}) = 3.03 \times 10^5 \text{J}$$

The heat supplied to the gas is approximately $3.03 \times 10^5 \text{J}$.

20. Equating the right-hand sides of $PV = \frac{1}{3}Nm\bar{c}^2$ and $PV = NkT$:

$$\frac{1}{3} Nm\bar{c}^2 = NkT$$

$$\bar{c}^2 = \frac{3kT}{m}$$

Taking the square root and using $m = M/N_A$ together with $R = k \times N_A$:

$$v_{\text{rms}} = \sqrt{\frac{3kT}{m}} = \sqrt{\frac{3RT}{M}}$$

For two gases A and B at the same temperature T, dividing one expression by the other gives:

$$\frac{v_{\text{rmsA}}}{v_{\text{rmsB}}} = \sqrt{\frac{M_B}{M_A}}$$

The ratio of rms speeds therefore depends only on the molar masses, not on temperature: raising the temperature accelerates both gases equally in the sense that their kinetic energies remain equal, so their speeds remain in the same mass-determined ratio.

21. The neon cylinder by a factor of 1.05.

22. The gas is helium. The identification is unique among common gases because helium is the only gas at A-Level with a molar mass of 4.0g/mol: the next heaviest common gas is neon at 20g/mol, five times heavier, far outside any reasonable rounding of the calculated value.

23. Newton's law of cooling, integrated, gives:

$$T(t) = T_s + (T_1 - T_s) e^{-\lambda t}$$

With $T_1 = 400\text{K}$, $T_s = 300\text{K}$, $\lambda = 2.0 \times 10^{-3}\text{s}^{-1}$, and $t = 600\text{s}$: $\lambda t = 1.2$ and $e^{-1.2} = 0.301$:

$$T(600\text{s}) = 300\text{K} + (100\text{K}) \times 0.301 = 330.1\text{K}$$

At constant volume and fixed n, the pressure is proportional to the Kelvin temperature:

$$P_2 = P_1 \times \frac{T_2}{T_1} = (5.0 \times 10^5\text{Pa}) \times \frac{330.1\text{K}}{400\text{K}} = 4.13 \times 10^5\text{Pa}$$

The pressure inside the container 10 minutes after placement is approximately $4.13 \times 10^5\text{Pa}$.

24. Initial total moles: $n = 1.5 + 2.0 + 0.5 = 4.0\text{mol}$. Final total moles (after removal of water vapour): $n' = 3.5\text{mol}$. Pressure at constant V and T depends only on total moles:

$$\frac{\Delta P}{P} = \frac{n_f - n}{n} = \frac{3.5 - 4.0}{4.0} = -0.125 = -12.5\%$$

The pressure decreases by 12.5%. Internal energy before removal, using $f = 3$ for He, $f = 5$ for N_2 , and $f = 6$ for water vapour:

$$U_i = \frac{RT}{2} \times (3 \times 1.5\text{mol} + 5 \times 2.0\text{mol} + 6 \times 0.5\text{mol}) = \frac{RT}{2} \times 17.5\text{mol}$$

$$U_f = \frac{RT}{2} \times (3 \times 1.5\text{mol} + 5 \times 2.0\text{mol}) = \frac{RT}{2} \times 14.5\text{mol}$$

Percentage change in internal energy:

$$\frac{\Delta U}{U} = \frac{14.5 - 17.5}{17.5} = -0.1714 = -17.1\%$$

The internal energy decreases by 17.1%, considerably more than the 12.5% pressure drop. The two percentages differ because pressure at fixed V and T depends only on the total moles, so removing any species reduces the pressure in proportion to the moles removed. Internal energy, however, counts each species weighted by its degrees of freedom; water vapour is polyatomic ($f = 6$) and carries more energy per mole than either helium or nitrogen, so removing it takes away a larger fraction of the total internal energy than of the total moles.

25. From $P = \alpha V$ and $PV = nRT$, substituting $P = \alpha V$:

$$\alpha V \times V = nRT, \text{ so } T = \frac{\alpha V^2}{nR}$$

The absolute temperature is therefore proportional to V^2 . When the volume doubles, the temperature becomes $(2V_0/V_0)^2 = 4$ times its original value:

$$\frac{T_f}{T_0} = (2)^2 = 4$$

That is, $T_f = 4T_0$. The rms speed scales as the square root of the Kelvin temperature:

$$\frac{v_{\text{rmsf}}}{v_{\text{rms0}}} = \sqrt{\frac{T_f}{T_0}} = \sqrt{4} = 2$$

The rms speed has doubled.

For a monatomic gas $U = (3/2)nRT$, so U is directly proportional to T:

$$\frac{U_f}{U_0} = \frac{T_f}{T_0} = 4$$

The final internal energy is four times the initial internal energy.

EXERCISE 5A

1. Both forms of the first law are correct statements of energy conservation; they differ only in the sign convention adopted for W. In the form $\Delta U = Q - W$, the symbol W denotes the work done by the gas on the surroundings: positive when the gas expands,

negative when it is compressed. In the form $\Delta U = Q + W$, the symbol W denotes the work done on the gas by the surroundings: positive when the gas is compressed, negative when it expands. The two values of W differ only by a sign, $W_{\text{by}} = -W_{\text{on}}$, so the two equations are algebraically identical: heat in minus work-done-by equals heat in plus work-done-on, and both equal ΔU . The students are therefore both correct, provided each is consistent within their chosen convention; the disagreement is bookkeeping, not physics.

2. The error is to assume that heat exiting a gas must reduce its internal energy. For an ideal gas the internal energy depends only on temperature, $U = U(T)$, so as long as T is held constant, $\Delta U = 0$ regardless of how much heat flows in or out. During the isothermal compression the surroundings do positive work on the gas, raising the internal energy, while heat simultaneously leaves the gas, lowering it; with $\Delta U = 0$, the two effects must cancel exactly. The first law $\Delta U = Q - W$ (with W done **by** the gas, hence negative in compression) gives $Q = W < 0$, the heat removed equals in magnitude the work done on the gas. The work added to the gas leaves the gas as heat, not as a rise in internal energy.

3. The internal energy U is a state function: its value at any equilibrium state is determined by the state alone (for an ideal gas, by the temperature alone), independently of how the gas reached that state. At state A the gas has a specific internal energy U_A ; at state B, a specific U_B ; the change $\Delta U = U_B - U_A$ is fixed by the two endpoints and is independent of the path joining them. The work done by the gas, $W = \int P dV$, is the area under the chosen path on the PV diagram; different paths enclose different areas and therefore yield different values of W . The heat exchanged then follows from the first law as $Q = \Delta U + W$, and since ΔU is path-independent while W is not, Q also depends on the path. Both W and Q are path functions; only ΔU is a state function.

4. In a gas, heating at constant pressure produces a substantial volume change, since gases expand strongly with temperature. Part of the heat input is therefore expended as work done by the gas on the surroundings, $W = P\Delta V$, on top of raising the internal energy. At constant volume no such work is done, so all of the heat goes into raising U . The difference between the two heat capacities, $C_p - C_v = R$ per mole for an ideal gas, accounts exactly for the $P\Delta V$ term. In a solid or liquid the thermal expansion coefficient is much smaller (of order 10^{-4}K^{-1} for liquids, 10^{-5}K^{-1} for solids), so the volume change associated with a temperature rise is small, and the work done against atmospheric pressure during heating is negligible compared with the heat absorbed. The difference $C_p - C_v$ does exist in principle for solids and liquids but is too small to matter for practical calculations, so a single specific heat capacity is quoted for each substance. The student's observation is therefore explained by the magnitude of thermal expansion, not by any difference in the underlying physics.

5. For any ideal gas the molar heat capacity at constant volume is $C_v = \frac{f}{2}R$, where f is the number of accessible degrees of freedom. Equal C_v therefore implies equal f . So far the student's chain of inference is sound: A and B share the same number of energy-storage modes per molecule. What this conclusion does **not** establish is the molecular identity of the gases, since many different molecules share the same f . Two distinct diatomic gases at room temperature, for example, both have $f = 5$ (three translational + two rotational) and therefore the same C_v : nitrogen and oxygen are both diatomic but are chemically distinct; the helium isotopes ^3He and ^4He are both monatomic with $f = 3$ but differ only in their nuclear mass. To distinguish A from B beyond their degree-of-freedom count, a further measurement is required: the molar mass (which follows from the density at known temperature and pressure via $\rho = PM/RT$), the absorption spectrum (which is unique to each molecule), or a chemical assay. The molar C_v tells you how many ways a molecule stores energy, not which molecule it is.

6. The vessel is welded, so its internal volume is fixed; the gas it contains undergoes an isochoric cooling. For a fixed mass of ideal gas at constant volume the equation of state $PV = nRT$ reduces to $P/T = \text{constant}$. Converting the temperatures to the Kelvin scale gives $T_1 = 305\text{K}$ and $T_2 = 295\text{K}$, hence:

$$\frac{P_2}{P_1} = \frac{T_2}{T_1} = \frac{295\text{K}}{305\text{K}} = 1 - \frac{10}{305}$$

The pressure falls by exactly 10/305 of its original value, neither more nor less. The assumption that underlies this exact result is the **ideal-gas** behaviour of the contents of the vessel over the modest 10K temperature interval. For typical industrial gases at temperatures well above their boiling points and at moderate pressures, this assumption is well justified.

7. By Avogadro's law, which follows from the ideal-gas equation of state $PV = nRT$, equal volumes of any gas at the same temperature and pressure contain the same number of moles, and therefore the same number of molecules. Both balloons have the same volume, the same temperature, and the same pressure, so both contain the same number N of molecules. The mass difference between the two balloons originates entirely in the molar mass of the gas filling them: helium is monatomic with molar mass about 4g mol^{-1} , while air is a mixture dominated by diatomic nitrogen (28g mol^{-1}) and oxygen (32g mol^{-1}), with an effective molar mass of about 29g mol^{-1} . The number of molecules per unit volume is the same; the mass of each molecule is not. The molecular property responsible for the lighter helium balloon is therefore the molar mass, not any difference in molecular count.

8. The error is to assume that any heat flow must change the temperature. Heat input raises the temperature only if the gas's internal energy is forced to rise; when the gas can simultaneously do work on its surroundings, the heat input emerges as work output rather than as a temperature rise. During an isothermal expansion of an ideal gas, $\Delta U = 0$ (since U depends only on T), and the first law reduces to $Q = W$: every joule of heat absorbed is converted to a joule of work done by the gas on the piston. Heat does flow, in substantial amounts; it just leaves the gas as mechanical work rather than appearing as kinetic energy of the molecules. In the

isothermal case the heat flow keeps the molecules at constant average kinetic energy while the gas expands, the heat input being passed straight through to the surroundings as work.

9. The claim is partially correct. A genuinely fast process does prevent heat from flowing through the container walls during the brief duration of the change, so to a good approximation $Q \approx 0$; this is what 'adiabatic' means in the heat-isolation sense. The relation $PV^\gamma = \text{constant}$, however, requires more. Its derivation assumes that at every instant during the process the gas is in thermodynamic equilibrium, so that a single pressure and temperature can be defined for the whole sample, and the work done in an infinitesimal step is $dW = PdV$. This is the quasi-static condition. A truly rapid event, such as a tyre suddenly bursting or a piston released from rest, can be heat-isolated ($Q \approx 0$) but is not quasi-static; the gas is far from equilibrium during the event, P and T vary across the sample, and the relation $PV^\gamma = \text{constant}$ does not apply. The condition that must be added to 'happens rapidly' is therefore that the process be quasi-static: slow enough on the equilibration timescale of the gas, but fast enough on the heat-conduction timescale of the walls.

10. Both compressions begin at the same initial point A on the PV diagram and end at the same final volume $V_2 < V_1$. The isothermal follows the curve $PV = \text{constant}$, a hyperbola, and ends at point B with pressure $P_B = P_A V_1/V_2$. The adiabat follows $PV^\gamma = \text{constant}$, a steeper hyperbola, and ends at point C with pressure $P_C = P_A (V_1/V_2)^\gamma$. Since $\gamma > 1$, $P_C > P_B$: the adiabatic compression reaches the higher final pressure.

Sketch description (the student's hand-drawn version): an axis labelled V horizontally, P vertically; a common starting point A; two downward curves drawn to the left, both ending at $V = V_2$; the upper curve is the adiabat (steeper, ending higher), the lower curve is the isotherm (shallower, ending lower); end-states marked B (isothermal) and C (adiabatic) with $P_C > P_B$.

The molecular reason: in the isothermal compression heat must leave the gas to keep T constant, so the molecular kinetic energy stays unchanged and the pressure rises only as much as the volume reduction itself compels through $PV = \text{constant}$. In the adiabatic compression no heat leaves, so the work done on the gas raises both the internal energy and the temperature; the higher temperature gives faster, more frequent wall collisions, and the pressure rises more steeply than the volume reduction alone would predict.

11. (a) The net work done by the gas over a complete cycle is $W_{\text{net}} = \oint PdV$ ($\oint$ means integration over a complete closed cycle), which equals the area enclosed by the cycle on the PV diagram with sign. By the standard sign convention (clockwise positive), counter-clockwise traversal gives $W_{\text{net}} < 0$: the surroundings do net work on the gas over the cycle. The sign is therefore, **negative**.

(b) Over a complete cycle the gas returns to its starting state, so $\Delta U_{\text{cycle}} = 0$ and the first law reduces to $Q_{\text{net}} = W_{\text{net}}$. Since $W_{\text{net}} < 0$, the net heat absorbed by the gas is also **negative**; on balance, heat flows out of the gas to the surroundings.

(c) The device consumes work and rejects net heat to the surroundings. This is the operating signature of a refrigerator (or, equivalently, a heat pump), **not** of a heat engine. (A heat engine traverses its cycle clockwise, producing positive net work and absorbing positive net heat; the counter-clockwise device does the opposite of both.)

12. A heat pump does not generate heat; it **moves** heat from one place to another, with the work input paying for the transport rather than for the heat itself. Applying the first law to the working substance over a complete cycle, $\Delta U_{\text{cycle}} = 0 = Q_{\text{cold}} - Q_{\text{hot}} + W_{\text{in}}$, where Q_{cold} is the heat absorbed from the cool exterior, Q_{hot} is the heat rejected to the warm interior, and W_{in} is the electrical work supplied. Rearranging, $Q_{\text{hot}} = Q_{\text{cold}} + W_{\text{in}}$. The heat delivered to the house is therefore the sum of the heat extracted from outside and the electrical work consumed; both contribute, and the total exceeds the work input by exactly the amount of heat extracted from the exterior. The first law is satisfied because energy is conserved: the outside cools slightly to compensate for the energy delivered to the inside. The Coefficient of Performance for heating, $\text{COP}_{\text{heating}} = Q_{\text{hot}}/W_{\text{in}}$, can therefore be considerably greater than unity (typically 3 to 5 for a domestic unit), without any violation of energy conservation.

13. The two statements are not unrelated; they are two faces of a single underlying principle, namely that there is a preferred direction for natural processes that energy conservation alone does not capture. Heat flows spontaneously from hot to cold (Clausius), and a cyclic engine cannot extract all of the heat absorbed as useful work without rejecting some to a cooler reservoir (Kelvin-Planck). Both are observed truths about the directionality of physical processes, and the two are logically equivalent: a hypothetical device that violated one could be combined with a normal device to violate the other.

For example, suppose a perfect engine existed that converted all heat absorbed from a single hot reservoir into useful work, in violation of Kelvin-Planck. Couple its work output to drive a Carnot refrigerator, which uses that work to pump heat from a cold reservoir to the same hot reservoir. The combination would extract heat from the cold reservoir, dump it into the hot reservoir, and produce no other effect, in violation of Clausius. The construction works in reverse too, starting from a hypothetical Clausius violator. Both statements therefore stand or fall together: they express a single physical principle in two convenient phrasings, not two separate rules.

EXERCISE 5B

14. The technician confuses energy conservation, which is the first law itself, with the much stronger claim that heat input equals work input. The first law states $\Delta U = Q - W$ (with W done by the gas), or equivalently $Q = \Delta U + W$; the heat exchanged with

the surroundings goes partly into raising the gas's internal energy and partly into work. During compression the surroundings do positive work on the gas (W done by gas is negative); the gas's internal energy rises and the heat exchanged is in general not equal in magnitude to the work, except in the special case of an isothermal process where the temperature happens to be held fixed.

The everyday observation that demonstrates the flaw: any air compressor at a factory or petrol station feels noticeably hot during operation, and a stream of compressed air leaving a fitting is warm to the touch. The warmth is the rise in internal energy that the work has produced; the heat exchanged with the surroundings is less than the work input by exactly this amount. If heat input and work input were always equal, the gas would never get warm during a compression, contradicting daily experience.

15. The two interpretations have opposite signs in the first-law equation, and lead to opposite conclusions about what is happening to the internal energy of the gas. Reading 'the gas doing 5000J of work' as $W_{by} = +5000J$ (work done by the gas on the surroundings) means the gas is expanding, and from $\Delta U = Q - W$ the internal energy is falling unless heat is supplied to compensate. Reading the same phrase as $W_{on} = +5000J$ (work done by the surroundings on the gas) means the gas is being compressed, and the internal energy is rising unless heat is removed. Whether the gas is heating up or cooling down therefore depends on which interpretation the reader assumes.

A careful technician records the direction explicitly: 'the gas expanded against the piston, doing 5000J of work on the piston' if the gas is the work source, or 'the piston compressed the gas, with the gas absorbing 5000J of work from the piston' if the surroundings are. Once the direction is specified, the sign of W in the chosen convention is fixed and the first-law bookkeeping follows without further ambiguity. The ambiguity is a labelling defect, not a physics one.

16. The cylinder's steel walls are rigid: its internal volume is fixed. Heating from the surrounding fire is therefore an isochoric process, and the first law at constant volume reduces to $Q = \Delta U = nC_v\Delta T$. All heat absorbed from the fire is converted into a rise in internal energy of the gas, which means a rise in temperature, which (via the equation of state $PV = nRT$ at constant V) means a proportional rise in pressure. The pressure can rise rapidly as the temperature climbs, eventually exceeding the cylinder's rated working pressure and rupturing the steel walls with explosive force. The high-pressure rating is set for normal use, not for fire conditions.

The safety feature manufacturers add is a pressure-relief valve, typically a frangible disc or a spring-loaded valve set to open well below the rupture pressure of the cylinder. As the pressure rises in the fire, the valve opens before the cylinder fails, allowing the gas to escape and ignite as a controlled flame rather than as a destructive explosion. The valve converts a potential bomb into a controlled torch, which trained firefighters can manage.

17. In an open saucepan the surface of the milk is exposed to atmospheric pressure, which is constant on the timescale of the heating. The milk therefore undergoes an isobaric thermal expansion: as its temperature rises, its volume increases at constant pressure, governed by the volumetric expansion coefficient. With the cross-sectional area of the pan fixed, the increase in volume shows up as an increase in depth, and the surface visibly rises.

Once the milk reaches its boiling point, additional heat input goes into the latent heat of vaporisation, with the temperature pinned at the boiling point rather than rising further. With the temperature no longer increasing, the isobaric thermal expansion of the bulk liquid stops, and the surface stops rising for that reason. (Boiling itself produces some apparent rise as bubbles lift the surface, but that is a different effect from thermal expansion of the liquid.)

18. The dominant effect is the adiabatic compression of the air trapped inside the pump barrel during each downstroke. The downstroke takes a small fraction of a second, too short for the air to lose appreciable heat to the metal walls, so the work done by the cyclist on the air is converted almost entirely into a rise in internal energy and hence in temperature. Successive strokes deposit heat into the metal barrel by conduction; over many strokes the barrel accumulates the deposited heat and becomes noticeably warm to the touch.

Friction in the pump seal is at most a secondary contributor for two reasons. First, the seal area is small (a thin annular contact between the piston and the cylinder wall), so the rate at which friction can generate heat is limited. Second, the relative speed at the seal is the speed of the piston, comparable to a metre per second, which is small compared with the speeds of the gas molecules; the energy dissipated per unit time at the seal is therefore much less than the energy delivered as compression work into the bulk gas.

Direct observation closes the case: the entire pump barrel becomes warm, including parts far from the seal. The friction-only explanation predicts heating localised at the seal alone, which is not what is observed. Compressed-air heating, distributed throughout the barrel and conducted outward through the cylinder wall, matches the observation; friction does not.

19. Mr. Akilikubwa would say: the four cornerstone processes are canonical idealisations, chosen for their pedagogical clarity, not for any requirement that real engines must use them equally. Real engines are built around mechanical and thermal constraints, not around a balanced curriculum, and the cycle that emerges reflects what the engineer can build.

In a four-stroke internal combustion engine, the compression and power strokes are fast: a few milliseconds each at typical operating speeds. The gas does not have time to lose heat to the cylinder walls during a single stroke, so each is well approximated by an adiabat. The intake and exhaust strokes occur at near-atmospheric pressure (set by the manifolds and the open valve to the atmosphere or exhaust system), so each is well approximated by an isobar.

Isothermal operation would require slow strokes with active heat exchange between the gas and a constant-temperature reservoir, which is impractical for an engine running at thousands of cycles per minute. An isochoric leg would require holding the piston motionless during heating; this happens briefly only at top dead centre during combustion (and is sometimes drawn as a near-vertical rise on the indicator diagram), but contributes essentially no useful work. The engine cycle is therefore structured by what the engineer can build with cams, crankshafts, and valves, not by a desire for symmetry of physics teaching.

20. The proposed device would extract heat from a single thermal reservoir (the surrounding warm atmosphere) and convert it entirely into useful mechanical work (turning the fan), with no waste heat rejected to a cooler reservoir. This is precisely what the Kelvin-Planck statement of the second law forbids: no device operating in a cycle can absorb heat from a single reservoir and convert it entirely into work. Some heat must always be rejected to a cooler reservoir, and the work output is bounded above by the Carnot efficiency between the two reservoir temperatures.

The technician's design has only one thermal reservoir available (the warm air, all at the same temperature) and so cannot operate, regardless of how the internal mechanism is constructed. The restriction is not a defect in any particular design but a fundamental limit on how heat can be converted to work; no engineering cleverness can circumvent it. The device is impossible in principle, in the same way that perpetual motion machines of the first kind (violating energy conservation) are impossible: both belong to a category of ambitious but doomed proposals.

21. Carnot's theorem fixes an upper bound on the efficiency of any heat engine operating between two reservoirs at temperatures T_{cold} and T_{hot} :

$$\eta_{\text{Carnot}} = 1 - \frac{T_{\text{cold}}}{T_{\text{hot}}} = 1 - \frac{300\text{K}}{500\text{K}} = 0.40 \text{ or } 40\%$$

The official's proposal claims 80%, double the Carnot ceiling between these two reservoirs. By Carnot's theorem, no heat engine, however cleverly designed, can exceed this ceiling for the given reservoir temperatures, so the proposal as stated is infeasible.

To make 80% theoretically achievable, the temperature ratio $T_{\text{cold}}/T_{\text{hot}}$ would have to fall to 0.20 or below. With $T_{\text{cold}} \approx 300\text{K}$ (set by the temperature of available cooling water in a tropical climate), this requires $T_{\text{hot}} \geq 1500\text{K}$, far above the operating temperature of typical coal boilers and beyond what conventional boiler materials can withstand.

22. As the cylinder enters the warmer kitchen, the gas inside warms from 25°C (298K) to 40°C (313K). The cylinder is sealed and rigid, so the gas undergoes an isochoric heating: at constant volume, $P/T = \text{const}$, and the new pressure is:

$$P_{\text{kitchen}} = P_{\text{outdoor}} \times \frac{313\text{K}}{298\text{K}} \approx 1.050 \times P_{\text{outdoor}}$$

So the internal pressure has risen by about 5%. The valve, calibrated assuming the cylinder is at 25°C , opens when the gas pressure exceeds the rated limit at that temperature. At 40°C , the same gas exerts a 5% higher pressure for the same amount, so the valve will open earlier than originally intended.

The calibration is therefore **conservative** in the new environment: it triggers a premature release, before the actual unsafe pressure level (which corresponds to a still-higher temperature, where the cylinder structurally fails) is reached. To use the cylinder safely indoors, the user can either accept earlier-than-needed valve activation, recalibrate the valve for kitchen temperature, or allow the cylinder to acclimatise to room temperature before any pressure-critical use. A risky calibration would be one where the valve fails to trigger early enough; conservative calibration errs in the opposite direction and is therefore safe.

23. The accountant's confusion arises from looking only at the heat balance and missing the work input that drives the compressor. The refrigeration cycle exchanges energy with three places: it absorbs heat Q_{cold} from the server room (the cold reservoir), it accepts work W from the electrical compressor, and it rejects heat Q_{hot} to the surroundings (the warm reservoir, typically the kitchen air or outside).

Applying the first law to the working substance over a complete cycle, $\Delta U_{\text{cycle}} = 0$, so the energy balance reduces to:

$$Q_{\text{cold}} + W = Q_{\text{hot}}$$

The heat dumped therefore exceeds the heat absorbed by exactly the amount of electrical work supplied to the compressor. Total energy in (heat absorbed plus electrical work) equals total energy out (heat rejected). Once the electrical input is included in the bookkeeping, energy conservation is intact; the engineer's claim is correct, and is in fact the operating principle of every refrigeration system. The accountant should be reassured that no perpetual motion is at play; the electricity bill is paying for the difference.

24. (a) The change in internal energy is **zero**. For an ideal gas the internal energy depends only on temperature, $U = U(T)$; with the temperature held constant at 300K, $\Delta U = 0$ regardless of how the volume or pressure change during the compression.

(b) From the first law, $\Delta U = Q - W = 0$, hence $Q = W$ where W is the work done by the gas. During compression W is negative because $dV < 0$, so Q is also negative; heat must leave the gas. Physically, the surroundings do positive work on the gas as they compress it, which would otherwise raise the internal energy and hence the temperature. To keep the temperature at 300K,

exactly the same amount of energy must leave the gas as heat to the surroundings. The work input is conducted away as heat, leaving the internal energy unchanged.

(c) The work done by the gas during an isothermal compression is:

$$W = nRT \ln \frac{V_f}{V_i} = 2.0 \text{ mol} \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1} \times 300 \text{ K} \times \ln \frac{1.0 \text{ L}}{4.0 \text{ L}}$$

$$W = 4988 \text{ J} \times (-1.386) = -6915 \text{ J}$$

So the magnitude of the heat removed from the gas is $|Q| = |W| \approx 6.9 \text{ kJ}$, in agreement with the technician's observation. The work added to the gas (+6915 J by the surroundings) is exactly compensated by the heat removed (-6915 J), leaving $\Delta U = 0$.

25. Water boils when its saturation vapour pressure equals the pressure of the gas above its surface. At standard atmospheric pressure (1 atm), the saturation pressure of water reaches 1 atm at 100°C, so 100°C is the open-pan boiling point. In a sealed pressure cooker, steam released by initial boiling is trapped above the water, and as more steam accumulates the gas pressure inside the cooker rises above atmospheric, typically to about 2 atm. The boiling point of water at 2 atm is approximately 120°C, so the water in the cooker now boils at this higher temperature. Food in contact with 120°C steam cooks much more rapidly than food in contact with 100°C steam, because the rate of biochemical softening reactions roughly doubles for every 10 K rise in temperature.

The safety valve role: the valve is set to open at a calibrated pressure threshold, typically about 2 atm. When the internal pressure reaches this threshold, the valve opens and steam escapes, removing energy from the cooker faster than the heater can supply it; the pressure (and therefore the boiling point) is held near the threshold value rather than rising indefinitely. Without the valve, continued heating would push the pressure and temperature above the safe limit, eventually rupturing the cooker.

EXERCISE 5C

26. (a) Path 1, after isothermal: $(P_0/2, 2V_0, T_0)$. After isochoric to $P = P_0$: $T = T_0 \times P_0/(P_0/2) = 2T_0$, so end state $(P_0, 2V_0, 2T_0)$.

Path 2, isobaric to $V = 2V_0$: $T = 2T_0$, end state $(P_0, 2V_0, 2T_0)$.

Both paths terminate at the same state B.

(b) For a monatomic gas, $\Delta U = \frac{3}{2} nR\Delta T = \frac{3}{2} nR(2T_0 - T_0) = \frac{3}{2} P_0 V_0$, using $nRT_0 = P_0 V_0$. Same value along both paths since U is a state function.

(c) Path 1: $W_1 = nRT_0 \ln 2 + 0 = P_0 V_0 \ln 2$.

Path 2: $W_2 = P_0(2V_0 - V_0) = P_0 V_0$.

$Q_1 = \Delta U + W_1 = \frac{3}{2} P_0 V_0 + P_0 V_0 \ln 2$.

$Q_2 = \Delta U + W_2 = \frac{3}{2} P_0 V_0 + P_0 V_0 = \frac{5}{2} P_0 V_0$.

$$Q_2 - Q_1 = P_0 V_0(1 - \ln 2) \approx 0.307 P_0 V_0$$

27. $f = 7$. Molar mass 16 g mol^{-1} identifies the gas as methane, CH_4 . Methane is a non-linear polyatomic molecule with 3 translational and 3 rotational degrees of freedom (6), with the additional fractional degree reflecting partial activation of vibrational modes at room temperature, consistent with $f \approx 7$ and $\gamma = 1.30$.

28.

$$P_f = P_0 \left(\frac{V_0}{V_f} \right)^\gamma = P_0 \times 4^{1.4} = 6.96 P_0$$

$$T_f = T_0 \left(\frac{V_0}{V_f} \right)^{\gamma-1} = T_0 \times 4^{0.4} = 1.74 T_0$$

$$W_{\text{on}} = nC_v(T_f - T_0) = \frac{5}{2} nR(0.74 T_0) = \frac{5}{2} (0.74) P_0 V_0 = 1.85 P_0 V_0$$

29.

$$\frac{dQ}{dt} = \frac{kA\Delta T}{L} = \frac{230 \text{ W m}^{-1} \text{ K}^{-1} \times 0.005 \text{ m}^2 \times 100 \text{ K}}{0.005 \text{ m}} = 23000 \text{ W}$$

$$U = mc\Delta T = 5.0 \times 10^{-4} \text{ kg} \times 1000 \text{ J kg}^{-1} \text{ K}^{-1} \times 100 \text{ K} = 50 \text{ J}$$

$$t_{\text{cond}} = \frac{U}{dQ/dt} = \frac{50 \text{ J}}{23000 \text{ W}} = 2.2 \text{ ms}$$

The compression stroke (1 ms) is shorter than the conduction timescale (2.2 ms) by a factor of about two. The adiabatic approximation is appropriate to first order, but with a non-negligible correction; most of the heat does not have time to leak during the stroke, but a small fraction does.

30. $W_{\text{net}} = 900 \text{ J}$, $Q_{\text{in}} = 3450 \text{ J}$, $\eta = 26.1\%$, the actual efficiency 26% is well below the Carnot ceiling 93.8%, as expected for a non-Carnot cycle.

31.

$$T_f = T_i \left(\frac{V_i}{V_f} \right)^{\gamma-1} = 300 \text{ K} \times 16^{0.4} = 909 \text{ K} = 636^\circ \text{C}$$

636°C is well above the autoignition temperature 480°C, so diesel fuel injected into the compressed air ignites spontaneously without an external spark. Petrol engines use lower compression ratios (typically 10-12), so the air temperature at the end of compression remains below the autoignition point of petrol-air mixture, and a spark plug is required to initiate combustion.

32. 28.8cmHg, -45.6°C

33.

$$\begin{aligned}\eta_1 &= 1 - \frac{T_c}{T_h} = 1 - \frac{300}{500} = 0.40 \\ Q_h &= \frac{W_1}{\eta_1} = \frac{200\text{J}}{0.40} = 500\text{J} \\ \eta_2 &= 1 - \frac{250}{500} = 0.50 \\ W_2 &= \eta_2 \times Q_h = 0.50 \times 500\text{J} = 250\text{J} \\ \frac{\eta_2 - \eta_1}{\eta_1} &= \frac{0.10}{0.40} = 25\%\end{aligned}$$

New work output is 250J per cycle; the percentage increase in efficiency is 25%.

34.

$$n = \frac{P_1 V_1}{RT_1} = \frac{1.0 \times 10^5 \text{Pa} \times 0.020 \text{m}^3}{8.314 \text{Jmol}^{-1} \text{K}^{-1} \times 300\text{K}} = 0.802 \text{mol}$$

Stage 1: adiabatic compression to $V_2 = V_1/4 = 5.0 \times 10^{-3} \text{m}^3$.

$$\begin{aligned}T_2 &= T_1 \times 4^{\gamma-1} = 300\text{K} \times 4^{0.4} = 522\text{K} \\ P_2 &= P_1 \times 4^\gamma = 1.0 \times 10^5 \text{Pa} \times 4^{1.4} = 6.96 \times 10^5 \text{Pa}\end{aligned}$$

$$W_{1,\text{on}} = nC_v(T_2 - T_1) = 0.802 \text{mol} \times \frac{5}{2} \times 8.314 \text{Jmol}^{-1} \text{K}^{-1} \times 222\text{K} \approx 3700\text{J}$$

Stage 2: isobaric cooling at P_2 from $T_2 = 522\text{K}$ to $T_3 = 300\text{K}$.

$$\begin{aligned}V_3 &= V_2 \times \frac{T_3}{T_2} = 5.0 \times 10^{-3} \text{m}^3 \times \frac{300\text{K}}{522\text{K}} = 2.87 \times 10^{-3} \text{m}^3 \\ W_{2,\text{on}} &= -P_2(V_3 - V_2) = -6.96 \times 10^5 \text{Pa} \times (-2.13 \times 10^{-3} \text{m}^3) \approx 1480\text{J} \\ W_{\text{on,total}} &= W_{1,\text{on}} + W_{2,\text{on}} = 3700\text{J} + 1480\text{J} = 5180\text{J}\end{aligned}$$

$$Q_2 = nC_p(T_3 - T_2) = 0.802 \text{mol} \times \frac{7}{2} \times 8.314 \text{Jmol}^{-1} \text{K}^{-1} \times (-222\text{K}) \approx -5180\text{J}$$

$$\Delta U_{\text{total}} = nC_v(T_3 - T_1) = 0$$

Total work done on the gas is approximately 5180J, the heat rejected during the cooling stage is approximately 5180J, and the net change in internal energy is zero.

35.

Before inversion (air below mercury): $P_{\text{air}} = P_{\text{atm}} + h_{\text{Hg}} = 76\text{cmHg} + 20\text{cmHg} = 96\text{cmHg}$, $L_1 = 40\text{cm}$.

After inversion (air above mercury): $P_{\text{air}} = P_{\text{atm}} - h_{\text{Hg}} = 76\text{cmHg} - 20\text{cmHg} = 56\text{cmHg}$.

Boyle's law (isothermal): $P_1 L_1 = P_2 L_2$.

$$L_2 = \frac{P_1 L_1}{P_2} = \frac{96\text{cmHg} \times 40\text{cm}}{56\text{cmHg}} = 68.6\text{cm}$$

The trapped air column has lengthened from 40cm to approximately 68.6cm.

SOLUTIONS TO EXAMINATION QUESTIONS

PART ONE

Question 1

(a) The error: the atoms in a stretched wire are not moving any faster than at zero strain; the temperature is unchanged. The strain energy is stored as the increase in the interatomic potential energy of the atoms, which sit slightly farther apart than at the equilibrium spacing of the unstretched wire. The work done by the applied force is converted into this potential energy.

(b) Under heavy load each leaf of the spring deflects further, so adjacent leaves slide against each other through a larger distance and the surface friction between them produces the audible squeak. With no load the leaves are in their natural curved shape, the contact between them is essentially static, and no sliding (and therefore no squeak) occurs.

(c)

$$\begin{aligned}\Delta L &= \frac{FL}{AE} = \frac{300\text{N} \times 2.0\text{m}}{1.5 \times 10^{-6} \text{m}^2 \times 2.0 \times 10^{11} \text{Pa}} = 2.0 \times 10^{-3} \text{m} \\ U &= \frac{1}{2} F \Delta L = \frac{1}{2} \times 300\text{N} \times 2.0 \times 10^{-3} \text{m} = 0.30\text{J}\end{aligned}$$

Extension is 2.0mm and strain energy stored is 0.30J.

Question 2

(a) The error: the proportional limit and the elastic limit need not coincide. The proportional limit is the highest stress at which Hooke's law (stress proportional to strain) holds. The elastic limit is the highest stress at which the wire still returns to its original length on unloading. Between these two limits the wire is still elastic but no longer linear in its response. For most metals the two limits are very close but not identical, and the elastic limit lies just above the proportional limit.

(b) Each downstroke loads the rod beyond its elastic limit by a small amount. The rod recovers most of its shape elastically after each stroke but retains a tiny plastic deformation. Over many cycles these tiny plastic deformations accumulate into the visible permanent bend, even though no single stroke applied a force large enough to fracture the rod.

(c)

$$A = \pi \left(\frac{d}{2} \right)^2 = \pi \times (0.5 \times 10^{-3} \text{m})^2 = 7.85 \times 10^{-7} \text{m}^2$$

$$F_{\text{yield}} = \sigma_y A = 2.5 \times 10^8 \text{Pa} \times 7.85 \times 10^{-7} \text{m}^2 \approx 196 \text{N}$$

$$F_{\text{break}} = \sigma_b A = 4.0 \times 10^8 \text{Pa} \times 7.85 \times 10^{-7} \text{m}^2 \approx 314 \text{N}$$

Plastic deformation begins at approximately 196N; fracture occurs at approximately 314N.

Question 3

(a) During an elastic stretch, the applied force builds up gradually from zero to its final value F . The work done is the area under the force-extension graph, which is the area of a triangle: $\frac{1}{2}FL$. Per unit volume, this becomes $\frac{1}{2}\sigma\epsilon$ where σ is stress and ϵ is strain. Multiplying final stress by final strain would assume the force was at its full final value throughout the stretch, which overcounts the work by a factor of two.

(b) The strain energy stored in a Hookean rubber band is proportional to the square of the extension, so doubling the extension stores four times the energy. On release this energy is converted to projectile kinetic energy $\frac{1}{2}mv^2$, so the launch velocity doubles. With doubled launch velocity at the same angle, the projectile range (proportional to v^2) becomes four times its previous value.

(c)

$$U = \frac{1}{2}kx^2 = \frac{1}{2} \times 200 \text{Nm}^{-1} \times (0.30 \text{m})^2 = 9.0 \text{J}$$

$$v = \sqrt{\frac{2U}{m}} = \sqrt{\frac{2 \times 9.0 \text{J}}{0.20 \text{kg}}} = 9.49 \text{ms}^{-1}$$

Energy stored 9.0J; launch velocity approximately 9.5m/s.

Question 4

(a) The isothermal bulk modulus B_T measures the resistance of a fluid to compression when the temperature is held constant; for an ideal gas $B_T = P$. The adiabatic bulk modulus B_S measures the same resistance when no heat flows during the compression; for an ideal gas $B_S = \gamma P$. The two differ because in an adiabatic compression the gas heats up, raising the pressure further than the volume reduction alone produces; the gas appears stiffer when compressed adiabatically.

(b) The bulk modulus, since the speed of sound in a fluid is $v = \sqrt{B/\rho}$. The bulk modulus of water is about $2.2 \times 10^9 \text{Pa}$, while that of air at atmospheric pressure is only about $1.4 \times 10^5 \text{Pa}$, a ratio of about 10^4 . Water molecules are tightly packed and resist compression strongly because the intermolecular separation is already close to the equilibrium value of the strong attractive forces. Air molecules are far apart and offer little resistance to compression beyond the kinetic-pressure term.

(c) (i) Isothermal bulk modulus:

$$B_T = P = 1.5 \times 10^5 \text{Pa}$$

(ii) Adiabatic bulk modulus:

$$B_S = \gamma P = \frac{5}{3} \times 1.5 \times 10^5 \text{Pa} = 2.5 \times 10^5 \text{Pa}$$

Question 5

(a) At every cross-section the same force F is transmitted, but stress is force per unit area; a smaller cross-section therefore carries a larger stress, and the strain $\epsilon = \sigma/E$ is correspondingly larger. The strain is greatest where the cross-section is narrowest, which is also where the bar is most likely to fail under load.

(b) Glass is a brittle material: under stress it stores elastic energy but cannot dissipate it through plastic deformation. When the elastic limit is reached, the stored energy is released suddenly through the propagation of cracks, shattering the tumbler. A metal cup is ductile: stress beyond the elastic limit causes plastic deformation, in which atoms slip past each other and absorb energy as heat through dislocation motion. The cup dents instead of shattering because plastic flow can absorb the impact energy without catastrophic crack propagation.

(c) Stress in main rod:

$$\sigma_0 = \frac{F}{A_0}$$

Nominal stress at the reduced section:

$$\sigma_{\text{nom}} = \frac{F}{A_0/4} = 4\sigma_0$$

Peak stress at the edge of the hole, with stress concentration factor 3:

$$\sigma_{\text{peak}} = 3\sigma_{\text{nom}} = 12\sigma_0$$

The peak local stress at the hole is 12 times the stress in the unaffected part of the rod.

Question 6

(a) The error: in a parallel-wire cable each wire carries only its share of the total load, not the entire load. Since the wires are identical and forced to extend by the same amount (they are joined at the ends), Hooke's law requires each to carry the same tension. The correct sharing is therefore F/n in each wire.

(b) First, redundancy: if one rope fails, the remaining ropes can still hold the lift (each rope is rated to carry the full load with a safety margin), avoiding a catastrophic fall. Second, fatigue and inspection: thinner ropes flex more easily over the drive sheave, reducing bending stresses and fatigue, and they can be inspected and replaced one at a time without taking the lift out of service for long. A single thick rope would be stiffer, would fatigue faster at the sheave, and would offer no redundancy.

(c) Since the beam is rigid and remains horizontal, all three wires extend by the same amount x . Each tension is $T = (AE/L)x$ for a wire of area A . The middle wire has area $2A$, so its tension is twice that of an outer wire.

$$T_1 = T_3 = \frac{AEx}{L}, \quad T_2 = \frac{2AEx}{L}$$

Vertical equilibrium: $T_1 + T_2 + T_3 = W$:

$$\frac{AEx}{L}(1 + 2 + 1) = W \Rightarrow \frac{AEx}{L} = \frac{W}{4}$$

$$T_1 = T_3 = \frac{W}{4}, \quad T_2 = \frac{W}{2}$$

Tensions in the ratio 1:2:1, with the middle (thicker) wire carrying half the load and each outer wire carrying a quarter.

Question 7

(a) Simple harmonic motion. The wire behaves as a Hookean spring of effective force constant $k = AE/L$, where A , L are the wire's cross-sectional area and natural length and E is its Young's modulus. The restoring force on the mass at displacement x from equilibrium is $-kx$, giving SHM with angular frequency $\omega = \sqrt{k/m}$ and period $T = 2\pi\sqrt{m/k}$, which depends on the wire's A , L , and E as well as on the mass.

(b) The suspension acts as a spring-mass system with its own natural frequency, set by the spring constant of the suspension and the sprung mass of the bus. At one particular cruising speed, the rotation frequency of the wheel (or some sub-multiple of it) coincides with this natural frequency, and small unbalanced forces from the wheel drive the suspension into resonance, producing the audible hum. Above and below that speed, the driving frequency is off-resonance and the response is too small to hear.

(c)

$$k = \frac{AE}{L} = \frac{5.0 \times 10^{-7} \text{ m}^2 \times 2.0 \times 10^{11} \text{ Pa}}{1.5 \text{ m}} = 6.67 \times 10^4 \text{ Nm}^{-1}$$

$$T = 2\pi\sqrt{\frac{m}{k}} = 2\pi\sqrt{\frac{2.0\text{kg}}{6.67 \times 10^4 \text{Nm}^{-1}}} = 3.44 \times 10^{-2}\text{s}$$

Period of oscillation is approximately 34ms.

Question 8

(a) The proportional limit has been passed: beyond this point the stress is no longer linearly proportional to the strain. Young's modulus is defined for the linear (Hookean) regime, where $\sigma = E\varepsilon$, so the slope of the straight portion of the stress-strain curve gives E directly. The slope of the non-linear portion does not represent any single material constant and cannot be used.

(b) A tensile test loads the sample to failure and records its yield stress, breaking stress and Young's modulus. A below-specification sample fails one or more of these tests directly and is detected. An above-specification sample passes the tension tests, but a bridge cable also fatigues under cyclic loading, suffers stress concentration at fittings, and is exposed to salt corrosion; tensile testing measures only the static-load properties and cannot detect defects that show up only under cyclic, environmental, or geometric stress. Above-specification by tension alone is therefore necessary but not sufficient.

(c)

$$A = \pi\left(\frac{d}{2}\right)^2 = \pi \times (5.0 \times 10^{-3}\text{m})^2 = 7.85 \times 10^{-5}\text{m}^2$$

$$\sigma = \frac{F}{A} = \frac{14 \times 10^3\text{N}}{7.85 \times 10^{-5}\text{m}^2} = 1.78 \times 10^8\text{Pa}$$

$$\varepsilon = \frac{\Delta L}{L} = \frac{0.180 \times 10^{-3}\text{m}}{0.200\text{m}} = 9.00 \times 10^{-4}$$

$$E = \frac{\sigma}{\varepsilon} = \frac{1.78 \times 10^8\text{Pa}}{9.00 \times 10^{-4}} = 2.0 \times 10^{11}\text{Pa}$$

Young's modulus of the steel is approximately $2.0 \times 10^{11}\text{Pa}$.

Question 9

(a) The error: a load placed gradually does work $\frac{1}{2}F\Delta L$ on the wire as the force builds up; a load suddenly released does work $F(\Delta L + h)$ as it falls through the wire's stretching distance plus any height h of free fall. The sudden load therefore stores more elastic energy in the wire and causes a larger maximum extension than the equilibrium static extension. After a sudden release the wire oscillates about the static extension, with the first downswing reaching the maximum dynamic extension, which can be much larger than the static value.

(b) In the controlled lowering, the wire stretches by its static extension only and the maximum tension is just the cement weight. In the dropping method, the falling sack converts gravitational potential energy into both elastic strain energy and kinetic energy of the sack just before contact; the wire must then absorb the total energy by stretching beyond its static extension, raising the peak tension well above the static value. The peak tension can exceed the breaking stress of the wire even though the static load alone is well within the safe range.

(c) Effective spring constant of the wire and weight of the load:

$$k = \frac{AE}{L} = \frac{1.0 \times 10^{-6}\text{m}^2 \times 2.0 \times 10^{11}\text{Pa}}{2.0\text{m}} = 1.0 \times 10^5\text{Nm}^{-1}$$

$$mg = 5.0\text{kg} \times 9.81\text{ms}^{-2} = 49.05\text{N}$$

Static extension:

$$x_s = \frac{mg}{k} = \frac{49.05\text{N}}{1.0 \times 10^5\text{Nm}^{-1}} = 4.91 \times 10^{-4}\text{m} \approx 0.49\text{mm}$$

Energy conservation from release to maximum extension (gravitational PE lost = strain energy stored):

$$mg(h + x_{\max}) = \frac{1}{2}kx_{\max}^2$$

$$\frac{1}{2} \times 1.0 \times 10^5 x_{\max}^2 - 49.05 x_{\max} - 4.905 = 0$$

Solving the quadratic (taking the positive root):

$$x_{\max} = 1.04 \times 10^{-2} \text{m} = 10.4 \text{mm}$$

$$\frac{x_{\max}}{x_s} = \frac{10.4 \text{mm}}{0.49 \text{mm}} \approx 21$$

Maximum dynamic extension is approximately 10.4mm, about 21 times the static extension.

Question 10

(a) The fundamental frequency of a stretched wire fixed at both ends is $f = \frac{1}{2L} \sqrt{T/\mu}$, which depends on the tension T , the length L , and the linear mass density μ , but does not contain Young's modulus explicitly. The Young's modulus enters the problem only through the relation between extension and tension; once a particular tension is established in the wire, the wire's vibration frequency is determined by T , L , and μ alone, regardless of which value of E produced that tension. A stiffer wire requires a larger applied force to reach a given tension, but the resulting frequency at that tension is the same as for a softer wire at the same tension.

(b) Tightening the tuning peg increases the tension T in the string. Since $f \propto \sqrt{T}$ at constant L and μ , raising T raises the fundamental frequency, which is heard as a rise in pitch.

(c)

$$f = \frac{1}{2L} \sqrt{\frac{T}{\mu}}$$

$$T = \mu(2Lf)^2 = 8.0 \times 10^{-3} \text{kgm}^{-1} \times (2 \times 0.75 \text{m} \times 220 \text{Hz})^2 \approx 871 \text{N}$$

Tension in the wire is approximately 871N.

Question 11

(a) The property must (i) vary continuously and reproducibly with temperature, so that a single value of the property unambiguously identifies a single temperature; (ii) vary monotonically (single-valued in the working range), so that the same temperature does not correspond to two different values of the property; (iii) vary by an amount large enough to be measured precisely with available instruments, so that small temperature differences produce a distinguishable change in the property.

(b) The reference thermometer is used only at the ice point, the steam point, and the calibration laboratory, so it accumulates very little wear, drift, or chemical contamination. The working thermometer accumulates these effects rapidly through daily handling. By keeping the reference thermometer in a controlled environment, the laboratory preserves a stable standard against which the working thermometer's drift can be detected and corrected.

(c)

$$\theta_{\text{Pt}} = \frac{R - R_0}{R_{100} - R_0} \times 100^\circ\text{C}$$

$$\theta_{\text{Pt}} = \frac{31.5\Omega - 25.0\Omega}{35.5\Omega - 25.0\Omega} \times 100^\circ\text{C} = 61.9^\circ\text{C}$$

Temperature on the platinum-resistance scale is approximately 62°C .

Question 12

(a) The pressure of an ideal gas at constant volume varies linearly with the absolute temperature, $P/T = \text{const}$, and this relation is independent of the gas used in the limit of low pressure. Different gases (helium, hydrogen, nitrogen) all give the same temperature reading when their thermometers are each calibrated at the triple point of water and used at low enough pressure that they behave ideally. The gas thermometer therefore defines a temperature scale that is essentially independent of the working substance, as required of a thermodynamic temperature scale.

(b) The gas thermometer's calibration is fixed by the laws of physics, not by the specific properties of any one gas, so it does not drift over time and does not need to be re-traced to a higher-tier reference. Platinum resistance thermometers are convenient and quick to read, but their resistance-temperature relation depends on the purity and prior heat-treatment of the platinum, both of which can change with use; they need periodic recalibration against an absolute reference. The gas thermometer fills that role.

(c)

$$T = T_{\text{tp}} \times \frac{P}{P_{\text{tp}}} = 273.16 \text{K} \times \frac{1.099 \times 10^5 \text{Pa}}{0.800 \times 10^5 \text{Pa}} = 375.3 \text{K} = 102.15^\circ\text{C}$$

Temperature of the bath is approximately 375K, or about 102°C .

Question 13

(a) The error: agreement at two fixed points does not imply agreement at intermediate points. The platinum-resistance scale is constructed by linear interpolation between the ice and steam points, but the actual relationship between the resistance of platinum and its temperature is not exactly linear; the ideal-gas scale, by contrast, is linear in absolute temperature. Wherever the platinum's true response curve deviates from the straight line that joins the two calibration points, the platinum-resistance scale and the ideal-gas scale disagree, with the largest discrepancy typically at the midpoint of the calibration range.

(b) First, the temperature 400°C is well above the boiling point of mercury (357°C); a mercury-in-glass thermometer cannot operate above this limit without the mercury vaporising and breaking the column. Second, the platinum thermometer can be sealed inside a metal sheath connected to its electronics by long leads, allowing remote reading from outside the steam chamber, while a mercury-in-glass thermometer would have to be visible at the chamber, requiring a window and an operator at the hot location.

(c)

$$\theta_{\text{Pt}} = \frac{R - R_0}{R_{100} - R_0} \times 100^\circ\text{C} = \frac{41.6 - 25.0}{35.6 - 25.0} \times 100^\circ\text{C} = 156.6^\circ\text{C}$$

On the ideal-gas scale, $T = 158.0^\circ\text{C}$. The platinum-resistance scale reads about 1.4°C below the ideal-gas value at this temperature. The discrepancy is consistent with the slight non-linearity of platinum's resistance-temperature curve over the 0 to 158°C range.

Question 14

(a) At each junction between the two metals, an electric potential difference arises (the Seebeck effect), with magnitude that depends on the temperature of the junction and on the pair of metals. When the two junctions are at different temperatures the two potential differences do not cancel, and a net EMF appears around the loop, driving a current if the loop is closed. The cold junction is necessary because only the temperature difference between the two junctions can be inferred from the EMF; without a reference junction at a known temperature, the hot-junction temperature cannot be deduced from the EMF reading alone.

(b) First, the thermocouple junction is small and has very low thermal mass, so it reaches thermal equilibrium with the meat quickly and does not cool the meat noticeably. Second, the junction can be inserted on a long thin probe into the centre of a piece of meat, while the EMF is read on a meter at a convenient distance; a glass thermometer would have to be inserted along its full length and read at the meat itself.

(c) Linear calibration: $E = k\theta$, with $k = 4.10\text{mV}/100^\circ\text{C} = 0.0410\text{mV}/^\circ\text{C}$:

$$\theta = \frac{E}{k} = \frac{7.40\text{mV}}{0.0410\text{mV}/^\circ\text{C}} = 180.5^\circ\text{C}$$

Temperature of the oil bath is approximately 180.5°C.

Question 15

(a) The ice point and the steam point are both pressure-dependent: the boiling point of water depends on the surrounding atmospheric pressure, and the melting point of ice changes slightly with pressure. Realising either point to the precision required for primary metrology demands controlling the pressure to high accuracy. The triple point of water, by contrast, is the unique combination of pressure and temperature at which all three phases of water coexist; it is fixed by the substance itself and is independent of external conditions, so a single fixed point is sufficient and is more reproducible than either of the historical pair.

(b) In a triple-point cell, water is sealed inside a glass cell with three phases (ice, liquid water, and water vapour) coexisting at the triple-point temperature 273.16K and pressure 611Pa. So long as all three phases are present, the temperature is fixed by the physics of the substance, independent of the surrounding atmospheric pressure (which cannot reach the sealed water inside the cell). An open ice-water bath, by contrast, is at the ice point at the surrounding atmospheric pressure, which varies from day to day and from one altitude to another; its temperature is therefore less reproducible by a few millikelvins.

(c)

$$T = T_{\text{tp}} \times \frac{P}{P_{\text{tp}}} = 273.16\text{K} \times \frac{28.0\text{kPa}}{100\text{kPa}}$$

$$T = 273.16\text{K} \times 0.280 = 76.5\text{K}$$

$$\theta = T - 273.15 \approx -196.7^\circ\text{C}$$

Boiling point of liquid nitrogen on the ideal-gas scale is approximately 76.5K or -196.7°C.

Question 16

(a) In a metal, the conduction electrons are essentially free at all temperatures, and the resistance arises mainly from electron scattering by lattice vibrations; rising temperature increases the lattice vibrations and so the resistance grows with temperature. In a semiconductor, the number of conduction electrons is itself temperature-dependent: thermal excitation liberates more charge carriers from the valence band as the temperature rises, and this rapid increase in carrier density dominates the resistance, which therefore falls steeply with temperature. The fall is exponential, so the fractional change in resistance per kelvin is large near room temperature, allowing thermistors to detect small temperature changes that a metal resistance thermometer would resolve only with difficulty.

(b) First, the thermistor's resistance changes more steeply with temperature than that of any practical metal element, so a small temperature change at the patient produces a large change in resistance and hence in the digital display, giving high sensitivity over the narrow clinical range 35 to 42°C. Second, thermistors can be made very small (sub-millimetre), so the probe reaches thermal equilibrium with the patient's tissue rapidly, giving a reading in a few seconds rather than the minute or more needed for a larger-thermal-mass platinum element.

(c) From $R = R_0 e^{(B/T)}$, taking the ratio at the two given temperatures:

$$\frac{R_1}{R_2} = e^{\left(\frac{B}{T_1} - \frac{B}{T_2}\right)} \Rightarrow \ln \frac{R_1}{R_2} = B \left(\frac{1}{T_1} - \frac{1}{T_2} \right)$$

$$B = \frac{\ln R_1/R_2}{1/T_1 - 1/T_2} = \frac{\ln 10.0/3.30}{1/298 - 1/323} = 4270\text{K}$$

At $T = 310\text{K}$ (37°C):

$$R(37^\circ\text{C}) = R_1 e^{\left(B\left(\frac{1}{T} - \frac{1}{T_1}\right)\right)} = 10.0\text{k}\Omega \times e^{(-0.554)} = 5.74\text{k}\Omega$$

$B = 4270\text{K}$ and $R(37^\circ\text{C}) = 5.74\text{k}\Omega$.

Question 17

(a) A clinical thermometer has a fine constriction in the capillary just above the bulb. As the bulb is heated, mercury expands and is forced through the constriction into the stem, where it rises to the appropriate temperature mark. When the thermometer is removed and cools, the mercury in the bulb contracts and pulls back; but the mercury thread in the stem cannot pass back through the narrow constriction by capillary attraction alone, so it remains in place at the maximum mark. The reading therefore stays at the maximum until the thermometer is shaken to force the mercury back through the constriction.

(b) Shaking accelerates the trapped mercury thread sharply along the stem; the resulting inertial force on the mercury column overcomes the capillary attraction holding it above the constriction, and the column is pushed back through the constriction into the bulb. Without this manual reset, the thermometer would still display the previous patient's maximum and would underread the next patient by however much the next temperature was below the previous reading.

(c)

$$\Delta V = \gamma_{\text{Hg}} V_0 \Delta T = 1.82 \times 10^{-4}\text{K}^{-1} \times 0.20\text{cm}^3 \times (39.5 - 24.0)\text{K} = 5.64 \times 10^{-4}\text{cm}^3$$

The volume of mercury increases by approximately $5.6 \times 10^{-4}\text{cm}^3$ (about 0.56mm^3) over the 15.5K rise. The narrow stem bore converts this small volume change into a visible rise of the mercury thread.

Question 18

(a) Mercury freezes at -39°C and boils at 357°C , giving a working range of about -38 to $+350^\circ\text{C}$; ethanol freezes at -114°C and boils at 78°C , giving a working range of about -110 to $+75^\circ\text{C}$. Mercury therefore covers high temperatures up to several hundred degrees, while alcohol covers very low temperatures (below -38°C) where mercury would solidify. Mercury has a smaller cubic expansion coefficient ($1.82 \times 10^{-4}\text{K}^{-1}$) than alcohol ($1.1 \times 10^{-3}\text{K}^{-1}$), so for the same bore an alcohol thermometer is roughly six times more sensitive (larger thread movement per kelvin), but covers a narrower range. The choice between mercury and alcohol therefore trades high-temperature reach against low-temperature reach and sensitivity.

(b) -20°C is well above the freezing point of ethanol (-114°C) but already cold enough to be approaching mercury's freezing point (-39°C). On a colder than average night, a mercury thermometer can solidify and stop responding, failing to record the actual minimum. Alcohol remains a free-flowing liquid throughout the temperature range encountered in Mbeya, so an alcohol thermometer is the safe choice for low-temperature work.

(c) Volume change per kelvin of temperature rise:

$$\frac{\Delta V}{\Delta T} = \gamma_{\text{Hg}} V_0 = 1.82 \times 10^{-4}\text{K}^{-1} \times 0.50\text{cm}^3 = 9.10 \times 10^{-5}\text{cm}^3/\text{K}$$

This volume change drives a $1.0\text{mm} = 0.10\text{cm}$ rise in the stem, occupying a length Δl of bore of cross-sectional area A :

$$\Delta V = A\Delta l \Rightarrow A = \frac{\Delta V}{\Delta l} = \frac{9.10 \times 10^{-5} \text{cm}^3}{0.10 \text{cm}} = 9.10 \times 10^{-4} \text{cm}^2$$

Cross-sectional area of the stem bore is approximately $9.1 \times 10^{-4} \text{cm}^2$ (about 0.091mm^2 , corresponding to a bore diameter of about 0.34mm).

Question 19

(a) Every body at non-zero absolute temperature radiates electromagnetic energy, and for a hot body in the visible-and-infrared range the spectrum and total flux of this radiation depend on the body's temperature, by Planck's law and the Stefan-Boltzmann law. A pyrometer collects this thermal radiation through a lens and infers the temperature of the body from the radiation flux or its spectrum, without any physical contact. Above about 600 to 700°C , mercury-in-glass thermometers fail (mercury boils above 357°C) and platinum resistance thermometers, though usable to about 1000°C , suffer contamination from hot gases and are no longer accurate; for body temperatures of order 1000°C and above, the pyrometer is the only practical instrument.

(b) The molten steel furnace operates at temperatures of order 1500°C , where any contact probe would be destroyed and where the body emits enough visible-and-infrared radiation for a pyrometer to read its temperature easily; the non-contact method is therefore both necessary (for survival of the instrument) and sufficient (for accuracy). The poultry incubator operates at modest temperatures of about 38°C , where a contact probe is survivable, and where a pyrometer would have to be calibrated at low temperatures with much weaker radiation; a thermistor or platinum probe gives a more accurate reading at room-temperature scales.

(c) The pyrometer reading is proportional to T^4 :

$$\frac{T^4}{1000^4 \text{K}^4} = \frac{4.00}{1.00} \Rightarrow T = 1000 \text{K} \times 4^{1/4} \approx 1414 \text{K} \approx 1141^\circ\text{C}$$

Temperature of the glowing body is approximately 1414K , or about 1141°C .

Question 20

(a) Thermal conductivity k is a property of the material alone, with units $\text{Wm}^{-1}\text{K}^{-1}$. Thermal conductance is a property of a particular component (slab, wall, window), $C = kA/L$ in units WK^{-1} , depending on the material's k , the cross-sectional area A through which heat flows, and the path length L . A wall with a window has two parallel paths for heat: the conductances of the wall and the window add (parallel rule), so the area of each portion enters the total conductance directly. A material of poor insulation occupying a large fractional area can dominate the heat loss, even if its thickness alone is favourable.

(b) The dominant heat-loss path is the window, because the glass window has both a much smaller thickness (5mm vs 100mm of brick) and a small but comparable conductivity to the brick; per square metre, the window conducts much more heat than the brick. Doubling the window area roughly doubles the window contribution, while leaving the small brick contribution unchanged; the overall heat lost therefore increases by approximately the window's share of the total, which can be more than doubling if the original window already carries most of the heat. In practice the increase is close to a doubling because the brick contribution is small to begin with.

(c) (i) Heat conduction through the brick portion:

$$P_{\text{brick}} = \frac{k_b A_b \Delta T}{L_b} = \frac{1.0 \text{Wm}^{-1}\text{K}^{-1} \times 8.5 \text{m}^2 \times 10 \text{K}}{0.10 \text{m}} = 850 \text{W}$$

(ii) Heat conduction through the glass window:

$$P_{\text{glass}} = \frac{k_g A_g \Delta T}{L_g} = \frac{0.80 \text{Wm}^{-1}\text{K}^{-1} \times 1.5 \text{m}^2 \times 10 \text{K}}{5.0 \times 10^{-3} \text{m}} = 2400 \text{W}$$

(iii) Total rate of heat conduction:

$$P_{\text{total}} = P_{\text{brick}} + P_{\text{glass}} = 850 \text{W} + 2400 \text{W} = 3250 \text{W}$$

Total heat-loss rate is approximately 3250W , of which the small glass window carries 2400W (about 74%) despite occupying only 15% of the wall area.

Question 21

(a) Heat flowing radially outward through a cylindrical layer crosses successive shells of increasing area $A = 2\pi rL$, so the local temperature gradient dT/dr falls as r increases at the same heat-flow rate P . Integrating Fourier's law $P = -kA(dT/dr)$ between the inner radius r_i and outer radius r_o gives $P = 2\pi kL(T_i - T_o)/\ln(r_o/r_i)$. The relevant geometric factor is therefore the logarithm of the radius ratio, not the thickness $r_o - r_i$ alone; a thin layer wrapped on a large-radius pipe and the same thickness on a small-radius pipe contribute differently to the insulation.

(b) The mineral wool, although a better thermal insulator per unit thickness, is not robust against compression and outdoor weather; the outer foil-faced foam, with a higher conductivity but better mechanical and weather protection, shields the inner layer from damage. The two-layer arrangement therefore combines the high-thermal-resistance of the inner wool with the durability of the outer foam, an option not available with a single-material lagging.

(c) For two cylindrical layers in series carrying the same heat per unit length, the total thermal resistance per unit length is:

$$\frac{R}{L} = \frac{1}{2\pi} \left(\frac{\ln r_2/r_1}{k_1} + \frac{\ln r_3/r_2}{k_2} \right)$$

With $r_1 = 2.0\text{cm}$, $r_2 = 4.0\text{cm}$, $r_3 = 5.0\text{cm}$:

$$\frac{R}{L} = \frac{1}{2\pi} \left(\frac{\ln 2}{0.05} + \frac{\ln 1.25}{0.10} \right) = 2.56\text{KmW}^{-1}$$

$$\frac{P}{L} = \frac{\Delta T}{R/L} = \frac{60\text{K}}{2.56\text{KmW}^{-1}} \approx 23.4\text{Wm}^{-1}$$

Heat loss is approximately 23W per metre of pipe.

Question 22

(a) Newton's cooling law assumes the heat-transfer coefficient h between the body and the fluid is independent of the temperature difference, so that the cooling rate is strictly linear in the excess temperature. In forced convection (a fan or wind), the fluid speed past the body is fixed by external means, h is determined by that speed and is independent of the temperature difference; the law holds rigorously. In natural convection, the buoyancy-driven fluid speed itself depends on the temperature difference (warmer body drives faster updraft), so h increases with ΔT , and the cooling rate depends on ΔT to a power slightly greater than one. At small ΔT the variation in h is small and the law is approximately linear; at large ΔT it is no longer linear.

(b) Fanning forces room-temperature air to flow rapidly past the meat, replacing the warm air layer next to the surface with cool air far more rapidly than buoyancy alone could; the heat-transfer coefficient for forced convection is much larger than for natural convection, so the meat loses heat at a higher rate. Radiation and conduction to the table are essentially unchanged by the fanning, so convection is the dominant mechanism enhanced.

(c) For Newton's law $\theta - \theta_s = (\theta_0 - \theta_s)e^{-kt}$, the cooling constant satisfies $kt = \ln[(\theta_0 - \theta_s)/(\theta - \theta_s)]$.

In free air, $\theta_0 - \theta_s = 60\text{K}$, $\theta - \theta_s = 40\text{K}$, $t = 5.0\text{min}$:

$$k_{\text{free}} = \frac{1}{5.0\text{min}} \ln \frac{60}{40} = 0.0811\text{min}^{-1}$$

In the draught, same temperature change in $t = 1.5\text{min}$:

$$k_{\text{forced}} = \frac{\ln 1.5}{1.5\text{min}} = 0.270\text{min}^{-1}$$

$$\frac{k_{\text{forced}}}{k_{\text{free}}} = \frac{5.0}{1.5} = 3.3$$

Forced-convection cooling constant is approximately 3.3 times the free-convection cooling constant under these conditions.

Question 23

(a) As the temperature of a blackbody rises, the wavelength at which the emission spectrum peaks shifts towards shorter wavelengths according to Wien's displacement law $\lambda_{\text{max}} T = b$, where $b \approx 2.898 \times 10^{-3}\text{mK}$ is the Wien constant. At 1000K the peak lies in the infrared (around $3\mu\text{m}$), invisible to the eye; at 2000K it has moved to around $1.5\mu\text{m}$, still infrared, but the visible tail of the curve makes the body glow dull red. At 5000K the peak is at 580nm , in the yellow-orange region, and the body glows yellow-white as our Sun does. The colour shift from invisible-infrared through dull red, orange, yellow, and finally bluish-white as temperature rises is a direct visualisation of Wien's law.

(b) The surface temperature of a star sets the wavelength at which its thermal emission peaks, by $\lambda_{\text{max}} = b/T$. Cool red giants have surface temperatures of around 3000K; their peak wavelength is $b/3000 \approx 970\text{nm}$, in the near-infrared, and the visible-spectrum tail is dominated by long red wavelengths. The Sun, at about 5800K, peaks at $b/5800 \approx 500\text{nm}$, in the green-yellow, with a broad visible spectrum that appears whitish-yellow. Hot young stars at 10000K or above peak in the ultraviolet ($b/10000 = 290\text{nm}$), with the visible-spectrum tail weighted to short blue wavelengths, appearing bluish-white. The hierarchy red-yellow-blue corresponds directly to increasing surface temperature.

(c)

$$T = \frac{b}{\lambda_{\max}} = \frac{2.898 \times 10^{-3} \text{mK}}{480 \times 10^{-9} \text{m}} \approx 6038 \text{K}$$

For comparison, the Sun's surface temperature from $\lambda_{\max} = 500 \text{nm}$:

$$T_{\text{sun}} = \frac{2.898 \times 10^{-3} \text{mK}}{500 \times 10^{-9} \text{m}} \approx 5796 \text{K}$$

The star is at approximately 6038K, which is 242K hotter than the Sun's surface.

Question 24

(a) At thermal equilibrium between a body and a cavity at temperature T , the body must emit and absorb radiation at every wavelength at exactly the same rate; otherwise its temperature would not stay equal to T . Since the cavity radiation is a perfect blackbody field, the body's emissivity ϵ_{λ} at each wavelength equals its absorptivity α_{λ} at the same wavelength: this is Kirchhoff's law. A poor emitter at a particular wavelength is therefore necessarily a poor absorber at the same wavelength, and conversely a good emitter at a wavelength is also a good absorber at that wavelength.

(b) During the day, the dominant incoming radiation is solar (peak around 500nm, in the visible), and a matt black surface absorbs strongly at these short wavelengths, heating the roof and the interior more than a reflective white surface would. At night, the roof is itself the warm body radiating to the sky; the dominant outgoing radiation is in the thermal infrared (peak around $10 \mu\text{m}$ at 300K), and the matt black surface emits strongly in the IR by Kirchhoff's law, cooling itself rapidly. The matt black paint therefore **heats** the roof during the day (bad) and **cools** it at night (good); on net, in tropical daytime climates, white or shiny aluminium-coated paint reduces interior temperatures more effectively. The matt black choice is therefore poorly matched to the daytime cooling goal.

(c)

$$\frac{P}{A} = \epsilon \sigma (T^4 - T_s^4)$$

$$\frac{P}{A} = 0.60 \times 5.67 \times 10^{-8} \text{Wm}^{-2} \text{K}^{-4} \times (400^4 - 300^4) \text{K}^4 \approx 595 \text{Wm}^{-2}$$

Net radiation loss per unit area is approximately 595Wm^{-2} .

Question 25

(a) The flask consists of a double-walled glass vessel with an evacuated gap between the walls and silvered inner surfaces. The vacuum eliminates both conduction and convection across the gap, since both mechanisms require a material medium. The silvered surfaces have very low emissivity ($\epsilon \approx 0.05$), reducing radiative heat exchange across the gap by about 20-fold compared with unsilvered glass. Conduction along the narrow neck of the flask cannot be fully eliminated (the lid is in contact with both the contents and the outer wall), and radiation across the silvered gap is reduced but not zero, so all three mechanisms contribute residual heat losses.

(b) Over the eight hours the contents lose only about 20K of temperature, corresponding to an average heat-loss rate of less than a watt for a half-litre of tea (heat capacity around 2.1kJ/K). The slow cooling reflects the success of the vacuum at suppressing conduction and convection across the gap, plus the silvered surfaces at suppressing radiation. The dominant residual loss is conduction along the neck of the flask through the stopper and the threaded fitting, paths that cannot be evacuated; at very long times this neck-conduction loss dominates and limits the flask's effectiveness.

(c) Radiation loss across the silvered vacuum gap ($T = 353 \text{K}$, $T_s = 293 \text{K}$):

$$P_{\text{rad}} = \epsilon \sigma A (T^4 - T_s^4) = 0.05 \times 5.67 \times 10^{-8} \times 0.010 \times (353^4 - 293^4) = 0.23 \text{W}$$

Free-convection loss from an unlagged cup of the same area:

$$P_{\text{conv}} = hA\Delta T = 10 \text{Wm}^{-2} \text{K}^{-1} \times 0.010 \text{m}^2 \times 60 \text{K} = 6.0 \text{W}$$

The vacuum-flask radiation loss (0.23W) is approximately 26 times smaller than the open-cup convection loss (6.0W), reflecting the design's effectiveness at suppressing the dominant convective mechanism.

Question 26

(a) A blackbody absorber and a blackbody emitter are necessarily the same surface, by Kirchhoff's law of thermal radiation: $\epsilon_{\lambda} = \alpha_{\lambda}$ at every wavelength. A surface that is black at visible-and-near-infrared wavelengths (where the Sun radiates) is automatically also black at the longer thermal-infrared wavelengths (where it radiates at its 350K operating temperature) unless the surface's optical properties differ between the two wavelength ranges. A 'selective surface' achieves this by being coated with a multi-layer film tuned to absorb strongly at solar wavelengths but reflect (and therefore emit poorly) at thermal infrared wavelengths;

commercially this is achieved by sputtered metal-on-oxide layers (e.g. nickel-on-tin oxide) and brings emissivity at $10\mu\text{m}$ down below 0.1 while keeping absorptivity at $0.5\mu\text{m}$ above 0.9.

(b) As the panel temperature rises, the rate at which it loses heat by infrared radiation increases as the fourth power of its absolute temperature, while the absorbed solar flux is fixed by the Sun's output. Steady state is reached when the radiative loss plus any convective and conductive losses just equal the absorbed solar power; any further temperature rise would cause net heat loss. The limiting factor on the steady-state temperature is therefore the balance between solar input (constant at noon) and combined infrared and convective output (rising rapidly with T). For a fixed-flux input of 800Wm^{-2} and ordinary panel emissivity, the equilibrium temperature is typically around 70 to 90°C in tropical sunshine.

(c) Solar absorbed:

$$P_{\text{abs}} = \alpha SA = 0.90 \times 800\text{Wm}^{-2} \times 2.0\text{m}^2 = 1440\text{W}$$

Radiative loss to surroundings ($T = 343\text{K}$, $T_s = 303\text{K}$):

$$P_{\text{rad}} = \epsilon \sigma A(T^4 - T_s^4) = 0.90 \times 5.67 \times 10^{-8} \times 2.0 \times (343^4 - 303^4) \approx 552\text{W}$$

Net useful collection rate:

$$P_{\text{net}} = P_{\text{abs}} - P_{\text{rad}} = 1440\text{W} - 552\text{W} = 888\text{W}$$

Net useful heat-collection rate is approximately 890W, an efficiency of about 62% relative to the absorbed solar input.

Question 27

(a) On a clear night, the surface radiates strongly in the thermal infrared to a sky whose effective radiative temperature is much lower than the surface temperature (the cold upper atmosphere and ultimately deep space radiate very little back). The radiative loss can exceed the slow conductive heat supply from the warmer air just above and the warmer ground beneath, allowing the surface temperature to fall to the dew point and then below freezing, even when the air at one metre is above 0°C . The local cooling of the surface is fast enough to deposit frost in a few hours of clear sky.

(b) Cloud cover reflects and re-emits thermal radiation back towards the ground, raising the effective sky temperature seen from the surface to a value close to the local air temperature. The net radiative loss $\sigma(T_{\text{ground}}^4 - T_{\text{sky}}^4)$ becomes small or even negative, and the surface stays warm enough that frost cannot form. The cloud layer acts as a radiation barrier blocking the surface's view of the cold upper atmosphere; under it, only the much weaker convective cooling operates.

(c)

Clear night, $T_{\text{ground}} = 278\text{K}$, $T_{\text{sky}} = 250\text{K}$:

$$\frac{P}{A} = \sigma(T_g^4 - T_{\text{sky}}^4) = 5.67 \times 10^{-8} \times (278^4 - 250^4) \approx 117\text{Wm}^{-2}$$

Cloudy night, $T_{\text{ground}} = 278\text{K}$, $T_{\text{sky}} = 280\text{K}$:

$$\frac{P}{A} = 5.67 \times 10^{-8} \times (278^4 - 280^4) \approx -10\text{Wm}^{-2}$$

(The negative sign means the ground gains a small amount of net radiation from the warmer cloud layer rather than losing it.)

Under the clear sky the surface loses about 117Wm^{-2} to space, easily enough to drop the surface below 0°C overnight and form frost. Under the cloudy sky the surface gains a small amount of radiation back from the warmer cloud and stays close to the air temperature; no frost forms.

Question 28

(a) The Earth presents a circular cross-section of area πR^2 to the incoming parallel sunlight, so the absorbed power is $(1 - A)S\pi R^2$, where A is the albedo and S the solar constant. The Earth radiates to space from its full spherical surface area $4\pi R^2$, isotropic in direction, with power $4\pi R^2\sigma T^4$ if treated as a blackbody at temperature T . Equating absorbed and emitted powers in steady state gives $\sigma T^4 = (1 - A)S/4$, with the factor $1/4$ reflecting the ratio of the cross-sectional intercepting area to the full spherical emitting area.

(b) The Earth's atmosphere is largely transparent to incoming solar radiation (concentrated in the visible-and-near-infrared) but strongly absorbing in the thermal infrared (around $10\mu\text{m}$) where the surface radiates. The atmosphere intercepts the surface's outgoing infrared, warms slightly, and re-radiates partly back towards the surface, reducing the net rate of heat loss to space. Steady-state is therefore achieved at a higher surface temperature than for a bare planet without atmosphere; this is the natural greenhouse effect, which raises the surface temperature from the calculated effective value of about -18°C to the observed average of about $+15^\circ\text{C}$. Without it, the Earth's surface would be permanently frozen.

(c)

$$\sigma T^4 = \frac{(1-A)S}{4} = \frac{(1-0.30) \times 1370 \text{ W m}^{-2}}{4} = 240 \text{ W m}^{-2}$$

$$T = \left(\frac{240}{5.67 \times 10^{-8}} \right)^{1/4} \approx 255 \text{ K}$$

Effective equilibrium temperature is approximately 255K, or about -18°C , well below the observed surface average of 288K and reflecting the absence of atmospheric greenhouse warming in this calculation.

Question 29

(a) Glass is highly transparent to visible and near-infrared light (roughly 0.4 to $2.5\mu\text{m}$), where solar radiation is concentrated, and admits the bulk of incoming sunlight to the greenhouse interior. At thermal-infrared wavelengths around $10\mu\text{m}$, where a body at about 300K emits its peak radiation, glass is essentially opaque and absorbs strongly. The glass therefore lets sunlight in (warming the interior) but blocks the escape of longer-wavelength thermal radiation from the warmed contents (preventing heat from leaving), a property known as spectral selectivity. The same trick works for a planetary atmosphere with selective greenhouse gases such as CO_2 and H_2O vapour.

(b) Polythene transmits both the incoming visible-and-near-IR radiation **and** the outgoing thermal IR, so it does not block the radiative loss as effectively as glass does. The polythene greenhouse retains heat mainly through its barrier against air-flow convection (it stops the warm interior air from escaping into the outside), while glass adds a radiative-loss barrier on top of the convective barrier. Polythene is therefore a partial improvement over no enclosure (it stops warm-air convection and provides shelter from rain), but its thermal performance is inferior to glass because it does not block the radiative heat-loss pathway.

(c)

$$P_{\text{rad}} = \sigma A(T_{\text{glass}}^4 - T_{\text{air}}^4) = 5.67 \times 10^{-8} \times 4.0 \times (293^4 - 278^4) \text{ W} \approx 107 \text{ W}$$

A perfectly transparent zero-emissivity film loses zero heat by radiation; the glass panel loses approximately 107W. The radiative loss through actual glass is therefore non-negligible on cool nights, but is the price paid for being opaque to incoming long-wave radiation from a warm interior emitting back to a colder exterior.

Question 30

(a) For an ideal gas, $\frac{1}{2}m\bar{c}^2 = \frac{3}{2}k_B T$, where m is the mass of one molecule and $\bar{c}^2$ the mean square speed. The right-hand side depends only on temperature, so the average translational kinetic energy per molecule is fixed by T alone, independently of the mass of the molecule. Solving for the speed: $v_{\text{rms}} = \sqrt{3k_B T/m} = \sqrt{3RT/M}$, which depends inversely on the square root of the molar mass. A lighter molecule moves faster, but with the same average kinetic energy, as a heavier molecule at the same temperature.

(b) Both gases have the same average translational kinetic energy per molecule (set by the common temperature), but the lighter hydrogen molecules ($M = 2\text{g mol}^{-1}$) move much faster than the heavier helium molecules ($M = 4\text{g mol}^{-1}$). Faster molecules collide with the rubber wall more frequently and have a larger probability of penetrating the microscopic pores of the rubber. Hydrogen therefore effuses out of the balloon faster than helium, and the hydrogen balloon deflates first.

(c)

$$v_{\text{rms},\text{H}_2} = \sqrt{\frac{3RT}{M_{\text{H}_2}}} = \sqrt{\frac{3 \times 8.314 \times 300}{2.0 \times 10^{-3}}} \text{ ms}^{-1} \approx 1934 \text{ ms}^{-1}$$

$$v_{\text{rms},\text{N}_2} = \sqrt{\frac{3RT}{M_{\text{N}_2}}} = \sqrt{\frac{3 \times 8.314 \times 300}{28 \times 10^{-3}}} \text{ ms}^{-1} \approx 517 \text{ ms}^{-1}$$

$$\frac{v_{\text{rms},\text{H}_2}}{v_{\text{rms},\text{N}_2}} = \sqrt{\frac{M_{\text{N}_2}}{M_{\text{H}_2}}} = \sqrt{14} \approx 3.74$$

The rms speed of hydrogen at 300K (about 1934ms^{-1}) is about 3.74 times that of nitrogen (about 517ms^{-1}), reflecting the inverse square-root dependence on molar mass at the same temperature.

Question 31

(a) The total translational kinetic energy of a sample of n moles of an ideal gas is $\frac{3}{2}nRT$, which scales with the number of molecules present. A small flask of gas contains of order 10^{22} molecules; each carries a tiny kinetic energy of order 10^{-21}J at room temperature, but the product is of order kilojoules. The gas appears at rest because the molecular velocities are randomly directed, so the vector sum of momenta is zero; the **scalar** kinetic energy is not zero, because energy adds without regard to direction.

(b) The kicked football carries the kinetic energy of a single object of ordinary mass moving at an ordinary speed. The stationary football sits in air, which itself contains an enormous number of molecules per unit volume (of order Avogadro's number per mole, with several moles in even a small volume of surrounding air). Each molecule's translational kinetic energy is tiny, but their sum over the entire surrounding gas vastly exceeds the bulk kinetic energy of any one sports ball. The stationary-football-and-air system therefore contains **more** translational kinetic energy than the kicked football. The gas's energy is distributed isotropically across the molecules, with the vector sum of momenta close to zero; this is why it does not push the football in any particular direction, and the stationary football appears at rest despite being immersed in this large pool of molecular kinetic energy.

(c)

$$KE_{\text{gas}} = \frac{3}{2}nRT = \frac{3}{2} \times 2.0\text{mol} \times 8.314\text{Jmol}^{-1}\text{K}^{-1} \times 300\text{K} \approx 7483\text{J}$$

$$KE_{\text{ball}} = \frac{1}{2}mv^2 = \frac{1}{2} \times 0.45\text{kg} \times (25\text{ms}^{-1})^2 \approx 141\text{J}$$

$$\frac{KE_{\text{gas}}}{KE_{\text{ball}}} = \frac{7483\text{J}}{141\text{J}} \approx 53$$

Total translational KE of the gas is approximately 7483J, about 53 times the kinetic energy of the kicked football. The KE per molecule is tiny, but N_A molecules per mole make the total substantial.

Question 32

(a) **The dashed grey curve**, peaked at higher speed and broadened, corresponds to the higher temperature. As T rises, the average translational kinetic energy per molecule rises in proportion to T , so the distribution is centred on a larger speed: the peak shifts to the right (higher v_p). The total area under each curve must remain equal to the total number of molecules, which is fixed; the peak height therefore falls and the distribution spreads to a wider range of speeds, giving the broader profile at higher temperature.

(b) Although the typical speed of a perfume molecule is around 400ms^{-1} (comparable to a fast bicycle), the molecule does not travel in a straight line through the room: it suffers about 10^{10} collisions per second with surrounding air molecules, and after each collision its direction changes randomly. The net displacement therefore grows much more slowly than the straight-line speed would predict, scaling as $\sqrt{t}$ rather than t (Brownian-like diffusion). The perfume effectively performs a slow random walk, taking many seconds to cross a few metres.

(c) From the diagram, the most probable speed of nitrogen at 300K is approximately 400ms^{-1} (peak position), and the rms speed is approximately 500ms^{-1} (slightly to the right of the peak). From the kinetic-theory formula:

$$v_{\text{rms}} = \sqrt{\frac{3RT}{M}} = \sqrt{\frac{3 \times 8.314 \times 300}{28 \times 10^{-3}}}\text{ms}^{-1} \approx 517\text{ms}^{-1}$$

The graph estimate ($\sim 500\text{ms}^{-1}$) is in good agreement with the calculated value (517ms^{-1}).

Question 33

(a) Effusion is the escape of gas molecules through a hole much smaller than the **mean free path** (the average distance a molecule travels between two successive collisions with other molecules), so molecules pass one by one without collisions in the hole; the rate is therefore proportional to the mean speed of the molecules and obeys Graham's law strictly. Diffusion is the spreading of one gas through another by random walk, in which molecules undergo many collisions per centimetre travelled; the rate depends on the diffusion coefficient D , which involves both molecular speed and mean free path. Graham's law applies approximately to diffusion (because both processes scale roughly as the molecular speed), but exactly only to effusion.

(b) By Graham's law, the ratio of effusion rates of the two isotopes is $\sqrt{352/349} \approx 1.0043$, a 0.4% enrichment per pass. Starting from natural uranium (0.7% ^{235}U), reaching the 4% enrichment needed for reactor fuel requires the lighter isotope to pass through several hundred porous-membrane stages. Each stage adds only a tiny enrichment, but the cumulative effect of hundreds of stages in series compounds the small per-stage advantage into a useable enrichment.

(c)

$$\frac{r_{\text{He}}}{r_{\text{O}_2}} = \sqrt{\frac{M_{\text{O}_2}}{M_{\text{He}}}} = \sqrt{\frac{32}{4.0}} = 2.83$$

Helium effuses at approximately 2.83 times the rate of oxygen under identical conditions, reflecting the inverse-square-root dependence on molar mass.

Question 34

(a) The pitch of a human voice is set by the resonant frequencies of the vocal-tract cavities (mouth, throat), which scale as $f_n \propto v_{\text{sound}}/L$, where v_{sound} is the speed of sound in the gas filling the cavity and L is the cavity length. The vocal cords vibrate at the same fundamental frequency in either gas; what changes is the resonant-cavity filtering. Helium has both a higher γ and a much smaller molar mass than air, so v_{sound} in helium is about three times that in air, and all of the formant frequencies of the voice rise by the same factor of about three, producing the audibly higher-pitched effect.

(b) Helium contains no oxygen; inhaling pure helium from a cylinder replaces the oxygen-containing air in the lungs with an inert gas, and the body cannot extract oxygen from helium. Within seconds the oxygen partial pressure in the lung drops to a level too low to support consciousness, and the worker can pass out without warning and (if the pure helium is continued) suffocate. The high-pitched-voice effect is harmless in itself, but the means of producing it (direct inhalation of pure helium) is dangerous because helium does not trigger the body's CO_2 -driven breathing reflex; the worker may not feel breathless before passing out.

(c)

$$v_{\text{sound,air}} = \sqrt{\frac{\gamma_{\text{air}}RT}{M_{\text{air}}}} = \sqrt{\frac{1.40 \times 8.314 \times 300}{29 \times 10^{-3}}} \text{ ms}^{-1} \approx 347 \text{ ms}^{-1}$$

$$v_{\text{sound,He}} = \sqrt{\frac{\gamma_{\text{He}}RT}{M_{\text{He}}}} = \sqrt{\frac{(5/3) \times 8.314 \times 300}{4.0 \times 10^{-3}}} \text{ ms}^{-1} \approx 1019 \text{ ms}^{-1}$$

$$\frac{v_{\text{sound,He}}}{v_{\text{sound,air}}} \approx \frac{1019}{347} \approx 2.94$$

Speed of sound in helium is approximately 1020 ms^{-1} , against approximately 347 ms^{-1} in air; the helium voice's resonant frequencies are therefore about 2.9 times higher than in air.

Question 35

(a) A diatomic molecule has its two atoms joined by a chemical bond along a single axis. Rotation about an axis perpendicular to this bond carries the two atoms in circles of large radius and stores energy as rotational kinetic energy; rotation about the bond axis itself, however, carries each atom in a circle of essentially zero radius (the atoms sit on the axis), so the moment of inertia about that axis is much smaller than the other two and the rotational quantum is correspondingly larger than $k_B T$ at room temperature. The bond-axis rotational mode is therefore frozen out, and only the two perpendicular rotational modes are accessible at room temperature, giving five active degrees of freedom in total (3 translational + 2 rotational).

(b) By equipartition, each accessible degree of freedom contributes $\frac{1}{2} k_B T$ of energy per molecule, so the total internal energy per molecule is $f \cdot \frac{1}{2} k_B T$. Argon, monatomic with $f = 3$, has total $U = \frac{3}{2} nRT$ per mole. Nitrogen, diatomic with $f = 5$, has total $U = \frac{5}{2} nRT$ per mole. For the same n and T , nitrogen contains $5/3$ times as much internal energy as argon, and is correspondingly more capable of doing work or absorbing heat. The same number of molecules can store more energy in the diatomic gas because each molecule has more accessible energy stores.

(c)

$$U_{\text{total}} = \frac{f}{2} nRT = \frac{5}{2} \times 1.0 \times 8.314 \times 300 \text{ J} \approx 6236 \text{ J}$$

Translational contribution ($f = 3$):

$$U_{\text{trans}} = \frac{3}{2} nRT = \frac{3}{2} \times 1.0 \times 8.314 \times 300 \text{ J} \approx 3741 \text{ J}$$

Rotational contribution ($f = 2$):

$$U_{\text{rot}} = \frac{2}{2} nRT = nRT = 1.0 \times 8.314 \times 300 \text{ J} \approx 2494 \text{ J}$$

$$\frac{U_{\text{trans}}}{U_{\text{total}}} = \frac{3}{5} = 60\%, \quad \frac{U_{\text{rot}}}{U_{\text{total}}} = \frac{2}{5} = 40\%$$

Total internal energy is 6236J, of which 60% (3741J) is translational and 40% (2494J) is rotational.

Question 36

(a) At constant volume, the gas does no work on the surroundings, and all heat input goes into increasing the internal energy: $Q_V = \Delta U = (f/2)nR\Delta T$, hence $C_V = (f/2)R$ per mole. At constant pressure, the gas expands as it heats: the heat input raises both the internal energy AND does work $P\Delta V = nR\Delta T$ against the surroundings (using the equation of state). The heat input per mole per kelvin is therefore $C_p = (f/2)R + R = ((f+2)/2)R$, with the additional R corresponding to the work done by the expanding gas at constant pressure. The result $C_p - C_V = R$ is independent of f and is known as Mayer's relation.

(b) Argon is monatomic ($f = 3$) so its $C_V = (3/2)R$, while nitrogen is diatomic ($f = 5$) so its $C_V = (5/2)R$. For the same heating power supplied per mole, the heat-up rate $dT/dt = P_{\text{heat}}/(nC_V)$ is larger for argon than for nitrogen, so argon's temperature rises faster but the total energy stored at any temperature is less. In the welding context, this means the argon shielding-gas envelope reaches its design temperature quickly and stays there, with less thermal mass to drag the arc temperature down; the arc actually runs **hotter** in argon, contradicting the premise. The benefit of argon for welding is its inertness (non-reactive with hot metal) rather than thermal-management.

(c)

$$C_V = \frac{f}{2}R = \frac{5}{2} \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1} = 20.79 \text{ J mol}^{-1} \text{ K}^{-1}$$

$$C_p = \frac{f+2}{2}R = \frac{7}{2} \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1} = 29.10 \text{ J mol}^{-1} \text{ K}^{-1}$$

$$\gamma = \frac{C_p}{C_V} = \frac{f+2}{f} = \frac{7}{5} = 1.40$$

For a diatomic ideal gas at room temperature: $C_V = 20.79 \text{ J mol}^{-1} \text{ K}^{-1}$, $C_p = 29.10 \text{ J mol}^{-1} \text{ K}^{-1}$, $\gamma = 1.40$.

Question 37

(a) From the first law $Q = \Delta U + W$, the heat input goes both into raising the internal energy of the gas and into doing work on the surroundings. At constant pressure the gas expands as it heats; the work done by the gas $W = P\Delta V$ is non-zero, so Q exceeds ΔU by exactly W . The relationship is general for any gas-on-piston experiment at constant pressure.

(b) The Sun supplies heat Q to the helium. By the first law $Q = \Delta U + W$, this heat is partitioned between an internal-energy rise ΔU (the helium gets hotter) and the work W done by the gas on the surroundings (the gas expands and pushes back the atmosphere by an amount V at atmospheric pressure P_{atm}). Q exceeds ΔU by $W = P_{\text{atm}}\Delta V$, in agreement with the first law of thermodynamics.

(c)

$$\Delta U = nC_V\Delta T = 0.50 \times \frac{3}{2} \times 8.314 \times 300 \text{ J} \approx 1871 \text{ J}$$

$$W = nR\Delta T = 0.50 \times 8.314 \times 300 \text{ J} \approx 1247 \text{ J}$$

$$Q = nC_p\Delta T = 0.50 \times \frac{5}{2} \times 8.314 \times 300 \text{ J} \approx 3118 \text{ J}$$

Heat supplied is approximately 3118J, internal energy change is 1871J, and the work done by the gas is 1247J.

Check: $\Delta U + W = 1871 + 1247 = 3118 \text{ J} = Q$, in agreement with the first law.

Question 38

(a) Internal energy is a state function: it depends only on the current state of the gas (its T , P , V), not on how the gas reached that state. The work done by the gas, by contrast, is $W = \int P dV$ along the chosen path on the PV diagram, which is the area under the chosen path. Different paths between the same two states enclose different areas, so the work done is different. Heat exchange $Q = \Delta U + W$ is also path-dependent, since ΔU is fixed and W varies.

(b) In cool weather the compression is approximately isothermal: heat escapes to the surroundings as the gas is compressed, keeping the temperature close to ambient and the pressure rise more gradual. In hot weather (or when the compression is fast enough that heat cannot escape), the compression is closer to adiabatic: the gas heats during compression and the pressure rises more steeply for the same volume reduction. The work done on the gas equals the area under the compression curve on the PV diagram; the steeper adiabatic curve requires more work to reach the same final pressure than the shallower isothermal curve, so the compressor uses more work in hot weather.

(c) Path 1: isochoric $A \rightarrow (P_0/2, V_0)$ has $W_{1a} = 0$. Then isobaric expansion at $P_0/2$ from V_0 to $2V_0$:

$$W_{1b} = \frac{P_0}{2}(2V_0 - V_0) = \frac{1}{2}P_0V_0$$

$$W_{\text{path 1}} = 0 + \frac{1}{2}P_0V_0 = 0.5 P_0V_0$$

Path 2: isobaric expansion at P_0 from V_0 to $2V_0$:

$$W_{2a} = P_0(2V_0 - V_0) = P_0V_0$$

Then isochoric drop at $2V_0$: $W_{2b} = 0$:

$$W_{\text{path 2}} = P_0V_0 + 0 = 1.0 P_0V_0$$

Path 2 delivers P_0V_0 of work, twice as much as path 1's $0.5 P_0V_0$, because path 2 expands at the higher pressure P_0 throughout while path 1 expands only at $P_0/2$.

Question 39

(a) The work done by the gas in any infinitesimal step is $dW = PdV$. Around a closed cycle, $W_{\text{net}} = \oint P dV$, which is the signed area enclosed by the path. The area-as-line-integral rule applies to any closed curve, regardless of shape: rectangular, triangular, oval, or any other. Clockwise traversal on the standard PV diagram gives a positive enclosed area (heat engine direction); counter-clockwise gives a negative area (refrigerator direction). The shape of the cycle determines the area; the direction of traversal determines the sign.

(b) The area enclosed by the indicator diagram on the PV plane equals the net work done by the gas per cycle. The shape of the diagram (triangular, rectangular, or rounded) does not matter for calculating the net work; only the enclosed area does. The technician measures the area (by counting squares on graph paper, by planimeter, or by digital integration) to obtain the work output per cycle, then multiplies by the engine's cycle frequency to get the power output.

(c) The triangle has base $V_C - V_A = 1.0 \times 10^{-3} \text{ m}^3$ and height $P_B - P_A = 1.0 \times 10^5 \text{ Pa}$:

$$W_{\text{net}} = \frac{1}{2} \times \text{base} \times \text{height} = \frac{1}{2} \times 1.0 \times 10^{-3} \text{ m}^3 \times 1.0 \times 10^5 \text{ Pa} = 50 \text{ J}$$

Net work done by the gas per cycle is 50J, equal to the area of the triangular cycle on the PV diagram, with positive sign because the cycle is traversed in the heat-engine direction.

Question 40

(a) For an ideal gas the internal energy depends only on temperature: $U = U(T)$. In an isothermal process, T is constant, so $\Delta U = 0$. The first law $\Delta U = Q - W$ then gives $Q = W$: every joule of heat absorbed from the surroundings is delivered as work done by the gas, with no fraction stored as internal energy. The gas acts as a perfect heat-to-work converter at the surrounding-temperature isotherm, with no temperature change.

(b) The pipeline gas, held by ground contact at the soil temperature, undergoes an essentially isothermal expansion as it flows from high to low pressure. By the first law for an isothermal expansion of an ideal gas, $\Delta U = 0$ and $Q = W$. The gas does positive work $W = nRT \ln(P_i/P_f)$ in expanding against the back pressure; to keep the temperature unchanged, an equal quantity of heat Q must flow into the gas from the surrounding soil. The gas does not visibly heat up because the inflow of heat exactly balances the work loss; the soil cools imperceptibly to compensate.

(c)

$$W = nRT \ln \frac{P_i}{P_f} = 1 \times 8.314 \times 290 \times \ln \frac{80}{5} \text{ J} \approx 6685 \text{ J}$$

Work done by the gas during the isothermal expansion is approximately 6685J per mole.

Question 41

(a) Carnot's theorem requires the cycle to be reversible: every infinitesimal step must occur quasi-statically, with no friction, no temperature gradients across heat-exchange surfaces, and no non-equilibrium expansions. Quasi-static means the process is slow enough that the system stays in equilibrium throughout, which implies an infinitely long cycle time and zero power output. Real engines run at finite speed, with finite temperature drops across heat exchangers, friction in moving parts, and finite-rate combustion; all of these introduce irreversibility and reduce the efficiency below Carnot. Practical engines therefore trade some efficiency for finite power output, reaching perhaps 50 to 70% of the Carnot ceiling.

(b) The Carnot ceiling between 150°C (423K) and 27°C (300K) is $\eta_{\text{Carnot}} = 1 - 300/423 \approx 29\%$. Two factors drive the realised efficiency below this: (i) finite temperature drops across the heat-exchanger surfaces (the working fluid cannot be at exactly

the reservoir temperature; some difference is needed to drive heat flow at the design rate), reducing the effective Carnot ceiling for the working fluid; and (ii) friction and viscous losses in the turbine, pump, and generator, dissipating some of the work output as heat. A realistic geothermal plant operates at 15 to 20% efficiency.

(c)

$$\eta_{\text{Carnot}} = 1 - \frac{T_{\text{cold}}}{T_{\text{hot}}} = 1 - \frac{300\text{K}}{423\text{K}} = 0.291 = 29.1\%$$

$$W_{\text{max}} = \eta_{\text{Carnot}} \times Q_{\text{hot}} = 0.291 \times 1000\text{J} = 291\text{J}$$

Carnot efficiency is approximately 29% for this temperature pair, with at most 291J of work extractable per kilojoule of heat absorbed.

Question 42

(a) The primary mechanism is convection: the warm calorimeter surface heats the air in contact with it, and the warm air rises and is replaced by cooler air. Radiation also contributes, but at the modest excess temperatures of a typical calorimetry experiment (10–30°C above room temperature), convection dominates. The error increases with longer experimental times because the heat loss accumulates: the system loses energy continuously to the surroundings, and the longer it takes for the water and metal to reach thermal equilibrium, the more energy escapes before the final temperature is read.

(b) $\text{COP}_{\text{max}} = T_{\text{cold}}/(T_{\text{warm}} - T_{\text{cold}}) = 277/(301 - 277) = 277/24 \approx 11.5$. The advertised value 12 slightly exceeds the Carnot ceiling and is therefore physically infeasible. Real refrigerators typically achieve 25 to 50% of the Carnot ceiling, so a realistic COP for these reservoir temperatures is about 3 to 6. The manufacturer's claim of 12 is too good to be true.

(c)

$$\text{COP}_{\text{max}} = \frac{T_{\text{cold}}}{T_{\text{warm}} - T_{\text{cold}}} = \frac{277\text{K}}{301\text{K} - 277\text{K}} = \frac{277}{24} \approx 11.5$$

$$\frac{Q_{\text{cold}}}{dt} = \text{COP}_{\text{max}} \times P = 11.5 \times 200\text{W} \approx 2308\text{W}$$

Maximum coefficient of performance is approximately 11.5, and the maximum rate of heat extraction is approximately 2300W.

Question 43

(a) After sunset, the iron roof loses heat by two dominant mechanisms. First, radiation: the roof, now at perhaps 50–60°C, radiates infrared energy upward toward the cold night sky (which can be as cold as –20°C in the upper atmosphere). The temperature difference between the roof and the sky is large, driving substantial radiative loss. Second, convection: any breeze passing over the roof carries heat away by forced convection. The roof cools faster than the concrete walls because iron has a much lower thermal mass per unit area than concrete: iron is thin (typically 0.5mm corrugated sheet) and has a relatively low specific heat capacity (450Jkg^{–1}K^{–1}), while concrete walls are thick (150–200mm) and have a large thermal mass. The roof's small thermal reservoir is emptied quickly by the same rate of radiative and convective loss that would take much longer to cool the massive concrete walls.

(b) The gas does no work on the surroundings as it expands, since the adjacent cylinder is evacuated; there is nothing to push back. The thermal insulation prevents heat exchange with the surroundings. By the first law $\Delta U = 0$, and for an ideal gas this gives $\Delta T = 0$. The temperature of the gas in the combined cylinders is therefore essentially the same as the initial temperature, regardless of the speed at which the expansion occurred. (For a real gas, intermolecular forces produce a small temperature change in a free expansion: slight cooling for most gases at room temperature, the Joule-Thomson effect, but this effect vanishes for an ideal gas.)

(c) For an ideal gas, free expansion has $Q = 0$, $W = 0$, hence $\Delta U = 0$ and $\Delta T = 0$:

$$T_f = T_i = 300\text{K}$$

Initial volume $V_i = 1.0 \times 10^{-3}\text{m}^3$ (half the container); final volume $V_f = 2.0 \times 10^{-3}\text{m}^3$. Constant temperature means $PV = \text{const}$:

$$P_f = P_i \times \frac{V_i}{V_f} = 1.0 \times 10^5\text{Pa} \times \frac{1}{2} = 5.0 \times 10^4\text{Pa}$$

$$\Delta U = nC_v\Delta T = nC_v \times 0 = 0$$

Final temperature is 300K, final pressure is $5.0 \times 10^4\text{Pa}$ (half the initial value), and the change in internal energy is zero, all in agreement with the qualitative argument from the first law of thermodynamics.