Chapter 12

CURRENT ELECTRICITY

INTRODUCTION

Late one cloudy afternoon Mr. Akilikubwa reached for the switch on the laboratory wall. The fluorescent tubes overhead came on instantly, with the small click and pink flicker that Miono students had come to know. Kipanga, sitting at the front bench, was paying attention this time.

Mr. Akilikubwa: *Kipanga, you closed the switch and the room lit up at once. How fast did the electricity travel?*

Kipanga (confident)**:** *At the speed of light, sir. Three hundred thousand kilometres per second.*

Mr. Akilikubwa: *The signal did, yes. But the electrons themselves crawl. In the copper wire above your head, the electrons are drifting toward the bulbs at about half a millimetre every second. Slower than an ant carrying a leaf.*

Kipanga (suddenly less confident)**:** *But the bulb lit up immediately, sir.*

Kipute (working it out before Mr. Akilikubwa can)**:** *Of course it did. Every electron in the wire was already there before the switch closed. When the switch closed, every electron started pushing the next one along at the same instant. The bulb lit up because the wire was already full of electrons, not because any one electron had reached the bulb.*

Mr. Akilikubwa: *Kipute has it. The wire is like a long queue of people passing buckets. No person walks the length of the queue, but a bucket arrives at the end almost the moment the first bucket is handed in. The signal travels at the speed of light. The electrons themselves move at the speed of an ant. The whole of this chapter is about that crawl: how fast it is, what slows it down, how it heats the wire, how we steer it through circuits, and what happens when there are no electrons already in place and we must make our own.*

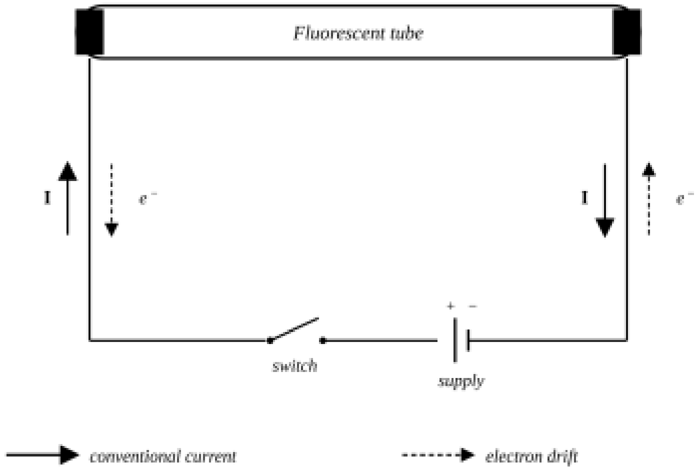


Figure: ***A fluorescent tube driven through a closed switch.*** *Solid arrows mark the conventional current I. Dashed arrows mark the electron drift, opposite to I on both wires.*

Chapter 11 ended with the time-varying current that flows briefly while a capacitor is charging or discharging. That current was the brief settling of charges into a new equilibrium; it lived for a few time constants and then died. The current of this chapter is different. It is what happens when the equilibrium is never reached at all, because a steady source keeps pushing charges through the circuit at the same rate forever. Once any initial transient has settled, every device in the world that runs on direct current is running on the steady current of this chapter, from the torch in a student's pocket to the immersion heater in a Nyamagana household to the power line that lit Mr. Akilikubwa's laboratory just now.

By the end of the chapter the reader will know what an electric current really is in microscopic terms, how Ohm's law arises from the collisions of drifting electrons with the lattice, what resistivity and conductivity and the temperature coefficient mean and where they come from, what an EMF is and why every battery has an internal resistance, how to combine resistors and cells in series and in parallel and in networks that bend under symmetry, how to convert a galvanometer into an ammeter or a voltmeter using only a shunt or a multiplier, how to apply Kirchhoff's two laws to any circuit that no series-and-parallel reduction can untangle, how the Wheatstone bridge and the potentiometer measure resistances and EMFs to a precision that no direct meter can match and why null methods are accurate at all, how electrical energy converts into heat and how to design a heating device that delivers exactly the power the design calls for, why TANESCO transmits power at hundreds of kilovolts rather than at household voltage, what conditions a load resistor must satisfy to receive the maximum power a source can deliver, how a gas at low pressure can be made to conduct when its molecules are ionised, what distinguishes an initiated from a self-sustaining discharge, and how the light from a gas discharge tube carries a fingerprint of the gas inside.

The next twelve sections of this chapter answer those questions in the order they are best learned. We begin with the crawl itself.

ELECTRIC CURRENT AND THE DRIFT OF ELECTRONS

The chapter opener finished with electrons crawling at half a millimetre every second while the bulb lit immediately. This section puts the mathematics on that scene. To make a wire carry a steady current, three numbers have to cooperate: *how fast each electron drifts, how thickly the electrons fill the wire, and how big the cross-section of the wire is.* Once those three numbers are in hand, the current itself follows in a single line of algebra. The job of the section is to identify each of the three, name it, and put it in its place in that line.

What an electric current is

Imagine a copper wire with a steady current flowing through it. Pick any cross-section of the wire, a single slice perpendicular to its length. In one second a certain quantity of charge crosses that slice. In the next second the same quantity crosses again, and the second after that, and so on, as long as the source keeps pushing. The current I is the quantity of charge that crosses the slice per second:

$$\mathbf{I = \frac{Q}{t}}$$

where Q is the charge that crosses in time t. The unit of current follows from the definition. One ampere is one coulomb per second:

$$1\text{A} = 1\frac{\text{C}}{\text{s}}$$

If the current is not steady, the same idea applies to a vanishingly short interval of time rather than a finite one:

$$\text{I} = \frac{\text{dQ}}{\text{dt}}$$

We will use the steady form for almost all of this chapter. The varying form returns in Chapter 13 when we deliberately make the current change with time and watch the magnetic field follow it.

Charge does not come in arbitrary amounts. It comes in lumps of electronic charge $e = 1.6 \times 10^{-19}$ C. If N electrons cross our slice in time t, the total charge that has crossed is:

$$\text{Q} = \text{Ne}$$

and the current is therefore:

$$\mathbf{I = \frac{Ne}{t}}$$

A 1A current carries one coulomb per second past every cross-section of the wire. The number of electrons that have to cross to deliver that coulomb in one second is $N = 1\text{C} / (1.6 \times 10^{-19}\ \text{C}) \approx 6.25 \times 10^{18}$ electrons every second, more than six quintillion of them. That is the first surprise of the chapter. Even the smallest current of any practical interest moves more electrons through every cross-section every second than there are stars in every galaxy in the observable universe.

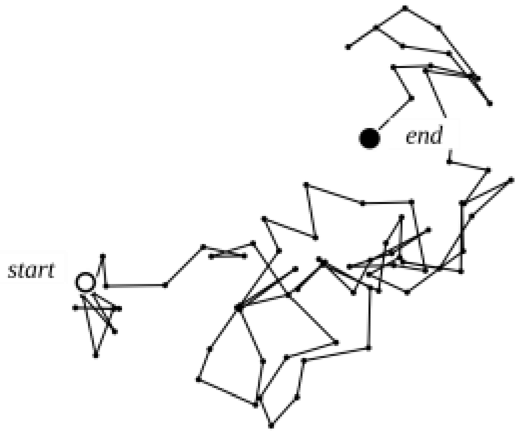
A note on direction:

Convention. Throughout this book, and on every figure that follows, the arrow drawn on a wire shows the direction of **conventional current**, which *is from the positive terminal of the source to the negative terminal through the external circuit.* Electrons in a metal drift in the **opposite** direction. We use the conventional direction in every formula. The electron-drift direction enters the discussion only when a question explicitly asks about it.

The convention is older than the discovery of the electron. Benjamin Franklin guessed in the 1750s that electric current was the flow of a single positive fluid from one place to another, and his choice stuck. We now know that in a metal the carriers are negative electrons. In a semiconductor the carriers may be electrons (negative) or holes (vacancies in the electron sea that behave as positive carriers and drift opposite to the electrons). In an electrolyte both positive and negative ions contribute, drifting in opposite directions. Regardless of what the carriers are, the conventional current measures the net flow of positive charge. The convention is therefore not wrong; it is simply about a different thing from the actual motion of the carriers.

The microscopic picture: random thermal speed and the small drift on top

Take a copper wire at room temperature with the switch open. The free electrons (about one per copper atom, the outermost electron that has parted with its atom) are not at rest. They fly about randomly at speeds of around 10^5 to 10^6 m/s, colliding with the lattice ions every few hundredths of a picosecond, ricocheting, colliding again, ricocheting again. Each electron is going somewhere very fast at every moment, but the swarm as a whole goes nowhere. The average velocity, taken over any single electron over any meaningful time, is exactly zero. The figure below sketches a typical electron's path through many collisions. The start and the end are close together; the motion in between is chaos.



no applied field: net displacement is small after many collisions

Figure: ***A single electron's path during many collisions with the lattice ions, in a metal with no applied electric field.*** *The motion is fast and chaotic, but the start and the end are close together: the net displacement is essentially zero.*

Now close the switch. An electric field E appears inside the wire, pointing along the wire's length. Each electron, in addition to its random thermal motion, picks up a tiny extra velocity in the direction opposite to E (because electrons are negative). This tiny added velocity is the **drift velocity**, written $\mathbf{v_d}$. The next figure shows what the same electron's path looks like now: the chaos is essentially unchanged, but a small systematic drift has been laid over it.

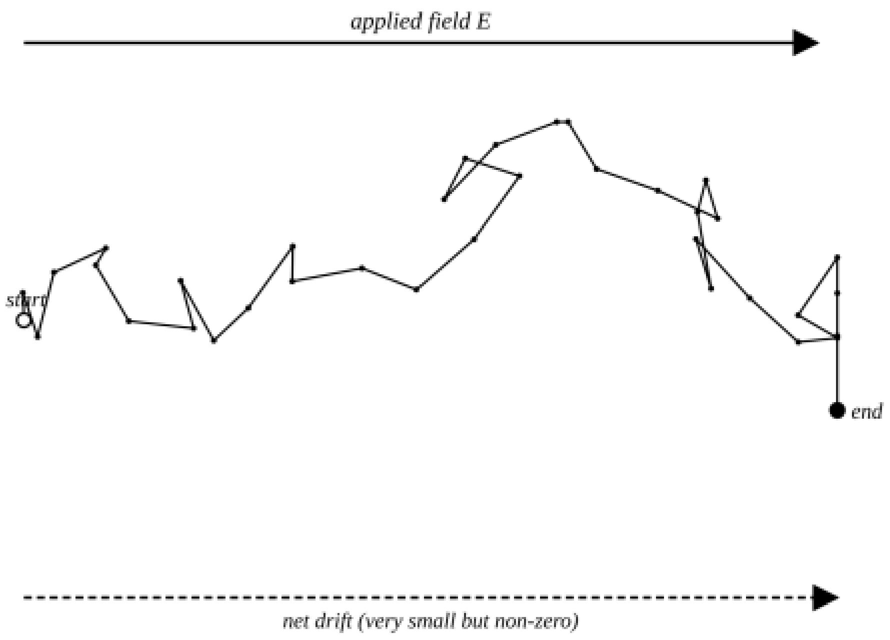


Figure: *With an applied field E along the wire, the same electron's chaos is now slightly biased***.** *After many collisions, the electron has drifted a small distance in one direction.* ***The drift is the current****.*

The drift velocity in a typical copper wire carrying a typical current is around 10^{-4} to 10^{-3} m/s, that is, a tenth of a millimetre to one millimetre every second. An ant carrying a leaf moves several centimetres every second. The drift is, almost literally, slower than the ant.

This is the central pedagogical point of the chapter opener and this section together. The drift is a tiny perturbation on top of a furious random motion. When the switch closes, the electric field fills the wire at the speed of light, and every electron everywhere in the wire begins drifting at the same instant. The bulb lights immediately because the wire is already full of electrons. No electron has to travel from the switch to the bulb; the field tells the electrons that already line the path between switch and bulb to start moving, and they do, **all at once.**

From the drift velocity to the current

We now have a picture of every free electron in the wire drifting uniformly with speed v_d in the conventional direction. To convert that picture into the current I, picture the wire as a cylinder of cross-sectional area A and length l, and call the number density of free electrons n, that is, the number of free electrons per cubic metre.

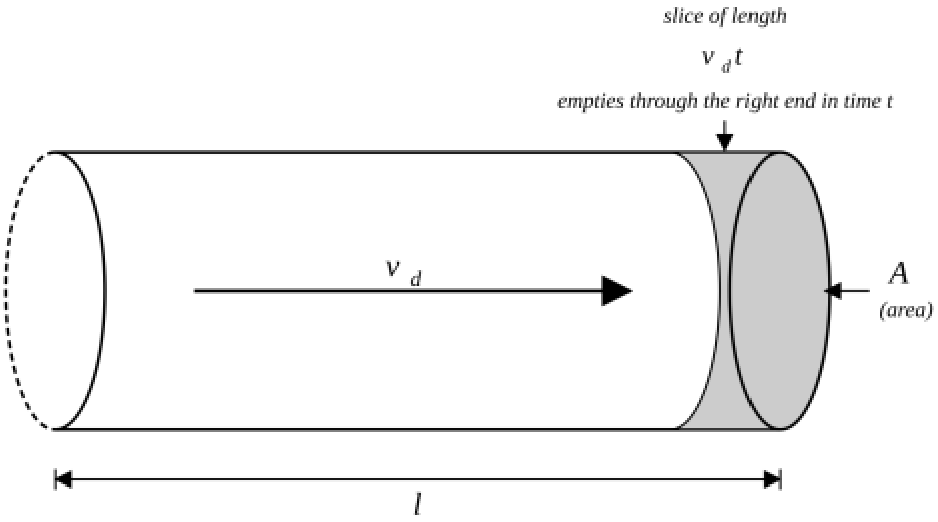


Figure: ***A cylindrical conductor used in the derivation.*** *Free electrons inside drift with speed v_d in the direction shown. The shaded slice at the right is the portion that empties through the right end during a time t.*

The total number of free electrons inside the cylinder is the number density multiplied by the volume of the cylinder:

$$N = nAl$$

In a time $t = \frac{l}{v_d}$ every electron in the cylinder has drifted out through the right end, assuming the drift is uniform throughout the cylinder, which to first order it is. The total charge that crosses the right end during that time is Ne. So the current is:

$$I = \frac{Ne}{t} = \frac{(nAl) \times e}{\left(\frac{l}{v_d}\right)} = neAv_d$$

The l in the numerator and the l in the denominator cancelled cleanly, and out fell a formula that does not depend on the length of the cylinder at all. It depends only on quantities that belong to the material (n and e) and to the cross-section of the wire (A) and to how fast the electrons are drifting (v_d). The result is the single most important formula of the section, the one every later calculation in this chapter starts from:

$$\mathbf{I = neAv_d}$$

Solving for the drift velocity:

$$\mathbf{v_d = \frac{I}{neA}}$$

Four ingredients sit on the right-hand side: how many free electrons per cubic metre the material has (n), the electronic charge (e), how big the cross-section is (A), and how much current you are driving (I). Three of the four belong to the wire itself; only one, the current, belongs to whatever is at the ends of the wire. Once we know what the wire is made of and how thick it is, the drift velocity is simply proportional to the current.

Current density

The combination I divided by A appears so often in what follows that it earns its own symbol and its own name. The **current density** is:

$$\mathbf{J = \frac{I}{A}}$$

with units of A/m². For our wire, substituting $I = neAv_d$:

$$\mathbf{J = nev_d}$$

Notice that J does not depend on the size of the wire. It depends only on what carriers are inside the material (n, e) and how fast they are drifting (v_d). *A thin wire and a thick wire made of the same material can carry different currents but the same J, provided their drift velocities are the* ***same****.* J is the intensive form of current; I is the extensive form. When in later sections we want to compare how easily a copper bar and a steel bar carry current, we will compare their values of J at the same applied field, never their values of I (which depend on the size of the bar).

A foreshadowing observation, to be earned properly in the next two sections: J is proportional to the electric field E that drives the current. In symbols:

$$\mathbf{J = \sigma E}$$

where σ is a property of the material called its conductivity. This is the local form of Ohm's law. It tells us that the current density at any point of any wire is set by the field at that point and by the conductivity of the material there, full stop. We will derive both of these from the microscopic collisions in the next section.

Drift versus thermal: the ratio that surprises everybody

We can now make the central comparison promised at the start of the section. The thermal speed of an electron in a metal at room temperature is about 10^5 to 10^6 m/s. The drift speed in the same metal carrying ordinary current is about 10^{-4} to 10^{-3} m/s. The ratio is:

$$\frac{v_{\text{thermal}}}{v_{\text{drift}}} \approx 10^8 \text{ to } 10^9$$

The drift is between a hundred million and a billion times slower than the random motion it is laid on top of. If you could film the electrons and slow the playback enormously, you would see chaotic flying in every direction; only after watching for a very long time, with a very steady eye, would you notice that on average the swarm has shifted by a tiny amount in one direction. That tiny average shift is the **current.**

And that completes the puzzle of the chapter opener. The bulb lights immediately because the wire was already full of electrons before the switch closed. When the switch closes, an electric field propagates through the wire at the speed of light, telling every electron at every point of the wire to start drifting. They all do, simultaneously. The bulb illuminates not because one electron rushed from the switch to it, but because the electrons that were already inside the bulb's filament started drifting at the same instant the electrons inside the switch did. The signal travels at the speed of light. The carriers crawl. The wire connects them because it was full all along.

Five examples follow. They each compute one of three things: how slow the electrons drift, how thickly the current packs through the cross-section, or how many electrons sneak past a given point each second. The numbers you are about to meet are absurd in both directions: miraculously slow and miraculously numerous. Once you have written them down, the queue of buckets from the chapter opener will have earned its keep.

BINDER Example 1

A current of 5A flows through a copper wire whose cross-sectional area is 2mm². The number density of free electrons in the copper is taken as 10^{28} m^{-3}. Find the drift velocity of the free electrons in the wire.

Solution

The drift velocity is given by:

$$v_d = \frac{I}{neA}$$

Substituting:

$$v_d = \frac{5\text{A}}{(10^{28}\text{ m}^{-3})(1.6 \times 10^{-19}\text{C})(2 \times 10^{-6}\text{ m}^2)} = 1.5625 \times 10^{-3}\text{ m/s}$$

Making Sense of the Answer: *A drift velocity of 1.56 × 10^{-3} m/s is roughly 1.5 millimetres every second. If you asked a single electron to walk from a wall socket to a lamp three metres away, at this drift speed the trip would take about 1900 seconds, more than half an hour. The lamp does not wait half an hour to light because the electrons inside its filament were already there when the switch was closed. The lamp lights at the speed of light.*

Think Like a Physicist: *Whenever a question asks for a drift velocity, expect an answer in the* 10^{-4} *to* 10^{-3} *m/s range for ordinary currents and ordinary wires. If you compute a drift velocity larger than 1m/s, somewhere in your arithmetic you have lost a factor of about* 10^3*, and you should hunt the missing thousand down before submitting the answer. Sanity always before substitution.*

BINDER Example 2

A copper wire of cross-sectional area 4mm² and length 4m carries a current of 10A. The number density of free electrons in the copper is $8 \times 10^{28}\ \text{m}^{-3}$. Find: (a) the drift velocity of the free electrons in the wire, (b) the time required for an individual electron to drift the entire length of the wire.

Solution

(a) From the same drift-velocity formula:

$$v_d = \frac{I}{neA}$$

Substituting:

$$v_d = \frac{10\text{A}}{(8 \times 10^{28}\ \text{m}^{-3})(1.6 \times 10^{-19}\text{C})(4 \times 10^{-6}\ \text{m}^2)} = 1.95 \times 10^{-4}\ \text{m/s}$$

(b) The transit time of an individual electron is the length of the wire divided by its drift velocity:

$$t = \frac{l}{v_d}$$

$$t = \frac{4\text{m}}{1.95 \times 10^{-4}\ \text{m/s}} = 2.05 \times 10^4\text{s} \approx 5.70\ \text{hours}$$

Making Sense of the Answer: *An individual electron takes nearly six hours to drift the four metres from one end of the wire to the other. Meanwhile the electric signal that drives the current propagates through that same four-metre wire in roughly* $4\text{m} / (3 \times 10^8\ \text{m/s}) \approx 1.3 \times 10^{-8}$ *s, about thirteen nanoseconds. The signal beats the electron by a factor of roughly* $(5.70 \times 3600) / (1.3 \times 10^{-8}) \approx 1.6 \times 10^{12}$*, about twelve orders of magnitude. The signal flies. The carriers crawl. The wire connects them because it was full all along.*

Think Like a Physicist: *Transit-time questions about drifting electrons are designed to surprise. The intuition for an ordinary copper wire is that an individual electron makes the trip in hours, while the electric signal that drives the current makes the trip in nanoseconds. If your answer for the transit time of an individual electron in a sensibly thick copper wire comes out in milliseconds, you have probably confused signal speed with drift speed; redo the calculation.*

BINDER Example 3

A copper wire has cross-sectional area 1.5mm² and carries a current of 6.0A. Taking the number density of free electrons in copper as $8.5 \times 10^{28}\ \text{m}^{-3}$, find:

(a) the current density J in the wire,

(b) the drift velocity of the free electrons.

Solution

(a) From the definition $J = I/A$:

$$J = \frac{I}{A} = \frac{6.0\text{A}}{1.5 \times 10^{-6}\ \text{m}^2} = 4.0 \times 10^6\text{A/m}^2$$

(b) From:

$$v_d = \frac{J}{ne}$$

Substituting:

$$v_d = \frac{4.0 \times 10^6\text{A/m}^2}{(8.5 \times 10^{28}\ \text{m}^{-3})(1.6 \times 10^{-19}\text{C})} = 2.94 \times 10^{-4}\ \text{m/s}$$

Making Sense of the Answer: *A current density of 4×10^6 A/m^2 sounds enormous until you remember that the cross-section is only 1.5mm^2, so the actual current of 6A is what a household lamp draws. The same J in a wire of cross-section 10mm^2 would still be 4×10^6 A/m^2 because J does not care how thick the wire is. The drift velocity is again in the familiar 10^{-4} m/s range, exactly the order of magnitude the section promised.*

Think Like a Physicist: *Whenever a question hands you a current and a cross-section together, compute J first, then v subscript d if you need it. J is the intensive quantity; it cleans up the algebra by removing the cross-sectional area from your formulas. After J, the drift velocity is one division away.*

REAL Example 4

It is past ten o'clock at night in Miono. Kipute has been reading her notes for two hours under the 100W ceiling lamp in her study room. She reaches up to adjust the angle of the shade, and her hand brushes the supply wire near the lamp holder. The wire is warm. Not hot, not dangerous, but warm enough to notice.

Kipute: *Sir, the wire to my reading lamp felt warm after two hours under it last night. Does that mean the electrons inside it are moving fast?*

Mr. Akilikubwa: *Let us answer that with a number rather than an opinion. Your lamp draws its 100W of power from the 240V mains; the supply wire is ordinary household copper with cross-sectional area 1.5mm^2 and free-electron number density 8.5×10^{28} m^{-3}.* ***Find the drift velocity of the free electrons in that wire, and then we shall talk about whether they are moving fast.***

Help Kipute:

(a) to calculate the drift velocity of the free electrons in the supply wire to her reading lamp, and
(b) from the answer in (a), explain to her why a wire that feels warm to the touch is not a wire full of fast electrons.

Solution

(a) The drift-velocity formula needs the current I, but the question hands us a power and a voltage instead. Power equals voltage times current, so:

$$I = \frac{P}{V}$$

Substituting to $v_d = \frac{I}{neA}$:

:

$$v_d = \frac{P}{VneA} = \frac{100\text{W}}{240\text{V} \times (8.5 \times 10^{28}\ \text{m}^{-3}) \times (1.6 \times 10^{-19}\text{C}) \times (1.5 \times 10^{-6}\ \text{m}^2)} = 2.04 \times 10^{-5}\ \text{m/s}$$

That is about twenty micrometres every second, twenty thousandths of a millimetre every second.

Have you noticed this? Twenty micrometres every second is slower than a snail crawling on tar at noon. An individual electron in the wire to Kipute’s reading lamp would need nearly an hour to drift the width of her fingertip. The electrons are not moving fast in any sense that the word ordinarily carries. And yet the wire is warm to the touch after two hours of operation. The warmth, therefore, cannot be a consequence of the electrons moving fast, because the electrons are not moving fast. It must come from somewhere else.

That somewhere else is the collisions. Now, let us answer part (b).

(b) The warming arises from collisions between the electrons and the lattice ions. Every time a drifting electron meets a lattice ion, it surrenders a tiny piece of the velocity it had just picked up from the field, and the lattice absorbs that velocity as a small vibration of the ion. With more than 10^{28} m^{-3} free electrons in every cubic metre of copper, slamming into the lattice ions many times every picosecond, those small vibrations add up to a wire that is detectably warm after long operation.

The mechanism is the collision-induced heating that we call Joule heating, and we will devote a whole section to it later in this chapter. For now Kipute can trade her worry for a corrected intuition: a warm wire is not a wire full of fast electrons. **It is a wire full of slow electrons colliding furiously.**

Making Sense of the Answer: *Both halves of the question concluded with the same moral: the carrier speeds inside an ordinary household wire are absurdly slow, and the macroscopic things the wire seems to be doing (lighting a room instantly when the switch closes, warming after long use) have nothing to do with electrons moving fast. The chapter opener settled the lighting; this example settles the warming. In each* *case the answer is the same: the wire was already full of carriers; the electric field set them all drifting simultaneously; their slow drift gets dissipated into heat by collisions with the lattice. Slow electrons, fast field, immediate response, slow heating. Four pieces, one story.*

Think Like a Physicist: *Whenever a question hands you a power rating and a supply voltage instead of a current, the first equation to reach for is I = P/V. From there, the drift velocity is exactly one substitution away. Power-voltage data hides the current; the current hides the drift; and the drift is usually what the question is really after. Three steps, not one.*

HOT Example 5

A steady current passes through a cross-section of a conducting wire. Over a period of 30 seconds, 7.5×10^{19} electrons cross the cross-section. Find:

(a) the total charge that has crossed the cross-section in the 30s,
(b) the magnitude of the current in the wire,
(c) explain briefly why a lamp connected in series with this wire would light up the instant the switch is closed, even though the individual electrons in the wire travel at drift velocities of only millimetres per second.

Solution

(a) The total charge is given by:

$$Q = Ne = (7.5 \times 10^{19}) \times (1.6 \times 10^{-19}\mathrm{C}) = 12\mathrm{C}$$

(b) The current is the charge that crossed per unit time:

$$I = \frac{Q}{t} = \frac{12\mathrm{C}}{30\mathrm{s}} = 0.4\mathrm{A}$$

(c) The electrons that are about to enter the lamp's filament were already there before the switch closed; they did not have to travel from the source to the filament. When the switch closes, the electric field fills the wire at the speed of light, telling every electron, including those already inside the filament, to start drifting at the same instant, so the filament heats and lights immediately because the electrons inside it began drifting immediately, not because an electron rushed across the room to reach it.

Making Sense of the Answer: *Twelve coulombs in thirty seconds gives 0.4A, roughly the steady current drawn by a small household lamp. The pedagogical part (c) is the real point of the question. The lamp's response is set by the speed of the field, not by the speed of the carriers, and that point separates students who understand current from students who only know how to compute it.*

Think Like a Physicist: *In any question that gives you a number of electrons over a time interval, the conversion is always Q = Ne and I = Q/t. The two facts merge into one statement: I times t gives the count N = It/e of electrons that have crossed. Many questions in this chapter can be unlocked by spotting which of these three quantities (Q, I, or N) is the implicit one in the question.*

HOT Example 6

A current of 0.5A flows through a 40W lamp. Calculate the number of electrons that flow through the filament of the lamp in one second.

Solution

The number of electrons that carried this charge is:

$$N = \frac{Q}{e} = \frac{It}{e}$$

Substituting:

$$N = \frac{(0.5\mathrm{A})(1\mathrm{s})}{1.6 \times 10^{-19}\mathrm{C}} = 3.125 \times 10^{18}\ \text{electrons}$$

Making Sense of the Answer: *This is more than three quintillion of electrons. The filament has three quintillion electrons entering and three quintillion leaving every second, and the difference between entering and leaving is exactly zero by charge conservation. The current is not the electrons going in (or out); it is the rate at which they cross.*

Think Like a Physicist: *When a question gives you both a current and a power rating, ask which of the two the question actually needs. Examiners regularly hand students decorative data to test whether they recognise the path to the answer. If the question asks for electrons per second, the answer depends only on I and t; the wattage is bait.*

Six examples done. The numbers should now sit comfortably in mind. Drift velocity in tenths of a millimetre every second. Current density in millions of amperes every square metre. Electron count past any cross-section in quintillions every second. The crawl is mathematically respectable; if anything, it is too slow to be respectable. The next question is what makes one wire crawl easier than another, why a copper wire does the job better than a steel one, and why a hot wire does it slightly worse than a cold one. That is Ohm's law, and the next section walks straight to it.

OHM'S LAW FROM MICROSCOPIC THEORY

The last section answered how a current I flows through a wire once the carriers are drifting. It left one question hanging, the one any honest student must ask next: *what makes the carriers drift in the first place, and why does one wire pass more current than another?* Two answers are on the table. The first is the answer Georg Simon Ohm wrung out of his apparatus in 1827, after wrapping platinum wires of different lengths around wooden frames and passing currents through them from a thermopile. It is empirical, brief, and astonishing in its simplicity. The second is the answer the microscopic picture of the previous section quietly built for us without our noticing. The two answers agree, and the agreement is the content of this section.

Ohm's law as an empirical regularity

Take a copper wire of fixed length and thickness. Connect it across a variable voltage source. Keep the wire in a beaker of cool water so its temperature does not change as current passes through it. Slowly raise the voltage from zero, and at each setting measure the current the wire carries. Plot the current I against the voltage V. The result is a straight line that passes through the origin and extends with the same slope into the negative-voltage quadrant if the source is reversed.

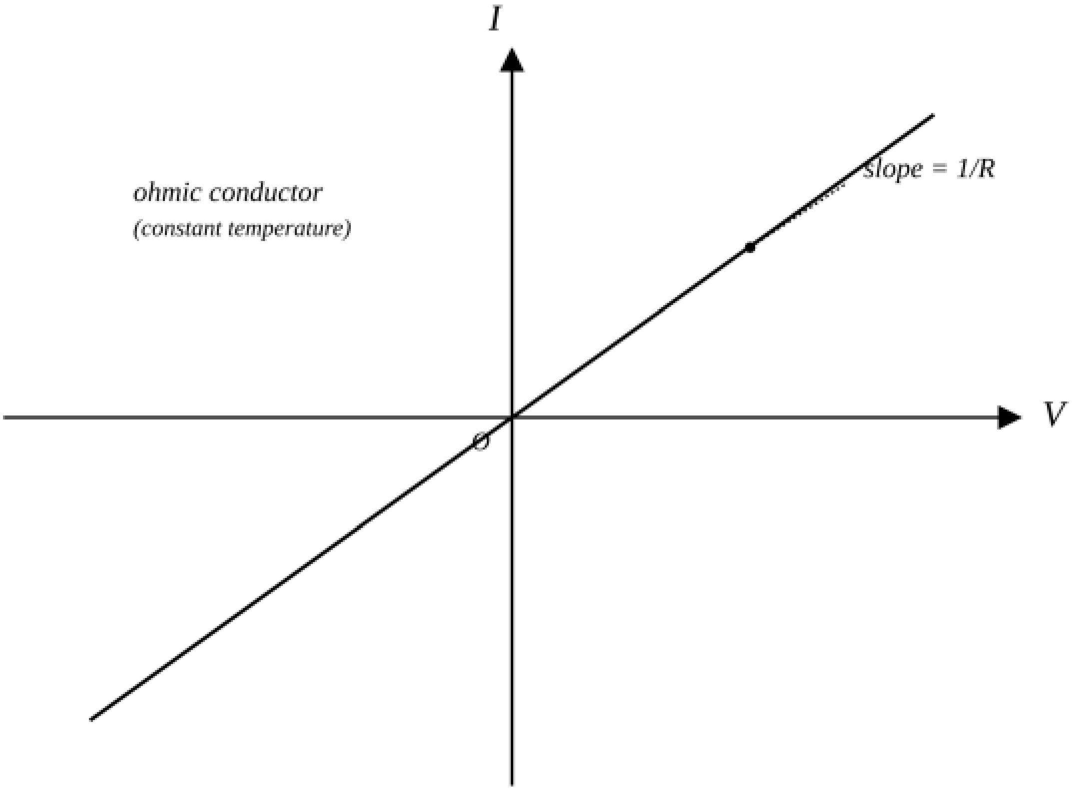


Figure: ***Current–voltage characteristic of an ohmic conductor at constant temperature.*** *The graph is a straight line through the origin in both polarities, and the reciprocal of its slope is the resistance of the conductor.*

Ohm's law states that:

The current that flows through a conductor between two points is directly proportional to the potential difference applied across the two points, provided physical conditions, especially the temperature, remain constant.

Mathematically:

$$\mathrm{V} \propto \mathrm{I}$$

Calling the constant of proportionality R, Ohm's law becomes:

$$\mathrm{V} = \mathrm{IR} \text{ or } \mathbf{I} = \frac{\mathbf{V}}{\mathbf{R}}$$

The constant **R** is called the **resistance** of the conductor, and Ohm's law itself defines it: the **resistance** of a piece of wire *is the voltage across the wire divided by the current through it, when the wire is held at constant temperature*. In symbols:

$$\mathrm{R} = \frac{\mathrm{V}}{\mathrm{I}}$$

The unit of resistance is the volt per ampere, given the special name ohm with the symbol Ω:

$$1\Omega = 1\mathrm{V/A}$$

A wire of resistance 1Ω carries one ampere of current under a voltage of one volt across its ends. A wire of resistance 10Ω carries only one tenth as much current under the same voltage. A wire of resistance 0.1Ω carries ten times as much. The resistance is the wire's individual contribution to the question of how much current the voltage pushes through it.

Two important warnings about the law itself before we proceed.

- The law is not a law of nature in the sense that Newton's laws of motion are. It is an empirical regularity that happens to be true for some materials and false for others. *Materials that obey it* are called **ohmic**. Most metals at fixed temperature are ohmic. Copper sulphate solution is ohmic. Carbon at fixed temperature is approximately ohmic. *The conductors that disobey it* are called **non-ohmic.** The filament of an incandescent lamp at varying current is non-ohmic, because the filament heats as the current rises, and its resistance changes with temperature. The junction of a semiconductor diode is strongly non-ohmic, because it lets current through easily in one polarity and almost not at all in the other.

We will sketch the non-ohmic curves at the end of the next section, after we have understood why temperature changes the resistance. For now, every conductor in the worked examples that follow is ohmic, and Ohm's law holds in the simple form just given.

- Ohm announced his result as an empirical regularity, true in the laboratory but lacking any deeper explanation. For the better part of seventy years the physics community took the law at face value, valid where the experiments said it was valid and silent on the question of why. The explanation arrived only at the very end of the nineteenth century, with the discovery of the electron and the construction of the microscopic model we developed in the previous section.

We will now follow that explanation through, in four steps. At the end of the four steps Ohm's law will reappear, not as an empirical truth but as a consequence of more elementary physics.

The microscopic question

From the previous section we have an exact relation between the current in a wire and the drift velocity of its free electrons:

$$\mathrm{I} = \mathrm{neAv_d}$$

Three of the four factors on the right are properties of the wire that do not depend on whether a voltage has been applied: the number density of free electrons n, the electronic charge e, and the cross-sectional area A. The fourth factor, the drift velocity $\mathrm{v_d}$, does depend on the applied voltage; it is in fact the only place the voltage can enter the expression. To understand how I depends on V, we therefore need to understand how the drift velocity depends on the applied electric field. The next two steps build that relationship.

An electron between collisions

Imagine a single free electron deep inside the wire, mid-flight between two collisions with the lattice. The applied electric field E points along the wire. The force on the electron is the field times the electron's charge, but because the electron is negative, the force points opposite to the field. In symbols (taking e for the magnitude of the electronic charge, so $e > 0$):

$$F = eE$$

where the equality is taken in magnitude, and the force points opposite to the field.

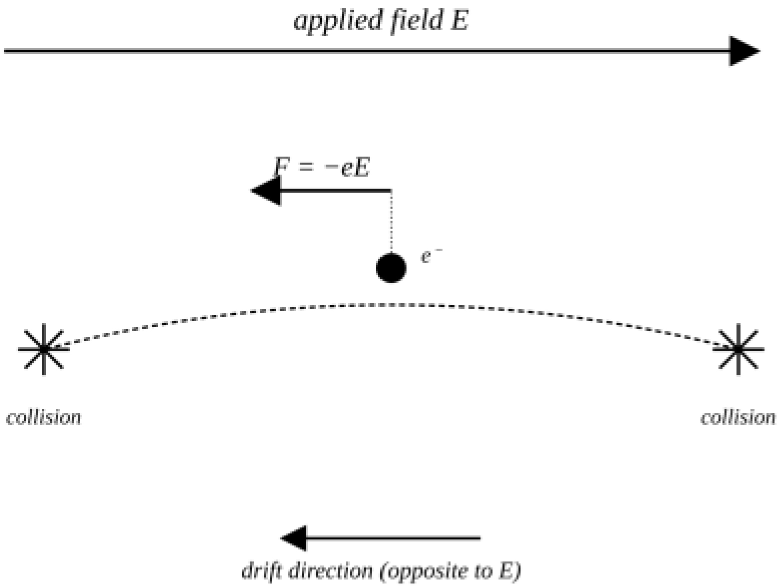


Figure: ***One free electron during a single flight between two collisions with the lattice****. The applied field points to the right, the force on the electron points to the left, and the electron picks up a small drift velocity to the left during the free flight before the next collision randomises its velocity.*

Step one: the acceleration. By Newton's second law, the acceleration of the electron during the free flight is the force divided by the electron's mass m:

$$a = \frac{F}{m} = \frac{eE}{m}$$

The acceleration points opposite to the field. The negative sign that this would carry in a vector treatment has been absorbed into our convention of always working with magnitudes when discussing drift; the direction is built into the conventional current picture and is not in dispute.

Step two: the relaxation time. If the electron were free to accelerate forever, it would reach speeds approaching the speed of light in a fraction of a second under any electric field of laboratory strength. It does not, because every so often it smashes into a lattice ion (or an impurity, or a vibrating lattice defect) and loses all memory of the drift it had just acquired. The average time between two such collisions is called the **relaxation time** of the metal, written $\boldsymbol{\tau}$. Starting from **zero** drift just after one collision and accelerating for an average time τ before the next, the electron acquires an average drift velocity:

$$v_d = a\tau = \frac{eE\tau}{m}$$

The drift velocity is the acceleration multiplied by the average time the acceleration is allowed to operate.

Putting it together

Step three: substitute into $I = neAv_d$. Replacing the drift velocity in the previous section's expression by the microscopic formula just derived:

$$\mathrm{I} = \mathrm{neA} \times \frac{\mathrm{eE\tau}}{\mathrm{m}} = \frac{\mathrm{ne^2A\tau}}{\mathrm{m}} \times \mathrm{E}$$

For a uniform conductor of length l carrying a steady current, the field inside it is the voltage across it divided by its length:

$$\mathrm{E} = \frac{\mathrm{V}}{\mathrm{l}}$$

Substituting once more:

$$\mathbf{I} = \mathrm{neA} \times \frac{\mathrm{eE\tau}}{\mathrm{m}} = \left(\frac{\mathrm{ne^2A\tau}}{\mathrm{ml}}\right) \times \mathbf{V}$$

Step four: read off the resistance. The current is now proportional to the voltage. The constant of proportionality is a single bracketed expression involving only properties of the material $(\mathrm{n, e, m, \tau})$ and of the geometry $(\mathrm{l, A})$. Comparing with the empirical Ohm's law $\mathrm{I} = \frac{\mathrm{V}}{\mathrm{R}}$, the resistance must be the reciprocal of that bracketed expression:

$$\mathbf{R} = \frac{\mathbf{ml}}{\mathbf{ne^2A\tau}}$$

This is the microscopic expression for the resistance of a wire. Notice that it has been built from quantities every one of which we have already met: the electron's mass, the length and cross-sectional area of the wire, the number density of free electrons in the material, the electronic charge, and the relaxation time. Nothing in the formula refers to V or I; *the resistance is a property of the wire by itself,* ***not*** *of the experiment we choose to do with it.* That is exactly what Ohm's law required.

Reading the microscopic formula

The microscopic resistance formula separates cleanly into two parts.

On the one hand it depends on the **geometry** of the wire through the length and the cross-sectional area. **A longer wire has more resistance**, in direct proportion to its length, *because the electron must traverse more lattice before reaching the other end.* **A thicker wire has less resistance**, in inverse proportion to its area, *because the same current can be carried by more electrons* each drifting more slowly.

On the other hand the formula depends on the **material** of the wire through four atomic constants: the free-electron number density, the electronic charge, the electron's mass, and the relaxation time. The first three are the same for any metal that we happen to consider, more or less; copper, silver, aluminium, and iron all have free-electron densities of order 10^{28} to 10^{29} per cubic metre. The fourth, the relaxation time, varies wildly from material to material and is the property that distinguishes good conductors from poor ones.

It is conventional to separate the material part of the resistance from the geometric part by defining the **resistivity ρ**:

$$\boldsymbol{\rho} = \frac{\mathbf{m}}{\mathbf{ne^2\tau}}$$

in terms of which the resistance of a wire is:

$$\mathbf{R} = \frac{\boldsymbol{\rho}\mathbf{l}}{\mathbf{A}}$$

The resistivity has units of Ωm and is a *property of the material alone*, independent of how the material is shaped into a wire. The resistance is the geometry multiplying the resistivity. The resistivity is the material; the resistance is the material wrapped in a particular shape. We will return to the resistivity in the next section, where we tabulate its values for the materials that appear in this book and study how it varies with temperature.

One more reading. We can now say precisely what makes a conductor ohmic. **A conductor is ohmic** *when τ depends on the temperature of the lattice but not on the field or the voltage applied across the conductor.* When that is so, the right-hand side of the microscopic formula is independent of V, and R is a constant; the current is then strictly proportional to the voltage. A conductor is non-ohmic when τ depends on the field as well, either because the field is so strong that it begins to heat the conductor (as in an incandescent lamp) or because the conductor's structure responds nonlinearly to the field (as in a semiconductor diode). Both effects produce a voltage-dependent resistance, and the I-V curve curves away from a straight line. The microscopic picture explains in one stroke why the ohmic class is so wide (most metals at fixed temperature, where τ is set by the lattice and not by the field) and why the non-ohmic class is so important (every device that does something interesting with current, from a transistor to a filament lamp).

Conductance

Conductance *is the reciprocal of resistance:*

$$G = \frac{1}{R}$$

Its unit is the siemens (S), defined as the reciprocal of the ohm: $1S = 1/\Omega = 1A/V$. Where **resistance answers the question** *"how much voltage do you need to push a given current through me*", **conductance answers** *"how much current do you get out of me for a given voltage"*. Both descriptions are equivalent; the choice is stylistic. Most of this book uses R, but conductance occasionally simplifies the algebra in parallel networks, where the total conductance is the sum of the individual conductances of the branches. We will meet that simplification later in the chapter, when we come to circuit networks.

Ohm's law is not Joule's law

A common student error is to merge Ohm's law with Joule's law into a single rule. They are different laws about different things. Ohm's law $I = \frac{V}{R}$ tells you how the current through a conductor relates to the voltage across it. **Joule's law:**

$$\mathbf{P = I^2R}$$

tells you *the rate at which the conductor dissipates electrical energy as heat*. Ohm's law is empirical and approximate; **Joule's law** is **exact** for any conductor, ohmic or not, *because it follows directly from the definition of work done against a potential difference.* A filament lamp is non-ohmic; Ohm's law fails for it; $P = I^2R$ still holds. The full treatment of Joule heating waits until later in the chapter; for now the only point is that the two laws are not the same.

Three examples follow.

BINDER Example 7

A 12V battery is connected across a resistor of resistance 4Ω. Find the current that flows through the resistor.

Solution

Direct application of Ohm's law:

$$I = \frac{V}{R} = \frac{12V}{4\Omega} = 3A$$

Making Sense of the Answer: *Three amperes from a 12V battery across a 4Ω resistor. A current of this size is what a small car interior light draws, a couple of orders of magnitude smaller than the starter motor (which pulls hundreds of amperes for a brief moment) and a couple of orders of magnitude larger than a transistor radio. Three amperes is also what a dry cell can comfortably deliver for a few minutes before its internal resistance starts to matter.*

Think Like a Physicist: *Whenever a question gives you a voltage and a resistance and asks for a current, the answer is one line of arithmetic. The trick is not to overthink it. Ohm's law is the shortest equation in the chapter and the most reliable.*

REAL Example 8

It is half past five in the evening in Morogoro. Kipanga's mother has just put the kettle on for tea, and within a minute the lamp on the ceiling dims to about half its usual brightness. A few seconds later the kettle's indicator light fades. Kipanga, who has been working through a physics holiday package at the kitchen table, looks up. The mains voltage has dropped; this is the third brownout this week.

Kipanga: *Mother, the kettle will still boil eventually, will it not? The voltage has only dropped, it has not been cut off entirely.*

Mother (half-smiling)**:** *Eventually, yes. But it will take longer. Ask the kettle why.*

Help Kipanga:

(a) to calculate the resistance of the kettle's heating element, given that it is rated 1500W and is designed to run on the 240V mains;
(b) to predict the current the kettle now draws when the brownout has dropped the mains voltage to 200V, treating the heating element as ohmic at constant temperature;
(c) to calculate the power the kettle now dissipates in the brownout, and explain to Kipanga, why his mother predicted that the water will boil but more slowly.

Solution

(a) From the rated power and voltage, the rated current is:

$$I = \frac{P}{V}$$

Applying Ohm's law in the form R = V/I:

$$R = \frac{V}{I} = \frac{V}{\frac{P}{V}} = \frac{V^2}{P} = \frac{(240V)^2}{1500W} = 38.4\Omega$$

(b) Treating the heating element as ohmic at constant temperature, the resistance does not change when the voltage changes. The new current is the new voltage divided by the same resistance:

$$I = \frac{V}{R} = \frac{200V}{38.4\Omega} = 5.21A$$

(c) The power dissipated by the heating element at the lower voltage is:

$$P = VI = 200V \times 5.21A = 1042W$$

This is about seventy percent of the kettle's rated 1500W output.

The kettle will still boil the water because it still supplies sufficient power to raise the temperature of the two litres to boiling point, but it does so more slowly since the time taken to boil is inversely proportional to the power; therefore, with reduced power, a longer time is required.

Making Sense of the Answer: *Drop the voltage to five sixths of its rated value and the current drops to five sixths of its rated value (because the kettle is ohmic), but the power drops to twenty-five over thirty-six, about seventy percent of its rated value. The power drops faster than the voltage because power is voltage multiplied by current, and both factors are dropping in proportion to the voltage. This is the general fact: ohmic conductors dissipate power as the square of the voltage applied across them.*

Think Like a Physicist: *Whenever a kettle, heater, geyser, immersion rod, or iron is fed a voltage different from its rating, expect three quantities to change in lockstep: current scales as the voltage, power scales as the square of the voltage, and time to boil or to heat scales inversely as the square of the voltage. The fastest sanity check on any such question is to ask what fraction of the rated voltage is being supplied; everything else follows by squaring or square-rooting that fraction.*

HOT Example 9

The free electrons in copper have an average relaxation time 2.5×10^{-14}s at room temperature. Taking the number density of free electrons as $8.5 \times 10^{28}\ m^{-3}$, the electron's charge as $e = 1.6 \times 10^{-19}$C, and the electron's mass as $m = 9.1 \times 10^{-31}$kg, calculate the resistivity of copper predicted by the microscopic model and compare it with the tabulated value of $1.72 \times 10^{-8}\ \Omega m$.

Solution

The microscopic formula for the resistivity is:

$$\rho = \frac{m}{ne^2\tau}$$

Substituting:

$$\rho = \frac{9.1 \times 10^{-31}\,\text{kg}}{(8.5 \times 10^{28}\,\text{m}^{-3}) \times (1.6 \times 10^{-19}\text{C})^2 \times (2.5 \times 10^{-14}\text{s})} = 1.67 \times 10^{-8}\,\Omega\text{m}$$

Compare with the tabulated value of 1.72 × 10^{-8} Ωm. The prediction is extremely close, differing by only about:

$$\left(\frac{1.72 \times 10^{-8}\,\Omega\text{m} - 1.67 \times 10^{-8}\,\Omega\text{m}}{1.72 \times 10^{-8}\,\Omega\text{m}}\right) \times 100\% = 2.9\%,$$

which shows that the microscopic model gives an excellent description of electrical conduction in copper.

Making Sense of the Answer: *About three percent agreement between a calculation built from four atomic-scale constants and a measurement taken in a copper wire at room temperature, with no fitting and no fudge factors, is one of the small triumphs of nineteenth-century physics. The free-electron model has its faults, particularly when temperatures get high or fields strong, but at room temperature in ordinary copper it is essentially right.*

Think Like a Physicist: *Whenever a derivation produces a number you can compare with a measurement, do the comparison. Numerical sanity checks are the fastest way to learn whether the microscopic model you have just spent a section building is worth believing. If a derivation predicts a quantity to within an order of magnitude of its measured value, the underlying physics is approximately right. If it predicts the quantity to within a few percent, the underlying physics is almost certainly right.*

Three examples done. Ohm's law has been stated, derived, applied to a battery and a resistor, applied to a kettle and a brownout, and finally checked against the resistivity of copper at the atomic scale. The microscopic formula has earned the chapter's trust. What remains is to follow it where it leads: *how does the resistivity differ between materials, why does copper conduct better than iron, and why does any metal's resistance rise as it heats up?* Those are the questions of the next section, and the answer in every case lives in the relaxation time.

RESISTIVITY, CONDUCTIVITY, AND TEMPERATURE DEPENDENCE OF RESISTANCE

The closing paragraph of the last section split the resistance of a wire into two halves and pointed at the second one. The first half belongs to the wire's geometry. The length and the cross-sectional area can be measured with a ruler and a calliper, and they are exactly the same numbers whether the wire is copper, steel, nichrome, or glass. The second half belongs to the wire's material. The number density of free electrons, the electronic charge, the electron's mass, and the relaxation time are properties of the substance the wire is made of, and no ruler will tell you their values. Copper conducts better than iron because of this second half. Silver conducts even better than copper, aluminium only slightly worse, and fused quartz very much worse. None of those facts depend on how thick or thin you have drawn the wire. They depend on what you have drawn the wire out of.

The cleanest way to handle this split is to write the material half down once, give it a name, and then forget the geometry until you need it again. The material half is called the resistivity of the substance, and the next several pages are about how to read it, where to find it, and how it changes when the wire is heated.

The resistivity of a substance

Imagine, first, a very fat and very short piece of the substance: a cube one metre on a side, with current entering through one face and leaving through the opposite face. The length of this block is one metre. Its cross-sectional area is one square metre. If you measured the resistance between the two faces, you would be measuring a quantity that depends only on what the block is made of. Lengthen the block to two metres, with the same cross-section, and the resistance doubles. Halve the cross-section, with the length still one metre, and the resistance doubles again. So the resistance of a block of given material is proportional to the length and inversely proportional to the cross-section. What the proportionality is, in ohms per metre divided by square metres, is a number that is set by the substance itself. That number is the **resistivity**, and we write it with the Greek letter **ρ**. It can be defined as *a property of a material that measures how strongly the material opposes the flow of electric current*

From the microscopic derivation of the last section, the resistance of a wire of length l and cross-sectional area A was:

$$R = \frac{ml}{ne^2A\tau}$$

Four of the six quantities on the right belong to the material (the electron's mass m, the number density of free electrons n, the electronic charge e, and the relaxation time τ); the other two belong to the geometry (the length l and the cross-sectional area A). Group the material quantities together by writing:

$$\rho = \frac{m}{ne^2\tau}$$

Then the resistance becomes simply:

$$\mathbf{R = \frac{\rho l}{A}}$$

The new quantity ρ is called the resistivity of the substance. Its unit follows directly from the formula by transposing for ρ:

$$\rho = \frac{RA}{l}, \qquad \text{unit: } \frac{\Omega m^2}{m} = \Omega m$$

The resistivity is therefore measured in ohm metres. A substance with a resistivity of 1Ωm would offer a resistance of one ohm to a block one metre long and one square metre in cross-section. In practice no ordinary conductor comes close to that; copper, the workhorse of household wiring, has a resistivity of about $1.7 \times 10^{-8}\Omega m$ (i.e. 1.7 nanoohm metres), ten billion times smaller. The unit is large because most useful conductors are very good at conducting.

Two things can now be said precisely. The resistance is what the wire offers; the resistivity is what the substance offers. The resistance depends on how the substance has been shaped into a wire; the resistivity does not. Two copper wires of different thicknesses have different resistances but the same resistivity. If you needed a single number to describe how good a conductor a substance is, independent of the shape it has been drawn into, the resistivity is that number.

Conductivity

When we want to say a substance is a good conductor, we often prefer a number that grows when the substance conducts well, rather than one that shrinks. The reciprocal of the resistivity does the job, and it is called the **conductivity**, written with the Greek letter **σ**:

$$\boldsymbol{\sigma = \frac{1}{\rho}}$$

Substituting the microscopic expression for ρ:

$$\boldsymbol{\sigma = \frac{ne^2\tau}{m}}$$

Conductivity *is a property of a material that measures how easily electric current flows through it.*

The unit of conductivity is the reciprocal of the unit of resistivity, which is the reciprocal ohm metre, more often called the siemens per metre and written S/m. A conductivity of one siemens per metre is rather poor; copper reaches roughly $6 \times 10^7 S/m$ (sixty million siemens per metre), which is the standard against which every other conductor is measured.

The conductivity gives Ohm's law its tidiest form. From the previous section we already had $J = \sigma E$ foreshadowed; we can derive it now in one line.

Starting from the resistance form $V = IR$ and substituting $V = El$ (for a uniform electric field along a wire of length l) and $R = \frac{\rho l}{A}$:

$$El = I\frac{\rho l}{A}, \quad \text{hence} \quad \frac{I}{A} = \frac{E}{\rho}, \quad \text{that is,}$$

$$\mathbf{J = \sigma E}$$

The current density at any point in any conductor equals the conductivity at that point multiplied by the electric field at that point. This is the local form of Ohm's law. It removes the geometry of the wire from the discussion altogether and is the form physicists prefer when working with materials whose cross-sections vary along the length, or whose fields vary in space. The integral form $V = IR$ is the same statement integrated over the length of a uniform conductor. Both versions of Ohm's law follow from the same microscopic picture; the difference is only which scale you choose to look at.

Reading the resistivity of common materials

The next figure tabulates the resistivity, at room temperature, of eleven materials that turn up in this book and in the engineering of any Tanzanian household. The axis is logarithmic because the values stretch across twenty-five orders of magnitude, from ten-billionths of an ohm metre at one end (silver) to ten million billion ohm metres at the other (fused quartz). No linear axis could show all eleven on one page.

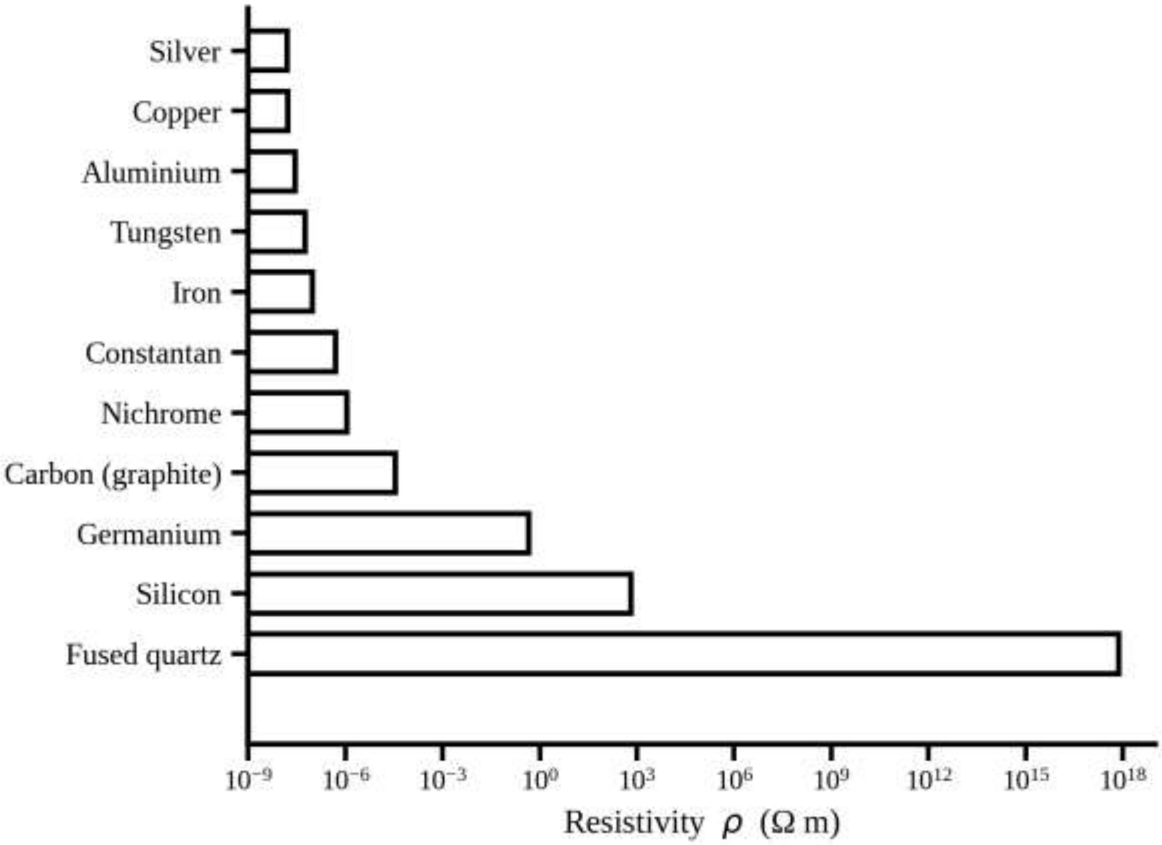


Figure: ***Resistivity of representative materials at 20°C, plotted on a logarithmic axis.*** *The eleven entries span twenty-five orders of magnitude. Silver, copper, aluminium, tungsten, and iron are the conductors. Constantan and nichrome are the designed-inconvenient alloys, used where resistance is wanted rather than avoided. Carbon graphite, germanium, and silicon are semiconductors, sitting between the conductors and the insulators. Fused quartz is an insulator.*

Four classes are visible at a glance.

First, the metallic conductors at the top of the chart (silver, copper, aluminium, tungsten, iron). Their resistivities lie within one order of magnitude of one another and all sit near $10 \times 10^{-9}\Omega$m (ten nanoohm metres). Silver is the best, but copper is cheaper and only seven percent worse, which is why almost every wire in your house is copper. Aluminium is sixty percent worse than copper, but it is also three times lighter, which is why long-distance transmission lines are aluminium and not copper. Tungsten is the metal of choice for filament lamps, not because it conducts best but because it stays solid at the high temperatures the filament must reach to glow visibly.

Second, the resistance-wire alloys (constantan, an alloy of copper and nickel; nichrome, an alloy of nickel and chromium). Their resistivities are about thirty to fifty times that of copper. This is not a flaw of the alloys; it is the whole point of using them. A laboratory rheostat or a kitchen kettle's heating element needs a wire whose resistance is several ohms or tens of ohms in a short length, rather than the thousandth of an ohm copper would give. Constantan and nichrome are designed to resist, not to conduct. They are also chosen because, as we shall see in a moment, their resistance changes very little when the temperature changes.

Third, the semiconductors (carbon graphite, germanium, silicon). Their resistivities lie between those of metals and those of insulators. Pure silicon at room temperature has resistivity about $640\Omega\text{m}$, fifty billion times that of copper. Doped silicon (the silicon out of which every chip in every phone is made) is many orders of magnitude better, but still nowhere near a metal. The middle of the chart belongs to the semiconductors, and the next book in this series devote a whole chapter to them.

Fourth, the insulators. Fused quartz is the example shown, with resistivity around $10^{17}\Omega\text{m}$. Glass, porcelain, dry wood, and most plastics are similar. Their resistivities are so much larger than those of the conductors that, for practical purposes, no current flows in them under any voltage a student is likely to apply to them. The whole of Chapter 11 quietly depended on this fact: the dielectric between the plates of a capacitor was assumed not to conduct at all, even when held at thousands of volts, and the chart shows why.

How resistance changes with temperature

Take any of the metal wires from the chart and warm it. Heat a copper coil from cool tap-water temperature to the temperature of boiling water and its resistance rises by about a quarter. Heat the tungsten filament of an incandescent lamp from the room temperature it sits at when the lamp is off to the 2500K it reaches when the lamp is on, and the resistance rises by a factor of about ten. Heating a metal does not just damage it; it also, and reliably, makes it conduct worse. The microscopic picture tells us at once why.

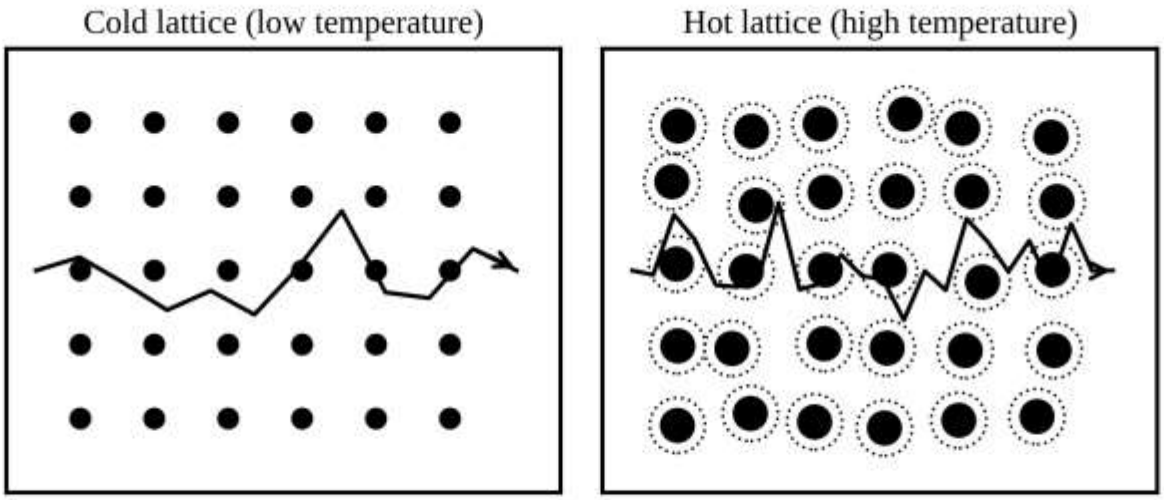


Figure: ***An electron threading through a metal lattice at two temperatures.*** *On the left, a cold lattice: the ions sit almost still and the electron makes its way past them with few collisions. On the right, a hot lattice: the ions vibrate vigorously about their lattice sites, sweeping out the dotted halos shown. The electron now has more frequent collisions, its relaxation time is shorter, and the wire's resistance is larger.*

From the last section, the resistivity is $\rho = \frac{m}{ne^2\tau}$. Of the four factors on the right, three of them (the electron's mass m, the electronic charge e, and the free-electron number density n) are essentially fixed for any given metal; they do not change when the metal is warmed. The fourth factor, the relaxation time τ, does change with temperature, and that is where the whole effect lives. The ions of the lattice vibrate more violently as the metal is heated, and an electron drifting through the lattice has a larger target to collide with at every site it passes. Collisions become more frequent, the relaxation time shortens, the resistivity rises, and (through $R = \frac{\rho l}{A}$) the resistance also rises.

If you do this experiment carefully over a moderate range, say from $0°\text{C}$ to a few hundred degrees, the rise of R with temperature is very close to linear. Write the temperature as θ (in degrees Celsius), call the resistance at the reference temperature θ_0 by the symbol R_0, and call *the fractional change in resistance per unit change in temperature* **coefficient of resistance**, written with the Greek letter $\boldsymbol{\alpha}$. Then, to a good first approximation:

$$\Delta R = \alpha R_0 \Delta\theta$$

Here ΔR is the change in resistance from R_0, and $\Delta\theta$ is the change in temperature from θ_0. The temperature coefficient α has units of per degree Celsius (equivalently per kelvin, since the **size** of a degree Celsius equals the size of a kelvin). For most pure metals α is small and positive, of order a few thousandths per degree Celsius.

Writing $R_\theta - R_0$ in place of ΔR, and $\theta - \theta_0$ in place of $\Delta\theta$, gives:

$$R_\theta - R_0 = \alpha R_0(\theta - \theta_0)$$

$$R_\theta = \alpha R_0(\theta - \theta_0) + R_\theta$$

Hence:

$$R_\theta = R_0(1 + \alpha(\theta - \theta_0))$$

With the reference temperature θ_0 of 0°C, the formula simplifies further to:

$$\mathbf{R_\theta = R_0(1 + \alpha\theta)}$$

Three things about it must be remembered together, because the formula loses its meaning if **any** of the three is forgotten.

First thing: The reference temperature is 0°C, not room temperature and not anything else. The value of α you read from a textbook table is the one that makes the formula above work with R_0 taken at exactly 0°C. Whenever a problem gives you 'the temperature coefficient of copper is $4.3 \times 10^{-3}\ {}^\circ C^{-1}$', that number is locked to the 0°C reference.

Second thing: The temperature θ in the formula is measured in degrees Celsius, not kelvin. The numerical value of α is the same in either scale (because the size of one degree Celsius equals the size of one kelvin), but the value of θ you substitute is the Celsius reading.

Third change: α has units of per degree Celsius (or equivalently per kelvin). For most pure metals α is small and positive, of order a few thousandths per degree Celsius, and a table of representative values is given in the next subsection.

There is one more form worth knowing.

In practice nobody can chill a wire to exactly 0°C just to read off R_0. Most laboratory experiments work at convenient temperatures well above zero. So we almost never use the strict formula $R = R_0(1 + \alpha\theta)$ directly. We use, instead, a working form derived from it. Apply the strict formula at two different temperatures θ_1 and θ_2 of the same wire and divide one equation by the other:

$$\frac{R_1}{R_2} = \frac{R_0(1 + \alpha\theta_1)}{R_0(1 + \alpha\theta_2)} = \frac{1 + \alpha\theta_1}{1 + \alpha\theta_2}$$

The R_0 at 0°C cancelled out completely. What remains is a clean relation between the two measured resistances, the two known temperatures, and the single material constant α. This is the ratio form of the temperature-coefficient relation, and it is the form that does almost all of the practical work in this section:

$$\mathbf{\frac{R_1}{R_2} = \frac{1 + \alpha\theta_1}{1 + \alpha\theta_2}}$$

Two facts about the ratio form are worth stating explicitly. First, it is mathematically exact (no approximation has been made beyond the original linear assumption); it is just the strict formula divided by itself at two temperatures. Second, the reference temperature 0°C is still present, but invisibly: the α inside the ratio form is the same α defined at 0°C, and the table values can be used directly. You never have to measure or know R_0.

A word of warning about the general form:

$$R(\theta_2) = R(\theta_1)[1 + \alpha(\theta_2 - \theta_1)]$$

This form looks like the strict formula with the reference moved from 0°C to a more convenient temperature θ_1. Be careful: it is not exact. It is an approximation that treats $R(\theta_1)$ as if it were R_0, which it is not (it is $R_0(1 + \alpha\theta_1)$). The approximation is good only when $\alpha\theta_1$ is much smaller than one. For copper at room temperature 20°C, $\alpha\theta_1 \approx 0.086$, which is not at all negligible, and the resulting error in the predicted resistance is several percent. Throughout this book we therefore use the ratio form (which is exact) and avoid the approximate general-reference form.

A graph of resistance versus temperature

From the exact formula:

$$R_\theta = R_0(1 + \alpha\theta),$$

$$\mathbf{R_\theta = R_0\alpha\theta + R_0}$$

So the graph of resistance R_θ versus temperature θ is a straight line with positive slope $R_0\alpha$ and y-intercept R_0.

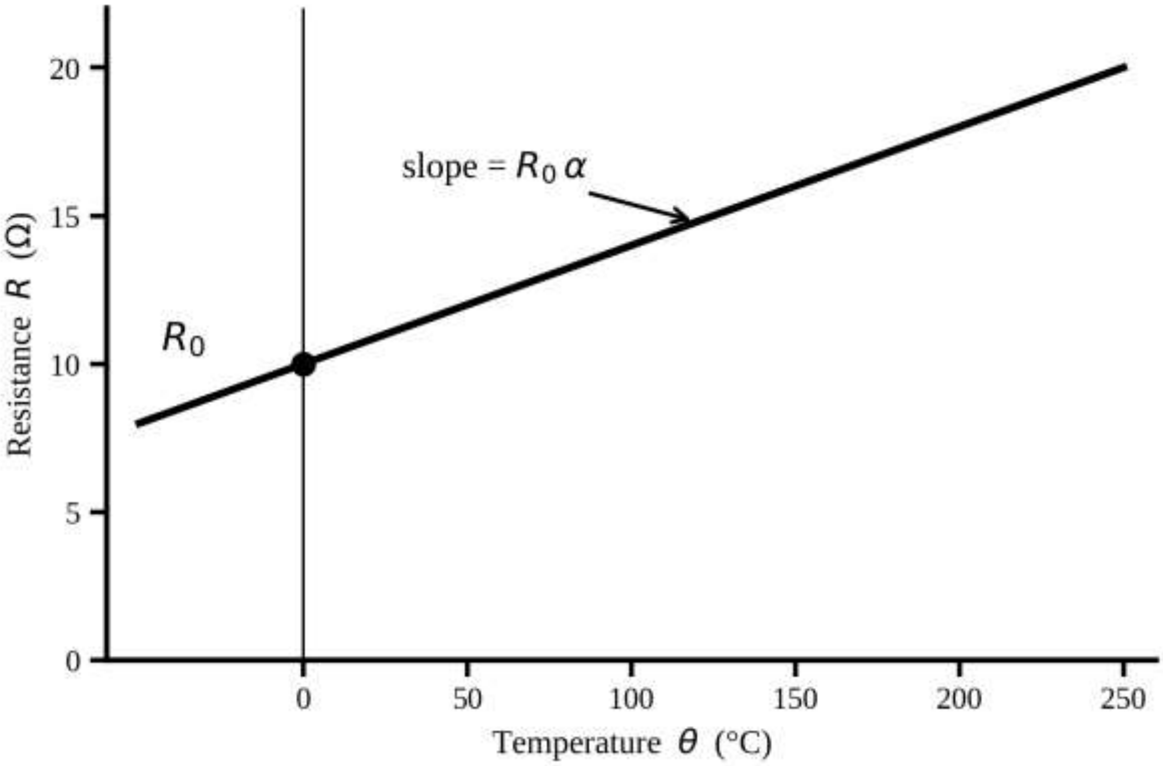


Figure: ***Resistance of a metal as a function of temperature.*** *The plot is a straight line over the range shown, with intercept R_0 at $\theta = 0°C$ and slope $R_0\alpha$. The graph would deviate from a straight line at very low temperatures (near absolute zero) and at very high temperatures (near the melting point); in the working range of any ordinary apparatus it is straight to within a fraction of a percent.*

Some representative values of α:

Material	**Temperature coefficient α at 0°C (per °C)**
Silver	4.0×10^{-3}
Copper	4.3×10^{-3}
Aluminium	4.0×10^{-3}
Tungsten	4.5×10^{-3}
Iron	5.0×10^{-3}
Constantan	4×10^{-4}
Nichrome	4×10^{-4}
Manganin	2×10^{-5}

Pure carbon (graphite)	-5×10^{-4}
Silicon (pure)	-75×10^{-3}

The table holds three lessons in plain view.

The first lesson: Pure metals (copper, aluminium, tungsten, iron, silver) have α values that cluster near 4×10^{-3} per degree Celsius. This is no coincidence; it tracks the way lattice scattering grows with temperature in a metal whose free-electron density is fixed and whose lattice ions vibrate more violently as the metal is warmed. Different metals have different ion masses and different lattice spacings, so the values are not identical, but they are all of the same order.

The second lesson: The alloys constantan and nichrome, and the better alloy manganin, have α values that are smaller than copper's by roughly a factor of ten to a thousand. Their resistance changes only a little when their temperature changes, which is exactly the property a standard laboratory resistance coil or a heating element needs. The alloys are designed for that purpose, and the design is reflected in their small α.

The third lesson: Not every α is positive. Pure carbon graphite and silicon both have negative temperature coefficients. Their resistance falls when they are heated. The microscopic reason in these substances is that the number density of carriers, n, is not constant with temperature as it is in metals. Heating a semiconductor frees more electrons into the conduction band, so n rises with temperature. The relaxation time τ still shortens with temperature, but the rise in n wins, and the net effect is that the resistance falls.

What α tells you about the metal

A large positive α (such as iron's, 5×10^{-3} per degree Celsius, or copper's, 4.3×10^{-3} per degree Celsius) means the lattice scatters drifting electrons strongly, and the strength of the scattering grows quickly with temperature. Copper between $0°C$ and $100°C$ raises its resistance by 43 percent; the same coil between $0°C$ and $50°C$ raises it by 22 percent. If you wanted to design an electric thermometer that worked by reading the resistance of a coil, copper would be a fine choice: its α is large enough that a modest temperature change gives a clearly measurable resistance change.

A small positive α (constantan's, nichrome's, or manganin's, of order 10^{-4} per degree Celsius) means the lattice scattering is nearly independent of temperature in the working range. Coils made of these alloys keep their resistance almost unchanged as their temperature changes. That is why every good-quality laboratory resistor in your physics laboratory is made of nichrome or constantan and not of copper, why the heating element of a kettle is nichrome, and why the same answers come out of the same resistance experiment whether the laboratory is cool at dawn or warm by mid-afternoon.

A negative α (carbon graphite, silicon, germanium) means the substance behaves opposite to a metal. Heat the substance, and its resistance falls. The microscopic mechanism is different: in these substances the number of carriers per cubic metre is itself a function of temperature, rising as the substance is heated, and that rise dominates everything else. We will meet the consequences again in the chapters that follow, where a gas at low pressure passes from being a perfect insulator to being a strong conductor once enough energy is poured into it to ionise its molecules.

How resistance changes with the geometry of the wire

We separated the resistance into a material part (ρ) and a geometric part $\left(\frac{l}{A}\right)$. The material part has now been fully discussed. The geometric part needs three short comments before we are done with this section.

Case one is the easy case. Hold the cross-sectional area fixed and change only the length. Then $R = \frac{\rho l}{A}$ and R is proportional to l. Doubling the length doubles the resistance. Halving the length halves it. Nothing surprising.

Case two is more delicate. Take a wire of length l and cross-section A, and stretch it (without breaking it) to a longer length while letting it thin out. The substance is incompressible, so the volume $V = Al$ is preserved as the wire is stretched. If the length becomes k times its original value, the cross-section must become $\frac{A}{k}$, so that the volume is still Al. The new resistance is:

$$\mathbf{R_{new}} = \frac{\rho(kl)}{\frac{A}{k}} = k^2\frac{\rho l}{A} = \mathbf{k^2R}$$

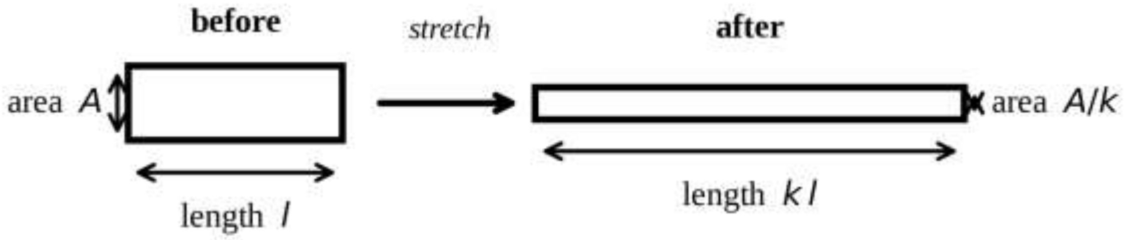


Figure: ***A wire stretched at constant volume.*** *The length grows from l to k l, the cross-sectional area shrinks from A to A divided by k, and the resistance grows by a factor of k squared, not just k.*

Stretching the wire to twice its length, at constant volume, multiplies the resistance by four, not by two. The stretching is doing two things to the resistance at once: *it is lengthening the path the electrons must travel,* ***and*** *it is narrowing the path they must travel through.* Both effects raise R by the same factor k, so the combined effect is $\mathbf{k^2}$. Forgetting the second of the two effects is the most common error students make with the stretched-wire question, and the worked Example 15 catches it directly.

Case three is fully general. The length and the cross-section are changed independently, perhaps because the wire is being redrawn through a die that gives it a different diameter as well as a different length. Then $R = \frac{\rho l}{A}$ reads off directly: if l becomes k times itself and A becomes j times itself, the new resistance is:

$$R_{new} = \frac{k}{j}R$$

with no constraint between k and j, because the volume is no longer being held fixed. The constant-volume case ($kj = 1$) recovers the stretched-wire result of case two, and the constant-area case ($j = 1$) recovers case one.

Non-ohmic conductors and their I-V characteristics

The last section made the point in passing, and the present section has now given us the tools to look properly at the non-ohmic conductors that section postponed. The figure below shows the current-voltage characteristics of four conductors on common axes.

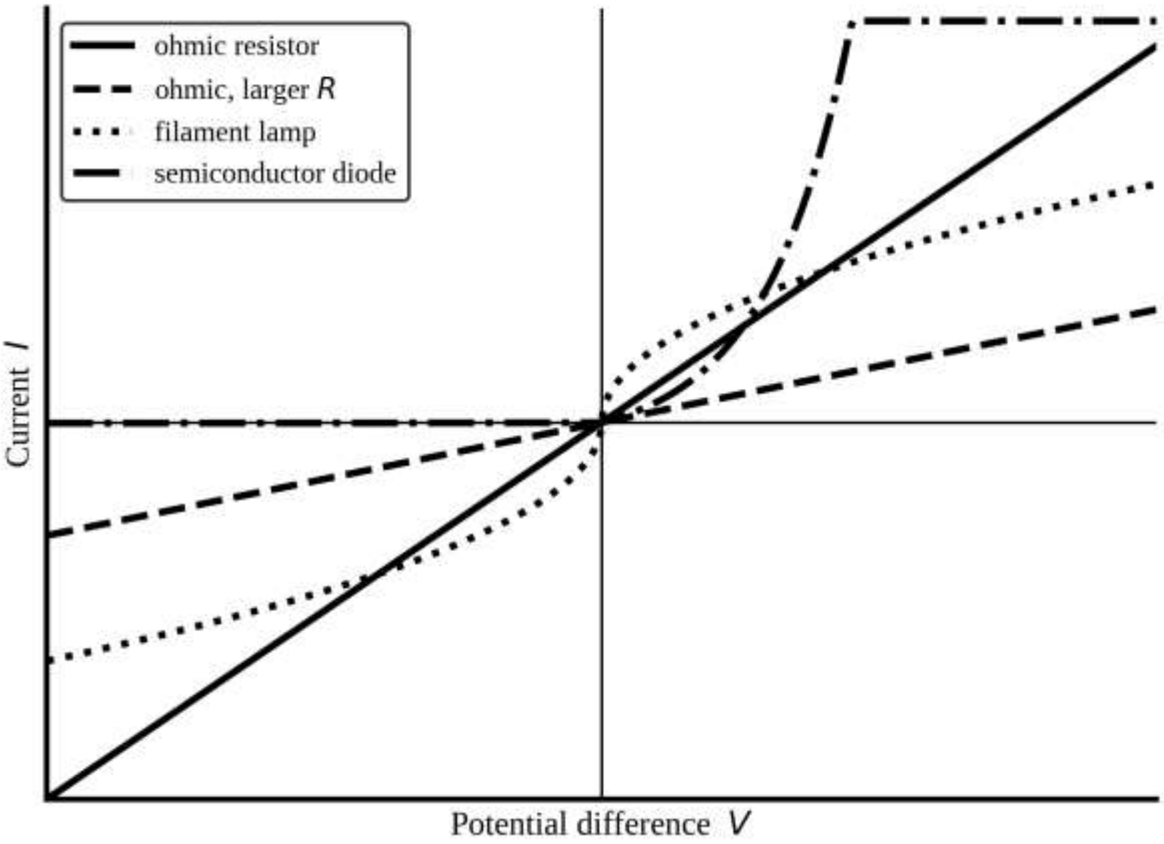


Figure: ***Current-voltage characteristics of four conductors.*** *Solid line: an ordinary ohmic resistor. Dashed line: a second ohmic conductor with larger resistance (smaller slope). Dotted line: a filament lamp, whose resistance rises as the filament heats with rising current. Dash-dot line: a semiconductor diode, which conducts easily in one polarity and almost not at all in the other.*

The first two conductors are ohmic, and their curves are straight lines through the origin with different slopes. The reciprocal of the slope is the resistance, so the steeper line belongs to the smaller resistance. Reversing the polarity of V reverses the polarity of I; the line is straight on both sides of the origin. Nothing here is new.

The dotted curve is the filament lamp. At very small V, the filament is cool, its resistance is small (about ten times smaller than its hot value), and the slope of the curve is correspondingly steep. As V is raised, the current rises, the I^2R heating raises the filament's temperature, the filament's resistance rises (because tungsten has a large positive α, 4.5×10^{-3} per degree Celsius), and the slope of the curve falls. By the time the filament is glowing at 2500K its resistance is roughly ten times the room-temperature value, and the curve has bent over noticeably. The filament is not disobeying Ohm's law at any moment; at any moment its resistance is well defined, and the current obeys $I = \frac{V}{R}$ at that resistance. But the resistance is changing as the filament heats up. The dotted curve is the trace of a moving Ohm's law, not the failure of one.

The dash-dot curve is the semiconductor diode. The diode's behaviour is asymmetric in polarity. With one polarity (forward bias) the diode passes current that rises exponentially with voltage; with the other (reverse bias) the diode passes a tiny saturation current that hardly grows at all until the diode is destroyed by breakdown at some large reverse voltage. The asymmetry has no analogue in any ohmic conductor; the device treats positive and negative voltages differently because its internal structure (the junction between two regions of doped silicon) breaks the symmetry. The physics behind the diode is the subject of a later chapter; for now the curve simply confirms that non-ohmic conduction is possible, common, and useful.

Seven examples now follow.

BINDER Example 10

A wire taken from a science-kit drawer at Miono Secondary School has length 10m and a uniform cross-sectional area of $1 \times 10^{-7}\text{m}^2$. A digital multimeter measures its resistance between the two ends as 50Ω at the laboratory temperature of $20°\text{C}$. Compute the resistivity of the wire material and identify the material by comparison with the bar chart of resistivities you read earlier in this section.

Solution

Take the geometric form of the resistance, $R = \frac{\rho l}{A}$, and rearrange it for the resistivity.

$$\rho = \frac{RA}{l}$$

Substituting the three measured values:

$$\rho = \frac{(50\Omega)(1 \times 10^{-7}\text{m}^2)}{(10\text{m})} = 5 \times 10^{-7}\Omega\text{m}$$

Compare this value with the bar chart of resistivities given earlier in this section. The closest entry is constantan, whose resistivity is $4.9 \times 10^{-7}\Omega\text{m}$ at $20°\text{C}$, identical to within the precision of the data. The wire is therefore constantan.

Making Sense of the Answer: *The number is large, by metal's standards. It is thirty times larger than copper's resistivity and ten times larger than aluminium's. That is exactly why constantan exists. A coil of constantan wire only ten metres long with a thin one-tenth-of-a-square-millimetre cross-section gives a useful resistance of* 50Ω*. The same coil made of copper, with the same length and the same cross-section, would give* 1.7Ω*, which is hardly noticeable in any circuit. The designed inconvenience of the alloy is what makes laboratory rheostats and standard resistors possible.*

Think Like a Physicist: *In identification problems of this kind, never compare the resistance of two objects. The resistance depends on the shape; two different samples of the same substance can have wildly different resistances. Always compute the resistivity first. The resistivity is a property of the substance, and that is what you compare against the table.*

BINDER Example 11

A short coil of metal wire has resistance 30Ω at $15°\text{C}$ and 36Ω at $65°\text{C}$. Find the temperature coefficient of resistance of the wire material, and use the value to suggest which of the metals in the table the wire is most likely made of.

Solution

The two resistance-temperature pairs lie on the same straight line $R = R_0(1 + \alpha\theta)$, but the value of R_0 is not given. The form to use is the ratio form, which has eliminated R_0 already:

$$\frac{R_1}{R_2} = \frac{1 + \alpha\theta_1}{1 + \alpha\theta_2}$$

Substitute $R_1 = 30\Omega$, $R_2 = 36\Omega$, $\theta_1 = 15°\text{C}$, $\theta_2 = 65°\text{C}$, and rearrange:

$$\frac{30}{36} = \frac{1 + 15\alpha}{1 + 65\alpha}$$

Solve for α:

$$\alpha = 4.26 \times 10^{-3}\ °\text{C}^{-1}$$

Compare this value with the table of α values given earlier in this section. The closest match is copper, whose temperature coefficient is 4.3×10^{-3} per degree Celsius. The wire is therefore very probably copper.

Making Sense of the Answer: *The answer is a single physical number with a clear physical meaning. An* α *of about* 4.3×10^{-3} *per degree Celsius says that every one-degree rise of the wire's temperature raises its resistance by 0.43 percent. Across the* $50°C$ *span of the experiment, the resistance rose by exactly 0.43 percent times 50, which is 21.5 percent. The actual rise was from* 30Ω *to* 36Ω*, which is 20 percent, and the small discrepancy is the result of taking* $15°C$ *rather than* $0°C$ *as the lower endpoint. The numbers are consistent with each other.*

Think Like a Physicist: *When you have two resistances at two temperatures and no third piece of information, the ratio form of the temperature coefficient formula is the form to use. It has already done the work of eliminating the unknown* R_0. *Trying to use the single-temperature form would force you to invent a value of* R_0, *propagate it through the calculation, and then (if you were lucky) cancel it at the end. The ratio form is doing exactly this cancelling for you, only quicker.*

BINDER Example 12

The Bagamoyo agricultural college is rewinding the field coil of an irrigation-pump motor, using bare copper wire. After the rewinding is complete and the motor is still cold, a workshop technician measures the resistance of the new coil at the morning temperature of 20°C and reads 45Ω. Under load on a hot afternoon the coil reaches a steady temperature of 75°C. Predict the resistance of the coil at that operating temperature. Take the temperature coefficient of copper to be $4.3 \times 10^{-3}{}^\circ\mathrm{C}^{-1}$.

Solution

Again the ratio form is the natural tool: it relates two resistances at two temperatures using only the table α (which is locked to the 0°C reference), and it never requires us to know the wire's resistance at 0°C:

$$\frac{\mathrm{R}_{75}}{\mathrm{R}_{20}} = \frac{1 + 75\alpha}{1 + 20\alpha}$$

$$\mathrm{R}_{75} = \left(\frac{1 + 75\alpha}{1 + 20\alpha}\right)\mathrm{R}_{20} = \left(\frac{1 + 75 \times 4.3 \times 10^{-3}}{1 + 20 \times 4.3 \times 10^{-3}}\right) \times 45\Omega = 54.8\Omega$$

Making Sense of the Answer: *The resistance of the operating coil is* 54.8Ω, *about 22 percent larger than the cold value of* 45Ω. *That is a substantial change, and it has real consequences for the pump. The current drawn from the supply (and hence the mechanical torque the coil can develop) is* $I = \frac{V}{R}$, *so the same supply voltage delivers about 18 percent less current to the hot coil than to the cold one (since* $1/1.218 \approx 0.82$*). Engineers who design winding-based motors must allow for this loss of grip when the motor is hot, by increasing the supply voltage or by accepting that the cold machine is the strong one.*

Think Like a Physicist: *The ratio form does two useful things at once. It eliminates the unknown* R_0 *(the resistance at* $0°C$*, which the technician would never measure in practice), and it keeps* α *honest with the table convention. The approximate general-reference form* $R(\theta_2) = R(\theta_1)[1 + \alpha(\theta_2 - \theta_1)]$ *would have given* 55.6Ω *here, an answer wrong by about 1.5 percent. Small, but real, and a sign that the wrong form was used. Whenever you find yourself with two temperatures and an* α *from a table, reach for the ratio form first.*

HOT Example 13

The nichrome heating element of a Mbeya bakery oven has resistance 20Ω at the cool temperature of 25°C. When the oven is pre-heated to 125°C the resistance of the element measures 22Ω. Find the temperature coefficient of the nichrome, and then predict the resistance of the element when the oven is operating at its full bread-baking temperature of 625°C.

Solution

The first step is to find α from the two known resistance-temperature pairs.

$$\frac{\mathrm{R}_1}{\mathrm{R}_2} = \frac{1 + \alpha\theta_1}{1 + \alpha\theta_2}$$

Substitute $\mathrm{R}_1 = 20\Omega$ at $\theta_1 = 25°\mathrm{C}$ and $\mathrm{R}_2 = 22\Omega$ at $\theta_2 = 125°\mathrm{C}$:

$$\frac{20}{22} = \frac{1 + 25\alpha}{1 + 125\alpha}$$

$$\alpha = 1.03 \times 10^{-3}\ {}^\circ\mathrm{C}^{-1}$$

The value $1.03 \times 10^{-3}{}^\circ\mathrm{C}^{-1}$ is roughly two to three times the value of pure nichrome listed in the table ($4 \times 10^{-4}\ {}^\circ\mathrm{C}^{-1}$). The discrepancy is characteristic of commercial heater wire, which is rarely pure nichrome but is usually a nichrome-based alloy with small amounts of other metals added to make it easier to draw and to weld. The value to use for this particular element, calibrated from this experiment, is the experimentally determined $1.03 \times 10^{-3}\ ^\circ C^{-1}$.

With α now known, apply the ratio form once more to find the resistance at 625°C. Pair the unknown R_{625} at 625°C with the known 20Ω at 25°C:

$$\frac{R_{625}}{R_{25}} = \frac{1 + 625\alpha}{1 + 25\alpha}$$

$$R_{625} = \left(\frac{1 + 625\alpha}{1 + 25\alpha}\right) R_{25} = \left(\frac{1 + 625 \times 1.03 \times 10^{-3}}{1 + 25 \times 1.03 \times 10^{-3}}\right) \times 20\Omega = 32\Omega$$

Making Sense of the Answer: *The element's resistance has risen by 60 percent between the cool oven and the hot oven. The current it draws from the mains, at constant voltage, has fallen by the same fraction (roughly speaking; in fact the mains voltage in Tanzania, $240V$, is held very steady by the grid). The element therefore self-regulates its temperature once it reaches operating heat: any further rise in temperature would raise the resistance, drop the current, and reduce the power dissipated, which would let the element cool back down. The Mbeya baker is benefiting from this self-regulation every time the oven holds a steady temperature without an explicit thermostat.*

Think Like a Physicist: *In any problem with three temperatures and two unknown quantities (here α and the third resistance), and no resistance value given at the reference temperature 0°C, the ratio form is the form to use. It eliminates R_0 cleanly in the calibration step (where two known pairs determine α), and it can be used again in the prediction step (where the now-known α and one of the original known pairs together determine the third resistance). Same equation, two substitutions, no fictitious reference value invented along the way. This is the lesson Example 11 made explicit, and it applies here word for word.*

HOT Example 14

A piece of copper wire has resistance 8Ω. The wire is melted down and redrawn through a die so that it emerges four times longer than before and three times thicker in diameter. Compute the resistance of the redrawn wire.

Solution

The wire is no longer the same wire; it has been remade. But it is still copper, so the resistivity ρ is unchanged. The length has been multiplied by $k = 4$, and the diameter has been multiplied by 3, which multiplies the cross-sectional area (proportional to the square of the diameter) by $j = 9$. From the geometric form of the resistance:

$$R = \frac{\rho l}{A}, \quad R_{new} = \frac{\rho(kl)}{jA} = \frac{k}{j}R$$

Substituting $k = 4$, $j = 9$, $R = 8\Omega$:

$$R_{new} = \frac{4}{9} \times 8\Omega = 3.56\Omega$$

Making Sense of the Answer: *The lengthening tries to raise the resistance by a factor of 4, but the thickening tries to lower it by a factor of 9, and the thickening wins. The redrawn wire has less than half the resistance of the original, even though the new wire is four times the length. Doubling the diameter of a wire is far more effective at lowering its resistance than is shortening the wire: the area depends on the square of the diameter, and the resistance depends inversely on the area.*

Think Like a Physicist: *Always work out separately what each geometric change does, then combine them. Length and area enter R differently: length in the numerator, area in the denominator. Mixing them up is the most common error in stretched-wire questions, and the remedy is the formula* $R = \frac{\rho l}{A}$ *written out in full, with every factor visible.*

HOT Example 15

A uniform wire of length 30cm has resistance 0.40Ω. The wire is then carefully stretched, without breaking, to a new length of 45cm. Throughout the stretching the wire keeps the same total volume of material. Compute the resistance of the stretched wire.

Solution

Two things happen to the wire when it is stretched at constant volume. The length grows; the cross-section shrinks. The ratio between the new length and the old length is:

$$\mathrm{k} = \frac{\mathrm{l_{new}}}{\mathrm{l_{old}}} = \frac{45\mathrm{cm}}{30\mathrm{cm}} = 1.5$$

Because the volume is preserved, the cross-section shrinks by exactly the same factor, so the new area is $\frac{\mathrm{A}}{\mathrm{k}}$. From the geometric form of the resistance:

$$\mathrm{R_{new}} = \frac{\rho(\mathrm{kl})}{\frac{\mathrm{A}}{\mathrm{k}}} = \mathrm{k}^2\frac{\rho \mathrm{l}}{\mathrm{A}} = \mathrm{k}^2\mathrm{R}$$

Substituting $\mathrm{k} = 1.5$ and $\mathrm{R} = 0.40\Omega$:

$$\mathrm{R_{new}} = (1.5)^2 \times (0.40\Omega) = 0.90\Omega$$

Making Sense of the Answer: *The resistance has more than doubled, even though the length has only grown by half. This is the classic stretched-wire result: at constant volume, a factor of* k *stretching of the length multiplies the resistance by* k^2*, not by* k*. Stretching the wire is doing two harmful things to the current path at the same time, lengthening it and narrowing it, and the two effects compound multiplicatively, not additively.*

Think Like a Physicist: *This is one of those problems that almost answers itself once you notice the constant-volume constraint. The whole answer lives in the substitution A becomes* $\frac{A}{k}$*. If the wire instead were stretched at constant cross-section (impossible in practice, but worth thinking about), the resistance would only have risen by a factor of 1.5. The compounding factor of* k^2 *is paid for by the conservation of metal.*

REAL Example 16

TANESCO is planning a 280km transmission line from Mwanza to Kahama. The engineering committee is choosing between two metals for the line, copper and aluminium. The line must have the same total resistance per kilometre regardless of which metal is chosen. For each metal, find the cross-section of the conductor required, and compare the total mass of the two candidate lines.

TANESCO is planning a 280km transmission line from Mwanza to Kahama. The engineering committee is choosing between copper and aluminium conductors. The line must be designed so that each conductor has a resistance of 0.10Ω per kilometre at $20°\mathrm{C}$.

(a) Determine the cross-sectional area required for each metal.
(b) Hence calculate the total mass of the complete 280km line for each metal.
(c) Compare the masses of the two lines and comment on which material is more practical for long-distance transmission.

Take the resistivities at $20°\mathrm{C}$ to be $1.7 \times 10^{-8}\Omega\mathrm{m}$ for copper and $2.8 \times 10^{-8}\Omega\mathrm{m}$ for aluminium.

Take the densities to be $8900\mathrm{kg/m^3}$ for copper and $2700\mathrm{kg/m^3}$ for aluminium.

Solution

First, find the cross-section required for each metal. The resistance of one kilometre of conductor is $\frac{\rho \mathrm{l}}{\mathrm{A}}$. Setting this equal to the specified $0.10\Omega/\mathrm{km}$ and rearranging:

$$\mathrm{A} = \frac{\rho \mathrm{l}}{\mathrm{R}}$$

For copper, with $\rho = 1.7 \times 10^{-8}\Omega\mathrm{m}, \mathrm{l} = 1000\mathrm{m}, \mathrm{R} = 0.10\Omega$:

$$\mathrm{A_{Cu}} = \frac{(1.7 \times 10^{-8}\Omega\mathrm{m})(1000\mathrm{m})}{(0.10\Omega)}$$

$$\mathrm{A_{Cu}} = 1.7 \times 10^{-4}\mathrm{m}^2$$

For aluminium, with $\rho = 2.8 \times 10^{-8}\Omega\mathrm{m}$, the same length, the same required resistance:

$$A_{Al} = \frac{(2.8 \times 10^{-8}\Omega\text{m})(1000\text{m})}{(0.10\Omega)}$$

$$A_{Al} = 2.8 \times 10^{-4}\text{m}^2$$

Now, for each metal, the mass of the full 280km line is given by:

$$m = \rho V = \rho AL$$

where ρ is the density of the metal and L is the total length of the line, 2.80×10^5m. For copper:

$$m_{Cu} = (8900\text{kg/m}^3)(1.7 \times 10^{-4}\text{m}^2)(2.80 \times 10^5\text{m}) = 4.24 \times 10^5\text{kg}$$

For aluminium:

$$m_{Al} = (2700\text{kg/m}^3)(2.8 \times 10^{-4}\text{m}^2)(2.80 \times 10^5\text{m}) = 2.12 \times 10^5\text{kg}$$

The mass ratio is therefore:

$$\frac{m_{Al}}{m_{Cu}} = \frac{2.12 \times 10^5}{4.24 \times 10^5} = 0.50$$

The aluminium line is half the mass of the copper line, even though it is two-thirds thicker. The mass saving wins on every count: every pylon must support a lighter cable, every footing must hold up a lighter pylon, and every truckload of wire delivered to the construction site costs half as much freight. Aluminium is therefore the metal TANESCO chooses for Mwanza-Kahama, as it does for every long-distance transmission line in the country.

Making Sense of the Answer: *The mass ratio of one-half is no accident. The ratio of the resistivities is roughly* $\frac{2.8}{1.7}$, *and the ratio of the densities is roughly* $\frac{2.7}{8.9}$. *The product of the two ratios is* $\frac{2.8 \times 2.7}{1.7 \times 8.9}$, *which works out to 0.50. The aluminium line is a little less efficient electrically (per cubic metre) than the copper line, but it compensates by being more than three times lighter (per cubic metre), and the compensation more than wins. Every transmission line in the world reflects the same compromise.*

Think Like a Physicist: *Engineering design is always a multi-criterion optimisation, and the right answer is the one that matters most for the application. For overhead transmission, mass and cost beat conductivity, and aluminium wins. For the wiring of a household, where the runs are short and the cross-section is constrained by the wall conduit, conductivity beats mass, and copper wins. The physics is the same in both cases (the formula* $R = \frac{\rho l}{A}$, *and the densities of the two metals); the choice changes with the context. A good physicist always asks not just for the answer but for the constraints that made the answer matter.*

Seven examples behind us. We have identified a science-kit wire by its resistivity, ambushed a coil into confessing its temperature coefficient, predicted what a Bagamoyo motor will do when the afternoon turns hot, watched a Mbeya oven calibrate itself by the very rise in resistance that the baker once cursed, made a copper wire shrink its own resistance by getting fatter faster than it got longer, made another wire double its resistance simply by agreeing to stretch, and walked a TANESCO committee through the single piece of arithmetic that decides whether 280km of transmission line are made of copper or of aluminium. None of those answers was a slogan; every one of them came out of a single formula, $R = \frac{\rho l}{A}$, combined with the way ρ itself responds to the temperature of the wire. The chapter so far has given you everything you need to read the resistance of any real conductor as a story about its material and its geometry, and to predict what that story will say next when the conductor is warmed, cooled, stretched, or remade.

With the resistance of the wire itself now thoroughly understood, one question remains, and it is the question we have been carefully sidestepping since the first paragraph of this chapter. Where does the current come from in the first place? Every example so far has begun with the words 'a battery of voltage V is connected to…', but no description of the battery itself has been offered. *What is inside a battery? What does it actually do to the charges in the wire? Why does the voltage across a battery's terminals fall a little when current is drawn from it, and rise back when current is stopped? Is there a hidden resistance inside every battery that no manufacturer advertises?* The next section answers all four questions, and the key idea is one that every dry cell, every car battery, every solar inverter, and every transmission-line generator have in common: **the electromotive force**.

ELECTROMOTIVE FORCE, INTERNAL RESISTANCE, AND TERMINAL VOLTAGE

The closing paragraph of the last section asked what is inside a battery, what it actually does to the charges flowing through a wire, and why the voltage across a battery's terminals falls a little when current is drawn from it. All three questions can be answered together with a single piece of vocabulary and a single short formula. The vocabulary comes first.

The electromotive force (EMF) defined

The electromotive force of a source *is the energy that the source supplies to every unit of charge that passes through it.* This is the plain-English definition, and it is worth saying twice in the two ways students most often need to hear it.

- Per coulomb of charge that the source pushes around the circuit, the EMF tells you how many joules of energy the source itself contributes.
- Equivalently, the EMF is the amount of work the source performs on each coulomb of charge in lifting that charge from its lower-potential terminal to its higher-potential terminal, ready to be sent into the external circuit.

A short analogy makes the definition vivid.

Imagine a battery as a tiny pump that lives inside its casing. The pump receives charges arriving at one of its terminals (the negative one), lifts each of them up through some internal step of energy, and releases them at the other terminal (the positive one) with more energy than they started with. The energy gained by each unit of charge during this lift is the **electromotive force** of the battery. The battery acts on charges in roughly the way a hand-pump acts on water: doing work on each unit that passes through, regardless of how much water (charge) passes through in total.

To turn the definition into a formula, suppose the source pumps a total charge Q through itself while doing total work W on those charges. By definition, the **EMF E** is the work done per unit charge:

$$\mathbf{E = \frac{W}{Q}}$$

The units of E follow at once: *joules per coulomb*, which is exactly the **volt.** (One volt is one joule per coulomb.) So when a manufacturer labels a battery 1.5V, the label is telling you that the battery supplies one and a half joules of energy to every coulomb of charge that passes through it.

The name electromotive force is a historical inconvenience and must not be taken literally. Despite the word 'force', the EMF is not a mechanical force. It is not measured in newtons. It has the units of energy per unit charge (joules per coulomb, i.e. volts), just like any other voltage. The word 'electromotive' captures the idea that the EMF is what sets charges into motion in the first place, but the word 'force' is a translation error from older European physics literature that has stuck for two centuries. Whenever you see EMF, mentally substitute 'driving voltage of the source'; that is closer to what it actually means.

Two things are worth saying about E before we go further.

First thing: *E is a property of the source alone, and not of the circuit the source is plugged into.* A AA battery and a car battery, both built from cells of the same chemistry, have nearly the same EMF per cell (about 1.5V for a dry cell, about 2V for each cell in a lead-acid car battery). What differs between them is the current the larger battery can supply, not the EMF. A car battery can deliver hundreds of amperes briefly; a torch cell would melt under the same load. But the energy delivered per coulomb is the same.

Second thing: *E is in general different from the voltage you would measure across the battery's terminals when the battery is in use.* That measured voltage has its own name and its own definition, given in a moment. The two voltages are equal only when no current is flowing through the battery. As soon as current flows, the measured voltage drops below the EMF by a small but definite amount. The reason has a name too, and that name is the second piece of vocabulary for this section.

Internal resistance demystified

The internal resistance of a source, written **r**, *is the total ohmic resistance offered by everything inside the source's casing to the current that the source drives.* The source is not an idealised pump operating in vacuum. Inside its casing live three real, physical things, and each of them resists the flow of current in the ordinary way.

First, the chemistry that does the lifting. Inside any wet cell, charge is carried partly by ions moving through a liquid or paste electrolyte, and the ions experience drag from the surrounding molecules of the electrolyte. That drag is, electrically speaking, an ohmic resistance. **Second,** the connectors that join the chemistry to the external terminals. These are made of metal, but the metal is neither infinitely thick nor perfectly conducting. **Third,** short lengths of wire inside the casing. Each of the three contributes its own small resistance. The sum of all three contributions is the internal resistance of the cell.

The internal resistance behaves like any other ohmic resistor. When the cell drives a current I through its own internal circuitry, it dissipates power I^2r as heat inside itself, exactly the formula for power dissipated in any resistor of resistance r carrying current I. That heat never reaches the external circuit; it warms the inside of the cell instead. If you have ever held a torch on for ten minutes and then felt the cells warm to the touch, you have felt this dissipation directly. The I^2r warming is the price the cell pays for being a real object rather than an idealised one.

By convention, *r is drawn as a small* ***zigzag*** *inside the rectangle that represents the cell*, rather than as a separate resistor outside. The physics is the same either way; the symbol is a shorthand for 'this resistance lives inside the casing'. When you solve a circuit, treat r as you would any other resistor, with the single understanding that it sits in **series** with the EMF source and inside the same casing as the source.

Terminal voltage and the V = E – Ir relation

The **terminal voltage** of a source, written **V**, *is the voltage that the external circuit sees across the source's two terminals while the source is in use.* Operationally, the terminal voltage is what a voltmeter would read if connected directly across the source's two terminals at the moment of measurement. The terminal voltage and the EMF are two different quantities, and one of the main tasks of this subsection is to derive the precise relationship between them.

Suppose the cell drives a steady current I through an external resistor R, as in the figure below.

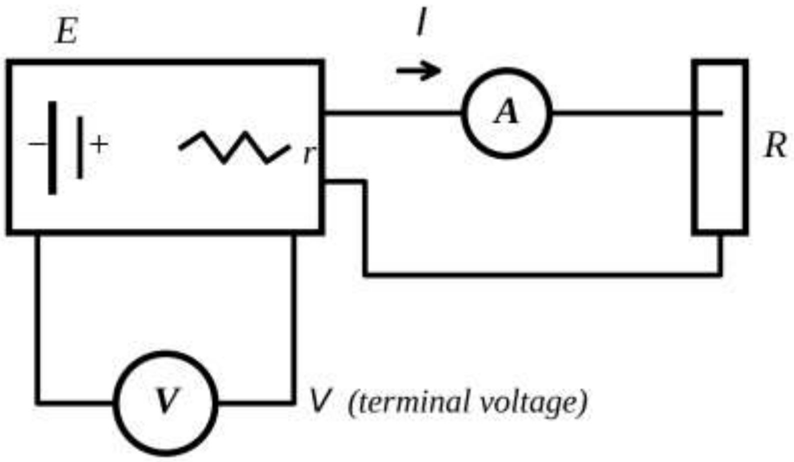


Figure: ***A battery of EMF E and internal resistance r driving a current I through an external resistor R****. The voltmeter, drawn below the battery, reads the terminal voltage V, which is what the external circuit sees. The ammeter in the top wire reads I.*

Trace one coulomb of charge round the circuit, and apply conservation of energy to it. Inside the cell, the chemistry does work E on the coulomb (raising it from the negative terminal up to the positive terminal). Some of that work, of magnitude Ir, is then immediately dissipated as the coulomb pushes through the internal resistance r on its way out. The remainder, E − Ir, is what the coulomb actually carries with it into the external circuit. By the operational definition above, that remainder is the terminal voltage of the cell:

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

This is the central equation of the section. It says that *the voltage available to the external circuit equals the EMF minus the drop across the internal resistance.* The external circuit never sees the EMF directly; what it sees is the EMF reduced by the internal-resistance loss.

A circuit is said to be in **open circuit** *when no current flows in it*, because the loop is broken somewhere (a switch is open, a wire is cut, or no external resistor is connected). In open circuit, I = 0, and the Ir loss in the central equation vanishes. The terminal voltage then equals the EMF exactly:

$$\mathbf{V = E \qquad (open\ circuit)}$$

Connect a voltmeter across an open battery, in other words, and the voltmeter reads the EMF directly. This is the cleanest way to measure E in the laboratory; the next figure shows how it is done.

More generally, when r is small and I is moderate, the Ir loss is small and V is only slightly less than E. A good-quality dry cell has r of the order of one-tenth of an ohm; drawing one ampere from such a cell drops the terminal voltage by only 0.1V below the 1.5V EMF, which is hardly noticeable. A worn or cheap cell has a much larger r, which is why an old torch goes dim long before its EMF has decayed.

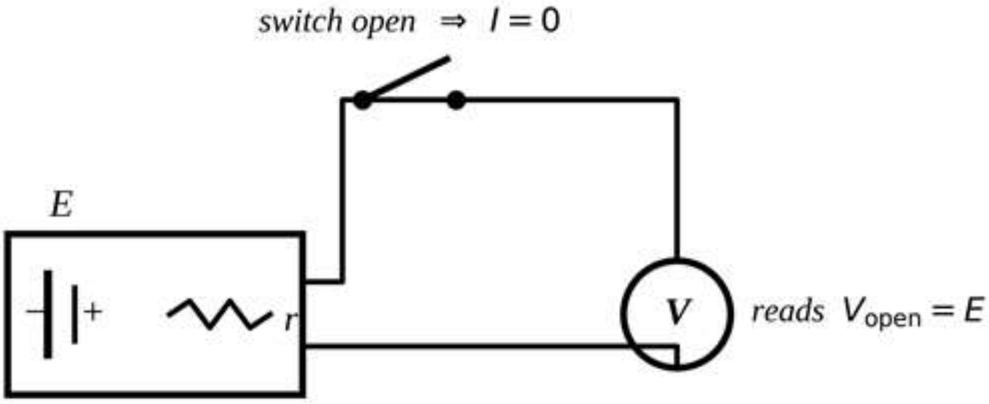


Figure: ***Measuring the EMF of a cell directly.*** *The switch in the top wire is open, so no current flows. The voltmeter then reads V-open, which equals E exactly. In practice the voltmeter must be of very high impedance (defined in the next subsection) so that the tiny current it draws is negligible.*

Plotting V against I for a real source gives a straight line, with two pieces of information on its face. The y-intercept is E (the value of V when I = 0), and the slope is −r (the terminal voltage falls by r for every ampere of current drawn). A single experiment that varies the load and records the resulting current and terminal voltage therefore reveals both E and r at once.

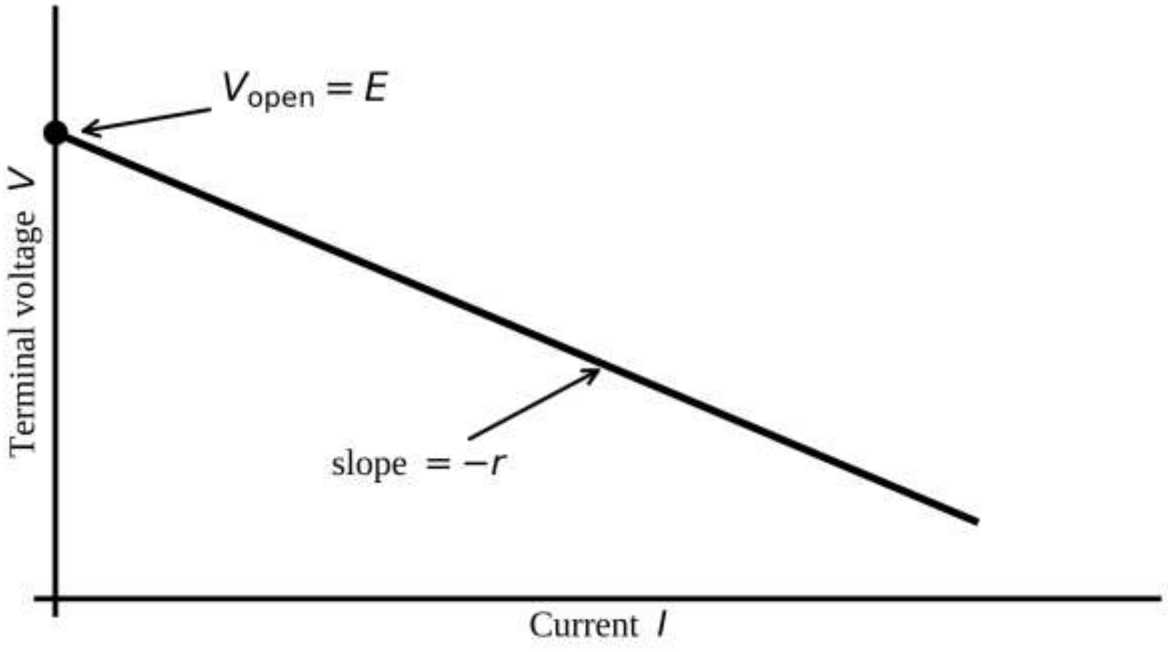


Figure: Terminal voltage as a function of the current drawn from a real source. The line is straight; the y-intercept is the EMF E; the negative slope has magnitude equal to the internal resistance r. The line crosses the I-axis at I = E/r, which is the short-circuit current.

The x-intercept of the line in the figure above, where $V = 0$, introduces another piece of vocabulary, short-circuit current. **The short-circuit current** of a source *is the current that would flow through it if its two terminals were directly joined by a wire of negligible resistance, so that the external resistance R is effectively zero.* Substituting $V = IR = 0$ into the central equation $V = E - Ir$ gives:

$$\mathbf{I_{short} = \frac{E}{r}}$$

The short-circuit current is determined by the EMF and the internal resistance alone, with no external resistance to limit it. It is the **largest current** the source can deliver, and for most sources it is large enough to be dangerous. A car battery (EMF 12V, r of order 0.01Ω) has a short-circuit current of about 1200A, enough to melt a spanner accidentally dropped across its terminals. Never short-circuit a real source deliberately.

Can the terminal voltage ever exceed the EMF? Yes, and the answer introduces one final piece of vocabulary, charging. A source is said to be **charging** *when another, larger source is connected to it in such a way that current is forced backwards through it, into its positive terminal and out of its negative terminal (the opposite of the direction it would drive on its own).* Charging is what happens to a car battery while the engine is running and the alternator is supplying current, or to a rechargeable phone cell while it is plugged into a wall charger.

During charging the current direction through the internal resistance is reversed, so the sign of the Ir term in the central equation reverses too. The terminal voltage of the source being charged is then:

$$\mathbf{V = E + Ir \qquad (during\ charging)}$$

The terminal voltage now sits above the EMF, by exactly the Ir drop across the internal resistance. This is the only way the terminal voltage of a real source can exceed its EMF; in any other operating regime, the terminal voltage is always less than or equal to E.

Measuring EMF and internal resistance

To measure EMF and internal resistance accurately, one piece of instrument vocabulary must come first, which is input impedance. **The input impedance** of a voltmeter *is the resistance that the voltmeter itself presents to the circuit it is measuring. (*In **direct-current** work the word ‘impedance’ means the **same** thing as ‘resistance’; we use the more general word to be consistent with the labels on commercial instruments, which often have alternating-current ranges where impedance and resistance differ.) A voltmeter's input impedance is important because, when the voltmeter is connected across a circuit element to measure its voltage, the voltmeter draws a small current of its own through its own input impedance, and that drawn current can disturb the reading.

A voltmeter is called high-impedance when its input impedance is so large compared with the resistances in the circuit being measured that the current it draws is negligible. For laboratory work, 'high' conventionally means at least a megaohm (a million ohms), and modern digital multimeters typically have input impedances of ten megaohm on their voltage ranges. A high-impedance voltmeter is the right instrument for measuring open-circuit EMFs because it disturbs the circuit it measures so little.

With that vocabulary in place, the cleanest way to measure the EMF of a source is the one already mentioned: connect a high-impedance (resistance) voltmeter across the source's terminals with no load attached. With $\mathrm{I = 0}$ the relation $\mathrm{V = E - Ir}$ collapses to $\mathrm{V = E}$, and the voltmeter reads the EMF directly.

The residual current drawn by a ten-megaohm voltmeter from a 1.5V cell is $\mathrm{I} = \frac{1.5\mathrm{V}}{1010^{6}\Omega} = 1.5 \times 10^{-7}\mathrm{A}$, which produces an Ir correction so small (a microvolt or so) that it cannot be detected on the voltmeter's own display. However, a more accurate method is the **potentiometer**, which we shall meet in a later section, and which draws no current at all at balance.

Measuring r requires a method known as the two-load method, which extracts both E and r from two pairs of (current, terminal-voltage) measurements taken at two different external loads. The method is built on the fact that the equation $\mathrm{V = E - Ir}$ contains two unknowns (E and r), so we need two independent equations to solve for them, and we obtain those two equations by taking two measurements at two different loads.

Connect the source through a known external resistance $\mathrm{R_1}$ and read the current $\mathrm{I_1}$ with an ammeter. The closed-circuit equation around the loop (EMF equals sum of IR drops) gives:

$$\mathrm{E = I_1(R_1 + r)}$$

Now change the external resistance to a different value $\mathrm{R_2}$ and read the new current $\mathrm{I_2}$:

$$\mathrm{E = I_2(R_2 + r)}$$

These are two equations in two unknowns (E and r). Solve them simultaneously by subtracting one from the other to eliminate E:

$$\mathrm{0 = I_2R_2 + I_2r - I_1R_1 - I_1r}$$

$$\mathrm{r(I_1 - I_2) = I_2R_2 - I_1R_1}$$

Hence:

$$\mathbf{r = \frac{I_2R_2 - I_1R_1}{(I_1 - I_2)}}$$

Substituting back into either of the originals then gives E.

Now we turn to worked examples.

BINDER Example 17

A battery of EMF 12V and internal resistance 0.5Ω drives a current of 2A through an external circuit. Find the terminal voltage of the battery while the current is flowing.

Solution

The terminal voltage of a real source carrying a current is given by V = E − Ir, with the minus sign because the current is flowing in the natural direction (out of the positive terminal).

$$\mathrm{V = E - Ir}$$

Substituting the given values:

$$\mathrm{V = 12V - (2A \times 0.5\Omega) = 11V}$$

Making Sense of the Answer: *The terminal voltage is $11V$, which is $1V$ less than the EMF. That missing volt is the price the battery pays for its own internal resistance: every coulomb of charge loses one joule of* *energy on the way out of the cell, which goes into* I^2r *heat inside the casing. The external circuit therefore sees* $11V$*, not* $12V$*, and any device powered by this battery will operate as though it were connected to an ideal* $11V$ *source.*

Think Like a Physicist: *Always carry units through the substitution. Volts minus amperes-times-ohms gives volts (because* $A \times \Omega = V$ *is what Ohm's law says), and the unit check is a free verification that the formula has been applied correctly.*

HOT Example 18

A voltmeter is connected in parallel with a variable resistor R, which is in series with an ammeter and a cell. For one value of R the meters read 0.3A and 0.9V. For another value of R the meters read 0.25A and 1.0V. Find the value of R in each case, the EMF of the cell, and the internal resistance of the cell. State the assumption made about the resistance of the meters in the calculation.

Solution

Begin with the assumption.

We assume that the ammeter has negligible resistance compared with R and r (so its presence in the loop does not affect the current that the cell drives), and that the voltmeter has very high impedance compared with R (so the current it draws is negligible compared with the current through R, and the voltmeter reads the same as if it were not there).

Under these assumptions, the voltmeter reads the voltage across R, which is also the terminal voltage of the cell (since the only other thing in the loop is the ammeter, whose voltage drop is negligible).

The terminal-voltage equation $\mathrm{V = E - Ir}$ applies to both meter readings.

For the first reading:

$$\mathrm{0.9V = E - (0.3A)r}$$

For the second reading:

$$\mathrm{1.0V = E - (0.25A)r}$$

Subtract the first equation from the second:

$$\mathrm{1.0V - 0.9V = (0.3A - 0.25A)r}$$

$$\mathrm{r = 2\Omega}$$

Substitute $\mathrm{r = 2\Omega}$ back into the first equation:

$$\mathrm{0.9V = E - (0.3A)(2\Omega)}$$

$$\mathrm{E = 0.9V + 0.6V = 1.5V}$$

The two external resistances follow from Ohm's law applied to R alone (each meter's reading on the voltmeter is the voltage across R, since the ammeter drop is negligible):

$$\mathrm{R_1 = \frac{0.9V}{0.3A} = 3\Omega}$$

$$\mathrm{R_2 = \frac{1.0V}{0.25A} = 4\Omega}$$

So the mystery cell has $\mathrm{E = 1.5V}$ and $\mathrm{r = 2\Omega}$, and the two external resistances are 3Ω and 4Ω.

Making Sense of the Answer: *All four numbers are physically sensible.* $E = 1.5V$ *matches a standard zinc-carbon dry cell. The internal resistance* $r = 2\Omega$ *is rather large, indicating either an old cell or a cheap one; the cell loses a substantial fraction of its EMF to internal dissipation as soon as any current is drawn. At* $I = 0.3A$ *the cell loses* $0.6V$ *inside itself and delivers only* $0.9V$ *to the external circuit, an efficiency of just sixty percent. A higher-quality cell with* $r = 0.1\Omega$ *would deliver* $1.47V$ *at the same current, an efficiency of ninety-eight percent. Battery quality and internal resistance are inversely related.*

Think Like a Physicist: *Whenever you have two measurements and two unknowns, the two-equation method is almost always the right approach. Write down the same physical equation for each measurement, subtract* *one from the other to eliminate one unknown, solve for the other, then substitute back. The strategy is mechanical, but it requires you to identify in advance which physical equation is the same for both measurements. Here that equation was $V = E - Ir$, and the two unknowns E and r are properties of the cell that are the same in both readings; only V and I change.*

HOT Example 19

A 9V battery is tested by a student. With no load, a high-impedance voltmeter reads the terminal voltage as 9.0V. The battery is then connected to an external resistor; the voltmeter reading drops to 7.5V while the ammeter reads 0.5A. Find (i) the internal resistance of the battery; (ii) the external resistance; (iii) comment on what this experiment reveals about the quality of the battery.

Solution

(i) With no load, means that the circuit is an open circuit.

The open-circuit reading is 9.0V with $\mathrm{I} = 0$, so the EMF is $\mathrm{E} = 9.0\mathrm{V}$.

With the external load connected, $\mathrm{V} = \mathrm{E} - \mathrm{Ir}$ gives r:

$$7.5\mathrm{V} = 9.0\mathrm{V} - (0.5\mathrm{A})\mathrm{r}$$

$$\mathrm{r} = 3.0\Omega$$

(ii) Under load the voltmeter reads the terminal voltage, which is the same as the voltage across the external resistor R. Apply Ohm's law to R alone:

$$\mathrm{R} = \frac{\mathrm{V}}{\mathrm{I}} = \frac{7.5\mathrm{V}}{0.5\mathrm{A}} = 15\Omega$$

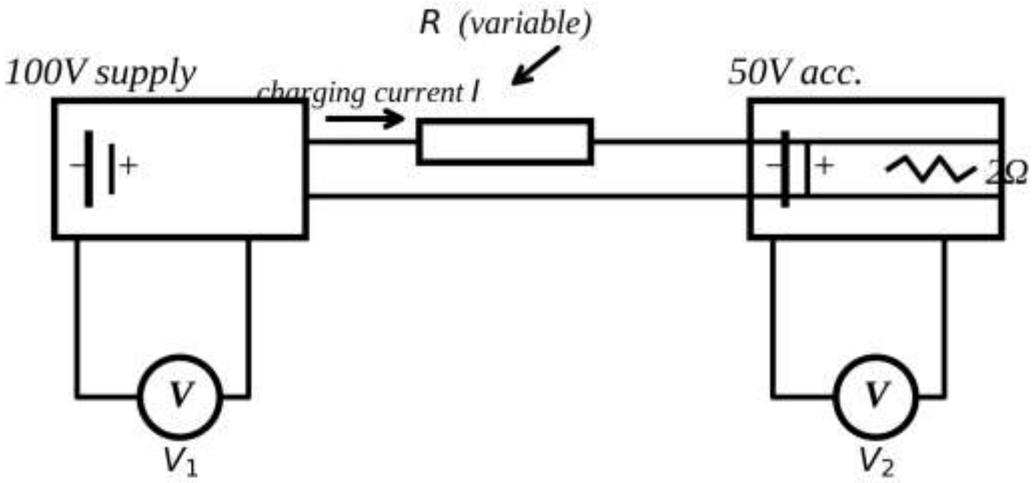
(iii) The experiment shows that the battery has a relatively high internal resistance, since its terminal voltage drops noticeably under load, meaning a significant amount of energy is lost inside the battery and it is not a very efficient or high-quality source for maintaining a steady voltage.

Making Sense of the Answer: *The internal resistance is $r = 3.0\Omega$, which is high for a $9V$ cell. Under the load tested, the battery delivers only $7.5V$ of its $9.0V$ EMF to the external circuit, which is $\frac{7.5}{9.0}$, or 83 percent efficiency. The remaining 17 percent of the available energy is dissipated inside the cell as heat. Battery quality scales inversely with r: the smaller r, the more of the EMF reaches the actual circuit.*

Think Like a Physicist: *The same battery, in different applications, can be 'good' or 'poor' depending on the load. A $9V$ smoke-detector battery with $r = 3\Omega$ might last years in service because a smoke detector draws only microamperes, at which current the Ir drop is negligible and the battery appears nearly ideal. The same battery in a $0.5A$ load shows its true weakness. Always evaluate r against the actual current drawn by the intended application; the number itself is meaningful only against a load context.*

REAL Example 20

An accumulator of EMF 50V and internal resistance 2Ω is being charged from a 100V DC supply, with a variable resistor R in series in the charging line. Two high-impedance voltmeters are connected, one across the supply terminals and one across the accumulator's terminals (see the figure below). Answer the following: (i) state the direction of the charging current; (ii) with $\mathrm{R} = 0$, calculate the charging current; (iii) calculate the terminal voltage of the accumulator under that condition; (iv) predict how each voltmeter reading changes as R is increased from zero.



Solution

(i) The supply is the more powerful source. It must force current into the positive terminal of the accumulator, against the direction the accumulator would drive on its own. The charging current therefore flows from the positive terminal of the supply, through the variable resistor R, into the positive terminal of the accumulator, out of the accumulator's negative terminal, and back to the negative terminal of the supply.

(ii) The cleanest way into this problem is to think of each component's terminal voltage separately. Three voltages are in play.

First, the supply. The phrase '100V DC supply' is the standard shorthand for an idealised voltage source: a source whose terminal voltage is held steady at 100V regardless of the current it delivers (equivalently, a source with negligible internal resistance). We adopt that idealisation throughout this problem.

Second, the accumulator. Being charged, its terminal voltage obeys the charging formula derived earlier in this section: $V = E + Ir = 50V + I(2\Omega) = 50V + (2\Omega)I$.

Third, the connecting wires. With $R = 0$, there is nothing else in the loop. The supply's positive terminal is connected directly to the accumulator's positive terminal by wires of negligible resistance, and similarly for the negative terminals. Two points joined by a wire of zero resistance must be at the same potential (because Ohm's law applied to that wire reads $V = I \times 0 = 0$ for any current). So the supply's terminal voltage equals the accumulator's terminal voltage:

$$V_{supply} = V_{acc}$$

$$100V = 50V + (2\Omega)I$$

$$I = \frac{100V - 50V}{2\Omega} = 25A$$

(iii) Substituting $I = 25A$ into the charging formula gives the accumulator's terminal voltage:

$$V = E + Ir = 50V + (25A)(2\Omega) = 100V$$

The accumulator's terminal voltage matches the supply voltage exactly, as we already noted in part (ii). With $R = 0$ there is nothing else in the loop, so the supply's $100V$ must appear directly across the accumulator's terminals.

(iv) When R is no longer zero, R sits in series between the supply and the accumulator. Ohm's law gives the voltage drop across R as $V_R = IR$. The supply's $100V$ is then shared (in series) between R and the accumulator's terminals:

$$V_{supply} = V_R + V_{acc}$$

$$100V = IR + (50V + (2\Omega)I)$$

Rearranging for I:

$$\mathrm{I} = \frac{50\mathrm{V}}{\mathrm{R} + 2\Omega}$$

which falls as R rises. The supply voltmeter V_1 reads $100\mathrm{V}$ regardless of R; the supply's own terminal voltage is independent of R (the supply is assumed to be a perfect $100\mathrm{V}$ source). The accumulator voltmeter V_2 reads $\mathrm{V} = \mathrm{E} + \mathrm{Ir} = 50\mathrm{V} + (2\Omega)\mathrm{I}$. As I falls toward zero with rising R, V_2 falls toward $\mathrm{E} = 50\mathrm{V}$. In the limit of very large R, almost no charging current flows, and the accumulator voltmeter reads its EMF directly.

Making Sense of the Answer: *The charging current at $R = 0$ is $25A$, which is dangerously large. No real charging room would allow this. Practical battery chargers always include a series resistance or, more commonly, an active current-limiting circuit that holds the charging current to a safe value (typically a few amperes for a small lead-acid cell, up to twenty or thirty for a large traction battery). The whole point of the variable R in this example is to let the operator dial the charging rate to a safe number. Increasing R slows the charge; the engineering trade-off is between charging quickly and the heat the cells can safely shed.*

Think Like a Physicist: *The cleanest way through any single-loop circuit containing two EMF sources is to think in terms of each component's terminal voltage, not in terms of energy balance. The supply's terminal voltage is a known property of the supply (here, just $100V$ since the supply is ideal). The accumulator's terminal voltage obeys $V = E + Ir$ during charging or $V = E - Ir$ during discharge. The voltage drop across any resistor is IR. The supply's voltage is then distributed across the loop's series elements, and the distributed voltages add up to the supply's voltage. This gives one equation in one unknown (the current), and the algebra is trivial.*

The Trap to Avoid: *The common error in charging problems is to forget that the accumulator's terminal voltage exceeds its EMF (it is $E + Ir$, not $E - Ir$). To decide which sign to use, ask: in which direction is the current passing through this source? In the direction the source would drive on its own (i.e. out of its positive terminal), use $V = E - Ir$. Forced backwards through the source (into its positive terminal, i.e. being charged), use $V = E + Ir$. Get the sign wrong and a $25A$ answer turns into a $75A$ answer, a factor of three off.*

HOT Example 21

A careless mechanic at a Dar es Salaam garage is leaning over the open bonnet of a car, working on the engine. He drops a metal spanner from his hand, and the spanner falls across the two terminals of the 12V car battery and rests there, with its body bridging the positive and negative terminals directly. The battery has internal resistance r = 0.04Ω. (a) Compute the current that flows through the spanner. (b) Compute the power dissipated inside the battery at that current. (c) Predict what happens physically in the next few seconds, and explain why a car battery has been designed in a way that makes this situation particularly dangerous.

Solution

(a) The spanner is a piece of metal of negligible resistance, so it presents essentially zero external resistance between the battery's two terminals. The voltage drop across the spanner is therefore zero (by Ohm's law applied to the spanner: $\mathrm{V} = \mathrm{I} \times 0 = 0$), and since the spanner is the only thing connected between the battery's terminals, the terminal voltage of the battery is $\mathrm{V} = 0$. Apply $\mathrm{V} = \mathrm{E} - \mathrm{Ir}$:

$$0 = 12\mathrm{V} - \mathrm{I}(0.04\Omega)$$

$$\mathrm{I} = \frac{12\mathrm{V}}{0.04\Omega} = 300\mathrm{A}$$

(b) Power dissipated inside the battery is $\mathrm{P} = \mathrm{I}^2\mathrm{r}$:

$$\mathrm{P} = (300\mathrm{A})^2(0.04\Omega) = 3600\mathrm{W}$$

(Equivalently, the total power delivered by the battery is $\mathrm{P} = \mathrm{EI} = (12\mathrm{V})(300\mathrm{A}) = 3600\mathrm{W}$, all of which is dissipated internally because the external resistance is zero.)

(c) The current of $300\mathrm{A}$ obtained in part (a) is what is called the **short-circuit current** of the battery: the current that flows when the external resistance is approximately zero. The spanner has acted as the short-circuiting wire.

The 3600W of power is concentrated inside the small volume of the battery's interior. In the first fraction of a second this heats the electrolyte rapidly. The point of contact between the spanner and the battery terminals carries the full 300A through a tiny surface area; the local resistance there (small but not exactly zero) dissipates further power, and the spanner glows red hot and can melt and weld itself to the terminal. The hydrogen and oxygen produced inside the casing by the rapid electrolysis can ignite. In the worst case the battery's casing ruptures and sprays sulphuric acid.

The reason this happens to a car battery and not, so spectacularly, to a torch cell is the design choice that gives the car battery its usefulness in the first place. A car's starter motor needs to draw hundreds of amperes for a couple of seconds while the engine is cranked. The battery must therefore have very low internal resistance, so that the Ir drop inside the battery does not collapse the terminal voltage when those high currents are drawn. That same low r is what permits the catastrophic 300A short-circuit current. Useful and dangerous are two sides of the same design.

Making Sense of the Answer: *The short-circuit current scales inversely with* r. *A car battery* ($E = 12V$, $r = 0.04\Omega$) *gives* $300A$. *A standard zinc-carbon dry cell* ($E = 1.5V$, $r = 0.5\Omega$) *gives only* $3A$. *The power dissipated scales as* $\frac{E^2}{r}$, *so the car battery dissipates* $3600W$ *in a short, while the dry cell dissipates only* $4.5W$. *That is a factor of 800 difference in danger, even though the EMFs differ by a factor of only 8. Internal resistance dominates the danger calculation.*

Think Like a Physicist: *The short-circuit current is a property of the source. It tells you the stiffness of the source: a source with small* r *holds its terminal voltage steady under load (because the* Ir *correction is small), and equivalently has a large short-circuit current. The two are inverse statements of the same property. Whenever a manufacturer quotes the internal resistance of a battery, they are giving you, implicitly, the short-circuit current, and through it the maximum power the battery could ever deliver to an external load.*

Five examples behind us.

Every example so far has had a single source driving a single resistor. Real circuits are not so simple. A torch has two cells stacked in series; a smartphone has parallel cells inside its casing; a household lighting circuit has dozens of lamps in parallel; and an oscillator on a teaching bench may have a Wheatstone bridge made of four resistors arranged with mirror symmetry. The next section opens that catalogue. Series resistors, parallel resistors, series cells, parallel cells, the symmetry methods that simplify complicated networks, and the shunt-and-multiplier modifications that turn a basic galvanometer into a useful ammeter or voltmeter, all follow. The same equations $V = IR$ and $V = E - Ir$ hold throughout; what changes is the bookkeeping.

RESISTORS, CELLS, SYMMETRY, AND METER MODIFICATION

Every circuit so far in this chapter has had a single source driving a single resistor. Real circuits are not so simple. The torch in your pocket has four cells stacked end to end. The phone in your other pocket has parallel cells inside its casing. The lighting circuit of your house has a dozen lamps in parallel between the live and neutral rails. The oscillator on your physics teacher's bench may have a Wheatstone bridge made of four resistors with a fifth one hung across its middle. None of these can be solved by the single-loop method of the last section.

The good news is that the underlying physics has not changed. Ohm's law still gives the voltage across each resistor, the EMF-plus-internal-resistance relation still gives the behaviour of each cell, and the conservation laws (charge in equals charge out at every junction, energy in equals energy out around every loop) still hold. What changes is the bookkeeping. This section catalogues the cases that yield to simple bookkeeping: resistors in series, resistors in parallel, cells in series, cells in parallel, networks that reduce by their geometric symmetry, and the meter modifications that turn a basic galvanometer into a useful ammeter or voltmeter. The next section, on Kirchhoff's laws, picks up everything that this one cannot reduce.

Series resistors

Three resistors of resistances R_1, R_2, and R_3 are connected end to end in a single closed loop, with a source of voltage V driving a current I round the loop. The three resistors are said to be in series. The following figure shows the configuration.

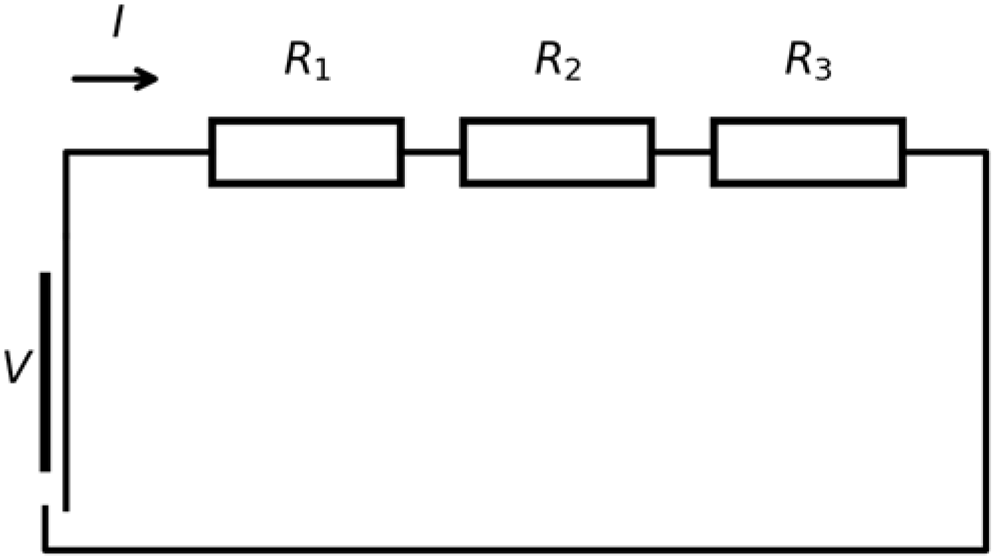


Figure: ***Three resistors connected in series with a source.*** *The same current I passes through each resistor. The total voltage V applied across the series combination equals the sum of the voltage drops across the individual resistors.*

Two physical facts about this arrangement are worth saying carefully, because everything else follows from them. The first is that the current is the same through every resistor. Charge cannot accumulate at any point in the loop (where would it go?), so the rate at which charge flows past any cross-section of the loop is the same as the rate at which it flows past any other cross-section. The second fact is that the total voltage drop around the loop equals the source voltage. This is conservation of energy: the work done per coulomb by the source must equal the energy lost per coulomb in the resistors, because the resistors are the only places the energy can be lost.

Apply Ohm's law to each resistor in turn, using the same current I because the current is the same everywhere:

$$V_1 = IR_1, \quad V_2 = IR_2, \quad V_3 = IR_3$$

Add the three equations together. The left-hand sides sum to the total voltage V (by conservation of energy); the right-hand sides factor I out:

$$V = V_1 + V_2 + V_3 = I(R_1 + R_2 + R_3)$$

Compare this with the Ohm's-law statement $V = IR_{eq}$ that would apply to a single equivalent resistor R_{eq} carrying the same total current I under the same source voltage V.

$$IR_{eq} = I(R_1 + R_2 + R_3)$$

$$R_{eq} = R_1 + R_2 + R_3$$

The same argument works for any number of resistors in series, not just three.

For n resistors R_1, R_2, ..., R_n in series:

$$\mathbf{R_{eq} = R_1 + R_2 + \cdots + R_n}$$

There is a useful consequence about how the source voltage is shared between the resistors. The voltage across each resistor is $V_i = IR_i$, and the current I is the same in every one. So V_i is directly proportional to R_i: the largest resistor gets the largest voltage drop, the smallest resistor gets the smallest.

Parallel resistors

Three resistors of resistances R_1, R_2, and R_3 are connected in parallel between two common points (call them the top rail and the bottom rail). A source of voltage V is connected across the rails. The following figure shows the configuration.

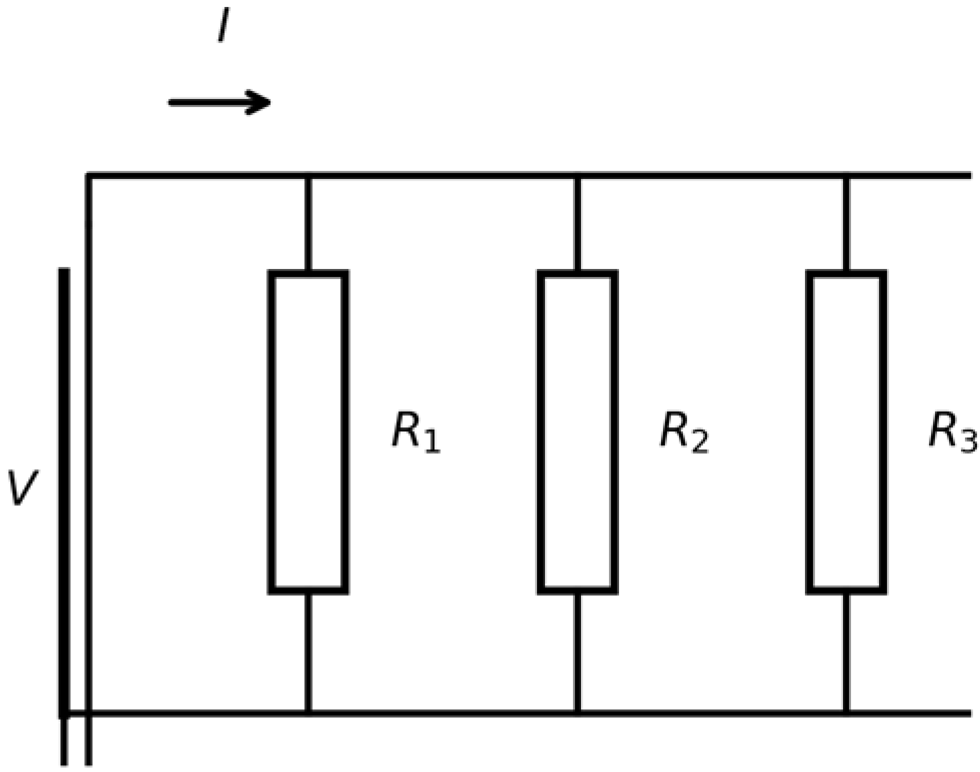


Figure: ***Three resistors connected in parallel with a source.*** *The same voltage V appears across each resistor. The total current I drawn from the source equals the sum of the currents through the individual resistors.*

The two physical facts to keep in mind are now the opposite of the series case. The voltage across each resistor is the same (because each resistor is connected between the same two rails, and the potential difference between two points cannot depend on which path you follow). The total current drawn from the source is the sum of the currents through the individual resistors (because at the junction where the rails split into the three branches, charge in equals charge out).

Apply Ohm's law to each resistor with the common voltage V:

$$\mathrm{I_1 = \frac{V}{R_1}, \quad I_2 = \frac{V}{R_2}, \quad I_3 = \frac{V}{R_3}}$$

Add the three currents to give the total drawn from the source:

$$\mathrm{I = I_1 + I_2 + I_3 = V\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right)}$$

Compare this with Ohm's law applied to a single equivalent resistor: $\mathrm{I = \frac{V}{R_{eq}}}$. The two read identically when:

$$\mathrm{\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}}$$

The same argument generalises to n parallel resistors in the obvious way.

$$\mathbf{\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots + \frac{1}{R_n}}$$

The special case of two resistors in parallel is so common that the formula is worth remembering in its compact form. Setting $\mathrm{\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2}}$ and rearranging gives:

$$\mathrm{R_{eq} = \frac{R_1 R_2}{R_1 + R_2}}$$

The current through each branch is inversely proportional to its resistance. The smallest resistor in a parallel combination takes the largest share of the current; the largest resistor takes the smallest share. For n equal resistors connected in parallel, the current is shared equally among them, and the equivalent resistance is equal to the resistance of one resistor divided by n, as shown below:

$$\frac{1}{R_{eq}} = \frac{1}{R} + \frac{1}{R} + \frac{1}{R} + \cdots = \frac{1 + 1 + 1 + \cdots}{R} = \frac{n}{R}$$

$$\mathbf{R_{eq} = \frac{R}{n}}$$

Series cells

Cells in series follow the same logic as resistors in series, with one extra fact: each cell adds its own EMF, with sign, to the total EMF of the chain. Suppose n identical cells, each of EMF E and internal resistance r, are connected in series with all the positive terminals joining to the negative terminals of their neighbours (so that the EMFs all push in the same direction round the loop). The figure below shows the arrangement.

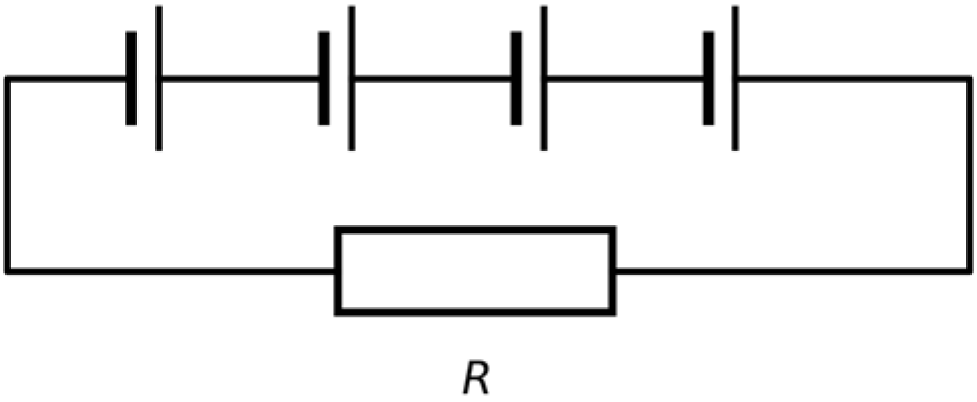


Figure: n identical cells in series, all driving in the same direction*. The total EMF is nE and the total internal resistance is nr.*

Because the cells are in series, the same current passes through every cell. Each cell's internal resistance r drops a voltage Ir along the chain, and each cell's EMF E lifts the potential by E. The two effects add cell by cell, giving:

$$\mathbf{E_{eq} = nE, \quad r_{eq} = nr}$$

The equivalent EMF is n times the EMF of a single cell; the equivalent internal resistance is n times the internal resistance of a single cell. Both scale with the number of cells. Series stacking is the right choice when you want a higher voltage than a single cell can supply, and you can afford the larger internal resistance that comes with it. The 6V torch in your house is a stack of four 1.5V cells in series; the 12V battery in a small motor-vehicle is six 2V lead-acid cells in series. The 1.5V is the natural EMF of a zinc-carbon cell and the 2V is the natural EMF of a lead-acid cell; cells are stacked to multiply the voltage by a small integer until the practical voltage required is reached.

Parallel cells of equal EMF

Cells in parallel work in the opposite direction. Suppose n identical cells (same EMF E, same internal resistance r) are connected in parallel, all positive terminals joining to a common top rail and all negative terminals joining to a common bottom rail. The next figure shows the arrangement.

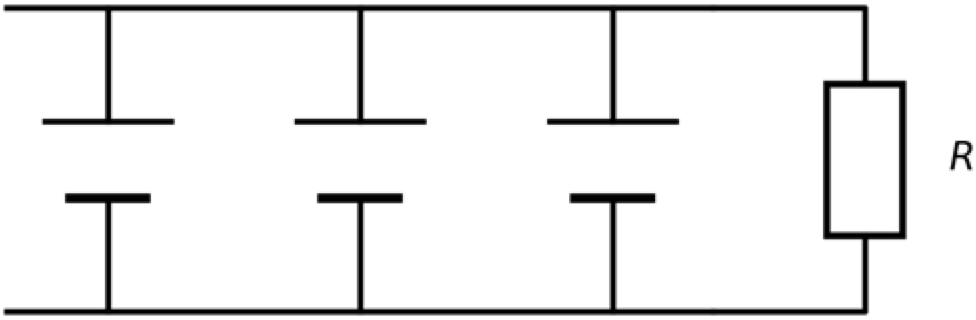


Figure: ***n identical cells in parallel.*** *The equivalent EMF is unchanged at E; the equivalent internal resistance is r divided by n. Voltage is unchanged, but the current capability is multiplied by n.*

Because all n cells share the same two rails, they all have the same terminal voltage; and because they all have the same EMF E, they all drive in the same direction. The equivalent EMF of the combination is therefore just E itself, the EMF of any single cell. The internal resistances, however, combine as resistors in parallel: n equal resistances r in parallel give $\frac{r}{n}$. So:

$$E_{eq} = E, \quad r_{eq} = \frac{r}{n}$$

The voltage stays the same, but the internal resistance falls by a factor of n. This is useful when you want to draw more current than a single cell can comfortably supply, at the cell's natural voltage. A **smartphone battery** pack is several lithium cells of identical EMF wired in parallel; the voltage stays at the cell's natural value (about 3.7V) while the current capability multiplies up.

The two strategies (series for voltage, parallel for current) are often combined: a large vehicle traction battery is a series-parallel grid of many cells together delivering both the high voltage and the high current demands of the motor.

If the parallel cells do not have equal EMFs, the analysis is less elementary. The cells then drive current internally from one to the other even with no external load, wasting energy. The general parallel-cell formula is derived in the next section using Kirchhoff's laws, and the practical advice is to avoid the configuration unless you specifically intend it.

Symmetry methods in networks

The series and parallel formulas of the last four subsections handle every network that can be broken into nested chunks of pure series or pure parallel structure. Many real circuits look exactly like this, and they reduce by repeated application of the two formulas alone. But not every circuit is so cooperative.

Consider a cubical lattice of twelve identical resistors, one on each edge of the cube, and connect a battery across the cube's main diagonal. *What is the equivalent resistance between the two diagonally opposite corners?* Stare at it as long as you like, you will not find any pair of edges that is in pure series with another, nor any group of edges in pure parallel with the rest. The series and parallel formulas alone are stuck.

There is, however, a way through. The cube has hidden order: it looks the same from several different points of view. That geometric symmetry leads, through a short argument we shall present next, to the conclusion that certain corners of the cube must be at the same potential. Once we know that, we may join those corners with a wire (because the wire carries no current) and redraw the network. The redrawn network turns out to be a simple series chain of parallel blocks, which the series and parallel formulas of the earlier subsections reduce in one line. That is what symmetry methods are. The next few paragraphs spell out the logic step by step.

Why symmetry forces two nodes to be at the same potential

Suppose you can perform some operation on the network (a rotation, a reflection, or a relabelling of corners) that has two properties. First, after the operation the picture looks identical to what it was before. Same wires, same resistors, same battery in the same place. Second, the operation swaps two specific nodes, call them X and Y. Then the operation has the unavoidable consequence that $V_X = V_Y$.

The argument is short. Before the operation, node X sits at some potential V_X and node Y sits at some other potential V_Y. After the operation the network is physically the same network as before (that is what the first property tells us), so the potential at any chosen point of it must be the same as before. But the operation has swapped X and Y, so the position now called 'X' used to be called 'Y'. That position must therefore have both potentials at the same time: the one it had before the operation (V_Y, from its old name) and the one it has after the operation (V_X, from its new name, since the picture is unchanged). The only way both numbers can describe the same point is if they are equal.

In plain words: *if a circuit looks the same after we swap two of its nodes, the two nodes cannot tell each other apart electrically, so they must be at the same potential.* This is the central lemma of symmetry methods.

Why we may merge equipotential nodes

Once we know that $V_X = V_Y$, we are allowed to add a wire of zero resistance between X and Y. The added wire does not change any current anywhere in the network. Why not? Because the current that would flow in the added wire would obey Ohm's law applied to the wire alone: $I = V/R$. The voltage V across the wire is the difference $V_X - V_Y$, which we have just shown is zero. The resistance R of the added wire is also zero (it is an ideal wire). So the current is $I = 0/0$ (indeterminate from this calculation alone), but the rest of the circuit has not changed at all, so the current actually carried by the added wire is whatever amount the rest of the network needed to be carried between X and Y anyway, namely none. Adding the wire is electrically the same as fusing the two nodes into one combined node. We say the two equipotential nodes have been merged.

Stated as a rule: *nodes proved equipotential by symmetry can be merged into a single node without changing anything that matters.*

A warm-up example: the balanced Wheatstone bridge

Before we attack the cube, here is a smaller example that uses the same two steps. The figure below shows four resistors of equal resistance R arranged as the four sides of a diamond ABCD, with a galvanometer G connected across the diagonal BD. A battery drives current in at A and out at C.

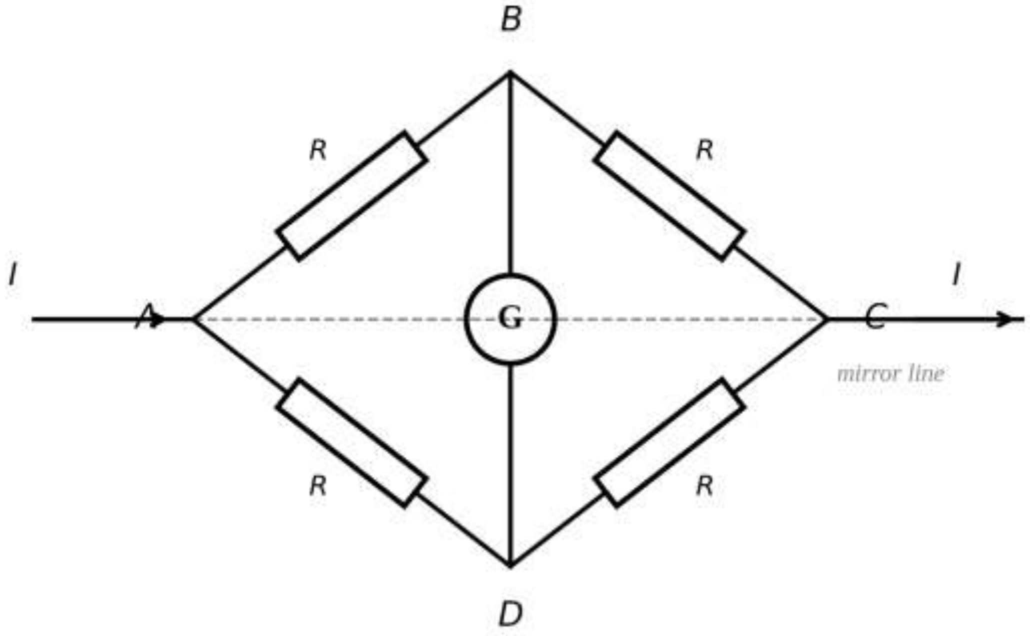


Figure: ***A balanced Wheatstone bridge****: four equal resistors R on the sides of a diamond, with a galvanometer G across the diagonal BD. The input is at A and the output is at C. The dashed line through A and C is the mirror axis of the network.*

Find a symmetry of this network. Reflect the whole figure across the horizontal dashed line through A and C. The top half (the path A→B→C through two R resistors) flips down to where the bottom half used to be (the path A→D→C through two R resistors). The bottom half flips up. After the reflection the picture looks identical to what it was before: the resistors are still all R, the battery is still at A and C, and the galvanometer is still on the vertical line. The reflection has swapped B with D, while leaving the picture unchanged.

The lemma above then tells us that B and D must be at the same potential: $V_B = V_D$. The merge rule above then tells us we are allowed to merge B and D into a single node. The galvanometer G, which sits between B and D, therefore has zero voltage across it and carries zero current. The galvanometer reads zero; this is the classic balanced-bridge condition that turns the Wheatstone bridge into a precision instrument for comparing resistances.

With B and D merged into one node, the four resistors are no longer a bridge but a much simpler combination: AB and AD are now in parallel between A and the merged BD node, and BC and DC are in parallel between the merged node and C. Each parallel pair has equivalent resistance $R/2$, and the two pairs are then in series along the path A to merged BD to C, giving the total resistance between A and C as:

$$\frac{R}{2} + \frac{R}{2} = R$$

Notice how the two-step recipe (find a symmetry, merge the swapped nodes) reduced the bridge to a straightforward series-parallel calculation. The same recipe, applied to a more elaborate network, is what we use on the cube.

The cube of twelve resistors

Now back to the cube. The cube has eight corners; label them A (the input), B, D, E (the three corners directly adjacent to A, joined to it by single edges), C, F, H (the three corners directly adjacent to G, joined to it by single edges), and G (the output, diagonally opposite A across the cube). Each of the twelve edges is a resistor of resistance R.

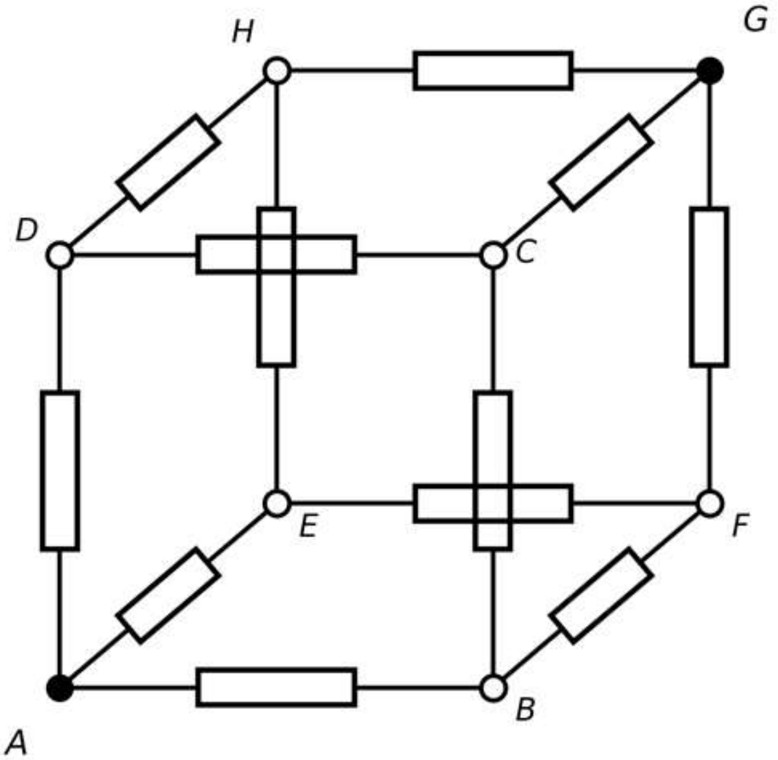


Figure: The cubical network of twelve equal resistors. The eight corners are labelled A, B, C, D, E, F, G, H. Current enters at corner A (filled circle, bottom-left) and leaves at the diagonally opposite corner G (filled circle, top-right). The three corners B, D, E adjacent to A and the three corners C, F, H adjacent to G are drawn as open circles. Each of the twelve edges is a resistor of resistance R.

Find the symmetry. Pick the cube up by the two corners A and G (the diagonal along which current enters and leaves) and rotate it through 120° about that diagonal. The three corners B, D, E cycle among themselves: B moves to where D was, D moves to where E was, E moves to where B was. The three corners C, F, H cycle similarly. A stays at A, G stays at G, every resistor moves to the position previously occupied by another resistor (with the same resistance R, since they are all identical). So after the rotation, the picture looks identical to what it was before.

Apply the lemma above. The rotation cycled B to D, D to E, E to B. So all three of B, D, E must sit at the same potential. By the same argument applied on the output side, all three of C, F, H must also sit at a (common) potential of their own. We now have two groups of three equipotential nodes each.

Apply the merge rule above. Merge B, D, E into a single node, call it P. Merge C, F, H into a single node, call it Q. The cube has now collapsed into a much simpler picture, which the next figure draws explicitly.

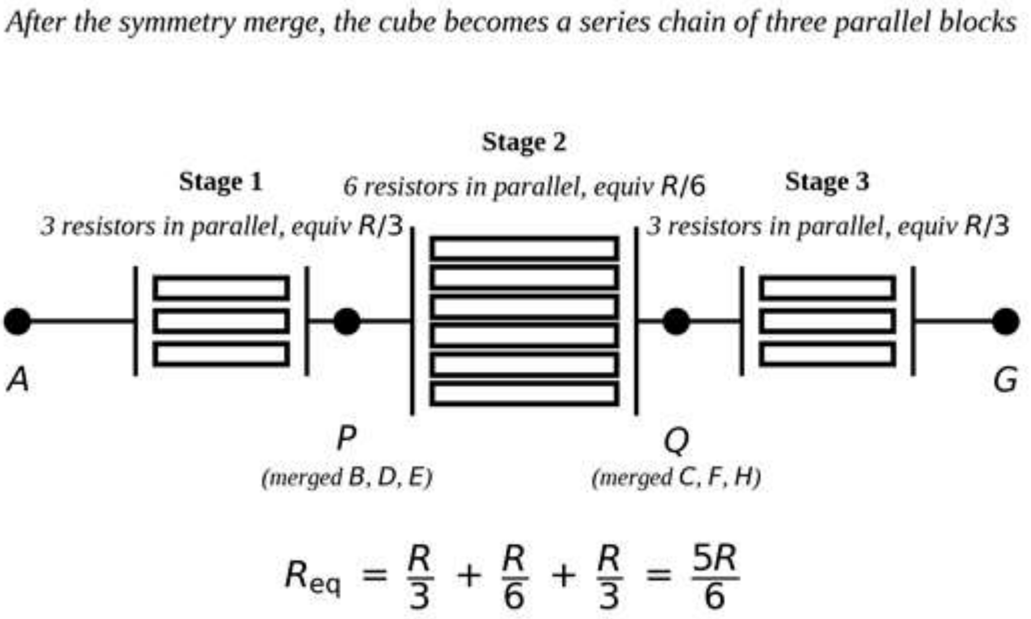


Figure: ***The cube AFTER the symmetry merge.*** *The three edges from A to its neighbours now connect A to the single merged node P, so they are three resistors R in parallel (equivalent R/3). The six edges that connected an A-neighbour to a G-neighbour now connect P to Q, so they are six resistors R in parallel (equivalent R/6). The three edges from the G-neighbours to G now connect Q to G, so they are three resistors R in parallel (equivalent R/3). These three blocks lie in series along the path A→P→Q→G, and their resistances add: total = R/3 + R/6 + R/3 = 5R/6.*

The figure makes the conclusion obvious: the cube, after merging, is just three parallel blocks in series. The first block has three equal R resistors in parallel (equivalent R/3), the middle block has six equal R resistors in parallel (equivalent R/6), and the last block has three equal R resistors in parallel (equivalent R/3). The series chain adds:

$$\mathbf{R_{eq}} = \frac{\mathbf{R}}{\mathbf{3}} + \frac{\mathbf{R}}{\mathbf{6}} + \frac{\mathbf{R}}{\mathbf{3}} = \frac{\mathbf{5R}}{\mathbf{6}}$$

The cube’s twelve-resistor network has been reduced to a single equivalent resistance of 5R/6. Example 22 below plugs in the numerical R = 2Ω and a 6V battery to find the total current drawn.

Summary recipe

Symmetry methods always follow the same four steps. Each step is set out below with its own short explanation of how to do it in practice.

Step 1: Find a symmetry of the network

Look for an operation you can perform on the picture that leaves it identical. The operation may be a rotation, a reflection across a line, or a permutation of node labels. The test of success is this: after the operation, every wire still goes between the same two corners as before, every resistor still has the same resistance as before, and the battery still sits at the same two terminals as before. The picture must look identical including the battery, not only the resistors.

In the balanced-bridge example above, the symmetry was the reflection across the horizontal line through A and C. In the cube example above, the symmetry was a 120-degree rotation about the diagonal from A to G. These are the two most common kinds of symmetry you will meet at this level: a mirror reflection and a rotational permutation.

Step 2: Identify the nodes that the symmetry moves

Track each node of the network through the operation. Some nodes (typically the input and output of the network) will stay in place. Other nodes will be swapped, or cycled, with one or more of their neighbours. List the groups of nodes that get cycled among themselves by the operation. Each such group has been proved equipotential by the lemma proved earlier.

In the bridge, the reflection swapped B with D, so {B, D} formed one equipotential group. In the cube, the 120-degree rotation cycled B to D to E to B, and cycled C to F to H to C, so {B, D, E} and {C, F, H} formed two equipotential groups. The number of groups equals the number of independent merges you will perform in the next step.

Step 3: Merge each equipotential group into a single node

For each equipotential group identified in Step 2, draw a fresh node and connect every member of the group to it with a zero-resistance wire. By the merge argument given earlier, these added wires carry no current, so no measurable quantity has been changed. The old nodes can now be erased and replaced by the single new node, which inherits all the resistors that previously hung off any member of the group.

Give each merged node a short fresh name (P, Q, and so on). This keeps the redrawn picture clean and avoids the confusion of carrying three labels for one electrical point. In the cube, the merged A-neighbour node became P and the merged G-neighbour node became Q.

Step 4: Redraw the simplified network, then reduce by series and parallel

Sketch the simplified network with the merged nodes in place. Count how many of the original resistors now run between each pair of nodes; resistors that previously ran between two members of the same equipotential group will have collapsed to short wires and may be ignored, while resistors that ran from one group to another are all now in parallel between the corresponding merged nodes.

Replace each parallel cluster by its equivalent resistance using the parallel-resistor formula introduced earlier. The resulting picture is a simple series chain of equivalent resistors, and the series-resistor formula introduced earlier finishes the job. The total equivalent resistance comes out in one or two further lines of arithmetic.

In the cube, after the merge there were three resistors in parallel between A and P, six resistors in parallel between P and Q, and three resistors in parallel between Q and G. The three parallel clusters had equivalent resistances R/3, R/6, and R/3 respectively, and these three sat in series along the path from A to G, summing to 5R/6.

Two practical warnings

First, the symmetry must respect the battery as well as the resistors. An operation that moves the battery to a different position on the network does not count, because the picture is no longer identical. This is why the balanced bridge has a mirror axis through A and C (the battery terminals) and not through B and D: a reflection across the BD axis would carry the battery to a different pair of terminals, which is a different physical setup.

Second, not every circuit has a useful symmetry. Networks like the unbalanced Wheatstone bridge of worked Example 28 have no symmetry that swaps the bridge nodes, so this method does not apply to them. For those circuits we shall need the more powerful method of the next section, Kirchhoff's laws.

Meter modification: shunts and multipliers

Two of the most common measuring instruments in any physics laboratory are the ammeter and the voltmeter. Both are built up from the same basic component: a sensitive **galvanometer**, which *is a current-measuring instrument whose pointer deflects in proportion to the current passing through its coil.*

A typical laboratory galvanometer gives full-scale deflection at a current of milliamperes, with the coil of resistance fifty to one hundred ohms. The galvanometer by itself is therefore neither an ammeter (which needs a range of amperes, not milliamperes) nor a voltmeter (which reads voltage, not current). To turn it into either, you add a single extra resistor in the right place.

To convert the galvanometer into an ammeter of full-scale range $\mathbf{I_{max}}$, connect a small resistor **S** (the **shunt**) in parallel with the galvanometer. The shunt diverts most of the current around the galvanometer, so that the galvanometer carries only its own small full-scale current $\mathbf{I_g}$ while the total current through the combination reaches $\mathrm{I_{max}}$. At full-scale deflection, $\mathrm{I_g}$ passes through the galvanometer and $\mathrm{I_{max}} - \mathrm{I_g}$ passes through the shunt. The voltage across the galvanometer is $\mathrm{I_g G}$ (with G the galvanometer's own resistance), and this is the same as the voltage across the shunt (they are in parallel), so:

$$\mathrm{I_g G = (I_{max} - I_g)S}$$

Solving for the required shunt:

$$\mathbf{S = \frac{I_g G}{I_{max} - I_g}}$$

Because $\mathrm{I_g}$ is small compared with $\mathrm{I_{max}}$, the shunt resistance S comes out very small (a fraction of an ohm in typical cases). The full ammeter (galvanometer in parallel with shunt) has effective resistance equal to the parallel combination of G and S, which is dominated by the smaller of the two, namely S. This is exactly what an ammeter needs: *a very low resistance, so that placing the ammeter into a circuit hardly changes the current already flowing there. A good ammeter is a near-perfect conductor.*

To convert the same galvanometer into a voltmeter of full-scale range $\mathbf{V_{max}}$, connect a large resistor **M** (the **multiplier**) in series with the galvanometer. At full-scale deflection the current through the combination is $\mathbf{I_g}$ (the galvanometer's full-scale current), and the voltage across the combination is $\mathrm{V_{max}}$. By Ohm's law applied to the (galvanometer + multiplier) series chain:

$$\mathrm{V_{max} = I_g(G + M)}$$

Solving for the required multiplier:

$$\mathbf{M = \frac{V_{max}}{I_g} - G}$$

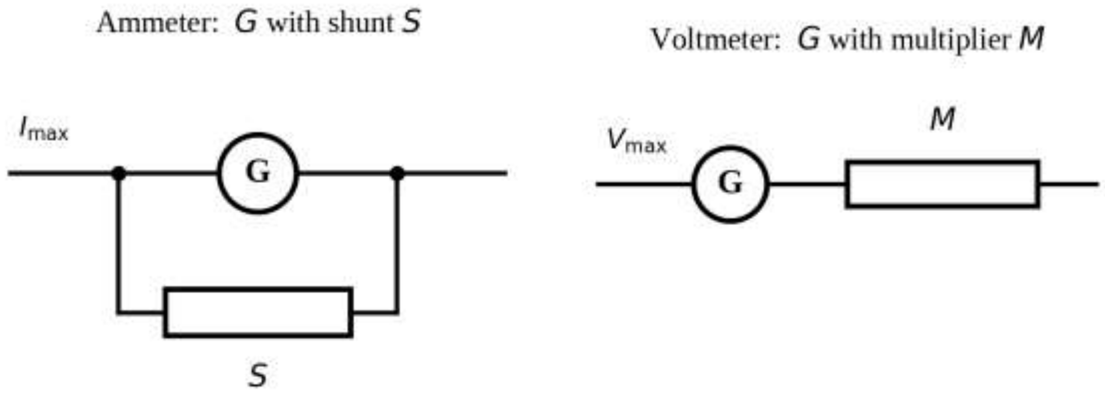


Figure: ***Galvanometer modifications.*** *Left: an ammeter, made by adding a small shunt S in parallel with the galvanometer G. Right: a voltmeter, made by adding a large multiplier M in series with G.*

Because $\mathrm{V_{max}}$ is typically several volts and $\mathrm{I_g}$ is only a milliampere, the multiplier resistance comes out large (thousands of ohms or tens of thousands of ohms). The full voltmeter has effective resistance $\mathrm{G + M}$, which is essentially M. This is what a voltmeter needs: *a very high resistance, so that placing the voltmeter across a circuit element draws almost no current from the circuit and hardly changes the voltage already present there. A good voltmeter is a near-perfect insulator.* Worked Examples 23 and 24 design a shunt and a multiplier respectively for a typical laboratory galvanometer.

Seven examples follow. Each one puts a different method of this section to the test, and the last anticipates a tool we shall need in the next section.

BINDER Example 22

A cubical network is built from twelve resistors of 2Ω each, one on each of the twelve edges of the cube. The network is connected across its main diagonal (from corner A to the diagonally opposite corner G) to a 6V battery of negligible internal resistance. Find the equivalent resistance of the network and the total current drawn from the battery.

(The cubical network is the same as the one used in the symmetry-methods subsection, with R = 2Ω on each edge.)

Solution

The full derivation was carried out in the symmetry-methods subsection under the heading 'The cube of twelve resistors'. The three-fold rotational symmetry about the main diagonal A → G forces the three A-neighbours B, D, E to share a common potential, and likewise the three G-neighbours C, F, H. Merging each group into a single node and redrawing collapses the cube into three parallel blocks in series, with equivalent resistance:

$$R_{eq} = \frac{R}{3} + \frac{R}{6} + \frac{R}{3} = \frac{5R}{6}$$

With $R = 2\Omega$:

$$R_{eq} = \frac{5 \times 2\Omega}{6} = 1.67\Omega$$

The total current drawn from the 6V battery (negligible internal resistance, so terminal voltage equals EMF) follows from Ohm's law applied to the equivalent resistance:

$$I = \frac{V}{R_{eq}} = \frac{6V}{1.67\Omega} = 3.6A$$

Making Sense of the Answer: *Symmetry has done all the work. The twelve-resistor cube, which has no obvious series or parallel decomposition when stared at directly, collapses in three short stages once the symmetry merge is performed. The current* $3.6A$ *is consistent with the small total resistance: a* $6V$ *battery into less than two ohms must drive a few amperes.*

Think Like a Physicist: *Always look for symmetry first when a network does not reduce by obvious series and parallel. The four-step recipe given earlier is general: find an operation that leaves the picture unchanged, identify the nodes it swaps or cycles, merge them, redraw and reduce. When no symmetry can be found, the Kirchhoff's laws of the next section give a universal method, but they are also the slower one. Symmetry first, Kirchhoff only as a fallback.*

HOT Example 23

A galvanometer has resistance 50Ω and gives full-scale deflection when a current of 1mA passes through it. Design a shunt that converts the galvanometer into an ammeter reading 1A at full scale. Find the resistance of the shunt and the effective resistance of the resulting ammeter.

Solution

Use the shunt formula derived above:

$$S = \frac{I_g G}{I_{max} - I_g}$$

Substitute $G = 50\Omega$, $I_g = 1mA = 0.001A$, $I_{max} = 1A$:

$$S = \frac{(0.001A) \times (50\Omega)}{1A - 0.001A} = 0.0501\Omega$$

The shunt has resistance about 0.05Ω (one twentieth of an ohm). The effective resistance of the ammeter is the parallel combination of $G = 50\Omega$ and $S = 0.05\Omega$:

$$\mathrm{R_{ammeter}} = \frac{50\Omega \times 0.05\Omega}{50\Omega + 0.05\Omega} = 0.05\Omega$$

Making Sense of the Answer: *The completed ammeter has an effective resistance of about* 0.05Ω*, which is dominated by the shunt because the shunt is very much smaller than the galvanometer's* 50Ω*. The small effective resistance is the whole point of the design: when you insert this ammeter into a circuit to measure the current, you add only* 0.05Ω *of extra series resistance, which is negligible in any circuit with sensible resistor values. The ammeter is, for practical purposes, a near-invisible piece of wire. The 1000:1 ratio between the current ranges (from* $1mA$ *to* $1A$*) is achieved by a 1000:1 ratio between* G *and* S*, which is the general design rule for ammeter shunts.*

Think Like a Physicist: *The shunt formula is conservation of charge written down with Ohm's law substituted in. At full-scale deflection,* I_g *passes through G and the rest,* $I_{max} - I_g$*, passes through S; since G and S are in parallel they share the same voltage; equating the two voltages gives the formula. Never memorise the shunt formula without remembering this short derivation, because the same physics will reappear in the next example with a different geometry.*

HOT Example 24

The same galvanometer of Example 23 (50Ω resistance, 1mA full-scale deflection) is now to be converted into a voltmeter reading 10V at full scale. Find the resistance of the required multiplier, and find the effective resistance of the resulting voltmeter.

Solution

Use the multiplier formula derived earlier:

$$\mathrm{M} = \frac{\mathrm{V_{max}}}{\mathrm{I_g}} - \mathrm{G}$$

Substitute $\mathrm{V_{max}} = 10\text{V}, \mathrm{I_g} = 0.001\text{A}, \mathrm{G} = 50\Omega$:

$$\mathrm{M} = \frac{10\text{V}}{0.001\text{A}} - 50\Omega = 9950\Omega$$

The multiplier has resistance of 9950Ω.

The effective resistance of the voltmeter is the series combination of G and M:

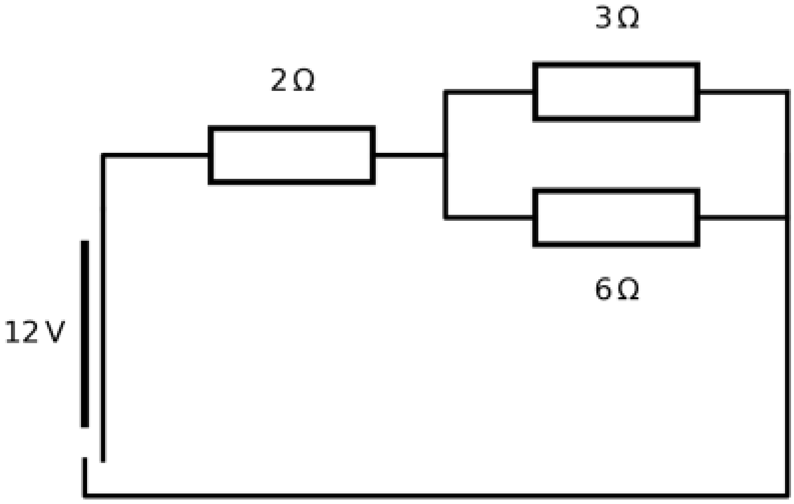
$$\mathrm{R_{voltmeter}} = \mathrm{G} + \mathrm{M} = 50\Omega + 9950\Omega = 10000\Omega$$

Making Sense of the Answer: *The completed voltmeter has an effective resistance of* $10k\Omega$*, which is dominated by the multiplier. This is the whole point of the design: when you connect this voltmeter across a circuit element to measure its voltage, you add a parallel path of* $10k\Omega$ *that draws only* $1mA$ *of current. In most circuits this current is small enough that the voltage being measured is not appreciably disturbed. The voltmeter is, for practical purposes, a near-invisible insulator. Compare this with the ammeter of Example 23, which was a near-invisible conductor: the two instruments use the same galvanometer but achieve opposite effective resistances by adding the helper resistor in opposite places.*

Think Like a Physicist: *Every measuring instrument disturbs the circuit it measures. The design challenge in both the ammeter and the voltmeter cases is to make the disturbance small. The strategy is different for the two instruments because they measure different quantities: an ammeter is in series with the load and so must have small resistance, a voltmeter is in parallel with the load and so must have large resistance. The shunt and the multiplier are not interchangeable. Mixing them up (putting the shunt in series with the galvanometer, say) gives a useless instrument. Always draw the circuit and think about which quantity, current or voltage, the instrument is intended to measure.*

BINDER Example 25

Using the figure below, determine: (a) the equivalent resistance of the network; (b) the current through each resistor; (c) the power dissipated in each resistor.



Solution

(a) First reduce the parallel combination. With $R_1 = 3\Omega$, $R_2 = 6\Omega$:

$$R_{\text{parallel}} = \frac{R_1 R_2}{R_1 + R_2} = \frac{(3\Omega)\times(6\Omega)}{3\Omega + 6\Omega} = 2\Omega$$

Then add the 2Ω series resistor on top of this parallel combination:

$$R_{\text{eq}} = 2\Omega + 2\Omega = 4\Omega$$

(b) The current drawn from the battery is by Ohm's law applied to the equivalent resistance:

$$I = \frac{12V}{4\Omega} = 3A$$

This is the current through the 2Ω resistor (which is in series with everything else). The voltage across the parallel combination is $V = IR = (3A)\times(2\Omega) = 6V$, and this same voltage appears across both branches of the parallel pair.

So the current through the 3Ω branch is $I_3 = \frac{6V}{3\Omega} = 2A$

And the current through the 6Ω branch is $I_6 = \frac{6V}{6\Omega} = 1A$.

Check: $2A + 1A = 3A$ matches the total drawn from the battery.

(c) Power in each resistor follows from $P = I^2R$ (or equivalently $P = \frac{V^2}{R}$):

$$P_{2\Omega} = (3A)^2 \times 2\Omega = 18W$$

$$P_{3\Omega} = (2A)^2 \times 3\Omega = 12W$$

$$P_{6\Omega} = (1A)^2 \times 6\Omega = 6W$$

Making Sense of the Answer: *Total power dissipated:* $18W + 12W + 6W = 36W$. *Check against* $P = IV = (3A)\times(12V) = 36W$ *of total power delivered by the battery. The two numbers agree, as they must by conservation of energy. The* 2Ω *resistor (which carries the full current) gets the largest share of the power; the* 6Ω *branch (which carries the smallest current) gets the smallest share. Both observations are general: in a series chain, the largest resistance dissipates the most; in a parallel cluster, the smallest resistance dissipates the most.*

Think Like a Physicist: *Mixed networks reduce by working from the innermost knot outwards. Spot the parallel block first, replace it by its equivalent, then the rest of the network is a simple series chain. Once the current through the equivalent resistor is known, work back into the parallel block to find how the current* *divides between its branches. The two-step strategy (outwards to find the total, inwards to find the parts) works for every mixed network this section can handle.*

REAL Example 26

A 6V torch is powered by four identical zinc-carbon cells connected in series, each of EMF 1.5V and internal resistance 0.2Ω. The torch bulb has resistance 4Ω when hot. Find the total EMF of the battery stack, the total internal resistance, the current that flows through the bulb, and the power dissipated in the bulb.

Solution

Four cells in series each contribute their EMF and internal resistance to the totals. Using the series-cells formulas:

$$\mathrm{E_{total}} = \mathrm{nE} = 4 \times 1.5\mathrm{V} = 6\mathrm{V}$$

$$\mathrm{r_{total}} = \mathrm{nr} = 4 \times 0.2\Omega = 0.8\Omega$$

The current through the bulb is determined by the total EMF, which is given by:

$$\mathrm{E_{total}} = \mathrm{I(R + r_{total})}$$

$$\mathrm{I} = \frac{\mathrm{E_{total}}}{\mathrm{R + r_{total}}} = \frac{6\mathrm{V}}{4\Omega + 0.8\Omega} = 1.25\mathrm{A}$$

Power dissipated in the bulb is $\mathrm{P = I^2R}$:

$$\mathrm{P} = (1.25\mathrm{A})^2 \times 4\Omega = 6.25\mathrm{W}$$

Making Sense of the Answer: *The bulb dissipates* $6.25W$*, which is a reasonable figure for a torch bulb. Some of the available electrical power is also lost inside the cells:* $P_{internal} = I^2 r_{total} = 1.25^2 \times 0.8 = 1.25W$*, which warms the cells while the torch is on. The total power drawn from the battery stack is* $6.25W + 1.25W = 7.5W$*, which matches* $I \times E_{total} = 1.25A \times 6V = 7.5W$ *as it must.*

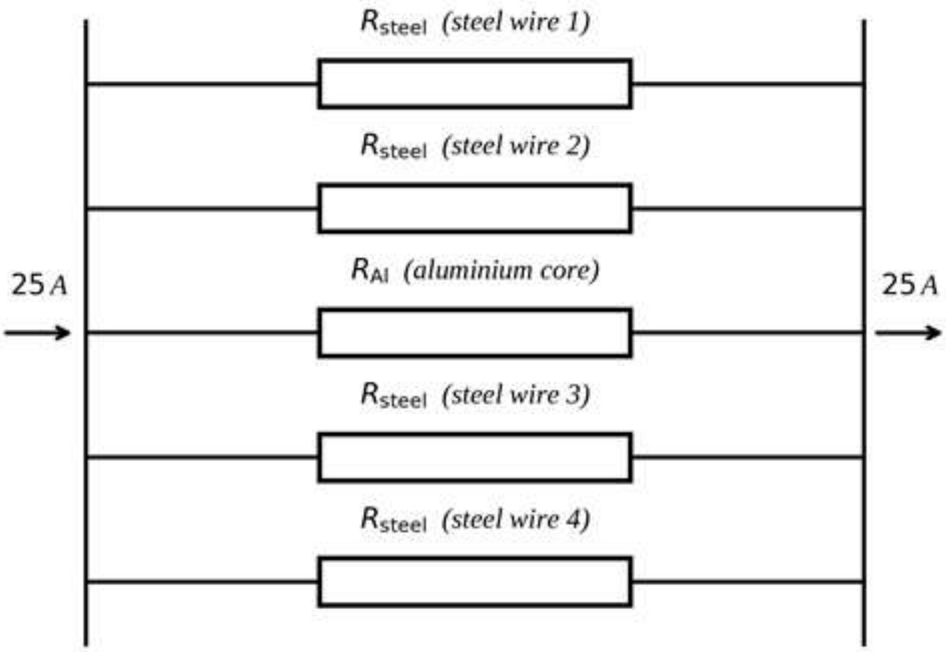
Think Like a Physicist: *Two facts about this torch are worth noticing. First,* 17% *of the chemical energy of the cells is lost as heat inside the cells themselves (the ratio* $\frac{r_{total}}{R+r_{total}}$ *of internal to total resistance equals the fraction of power wasted internally). This is an inevitable loss; nothing about circuit design can recover it. Second, as the cells age, their internal resistance r grows. The total internal resistance grows with it, and a larger fraction of the total power is wasted inside the cells, leaving less for the bulb. This is why old torch cells produce a dim light: not because the EMF has dropped much, but because too much of the available power is now being wasted internally.*

HOT Example 27

A composite TANESCO transmission cable consists of an aluminium core surrounded by four steel reinforcing wires, each of the same length and the same cross-sectional area as the aluminium core. The total current carried by the cable is 25A. The resistivities are $\rho_{\mathrm{Al}} = 2.82 \times 10^{-8}\Omega\mathrm{m}$ and $\rho_{\mathrm{steel}} = 1.6 \times 10^{-7}\Omega\mathrm{m}$. Find the current carried by the aluminium core and the current carried by each of the four steel wires.

Solution

The five wires (one aluminium core, four steel reinforcement wires) run the whole length of the cable and meet at the cable's two ends. So all five share the same pair of endpoints, which means they are connected in parallel. The following circuit diagram sets the parallel network out explicitly.



Since all five wires have the same length and the same cross-section, their geometric factor l/A is the same for each wire, and hence the resistance R = ρl/A is proportional to the resistivity ρ alone. A ratio of two resistances therefore reduces to a ratio of two resistivities.

Computing that ratio:

$$\frac{R_{steel}}{R_{Al}} = \frac{\rho_{steel}}{\rho_{Al}} = \frac{1.6 \times 10^{-7}\Omega m}{2.82 \times 10^{-8}\Omega m} = 5.67 \text{ or } R_{steel} = 5.67R_{Al}$$

Since the five wires are in parallel, the voltage across each of them is the same value V. The current in each wire is given by Ohm's law applied to that wire alone:

$$I_{Al} = \frac{V}{R_{Al}}, \quad I_{steel} = \frac{V}{R_{steel}}$$

Taking the ratio of these two equations cancels V and gives the current ratio in terms of the resistance ratio:

$$\frac{I_{steel}}{I_{Al}} = \frac{R_{Al}}{R_{steel}} = \frac{R_{Al}}{5.67R_{Al}} = \frac{1}{5.67} = 0.176 \text{ or } I_{steel} = 0.176I_{Al}$$

The total current is the sum across all five wires:

$$25A = I_{Al} + 4 \times I_{steel} = I_{Al} + 4 \times (0.176I_{Al}) = 1.706I_{Al}$$
$$I_{Al} = 14.7A$$

And the current in each steel wire follows:

$$I_{steel} = 0.176 \times 14.7A = 2.6A$$

Checking the totals: aluminium 14.7A plus four steel wires at 2.6A each gives 14.7A + 4 × 2.6A = 25.1A, which agrees with the given total of 25A to within the rounding.

Making Sense of the Answer: *The aluminium core carries* 14.7*A out of the total* 25*A, which is about* 59% *of the cable's current. The four steel wires together carry the remaining* 41%*, split equally as* 2.6*A each. The aluminium is a single wire out of five and carries more current than all four steel wires combined; this is because aluminium is* 5.67 *times more conductive than steel, and the steel wires divide the load equally among themselves, four ways.*

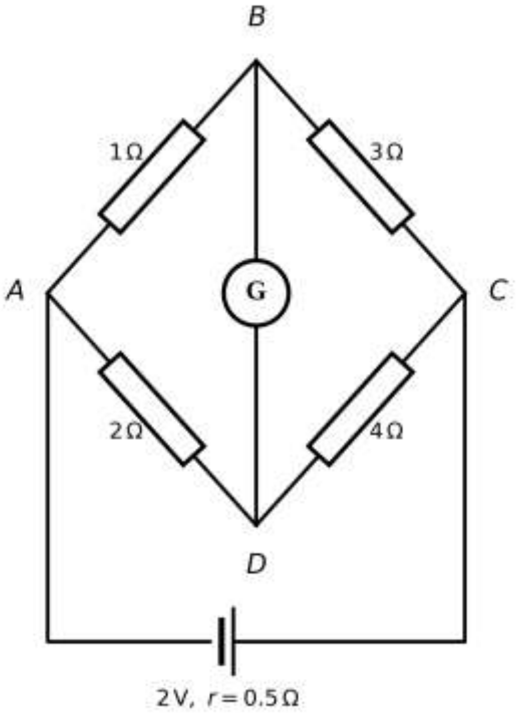
The Physics Behind the Design: *The steel wires are not in the cable to carry current efficiently. They are in the cable to keep it from stretching under its own weight as it hangs between distant pylons. Aluminium is a soft metal that creeps and sags; steel is strong but a poor conductor. The composite cable combines the two properties: the aluminium does most of the conducting and the steel does most of the holding.*

Engineering trade-offs of exactly this kind are why real transmission cables are rarely made of a single material.

Think Like a Physicist: *When a parallel-resistor problem has all the geometric factors (length, area, shape) the same across every branch, do not compute the absolute resistances. Work with the ratio of the remaining quantity (here the resistivity) and let it carry the whole calculation. The geometric factors cancel out of every ratio, and the arithmetic stays close to numbers of order unity rather than scientific notation. This habit, working with ratios rather than absolute values whenever the problem allows, is one of the most powerful tools a physicist has, and not just in circuit problems.*

HOT Example 28

Four resistors of resistances 1Ω, 3Ω, 4Ω, and 2Ω are connected along the four sides of a quadrilateral ABCD in that order (so AB = 1Ω, BC = 3Ω, CD = 4Ω, DA = 2Ω). An accumulator of EMF 2V and internal resistance 0.5Ω is connected across A and C. A galvanometer of resistance 5Ω is connected between B and D. Find the current through the galvanometer and the effective resistance of the network between A and C.



Solution

Try first the methods of this section. The four resistors form two paths from A to C: a top path through B (with AB = 1Ω then BC = 3Ω in series, total 4Ω) and a bottom path through D (with AD = 2Ω then DC = 4Ω in series, total 6Ω). If the network had no galvanometer between B and D, the two paths would be a simple parallel pair, and the equivalent resistance between A and C would be (4Ω)(6Ω)/(4Ω + 6Ω) = 2.4Ω.

But the galvanometer disturbs this simple picture. The resistor ratio AB:BC is 1:3, and the resistor ratio AD:DC is 2:4 = 1:2. The two ratios are **not** equal. So the potential at B (which is shared between AB and BC in ratio 3:1 of the A-to-C voltage drop) is **not** equal to the potential at D (which is shared in ratio 2:1 of the same drop). Because the potentials at B and D are different, current flows through the galvanometer connecting them, and this current alters how the original two paths share the total.

This is the unbalanced Wheatstone bridge, in its simplest form. It cannot be reduced by series and parallel alone (the galvanometer branch creates a fifth connection that breaks the simple parallel topology). It does not have the symmetry of the cube either (the four resistor values are all different from one another). The methods of this section are therefore not enough.

The unbalanced bridge yields cleanly to Kirchhoff's laws, which the next section introduces. We shall return to the exact numerical solution after Kirchhoff is in our hands. For now, note that the diagnosis (unbalanced bridge, not reducible by the methods of this section) is the most important conceptual step. Identifying the right tool for the job is half the physics.

Making Sense of the Anticipation: *If the four resistor values had been in balanced ratio AB:BC = AD:DC (for example, 1Ω, 2Ω, 2Ω, 4Ω rather than 1Ω, 3Ω, 4Ω, 2Ω), the potentials at B and D would have been equal, no current would flow through the galvanometer, and the network would have reduced to a simple parallel pair. The fact that the actual values are unbalanced is what makes the problem non-trivial. The balanced case (galvanometer reads zero) is the working principle of the Wheatstone bridge as a measurement device, and we shall meet it again in a later section.*

Think Like a Physicist: *Recognising which method applies to which network is the most important skill of this section. Use series and parallel reductions whenever they work; they always work fastest. Use symmetry methods when geometric symmetry forces equipotentials. Use Kirchhoff's laws (the next section) as a fallback for everything else. The fallback is the most general tool, but it is also the slowest, so never reach for it first.*

Seven examples and one escape. Series, parallel, symmetry, and meter modification carried us through the first six. The last one stayed beyond our reach. For that we need a sharper tool.

That tool is Kirchhoff's laws, two short statements about conservation in circuits that solve every network in principle by reducing it to a set of simultaneous linear equations. The next section sets them out and then applies them in earnest.

KIRCHHOFF'S LAWS

Every method in this chapter so far has had a limited reach. Ohm's law on its own handles a single resistor. Series and parallel reduction handles any network that can be untangled into nested series and parallel pieces. Symmetry methods handle the special networks whose geometry forces certain nodes to the same potential. Each method is fast within its own territory, and each fails outside it. The unbalanced bridge that closed the last section defeated all three.

This section introduces the method that has no such limit. Kirchhoff's two laws apply to every network without exception, however many sources it contains, however the resistors are wired, and whether or not it has any symmetry. The price for this generality is labour: Kirchhoff's laws turn a circuit into a set of simultaneous equations that must then be solved. They are the method of last resort, to be reached for only when the faster methods fail, but they never fail themselves.

Both laws are conservation statements that you already know from earlier physics, applied now one piece of a circuit at a time. The first law is conservation of charge applied at a junction. The second law is conservation of energy applied around a loop. Before stating them, we need three pieces of vocabulary.

Junctions, branches, and loops

A junction (also called a **node**) *is a point in a circuit where three or more wires meet.* At a junction the current has a choice of paths, so the current can **split** or **merge** there. A point where only two wires meet is **not** a junction; it is just a point along a single wire, and the same current flows through it on both sides.

A branch *is a single path between two junctions, along which the current is the same all the way.* A branch may contain several components in series (resistors, cells, or both), but because it has no junction in its interior, the current cannot change anywhere along it. Each branch carries one current, and that current is one of the unknowns we shall solve for.

A loop *is any closed conducting path in the network:* a path that starts at one point, travels through some branches, and returns to its starting point without retracing any branch. A network with several junctions usually has several possible loops, and part of the skill of applying Kirchhoff's laws is choosing a convenient set of them.

Kirchhoff's current law

Kirchhoff's current law, abbreviated **KCL**, states that: *The algebraic sum of the currents at any junction is zero.* Said more plainly: *the total current flowing into a junction equals the total current flowing out of it.* Whatever charge arrives at the junction each second must leave it again the same second.

The reason is conservation of charge. A junction is just a point, and a point has no room to store charge. Charge cannot pile up there and cannot be created or destroyed there, so every coulomb that flows in must immediately flow out along one path or another. If this were not so, charge would accumulate at the junction without limit, which never happens in a steady circuit.

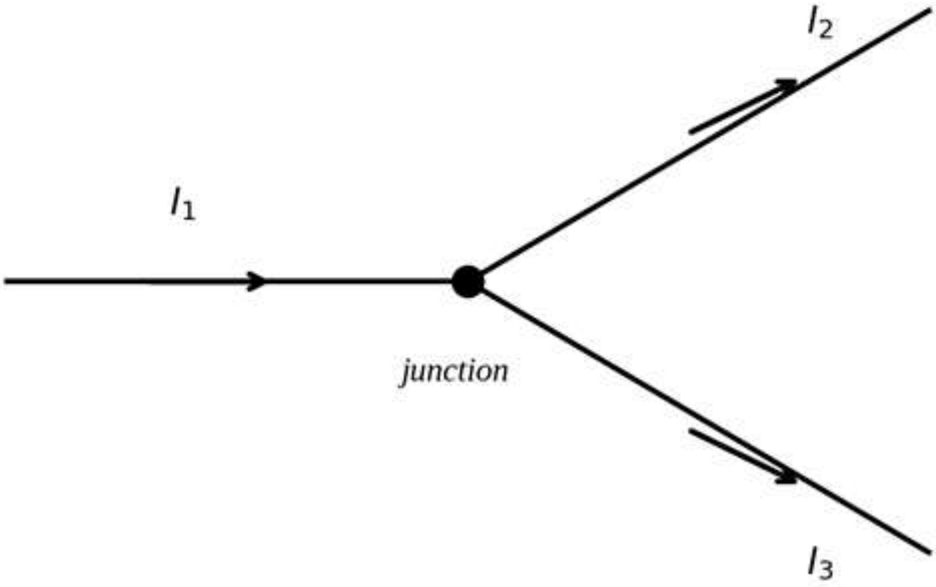


Figure: ***A junction where three wires meet.*** *Current I_1 flows in; currents I_2 and I_3 flow out. Kirchhoff's current law requires $I_1 = I_2 + I_3$, the current in equals the current out.*

For the junction in the figure, with I_1 flowing in and I_2 and I_3 flowing out, the law reads:

$$\mathbf{I_1 = I_2 + I_3}$$

To apply the law systematically, adopt a sign convention: *count currents flowing into the junction* as **positive** and *currents flowing out* as **negative**. The algebraic sum is then zero. With this convention the same junction reads:

$$\mathbf{I_1 - I_2 - I_3 = 0},$$

which rearranges to the same statement. The two forms are identical; use whichever is clearer for the problem at hand.

Kirchhoff's voltage law

Kirchhoff's voltage law, abbreviated **KVL**, states that *around any closed loop in a network, the algebraic sum of the EMFs equals the algebraic sum of the potential drops IR.* In symbols, taken around one complete loop:

$$\mathbf{\Sigma E = \Sigma IR}$$

The reason is conservation of energy. Follow one coulomb of charge once around the loop and back to where it started. Along the way the EMF sources do work on the coulomb (raising its energy) and the resistors take energy from it (lowering its energy as heat). Since the coulomb returns to exactly where it began, its energy must be back to its starting value, so the total energy gained from the sources must equal the total energy lost in the resistors. Energy per coulomb is voltage, so the sum of the EMFs equals the sum of the IR drops.

This is the same coulomb-tracing argument used earlier in the chapter to derive the terminal-voltage relation for a single cell. Kirchhoff's voltage law is that argument promoted to a general principle and applied to any loop, not just a single-cell circuit.

Applying the law requires a sign convention, because both the EMFs and the IR drops can act in either direction around the loop. Begin by choosing a direction to traverse the loop, either clockwise or anticlockwise; the choice is free, but once made it must be held for the whole loop. Then:

An ***EMF*** *counts as positive when the chosen loop direction passes through the source from its negative terminal to its positive terminal (the direction in which the source raises the potential). It counts as negative when the loop direction passes through it the other way.*

An ***IR drop*** *counts as positive when the chosen loop direction matches the assumed direction of the current through that resistor. It counts as negative when the loop direction opposes the assumed current. The product IR uses the branch current I with its sign, so a branch current that later turns out negative takes care of itself automatically.*

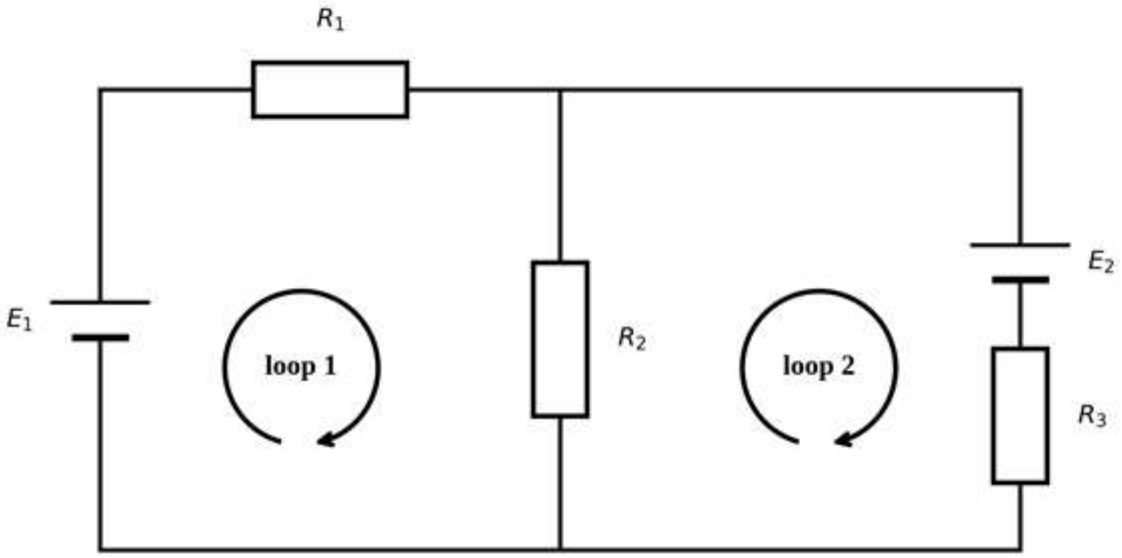


Figure***: A two-loop network.*** *Each loop is traversed in a chosen direction (here both clockwise, shown by the circular arrows). Kirchhoff's voltage law is written once for each loop, and Kirchhoff's current law once for each independent junction. Together they give enough equations to solve for all the branch currents.*

The procedure, step by step

Solving any network by Kirchhoff's laws follows four steps. Each is explained below with what to do in practice.

Step 1: Label every branch current with an assumed direction

Mark a current symbol and an arrow on every branch. The direction of each arrow is a guess; you do not need to guess correctly. If a current turns out to flow the other way, its value will simply come out negative at the end, and that negative sign is the circuit telling you the true direction is opposite to your guess. Use Kirchhoff's current law at the junctions immediately to reduce the number of distinct unknowns: if three branches meet at a junction, the third current is fixed once the first two are named.

Step 2: Write the current law at each independent junction

A network with $\mathbf{n}$ junctions gives $\mathbf{n-1}$ independent current-law equations. (The last junction gives no new information, because charge conservation across the whole network makes it follow automatically from the others.) Write the current law at all but one of the junctions.

Step 3: Write the voltage law around enough independent loops

Choose a set of loops so that every branch is included in at least one loop, and write Kirchhoff's voltage law for each, using the sign convention above. You need exactly as many loop equations as there are remaining unknown currents after the junction equations have been used. A loop is independent if it includes at least one branch that no previous loop has used.

Step 4: Solve the simultaneous equations

The junction and loop equations together form a system of linear simultaneous equations in the unknown currents. Solve them. A negative answer for any current means the actual current in that branch flows opposite to the direction you assumed in Step 1; the magnitude is still correct. Always finish by checking the solution in one equation you did not use to find it, or by confirming that the power delivered by the sources equals the power dissipated in the resistors.

Parallel cells with different EMFs

We can now settle a question left open earlier in the chapter: *what happens when two cells of different EMF are connected in parallel across a load?* The equal-EMF case was simple, but the unequal case needs Kirchhoff's laws, and it makes an instructive first application of the procedure.

Two cells are connected in parallel across a load resistor R. The first cell has EMF E_1 and internal resistance r_1; the second has EMF E_2 and internal resistance r_2. The figure shows the circuit, with the cell currents I_1 and I_2 both assumed to flow out of their cells toward the load, and the load current I flowing down through R.

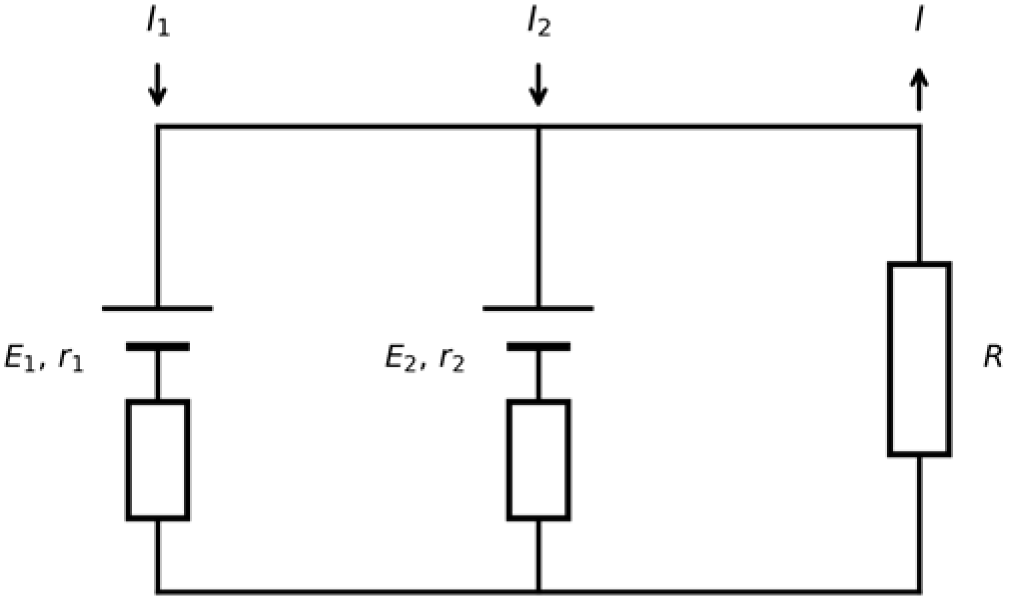


Figure: ***Two cells of different EMF in parallel, driving a common load R.*** *Cell 1 supplies current I_1, cell 2 supplies current I_2, and the load carries $I = I_1 + I_2$.*

Apply the current law (KCL) at the top junction, where both cell currents meet and the load current leaves:

$$I = I_1 + I_2$$

Apply the voltage law (KVL) around the loop containing cell 1 and the load. The EMF E_1 drives the current; it is spent on the internal-resistance drop $I_1 r_1$ inside cell 1 and the load drop IR across R:

$$E_1 = I_1 r_1 + IR$$

Apply the voltage law around the loop containing cell 2 and the load:

$$E_2 = I_2 r_2 + IR$$

These are three equations in the three unknowns I_1, I_2, and I. Solve them.

From the two loop equations, express each cell current in terms of I:

$$I_1 = \frac{E_1 - IR}{r_1}, \quad I_2 = \frac{E_2 - IR}{r_2}$$

Substitute both into the junction equation $I = I_1 + I_2$ and collect the terms in I:

$$I = \frac{E_1 - IR}{r_1} + \frac{E_2 - IR}{r_2} = \frac{r_2(E_1 - IR) + r_1(E_2 - IR)}{r_1 r_2} = \frac{E_1 r_2 + E_2 r_1 - IR(r_2 + r_1)}{r_1 r_2}$$

$$I r_1 r_2 = E_1 r_2 + E_2 r_1 - IR(r_2 + r_1);\ I\big(r_1 r_2 + R(r_2 + r_1)\big) = E_1 r_2 + E_2 r_1$$

Hence:

$$\mathbf{I = \frac{E_1 r_2 + E_2 r_1}{R(r_1 + r_2) + r_1 r_2}}$$

This is the load current for any two cells in parallel. The formula is worth reading for its structure rather than memorising. Compare it with the single-source result, load current equals EMF over (load plus internal resistance), and it suggests that the parallel pair behaves like one equivalent cell. Indeed it does.

Rewriting the result in the single-source form $I = \frac{E_{eq}}{R + r_{eq}}$ identifies the equivalent EMF and equivalent internal resistance of the parallel pair:

$$\mathbf{E_{eq} = \frac{E_1 r_2 + E_2 r_1}{r_1 + r_2}, \quad r_{eq} = \frac{r_1 r_2}{r_1 + r_2}}$$

The equivalent internal resistance is just r_1 and r_2 in parallel, as expected for two resistances side by side. The equivalent EMF is a weighted average of the two EMFs, each weighted by the other cell's internal resistance.

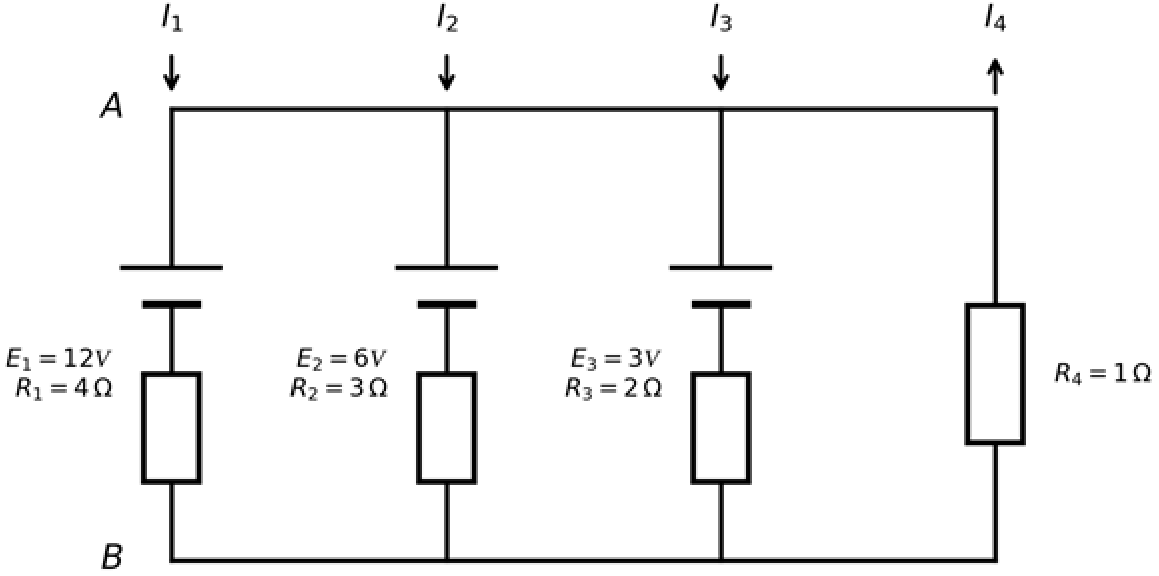
Check the formula against the equal-cell case. Put $E_1 = E_2 = E$ and $r_1 = r_2 = r$. Then $E_{eq} = \frac{Er+Er}{2r} = E$ and $r_{eq} = \frac{r^2}{2r} = \frac{r}{2}$, which recovers exactly the equal-cell result found earlier in the chapter. The general formula contains the simple one as a special case, which is a good sign that it is right.

A practical warning: *Connecting cells of different EMF in parallel is generally a bad idea. Even with no external load connected, the two cells form a closed loop by themselves, and the higher-EMF cell drives a current backwards through the lower-EMF cell, charging it and wasting energy as heat in both internal resistances. Worked Example 29 shows this directly: one of the two cell currents comes out negative, meaning that cell is being charged by the other rather than helping to drive the load. Parallel cells should have matched EMFs.*

Five examples follow. Each is solved by the same four-step procedure, and several contain a current that comes out negative, so watch how the sign is read at the end.

HOT Example 29

Three cells are connected in parallel with a load resistor across a pair of common rails, as shown. The cells are $E_1 = 12V$ with $R_1 = 4\Omega$, $E_2 = 6V$ with $R_2 = 3\Omega$, and $E_3 = 3V$ with $R_3 = 2\Omega$. The load is $R_4 = 1\Omega$. Find the current in each of the four branches.



Solution

There are only two junctions, the top rail A and the bottom rail B. Let V be the potential difference between them, $V = V_A - V_B$. This single unknown is the key: once V is known, every branch current follows from Ohm's law applied to that branch.

For each cell branch, the terminal voltage equals the EMF minus the internal drop:

$$\mathrm{V} = \mathrm{E_i} - \mathrm{I_i R_i} \text{ or } \mathrm{I_i} = \frac{(\mathrm{E_i} - \mathrm{V})}{\mathrm{R_i}}$$

For the load branch:

$$\mathrm{I_4} = \frac{\mathrm{V}}{\mathrm{R_4}}$$

The KCL at A says the three cell currents flowing in equal the load current flowing out:

$$\frac{\mathrm{E_1 - V}}{\mathrm{R_1}} + \frac{\mathrm{E_2 - V}}{\mathrm{R_2}} + \frac{\mathrm{E_3 - V}}{\mathrm{R_3}} = \frac{\mathrm{V}}{\mathrm{R_4}}$$

$$\frac{12 - \mathrm{V}}{4} + \frac{6 - \mathrm{V}}{3} + \frac{3 - \mathrm{V}}{2} = \frac{\mathrm{V}}{1}$$

Solving for V:

$$\mathrm{V} = 3.12\mathrm{V}$$

Now back-substitute to get each branch current:

$$\mathrm{I_1} = \frac{12\mathrm{V} - 3.12\mathrm{V}}{4\Omega} = 2.22\mathrm{A}$$

$$\mathrm{I_2} = \frac{6\mathrm{V} - 3.12\mathrm{V}}{3\Omega} = 0.96\mathrm{A}$$

$$\mathrm{I_3} = \frac{3\mathrm{V} - 3.12\mathrm{V}}{2\Omega} = -0.06\mathrm{A}$$

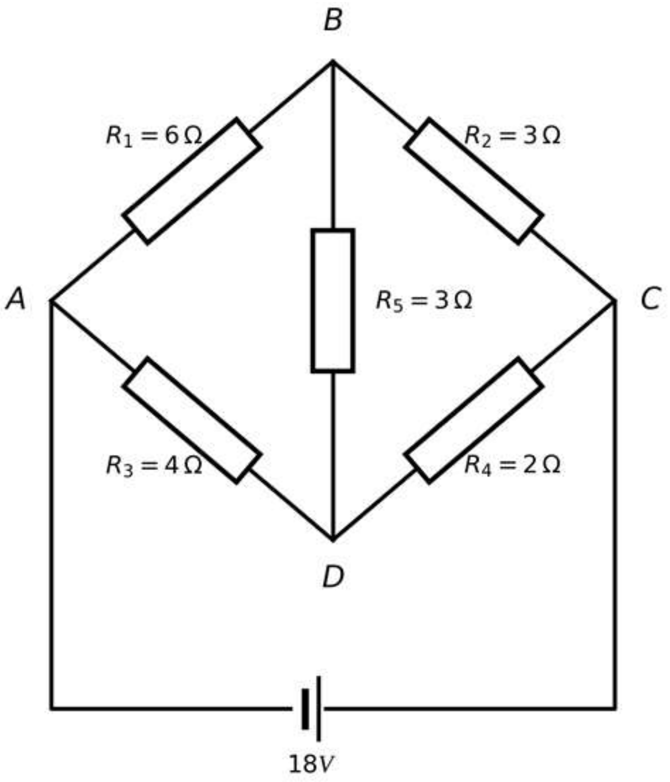
$$\mathrm{I_4} = \frac{3.12\mathrm{V}}{1\Omega} = 3.12\mathrm{A}$$

Making Sense of the Answer: ***Check the current law:*** $2.22A + 0.96A + (-0.06A) = 3.12A$, *which equals* I_4, *as it must. The interesting result is* I_3, *which came out negative. The assumed direction for* I_3 *was upward, out of the 3V cell. The negative sign means the actual current is downward: the 3V cell is too weak to push against the common rail voltage of 3.12V, so instead of delivering current it receives a small current of 0.06A and is being charged by the other two cells. This is exactly the behaviour the parallel-cell warning predicted.*

Think Like a Physicist: *This network had only two junctions, so a single unknown (the rail voltage* V*) captured the whole problem. Whenever a network reduces to two nodes joined by several parallel branches, the node-voltage shortcut used here is faster than writing a separate current symbol for every branch and solving the full system. Recognising when a circuit has this two-node form can save a great deal of algebra.*

HOT Example 30

In the bridge network shown, four resistors form a diamond ABCD: $\mathrm{R_1} = 6\Omega$ on AB, $\mathrm{R_2} = 3\Omega$ on BC, $\mathrm{R_3} = 4\Omega$ on AD, and $\mathrm{R_4} = 2\Omega$ on DC. A fifth resistor $\mathrm{R_5} = 3\Omega$ bridges B to D. A 18V source of negligible internal resistance is connected across A and C. Find the current in all five branches.



Solution

Before grinding through Kirchhoff's laws, look at the arm resistances. Going down the upper path A → B → C, the ratio of the two arms is $\frac{R_1}{R_2} = \frac{6\Omega}{3\Omega} = 2$. Going down the lower path A → D → C, the ratio is $\frac{R_3}{R_4} = \frac{4\Omega}{2\Omega} = 2$. The two ratios are equal. This is the balance condition of the bridge, and it has a striking consequence.

When the arm ratios are equal, the potentials at B and D are equal. To see why, note that the upper path divides the full A-to-C voltage between R_1 and R_2 in the ratio 6 to 3, while the lower path divides the same voltage between R_3 and R_4 in the ratio 4 to 2. Both ratios are 2 to 1, so the point B sits at the same fraction of the way down from A to C as the point D does. They are at the same potential.

Since $V_B = V_D$, the voltage across the bridge resistor R_5 is zero, and so by Ohm's law the current through it is zero:

$$I_5 = \frac{V_B - V_D}{R_5} = 0$$

With no current in the bridge, the network becomes two simple series paths in parallel. The upper path has resistance $R_1 + R_2 = 6\Omega + 3\Omega = 9\Omega$, carrying current $I_{upper} = \frac{18V}{9\Omega} = 2A$.

The lower path has resistance $R_3 + R_4 = 4\Omega + 2\Omega = 6\Omega$, carrying current $I_{lower} = \frac{18V}{6\Omega} = 3A$.

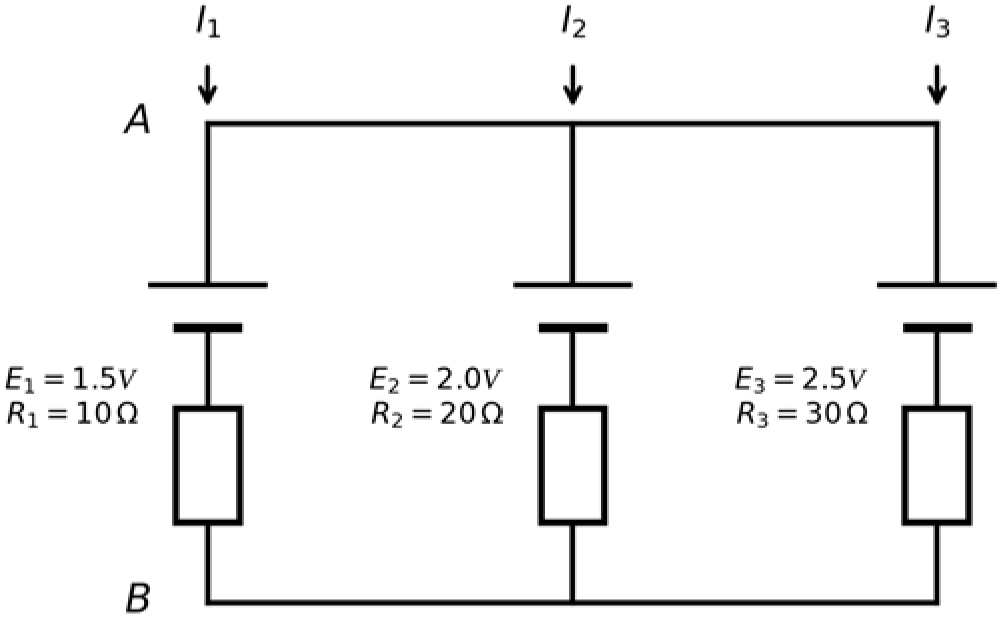
So the five branch currents are: 2A through R_1, 2A through R_2 (the whole upper path carries the same 2A), 3A through R_3, 3A through R_4 (the whole lower path carries 3A), and zero through the bridge resistor R5.

Making Sense of the Answer: *The total current drawn from the source is $2A + 3A = 5A$, which leaves the source, splits at A into the two paths, and recombines at C. The zero bridge current is the heart of the Wheatstone bridge as a measuring instrument: when the arms are balanced, a galvanometer placed across BD reads exactly zero, and that null reading can be detected with great precision. The next section builds an entire measurement technique on this fact.*

Think Like a Physicist: *Always test a bridge for balance before reaching for the full Kirchhoff machinery. The balance check is one division: compare $\frac{R_1}{R_2}$ with $\frac{R_3}{R_4}$. If they are equal, the bridge branch carries no current and the network collapses to two series paths in parallel, which you can solve in your head. Only when the ratios differ (the unbalanced case of Example 33) do you need the full simultaneous equations.*

HOT Example 31

Three cells of negligible internal resistance are connected in parallel between two points A and B, each cell in series with its own resistor: $E_1 = 1.5V$ with $R_1 = 10\Omega$, $E_2 = 2.0V$ with $R_2 = 20\Omega$, and $E_3 = 2.5V$ with $R_3 = 30\Omega$. All three positive terminals face A. Find the current in each resistor.



Solution

As in Example 29, there are only two junctions, so let $V = V_A - V_B$ be the single unknown. Each branch current, assumed to flow from B up to A, is $I_i = (E_i - V)/R_i$.

Since A and B are the only nodes and these three branches are the only paths between them, the current law requires the three currents into A to sum to zero:

$$\frac{1.5 - V}{10} + \frac{2.0 - V}{20} + \frac{2.5 - V}{30} = 0$$

Solve for V:

$$V = 1.818V$$

Substitute back to find each current:

$$I_1 = \frac{1.5V - 1.818V}{10\Omega} = -0.032A = -32mA$$

$$I_2 = \frac{2.0V - 1.818V}{20\Omega} = 0.0091A = 9.1mA$$

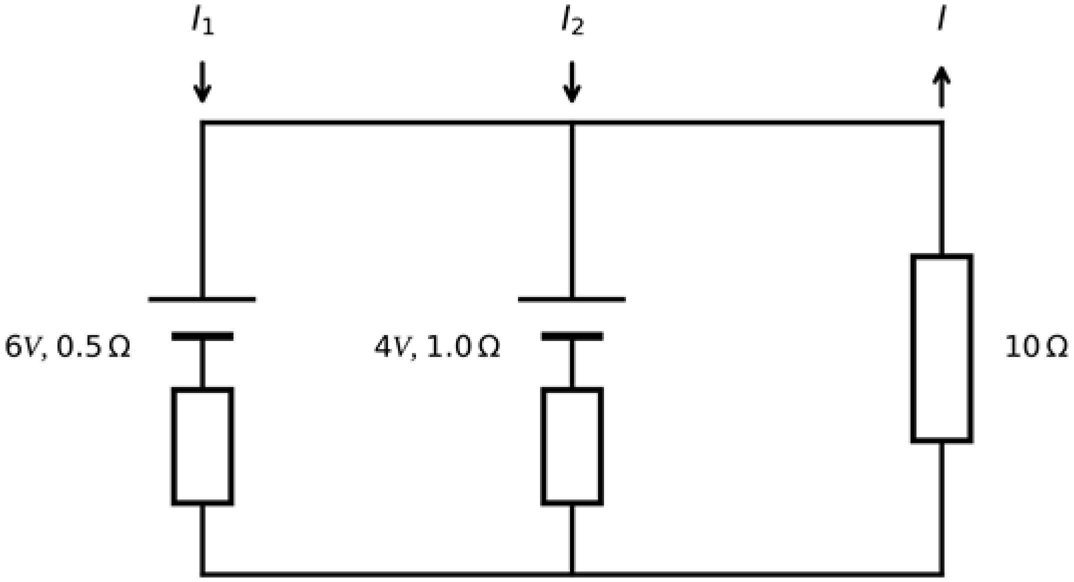
$$I_3 = \frac{2.5V - 1.818V}{30\Omega} = 0.0227A = 22.7mA$$

Making Sense of the Answer: ***Check:*** *$-32mA + 9.1mA + 22.7mA \approx 0mA$, confirming the current law. The current I_1 came out negative. The 1.5V cell has the lowest EMF of the three, lower than the common rail voltage of 1.818V that the other two establish, so it cannot push current out into the rail. Instead it absorbs current: the actual flow in branch 1 is downward at 32mA, and the 1.5V cell is being charged by the 2.0V and 2.5V cells. The higher-EMF cells, being above the rail voltage, deliver current as expected.*

Think Like a Physicist: *The common rail voltage 1.818V is a kind of compromise between the three EMFs, pulled toward the cells with the smaller series resistance (which have more influence). Any cell whose EMF lies above this compromise voltage delivers current; any cell whose EMF lies below it absorbs current. This is the general fate of mismatched cells wired in parallel, and the reason the practice is discouraged.*

REAL Example 32

Two cells are connected in parallel across a 10Ω load. The first cell has $E_1 = 6V$, $r_1 = 0.5\Omega$; the second has $E_2 = 4V$, $r_2 = 1.0\Omega$. Find (a) the load current; (b) the current through each cell; (c) the equivalent EMF and internal resistance of the parallel combination.



Solution

(a) Start with the equivalent source, using the formulas derived in this section. The equivalent internal resistance is r_1 and r_2 in parallel:

$$r_{eq} = \frac{r_1 r_2}{r_1 + r_2} = \frac{0.5\Omega \times 1.0\Omega}{0.5\Omega + 1.0\Omega} = 0.333\Omega$$

The equivalent EMF is the weighted average:

$$E_{eq} = \frac{E_1 r_2 + E_2 r_1}{r_1 + r_2} = \frac{6V \times 1.0\Omega + 4V \times 0.5\Omega}{0.5\Omega + 1.0\Omega} = 5.33V$$

(b) The load current follows from the equivalent source driving the 10Ω load:

$$I = \frac{E_{eq}}{R + r_{eq}} = \frac{5.33V}{10\Omega + 0.333\Omega} = 0.516A$$

(c) The voltage across the load is $V = IR = 0.516A \times 10\Omega = 5.16V$. This same voltage appears across each cell's terminals (they are all in parallel). The current from each cell is then $I_i = (E_i - V)/r_i$:

$$I_1 = \frac{6V - 5.16V}{0.5\Omega} = \frac{0.84}{0.5} = 1.68A$$

$$I_2 = \frac{4V - 5.16V}{1.0\Omega} = \frac{-1.16}{1.0} = -1.16A$$

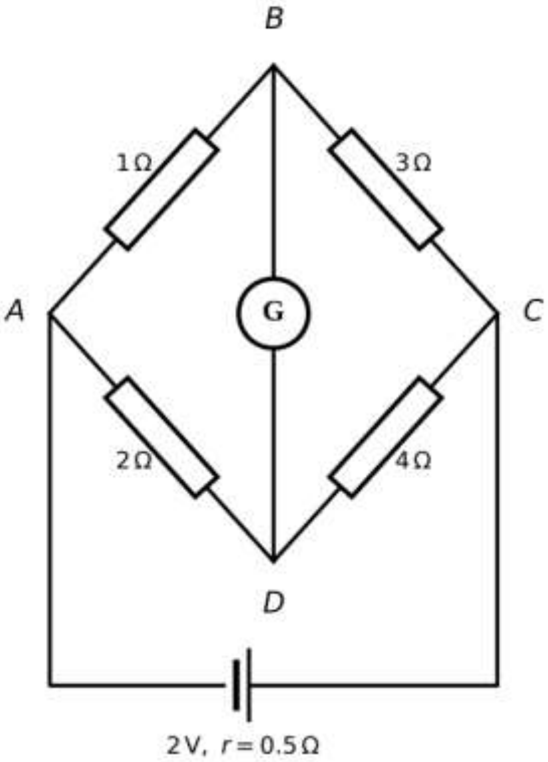
Making Sense of the Answer: ***Check the current law:*** $1.68 + (-1.16) = 0.52A$, *which matches the load current* $0.516A$ *to within rounding. The second cell's current is negative. Its EMF of 4V is below the common terminal voltage of 5.16V that the stronger 6V cell establishes, so the 4V cell cannot deliver current. Instead it absorbs 1.16A: it is being charged by the 6V cell. The 6V cell is supplying 1.68A, of which 0.52A goes to the load and 1.16A goes uselessly into charging its partner. This is the concrete cost of mismatched parallel cells: more than two thirds of the strong cell's output is wasted on the weak one.*

Think Like a Physicist: *The equivalent-source method turned a three-unknown problem into two short formulas plus one Ohm's-law division. Once the parallel pair is replaced by its single equivalent EMF and internal resistance, the rest of the circuit sees nothing but an ordinary cell. Replacing a complicated sub-network by its equivalent source is one of the most useful simplifications in circuit analysis, and it is worth* *doing whenever a group of sources and resistors feeds the rest of the circuit through a single pair of terminals.*

We can now finish the problem left unsolved at the end of the previous section.

HOT Example 33

Four resistors lie along the sides of a quadrilateral ABCD: $AB = 1\Omega$, $BC = 3\Omega$, $CD = 4\Omega$, $DA = 2\Omega$. A galvanometer of resistance 5Ω connects B to D. An accumulator of EMF 2V and internal resistance 0.5Ω is connected across A and C. Find the galvanometer current and the resistance of the network between A and C.



Solution

First confirm that the bridge is unbalanced, so that the shortcut of Example 30 does not apply. The upper ratio is $\frac{AB}{BC} = \frac{1}{3}$; the lower ratio is $\frac{AD}{DC} = \frac{2}{4} = \frac{1}{2}$. These are not equal, so $V_B \neq V_D$ and current does flow through the galvanometer. We use the node-voltage form of Kirchhoff's laws.

Take C as the zero of potential ($V_C = 0$). The unknowns are the potentials V_A, V_B, V_D.

Write the current law at each of the three nodes, expressing every branch current as a potential difference over a resistance.

At node A, the accumulator delivers current $(2 - V_A)/0.5$ into the node, and this leaves through AB and AD:

$$\frac{2 - V_A}{0.5} = \frac{V_A - V_B}{1} + \frac{V_A - V_D}{2}$$

$$7V_A - 2V_B - V_D = 8 \ldots \text{(i)}$$

At node B, current arrives from A and leaves through BC and the galvanometer BD:

$$\frac{V_A - V_B}{1} = \frac{V_B}{3} + \frac{V_B - V_D}{5}$$

$$15V_A - 23V_B + 3V_D = 0 \ldots \text{(ii)}$$

At node D, current arrives from A and from the galvanometer, and leaves through DC:

$$\frac{V_A - V_D}{2} + \frac{V_B - V_D}{5} = \frac{V_D}{4}$$

$$10V_A + 4V_B - 19V_D = 0 \ldots (iii)$$

These are three linear equations in the three unknown potentials. Solving with a calculator gives:

$$V_A = 1.654V, \quad V_B = 1.226V, \quad V_D = 1.129V$$

The galvanometer current flows from B to D (since $V_B > V_D$):

$$I_g = \frac{V_B - V_D}{5} = \frac{1.226V - 1.129V}{5\Omega} = 0.0195A = 19.5mA$$

For the resistance between A and C, first find the total current the accumulator drives:

$$I = \frac{2 - V_A}{0.5} = \frac{2V - 1.654V}{0.5\Omega} = 0.692A$$

The network resistance between A and C (not counting the accumulator's own internal resistance) is the terminal voltage divided by this current:

$$R_{AC} = \frac{V_A}{I} = \frac{1.654V}{0.692A} = 2.39\Omega$$

Making Sense of the Answer: *The galvanometer carries about* $19.5mA$*, a small but definitely non-zero current, exactly as the unbalanced arm ratios predicted. Had the bridge been balanced, this current would have been zero and the network resistance would have been the simple parallel combination of the two arm paths. The unbalanced current is what makes this problem need the full Kirchhoff treatment. As a sanity check,* $I = 2/(R_{AC} + 0.5) = 2/2.89 = 0.692A$*, which matches the total current found above.*

Think Like a Physicist: *The node-voltage method used here is often the quickest form of Kirchhoff's laws for a network with few nodes. Instead of naming a separate current for every branch and juggling both laws, name the node potentials, write the current law once per node with each current expressed as a voltage difference over a resistance, and the voltage law takes care of itself. Three nodes gave three equations and the whole bridge fell out. The branch-current method and the node-voltage method always agree; choose whichever gives fewer unknowns for the network in front of you.*

Five worked examples and five networks now lie behind us.

One particular network in this section, the balanced bridge of Example 30, hid a special gift: when its arms are in equal ratio, the bridge branch carries no current at all. That null condition is the basis of one of the most precise measuring instruments in the elementary laboratory. The next section takes the balanced Wheatstone bridge and its practical cousin, the metre bridge, and turns the null condition into a method for measuring an unknown resistance to high accuracy.

THE WHEATSTONE BRIDGE AND THE METRE BRIDGE

Direct measurement always affects the circuit being measured. An ohmmeter sends a small test current through the unknown resistor and reads back the voltage that develops; a voltmeter taps a tiny current out of the circuit it is supposed to be observing. Both instruments are reasonably good, but neither is exact, because the current they draw alters the very quantity they are trying to read.

A **null method** avoids this fault. The instrument is wired so that, when correctly adjusted, no current flows through the detector. There is then nothing being drawn from the circuit, and the balance condition gives the unknown quantity in terms of known ones. The price you pay is a small loss of speed: you must adjust until the detector reads zero. The gain is precision better than any direct meter reading.

This section develops the two most important null-method instruments for resistance measurement. The Wheatstone bridge in its laboratory form uses four resistors and a galvanometer; the metre bridge is the same circuit rebuilt around a uniform resistance wire, which makes the balance length easy to read directly from a centimetre scale. The closing subsection treats the unbalanced bridge, showing how the galvanometer current changes when the balance condition is not met. The unbalanced case needs the full Kirchhoff machinery of the previous section.

The Wheatstone bridge

Four resistors R_1, R_2, R_3, R_4 are arranged as the four sides of a diamond, with corners A, B, C, D. An EMF source is connected across one diagonal AC, and a sensitive galvanometer is connected across the other diagonal BD. The next figure shows the layout.

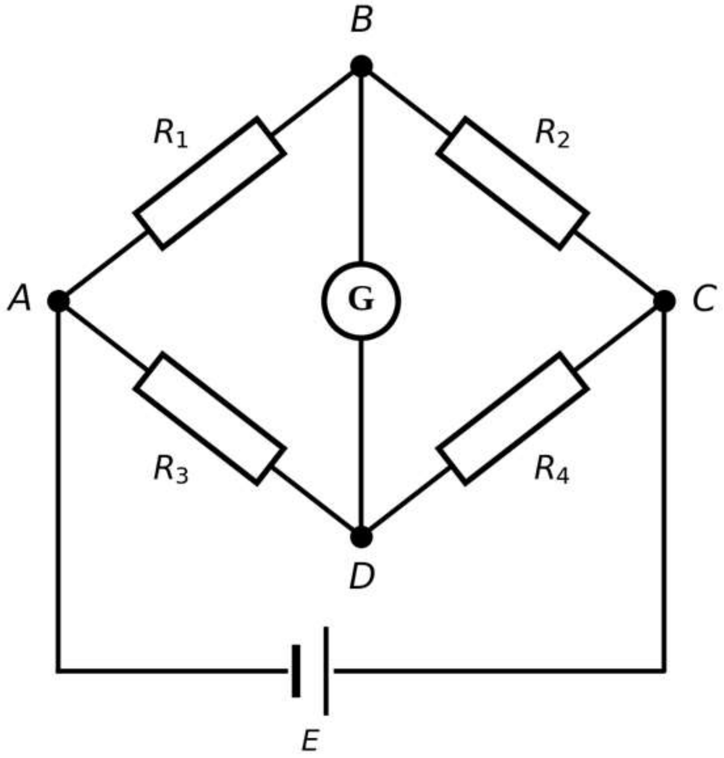


Figure: ***The Wheatstone bridge.*** *Four resistors R_1, R_2, R_3, R_4 sit on the sides of the diamond ABCD. The EMF E drives current in at A and out at C. A sensitive galvanometer G connects B to D and detects any imbalance.*

Three of the four resistors are known precisely. The fourth, the unknown, is what we are trying to measure. One of the three known resistors is a calibrated variable resistance box. Adjust this variable resistance until the galvanometer reads exactly zero. When that happens, the bridge is said to be balanced, and the unknown resistance follows from the balance condition we are about to derive.

At balance, no current flows through the galvanometer. The current must therefore flow only along the two parallel paths $A \to B \to C$ through R_1 and R_2, and $A \to D \to C$ through R_3 and R_4. Call these currents I_{12} (through R_1 and R_2 in series) and I_{34} (through R_3 and R_4 in series).

Zero galvanometer current also means $V_B = V_D$ (no current flows through R_g, so by Ohm's law applied to the galvanometer there is zero voltage across it). This gives us two equations. The voltage drop across R_1 (from A to B) equals the voltage drop across R_3 (from A to D):

$$I_{12}R_1 = I_{34}R_3$$

The voltage drop across R_2 (from B to C) equals the voltage drop across R_4 (from D to C):

$$I_{12}R_2 = I_{34}R_4$$

Divide the first equation by the second; the currents cancel and we are left with the balance condition:

$$\mathbf{\frac{R_1}{R_2} = \frac{R_3}{R_4}}$$

This is the working equation of the Wheatstone bridge. If three of the four resistances are known, the fourth is determined. For instance, if R_1 is the unknown:

$$R_1 = \frac{R_3}{R_4} \times R_2$$

The precision of the measurement is the precision of the three known resistors plus the sensitivity of the galvanometer to small off-balance currents. The EMF source need not be precisely known: the balance condition does not involve E. This is one reason the method is so accurate. The voltage of the driving battery may drift, the temperature of the room may change, even the galvanometer's own resistance does not enter the answer. Only the three known resistors set the precision.

The metre bridge

The Wheatstone bridge with four resistors is theoretically perfect but practically clumsy. To balance it, you have to keep adjusting a calibrated variable resistance box, which is bulky and expensive. A clever rearrangement of the same circuit replaces two of the resistors with a uniform resistance wire one metre long, and replaces the variable resistance with a sliding contact (jockey) that touches the wire at any chosen position. This is the metre bridge.

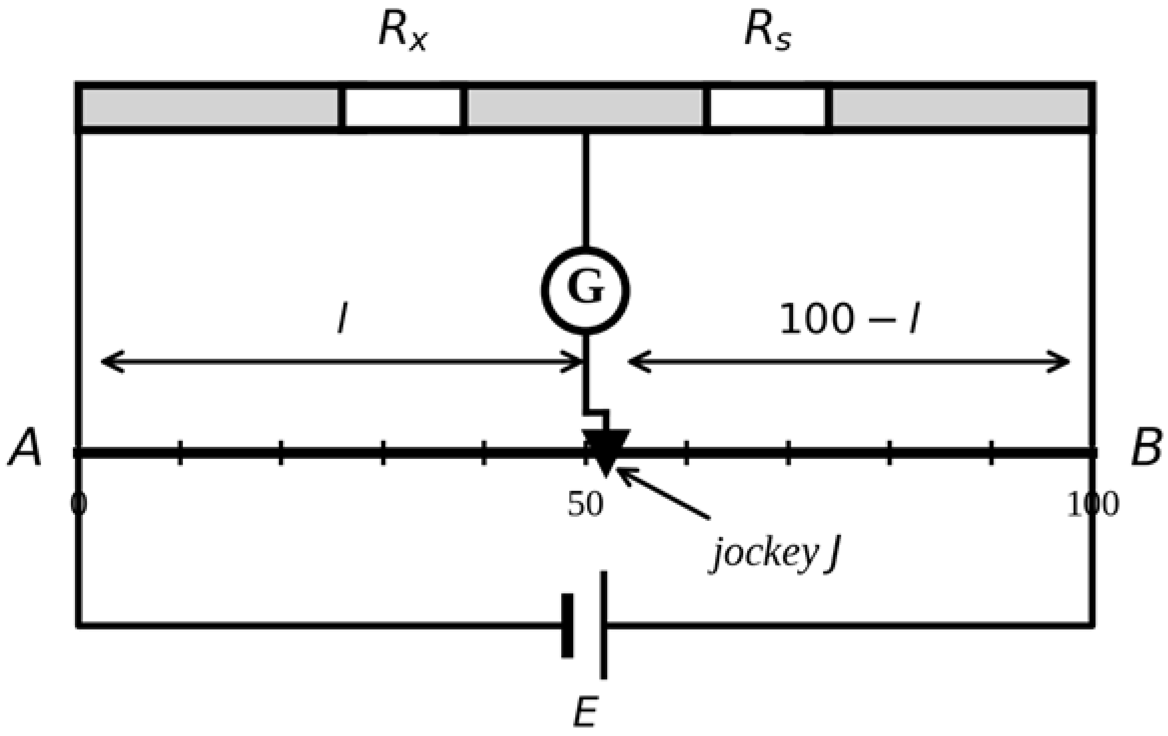


Figure: ***The metre bridge.*** *A uniform 1m resistance wire is stretched between two metallic strips A and B. The unknown R_x is clamped into one gap; a known standard resistor R_s is clamped into the other. A galvanometer G is connected between the middle strip and a sliding jockey J that can touch the wire at any position. At balance, the jockey is at distance l from A. The driving EMF E sits below the wire.*

The unknown resistance R_x sits in one gap and a known standard R_s sits in the other. The two segments of the metre wire (from A to the jockey, and from the jockey to B) play the roles of R_3 and R_4 of the four-resistor bridge.

Why a uniform wire is so useful: because the wire has uniform cross-section and uniform material, its resistance is proportional to its length.

$$R = \frac{\rho l}{A}; \text{for constant } \rho \text{ and } A, R = \text{constant} \times l, \text{and therefore:}$$

$$R \propto l$$

So the resistance of the segment from A to the jockey at position l centimetres is proportional to l, and the resistance of the segment from the jockey to B is proportional to $(100 - l)$ centimetres. The constant of proportionality (the resistance per centimetre) is the same for both segments and cancels out when we take their ratio.

Move the jockey along the wire until the galvanometer reads zero. At that position the bridge is balanced. Apply the balance condition from the four-resistor analysis:

$$\frac{R_x}{R_s} = \frac{R_l}{R_{100-l}} = \frac{l}{100 - l}$$

(both l and $100 - l$ in centimetres). Solving for the unknown:

$$\mathbf{R_x = \frac{l}{100 - l} \times R_s}$$

The balance length l is read directly off a centimetre scale running underneath the wire, so a single measurement of length (plus the known standard resistor) gives the unknown resistance. No variable resistance box is needed, which is the practical advantage of the metre bridge over its laboratory cousin.

Practical tip on choosing the standard: *For best accuracy, choose the standard R_s so that the balance point falls near the middle of the wire. Reading l with a metre rule has the same absolute error wherever the jockey sits (a fraction of a millimetre, say), but the fractional error in $l/(100 - l)$ is smallest when l and $100 - l$ are both large. If the unknown is small (a few ohms), use a small standard. If it is large, use a large standard. The aim is to keep the balance length between about 40 cm and 60 cm.*

The unbalanced Wheatstone bridge

What happens when the bridge is not balanced? The galvanometer no longer reads zero; it deflects by an amount that depends on the four bridge resistors, the galvanometer resistance, and the supply EMF. The deflection is not the unknown we want, but it tells us how far the bridge is from balance and which way to adjust. Finding the size of the galvanometer current in an unbalanced bridge needs the full method of the previous section, Kirchhoff's laws.

In an unbalanced bridge five currents flow: four in the bridge arms and one in the galvanometer. The five currents are linked by KCL at the two junctions B and D and by KVL around two independent loops. Solving the resulting system of equations gives all five currents, including the galvanometer current I_g of greatest interest.

The balanced bridge is a special case of this calculation: if the balance condition R1/R2 = R3/R4 holds, the Kirchhoff equations yield $I_g = 0$ automatically. The balance condition is the algebraic statement of that vanishing. Conversely, any non-zero galvanometer current is the bridge telling you it is unbalanced; the size of I_g measures by how much.

Worked Example 36 demonstrates the full Kirchhoff treatment for an unbalanced bridge with specific numerical values. The node-voltage method works particularly well here, and we shall use it.

Three examples follow.

HOT Example 34

In a metre bridge experiment, two unknown resistors R_1 and R_2 are placed in the left and right gaps respectively. The first null point is found at 25 cm from end A. A 12Ω resistor is then connected in parallel with R_2, and the null point shifts to 40 cm from A. Find R_1 and R_2.

Solution

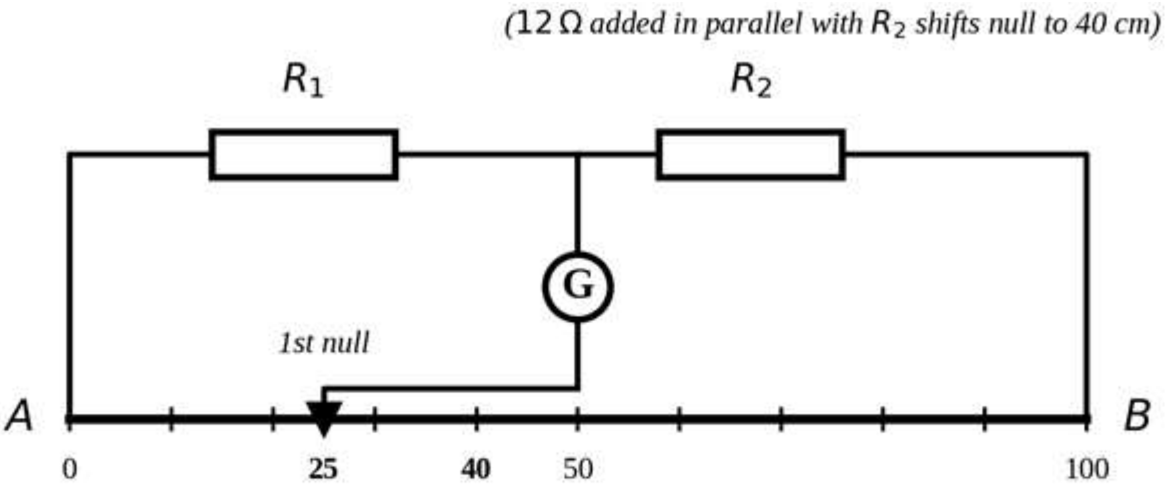


Figure: ***Example 34.*** *R_1 in the left gap and R_2 in the right gap. The first null is found at 25 cm; when a 12Ω resistor is added in parallel with R_2, the null shifts to 40 cm.*

Two balance conditions, two unknowns. Write each balance condition using the derived metre bridge formula.

First balance, null at 25 cm:

$$\frac{R_1}{R_2} = \frac{25\text{cm}}{100\text{cm} - 25\text{cm}} = \frac{25\text{cm}}{75\text{cm}} = \frac{1}{3}$$

So:

$$R_2 = 3R_1 \quad \ldots\text{(i)}$$

For the second balance, a 12Ω resistor is now in parallel with R_2. The effective resistance in the right gap is the parallel combination:

$$R_{2,\text{eff}} = \frac{12\Omega \times R_2}{12\Omega + R_2}$$

The second balance, with the null now at 40 cm, reads:

$$\frac{R_1}{R_{2,\text{eff}}} = \frac{40\text{cm}}{100\text{cm} - 40\text{cm}} = \frac{40\text{cm}}{60\text{cm}} = \frac{2}{3}$$

Rearranging:

$$R_{2,\text{eff}} = \frac{3}{2}R_1 \quad \ldots\text{(ii)}$$

Substitute the expression for $R_{2,\text{eff}}$ and use equation (i) to eliminate R_2:

$$\frac{12\Omega \times 3R_1}{12\Omega + 3R_1} = \frac{3}{2}R_1$$

Solve for R_1:

$$R_1 = 4\Omega$$

And from equation (i):

$$R_2 = 3 \times 4\Omega = 12\Omega$$

Checking both balances: First balance: $\frac{4\Omega}{12\Omega} = \frac{1}{3} = \frac{25}{75}$, so null at 25 cm. ✓

Second balance, with $R_{2,\text{eff}} = \frac{12\Omega\times12\Omega}{12\Omega+12\Omega} = 6\Omega$: $\frac{4\Omega}{6\Omega} = \frac{2}{3} = \frac{40}{60}$, so null at 40 cm. ✓

Making Sense of the Answer: *Adding a 12Ω resistor in parallel with the 12Ω right-gap resistor halves the right-gap resistance from 12Ω to 6Ω. The balance ratio $l/(100 - l)$ must double to compensate, going from*

$1/3$ *(which is* $25/75$*) to* $2/3$ *(which is* $40/60$*). The null point has shifted rightward from 25 cm to 40 cm, which is exactly what halving the right-gap resistance demands.*

Think Like a Physicist: *This problem is a two-stage application of the same balance equation. The trick is to use the simplest possible form of the balance condition at each stage and to write each as a fraction before doing any algebra. The fractions* $1/3$ *and* $2/3$ *carry the information of both null lengths in a compact form, and they combine cleanly into a quadratic in* R_1 *that has only one positive root.*

HOT Example 35

A resistor X is placed in the left gap of a metre bridge and a fixed resistor S is placed in the right gap. The resistor X is gradually heated. At 30°C the balance point is at 51.5cm from the left end; at 100°C the balance point is at 54.6cm. Find the temperature coefficient of X and the balance point if X were cooled to 0°C.

Solution

Use the metre bridge balance equation at each temperature. At temperature θ, with balance length l:

$$\frac{\mathrm{X}(\theta)}{\mathrm{S}}=\frac{\mathrm{l}}{100\mathrm{cm}-\mathrm{l}}$$

Substitute the two given balance lengths:

$$\frac{\mathrm{X}(30°\mathrm{C})}{\mathrm{S}}=\frac{51.5\mathrm{cm}}{48.5\mathrm{cm}}$$

$$\frac{\mathrm{X}(100°\mathrm{C})}{\mathrm{S}}=\frac{54.6\mathrm{cm}}{45.4\mathrm{cm}}$$

Take the ratio of the two equations; S cancels, leaving a direct ratio of resistances at the two temperatures:

$$\frac{\mathrm{X}(100°\mathrm{C})}{\mathrm{X}(30°\mathrm{C})}=\frac{54.6\mathrm{cm}\times 48.5\mathrm{cm}}{45.4\mathrm{cm}\times 51.5\mathrm{cm}}=1.1326$$

Now apply the ratio form of the temperature dependence introduced earlier in the chapter (referenced to 0°C):

$$\frac{\mathrm{X}(\theta_1)}{\mathrm{X}(\theta_2)}=\frac{1+\alpha\theta_1}{1+\alpha\theta_2}$$

With $\theta_1 = 100°\mathrm{C}$ and $\theta_2 = 30°\mathrm{C}$:

$$1.1326=\frac{1+100°\mathrm{C}\times\alpha}{1+30°\mathrm{C}\times\alpha}$$

Solve for α:

$$\alpha = 2.01\times 10^{-3}/°\mathrm{C}$$

For the balance length at 0°C, first compute $\mathrm{X}(0°\mathrm{C})/\mathrm{X}(30°\mathrm{C})$ using the ratio form:

$$\frac{\mathrm{X}(0°\mathrm{C})}{\mathrm{X}(30°\mathrm{C})}=\frac{1}{1+30°\mathrm{C}\times\alpha}=\frac{1}{1+30°\mathrm{C}\times 2.01\times 10^{-3}/°\mathrm{C}}=0.9431$$

Then the ratio $\frac{\mathrm{X}(0°\mathrm{C})}{\mathrm{S}}$ follows from $\frac{\mathrm{X}(30°\mathrm{C})}{\mathrm{S}}=\frac{51.5\mathrm{cm}}{48.5\mathrm{cm}}=1.0619$:

$$\frac{\mathrm{X}(0°\mathrm{C})}{\mathrm{S}}=\frac{\mathrm{X}(0°\mathrm{C})}{\mathrm{X}(30°\mathrm{C})}\times\frac{\mathrm{X}(30°\mathrm{C})}{\mathrm{S}}=0.9431\times 1.0619=1.0015$$

At balance, the metre bridge equation gives:

$$\frac{\mathrm{l}}{100\mathrm{cm}-\mathrm{l}}=\frac{\mathrm{X}(0°\mathrm{C})}{\mathrm{S}}=1.0015$$

Solving for l:

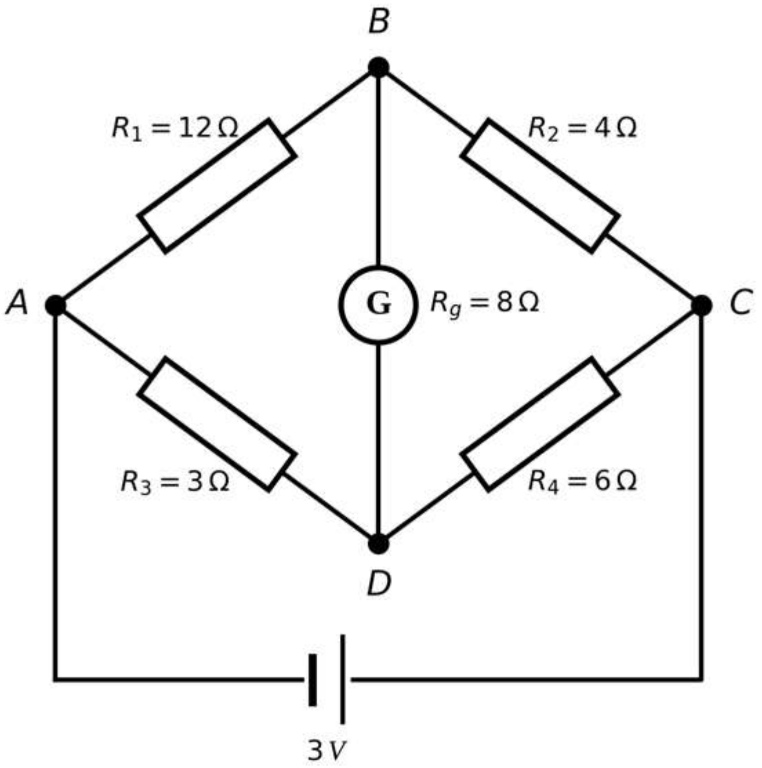
$$\mathrm{l}=50.04\mathrm{cm}$$

Making Sense of the Answer: *The temperature coefficient* $2.0 \times 10^{-3}/°C$ *is in the range expected for ordinary metals (copper is* $4.3 \times 10^{-3}/°C$*, aluminium is* $4.0 \times 10^{-3}/°C$*, nichrome is* $4 \times 10^{-4}/°C$*). The value* $2 \times 10^{-3}/°C$ *sits between these, suggesting an alloy of moderate temperature sensitivity. At* $0°C$ *the balance length is very nearly 50 cm, which makes physical sense: at this temperature* X *and* S *turn out to be almost equal in resistance, and equal-resistance gaps put the null at the midpoint of the wire.*

Think Like a Physicist: *Taking the ratio of the two balance equations was the key step. It eliminated both* S *(which is just a standard reference) and the unknown* X_0 *(the cold resistance of* X*). The remaining equation had only one unknown, α, which then dropped out cleanly. When a problem gives you two balance conditions or two measurements at different conditions, consider their ratio as the first move; most of the constants cancel.*

HOT Example 36

A Wheatstone bridge has arms $R_1 = 12\Omega$, $R_2 = 4\Omega$, $R_3 = 3\Omega$, $R_4 = 6\Omega$, supplied by a 3V battery of negligible internal resistance. A galvanometer of resistance $R_g = 8\Omega$ is connected between B and D. Find the galvanometer current and interpret what its value indicates about the balance condition.



Solution

First check the balance condition: $R_1/R_2 = 12\Omega/4\Omega = 3$, but $R_3/R_4 = 3\Omega/6\Omega = 0.5$. The two ratios differ by a factor of 6, so the bridge is very far from balance and a substantial current will flow through the galvanometer.

Use the node-voltage form of Kirchhoff's laws. Take $V_C = 0$ as the reference, so $V_A = 3V$ (the source has no internal resistance). The unknowns are V_B and V_D. Write KCL at each.

At node B, current arrives from A through R_1 and leaves through R_2 to C and through R_g to D:

$$\frac{3 - V_B}{12} = \frac{V_B}{4} + \frac{V_B - V_D}{8}$$

At node D, current arrives from A through R_3 and (possibly) from the galvanometer, and leaves to C through R_4:

$$\frac{3 - V_D}{3} + \frac{V_B - V_D}{8} = \frac{V_D}{6}$$

Multiply the node-B equation by 24:

$$2(3\mathrm{V}-\mathrm{V_B})=6\mathrm{V_B}+3(\mathrm{V_B}-\mathrm{V_D})$$

$$11\mathrm{V_B}-3\mathrm{V_D}=6 \quad \ldots(\mathrm{i})$$

Multiply the node-D equation by 24:

$$8(3\mathrm{V}-\mathrm{V_D})+3(\mathrm{V_B}-\mathrm{V_D})=4\mathrm{V_D}$$

$$\mathrm{V_B}-5\mathrm{V_D}=-8\mathrm{V} \quad \ldots(\mathrm{ii})$$

Solving with a calculator gives:

$$\mathrm{V_D}=1.808\mathrm{V}$$

$$\mathrm{V_B}=1.040\mathrm{V}$$

Since $\mathrm{V_D}>\mathrm{V_B}$, the current through the galvanometer flows from D to B (against the direction we initially assumed). Its magnitude is:

$$\mathrm{I_g}=\frac{\mathrm{V_D}-\mathrm{V_B}}{\mathrm{R_g}}=\frac{1.808\mathrm{V}-1.040\mathrm{V}}{8\Omega}=0.096\mathrm{A}=96\mathrm{mA}$$

Making Sense of the Answer: *Nearly a tenth of an ampere flows through a device meant to detect microamperes. A real laboratory galvanometer would be instantly destroyed by such a current. The problem is academic rather than experimental: it shows that when the arm ratios are highly mismatched (*3 *versus* 0.5*, a factor of* 6*), the galvanometer current is large. In real Wheatstone experiments, the bridge is brought near balance first (often by protective resistors in series with the galvanometer that are shorted out as the balance is approached), and the final fine adjustment is made when I_g is already small.*

Think Like a Physicist: *The size of the galvanometer current is a quantitative measure of how far the bridge is from balance. A balanced bridge gives $I_g=0$; a slightly unbalanced bridge gives a small I_g; a wildly unbalanced bridge gives a huge I_g. The direction of I_g tells you which way to adjust: if it flows from D to B, you need to lower R_1/R_2 or raise R_3/R_4. The bridge gives both magnitude and direction information at the same time, which is what makes it so useful as a precision comparator.*

Three examples illustrate the bridge in three modes. In every case the precision of the result rests on a length measured along a wire or a ratio of known resistances, never on the EMF of the driving cell or the sensitivity of the galvanometer. That is the strength of a null method.

The next section extends the null-method principle from resistance measurement to EMF measurement. The instrument is called a potentiometer, and it answers a question no voltmeter can answer exactly: *what is the true EMF of a cell, separately from the voltage drop across its internal resistance?*

THE POTENTIOMETER

The previous section gave us a null method for resistance. This section gives us a null method for EMF. The instrument is called the potentiometer, and it does for cells what the Wheatstone bridge does for resistors: it measures the unknown quantity without drawing any current from the source being measured. The consequence is that a potentiometer reads the true EMF of a cell, not the terminal voltage. This distinction matters whenever a cell has appreciable internal resistance, which is to say, almost always.

Recall the terminal-voltage relation derived earlier in the chapter: when a cell of EMF E and internal resistance r drives a current I through a load, the voltage across its terminals is $\mathrm{V}=\mathrm{E}-\mathrm{Ir}$. A voltmeter, however good, always draws some current; it always reads V, never E. The error is small when I is small or r is small, but it is never zero. The potentiometer escapes this trap by making I exactly zero through the cell at the moment of measurement.

Why the potential gradient along the wire is uniform

A potentiometer is built around a long uniform resistance wire, stretched along a metre scale. A separate driving cell pushes a steady current along the whole length of the wire. Because the wire is uniform in cross-section and uniform in material, the current density is the same everywhere along it, and Ohm's law applied locally gives a uniform potential gradient: equal lengths of wire carry equal voltage drops.

Quantitatively, let the wire have length L, cross-section A, and resistivity ρ. Its resistance per unit length is:

$$r_L = \frac{\rho}{A}, \text{a constant } \left(\text{From } R = \frac{\rho l}{A};\ r_L = \frac{R}{l} = \frac{\rho}{A}\right)$$

With a steady driving current I_d flowing, the potential drop per unit length is:

$$\mathbf{k} = I_d \times r_L = \frac{\mathbf{I_d \rho}}{\mathbf{A}}$$

This quantity **k** is the **potential gradient** along the wire, in volts per metre or volts per centimetre depending on the unit of length. Because all three factors on the right are constants, k is constant along the wire.

Equivalently, the total potential drop across the wire is k × L, which (if no other resistances are in the driving loop) equals the **EMF of the driving cell $\mathbf{E_d}$**. So:

$$E_d = k \times L \text{ or } k = \frac{E_d}{L}$$

Once the potential gradient is known, the potential at any point along the wire (measured from end A) is k times the distance. Touching a jockey at distance l from A picks up a potential of k × l. This is the working principle of every potentiometer measurement: balance an unknown voltage against k × l for some length l, and the unknown voltage is determined by the balance length.

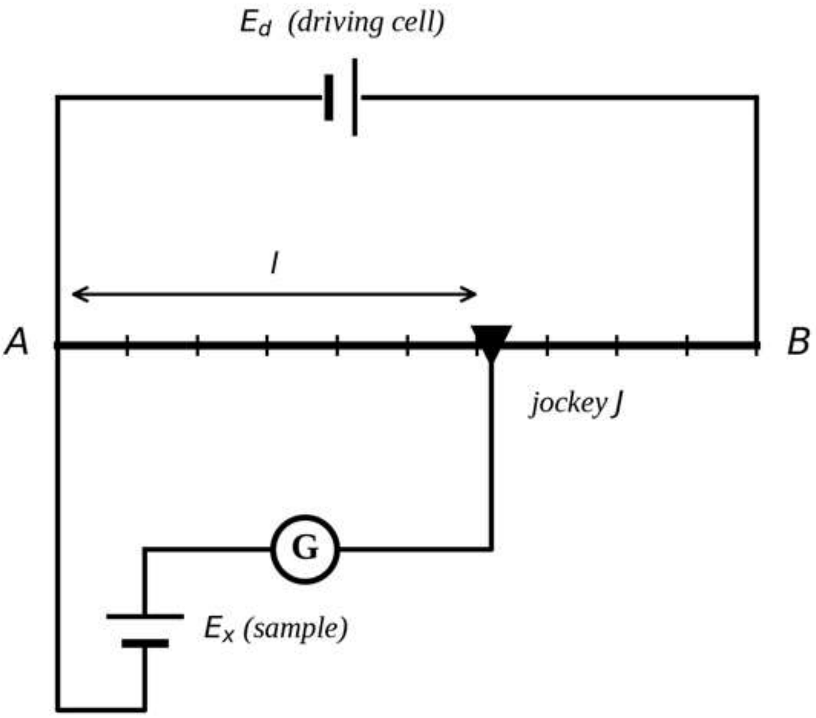


Figure: ***The basic potentiometer.*** *A driving cell E_d pushes a steady current along the uniform wire AB. The jockey J taps the wire at distance l from A. The galvanometer G connects the jockey to one terminal of the sample cell E_x; the other terminal of E_x returns to A. When the jockey is at the balance position, no current flows through G and the sample cell's EMF equals k×l.*

At the balance point, the potential of the jockey (measured relative to A, along the wire) exactly equals the EMF of the sample cell. The galvanometer therefore has zero voltage across it, so by Ohm's law no current flows through it. Because no current flows through the galvanometer arm, no current flows through the sample cell either; the sample cell is instantaneously on open circuit, and its terminal voltage equals its EMF. The balance reading is therefore the true EMF, not the terminal voltage under load.

Three measurements follow from this principle: comparison of EMFs, measurement of internal resistance, and calibration of the instrument. Each is treated separately below.

Comparison of EMFs

Suppose we have a standard cell of accurately known EMF Es, and an unknown cell of EMF Ex. We can connect them in turn to the potentiometer (using a two-position switch) and find the balance length for each in turn. Let Ls be the balance length for the standard cell, and Lx be the balance length for the unknown cell.

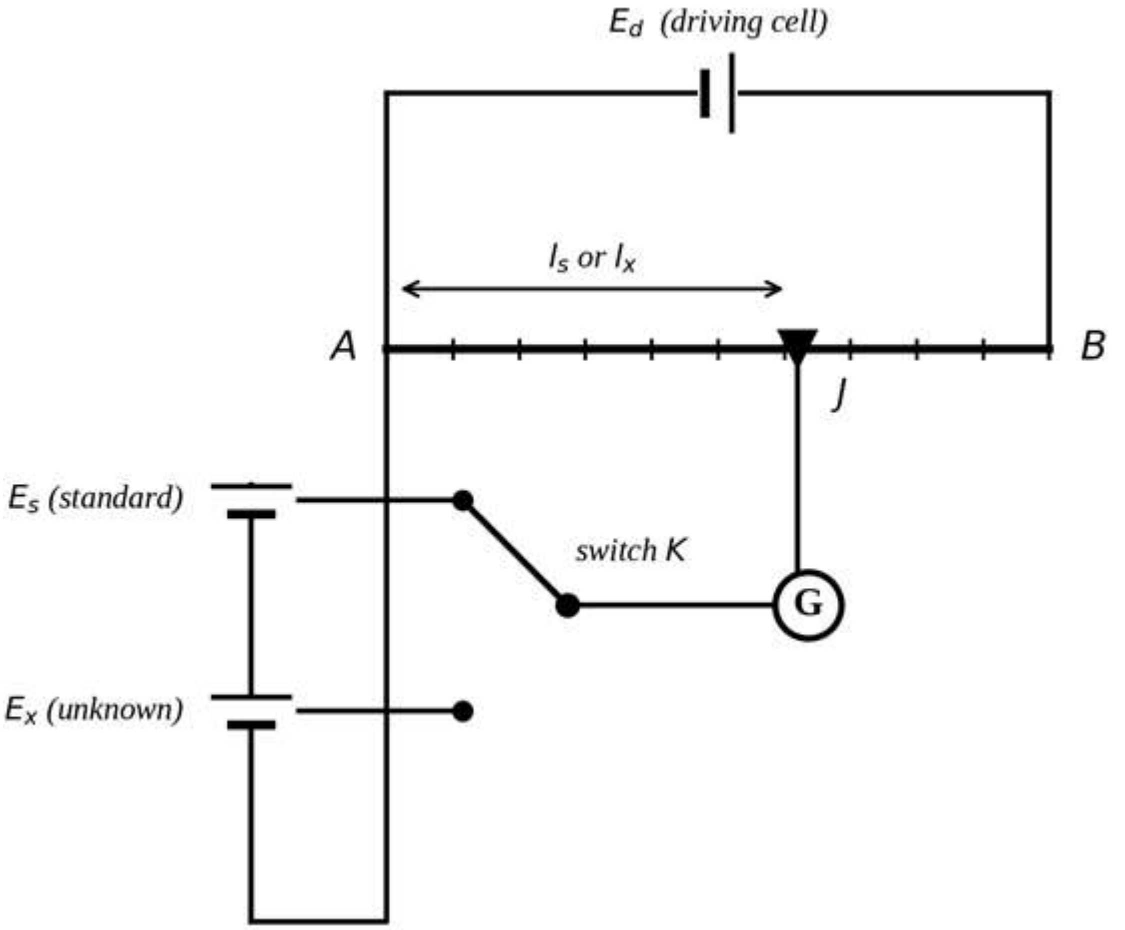


Figure: ***Potentiometer arrangement for comparing two EMFs.*** *A two-position switch K selects either the standard cell E_s or the unknown cell E_x; the chosen cell is balanced on the wire. The balance length for each is measured.*

At balance:

$$E_s = k \times L_s \text{ and } E_x = k \times L_x.$$

Dividing the second by the first cancels k:

$$\frac{E_x}{E_s} = \frac{L_x}{L_s}$$

Rearranging for the unknown EMF:

$$\mathbf{E_x = \frac{L_x}{L_s} \times E_s}$$

The unknown EMF is determined entirely by the two balance lengths and the standard EMF. The driving current I_d, the resistance per unit length, and any other property of the wire do not enter; they all cancelled with k. The same is true of the galvanometer resistance and the switch resistance, neither of which carries any current at balance.

Measurement of internal resistance

The internal resistance of a cell cannot be measured directly with an ohmmeter (the cell's own EMF would interfere with the test current). The potentiometer offers an indirect route. Take two balance readings of the same cell: one on open circuit (no external load), and one with a known external resistor R connected across the cell.

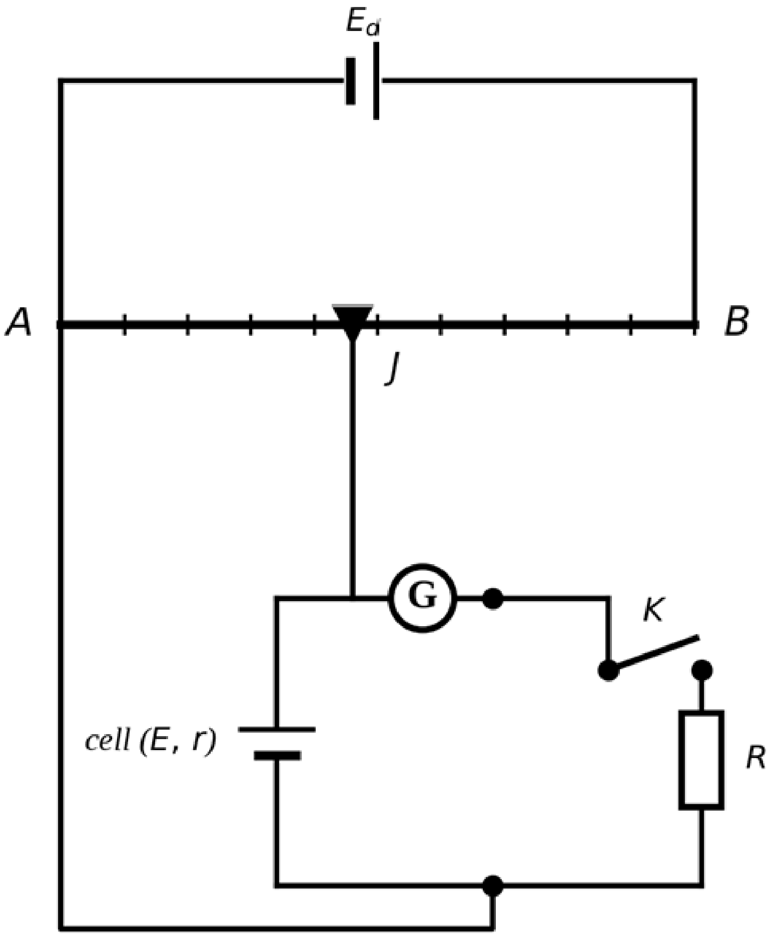


Figure: ***Potentiometer arrangement for measuring the internal resistance of a cell.*** *With the switch K open, the cell is on open circuit and its EMF is balanced at length L. With K closed, the resistor R draws current from the cell, the terminal voltage drops below the EMF, and the new balance length L1 is smaller. The internal resistance follows from the two lengths and R.*

Open-circuit reading first. With the switch K open, the cell delivers no current (the galvanometer arm carries none at balance, and the load R is disconnected). The terminal voltage therefore equals the EMF E. Let L be the balance length under this condition. Then:

$$\mathrm{E = k \times L}$$

Now close the switch K. Current flows from the cell through R (this is the cell's load current). The cell's terminal voltage drops below the EMF by the internal-drop Ir:

$$\mathrm{V = E - Ir}$$

The current flowing through R is $\mathrm{I = \frac{V}{R}}$, so:

$$\mathrm{V = E - \frac{V}{R} \times r}$$

Multiply through by R and rearrange:

$$\mathrm{VR = ER - Vr}$$

$$\mathrm{V(R + r) = ER}$$

$$\mathrm{\frac{E}{V} = \frac{R + r}{R}}$$

Re-balance the potentiometer with the load now connected. The new balance length L_1 balances the terminal voltage V (the galvanometer arm carries no current at balance, so the voltage across the cell's terminals at balance equals the potential of the jockey, which is $k \times L_1$). So:

$$V = k \times L_1$$

Taking the ratio E/V using $E = k \times L$ and $V = k \times L_1$:

$$\frac{E}{V} = \frac{L}{L_1} = \frac{R + r}{R}$$

Solve for r:

$$r = \frac{RL}{L_1} - R = R\left(\frac{L}{L_1} - 1\right)$$

$$\mathbf{r = R\left(\frac{L - L_1}{L_1}\right)}$$

The internal resistance is determined by the two balance lengths and the known load R. The driving current and the potential gradient k cancel out of the answer, just as in the EMF comparison.

Calibration of the potentiometer

Before the potentiometer can be trusted to measure an unknown EMF directly (rather than by comparison), the potential gradient k must be known. Calibration consists of balancing a standard cell of known EMF E_s at length L_s and computing:

$$k = \frac{E_s}{L_s}$$

Once k is known, any unknown EMF balanced at length L_x follows from:

$$E_x = k \times L_x$$

Verification: A second standard cell of different EMF balanced at a different length should give the same k. If it does not, the driving cell has drifted, the wire is not uniform, or there is a loose contact in the driving loop. A potentiometer that fails this check should not be trusted until it passes again.

Three examples now follow.

BINDER Example 37

A potentiometer is first balanced with a standard cell of EMF $E_s = 1.50V$ at length $L_s = 40cm$. The same potentiometer then balances a second cell A at length $L_A = 60cm$. Find the EMF of cell A, and explain briefly why the potentiometer ensures an accurate comparison.

Solution

Apply the EMF-comparison formula derived in this section:

$$\frac{E_A}{E_s} = \frac{L_A}{L_s}$$

Substitute the given numerical values:

$$E_A = \frac{L_A}{L_s} \times E_s = \frac{60cm}{40cm} \times 1.50V = 2.25V$$

Explanation: The potentiometer ensures an accurate comparison because at balance no current flows from cell A through the galvanometer arm, so cell A is momentarily on open circuit and the potential picked off the wire equals the true EMF of cell A, not the terminal voltage it would show while delivering current to a load.

Making Sense of the Answer: *Cell A has a larger EMF than the standard cell, so it requires a longer length of wire to balance. The ratio of balance lengths,* $\frac{60cm}{40cm} = 1.5$, *is exactly the ratio of the two EMFs, which is* *why $E_A = 1.5 \times 1.50V = 2.25V$. The number is reasonable: 2.25V is a little more than a single dry cell's nominal 1.5V, the sort of value a fresh cell or a small two-cell source might give.*

Think Like a Physicist: *The whole comparison rests on a single fact: the potential along a uniform wire grows in direct proportion to the distance from the start. So a ratio of two EMFs becomes a ratio of two lengths, and the potential gradient* k *cancels without ever needing to be measured. A voltmeter could give the same answer in principle, but only by drawing a small current and so reading each cell's terminal voltage rather than its EMF. Whenever you need a true EMF and not a loaded voltage, reach for a balance method, not a meter.*

HOT Example 38

In measuring the internal resistance of a cell with a potentiometer, the EMF is first balanced by a length of 90cm on the potentiometer wire (with the cell on open circuit). When a 5Ω resistor is connected across the cell's terminals, the terminal voltage now balances at 45cm. (a) Explain why the balance length is smaller when the resistor is connected. (b) Calculate the internal resistance of the cell.

Solution

(a) When the resistor R is connected, the cell drives a current I through R, and the terminal voltage of the cell drops below its EMF by the internal-drop Ir. A smaller voltage requires a shorter length of wire to balance.

(b) Apply the internal-resistance formula:

$$\mathrm{r} = \mathrm{R} \times \frac{\mathrm{L} - \mathrm{L_1}}{\mathrm{L_1}},$$

with $\mathrm{L} = 90\text{cm}$, $\mathrm{L_1} = 45\text{cm}$, and $\mathrm{R} = 5\Omega$:

$$\mathrm{r} = 5\Omega \times \frac{90\text{cm} - 45\text{cm}}{45\text{cm}} = 5\Omega \times \frac{45\text{cm}}{45\text{cm}} = 5\Omega$$

Making Sense of the Answer: *The balance length halved when the load was connected, dropping from* $90cm$ *to* $45cm$. *This means the terminal voltage under load is exactly half the EMF:* $V = E/2$. *From* $V = E - Ir$ *with* $V = E/2$, *we get* $E - Ir = E/2$, *so* $Ir = E/2 = V$. *The internal drop equals the load drop, which means the internal resistance equals the external load:* $r = R = 5\Omega$. *This is a special case worth recognising:* ***whenever the balance length halves on adding a load, the internal resistance equals the load resistance, regardless of the actual EMF value.***

Think Like a Physicist: *Notice how everything in the answer cancelled except R and the ratio of lengths. The EMF E never had to be known, the potential gradient* k *never had to be known, and the driving current* I_d *never had to be known. This is what* ***null methods*** *give you: answers that depend only on the known quantities the experimenter actually controls, with all the inconvenient ones cancelling out.*

HOT Example 39

A potentiometer is calibrated using a standard cell of EMF 1.0V; the balance length is 40cm. An unknown dry cell balances at 30cm under open-circuit conditions. The dry cell is then connected to an external resistor $\mathrm{R} = 4\Omega$; the new balance length under load is $\mathrm{L_t} = 20\text{cm}$. Find (a) the potential gradient k in V/cm, (b) the EMF of the dry cell, (c) the terminal voltage of the dry cell under load, (d) the current through R, and (e) the internal resistance of the dry cell.

Solution

(a) The potential gradient follows from the calibration. The standard cell's EMF balances at 40cm, so the potential drop per centimetre of wire is:

$$\mathrm{k} = \frac{\mathrm{E_s}}{\mathrm{L_s}} = \frac{1.0\mathrm{V}}{40\text{cm}} = 0.025\mathrm{V/cm}$$

(b) The EMF of the dry cell is the open-circuit balance reading. At $\mathrm{L_x} = 30\text{cm}$:

$$\mathrm{E_x} = \mathrm{k} \times \mathrm{L_x} = 0.025\mathrm{V/cm} \times 30\text{cm} = 0.75\mathrm{V}$$

(c) The terminal voltage under load balances at $\mathrm{L_t} = 20\text{cm}$:

$$V = k \times L_t = 0.025\text{V/cm} \times 20\text{cm} = 0.50\text{V}$$

(d) The current through R follows from Ohm's law applied to the load:

$$I = \frac{V}{R} = \frac{0.50\text{V}}{4\Omega} = 0.125\text{A}$$

(e) The internal resistance follows from the terminal-voltage relation $V = E - Ir$, which rearranges to:

$$r = \frac{E_x - V}{I} = \frac{0.75\text{V} - 0.50\text{V}}{0.125\text{A}} = 2.0\Omega$$

Making Sense of the Answer: *The five quantities form a tightly linked story. The driving wire carries a potential gradient of* 0.025V/cm*, which converts every length into a voltage. The unknown cell has EMF* 0.75V *(open circuit,* 30cm*) and terminal voltage* 0.50V *(under* 4Ω *load,* 20cm*). The drop from* 0.75V *to* 0.50V *is the internal-drop* $Ir = 0.25\text{V}$*, and with the current* $I = 0.125\text{A}$ *we recover* $r = 2.0\Omega$*. The internal resistance is half the external load, which means two-thirds of the cell's EMF appears across the load and one-third is lost as the internal drop. A fresh dry cell would have a smaller internal resistance;* 2Ω *suggests this one is approaching the end of its useful life.*

Think Like a Physicist: *This problem strings together five separate physical quantities, but every one is derived from a length measurement on the potentiometer wire and an Ohm's-law step. The potentiometer has done two distinct things in one experiment: measured the EMF (by open-circuit balance) and measured the terminal voltage under load (by closed-circuit balance). Their difference, divided by the load current, gives the internal resistance. The whole experiment uses only lengths, a known standard EMF, and one known load resistor; no voltmeter or ammeter is needed anywhere.*

The potentiometer turns three separate measurements (EMF, terminal voltage, internal resistance) into a single experiment based on balance lengths along a wire. The accuracy of every answer rests on the same two pillars: a uniform wire (giving a uniform potential gradient) and a null detection (giving zero current from the cell being measured at the moment of measurement). When both conditions hold, the answers are independent of the driving cell, the galvanometer, and every property of the apparatus except the lengths and the standard reference.

The next section turns from measurement to the energy budget of an electrical circuit: where the power comes from, where it goes, and how much of it can be made to do useful work.

ENERGY, POWER, MAXIMUM POWER TRANSFER, AND JOULE HEATING

Every section so far has treated the circuit as a place where charge moves and voltage falls. This section treats it as a place where energy is spent, which, as it happens, is the only thing your electricity bill has ever cared about. A source pours energy into the circuit at some rate; the circuit hands that energy out to its components at the same rate; and in a resistor the energy handed out does not pile up or get stored, it turns at once into heat. The questions that matter from here on are accounting questions: *how fast the source supplies energy, how fast each component spends it, what fraction of the supplied energy reaches the load the circuit was built to serve, and what is lost on the way.*

Those four questions decide why a kettle element glows, why a fuse melts at exactly the right moment, why an amplifier runs hot enough to warm your hands, and why TANESCO sends power across the country at 132kV instead of the 240V that reach a house. All of it follows from one idea: a resistor is a device that turns electrical energy into heat, and it does so at a rate we can write down exactly.

Where the heat comes from

Imagine a crowd of people shoving their way down a long corridor crammed with furniture. They do not glide through; they bump, they jostle, they lose a little energy at every collision, and the furniture and the air grow warmer for it. That corridor is a resistor, the crowd is the conduction electrons, and the warmth left behind is the entire subject of this section. Nothing in the picture changes when we make it quantitative. We only count the energy that the crowd leaves behind.

So count it. A resistor R carries a steady current I, with a voltage V across its ends. In a time t, a quantity of charge $Q = It$ passes through it. Each coulomb of that charge falls through the potential difference V as it crosses the resistor, and falling through V joules per coulomb means each coulomb gives up V joules. The total energy given up is the charge multiplied by the voltage drop:

$$\mathbf{U = QV}$$

Now substitute the two facts we already know about a resistor: the charge that flowed is $Q = It$, and the voltage across it obeys Ohm's law, $V = IR$. Putting both into $U = QV$:

$$\mathbf{U} = QV = (It)(IR) = \mathbf{I^2Rt}$$

This is the energy delivered to the resistor in time t, and because a resistor stores nothing, it is also the heat produced in the resistor in that time. *The conversion of electrical energy into heat inside a current-carrying resistor* is called **Joule heating**, and the quantity I^2Rt is the **Joule heat**.

The earlier warm lamp wire was Joule heating seen from the outside; this is the same effect written as a formula. Every drifting electron that collides with the lattice surrenders a little of the energy the field gave it, the lattice takes that energy up as vibration, and the wire grows warm. The formula I^2Rt simply counts the total of all those tiny surrenders over the whole time t.

The three forms of electrical power

Energy is the whole story, but rate is what we usually feel. A heater that delivers a megajoule over a week keeps no one warm; the same megajoule in ten minutes is a serious fire. So before any formula, fix the idea: ***power is the rate of spending energy****, and a resistor spends it as heat.*

Dividing the Joule heat by the time gives that rate, the power dissipated in the resistor:

$$P = \frac{U}{t} = I^2R$$

This is the first of three forms. Because the resistor also obeys $V = IR$, the same power can be written two other ways, depending on which pair of quantities you happen to know. Replacing one factor of I by $\frac{V}{R}$, or replacing IR by V, gives the full set:

$$P = I^2R = VI = \frac{V^2}{R}$$

These are not three different powers. They are one power wearing three outfits, and you pick the outfit that matches the data. If a problem hands you current and resistance, use I^2R. If it hands you voltage and current, use VI. If it hands you voltage and resistance, use $\frac{V^2}{R}$. Choosing the form that fits the data saves a step of algebra and avoids inventing a quantity the problem never gave you. The unit of power is the watt, one joule per second, and from $P = VI$ one watt is also one volt-ampere.

Two of the forms carry a warning worth stating once. In $P = I^2R$ the power rises with the **square** of the current, so doubling the current through a fixed resistor quadruples the heat; a small rise in current can cook a wire that felt cool a moment earlier. In $P = \frac{V^2}{R}$ the power rises with the square of the voltage but **falls** as the resistance rises, so for a fixed supply voltage a smaller resistance draws more power, not less. The kettle element with the lower resistance is the more powerful one, a fact that feels backwards right up until you decide to trust the formula over your gut.

The energy budget of a source

A resistor only consumes. A source supplies, and it is not quite as generous as it looks: it skims a little off the top to warm itself. To see the whole budget, return to the cell of EMF E and internal resistance r driving a current I through an external load R, the circuit met earlier in the chapter. The EMF is the work the source does per coulomb, so the rate at which the source delivers energy to the entire circuit, the total power generated, is:

$$P_{\text{source}} = EI$$

That total splits in two, because the current must fight its way through the internal resistance before it ever reaches the load. The power wasted inside the source, heating the cell itself, is the Joule heating in r:

$$P_{\text{internal}} = I^2r$$

What is left over reaches the external load. Subtracting, and using $E - Ir = V$, the terminal voltage:

$$\mathbf{P_{external}} = P_{source} - P_{internal} = EI - I^2r = I(E - Ir) = \mathbf{VI}$$

So the power reaching the load is the terminal voltage times the current, exactly the VI you would compute at the load's own terminals, with no reference to what goes on inside the source. The budget closes: every joule the EMF supplies either heats the internal resistance or reaches the load, and nothing goes missing.

The fraction that reaches the load is the **efficiency** of the transfer, *the external power divided by the total power generated:*

$$\eta = \frac{P_{external}}{P_{source}} = \frac{VI}{EI} = \frac{V}{E} = \frac{R}{R + r}$$

The last step uses $V = IR$ and $E = I(R + r)$, which lets the current cancel and leaves efficiency as a pure ratio of resistances. Efficiency is high when R is large compared with r, because then most of the EMF lands on the load and little is lost inside. It sinks toward one half, and below, as R shrinks toward r. This one ratio, $\frac{R}{R+r}$, is the hinge on which the rest of the section turns.

When is the most power delivered to the load?

Efficiency is one question; raw power is another, and they do not share an answer. Imagine the source is fixed, its EMF E and internal resistance r beyond your control, and the only knob you may turn is the load R. You want the load to receive as much power as it possibly can. Turn R too low and the load is nearly a short circuit with almost no voltage across it; turn R too high and you choke the current to a trickle. Somewhere between greed and restraint there is a best setting, and we can find it exactly.

Write the load power in terms of R alone. The current is $I = \frac{E}{R+r}$, and the power in the load is I^2R:

$$P = I^2R = \frac{E^2R}{(R + r)^2}$$

As R climbs from zero, this power starts at zero, rises to a peak, and falls back toward zero. To pin the peak, differentiate P with respect to R and set the derivative to zero. Using the quotient rule on $\frac{E^2R}{(R+r)^2}$:

$$\frac{dP}{dR} = \frac{E^2[(R + r)^2 - R \cdot 2(R + r)]}{(R + r)^4}$$

The numerator simplifies. Cancel one factor of $(R + r)$ from top and bottom, and collect the bracket $(R + r) - 2R = r - R$:

$$\frac{dP}{dR} = \frac{E^2(r - R)}{(R + r)^3}$$

The denominator is always positive, so the derivative vanishes only when its numerator does, that is when $r - R = 0$:

$$R = r$$

The load receives the most power when its resistance equals the internal resistance of the source. This is the **maximum power transfer theorem**. The derivative is positive for $R < r$ (power still climbing) and negative for $R > r$ (power now falling), which confirms the turning point is a maximum, not a minimum. Substituting $R = r$ back into the power gives the height of the peak:

$$\mathbf{P_{max}} = \frac{E^2r}{(r + r)^2} = \frac{E^2r}{4r^2} = \mathbf{\frac{E^2}{4r}}$$

And the efficiency at the matched point follows from the ratio above, with $R = r$:

$$\eta = \frac{R}{R + r} = \frac{r}{2r} = \frac{1}{2} = 50\%$$

Here is the sting in the tail. At the very setting that pours the most power into the load, exactly half of the source's total power is being burned inside the source itself. Maximum power and high efficiency are not the same goal; they sit at opposite ends of one curve. The matched load grabs the largest possible slice, but it pays by forcing the source to waste an equal slice internally. To run efficiently instead, you must make the load far larger than r, accepting less power in exchange for wasting less.

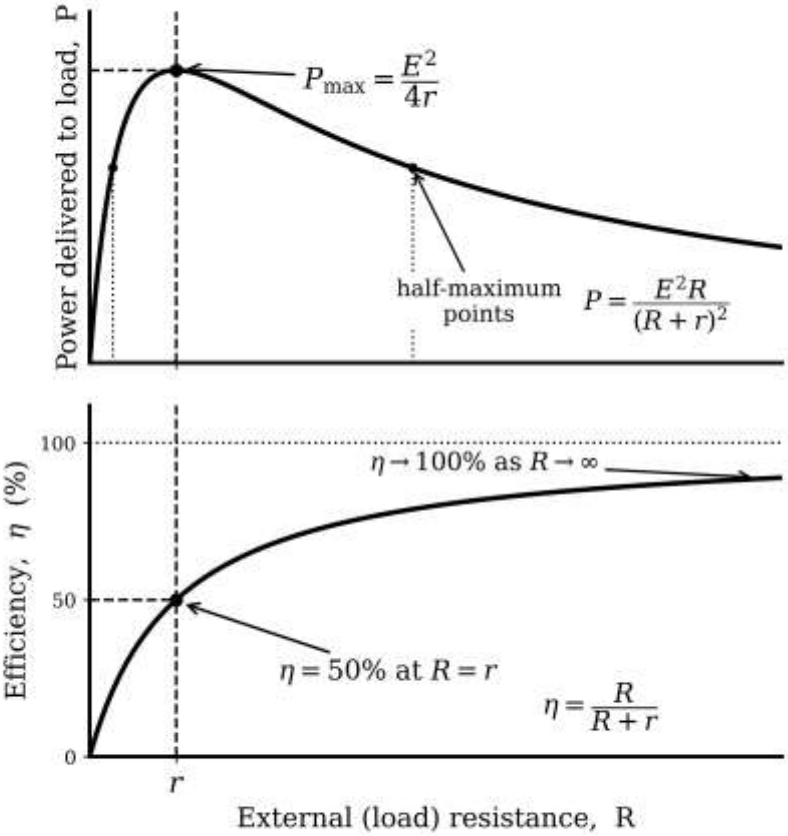


Figure: ***The power delivered to the load and the efficiency of the transfer, both plotted against the load resistance R, with the internal resistance r marked on the axis.*** *The power (upper curve) climbs from zero to a single peak directly above $R = r$ at the height $\frac{E^2}{4r}$, then falls away slowly; its half-maximum points are marked for context. The efficiency (lower curve) rises steadily, crosses exactly 50% at $R = r$, and climbs toward 100% as R grows without bound. At the one value $R = r$ the power is greatest while the efficiency is only one half.*

Kipute saw the contradiction at once.

Kipute: *Sir, if the maximum-power condition wastes half the source's energy as heat inside the source, why is anyone ever interested in maximum power transfer? Fifty per cent efficiency sounds like a failure, not a target.*

Mr. Akilikubwa: *It is a failure in a power station, and a triumph in a loudspeaker. The trick is to ask what is scarce. In an audio amplifier the source's power is cheap and plentiful; what is precious is the power pushed into the speaker, because that is what becomes sound. So the amplifier's output is deliberately matched to the speaker's four or eight ohms, accepting fifty per cent efficiency to win the loudest possible note. The wasted half is exactly why the back of an amplifier is warm.*

Kipute: *And TANESCO does the opposite.*

Mr. Akilikubwa: *It must. The source there is the whole Mtera plant, and wasting half of it would be a national disaster, so the grid runs at well over ninety-five per cent efficiency, with the load resistance enormously larger than the line's. Same theorem, opposite design. The maximum power theorem tells you the trade-off; it never tells you which side of it to stand on. That choice is engineering, not physics.*

Designing things that get hot on purpose

Joule heating is a nuisance in a transmission line and the entire point of a heater. A heating appliance is nothing more than a resistor chosen so that I^2R, or equivalently $\frac{V^2}{R}$, comes out to the wattage the appliance is meant to deliver at the voltage it is meant to run on. Designing one is choosing a resistance, then choosing a wire that has that resistance, runs hot without melting or burning away, and barely changes its resistance as it heats.

That last requirement quietly rules out copper, the very metal used everywhere else in the circuit. Copper's resistivity is low, so a copper element would need an absurd length of wire to reach a useful resistance, and copper's resistance climbs steeply with temperature, so it would draw one power cold and a different one hot, and oxidise away at red heat.

The standard answer is **nichrome**, an alloy of **ni**ckel and **chrom**ium whose resistivity is roughly sixty times that of copper, so a short coil suffices, and whose temperature coefficient of resistance is small, so its power barely drifts from cold to glowing. Nichrome also grows a protective oxide skin that lets it sit at bright red heat in open air for years. Every exposed heating coil you have ever seen, in a room heater, a hair dryer, a toaster, is nichrome for these three reasons at once.

Around that core idea sit the familiar appliances.

- **An electric kettle** hides a sealed element in the water and is rated by the power it must deliver, from which its resistance follows by $R = \frac{V^2}{P}$; a 2.5kW kettle on 240V mains is a resistance of about 23Ω.
- **An electric iron** uses the same kind of element but adds a thermostat, a bimetallic strip that bends as it warms and breaks the circuit at a set temperature, then remakes it as it cools, holding the soleplate at the chosen heat by switching the Joule heating on and off.
- **A fuse** is the same physics turned to protection: a short length of low-melting-point wire, often a lead-tin alloy, sized so that at the rated current its I^2R heating keeps it just below melting, but at any larger current the extra heat melts it and breaks the circuit before the fault can set the rest of the wiring alight.
- **An immersion heater** is a long nichrome wire wound on a ceramic former and sheathed in a metal tube, sized by exactly the same $\frac{V^2}{R}$ sum that sizes the kettle.

Why power is sent across the country at high voltage

The last application is the largest of all. A power station sits far from the towns it feeds, and the line that carries its power has resistance, perhaps a hundred ohms over a hundred kilometres of conductor. That line resistance sits in series with everything, so the current that feeds the load must also pass through the line, heating it by I^2R. This heat is pure waste: it warms the open countryside and serves no one. The whole art of transmission is to make that waste small.

Here is the key, and it is worth slowing down for. The same power can be carried either as a small current at a high voltage or a large current at a low voltage, because power is the product VI. But the line loss depends only on the current, not on the voltage. Write the loss as I^2R_{line} and replace the current by $I = \frac{P_{\text{load}}}{V_{\text{load}}}$:

$$\mathbf{P_{line}} = I^2R_{\text{line}} = \left(\frac{\mathbf{P_{load}}}{\mathbf{V_{load}}}\right)^{\mathbf{2}} \times \mathbf{R_{line}}$$

Read this slowly, because it contains the entire reason for the national grid. *For a fixed power delivered and a fixed line, the loss falls as the* ***square*** *of the transmission voltage.* Double the voltage and the loss drops to a quarter; raise it tenfold and the loss drops to a hundredth.

To carry a fixed power, a higher voltage means a proportionally smaller current, and since the loss goes as the current squared, a high voltage with a small current is dramatically cheaper to transmit than a low voltage with a large current. This is why the grid steps the voltage up to hundreds of kilovolts for the long haul, then steps it back down with transformers to the safe 240V at the point of use.

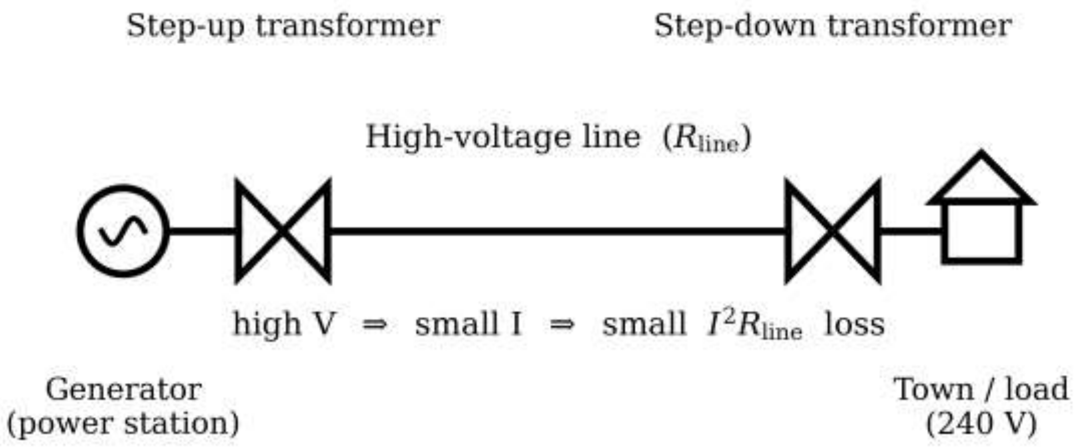


Figure: ***A high-voltage transmission line drawn from left to right.*** *A generator at the power station feeds a step-up transformer that raises the voltage to a high transmission value. A long line of resistance* R_{line} *carries the power across country at high voltage and therefore small current, keeping the* I^2R_{line} *loss small. A step-down transformer at the far end lowers the voltage again, delivering it to the town at* $240V$.

The theory is now complete, which means it is time to discover whether it survives contact with actual numbers. Six worked examples follow. Work each one before you read the solution; physics, like swimming, is not learned by watching other people do it. The solutions will wait for you. They are extremely patient, being made entirely of ink.

BINDER Example 40

An electric lamp is marked 100W, 240V. It is switched on for one hour. Find (a) the current it draws, (b) the resistance of its filament at working temperature, and (c) the total electrical energy it converts to heat and light in the hour, in joules and in kilowatt-hours.

Solution

(a) The marking tells us the lamp dissipates 100W on 240V. Power, voltage and current are tied by $\mathrm{P = VI}$, so:

$$\mathrm{I} = \frac{\mathrm{P}}{\mathrm{V}} = \frac{100\mathrm{W}}{240\mathrm{V}} = 0.417\mathrm{A}$$

(b) The resistance follows from the form of the power equation that uses the two quantities we were handed, voltage and power:

$$\mathrm{R} = \frac{\mathrm{V}^2}{\mathrm{P}} = \frac{(240\mathrm{V})^2}{100\mathrm{W}} = 576\Omega$$

(c) Energy is power times time. One hour is 3600s:

$$\mathrm{U = Pt = 100W \times 3600s = 360000J = 360kJ}$$

In the unit your bill actually uses, one kilowatt-hour is one kilowatt sustained for one hour, so a 100W lamp burning for one hour uses 0.1kWh.

Making Sense of the Answer: *The three parts are three readings of one lamp. It draws less than half an ampere, which is why the thin flex to a reading lamp never feels more than warm. Its working resistance of* 576Ω *is far above the cold resistance a meter would show, because tungsten resistance climbs steeply with temperature and the filament runs well past two thousand degrees. And the* $360kJ$ *it spends in an hour is mostly heat, not light: an incandescent lamp is a heater that happens to glow, which is the whole reason it has been retired in favour of cooler technologies.*

Think Like a Physicist: *Notice how the choice of formula followed the data. Part (a) had power and voltage and wanted current, so* $P = VI$ *was the natural form. Part (b) had voltage and power and wanted resistance, so* $\frac{V^2}{R}$ *was natural. At no point was it necessary to find the current first and then the resistance from it, though that route would also work. Reading the data before reaching for a formula is half the skill; the three forms of the power equation exist precisely so that whatever pair you are handed, one of them fits without a detour.*

HOT Example 41

A battery of EMF 12V has an internal resistance of 0.5Ω. (a) Find the power delivered to an external resistor of 4Ω, and the efficiency of the transfer. (b) Repeat for an external resistor of 0.5Ω. (c) Show that 0.5Ω is the load that draws the maximum power, and state that maximum. (d) Comment on the efficiency at the matched load compared with part (a).

Solution

(a) With $\mathrm{R} = 4\Omega$ and $\mathrm{r} = 0.5\Omega$, the current is:

$$\mathrm{I} = \frac{\mathrm{E}}{\mathrm{R} + \mathrm{r}} = \frac{12\mathrm{V}}{4\Omega + 0.5\Omega} = 2.67\mathrm{A}$$

The power in the load is $\mathrm{I}^2\mathrm{R}$, and the efficiency is $\frac{\mathrm{R}}{\mathrm{R+r}}$:

$$\mathrm{P} = \mathrm{I}^2\mathrm{R} = (2.67\mathrm{A})^2 \times 4\Omega = 28.4\mathrm{W}$$

$$\eta = \frac{\mathrm{R}}{\mathrm{R} + \mathrm{r}} = \frac{4\Omega}{4\Omega + 5\Omega} = 0.889 = 88.9\%$$

(b) With $\mathrm{R} = 0.5\Omega$ the current is larger, because the total resistance is smaller:

$$\mathrm{I} = \frac{12\mathrm{V}}{0.5\Omega + 0.5\Omega} = 12\mathrm{A}$$

$$\mathrm{P} = \mathrm{I}^2\mathrm{R} = (12\mathrm{A})^2 \times 0.5\Omega = 72\mathrm{W} \qquad \eta = \frac{0.5\Omega}{0.5\Omega + 0.5\Omega} = 0.50 = 50\%$$

(c) The maximum power transfer theorem says the load draws the most power when $\mathrm{R} = \mathrm{r}$. Here $\mathrm{r} = 0.5\Omega$, so the matched load is precisely the 0.5Ω resistor of part (b). The maximum power is:

$$\mathrm{P}_{\mathrm{max}} = \frac{\mathrm{E}^2}{4\mathrm{r}} = \frac{(12\mathrm{V})^2}{4 \times 0.5\Omega} = 72\mathrm{W},$$

which agrees with the 72W found directly in part (b), confirming that 0.5Ω is the matched load.

(d) At the matched load the efficiency is only 50%: half the battery's power is burned inside the battery. The 4Ω load of part (a) delivers less power, 28.4W against 72W, but it does so at 88.9% efficiency, wasting only a ninth of the battery's energy internally. The matched load wins on power and loses on efficiency, exactly the trade-off the theorem describes.

Making Sense of the Answer: *The two loads tell the two halves of the story in numbers. The 4Ω load is the efficient regime: it is eight times the internal resistance, so most of the EMF lands on the load and only one part in nine is lost inside. The 0.5Ω load is the matched regime: it equals the internal resistance, so the EMF splits evenly and the battery heats up as much as the load does. The matched load draws more than twice the power but throws away half of everything to do it. Which one is right depends entirely on whether the battery is feeding a power-hungry device whose supply is cheap, or a circuit where battery life is precious.*

Think Like a Physicist: *When a problem mentions both maximum power and a specific internal resistance, the matched load is handed to you for free: it is simply* $R = r$*, with no calculation. The maximum power itself is then* $\frac{E^2}{4r}$ *without ever finding the current. Recognising the matched condition by sight, rather than re-deriving it each time, turns part (c) into a single substitution. The derivation is there to be understood once; the result is there to be used on sight ever after.*

HOT Example 42

A resistor wound from copper wire forms part of the control circuitry in a Tanzanian textile factory. Measured cold, in melting ice at 0°C, its resistance is 10Ω; in boiling water at 100°C its resistance is 14Ω.

In operation the resistor is connected across a 12V supply and is found to dissipate 6W. (a) Find the temperature coefficient of resistance of the copper. (b) Find the resistance under operating conditions. (c) Estimate the operating temperature, and comment on whether the resistor is safe.

Solution

(a) The resistance of a metal rises with temperature as $R_\theta = R_0(1+\alpha\theta)$, where R_0 is the resistance at $0^\circ C$ and θ is the temperature in degrees Celsius.

Using the boiling-water reading, with $R_0 = 10\Omega$ and $R_{100} = 14\Omega$ at $\theta = 100^\circ C$:

$$14\Omega = 10\Omega \times (1 + \alpha \times 100^\circ C)$$

$$1.4 = 1 + 100\alpha \;\Rightarrow\; \alpha = 4 \times 10^{-3}{}^\circ C^{-1}$$

(b) Under operation the resistor dissipates 6W across 12V. The resistance follows from the form that uses voltage and power:

$$R = \frac{V^2}{P} = \frac{(12V)^2}{6W} = 24\Omega$$

(c) Feeding $R = 24\Omega$ back into $R_\theta = R_0(1+\alpha\theta)$, with $R_0 = 10\Omega$ and $\alpha = 4 \times 10^{-3}{}^\circ C^{-1}$:

$$24\Omega = 10\Omega \times (1 + 0.004\,\theta)$$

$$\theta = 350^\circ C$$

A copper component sitting at $350^\circ C$ is in serious trouble. Copper oxidises rapidly in air well below that, and $350^\circ C$ is far above the temperature any insulation around the wire could survive. The resistor is not safe; it is running close to destruction, and the design that put a bare copper resistor where it must dissipate 6W has badly underestimated the heat.

Making Sense of the Answer: *The number tells a cautionary tale. The resistance has more than doubled from its ice-point value, from 10Ω to 24Ω, and for copper a doubling of resistance means a temperature rise of hundreds of degrees, because the coefficient is small. The same small coefficient that makes copper a stable conductor at room temperature makes a large resistance rise the signature of a dangerously large temperature rise. Six watts sounds trivial, the dim glow of an indicator lamp, but concentrated in a small copper coil with nowhere to shed its heat it drives the temperature to a level that would melt solder and char insulation.*

Think Like a Physicist: *This problem chains two separate pieces of the chapter: the temperature dependence of resistance from the earlier section, and the power dissipation of this one. The bridge between them is the operating resistance, found from the power data and then read backwards through the temperature law to recover the temperature. Whenever a problem gives you both a power-and-voltage operating condition and a resistance-temperature characteristic, that operating resistance is the hinge that joins them. Find it from the power side, then turn it into a temperature on the resistance side.*

REAL Example 43

A small heater for a college laboratory in Bagamoyo is to deliver 8W when connected across a 24V supply. It is to be wound from nichrome wire of cross-sectional area $2 \times 10^{-7} m^2$ and resistivity $1.1 \times 10^{-6}\Omega m$. (a) Find the resistance the element must have. (b) Find the length of nichrome wire required. (c) A technician decides to use wire of half the diameter instead, keeping the same power and voltage. Find the new length, and explain why it is shorter even though the wire is thinner.

Solution

(a) The element must dissipate 8W at 24V, so its resistance follows from voltage and power:

$$R = \frac{V^2}{P} = \frac{(24V)^2}{8W} = 72\Omega$$

(b) The resistance of a wire is $R = \frac{\rho l}{A}$, so the length is $l = \frac{RA}{\rho}$:

$$l = \frac{RA}{\rho} = \frac{72\Omega \times 2 \times 10^{-7} m^2}{1.1 \times 10^{-6}\Omega m} = 13.1m$$

(c) Halving the diameter quarters the cross-sectional area, because area depends on the square of the diameter; the new area is $\frac{A}{4} = 0.5 \times 10^{-7}\text{m}^2$. The power and voltage are unchanged, so the required resistance is still 72Ω. The new length is:

$$l = \frac{RA}{\rho} = \frac{72\Omega \times 0.5 \times 10^{-7}\text{m}^2}{1.1 \times 10^{-6}\Omega\text{m}} = 3.27\text{m}$$

The thinner wire needs only a quarter of the length, 3.27m against 13.1m. This seems backwards until the formula is read carefully. The design fixes the resistance at 72Ω. Thinning the wire to a quarter of its area would, on its own, quadruple the resistance; to bring it back down to 72Ω, the length must be cut to a quarter. A thinner wire reaches the same resistance in a shorter run.

Making Sense of the Answer: *The result overturns a common guess. People expect a thinner wire to need more length, picturing thin wire as more resistive per metre, which is true: the thinner wire has four times the resistance per metre. But the design fixes the total resistance, not the resistance per metre, and four times the resistance per metre means a quarter of the metres to reach the same total. Both elements are* 72Ω*; the thin one packs it into* 3.27m, *the thick one spreads it over* 13.1m. *For a coil that must fit inside a small heater, the thinner wire is the more compact choice, which is one reason heating elements are wound from fine wire.*

Think Like a Physicist: *The safe way through a problem like part (c) is to hold fixed what the design holds fixed, and let the formula tell you the rest. Here the power and voltage are fixed, so the resistance is fixed at* 72Ω*; only then do you ask what length the new area demands. The error to avoid is reasoning about resistance per metre, which is not what the design controls. Anchor on the quantity that is actually constrained, the total resistance, and the counter-intuitive answer collapses into a single line of arithmetic.*

REAL Example 44

An immersion heater at a Tanga substation rest-house is rated 2.5kW and runs on the 240V mains. (a) Find the current it draws and the resistance of its element. (b) Find the energy it delivers to the water in thirty minutes, in joules and in kilowatt-hours. (c) Explain why the element is wound from nichrome rather than copper.

Solution

(a) The current follows from $P = VI$, and the resistance from $\frac{V^2}{P}$:

$$I = \frac{P}{V} = \frac{2500\text{W}}{240\text{V}} = 10.4\text{A}$$

$$R = \frac{V^2}{P} = \frac{(240\text{V})^2}{2500\text{W}} = 23\Omega$$

(b) Energy is power times time. Thirty minutes is 1800s:

$$U = Pt = 2500\text{W} \times 1800\text{s} = 4500000\text{J} = 4.5\text{MJ}$$

In the billing unit, 2.5kW running for half an hour is:

$$U = 2.5\text{kW} \times 0.5\text{h} = 1.25\text{kWh}$$

(c) Nichrome is chosen for three reasons copper cannot meet at once. Its resistivity is about sixty times that of copper, so the 23Ω element is a short, compact coil rather than an impractically long wire. Its temperature coefficient of resistance is small, so the element draws nearly the same 2.5kW whether it starts cold or runs at full heat, instead of drifting as it warms. And it grows a protective oxide layer that lets it sit at high temperature in hot water and steam for years without corroding away. Copper fails all three: it would need an enormous length, its resistance and power would swing wildly from cold to hot, and it would oxidise quickly.

Making Sense of the Answer: *The figures are the everyday arithmetic of a power bill. Ten and a half amperes is a heavy household current, the reason immersion heaters need their own thick cable and a dedicated switch rather than sharing a thin lighting circuit. And* $1.25kWh$ *for half an hour of heating is, at any realistic tariff, a noticeable cost, which is why heating water is one of the larger lines on a domestic bill.*

The $4.5MJ$ *and the* $1.25kWh$ *are the same energy in two units: the joule for physics, the kilowatt-hour for the meter on the wall.*

Think Like a Physicist: *Energy problems reward keeping the units in plain sight. Working in joules means seconds, so thirty minutes had to become* $1800s$ *before multiplying by watts. Working in kilowatt-hours means hours and kilowatts, so the same calculation in those units is the cleaner one for a bill: kilowatts times hours, with no large powers of ten to carry. Knowing both routes, and which unit the question wants, is part of reading the problem correctly.*

REAL Example 45

TANESCO must deliver 1MW of power to a town 100km from a generating station. The transmission line has a total resistance of 100Ω. (a) If the power were sent at the household voltage of 240V, find the current in the line and the power lost in it, and comment. (b) If instead the power is sent at 132kV, find the current, the line loss, and the fraction of the transmitted power that is lost. (c) State the principle the comparison illustrates.

Solution

(a) At 240V, the current needed to carry 1MW is:

$$\mathrm{I} = \frac{\mathrm{P}}{\mathrm{V}} = \frac{1 \times 10^6\mathrm{W}}{240\mathrm{V}} = 4167\mathrm{A}$$

The power dissipated in the 100Ω line at that current is:

$$\mathrm{P_{line}} = \mathrm{I}^2\mathrm{R} = (4167\mathrm{A})^2 \times 100\Omega = 1.7 \times 10^9\mathrm{W}$$

Comment: This is practically impossible. The line would have to dissipate 1700MW of heat to deliver 1MW to the town, more than a thousand times the power being delivered. No conductor could survive it and no station could supply it. Sending bulk power at 240V is not merely wasteful; it cannot be done at all.

(b) At 132kV, the current carrying the same 1MW is far smaller:

$$\mathrm{I} = \frac{\mathrm{P}}{\mathrm{V}} = \frac{1 \times 10^6\mathrm{W}}{132000\mathrm{V}} = 7.6\mathrm{A}$$

$$\mathrm{P_{line}} = \mathrm{I}^2\mathrm{R} = (7.6\mathrm{A})^2 \times 100\Omega = 5.8 \times 10^3\mathrm{W} = 5.8\mathrm{kW}$$

The fraction of the transmitted power lost in the line is:

$$\frac{\mathrm{P_{line}}}{\mathrm{P}} = \frac{5800\mathrm{W}}{1 \times 10^6\mathrm{W}} = 0.0058 = 0.58\%$$

Less than one per cent is lost on the way. Same line, same distance, same power delivered, but now the transmission is entirely practical.

(c) The comparison illustrates that line loss falls as the **square** of the transmission voltage for a fixed delivered power, because raising the voltage lowers the current in proportion and the loss goes as the current squared. High-voltage transmission is not a refinement or an economy; it is the only way bulk power can cross a long line at all. This is why TANESCO transmits at 132kV and above, and why every long line begins and ends at a transformer.

Making Sense of the Answer: *The two parts are separated by a factor that looks impossible until the squares are counted. The voltage rose from* $240V$ *to* $132kV$*, a factor of* 550*. The current fell by the same factor of* 550*, from* $4167A$ *to* $7.6A$*. But the loss, going as the current squared, fell by* 550 *squared, more than three hundred thousand times, from* 1.7×10^9W *to* 5.8×10^3W*. A factor of* 550 *in voltage buys a factor of* 300000 *in reduced loss. That savage leverage of the square is the single fact that makes a national grid possible, and it is the same square that appeared in* $P = I^2R$ *at the very start of the section.*

Think Like a Physicist: *When a transmission problem gives you a delivered power and asks about line loss, do not start from the line resistance and a current you have not found yet. Start from* $P_{line} = \left(\frac{P}{V}\right)^2 R_{line}$*, which puts the loss directly in terms of the two things the problem fixes, the power delivered and the voltage of transmission, with the line resistance a constant. The current is an intermediate you can compute if asked,* *but the shape of the answer, loss falling as the inverse square of voltage, is clearest when the current is never made the hero of the calculation.*

Six problems down. If your answers matched ours, congratulations: you can now, in principle, bill an entire town for electricity. If they did not, the gap between your number and ours is precisely where the learning is hiding, so go back and hunt it down before reading on. Either way, the resistor has now surrendered every joule it intends to surrender.

CONDUCTION OF ELECTRICITY IN GASES

Every section until now has run its current through a metal, and a metal is a generous host: it comes with its conduction electrons already in place, a standing crowd waiting only for a field to tell them which way to drift. Air is not generous. Switch on a voltage across an ordinary air gap and, almost always, nothing happens at all. The air sits there as a perfect insulator, which is lucky, because it is the air between the wires in your house that stops every socket from short-circuiting through the room. Yet the same air, pushed hard enough, becomes the most violent conductor in nature and throws a lightning bolt across a kilometre of sky. This section is about that switch from insulator to conductor: why a gas refuses to carry current, what it takes to change its mind, and what the current looks like once it flows.

The empty queue

Think back to the picture from the start of the chapter, where a current was a queue of charge carriers passing buckets of energy down a line. In a metal the queue is already full. Every atom has donated an electron or two to a common pool, so the carriers are standing shoulder to shoulder before you apply any voltage; close the switch and the bucket is passed along the line at once.

A gas is the opposite. Its molecules are electrically neutral, each one holding tightly to all of its own electrons, so the queue is simply empty. There are no free charges to push, and a voltage with nothing to push moves no current. Before a gas can conduct, the carriers have to be created, and creating them means tearing electrons loose from neutral molecules. That act is called ionisation, it leaves behind a positive ion and a freed electron, and it does not happen for free: every electron is bound to its molecule, and prying one loose costs energy. The whole of this section follows from that single sentence. *A gas conducts only when something supplies the energy to fill its empty queue with ions.*

The breakdown potential

A gas, then, will not conduct until it is ionised, and the most direct way to ionise it is to turn up the voltage until the electric field itself does the tearing. Imagine the field as a hand pulling on the molecules' electrons. At low voltage the pull is gentle and the electrons stay put. Raise the voltage and somewhere there is a threshold where the pull wins, an electron is ripped free, and that electron, accelerated by the same field, slams into the next molecule and frees another. The gap suddenly fills with carriers and the gas, an insulator a volt ago, becomes a conductor.

The voltage at which this begins is the **breakdown potential** of the gap. It can be defined as *the minimum potential difference that must be applied across a gas to cause electrical breakdown, allowing the gas to become conducting and a sustained electric current to flow through it.*

Below the breakdown potential, the gas is an insulator; above it the gas conducts, sometimes quietly in a lamp and sometimes catastrophically in a spark. The value depends on:

- the nature of the gas,
- the spacing (distance) between the electrodes,
- the shape and condition of the electrodes,
- the gas pressure, and
- the temperature of the gas.

For two rounded electrodes about 1cm apart in air it takes roughly 30kV to strike a spark. Sharpen those electrodes to points at the same separation and the figure falls to about 12kV, because the field crowds together at a sharp point and reaches the breakdown value there while the average field is still modest. The same effect sends lightning to the tip of a tall tree rather than the flat ground beside it. One more lever matters more than any of these: lowering the pressure lowers the breakdown potential sharply, which is the entire reason the discharge tubes we are about to meet are run **not** at the pressure of the room but at a small fraction of it.

What does the ionising

Turning up the voltage is only one way to fill the empty queue. Anything that can deliver enough energy to a molecule to knock an electron off it will serve as an **ionising agent**, and they fall into three families.

1) **Nuclear radiation:** The alpha, beta, and gamma rays streaming from a radioactive source each carry far more energy than the dozen or so electron-volts it takes to ionise a gas molecule, so a single ray ionises a long trail of molecules as it passes.
2) **Electromagnetic radiation from the high-energy end of the spectrum:** X-rays, carrying from about 10^2 to 10^6eV per photon, and the more energetic ultraviolet, carrying from a few up to about 10^2eV, both ionise by the photoelectric effect, the photon spending its energy to eject an electron.
3) **Accelerated charged particles:** A free electron or proton that has fallen through a potential difference of only a few tens of volts picks up enough kinetic energy that, when it collides with a neutral molecule, it ionises it on impact. This third family is the important one, because it is the seed of the avalanche: *the very electrons that one ionising event sets free are themselves accelerated by the field into ionising agents for the next collision.*

Inside a discharge tube

The cleanest place to watch a gas conduct is a **discharge tube**: a glass tube with a metal electrode sealed into each end, the cathode and the anode, wired to a high-voltage supply, with a side tube to a vacuum pump so the pressure inside can be lowered step by step. At atmospheric pressure nothing is seen. But as the pump draws the pressure down, the gas lights up, and the pattern of light marches through a remarkable sequence of stages that is worth watching closely, because each stage is a snapshot of the same physics at a different density of molecules.

At about 20mmHg the discharge first appears as ragged brushes of light at the electrodes and faint violet streamers reaching between them.

Pump down to about 5mmHg and the tube organises itself: a long pink **positive column** fills most of the length and ends in a glow at the anode, a blue **negative glow** sits near the cathode, and between them lies a dark gap called the **Faraday dark space**.

Lower the pressure again to about 0.1mmHg and the positive column retreats toward the anode and breaks into a row of bright bands called **striations**; the dark spaces stretch out, and a new dark region, the **Crookes dark space**, opens up around the cathode.

Finally, near 0.01mmHg, the positive column and the negative glow have vanished altogether, the Crookes dark space has grown to fill the whole tube, and the glass itself fluoresces with a soft green light where invisible rays from the cathode strike it.

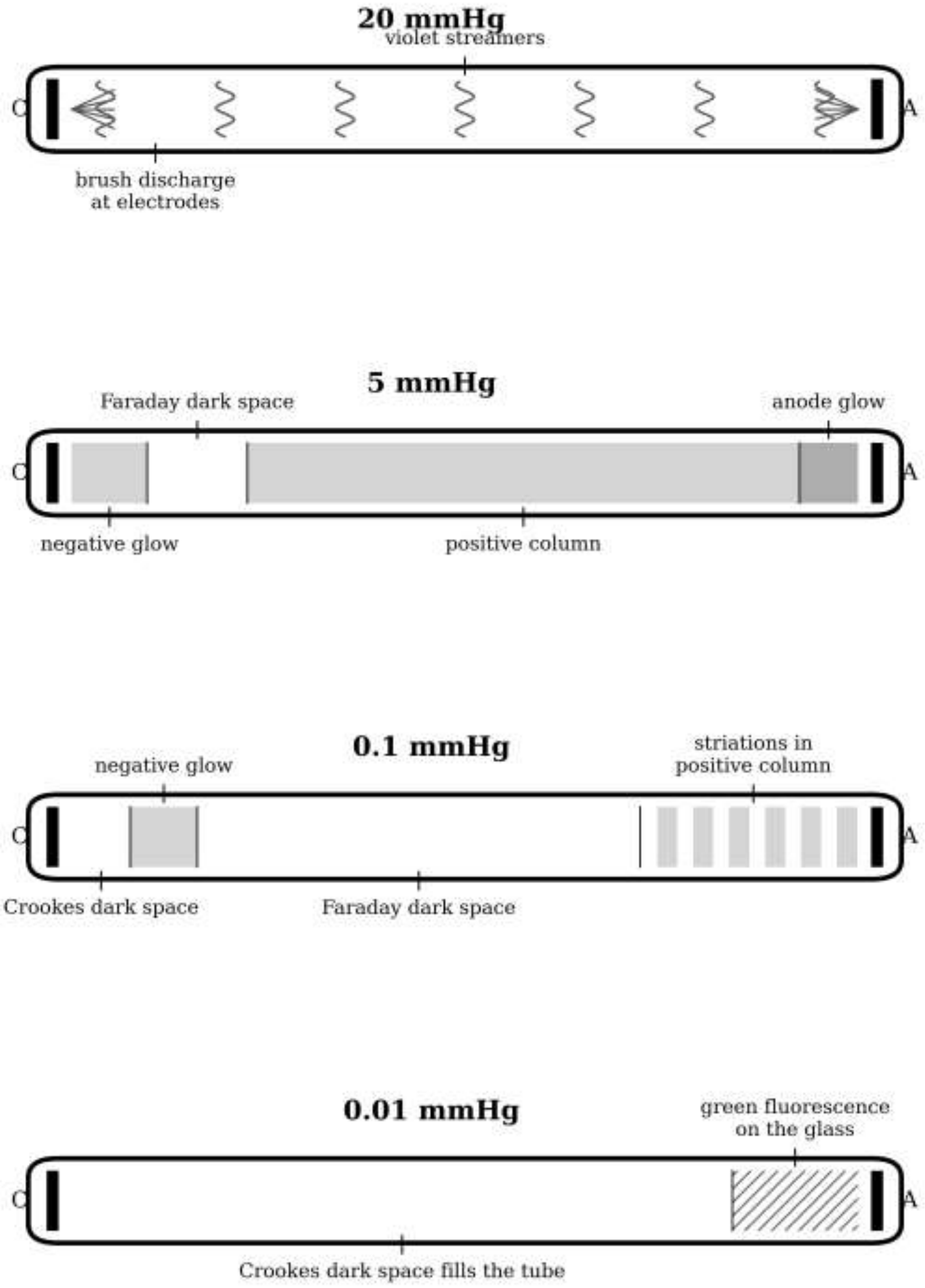


Figure: ***The appearance of the discharge in a tube as the pressure is lowered in four stages.*** *At* $20mmHg$ *the discharge is brush-like at the electrodes with violet streamers between. At* $5mmHg$ *a long positive column ends in an anode glow, with a negative glow near the cathode and a Faraday dark space between. At* $0.1mmHg$ *the positive column breaks into striations and the Crookes dark space appears at the cathode. At* $0.01mmHg$ *the Crookes dark space fills the tube and the glass walls fluoresce green. The cathode (C) is at the left and the anode (A) at the right of each tube.*

There is one experimental fact in this parade that is easy to miss and worth fixing in mind: *the length of the dark spaces depends only on the pressure, not on the length of the tube.* A longer tube does not stretch the Crookes dark space; it only lengthens the positive column to fill the extra room. The dark spaces are set by how far an electron travels between collisions, which is a matter of how crowded the gas is, and that is why the same demonstration tube can be pumped through every stage above and show each one faithfully.

Two kinds of discharge

Underneath all that scenery, a gas discharge is running in one of two fundamentally different ways, and telling them apart is the key idea of the section.

In an **initiated discharge** *the gas conducts only because some outside ionising agent, an ultraviolet lamp, an X-ray set, a radioactive source, is steadily manufacturing ions for it.* The applied voltage does an important job, sweeping those ions to the electrodes to make a current, but it does not itself create a single new ion. The proof is brutal and simple: switch off the external agent and the current dies within an instant, because the existing ions reach the electrodes or recombine and nothing replaces them. The gas is a conductor only on borrowed carriers.

In a **self-sustaining discharge** *the voltage has been raised so high that it takes over the manufacturing itself.* Now an electron accelerated by the field gains, in the short hop between one collision and the next, enough kinetic energy to ionise the molecule it strikes. That collision frees a second electron; the field accelerates both; each ionises another; and the numbers double and redouble in a runaway called **chain ionisation**. This is a positive-feedback avalanche, and once it is running the gas makes all its own carriers. You can switch the external agent off entirely and the discharge carries on, feeding itself.

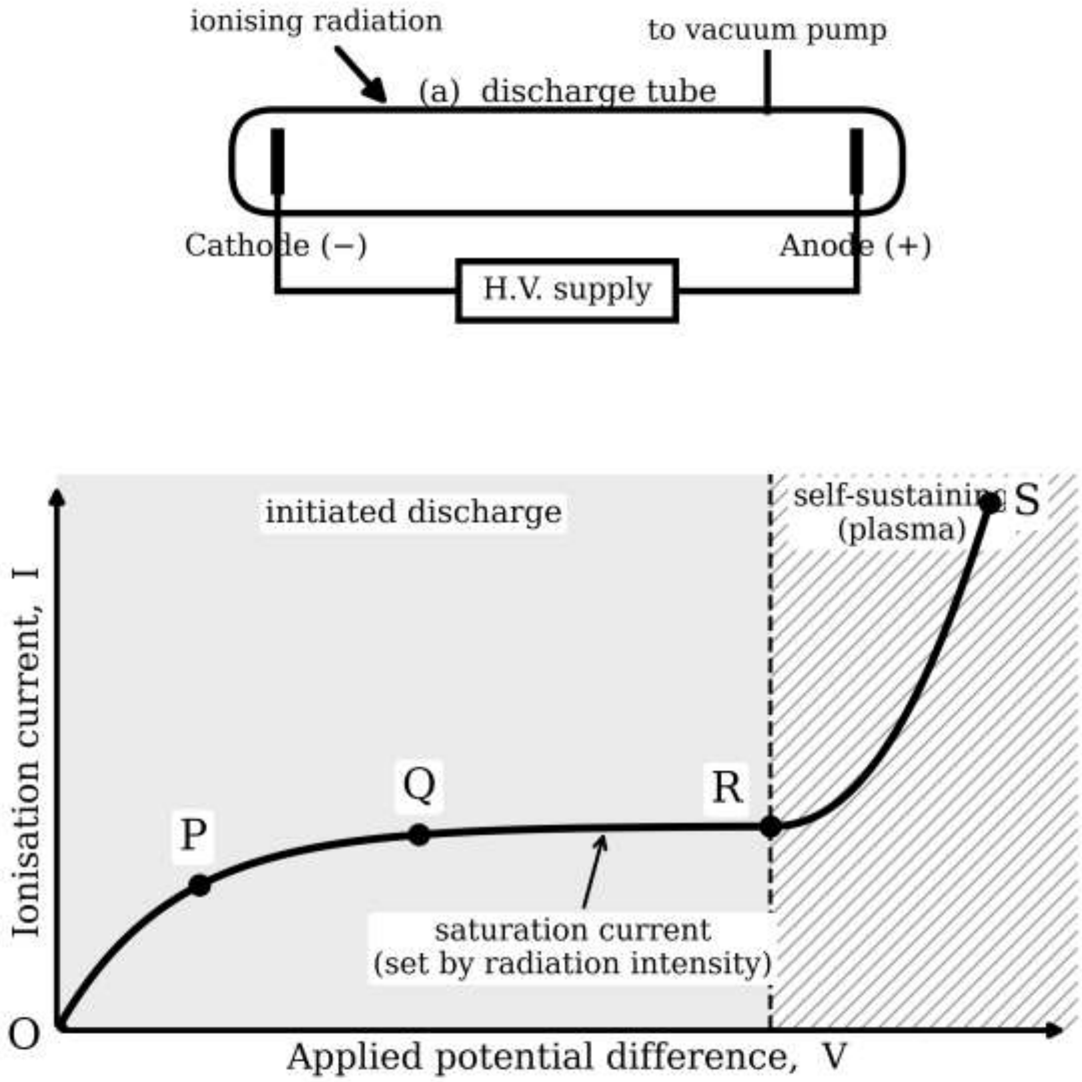
The difference between the two regimes is exactly the difference between a fire you must keep lighting with a match and a fire that has caught and now spreads on its own.

The distinction is not academic. A Geiger-Müller counter is deliberately run in the self-sustaining regime, so that one single ionising particle entering the tube is enough to trigger the avalanche and produce a pulse of current large enough to hear as a click. A fluorescent lamp, by contrast, runs just short of runaway, in the region where the chain barely begins, enough to keep the current flowing steadily without the avalanche tearing away.

The ionisation curve

Both regimes, and the boundary between them, show up on a single graph. Hold the intensity of an ionising agent fixed, raise the applied voltage from zero, and plot the current that flows. The result is the ionisation curve, traditionally labelled at five points from O to S.

- From O to P the voltage is low. The agent is making ions, but the field is too weak to sweep them across before many of them meet an opposite ion and recombine, so only a fraction reach the electrodes; raising the voltage rescues more of them, and the current climbs.
- From P to Q the field is now strong enough to collect ions faster than they can recombine, and the current climbs further.
- From Q to R something new happens: every single ion the agent produces is now being collected, none are left to rescue, and so raising the voltage further does nothing at all. The curve goes flat. This plateau is the **saturation current**, and its height is set not by the voltage but by the intensity of the ionising agent, a brighter source giving a higher plateau.
- Beyond R, the voltage becomes large enough for chain ionisation to ignite, the gas begins to make its own carriers, and the current turns sharply upward toward S with no sign of levelling off. To the left of R the discharge is initiated and borrowing its carriers; to the right of R it is self-sustaining and a true plasma.



(b) ionisation curve

Figure: ***(a) A discharge tube, with cathode and anode wired to a high-voltage supply, a port to the vacuum pump, and an external ionising agent. (b) The ionisation curve, the current* I *plotted against the applied potential difference* V *at fixed radiation intensity.*** *From O to P to Q the current rises as fewer ions recombine; from Q to R it saturates, every ion produced now being collected, at a height fixed by the radiation intensity; beyond R, toward S, chain ionisation sets in and the current climbs steeply. The region up to R is the initiated discharge; beyond R the discharge is self-sustaining.*

The physics of this section is told in pictures and words rather than equations, so these two examples ask you to reason rather than to calculate, which is its own kind of difficulty. Read each question, look away, and try to build the answer in your own words before you check it against the provided one. A gas-discharge device you can explain out loud is a gas-discharge device you understand.

REAL Example 46

Explain why a fluorescent tube in the Miono physics laboratory flickers for a moment when it is first switched on, but then settles into a steady glow.

Solution

Two things have to come together before the tube can light, and at switch-on neither is ready. The tube is filled with mercury vapour at low pressure, and while it is cold that vapour is sparse and its breakdown potential is high, higher than the mains voltage alone can supply. Meanwhile the carriers have to be seeded: small filaments at each end of the tube are heated by the initial current and begin to emit electrons, but it takes a moment for them to warm. To bridge the gap, a starter switch in the circuit closes and then snaps open, and at the instant it opens the choke, a coil in series with the tube, responds to the sudden collapse of current with a brief high-voltage pulse. That pulse, added to the mains, finally exceeds the cold breakdown potential and strikes the discharge. The visible flicker is exactly this struggle: the gas being driven across its threshold, perhaps stuttering through a failed attempt or two before it catches.

Once the discharge is running it warms the tube, and a warmer tube is a different device. The mercury vapour pressure rises, ionisation becomes easy, and the voltage needed to sustain the current falls well below the value needed to start it. The tube settles to a steady operating point partway along the ionisation curve, where the chain just sustains itself without running away, and the starter, no longer needed, stays open. The flicker belongs to starting; the steady glow belongs to running; and the difference between them is the difference between forcing a cold gas across its breakdown potential and coasting a warm one along its operating point.

Making Sense of the Answer: *The flicker is not a fault but a fingerprint of the physics. A cold gas has a high breakdown potential and a warm one a low operating voltage, so every fluorescent tube must briefly be over-driven to start and then runs comfortably under-driven, and the flicker is the visible seam between those two states. It is also why a tube near the end of its life flickers endlessly: its filaments no longer seed electrons well, so it never quite settles past the starting struggle.*

Think Like a Physicist: *Every gas-discharge device is a two-part problem, and naming the parts is the whole skill. One part supplies the carriers, here the heated filaments and the choke's voltage pulse; the other supplies the steady field that keeps them moving, here the mains across a now-warm vapour. When you meet any such device, sort its components into those that make carriers and those that drive them, and the operation explains itself.*

HOT Example 47

A Geiger-Müller counter operates with the gas between its electrodes held at the upper end of the R-to-S region of the ionisation curve. (a) Explain why this regime is required for the counter to register a single ionising particle. (b) Explain why the discharge extinguishes after each pulse, so that the next particle can be counted. (c) Identify the mechanism by which the discharge is quenched.

Solution

(a) A single ionising particle, on its own, frees only a handful of ions, far too few to drive a current any meter could notice. The counter therefore works in the self-sustaining region beyond R, where chain ionisation rules. Here that handful of initial ions is not the end of the story but the spark of an avalanche: each freed electron is accelerated hard enough to ionise again, the numbers double and redouble, and the single particle's faint mark is amplified into a current pulse large enough to register as an audible click. The counter does not detect the particle; it detects the avalanche the particle triggers.

(b) An avalanche that never stopped would leave the tube permanently conducting, blind to every particle after the first. So each pulse must be cut short. After the avalanche has produced its pulse, the discharge has to be extinguished, the gas returned to its non-conducting state, and the tube made ready for the next particle. Only then is the counter a counter rather than a one-shot switch.

(c) The extinguishing is done by a **quenching agent**, a small amount of a halogen or an organic vapour mixed into the tube's gas. During the avalanche, excited atoms emit ultraviolet photons, and those photons, striking the cathode or other atoms, would ignite fresh avalanches and keep the discharge alive indefinitely. The quenching vapour absorbs those ultraviolet photons before they can do so, starving the discharge of its secondary ignition and letting it die out after each pulse. Without a quenching agent the counter would latch on after a single count and never recover.

Making Sense of the Answer: *Amplification is the entire point of the device. One particle is unmeasurable and one avalanche is unmistakable, so the counter is built to convert the first into the second by living in the chain-ionisation regime. But raw amplification with no off-switch is useless, because a detector that fires once and sticks counts nothing, and the quenching gas is precisely the off-switch that turns a runaway into a repeatable pulse.*

Think Like a Physicist: *A detector of single events needs enormous gain, and enormous gain wants to run away, so every such device pairs its amplifier with a brake. Gain and reset are the two requirements, and* *they pull against each other exactly as raw power and efficiency did in the previous section. When you study any detector, ask the same two questions in turn: what makes the signal large, and what makes it stop, so the next signal can be seen.*

Two devices, no equations, and yet the same idea ran through both: a gas carries no current until something fills its empty queue with ions, and everything else, the flicker, the click, the colour, the quench, is a detail of how those ions are made and unmade. If you can now look at a glowing tube and say which part is making carriers and which part is driving them, the section has done its work.

The next section turns from how the gas conducts to the light it gives off while conducting. That glow is not a single colour but a precise pattern of colours, and the pattern is a fingerprint: unique to each gas, sharp enough to name an unknown sample in a laboratory, and steady enough to read the composition of a star from its light alone.

OPTICAL SPECTRA OF GASES

A discharge tube does not merely conduct; it glows, and the colour of that glow is the whole subject of this section. Neon signs burn orange-red, sodium street lamps a flat insistent yellow, mercury lamps a cold blue-white, and the colour is so dependable that a physicist can name the gas in a sealed tube from across a dark room. The previous section explained how a gas carries current; this one asks what the light it gives off while conducting can tell us. The answer is, astonishingly, almost everything about the gas, and the same trick read backwards tells us the make-up of stars no one will ever visit.

Why the glow has a colour

The first thing to clear away is a natural but wrong idea: that the colour of the glow is the colour of the gas, the way paint has a colour. It is not. The gas at rest is invisible. The colour appears only when the discharge pumps energy into the atoms, and it is the energy coming back out that we see.

Here is the mechanism, in the smallest number of moving parts. Inside every atom the electrons may occupy only certain fixed energy levels, never the gaps between them. The discharge knocks an electron up to a higher level; a moment later it falls back down, and as it falls it must shed exactly the energy it gained, which it does by emitting a single packet of light, a photon. The energy of that photon is fixed by the size of the drop, and the energy of a photon fixes its colour. A big drop gives a blue or violet photon; a small drop gives a red one. Because each element has its own private ladder of energy levels, unlike that of any other element, the set of drops it can make is unique, and so the set of colours it emits is unique. The glow is not a colour but a **pattern** of colours, and the pattern belongs to one element and no other. This is why the colour is a fingerprint: read the pattern and you have named the gas.

Emission and absorption

Spectra come in two complementary families, and the difference is simply whether the gas is giving light out or taking it in.

An **emission spectrum** is what we have been describing: a hot, excited gas radiating its own characteristic photons, seen as bright lines of colour against darkness.

An **absorption spectrum** is the mirror image. Shine white light, which contains every colour, through a cool gas, and the gas absorbs exactly the photons it would itself emit, lifting its electrons up the very same ladder. The light that comes through is then missing those colours, and the spectrum shows dark lines at precisely the wavelengths the gas would have emitted if it were hot.

The fingerprint is the same in both cases, printed in bright ink or in shadow. This mirror symmetry is the engine of astronomy: the cool outer gas of a star absorbs its own lines out of the hot light beneath, and reading those dark lines tells us which elements the star is made of, from across distances light itself takes thousands of years to cross.

Three kinds of emission spectra

Not every glowing source gives the same kind of spectrum, and the kind reveals what is doing the glowing.

A line spectrum *is a set of separate, narrow, bright lines on a dark background.* It is produced by a single-atom (monatomic) gas at low pressure, where each atom radiates alone and undisturbed, so only its own sharp transitions appear. This is the spectrum of the discharge tubes in this chapter, and the one drawn in the next figure.

A band spectrum *is a set of groups of lines, each group so closely crowded that it looks like a band.* It is produced by gases whose particles are molecules of more than one atom, such as oxygen or carbon dioxide, because a molecule can store energy not only in its electrons but in vibrations and rotations, splitting each transition into a dense cluster.

A **continuous spectrum** *is an unbroken smear of colour with no gaps at all, the full rainbow.* It is produced by hot solids, hot liquids, and gases dense enough that the atoms are packed shoulder to shoulder, their crowded energy levels blurring into one continuous span. The glowing filament of an old lamp gives a continuous spectrum; a low-pressure gas gives lines.

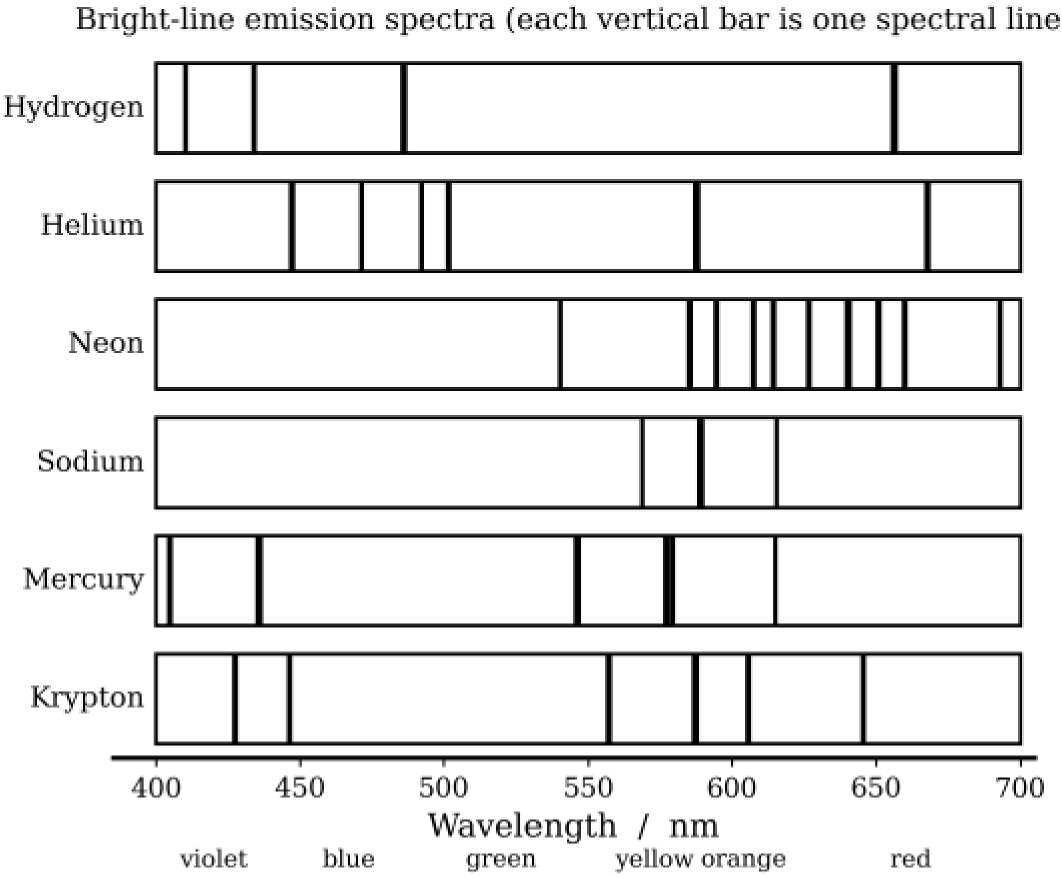


Figure: ***The bright-line emission spectra of six gases on a common wavelength axis from $400nm$ to $700nm$.*** *Each vertical bar is one spectral line, and the thicker bars are the stronger lines. No two patterns are alike: hydrogen shows its four Balmer lines, sodium is dominated by a single yellow line near* $589nm$*, mercury by a green line at* $546nm$*, and neon by a thicket of orange-red lines. The pattern, not any single line, identifies the gas.*

Reading the lines: the spectrometer

To turn that pattern into numbers we need an instrument that sorts light by wavelength and lets us measure each piece, and that instrument is the **spectrometer**. Before any formula, picture what it must do: take the mixed light from a discharge tube and fan it out so that each colour leaves in its own direction, then measure those directions.

The fanning out is done by a diffraction grating, a plate ruled with many thousands of fine parallel lines, met already in the chapter on waves. Light from the tube is passed through a narrow slit and a collimator that makes it into a parallel beam, the beam strikes the grating, and the grating sends each wavelength off at its own angle. A telescope mounted on a calibrated angular scale is then swung round to find each line and read the angle at which it sits. The angle and the wavelength are tied together by the grating equation from the wave chapter,

$$m\lambda = d\sin\theta$$

where d is the spacing between neighbouring rulings of the grating, θ is the angle of the line from the straight-through direction, λ is the wavelength, and m is the order, a whole number counting how many times the pattern repeats out from the centre. Measure θ for a line, and the equation hands you λ. The spectrometer turns a colour you can only describe into a wavelength you can write down.

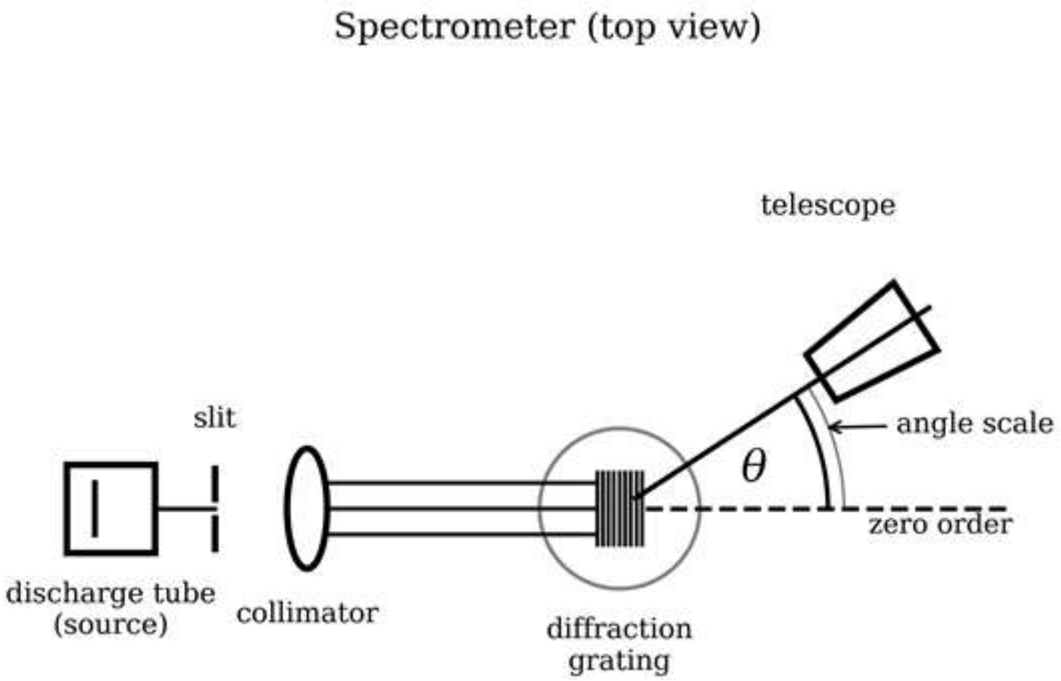


Figure: ***A spectrometer seen from above.*** *Light from the discharge tube passes through a slit and a collimator, which forms it into a parallel beam, and falls on a diffraction grating mounted on a turntable. The grating sends each wavelength off at its own angle; the straight-through direction is the zero order. A telescope on a calibrated angular scale is swung to the angle* θ *of each line, and the grating equation* $m\lambda = d\sin\theta$ *converts that angle into a wavelength.*

A table for naming the gas

Once the wavelengths are measured they are compared against known values, and identification is simply matching the pattern. The brightest visible lines of a few common gases are collected below; an unknown tube is named by measuring its lines and finding the row they fit.

Gas	**Brightest lines in the visible range: wavelength / nm (colour)**
Hydrogen	410 (violet), 434 (violet), 486 (blue-green), 656 (red)
Helium	447 (blue), 502 (green), 588 (yellow), 668 (red)
Neon	585 to 640 (a thicket of orange-red lines), 640 (red, strong)
Mercury	405 (violet), 436 (blue), 546 (green, strong), 577 and 579 (yellow)
Krypton	427 (violet), 557 (green), 587 (yellow), 606 (orange)

After a section told almost entirely in pictures, here is the one place the grating earns its keep with a calculation. Work it before you read the solution; the grating equation is short, but it has a way of hiding an extra answer or two from anyone in a hurry.

HOT Example 48

Light from a hydrogen discharge tube is examined with a diffraction grating ruled with 600 lines per millimetre. The blue-green hydrogen line, of wavelength 486nm, is one of the lines seen. (a) Find the spacing of the grating. (b) Find the angle at which the blue-green line appears in the first order. (c) Determine whether the same line can also be seen in the second order, and if so, at what angle.

Solution

(a) The grating is ruled with 600 lines in every millimetre, so the spacing between neighbouring lines is one millimetre divided by 600:

$$\mathrm{d} = \frac{1\mathrm{mm}}{600} = \frac{1 \times 10^{-3}\mathrm{m}}{600} = 1.67 \times 10^{-6}\mathrm{m} = 1.67\mu\mathrm{m}$$

(b) The grating equation is $\mathrm{m}\lambda = \mathrm{dsin}\theta$. Rearranged for the angle and used in the first order, with $\mathrm{m} = 1$:

$$\sin\theta = \frac{\mathrm{m}\lambda}{\mathrm{d}} = \frac{1 \times 486 \times 10^{-9}\mathrm{m}}{1.67 \times 10^{-6}\mathrm{m}} = 0.292$$

$$\theta_1 = \sin^{-1}(0.292) = 17.0°$$

(c) For the second order, set $\mathrm{m} = 2$. The right-hand side simply doubles:

$$\sin\theta = \frac{\mathrm{m}\lambda}{\mathrm{d}} = \frac{2 \times 486 \times 10^{-9}\mathrm{m}}{1.67 \times 10^{-6}\mathrm{m}} = 0.583$$

Since 0.583 is less than 1, an angle exists and the line is indeed visible in the second order:

$$\theta_2 = \sin^{-1}(0.583) = 35.7°$$

So the blue-green line appears twice, once at 17.0° and again, more widely deflected, at 35.7°.

Making Sense of the Answer: *The same line turning up at two angles is not a contradiction but the meaning of the order number m. The grating repeats its pattern outward from the centre, and each repeat throws every wavelength a little further out, so a single colour appears once in the first order and again, further from the centre, in the second. In fact this line also satisfies the equation for the third order, where* $sin\theta = 0.875$ *and* $\theta_3 = 61.0°$*; only when* $\frac{m\lambda}{d}$ *would exceed* 1 *does the line run out of orders, because no real angle has a sine greater than one.*

Think Like a Physicist: *Whenever a grating problem asks for "the angle," pause and ask "in which order?", because the honest answer is usually several angles, not one. The quantity* $\frac{m\lambda}{d}$ *is a sine, and a sine is capped at* 1*, so the test for whether an order exists is simply whether* $\frac{m\lambda}{d} \leq 1$*. Counting the orders this way, rather than stopping at the first, is what separates a full answer from half of one, and it is the same discipline of checking the limits of a formula that served us with maximum power and with high-voltage transmission earlier in the chapter.*

One short calculation, and the loop is closed: the colour we could only point at across a dark room has become a wavelength in nanometres, and a wavelength in nanometres has become the name of an element. The pattern that the eye reads as "orange-red" or "cold blue" is, to a spectrometer, a list of numbers as specific as a signature, and that is why the light of a flame, a lamp, or a distant star can be made to confess what it is made of.

The chapter is nearly done, and it has now built three things in turn: how charge moves through metals, how it forces its way through gases, and how the light from those glowing gases can be read. The final section spends all of it at once, walking through the everyday devices, the kettle and the fuse, the fluorescent tube and the lightning overhead, in which this chapter's physics is quietly at work.

APPLICATIONS OF CURRENT ELECTRICITY

Current electricity is at its most convincing when it stops being equations and starts being objects. This section takes the devices of everyday Tanzanian life and reads each one through the physics the chapter has built, sorting them into three families: those that work by conduction through a metal, those that work by conduction through a gas, and a few that quietly need both at once. The aim is not to add new theory but to watch the old theory earn its keep, in a kettle, a street lamp, a welder's arc, and the storm overhead.

Applications of conduction in metals

1. Electric kettle and immersion heater

The most familiar electrical device is also the simplest: a coil of nichrome or constantan, chosen through the relation $R = \frac{V^2}{P}$ so that it draws its rated power from the mains and pours that power into the water as heat at the rate $P = \frac{V^2}{R}$ until it boils. The kettle that warms a household's first cup of the day, and the immersion heater that fills the bath, are exactly the kind of element we sized earlier in the chapter; the only design decision is the resistance, and the resistance is fixed by the power and the voltage.

2. Electric iron

An electric iron is that same heated coil with one clever addition, a thermostat made from a bimetallic strip. The strip bends as it warms and opens the circuit when the soleplate reaches the chosen temperature, then closes it again as it cools, switching the Joule heating on and off so that a school uniform can be pressed at a steady heat without scorching. The element supplies the warmth; the bimetallic strip supplies the judgement.

3. Filament lamp

The oldest way of making light from electricity pushes the heating to its limit. A fine tungsten coil is run white-hot, at about 2500K, glowing simply because it is too hot not to; tungsten is chosen for its enormous melting point, above $3400\,^{\circ}\text{C}$, and the coil is sealed in a bulb of inert argon so that it cannot burn away in air. Its weakness is that most of its energy escapes as invisible infrared heat rather than light, which is why the filament lamp is steadily giving way across Tanzania to the cooler fluorescent and LED lamps that do the same work for a fraction of the current.

4. Fuse and circuit breaker

Not every use of heating is meant to be seen. A fuse is a deliberate weak point built into a circuit for its own protection, a short strand of low-melting lead-tin alloy sized so that its $P = I^2R$ heating keeps it just below melting at the safe current but melts it the instant the current climbs too high, breaking the circuit before a fault can set the wiring alight; Tanzanian homes carry fuses rated at 5A, 13A, and 30A for the lighting, the sockets, and the cooker. A circuit breaker does the same duty more conveniently, tripping a switch with an electromagnet or a bimetallic strip and resetting once the fault is cleared, instead of being replaced.

5. High-voltage power transmission

On a national scale the same concern over heating decides whether power reaches the town at all. Because the loss in a line grows with the square of its current, the electricity TANESCO generates is stepped up to 132kV, and on the newer southern interconnection to 220kV, for its long journey across the country, then stepped back down at the substations to the gentle voltage that enters the house. As the chapter showed in full, low-voltage transmission over Tanzanian distances is not merely wasteful but physically impossible; the high-voltage line and its transformers are the only way bulk power can travel.

6. Multimeter shunts and multipliers

Electricity also measures itself. The multimeter on the laboratory bench reads current and voltage across wide ranges by switching extra resistors around its sensitive moving coil: a small resistor in parallel, a shunt, diverts most of the current and widens the ammeter, while a large resistor in series, a multiplier, drops most of the voltage and widens the voltmeter. Turning the range knob simply selects which shunt or multiplier is in play, exactly as the chapter developed.

7. Electroplating

A steady current can also build. In the metal workshops of every town, a direct current passed through a solution carries dissolved metal onto the surface of an object held at the cathode, plating jewellery in silver or gold and tools in chromium. The mass laid down in a given time follows Faraday's law of electrolysis, $m = \frac{MIt}{nF}$, in which M is the molar mass, n the charge on each ion, F the Faraday constant, and It the total charge; the very charge we have tracked all chapter is here, quite literally, ferrying metal from one place to another.

Applications of conduction in gases

1. Fluorescent lighting

The second family makes light not by heating a wire but by lighting a gas, and its brightest member hangs on the wall of the laboratory itself. A fluorescent lamp is a low-pressure mercury-vapour tube whose discharge floods out invisible ultraviolet, which a phosphor painted inside the glass drinks in and returns as visible white light, far more efficiently than any glowing filament.

It is worth taking apart in the mind, because it holds nearly the whole chapter inside one glass tube, as Mr. Akilikubwa was fond of showing.

Mr. Akilikubwa (taking the spare tube down from the wall): *The lamp that lit up at the very start of the chapter will now show you the difference between metal conduction and gas conduction inside a single object.*

Kipanga: *It is only a long lamp, sir.*

Mr. Akilikubwa: *Look at the small coils sealed into its two ends. They are metal filaments, like the inside of a torch bulb. When the switch closes, current heats them, and the hot metal boils off electrons by thermionic emission, spilling free electrons into the low-pressure gas. Only now does the gas have carriers. A high-voltage pulse from the choke ionises a path between the ends, the discharge strikes, and ultraviolet light pours from the glowing mercury vapour. That light is invisible, so the glass is coated with a phosphor that swallows it and gives it back as the white you read by.*

Kipute: *So the one tube uses metal conduction in the filaments, gas conduction in the vapour, and something else again in the phosphor?*

Mr. Akilikubwa: *All three at once. The chapter has prepared you for the first two; the third, the way the phosphor turns one colour into another, waits for you in the atomic-physics chapter.*

2. Vapour discharge lamps

Strip away the phosphor and the bare colour of the gas shines through. A sodium lamp burns the hard orange of its own 589nm line over the trunk roads at night, a neon sign glows red in a Dar es Salaam shop window, and a mercury lamp throws a cold bluish-white across a factory yard. In each, the colour is nothing but the emission spectrum of the chosen gas, the fingerprint of the previous section pressed into service as signboard and street light.

3. Radiation detection: The Geiger-Müller counter

Pushed to its self-sustaining extreme, the same discharge becomes a counter rather than a lamp. The Geiger-Müller tube, met earlier, runs so near to runaway that a single ionising particle entering it triggers a whole avalanche and a current pulse loud enough to count as an audible click. This amplification of one particle into one click is why such tubes guard the staff at the Ocean Road Cancer Institute in Dar es Salaam and the radiotherapy unit at Bugando in Mwanza, wherever radioactive material is handled.

3. Plasma display panels

The same gas-discharge idea, shrunk to the size of a dot and repeated a million times across a screen, once lit every pixel of a plasma television: each cell is a microscopic discharge whose ultraviolet excites a coloured phosphor, the fluorescent lamp in miniature. Newer screens have largely replaced it, but the physics still glows in some industrial display panels.

Bridging applications: Devices that need both

A few devices put metal conduction and gas conduction to work together, in series, each doing what the other cannot.

1. Arc welding

In arc welding a fierce arc, a discharge through the open air at thousands of degrees, leaps between a metal electrode and the metal work and fuses them where it lands. The electrodes are metal conduction and the arc is gas conduction, so the welder crouched behind a dark visor at a roadside workshop is running both halves of this chapter at once.

2. Sodium street lamps with a metallic ballast

A sodium street lamp cannot run on its discharge alone. Wired in series with the glowing vapour is a metal ballast, a resistor or coil, that throttles the current the discharge would otherwise let run away and destroy itself the instant it strikes. The gas half makes the light; the metal half keeps it alive.

3. Electric arc furnace

In the electric arc furnaces of the scrap-metal yards, an arc struck between great graphite electrodes and a charge of old iron pours its heat into the metal until it runs liquid. The graphite electrodes carry the current in, and the ionised gas of the arc carries the heat, thousands of degrees of it, into the charge. Metal conduction delivers the energy; gas conduction spends it.

The physics of the chapter is now fully spent, so the last example asks the one question every household actually cares about: not how the current flows, but what it costs by the end of the month. Add the numbers yourself first, because this is the rare example you can check against the bill on your own wall.

REAL Example 49

A Tanzanian household runs the following from the 240V mains each day: a 2.5kW kettle for 12 minutes, three 9W LED bulbs for 6 hours, a 1.8kW iron for 25 minutes, and a 150W refrigerator that runs for a total of 8 hours. (a) Find the total electrical energy used each day, in kilowatt-hours. (b) At a TANESCO tariff of 290TZS for each kilowatt-hour, find the bill for a 30-day month.

Solution

(a) Energy is power times time, and if the power is in kilowatts and the time in hours, the product comes out directly in kilowatt-hours. Taking the appliances one at a time:

$$\mathrm{E_{kettle}} = 2.5\mathrm{kW} \times \frac{12}{60}\mathrm{h} = 0.5\mathrm{kWh}$$

$$\mathrm{E_{bulbs}} = 3 \times 9\mathrm{W} = 27\mathrm{W} = 0.027\mathrm{kW} \times 6\mathrm{h} = 0.162\mathrm{kWh}$$

$$\mathrm{E_{iron}} = 1.8\mathrm{kW} \times \frac{25}{60}\mathrm{h} = 0.75\mathrm{kWh}$$

$$\mathrm{E_{fridge}} = 0.15\mathrm{kW} \times 8\mathrm{h} = 1.2\mathrm{kWh}$$

Adding the four gives the energy used in one day:

$$\mathrm{E_{day}} = 0.5\mathrm{kWh} + 0.162\mathrm{kWh} + 0.75\mathrm{kWh} + 1.2\mathrm{kWh} = 2.61\mathrm{kWh}$$

(b) A 30-day month uses thirty times the daily energy, and the bill is that energy multiplied by the tariff:

$$\mathrm{E_{month}} = 2.612\mathrm{kWh} \times 30 = 78.4\mathrm{kWh}$$

$$\mathrm{Bill} = 78.4\mathrm{kWh} \times 290\mathrm{TZS/kWh} = 22720\mathrm{TZS}$$

Making Sense of the Answer: *The bill is dominated by exactly the appliances that turn electricity into heat. The kettle and the iron between them, running for barely half an hour a day, account for almost half the energy, while three LED bulbs glowing all evening cost less than either, because a watt is a watt and the bulbs simply draw very few of them. The refrigerator is the quiet surprise: low-powered, but running a third of the day, it climbs to the largest single share. Energy is power multiplied by time, and time is the variable households forget.*

Think Like a Physicist: *The whole calculation is the kilowatt-hour taken seriously: kilowatts multiplied by hours, with no large powers of ten to carry, which is precisely why the electricity meter uses that unit rather than the joule. When an energy question quotes powers in watts and kilowatts and times in minutes and hours, convert everything to kilowatts and hours first, and the arithmetic becomes the simple multiplication you see here. The hardest part of a real bill is not the physics; it is remembering to count the hours honestly.*

There it is, the chapter delivered to your own front door, with a number on it in shillings. Every line of that bill is a resistor, a current, or a discharge doing exactly what the chapter said it would. Yet the story of electric charge is not confined to wires, meters, and household appliances.

Mr. Akilikubwa: *Before we close the chapter, look out of the window. A storm has gathered over the Indian Ocean to the east of Pwani, and if you watch closely you will see a bright fork of lightning streak from cloud* *to ground in a fraction of a second. The charge between the cloud's base and the ground has built until the electric field reached the breakdown value of air, around three million volts per metre. At that point the air ionises in a thin channel, becomes a plasma, and carries a current of tens of thousands of amperes in under a millisecond.*

From the bill on your table to the lightning in the sky, the same ideas have been at work throughout this chapter. That is the chapter's tidy account of itself, and it reads beautifully. Whether any of it has actually lodged in you is a different question entirely, and the miscellaneous worked examples that follow are where the two of us find out together.

MISCELLANEOUS WORKED EXAMPLES ON CURRENT ELECTRICITY

Example 50

(a) Explain why a copper wire conducts electricity readily whereas a glass rod does not.
(b) A conductor carries a steady current of 0.50A when a potential difference of 2.0V is maintained across it. Determine its conductance and its resistance.

Solution

(a) A metal such as copper contains a large number of free electrons that are not bound to any particular atom but are free to move throughout the lattice. When a potential difference is applied across the metal, these free electrons drift through it and so constitute a current. In glass, by contrast, every electron is tightly held within its atoms and bonds, so there are almost no free charge carriers available to move. Since a current requires mobile charge carriers, copper conducts readily whereas glass does not.

(b) Conductance is the reciprocal of resistance, and follows directly from the current and voltage:

$$G = \frac{I}{V} = \frac{0.50\text{A}}{2.0\text{V}} = 0.25\text{S}$$

The resistance is the reciprocal of the conductance:

$$R = \frac{1}{G} = \frac{1}{0.25\text{S}} = 4.0\Omega$$

The conductor has a conductance of 0.25S and a resistance of 4.0Ω.

Example 51

(a) Distinguish between an ohmic and a non-ohmic conductor, giving one example of each.
(b) A filament lamp carries a current of 0.5A when 2.0V is applied across it, and 1.0A when 8.0V is applied. Determine its resistance at each setting and account for the difference.

Solution

(a) An ohmic conductor obeys Ohm's law, so the current through it is directly proportional to the potential difference across it at constant temperature; its current-voltage graph is a straight line through the origin and its resistance stays constant. A non-ohmic conductor does not show this proportionality, so its current-voltage graph is curved and its resistance changes with the operating conditions. A metallic resistor at constant temperature is ohmic, whereas a filament lamp is non-ohmic.

(b) At the first setting the resistance follows from Ohm's law:

$$R_1 = \frac{V_1}{I_1} = \frac{2.0\text{V}}{0.5\text{A}} = 4.0\Omega$$

At the second setting:

$$R_2 = \frac{V_2}{I_2} = \frac{8.0\text{V}}{1.0\text{A}} = 8.0\Omega$$

The resistance rises from 4.0Ω to 8.0Ω because the larger voltage drives a larger current, which dissipates more power and raises the filament's temperature; and since the resistance of a metal increases with temperature, the hotter filament has the greater resistance. The lamp is therefore non-ohmic.

Example 52

(a) Explain why the resistance of a metal increases with temperature while that of a semiconductor decreases.

(b) The platinum coil of a resistance thermometer has resistance 20.0Ω at 0°C and 27.8Ω at 100°C. When placed in a furnace its resistance is 51.2Ω. Determine the temperature of the furnace.

Solution

(a) In a metal the number of free electrons is essentially fixed, so conduction depends only on how freely those electrons move. As the temperature rises, the lattice ions vibrate more vigorously, the electrons collide with them more frequently, the relaxation time falls, and so the resistance increases. In a semiconductor, however, raising the temperature releases many additional charge carriers, and this large increase in the number of carriers outweighs the increased scattering, so the resistance decreases.

(b) The temperature coefficient of resistance follows from the two fixed-point readings:

$$\alpha = \frac{R_{100} - R_0}{R_0\theta} = \frac{27.8\Omega - 20.0\Omega}{20.0\Omega \times 100^\circ\text{C}} = 3.9 \times 10^{-3\circ}\text{C}^{-1}$$

Applying $R_\theta = R_0(1 + \alpha\theta)$ to the furnace reading:

$$51.2\Omega = 20.0\Omega \times (1 + 3.9 \times 10^{-3}\theta)$$

$$2.56 = 1 + 3.9 \times 10^{-3}\theta \ \Rightarrow \ \theta = \frac{1.56}{3.9 \times 10^{-3}} = 400^\circ\text{C}$$

The furnace temperature is 400°C.

Example 53

(a) Explain why the headlamps of a car dim momentarily when the starter motor is engaged.

(b) A 12V car battery of internal resistance 0.04Ω supplies 150A to the starter motor during cranking. Determine the terminal voltage during cranking and the power dissipated inside the battery.

Solution

(a) The starter motor draws a very large current from the battery. Because the battery has internal resistance, this large current produces a large voltage drop inside the battery, so the terminal voltage falls well below the EMF. The headlamps are connected across these same terminals, so the reduced terminal voltage supplies less power to them and they dim. Once the engine fires and the starter disengages, the current drops sharply, the internal voltage drop becomes small again, and the lamps return to full brightness.

(b) The terminal voltage during cranking follows from $V = E - Ir$:

$$V = E - Ir = 12\text{V} - (150\text{A})(0.04\Omega) = 6.0\text{V}$$

The power dissipated inside the battery is the Joule heating in the internal resistance:

$$P = I^2r = (150\text{A})^2(0.04\Omega) = 900\text{W}$$

The terminal voltage falls to 6.0V during cranking, and 900W is dissipated as heat inside the battery.

Example 54

(a) A network containing more than one source cannot be solved by series and parallel reduction alone. Name the two laws of Kirchhoff that make such a network solvable, and identify the conservation principle that each one expresses.

(b) Two batteries are connected with their positive terminals joined together and their negative terminals joined together, so that they drive a common external resistor of 2Ω. The first has EMF 8V and internal resistance 1Ω; the second has EMF 7V and internal resistance 1Ω. Using Kirchhoff's laws, determine the current supplied by each battery and the current in the external resistor.

Solution

(a) The first is Kirchhoff's current law, which states that the sum of the currents entering any junction equals the sum of the currents leaving it; it expresses the conservation of electric charge, since charge can neither accumulate at nor be created at a junction.

The second is Kirchhoff's voltage law, which states that around any closed loop the sum of the EMFs equals the sum of the potential drops; it expresses the conservation of energy, since a unit charge taken once round a loop must return to its starting potential.

(b) Let I_1 and I_2 be the currents from the first and second batteries, and let I_3 be the current in the external resistor. Applying Kirchhoff's current law at the junction:

$$I_1 + I_2 = I_3 \ldots \text{(i)}$$

Applying Kirchhoff's voltage law to the loop containing the first battery and the resistor, with $r_1 = 1\Omega$ and $R = 2\Omega$:

$$E_1 = I_1 r_1 + I_3 R \;\Rightarrow\; 8\text{V} = (1\Omega)I_1 + (2\Omega)I_3 \ldots\ldots \text{(ii)}$$

Applying it to the loop containing the second battery and the resistor, with $r_2 = 1\Omega$:

$$E_2 = I_2 r_2 + I_3 R \;\Rightarrow\; 7\text{V} = (1\Omega)I_2 + (2\Omega)I_3 \ldots\ldots \text{(iii)}$$

Solving the three equations gives:

$$I_1 = 2\text{A}, \qquad I_2 = 1\text{A}, \qquad I_3 = 3\text{A}$$

The first battery supplies 2A, the second supplies 1A, and the external resistor carries their sum, 3A.

Example 55

(a) Joule's law of heating and Ohm's law are sometimes confused. Explain what each law describes and how the two differ.

(b) A battery of EMF 6V and internal resistance 0.5Ω is short-circuited by a length of wire of negligible resistance. Determine the current that flows and the rate at which heat is produced inside the battery, and comment on the hazard.

Solution

(a) Ohm's law relates current to voltage: it states that the current through an ohmic conductor is directly proportional to the potential difference across it at constant temperature, so that $V = IR$. Joule's law of heating, by contrast, relates power to current: it states that the rate at which a resistor converts electrical energy into heat is proportional to the square of the current, so that $P = I^2R$. The two therefore answer different questions, Ohm's law telling us how much current flows and Joule's law telling us how much heat that current produces, and while Ohm's law holds only for ohmic conductors, Joule's law holds for any resistor whatever.

(b) With the external resistance negligible, the only resistance limiting the current is the internal resistance of the battery, so the short-circuit current is:

$$I = I_{\text{short}} = \frac{E}{r} = \frac{6\text{V}}{0.5\Omega} = 12\text{A}$$

The rate of heat production inside the battery is the Joule heating in its internal resistance:

$$P = I^2 r = (12\text{A})^2(0.5\Omega) = 72\text{W}$$

A current of 12A flows and 72W is produced inside the battery. This is dangerous: the whole of the battery's power is now dumped as heat within the cell itself, so it overheats rapidly and may leak, rupture, or explode, which is why a short circuit must never be allowed.

Example 56

(a) In everyday speech the words "energy" and "power" are used loosely. Distinguish between electrical energy and electrical power as the terms are used in physics, and give the SI unit of each.

(b) A 2.0kW electric kettle heats 1.5kg of water from 25°C to 100°C. Assuming no heat is lost to the surroundings, determine the time taken. Take the specific heat capacity of water as $4200\text{Jkg}^{-1}\text{K}^{-1}$.

Solution

(a) Electrical energy is the total work done by an electric current in transferring or converting charge's energy, and its SI unit is the joule. Electrical power is the rate at which that energy is supplied or converted, that is the energy per unit time, and its SI unit is the watt, where one watt equals one joule per second. Energy is therefore the total amount delivered, whereas power is how fast it is delivered.

(b) The heat the water must absorb follows from the specific heat capacity, with a temperature rise of $\Delta\theta = 100^\circ\mathrm{C} - 25^\circ\mathrm{C} = 75\mathrm{K}$:

$$\mathrm{Q} = \mathrm{mc}\Delta\theta = (1.5\mathrm{kg})(4200\mathrm{Jkg^{-1}K^{-1}})(75\mathrm{K}) = 472500\mathrm{J}$$

Since the kettle supplies energy at the steady rate of $2.0\mathrm{kW}$, the time to deliver this heat is:

$$\mathrm{t} = \frac{\mathrm{Q}}{\mathrm{P}} = \frac{472500\mathrm{J}}{2000\mathrm{W}} = 236\mathrm{s}$$

The kettle takes about $236\mathrm{s}$, which is approximately $3.9\mathrm{min}$, to bring the water to the boil.

Example 57

(a) Explain what is meant by the figure of merit of a galvanometer, and state how it is related to the sensitivity of the instrument.
(b) A galvanometer gives a full-scale deflection of 25 divisions when a current of $2.0\mathrm{mA}$ passes through it. Determine its figure of merit. Given that its coil has a resistance of 50Ω, determine the potential difference across it at full-scale deflection.

Solution

(a) **The figure of merit** of a galvanometer *is the current required to produce a deflection of one division on its scale*. It is the reciprocal measure of sensitivity: a galvanometer with a small figure of merit needs only a tiny current for each division and is therefore highly sensitive, whereas a large figure of merit means a large current is needed per division and the instrument is less sensitive.

(b) The figure of merit is the current divided by the number of divisions it deflects:

$$\mathrm{k} = \frac{\mathrm{I}}{\mathrm{n}} = \frac{2.0 \times 10^{-3}\mathrm{A}}{25} = 8.0 \times 10^{-5}\mathrm{A}\ \text{per division}$$

The potential difference across the coil at full-scale deflection follows from Ohm's law:

$$\mathrm{V} = \mathrm{IR} = (2.0 \times 10^{-3}\mathrm{A})(50\Omega) = 0.1\mathrm{V}$$

The galvanometer has a figure of merit of $8.0 \times 10^{-5}\mathrm{A}$ per division, and the potential difference across it at full-scale deflection is $0.1\mathrm{V}$.

Example 58

(a) A gas is normally an insulator, yet under the right conditions it will carry a current. Explain the conditions under which a gas conducts, and explain what is meant by ionisation.
(b) A spark first jumps across a $2.0\mathrm{cm}$ air gap when the potential difference between the electrodes reaches $60\mathrm{kV}$. Determine the breakdown field of the air, and hence find the voltage at which a spark would jump across a $3.0\mathrm{cm}$ gap between the same electrodes.

Solution

(a) A gas conducts only when it contains free charge carriers, and a neutral gas has none, so carriers must first be created by ionisation. Ionisation is the removal of one or more electrons from a neutral gas molecule, which leaves a positive ion and a free electron, both of which can carry current. It requires energy, supplied either by a strong electric field or by an ionising agent such as ultraviolet light, X-rays, or radiation, and the gas conducts only once the applied voltage reaches the breakdown value at which ionisation begins.

(b) The breakdown field is the breakdown voltage divided by the gap across which it acts:

$$\mathrm{E} = \frac{\mathrm{V}}{\mathrm{d}} = \frac{60 \times 10^{3}\mathrm{V}}{2.0 \times 10^{-2}\mathrm{m}} = 3.0 \times 10^{6}\mathrm{Vm^{-1}}$$

For the same air the breakdown field is unchanged, so the voltage needed across the wider gap follows from $V = Ed$:

$$V = Ed = (3.0 \times 10^{6}\mathrm{Vm^{-1}})(3.0 \times 10^{-2}\mathrm{m}) = 90\mathrm{kV}$$

The breakdown field of the air is $3.0 \times 10^{6}\mathrm{Vm^{-1}}$, and the wider 3.0cm gap requires the larger voltage of 90kV to spark.

Example 59

(a) Distinguish between thermionic emission and photoelectric emission.
(b) In a cathode-ray tube an electron starts from rest and is accelerated through a potential difference of 2.0kV. Determine the kinetic energy it gains, in joules and in electron-volts, and its final speed. Take the electronic charge as $1.6 \times 10^{-19}\mathrm{C}$ and the electron mass as $9.1 \times 10^{-31}\mathrm{kg}$.

Solution

(a) Both processes release electrons from the surface of a metal, but they differ in the energy that frees them. In thermionic emission the metal is heated, and the electrons gain enough thermal energy to escape from its surface. In photoelectric emission the metal is illuminated, and an electron escapes by absorbing a photon of sufficiently high frequency. The one is therefore driven by heat and the other by light.

(b) The kinetic energy gained equals the work done on the electron by the accelerating voltage:

$$E_k = qV = (1.6 \times 10^{-19}\mathrm{C})(2000\mathrm{V}) = 3.2 \times 10^{-16}\mathrm{J}$$

Expressed in electron-volts, the energy gained in falling through 2000V is simply:

$$E_k = 2000\mathrm{eV} = 2.0\mathrm{keV}$$

The final speed follows from $E_k = \frac{1}{2}mv^2$, so that:

$$v = \sqrt{\frac{2E_k}{m}} = \sqrt{\frac{2(3.2 \times 10^{-16}\mathrm{J})}{9.1 \times 10^{-31}\mathrm{kg}}} = 2.7 \times 10^{7}\mathrm{ms^{-1}}$$

The electron gains $3.2 \times 10^{-16}\mathrm{J}$, equal to 2.0keV, and reaches a final speed of $2.7 \times 10^{7}\mathrm{ms^{-1}}$.

Example 60

(a) Faraday summarised electrolysis in two laws. Explain what each law tells us, and explain what is meant by the electrochemical equivalent of a substance.
(b) A steady current of 2.0A is passed through copper sulphate solution for 30min. Determine the mass of copper deposited on the cathode. Take the electrochemical equivalent of copper as $3.3 \times 10^{-7}\mathrm{kgC^{-1}}$.

Solution

(a) Faraday's first law tells us that the mass of substance liberated at an electrode is directly proportional to the quantity of charge passed through the electrolyte. His second law tells us that, for the same quantity of charge, the mass liberated is proportional to the chemical equivalent of the substance, so that different substances are deposited in proportion to their equivalent weights. The electrochemical equivalent of a substance is the mass of it liberated by one coulomb of charge.

(b) The charge passed is the product of the current and the time, with the time first converted to seconds:

$$Q = It = (2.0\mathrm{A})(30 \times 60\mathrm{s}) = 3600\mathrm{C}$$

By Faraday's first law the mass deposited is the electrochemical equivalent multiplied by the charge:

$$m = zQ = (3.3 \times 10^{-7}\mathrm{kgC^{-1}})(3600\mathrm{C}) = 1.2 \times 10^{-3}\mathrm{kg}$$

The mass of copper deposited on the cathode is $1.2 \times 10^{-3}\mathrm{kg}$, that is about 1.2g.

Example 61

(a) Distinguish between line, band, and continuous spectra, and give one source of each.

(b) A diffraction grating is ruled with 5000 lines per centimetre. Determine the spacing of its lines, and hence the highest order in which light of wavelength 600nm can be observed with it.

Solution

(a) A line spectrum consists of separate bright lines on a dark background and is produced by an excited monatomic gas at low pressure, such as a sodium lamp. A band spectrum consists of groups of closely spaced lines, each group crowding together into a band, and is produced by an excited molecular gas, such as nitrogen. A continuous spectrum is an unbroken range of colour with no gaps and is produced by a hot dense source, such as the glowing filament of a lamp.

(b) The spacing of the lines is one centimetre divided by the number of lines in it:

$$\mathrm{d} = \frac{1 \times 10^{-2}\mathrm{m}}{5000} = 2.0 \times 10^{-6}\mathrm{m}$$

The highest order is fixed by the requirement that $\sin\theta$ cannot exceed 1 in the grating equation $\mathrm{m}\lambda = \mathrm{d}\sin\theta$, so that $\mathrm{m} \le \frac{\mathrm{d}}{\lambda}$:

$$\frac{\mathrm{d}}{\lambda} = \frac{2.0 \times 10^{-6}\mathrm{m}}{600 \times 10^{-9}\mathrm{m}} = 3.3$$

Since the order must be a whole number not exceeding 3.3, the highest observable order is the third. The grating has a line spacing of 2.0×10^{-6}m, and 600nm light can be seen up to the third order.

Example 62

(a) Using the energy changes that take place in each, explain why an incandescent filament lamp is far less efficient than a fluorescent or LED lamp.

(b) A shop replaces ten 60W incandescent bulbs with ten 8W LED bulbs, each lit for 10h each day. Determine the energy saved in a 30-day month, and the money saved if TANESCO charges 290TZS for each kilowatt-hour.

Solution

(a) In an incandescent lamp the current heats a filament until it glows, so the energy is converted first into heat and only the small part of it that falls in the visible range leaves as light; the great majority is wasted as infrared heat. In a fluorescent or LED lamp the electrical energy is converted into light far more directly, with very little spent on heating, so a much larger fraction of the input energy emerges as visible light. The incandescent lamp is therefore far less efficient, because it is essentially a heater that happens to glow.

(b) The energy used by the ten incandescent bulbs in the month is the total power multiplied by the running time:

$$\mathrm{E_{inc}} = (10 \times 60\mathrm{W})(10\mathrm{h})(30) = 180\mathrm{kWh}$$

The energy used by the ten LED bulbs over the same period is:

$$\mathrm{E_{LED}} = (10 \times 8\mathrm{W})(10\mathrm{h})(30) = 24\mathrm{kWh}$$

The energy saved is the difference between the two:

$$\mathrm{E_{saved}} = 180\mathrm{kWh} - 24\mathrm{kWh} = 156\mathrm{kWh}$$

The money saved is this energy multiplied by the tariff:

$$\mathrm{Saving} = 156\mathrm{kWh} \times 290\mathrm{TZS/kWh} = 45240\mathrm{TZS}$$

The change saves 156kWh and 45240TZS each month.

Example 63

(a) Explain how a fuse protects an electrical circuit, and explain why it is connected in the live wire rather than the neutral.

(b) A 13A fuse protects a 240V socket. An electric iron rated 1.0kW, a kettle rated 2.0kW, and a heater rated 1.5kW are all run from that socket at the same time. Determine the total current drawn and state whether the fuse blows.

Solution

(a) A fuse is a short piece of thin wire of low melting point connected in series with the circuit. If the current rises above the safe value, the heating produced in the fuse melts it, the circuit breaks, and the rest of the wiring is protected from overheating and fire. It is placed in the live wire so that, when it blows, the appliance is cut off from the dangerous live potential; were it placed in the neutral, the appliance would remain connected to the live supply even after the fuse had melted, leaving it a shock hazard.

(b) The three appliances together draw a total power that is the sum of their ratings:

$$P = 1.0\text{kW} + 2.0\text{kW} + 1.5\text{kW} = 4.5\text{kW}$$

The total current drawn from the 240V supply follows from $P = VI$:

$$I = \frac{P}{V} = \frac{4500\text{W}}{240\text{V}} = 18.75\text{A}$$

Since 18.75A exceeds the 13A rating of the fuse, the fuse blows and the circuit is broken.

Example 64

(a) Explain why the appliances in a house are connected in parallel across the mains rather than in series.
(b) Two identical bulbs, each of resistance 6Ω, are connected across a 12V supply of negligible internal resistance, first in series and then in parallel. Determine the power dissipated in one bulb in each arrangement, and state which arrangement makes the bulbs brighter.

Solution

(a) Appliances are connected in parallel so that each receives the full mains voltage and can be switched on and off independently of the others. Were they in series, they would share the supply voltage between them, so each would receive less than its rated voltage and run dimly or weakly, and a break anywhere in the chain, or the switching off of any one appliance, would interrupt the current to them all. Parallel connection avoids both faults.

(b) In series the two bulbs add to a total resistance of 12Ω, so the current is:

$$I = \frac{V}{R_{\text{series}}} = \frac{12\text{V}}{12\Omega} = 1.0\text{A}$$

and the power in one bulb is:

$$P_{\text{series}} = I^2R = (1.0\text{A})^2(6\Omega) = 6.0\text{W}$$

In parallel each bulb has the full 12V across it, so the power in one bulb is:

$$P_{\text{parallel}} = \frac{V^2}{R} = \frac{(12\text{V})^2}{6\Omega} = 24\text{W}$$

Each bulb dissipates 6.0W in series but 24W in parallel, so the bulbs are brighter in the parallel arrangement.

Example 65

(a) Explain the difference between an initiated and a self-sustaining gas discharge, and explain the part that chain ionisation plays in the self-sustaining case.
(b) The gas in a discharge tube carries a current of 8.0mA when a potential difference of 800V is maintained across it. Determine the power dissipated in the tube and the energy it delivers to the gas in 5.0min.

Solution

(a) An initiated discharge conducts only while an external ionising agent keeps supplying ions, so that the current dies the moment the agent is removed; the applied voltage merely sweeps the ready-made ions to the electrodes. A self-sustaining discharge continues on its own without any external agent, because the voltage is now large enough for the discharge to create its own ions. The mechanism that makes this possible is chain ionisation: each accelerated electron gains enough energy between collisions to ionise the molecule it strikes, the freed electrons are accelerated in turn and ionise further molecules, and this avalanche keeps the supply of carriers going without outside help.

(b) The power dissipated in the tube is the product of the voltage and the current:

$$\mathrm{P} = \mathrm{VI} = (800\mathrm{V})(8.0 \times 10^{-3}\mathrm{A}) = 6.4\mathrm{W}$$

The energy delivered is the power multiplied by the time, with the time converted to seconds:

$$\mathrm{E} = \mathrm{Pt} = (6.4\mathrm{W})(5.0 \times 60\mathrm{s}) = 1920\mathrm{J}$$

The tube dissipates 6.4W and delivers 1920J, that is 1.92kJ, to the gas in five minutes.

Sixteen miscellaneous worked examples now lie behind you, and between them they have called on every corner of the chapter, from the drift of an electron to the breakdown of air, and demanded that you tell each one apart by sight. If they have left you wanting more, you are in luck, and if they have left you wanting mercy, you are out of it: the Digging Deeper questions that follow give the scenario and the instruction and nothing else, and this time there is no worked solution waiting at the foot of the page. Pen ready.

DIGGING DEEPER EXERCISE 12

EXERCISE 12A: BINDER QUESTIONS

Question 1

Explain why the drift velocity of electrons in a copper wire is of the order of a millimetre per second, while the signal that switches on a lamp travels almost at the speed of light. Use the queue-and-buckets argument.

Question 2

Distinguish between the EMF and the terminal voltage of a cell, and give the condition under which the two are equal.

Question 3

Ohm’s law is often quoted in a single sentence. Express it in its simplest form, identify two materials that obey it and two that do not, and explain why it is described as empirical rather than as a fundamental law of nature.

Question 4

Explain what is meant by the conductance of a conductor, give its SI unit, and suggest why conductance is seldom used in school-level circuit problems.

Question 5

Explain why the resistance of a metal increases with temperature, referring both to the relaxation time and to the microscopic expression.

Question 6

Distinguish between resistance and resistivity, and give the SI unit of each.

Question 7

Explain why nichrome or constantan, rather than copper, is used for heating elements and standard resistors.

Question 8

Kirchhoff gave two laws for electrical networks. Explain what each one says, and name the conservation principle that each embodies.

Question 9

Explain why a Wheatstone bridge gives a more accurate value of a resistance than a direct ohmmeter reading.

Question 10

Explain why the potential gradient along a potentiometer wire must be uniform, and give the conditions on the wire and on the driving current that guarantee it.

Question 11

Explain how a fuse made of a low-melting-point alloy protects a household circuit.

Question 12

Explain what the maximum power transfer theorem says, and why power transmission systems such as TANESCO’s national grid deliberately avoid the maximum-power condition.

Question 13

Distinguish between an initiated and a self-sustaining gas discharge, and explain the part that chain ionisation plays in the self-sustaining regime.

Question 14

Explain why a sodium street lamp glows yellow-orange while a neon advertising sign glows red. Refer to the discrete energy levels of the gas atoms.

EXERCISE 12B: REAL QUESTIONS

Question 15

Explain why the metal casing of an electric appliance, such as a refrigerator, is connected to the earth pin of its three-pin plug.

Question 16

TANESCO transmits power between its substations at 132kV but steps it down to 240V before it reaches a house. Explain why such a high voltage is used for long-distance transmission rather than the household voltage.

Question 17

Explain why a torch bulb almost always blows at the instant it is switched on, rather than after it has been glowing steadily for some time.

Question 18

A torch fitted with dry cells is noticeably brighter when the cells are new than after a week of use. Explain this, referring to the EMF, the internal resistance, and the terminal voltage of the cells.

Question 19

Explain why a lamp designed to run on the 240V mains burns out almost at once if it is mistakenly connected to a supply of much higher voltage.

Question 20

Explain why it is dangerous to run a high-power electric heater from a long, thin extension lead that has been left coiled up on its reel.

Question 21

A student measures the resistance of a copper coil with a multimeter and notices that the reading slowly increases the longer the meter is left connected. Explain this observation.

Question 22

A Geiger-Müller counter clicks more rapidly when it is brought closer to a radioactive source. Explain this, referring to the operating region of the ionisation curve.

Question 23

The arc of an electric welder appears bluish-white at its centre but more orange towards its edges. Explain this, referring to the temperature across the arc and the light each part gives off.

Question 24

Explain why the cathode filaments at the two ends of a fluorescent tube are briefly heated before the main discharge begins.

Question 25

A person who touches a faulty switch receives a more severe shock when their hands are wet than when they are dry. Explain why.

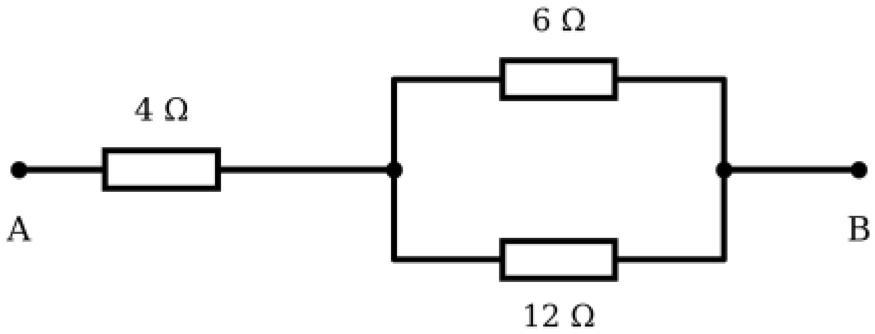
EXERCISE 12C: HOT QUESTIONS

Question 26

A uniform wire of resistance 12Ω is cut into three equal pieces, and the three pieces are then joined side by side in parallel. Determine the resistance of the combination.

Question 27

For the network of resistors in the figure, determine the equivalent resistance between the terminals A and B.



Question 28

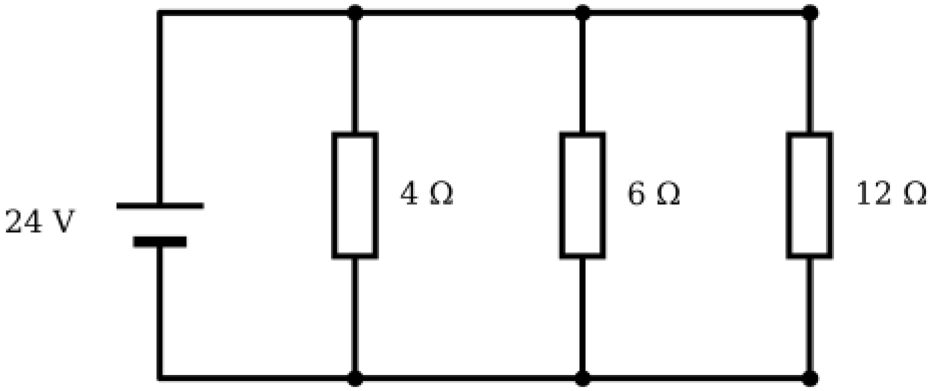
A coil of copper wire has a temperature coefficient of resistance of $4.0 \times 10^{-3}{}^{\circ}\mathrm{C}^{-1}$, referred to its resistance at $0^{\circ}\mathrm{C}$. Determine the temperature at which its resistance is 20% greater than its value at $0^{\circ}\mathrm{C}$.

Question 29

When a resistor of 4Ω is connected across a cell, the current is 1.0A; when the 4Ω resistor is replaced by one of 9Ω, the current falls to 0.5A. Determine the EMF and the internal resistance of the cell.

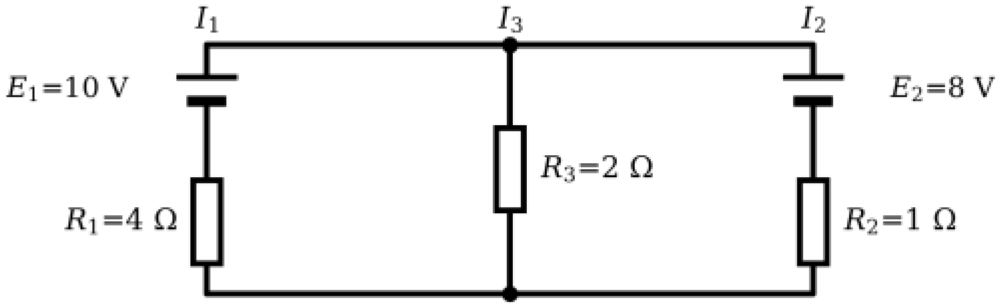
Question 30

Three resistors of 4Ω, 6Ω and 12Ω are connected in parallel across a 24V supply of negligible internal resistance, as shown. Determine the equivalent resistance of the combination and the current in each resistor.



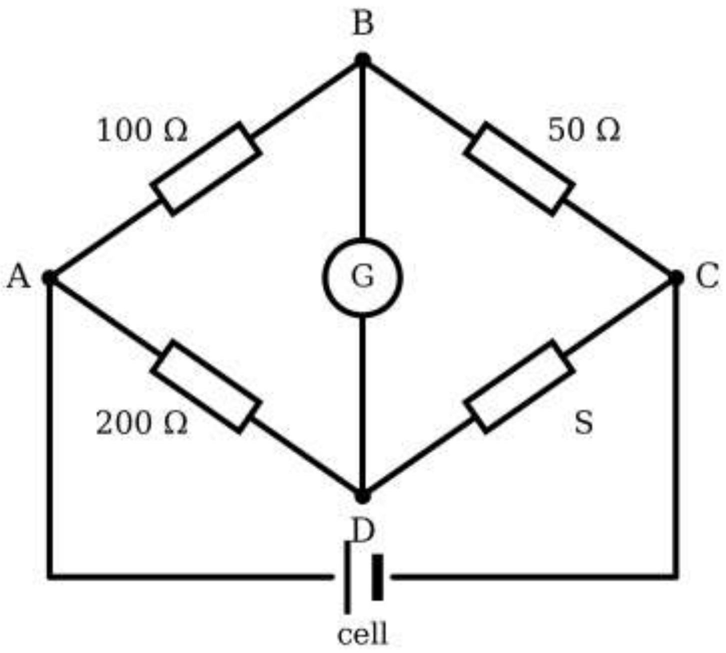
Question 31

In the network of the figure the cells have negligible internal resistance. Using Kirchhoff's laws, determine the current supplied by each cell, the current in the central resistor R_3, and the power dissipated in R_3.



Question 32

The Wheatstone bridge in the figure is balanced, so that no current flows through the galvanometer. Determine the value of the unknown resistance S.

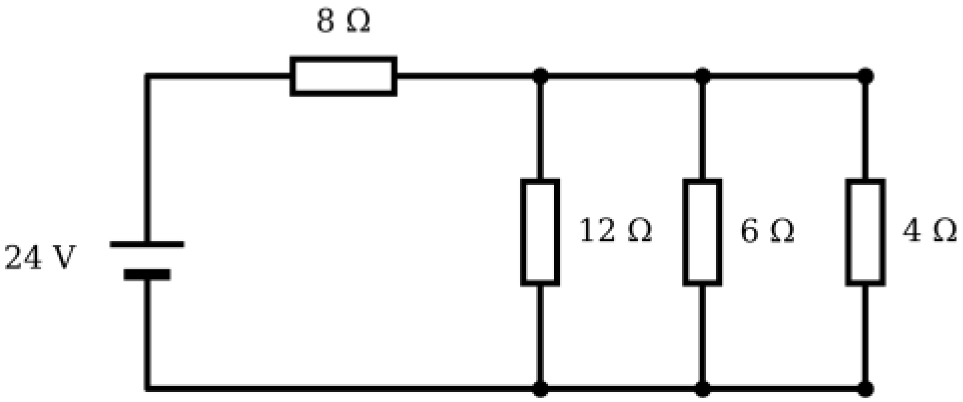


Question 33

Four identical cells, each of EMF 1.5V and internal resistance 0.5Ω, are connected to an external resistor of 2.0Ω, first all in series and then all in parallel. Determine the current through the external resistor in each arrangement.

Question 34

The 24V battery in the figure has negligible internal resistance. Determine the total current drawn from the battery and the current in each of the parallel resistors.



Question 35

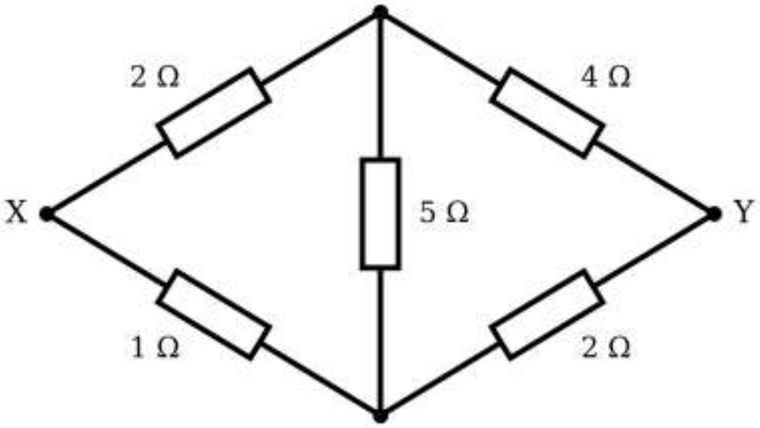
An aluminium wire of cross-sectional area $1.0 \times 10^{-6}\text{m}^2$ carries a steady current of 5.0A. Taking the number of free electrons per unit volume in aluminium as $1.8 \times 10^{29}\text{m}^{-3}$ and the electronic charge as $1.6 \times 10^{-19}\text{C}$, determine the drift velocity of the electrons.

Question 36

Three resistors of 2Ω, 3Ω and 5Ω are joined in series across a supply of 20V with negligible internal resistance. Determine the current in the circuit and the power dissipated in each resistor.

Question 37

For the bridge network in the figure, determine the equivalent resistance between the terminals X and Y.

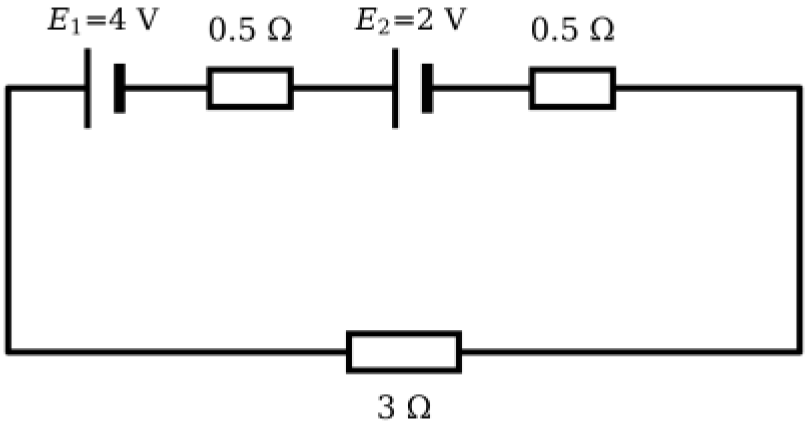


Question 38

Two resistors give a combined resistance of 10Ω when joined in series and 2.4Ω when joined in parallel. Determine the resistance of each.

Question 39

The two cells in the figure each have an internal resistance of 0.5Ω and are joined in series with a 3Ω resistor. Determine the current in the circuit when the cells are connected so that they assist each other, and again when one is reversed so that they oppose each other.

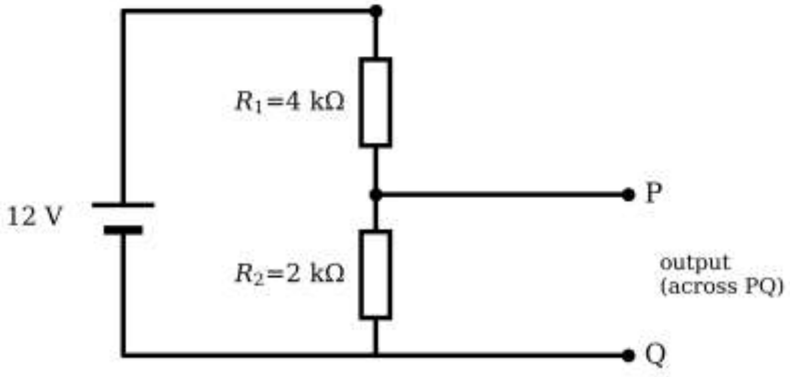


Question 40

A 6Ω resistor and a 3Ω resistor are connected in parallel, and the combination dissipates a total power of 18W. Determine the supply voltage across the combination and the current in each resistor.

Question 41

In the potential divider of the figure a 12V supply is connected across a 4kΩ resistor in series with a 2kΩ resistor, and the output is taken across the 2kΩ resistor at the terminals P and Q. Determine the output voltage when nothing is connected across PQ, and again when a load of 2kΩ is connected across PQ.

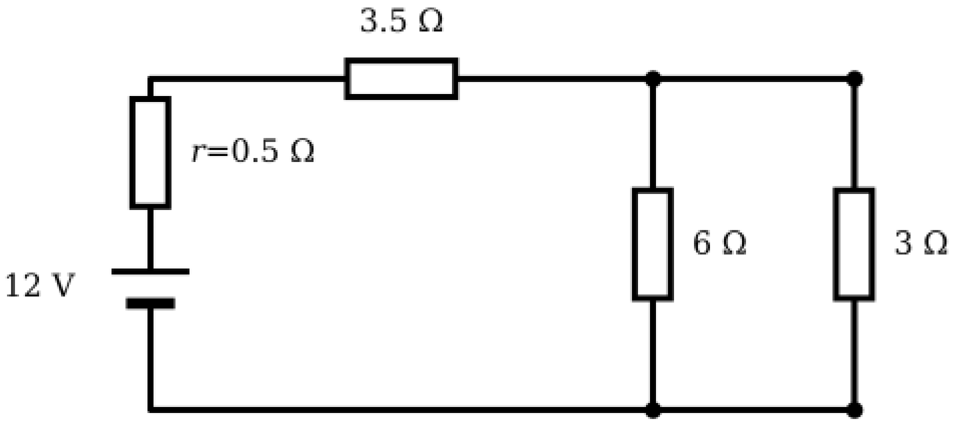


Question 42

A single stroke of lightning transfers a charge of 15C from the cloud to the ground in a time of 0.20ms, through a potential difference of $3.0 \times 10^8\text{V}$. Determine the average current in the stroke, the energy released, and the average power delivered.

Question 43

The cell in the figure has an EMF of 12V and an internal resistance of 0.5Ω, and it drives current through a 3.5Ω resistor in series with a parallel combination of a 6Ω and a 3Ω resistor. Determine the current drawn from the cell, the terminal voltage of the cell, the power delivered to the external circuit, and the power wasted inside the cell.



ANSWERS

EXERCISE 12A

1. The electrons are already present all along the wire before the switch is closed, like a queue of people filling a corridor. When the switch closes, the electric field is set up along the whole wire almost at the speed of light, so every electron begins to drift at once and the lamp lights immediately. The electrons themselves only crawl forward at about a millimetre per second; it is the field, not the electrons, that travels fast, just as the bucket is passed down a standing queue at once although each person barely moves.

2. The EMF is the energy supplied per coulomb by the source, equal to the work done per unit charge when no current flows. The terminal voltage is the potential difference actually available at the terminals when current flows, and is smaller than the EMF by the internal drop Ir. The two are equal only when no current is drawn, that is on open circuit, because then $\text{Ir} = 0$.

3. Ohm's law states that the current through a conductor is directly proportional to the potential difference across it, provided the temperature and other physical conditions remain constant, so that $\text{V} = \text{IR}$ with R constant. A metal wire at constant temperature and constantan obey it; a filament lamp and a semiconductor diode do not. It is called empirical because it is an experimental result that holds only for certain materials under certain conditions, rather than a universal law derived from first principles.

4. The conductance of a conductor measures how easily it allows current to pass, and is the reciprocal of its resistance, $\text{G} = \frac{1}{\text{R}}$. Its SI unit is the siemens, equal to one ampere per volt. It is seldom used in school problems because series resistances simply add and the familiar form $\text{V} = \text{IR}$ is more direct, so conductance offers no real advantage there.

5. In a metal the number of free electrons is essentially fixed, so the resistance depends only on how often the electrons are scattered. Raising the temperature makes the lattice ions vibrate more violently, so the electrons collide with them more often and the relaxation time τ falls. Since $\text{R} = \frac{\text{ml}}{\text{ne}^2\text{A}\tau}$, a smaller τ gives a larger R, so the resistance of a metal rises with temperature.

6. Resistance is the opposition that a particular specimen offers to current; it depends on the material and on the dimensions of the specimen, and its unit is the ohm. Resistivity is a property of the material alone, independent of shape or size, numerically equal to the resistance of a unit cube of the material, and its unit is the ohm-metre.

7. Nichrome and constantan have a resistivity about sixty times that of copper, so a short length gives a usefully large resistance, whereas copper would need an impossibly long wire. They also have a very small temperature coefficient of resistance, so their resistance barely changes as they heat, keeping a heater's power or a standard resistor's value steady. Copper fails on both counts, having a low resistivity and a large temperature coefficient.

8. Kirchhoff's current law states that the total current entering any junction equals the total current leaving it, and it embodies the conservation of electric charge. Kirchhoff's voltage law states that around any closed loop the sum of the EMFs equals the sum of the potential drops, and it embodies the conservation of energy.

9. An ordinary ohmmeter passes a current through the unknown resistor and reads it on a calibrated scale, so its accuracy is limited by that calibration and by the current drawn. A Wheatstone bridge is instead adjusted until the galvanometer reads exactly zero, and at balance no current flows through the galvanometer, so the unknown resistance is fixed by a ratio of the other three resistances and does not depend on the galvanometer's calibration at all, only on its ability to detect zero. This null method is what makes the bridge more accurate.

10. A potentiometer assumes that the potential falls evenly along its wire, so that the balance length is directly proportional to the potential difference being measured. If the gradient were not uniform, equal lengths would not correspond to equal voltages and the readings would be meaningless. A uniform gradient is guaranteed when the wire is of uniform cross-section and material, so its resistance per unit length is constant, and the driving current is kept steady.

11. A fuse is a short length of low-melting-point alloy connected in series with the circuit. The current through it heats it at the rate $P = I^2R$. At normal currents this heating is mild and the fuse survives, but if a fault makes the current rise too high, the increased I^2R heating melts the fuse, so the circuit breaks and the current stops before the wiring or appliance can overheat and catch fire.

12. The maximum power transfer theorem states that *a source of fixed EMF and internal resistance r delivers the greatest power to a load when the load resistance equals the internal resistance, $R = r$.* At that matched point the efficiency is only 50%, because half the power is wasted inside the source. A transmission grid must deliver almost all its power to consumers, so it deliberately makes the load resistance far larger than the line resistance and runs at high efficiency, well away from the matched condition; matching would waste half the generated power as heat in the lines.

13. In an initiated discharge the gas conducts only while an external ionising agent keeps producing ions, so the current stops the moment the agent is removed, since the applied voltage merely collects ready-made ions. In a self-sustaining discharge the voltage is large enough that the discharge makes its own ions and continues without any external agent. The mechanism is chain ionisation: each accelerated electron gains enough energy between collisions to ionise the next molecule, those freed electrons ionise still more, and the resulting avalanche keeps the gas supplied with carriers.

14. The colour a gas emits is set by the discrete energy levels of its atoms, because when an excited electron falls between two levels it emits a photon whose energy, and hence colour, equals the gap between them. Sodium atoms have an energy gap corresponding to yellow-orange light, while neon atoms have gaps corresponding to red light. Since each element has its own unique set of energy levels, each glows with its own characteristic colour.

EXERCISE 12B

15. The earth pin joins the metal casing to the ground through a path of very low resistance, and in normal use the casing carries no current. Should a fault let the live wire touch the casing, the casing would otherwise rise to the live potential and shock anyone who touched it. Because the earth connection offers an easy path to ground, a large fault current instead flows through it and blows the fuse, cutting off the supply. The casing is therefore never left at a dangerous potential.

16. The power delivered is $P = VI$, so for a fixed power a higher transmission voltage means a proportionally smaller current. The power wasted as heat in the line is I^2R_{line}, which grows with the square of the current. Raising the voltage therefore lowers the current and cuts the line loss sharply, so that far more of the generated power reaches the consumer. Transmitting at the household voltage would demand an enormous current and waste most of the power as heat in the lines.

17. The resistance of a metal filament rises with temperature, so a cold filament has its lowest resistance. At the instant of switching on the filament is still cold, so it offers little resistance and a large surge of current passes through it. This sudden surge produces a burst of heating that can snap the already thin and weakened filament. Once the filament warms, its resistance rises and limits the current, so failure is far more likely at the cold moment of switch-on than during steady glowing.

18. The brightness depends on the terminal voltage supplied to the bulb, given by $V = E - Ir$. As the cells are used, chemical changes within them increase the internal resistance r, so the internal drop Ir grows and the terminal voltage falls. Less power therefore reaches the bulb and it glows more dimly. New cells have a small internal resistance and deliver almost the full EMF to the bulb, so they make it brighter.

19. A lamp's filament has a roughly fixed resistance, and the power it dissipates is $P = \frac{V^2}{R}$, which rises with the square of the applied voltage. Connecting it to a much higher voltage therefore drives its power up steeply, so the filament is heated far beyond its design temperature and melts almost at once. The lamp is rated for one particular voltage precisely because a higher one destroys it.

20. The extension lead has its own resistance, and the heavy current drawn by the heater heats the lead itself at the rate I^2R. When the lead is coiled on its reel, the turns trap this heat and prevent it from escaping, so the temperature of the insulation climbs steadily. The insulation may then soften, melt, and short, starting a fire. The same lead run out straight sheds its heat easily and stays cool, which is why a heavy load must never be drawn through a coiled lead.

21. To measure resistance the meter passes a small current through the coil, and this current gently heats the copper. Since the resistance of copper rises with temperature, the coil's resistance increases as it warms. The longer the meter is left connected, the warmer the coil grows, so the displayed reading slowly climbs.

22. The counter operates in the self-sustaining region of the ionisation curve, where each ionising particle that enters the tube triggers a full avalanche and so produces one current pulse, heard as a click. Bringing the counter closer to the source means more particles reach the tube each second, so more avalanches are triggered per second and the clicks come more rapidly. The count rate therefore measures how many particles arrive, not how energetic each one is.

23. The light given off by a hot body shifts towards shorter, bluer wavelengths as its temperature rises. The centre of the arc is the hottest region, so it emits the high-energy bluish-white light, while towards the edges the arc is cooler and emits longer-wavelength orange light. The change of colour across the arc therefore maps the change of temperature across it.

24. A cold tube contains few free carriers and has a high breakdown potential, so the mains voltage alone cannot strike the discharge. Heating the cathode filaments makes them release electrons by thermionic emission, seeding the gas with free carriers. With these seed electrons present, the high-voltage pulse from the choke can then ionise a path along the tube and start the discharge. The preheating therefore supplies the carriers the tube needs before it can light.

25. The current driven through a person is set by the voltage and the resistance of the body, $I = \frac{V}{R}$. Dry skin has a high resistance, which limits the current to a small and usually harmless value. Wet skin, however, has a much lower resistance, so for the same voltage a far larger current flows through the body, and the shock is correspondingly more severe.

EXERCISE 12C

26. The resistance of the parallel combination is 1.33Ω.

27. The equivalent resistance between A and B is 8Ω.

28. The resistance is 20% greater at a temperature of $50^\circ C$.

29. The cell has an EMF of $5.0V$ and an internal resistance of 1.0Ω.

30. The equivalent resistance is 2Ω, and the currents are $6A$, $4A$ and $2A$.

31. The cells supply $1.0A$ and $2.0A$, the central resistor carries $3.0A$, and it dissipates $18W$.

32. The unknown resistance is $S = 100\Omega$.

33. The series arrangement gives $1.5A$ and the parallel arrangement $0.71A$.

34. The battery delivers $2.4A$, which divides into $0.4A$, $0.8A$ and $1.2A$ through the 12Ω, 6Ω and 4Ω resistors respectively.

35. The drift velocity of the electrons is $1.7 \times 10^{-4} ms^{-1}$.

36. The current is $2.0A$, and the powers are $8W$, $12W$ and $20W$.

37. The equivalent resistance between X and Y is 2Ω.

38. The two resistances are 4Ω and 6Ω.

39. When the cells assist each other, their EMFs add and so do their internal resistances, giving a total EMF of $4V + 2V = 6V$ against a total resistance of $0.5\Omega + 0.5\Omega + 3\Omega = 4\Omega$:

$$I_{aiding} = \frac{6V}{4\Omega} = 1.5A$$

When one cell is reversed, the EMFs subtract while the resistances still add, giving a net EMF of $4\text{V} - 2\text{V} = 2\text{V}$ against the same 4Ω:

$$\text{I}_{\text{opposing}} = \frac{2\text{V}}{4\Omega} = 0.5\text{A}$$

The current is 1.5A when the cells assist each other and 0.5A when they oppose.

40. The two resistors in parallel have a combined resistance of:

$$\text{R}_\text{p} = \frac{(6\Omega)(3\Omega)}{6\Omega + 3\Omega} = 2\Omega$$

The total power dissipated is $\text{P} = \frac{\text{V}^2}{\text{R}_\text{p}}$, so the supply voltage is:

$$\text{V} = \sqrt{\text{PR}_\text{p}} = \sqrt{(18\text{W})(2\Omega)} = 6.0\text{V}$$

This voltage acts across each resistor, so the branch currents are:

$$\text{I}_6 = \frac{6.0\text{V}}{6\Omega} = 1.0\text{A}, \qquad \text{I}_3 = \frac{6.0\text{V}}{3\Omega} = 2.0\text{A}$$

The supply voltage is 6.0V, and the currents are 1.0A in the 6Ω resistor and 2.0A in the 3Ω resistor.

41. With nothing connected across PQ, no current is drawn from the tap, so the same current flows through both divider resistors, and the output is the fraction of the supply that falls across the lower resistor:

$$\text{V}_{\text{out}} = \text{V} \times \frac{\text{R}_2}{\text{R}_1 + \text{R}_2} = 12\text{V} \times \frac{2\text{k}\Omega}{4\text{k}\Omega + 2\text{k}\Omega} = 4.0\text{V}$$

When a 2kΩ load is connected across PQ, it is in parallel with the lower 2kΩ resistor, and the two combine to:

$$\text{R}_{\text{lower}} = \frac{(2\text{k}\Omega)(2\text{k}\Omega)}{2\text{k}\Omega + 2\text{k}\Omega} = 1\text{k}\Omega$$

The output now falls across this smaller resistance:

$$\text{V}_{\text{out}} = 12\text{V} \times \frac{1\text{k}\Omega}{4\text{k}\Omega + 1\text{k}\Omega} = 2.4\text{V}$$

The output is 4.0V on open circuit but drops to 2.4V under load, which shows that drawing current from a potential divider lowers its output voltage.

42. The average current is the charge transferred divided by the time it takes:

$$\text{I} = \frac{\text{Q}}{\text{t}} = \frac{15\text{C}}{0.20 \times 10^{-3}\text{s}} = 7.5 \times 10^4\text{A}$$

The energy released is the charge multiplied by the potential difference through which it falls:

$$\text{E} = \text{QV} = (15\text{C})(3.0 \times 10^8\text{V}) = 4.5 \times 10^9\text{J}$$

The average power is the energy delivered divided by the time:

$$\text{P} = \frac{\text{E}}{\text{t}} = \frac{4.5 \times 10^9\text{J}}{0.20 \times 10^{-3}\text{s}} = 2.25 \times 10^{13}\text{W}$$

The stroke carries an average current of 7.5×10^4A, releases 4.5×10^9J of energy, and delivers it at an average power of 2.25×10^{13}W, which is why lightning is so destructive in the brief instant it lasts.

43. The cell drives 2.0A, holds a terminal voltage of 11V, delivers 22W to the external circuit, and wastes 2W within itself.