

PART THREE

ELECTRICITY

Having traced the behaviour of heat in Part One and the motions and waves that carry energy through matter and space in Part Two, we now turn to one of the most far-reaching forces in nature: the electric force. Electricity examines charges at rest and in motion, beginning with electrostatic fields and potential, moving through the storage of charge in capacitors, the steady flow of current in conductors and circuits, and finally the alternating currents that power modern life. These ideas explain how charge gives rise to forces and fields, how energy is stored and delivered through electrical networks, and how the same principles govern systems ranging from a spark across a doorknob to national power transmission, radio tuning, and electronic devices. Mastering this part will deepen your understanding of charge, field, and circuit behaviour, and prepare you for more advanced studies in electronics, electromagnetism, and applied technology.

Chapter 10

ELECTROSTATICS

INTRODUCTION

Mid-afternoon at Miono Secondary School. The corridor outside the physics laboratory has been polished so aggressively that students walk through it like people crossing a frozen lake. For two weeks the brass door handle has been famous. The Form Sixes call it *the snake* because every afternoon somebody touches it and jumps backwards with wounded dignity.

Today it is Kipanga's turn.

He drags his black rubber shoes slowly across the polished floor while several Form Sixes watch from a safe distance with the excitement normally reserved for football penalties and public disasters. Kipanga approaches the brass handle carefully and stretches out one finger.

Nothing happens.

The Form Sixes groan in disappointment.

Kipanga frowns, walks back to the far wall, and rubs his shoes across the corridor again, harder this time. Even the bougainvillea beside the laboratory seems interested. He returns to the handle and lowers his finger toward it with the caution of a man touching a sleeping cobra.

SNAP!

A tiny blue spark jumps from the brass before he even touches it.

Kipanga yelps and nearly throws himself into the notice board while the Form Sixes explode with laughter.

"The snake has bitten another one!"

Mwajuma, passing with a flask from the chemistry prep room, jumps so badly she almost drops it. Kipute, wearing open sandals and carrying absolutely no fear, walks straight to the same handle, grabs it, opens the door, and walks inside unharmed.

The Form Sixes look cheated.

From the louvre window, Mr. Akilikubwa finally steps into the corridor with the calm expression of a man who has been waiting all week for this exact foolishness.

Mr. Akilikubwa: *"Kipanga, congratulations. You have just produced several thousand volts using only your shoes and poor decision-making."*

Kipanga (staring at his finger suspiciously): *"Sir, there is no socket anywhere near me."*

Mr. Akilikubwa: *"You do not need a socket. Every step moved electrons from the floor onto your body. Your rubber shoes trapped the charge instead of letting it escape. By the time you reached the brass handle, your finger and the metal had a huge potential difference between them. The air finally gave up and the charge jumped across."*

Kipute (lifting one sandal): *"Then why did the snake ignore me?"*

Mr. Akilikubwa: *"Because your sandals leak charge away continuously. Kipanga stored charge like a battery. You leaked it away like a cracked bucket."*

Mwajuma (edging closer): *"Sir... is that the same physics as lightning?"*

Mr. Akilikubwa (pointing upward dramatically): *"Exactly the same physics. A cloud stores charge the way Kipanga did. When the electric field becomes strong enough, the air breaks down and the charge jumps through the sky. Lightning is simply the giant version of Kipanga embarrassing himself in public."*

The Form Sixes cheer loudly at this scientific interpretation.

Kipanga rubs his finger carefully.

Kipanga: *"Sir, so the snake is just electrostatics?"*

Mr. Akilikubwa (smiling): *"Correct. By the end of this chapter, you will calculate the voltage in that spark, the electric field in the air gap, and the energy released into Kipanga's unfortunate finger. You will also discover that electrostatics powers photocopiers, removes smoke from factory chimneys, and occasionally attacks Form Five students in corridors."*

Kipute is already opening her notebook.

Kipanga, meanwhile, keeps a respectful distance from the brass handle.

Welcome to Chapter 10. We are studying charges at rest: where they come from, what forces they exert on one another, what invisible field they fill the space around them with, what potential and potential energy they carry, what happens when they meet a perfect conductor, and what they do when an outside agent pushes them around.

Two debts from earlier chapters are paid here. The first is to *Foundations of Physics*, Chapter 7. Newton's law of universal gravitation,

$$F = \frac{Gm_1m_2}{r^2}$$

had a twin all along, and you are about to meet it as Coulomb's law,

$$F = \frac{kQ_1Q_2}{r^2}.$$

Everything that worked for planets carries across with the labels changed.

The second debt is to Chapter 1 of this book. The $U(r)$ curve you met for two molecules will reappear as the $U(r)$ of two point charges, which is the cleanest version of that picture you will ever see, with one smooth curve and a sign that comes from the charges.

We close the chapter the way we opened it, with a static-electric story. By the last pages the brass handle, the photocopier, the lightning conductor, the smoke at Tanga Cement, and the cassava-yard conveyor will all sit on one page of running prose, each explained by the same handful of equations. By that page Kipanga will know exactly which electrons crossed the air gap this afternoon, why, and at what speed.

Let us begin.

ELECTRIC CHARGE

Two thousand six hundred years ago, a Greek philosopher called Thales noticed that a piece of amber rubbed with fur would pick up small bits of dry grass and chaff. The Greek word for amber is *elektron*, and that single observation has, by a chain of inheritances, given us the word electricity and the entire industry that lights our towns, charges our phones, and runs the door-handle spark that opened this chapter. Thales did not know what was happening; nobody did, for the next two thousand four hundred years. But the experiment is real, repeats every time, and is the door into a body of physics that this chapter is about to open.

Before we can write down a single equation about electric forces, we have to understand what the rubbed amber actually has more of, and the rubbed fur actually has less of, after the rubbing. That common quantity, the one that flows from one to the other, is what we call **electric charge**. This first section answers four questions about it. *What kinds of charge are there? Can charge be created or destroyed, or does it only move around? Is charge a smooth quantity like length, or does it come in fixed indivisible packets like a brick wall? And finally, by what physical processes does a neutral body become a charged one?* Once those four questions are answered, the machinery of Coulomb, of fields, of Gauss and of potential will follow without resistance.

The two kinds of charge

Take two glass rods, rub each of them with a piece of silk, and hang one of them from a thread so it can swing freely. Bring the second glass rod close to the first. The hanging rod swings away. The two rubbed glass rods repel each other. Now repeat the experiment, but with two plastic rods rubbed with fur instead of two glass rods rubbed with silk. The two rubbed plastic rods also repel each other. Finally, bring a rubbed plastic rod close to the rubbed glass rod hanging on its thread. The hanging rod swings towards the plastic rod, not away from it. Glass and plastic attract.

This humble experiment, which any reader can perform in a school laboratory in fifteen minutes, contains the first big result of electrostatics. The rubbing has put something onto each rod, and that something comes in exactly two flavours. Two glasses repel, two plastics repel, but glass and plastic attract. There is no third flavour and there is no neutral flavour, because all the rods involved have been rubbed. We name the two flavours after the signs we give them on paper. The charge that builds up on the glass rod rubbed with silk is called **positive charge**. The charge that builds up on the plastic rod rubbed with fur is called **negative charge**. The choice of which one to call positive was made by Benjamin Franklin in the eighteenth century

and is, in honesty, an accident of history; if Franklin had flipped his coin the other way, the electrons in your phone would today be the positive carriers. The physics would be unchanged.

The two-flavour rule that the experiment uncovered can now be stated as a working principle of the chapter. *Like charges repel each other, and unlike charges attract each other.* That sentence will be the seed from which every Coulomb force in the chapter grows, and the rubbed glass rod hanging on its thread is the diagnostic test that decides, in any new experiment, whether the unknown sample is positively charged, negatively charged, or not charged at all. If the unknown attracts the glass, the unknown is negative. If the unknown repels the glass, the unknown is positive. If the unknown does nothing measurable to the glass, the unknown is neutral.

Conservation of charge

Where did the charge on the glass rod come from? The glass rod was neutral before the rubbing, the silk was neutral before the rubbing, and yet after a few seconds of contact the glass had become positive. Either charge was made out of nothing during the rubbing, which would be an extraordinary thing to claim, or some charge that was already there moved from one object to the other. Careful experiments performed over the last three centuries have settled the question decisively in favour of the second answer. The silk did not create the charge on the glass. The silk supplied it.

In every experiment that has ever been performed, the total electric charge of an isolated system, that is, of a system with no exchange of matter with anything outside it, has been observed to remain constant. Charge can move from one body to another, can be hidden inside a neutral object by being equal and opposite to an unseen partner, and can be moved around at great speed inside wires; but it cannot be created and it cannot be destroyed. The total stays the same forever.

This experimental fact has been elevated, by its perfect record, to the status of a named principle. We call it the **principle of conservation of charge**:

In any isolated system the algebraic sum of the positive and negative charges is constant, no matter what physical or chemical processes take place inside the system.

When the glass rod was rubbed with silk, the silk gave up some of its electrons to the glass, and ended that exchange with a deficit of electrons of exactly the same magnitude as the deficit the glass had gained. The glass rod is positive by exactly as many micro-coulombs as the silk is negative; the books balance.

A second, more dramatic illustration is the reaction in which a single high-energy gamma photon, carrying no charge at all, materialises into an electron of charge minus e and a positron of charge plus e . The total charge before the event is zero. The total charge after the event is also zero, because the electron and positron carry equal and opposite charges. Charge has not been made out of nothing, even in this most exotic process; it has been made out of an equally large pair that adds to zero. The book of electric accounts is never, ever overdrawn.

Quantisation of charge

A neutral piece of glass can be rubbed against silk to give it a charge of two microcoulombs. *Can it then be rubbed a little less to give it a charge of, say, exactly one half of a microcoulomb? Or one third? Or any tiny fraction at all?* The answer, discovered in 1909 by the American physicist Robert Millikan in his famous oil-drop experiment and confirmed by every careful measurement made since, is **no**. There is a smallest unit of free charge in nature, and the charge on any isolated object is always a **whole-number** multiple of that smallest unit.

The smallest free charge is the magnitude of the charge on a single electron, and is given the symbol **e** . Its measured value, accurate to many decimal places, is $e \approx 1.6 \times 10^{-19} \text{C}$. An electron carries a charge of $-e$, that is $-1.6 \times 10^{-19} \text{C}$, and a proton carries a charge of $+e$, that is $+1.6 \times 10^{-19} \text{C}$. Every free electric charge that has ever been measured is an integer number n of these elementary units, and the total charge Q on an object that has gained or lost n electrons is therefore:

$$Q = ne,$$

where n is a positive whole number when the body has lost electrons (and is therefore positive) and a negative whole number when the body has gained electrons (and is therefore negative). This property, that free charge always comes in whole-number multiples of e , is called the **quantisation of charge**. The word quantisation means that the quantity is not a smooth river but a stack of identical bricks, and you can have one brick, two bricks, a million bricks, but never one and a half bricks.

For ordinary laboratory objects the brick is so small that the staircase looks like a ramp. A charge of one microcoulomb is six and a quarter trillion electrons, so adding or removing one more electron is a step too small to notice. Quantisation matters very much, though, at atomic scales, where individual electrons are the actors. This is not as a peculiar rule that physicists invented but as the simple statement that you cannot have half an electron, just as you cannot have half a brick.

How a body becomes charged

A neutral body has equal amounts of positive and negative charge, the positives bolted onto the nuclei and unable to move, the negatives carried by the outer electrons of the atoms and able to move with relative ease through some materials. To charge a body is to give it a net excess of one kind, which by conservation requires that an equal amount of the other kind ends up somewhere else. Three standard procedures do this, namely: charging by friction, charging by conduction, and charging by induction. Each one is best understood by following the electrons.

1) Charging by friction

When two different insulating materials are rubbed firmly against each other, electrons from the surface atoms of one material are transferred mechanically to the surface atoms of the other. Which material gives up its electrons and which one receives them depends on the chemical nature of the two surfaces, and the order has been catalogued experimentally into a list called the **triboelectric series**. Near one end of the series sit *glass, hair, and wool*, which **lose electrons easily**. Near the other end sit *rubber, amber, and plastic*, which **accept electrons easily**. The **further apart** two materials sit on the series, *the more charge each one acquires when they are rubbed together*.

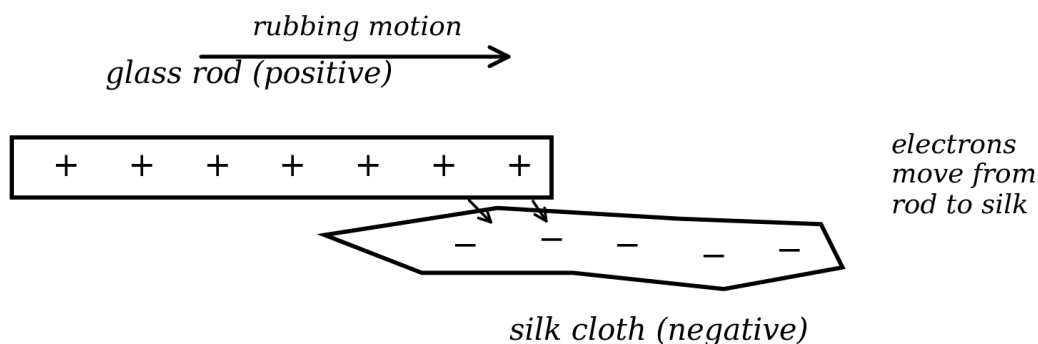


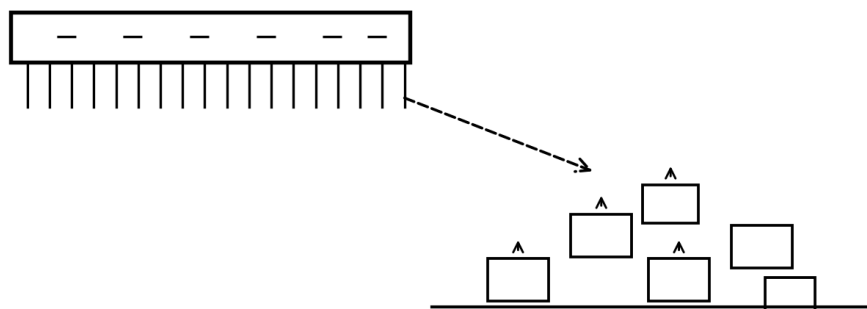
Figure: Charging by friction. Rubbing a glass rod with silk transfers electrons from the surface of the glass to the surface of the silk. The glass rod is left with a deficit of electrons and is therefore positive; the silk is left with an excess of electrons and is therefore negative. The total charge of the rod-and-silk system is zero before and after the rubbing, as required by the conservation of charge.

Charging by friction is the procedure by which the door handle of the introduction acquired its charge. Kipanga's rubber soles dragged across the polished concrete of the corridor and exchanged electrons with the floor; the shoes (and Kipanga's body, which is in good electrical contact with the shoes) became charged. The procedure works best on dry days. On a humid day the moisture in the air forms a thin conducting film over the surfaces, and the charge leaks away through that film as fast as the rubbing puts it on. This is why the famous door handle is famous in the dry season and unremarkable in the rainy one.

The reader who would like to see **triboelectric charging** in their own hand can do so in two minutes. Tear a sheet of writing paper into pieces no larger than a fingernail, scatter them on a dry tabletop, and pass a plastic comb a few times briskly through clean dry hair. Held two centimetres above the scattered paper, the comb will lift a handful of the pieces up to its teeth and hold them there. Hair sits above plastic on the triboelectric series, so the comb gains electrons during the rubbing and becomes negatively charged. The paper, although neutral overall, is polarised by the field of the comb (the same kind of induced separation we shall see in the next subsection), and the net force on each polarised piece is upward towards the comb. After a few seconds some of the lifted pieces fall back to the table, because the comb has lost a little of its

charge to the surrounding air. On a humid day almost nothing rises; on a dry day the whole demonstration works first time.

plastic comb (rubbed with hair)



torn paper bits on a table

Figure: A plastic comb that has been rubbed through dry hair gains a small negative charge and lifts torn pieces of paper from a flat surface. The paper itself has no net charge, but the comb induces a separation of charge inside each piece (negative side away, positive side towards the comb), and the resulting attraction is sufficient to lift the piece against gravity.

2) Charging by conduction

A second way to charge a previously neutral body is to put it into direct contact with a body that is already charged. If the two bodies are conductors, the excess charge on the charged body redistributes itself between the two until both objects sit at the same electric potential, and they share the available charge between them. The previously neutral body now carries a charge of the same sign as the original charged body, by exactly the amount that has flowed across. The original body has lost an equal amount, and the total charge of the two-body system is unchanged. **Charging by conduction** therefore *always produces a body of the same sign as the body that did the charging*.

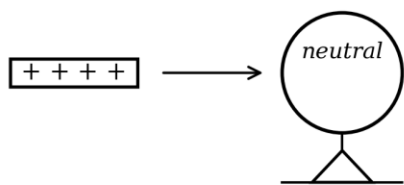
A small subtlety of conduction is worth flagging. *If the previously neutral body is large compared to the charged body that touches it, the charged body loses almost all of its charge to the larger one.* In the limit of a very large body, such as the Earth itself, the charged body loses essentially all of its charge and ends up neutral. This is why touching a charged metal object to the ground discharges it almost completely, and it is the principle behind every earth wire in every electrical installation in your home.

3) Charging by induction

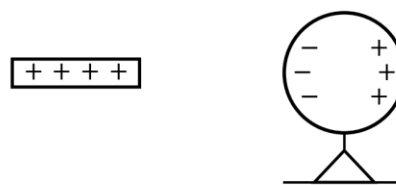
The third procedure is the cleverest, because it produces a charged body without the charging body ever touching it. It is called charging by induction, and it relies on the fact that the free electrons in a conductor are free to move when they feel an electric force.

The procedure has four steps and is best followed on the figure that accompanies this paragraph. In Step 1 a charged rod, say a positively charged glass rod, is brought close to a neutral conducting sphere mounted on an insulating stand. The sphere does not touch the rod. In Step 2 the free electrons inside the sphere, being negative, are attracted by the positive rod and crowd onto the near side of the sphere, leaving the far side with an equal deficit of electrons (that is, a positive surface region). The sphere as a whole is still neutral, but its charge has rearranged itself.

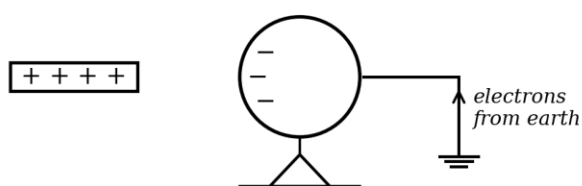
Step 1: a charged rod is brought near



Step 2: charges separate inside the sphere



Step 3: the sphere is briefly earthed



Step 4: earth wire then rod removed

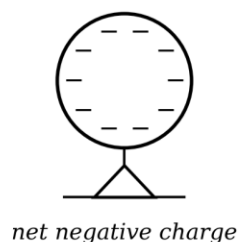


Figure: Charging by induction. In Step 1 a positively charged rod is brought near a neutral conducting sphere on an insulating stand. In Step 2 the free electrons in the sphere are attracted to the near side, leaving the far side with an equal positive region. In Step 3 the sphere is briefly earthed, so the unwanted far-side positive region is neutralised by electrons that flow up from the earth. In Step 4 the earth connection is removed first and then the rod is removed, leaving the sphere with a net negative charge of the opposite sign to the rod.

In Step 3 a connecting wire briefly joins the sphere to the earth. The earth is effectively a very large reservoir of charge, and electrons flow up the wire from the earth into the sphere to neutralise the positive far-side region. (One can equivalently say that the positive charges flow down the wire into the earth; the two pictures are equally valid and produce the same final state.) The sphere now carries a net excess of negative charge, all of it bound to the near side by the still-present positive rod. In Step 4 the earth wire is removed first, and only then is the rod removed. With the rod gone, the trapped electrons on the near side are free to redistribute, and they spread out over the entire surface of the sphere. The sphere is left with a net negative charge, opposite in sign to the rod that caused all the rearrangement.

Three small details of induction repay attention.

- The rod never touched the sphere, so no charge crossed between them; the rod's own charge is unchanged at the end of the procedure.
- The order of removal matters: the earth wire must be lifted before the rod, because lifting the rod first would simply let the bound electrons on the sphere escape back to earth before they can be trapped.
- The sphere ends up with a charge of the sign opposite to the rod, in contrast to charging by conduction which always produces the same sign.

Examination questions that ask students to predict the final sign always reward those who remember which procedure was used.

The book has so far been all words and no numbers, like a football commentator who never tells you the score. Two worked examples now arrive to fix that.

BINDER Example 1

A small metal sphere carries a positive charge of $+5.0\mu\text{C}$ after being rubbed with silk. Calculate the number of electrons that the sphere has lost during the rubbing. Take the elementary charge to be $e = 1.6 \times 10^{-19}\text{C}$.

Solution

The charge on the sphere is positive, which by the quantisation rule means the sphere has lost a whole number of electrons rather than gained any. Calling that whole number n , the magnitude of the charge on the sphere is the product of n and the elementary charge:

$$Q = ne$$

Make n the subject and substitute:

$$n = \frac{Q}{e} = \frac{5.0 \times 10^{-6} \text{C}}{1.6 \times 10^{-19} \text{C}} = 3.125 \times 10^{13}$$

The sphere has therefore lost about 3.125×10^{13} electrons (about thirty-one trillion electrons).

Making Sense of the Answer: *Thirty-one trillion sounds like an unimaginable number to lose, and yet five microcoulombs is a perfectly ordinary amount of charge to find on a rubbed object in a school laboratory. The two facts together tell us that a single electron carries an extremely small charge, so an ordinary rubbed body has to lose a colossal number of electrons before the total charge becomes laboratory-sized. The sphere as a whole, after losing those thirty-one trillion electrons, has not lost any measurable mass and has not lost any measurable size.*

Think Like a Physicist: *Whenever you are given a charge in coulombs and asked for a count of electrons, the answer is always going to be the charge divided by e , and the only question is where the decimal point lands. The trick is to remember that e is a very small number, of order 10^{-19} , so dividing by it produces a very large count. Conversely, whenever you are asked for the charge produced by a given number of electrons, you will multiply by e , and the count will collapse into a tiny charge in coulombs.*

REAL Example 2

On a dry afternoon a child rubs an inflated rubber balloon briskly against her woollen jumper and presses it lightly against the painted plaster wall of her classroom. The balloon sticks to the wall and remains there for several minutes without any tape or glue. Explain in terms of charge conservation and induction:

- what sign of charge the balloon acquires from the rubbing,
- why the balloon sticks to a wall that is itself electrically neutral, and
- why the balloon eventually falls off after a long enough wait.

Solution

(a) Rubber sits well below wool on the triboelectric series, which means that when the two materials are rubbed together electrons are transferred from the wool to the rubber. The balloon, made of rubber, therefore gains electrons during the rubbing and ends up with a negative net charge. The jumper loses an exactly equal number of electrons during the same rubbing and ends up with a positive net charge of the same magnitude. The total charge of balloon plus jumper is zero before the rubbing and zero after, in obedience to the conservation of charge.

(b) The painted plaster wall is neutral overall, with equal amounts of positive and negative charge tied up in its atoms. When the negatively charged balloon is brought close to the wall, its negative surface charge pushes the negative electrons in the surface layer of the wall slightly inward, and pulls the positive nuclei of those atoms slightly outward. This is induction in an insulator: no charge is free to travel across the wall, but each atom is slightly distorted, so the side of every atom nearer the balloon becomes a little positive, and the side further from the balloon becomes a little negative. The result is a thin layer of induced positive charge on the wall at the spot directly under the balloon. This thin induced positive layer attracts the negative balloon. Because the induced positive layer is closer to the balloon than the induced negative layer behind it, the attraction is stronger than the repulsion, and the net force on the balloon is towards the wall. That net force is what holds the balloon up against its own weight.

(c) The induced separation of charge in the wall is not maintained by the wall itself; it is maintained by the continued presence of the charged balloon. Over time, two things happen. The first is that some of the negative charge on the balloon's surface slowly leaks away into the air, especially if the air carries a little humidity, and the balloon's charge falls. The second is that the polarised molecules in the wall slowly relax back to their unbiased positions, partly because of thermal motion and partly because their internal electric forces resist being held apart. As the balloon's charge falls and the wall's polarisation relaxes, the attractive force weakens until it is no longer large enough to support the balloon's weight, at which point gravity wins and the balloon slides down the wall and falls. On a humid day this happens in a shorter time (within a minute or so); on a dry day it takes longer time.

Making Sense of the Answer: *The balloon-on-the-wall trick, which most readers will have performed as small children long before they took a physics class, is in fact a beautifully complete electrostatics*

experiment. It uses friction to charge the balloon, induction to charge the wall, and the resulting attraction to hold the balloon up against gravity. Three of the four big ideas of this section (the two kinds of charge, conservation of charge, and induction) appear in a single household demonstration. The fourth, quantisation, is hidden inside the calculation of just how many electrons moved, and is the only one that requires a number to be revealed.

Think Like a Physicist: An attractive force between a charged object and an apparently neutral one is almost always a sign that induction is at work. The neutral object is responding to the charged one by separating its own internal charges slightly, and the separated charges then experience unequal forces because they are at slightly different distances from the charged object. You will meet this same idea three more times in the chapter, once for the rubbed comb earlier in this section, once for the bird on the high-voltage power line, and once for the electrostatic precipitator that cleans Twiga Cement's chimney exhaust.

Two worked examples now lie behind us, and four big results have been bagged: that charge comes in two flavours, that the books always balance, that the smallest brick in the wall has the value of e , and that a body can be charged by three different procedures. The balloon you stuck to the wall in primary school turns out to have been a working electrostatics laboratory, and you can now charge for the rent.

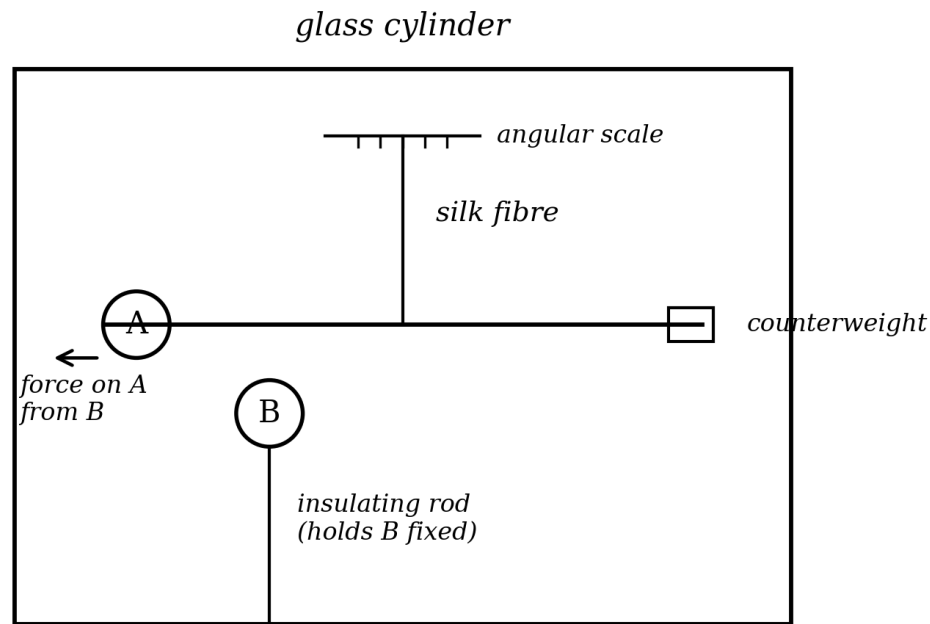
With two charged objects in hand and a clear list of the rules that govern their charges, the next obvious question is the size of the force one of them exerts on the other. That question was answered, in elegant quantitative form, by a French military engineer in 1785. His result is the subject of the next section, and the inverse-square law it contains will turn out to be the electrostatic twin of the gravitational law you met in *Foundations of Physics*.

COULOMB'S LAW AND THE SUPERPOSITION PRINCIPLE

The previous section settled what charge is, where it comes from, and how it moves between bodies. It did not say how much force two charged bodies exert on each other. The man who answered that question lived two hundred and forty years ago in revolutionary France, and the answer he found is one of those rare results in physics that fits on a single line, applies from the size of an atom to the size of a thunderstorm, and never has a known exception in any laboratory experiment. We meet his result, meet the law of superposition that turns it into a tool for arbitrarily many charges, and find, at the end of the section, that the law we have just learned has a twin in another part of this series of books.

What Coulomb measured

In 1785, the French military engineer Charles-Augustin de Coulomb hung a light horizontal insulating rod from a thin silk fibre inside a tall glass cylinder. At one end of the rod sat a small metal ball; at the other a small paper counterweight. Coulomb gave the ball a known electric charge, and then brought a second charged ball, fixed on its own insulating rod, very close to the first one without touching it. The repulsion between the two charged balls twisted the suspending fibre by a small angle, and Coulomb read the angle on a scale at the top of the cylinder. Because the twist of a thin fibre is known to be proportional to the torque applied to it, the read angle told him, directly, the size of the repulsive force.



Coulomb's torsion balance (1785)

Figure: Coulomb's torsion balance. A light insulating rod, suspended at its centre by a thin silk fibre inside a tall glass cylinder, carries a small charged ball A at one end and a paper counterweight at the other. A second charged ball B, fixed on its own insulating stem, is brought close to A. The repulsion between the two charged balls twists the fibre, and the angle of twist, read on the scale at the top, measures the force between them.

Coulomb repeated the experiment many times. He doubled the distance between the balls and saw the force fall to one quarter of its earlier value. He tripled the distance and saw it fall to one ninth. The pattern was unmistakable. The force between two charged balls falls as the square of the distance between them increases:

$$F \propto \frac{1}{r^2}$$

Coulomb then varied the amount of charge on the two balls, using the trick of sharing charge by contact (a charged ball briefly touched to an identical neutral ball shares its charge equally between them, so dividing the original charge by two; repeating the contact divides by four; and so on). He found that the force is proportional to each of the two charges:

$$F \propto Q_1 Q_2$$

Both proportionalities are observed at the same time, so the force is jointly proportional to the product of the two charges and inversely proportional to the square of the distance between them. Combining the two relations into a single statement gives the inverse-square law of electrostatics, which is universally known today as **Coulomb's law**:

The electrostatic force between two stationary point charges is directly proportional to the product of their charges and inversely proportional to the square of the distance between them, acting along the line joining the charges.

Mathematically:

$$F = \frac{kQ_1 Q_2}{r^2}$$

The constant k is the proportionality constant whose numerical value depends on the units in which charge, distance, and force are measured. In the SI system, in which charge is measured in coulombs, distance in metres, and force in newtons, the experimentally measured value of this constant is:

$$k = 9.0 \times 10^9 \text{ Nm}^2/\text{C}^2,$$

to two significant figures, and is the value we shall use throughout every numerical calculation in this chapter.

Why the constant is written as $1/(4\pi\epsilon_0)$

In nearly every physics textbook that goes beyond an introductory level, the constant k in Coulomb's law is rewritten in the slightly odd-looking form:

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

so that the law is written as

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0} \frac{Q_1 Q_2}{r^2},$$

in which ϵ_0 is called the permittivity of free space. The first reaction on meeting this expression is to ask, very reasonably, why anyone would replace a single clean letter k with a fraction containing a 4π factor. The reason is geometric, and it foreshadows a powerful theorem we shall meet later in this chapter (Gauss's law).

Imagine a single point charge sitting at the centre of an imaginary sphere of radius r . The surface area of that sphere is $4\pi r^2$. If we picture the lines of electric influence that emerge from the charge as fanning out uniformly into all directions, then the density of those lines on the surface of the sphere is the total number of lines divided by $4\pi r^2$. Writing the constant in Coulomb's law as $\frac{1}{4\pi\epsilon_0}$ folds this geometric 4π factor directly into the law, and leaves the new constant ϵ_0 to carry only the physics of how strongly empty space resists being threaded by electric influence. It is, in the end, an elegant book-keeping choice that pays its dividends in three sections' time, when Gauss's law uses exactly this rearrangement to make the algebra of fields and fluxes work in one line. For the moment, take it on trust, and remember the numerical value:

$$\epsilon_0 = 8.854 \times 10^{-12} \text{ C}^2/(\text{Nm}^2)$$

Permittivity and relative permittivity

The constant ϵ_0 that appears above is, strictly speaking, the permittivity of a vacuum: the space between the two charged balls of Coulomb's torsion balance contained almost no matter, and the value of ϵ_0 just quoted describes that empty space. If the two charges are immersed in a different material, the force between them changes, and almost always falls. A drop of water placed between two charged spheres reduces the force between them by a factor of about eighty, because the polar water molecules rearrange themselves to oppose the field, weakening it.

To capture this change quantitatively, we introduce a dimensionless number ϵ_r called the **relative permittivity** (also termed as the **dielectric constant**) of the material. The permittivity of the material is then:

$$\epsilon = \epsilon_r \epsilon_0 \quad \left(\text{from } \epsilon_r = \frac{\epsilon}{\epsilon_0} \right)$$

and Coulomb's law inside the material reads:

$$F = \frac{1}{4\pi\epsilon} \frac{Q_1 Q_2}{r^2} = \frac{1}{4\pi\epsilon_r \epsilon_0} \frac{Q_1 Q_2}{r^2} = \left(\frac{1}{4\pi\epsilon_0} \right) \frac{Q_1 Q_2}{\epsilon_r r^2} = \frac{k Q_1 Q_2}{\epsilon_r r^2}$$

For ordinary air at sea level, ϵ_r is almost exactly one (about 1.0006), and we shall treat the force between two charges in air as identical to the force between the same charges in a vacuum, so that no correction is needed when laboratory problems specify "in air". For paraffin, ϵ_r is about 2.2; for glass, between 4 and 10 depending on the kind; for distilled water, about 80; for the metals, the concept does not apply because free charges inside a conductor simply screen the field and reduce the force to zero. We shall return to this case later.

Coulomb force has a direction as well as a magnitude

The scalar form of Coulomb's law gives the magnitude of the force, but every force has a direction, and the direction of the Coulomb force is fixed by the geometry of the two charges. Each charge feels a force directed along the line joining it to the other charge. If the two charges have the same sign (both positive or both

negative), the force on each is directed away from the other, and the two charges repel. If the two charges have opposite signs, the force on each is directed towards the other, and they attract.

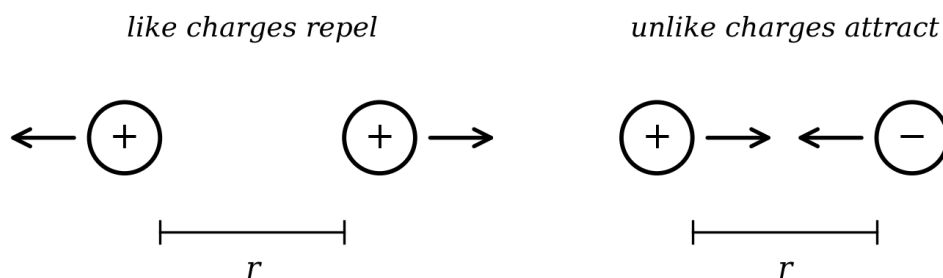


Figure: Like charges repel along the line joining their centres; unlike charges attract along the same line. In both cases the two forces on the pair are equal in magnitude and opposite in direction, in obedience to Newton's third law. The common separation between the two centres is r .

The force that charge 1 exerts on charge 2, and the force that charge 2 exerts on charge 1, are equal in magnitude and opposite in direction: this is simply Newton's third law, applied to a pair of charged bodies. Algebraically, writing F_{12} for the force on body 1 due to body 2 and F_{21} for the force on body 2 due to body 1, we have:

$$\mathbf{F}_{12} = -\mathbf{F}_{21},$$

and these two opposite vectors share the same line of action, which is the line joining the centres of the two charges. The sign of $Q_1 Q_2$ in the formula encodes the rest: when the product is positive (two like charges) the force is repulsive, and when the product is negative (two unlike charges) the force is attractive. In a problem in which the directions matter, the practical procedure is to compute the magnitude of the force using the magnitudes of the charges, then add the direction by sketching the geometry and applying the like-repels, unlike-attracts rule.

Both pieces of information about the Coulomb force, its size and its direction, can be wrapped into a single **vector equation**. Introduce a unit vector $\hat{r}$, a vector of length one whose only purpose is to mark the direction along the line joining the two charges, pointing from the first charge to the second. The force on the second charge due to the first can then be written in one line as

$$\mathbf{F} = k \frac{Q_1 Q_2}{r^2} \hat{r}$$

Read the equation in two halves. The fraction $k \frac{Q_1 Q_2}{r^2}$ carries the size of the force, exactly as it did in the scalar form, with a sign attached to it that depends on the signs of the two charges. The unit vector $\hat{r}$ carries the direction. When the product $Q_1 Q_2$ is positive, the two charges sharing the same sign, the equation hands back a vector pointing along $\hat{r}$, away from the first charge, and the force is the repulsion the rule of signs has always predicted. When the product is negative, the two signs opposite, the equation hands back a vector pointing the other way, along $-\hat{r}$, back towards the first charge, and the force is an attraction. The geometry of attraction and repulsion is no longer a separate rule applied by hand after the calculation; it lives inside the formula itself, governed by the sign of the product of the charges.

The force on the first charge due to the second is, by Newton's third law already noted above, the same vector reversed in direction:

$$\mathbf{F}_{12} = -\mathbf{F}_{21}$$

Either form of Coulomb's law, the scalar one with directions decided by the rule of signs, or the vector one with directions hidden inside the symbols, is correct. The two are not in competition; they are two ways of saying the same thing.

The principle of superposition

Coulomb's law concerns a pair of charges. Real problems concern collections of three, four, or many more charges. To find the total force on one of the charges produced by all the others, we appeal to a separate experimental fact, called the **principle of superposition**. Stated simply, the principle says that:

The total electric force on a particular charge is the vector sum of the individual Coulomb forces that each of the other charges, taken one at a time, would exert on it.

The principle is not an automatic consequence of Coulomb's law. It is a separate experimental discovery, although a deep one: it amounts to the statement that the Coulomb force between two charges does not depend on the presence or absence of any third charge. The two charges do their own business with each other regardless of the company they keep, and the company's contributions add up by ordinary vector addition. If the principle had failed in some way (if a third charge had caused the force between the first two to shrink, for instance), all of the chapter that follows this paragraph would have to be rewritten.

force on q from a collection = vector sum of the pair forces

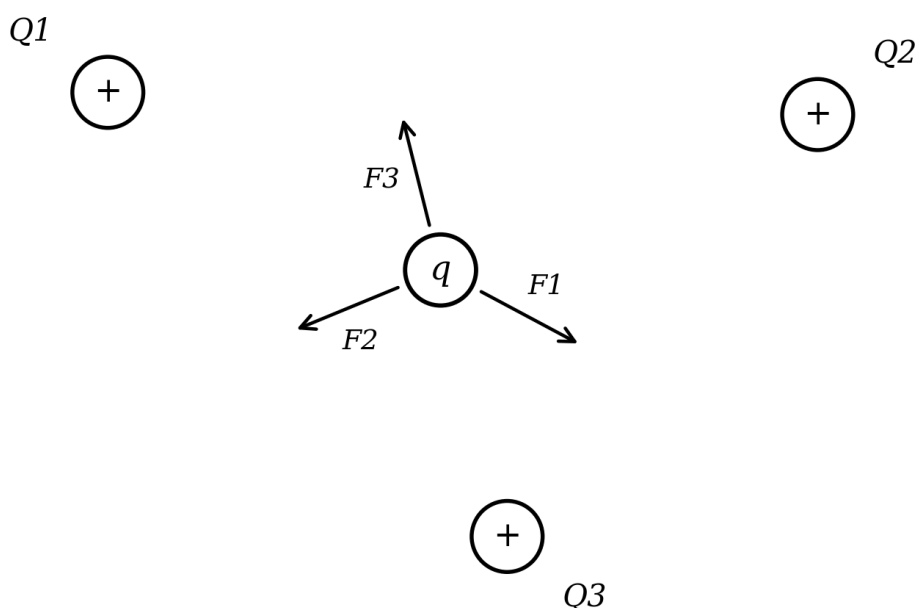


Figure: Superposition. Three source charges Q_1 , Q_2 , and Q_3 each exert a Coulomb force on the target charge q . The total force on q is the vector sum of the three individual forces F_1 , F_2 , and F_3 . Each component force is computed using Coulomb's law as if the other two were not there.

Symbolically, if q is the charge on which we want the force and Q_1 , Q_2 , Q_3 , and so on are the charges that act on it, then the total force on q is:

$$F_{\text{total}} = F_1 + F_2 + F_3 + \dots,$$

where each F_i is the Coulomb force that the source Q_i alone would exert on q , computed by Coulomb's law and directed along the line joining Q_i to q . The addition is vector addition, with magnitudes and directions both honoured. In problems with symmetric arrangements (charges at the corners of a square, an equilateral triangle, or a regular ring), this addition often collapses by symmetry to a single non-zero component along an axis of symmetry; in less symmetric arrangements, the addition is straightforward but tedious, as we will see later.

In the actual numerical practice of these problems, the vector sum is almost always carried out by components. Choose a pair of axes, the natural pair being the horizontal x and the vertical y , resolve each individual Coulomb force F_i into its components along those axes, and add the components along each axis separately:

$$\Sigma F_x = F_{1x} + F_{2x} + F_{3x} + \dots$$

$$\Sigma F_y = F_{1y} + F_{2y} + F_{3y} + \dots$$

The two component-sums then combine into a single resultant vector whose magnitude is given by,

$$F = \sqrt{(\Sigma F_x)^2 + (\Sigma F_y)^2}$$

and whose direction makes an angle

$$\theta = \tan^{-1} \left(\frac{\Sigma F_y}{\Sigma F_x} \right)$$

with the chosen x axis. This four-step procedure, resolve, sum along each axis, combine into a single magnitude, then find the angle, is the procedure every examination-grade worked example follows, and the one that should be reached for the moment three or more charges enter a question.

The twin of Newton's law of gravitation

A reader of our previous book, Foundations of Physics, will feel a strong sense of recognition on meeting Coulomb's law for the first time. The reason is that this is not the first inverse-square law of nature you have met. Newton's universal law of gravitation, which described the force between two masses m_1 and m_2 separated by a distance r , reads:

$$F = \frac{Gm_1m_2}{r^2},$$

where G is the gravitational constant. Set this beside Coulomb's law,

$$F = \frac{kQ_1Q_2}{r^2},$$

and the two structures are identical: a constant of nature, multiplied by the product of two source quantities (masses for Newton, charges for Coulomb), divided by the square of the distance between them. The two laws are twins, and the machinery you will build for one of them transfers, with almost no change of notation, to the other. Gravitational fields, gravitational potentials, gravitational potential energies, and the principle of superposition for gravity all reappear in this chapter under electric names, and the derivations look the same on both sides.

Two important disanalogies must be flagged, however, because every examination paper that touches Coulomb's law tests them.

The first is that *charge has two signs, positive and negative, whereas mass has only one*. Gravity is therefore always attractive, while the electric force is sometimes attractive and sometimes repulsive depending on the signs of the two charges.

The second is that *the electric force is enormously stronger than gravity at the atomic scale*. For an electron and a proton separated by one angstrom, the ratio of the electric force to the gravitational force between them is about 10^{39} , a number so large that gravity is utterly negligible in every atomic and molecular process. The reason a rubbed comb can lift a piece of paper against the entire gravitational pull of the Earth is hidden in that one ratio: at the molecular scale, electric forces are not just stronger than gravity, they are **very much stronger**.

A formula in a frame is one thing; a formula that has computed a numerical answer is another, and is usually the one that sticks. Two worked examples now exercise Coulomb's law on a pair of perfectly ordinary numbers, and then on a pair of numbers so small that the result changes the way you think about both gravity and atoms.

BINDER Example 3

Two small spheres carry charges of $+10\text{nC}$ and $+10\text{nC}$ respectively, and their centres are 5.0cm apart in air. Calculate the magnitude of the electric force on each sphere. Take $k = 9.0 \times 10^9 \text{ Nm}^2/\text{C}^2$.

Solution

Both charges are positive, so the force on each is repulsive and directed along the line joining their centres, away from the other sphere. The magnitudes are equal by Newton's third law, so we need only compute one of them. The magnitude of the Coulomb force between the two spheres in air is given directly by Coulomb's law:

$$F = \frac{kQ_1Q_2}{r^2}$$

Substitute the given values, with every quantity expressed in SI units:

$$F = \frac{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2)(10 \times 10^{-9} \text{ C})(10 \times 10^{-9} \text{ C})}{(0.050 \text{ m})^2} = 3.6 \times 10^{-4} \text{ N}$$

The magnitude of the force on each sphere is therefore $3.6 \times 10^{-4} \text{ N}$, or about 0.36mN.

Making Sense of the Answer: One third of a millinewton sounds like a small force, and on human scales it is. To feel it in your hand you would have to put the equivalent of about a thirty-fifth of a gram on a sensitive kitchen balance, which is below the resolution of most balances available in a school laboratory. Yet ten nanocoulombs is a tiny amount of charge (it is $\frac{1}{10^{11}}$ of a coulomb), and the two spheres were five centimetres apart, which is far enough that the inverse-square law has already cut the force back hard. The conclusion to draw is that even very small charges, in a small piece of ordinary laboratory apparatus, produce forces that a precise instrument like a torsion balance can read clearly.

Think Like a Physicist: Whenever you read off the answer to a Coulomb-force calculation, look at the three powers of ten in the original data and check that they have arrived at the answer in the right combination. In this problem we had two factors of 10^{-9} from the charges in coulombs, one factor of 10^9 from the constant k , and two factors of 100 from converting centimetres to metres in the denominator. Add up the exponents and you get -4 : the answer has to be of order 10^{-4} newtons. A rough exponent check like this catches most arithmetic slips long before the final number arrives, and it is the single fastest way to know whether you have the order of magnitude right.

REAL Example 4

In the laboratory of Ilboru Secondary School, two small identical pith balls of mass $m = 0.10\text{g}$ each are hung from a common point by two light insulating threads of length $L = 50\text{cm}$. Each ball is given the same positive charge Q , and the two threads settle in equilibrium at an angle of $\theta = 8^\circ$ from the vertical. Calculate the charge Q on each ball. Take $g = 9.8\text{m/s}^2$ and $k = 9.0 \times 10^9 \text{ Nm}^2/\text{C}^2$.

Solution

Each ball is in equilibrium under three forces: its weight mg acting straight down, the tension T in its supporting thread acting along the thread back to the common pivot, and the Coulomb repulsion F from the other ball, acting horizontally outward. We start with a picture of the geometry and the forces, then write down the two equilibrium equations (one for the horizontal direction, one for the vertical direction), and finally substitute Coulomb's law for F .

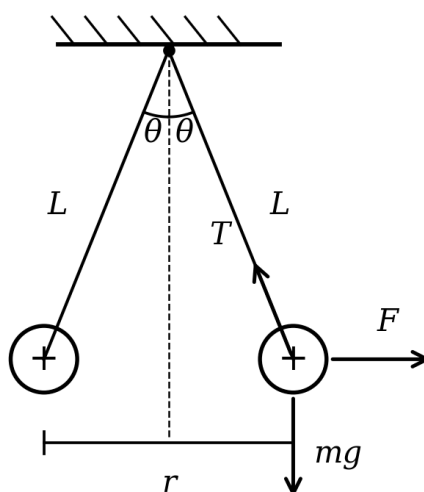


Figure: Two pith balls of mass m hang from a common point by threads of length L , each carrying charge Q . In equilibrium each thread makes an angle θ with the vertical, the centre-to-centre separation is $r = 2L \sin\theta$, and each ball feels three forces: its weight mg , the thread tension T , and the Coulomb repulsion F .

Resolving the forces on (say) the right-hand ball, the vertical equilibrium condition reads:

$$T \cos\theta = mg,$$

and the horizontal equilibrium condition reads:

$$T \sin\theta = F.$$

Dividing the horizontal equation by the vertical one, T cancels and we are left with:

$$F = mg \tan\theta$$

The geometry fixes the centre-to-centre separation of the two balls. Each ball sits a horizontal distance $L \sin\theta$ from the vertical line through the pivot, so the two balls are separated by:

$$r = L \sin\theta + L \sin\theta = 2L \sin\theta$$

Now write Coulomb's law for the repulsion between the two balls, each carrying the same charge Q :

$$F = \frac{kQ^2}{r^2}$$

Equating the two expressions for F and making Q^2 the subject of the formula:

$$\begin{aligned} mg \tan\theta &= \frac{kQ^2}{r^2} \\ Q^2 &= \frac{mgr^2 \tan\theta}{k} = \frac{mg(2L \sin\theta)^2 \tan\theta}{k} \end{aligned}$$

Substitute the given values:

$$\begin{aligned} Q^2 &= \frac{(1.0 \times 10^{-4} \text{ kg})(9.8 \text{ m/s}^2) \times (2(0.50 \text{ m}) \sin 8^\circ)^2 \times \tan 8^\circ}{9.0 \times 10^9 \text{ Nm}^2/\text{C}^2} = 2.97 \times 10^{-16} \text{ C}^2 \\ Q &= \sqrt{2.97 \times 10^{-16} \text{ C}^2} = 1.72 \times 10^{-8} \text{ C} \end{aligned}$$

The charge on each ball is therefore $1.72 \times 10^{-8} \text{ C}$ or about 17 nC .

Making Sense of the Answer: Seventeen nanocoulombs sits comfortably between the value used in BINDER Example 3 (ten nanocoulombs) and the rough amount of charge a rubbed plastic comb acquires during a few brisk strokes through dry hair. The experiment described here is, in fact, a routine school laboratory demonstration: any teacher with a pair of pith balls, a length of silk thread, and the ability to read a small angle from a protractor can run the experiment in ten minutes and reproduce the calculation we have just done. The pith-ball experiment is the modern descendant of Coulomb's torsion balance, and the formula $F = mg \tan\theta$ is the cleanest way ever found to convert a sticky question (how do I weigh an electrostatic force?) into a tidy one (how do I measure an angle?).

Think Like a Physicist: When the equilibrium of a charged body involves gravity and a string, divide the horizontal and vertical equations before substituting anything numerical. The tension T is usually a quantity nobody cares about, and dividing the two equilibrium equations cancels T immediately, replacing two unknowns with one. The same trick applies whenever a string holds up a charged or massed object at an angle, and you will see it again in the lightning-rod problems and the suspended oil-drop problems later in this chapter.

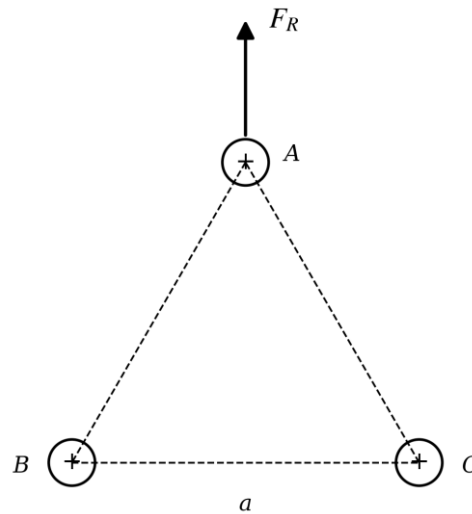
HOT Example 5

Three identical positive point charges, each of magnitude $2.0 \mu\text{C}$, are placed at the vertices of an equilateral triangle of side 30 cm in air. Calculate the magnitude and direction of the resultant Coulomb force on any one of the charges. Take the Coulomb constant to be $9.0 \times 10^9 \text{ Nm}^2/\text{C}^2$.

Solution

Label the three vertices A, B, and C, and consider the force on the charge at A. By the principle of superposition, this force is the vector sum of the Coulomb forces exerted on the charge at A by the charges

at B and at C. Both contributing forces have the same magnitude, because all three charges are equal and B and C are at the same distance from A (the side of the equilateral triangle). The geometry of the problem, with the resultant force on the charge at A indicated, is shown in the figure.



Let Q be the magnitude of each charge and a be the side of the triangle:

$$Q = 2.0\mu\text{C} = 2.0 \times 10^{-6} \text{ C}$$

$$a = 30\text{cm} = 0.30\text{m}$$

The magnitude of the Coulomb force on the charge at A due to a single neighbour is given by Coulomb's law:

$$F = \frac{kQ^2}{a^2}$$

Substitute the values:

$$F = \frac{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2)(2.0 \times 10^{-6} \text{ C})^2}{(0.30\text{m})^2} = 0.40 \text{ N}$$

Each of the two contributions on the charge at A has therefore magnitude $F = 0.40\text{N}$. The directions of these two contributions point along the lines BA and CA, extended outward past A (repulsion is from like to like). The two vectors are separated by the interior angle of the equilateral triangle at the vertex A, which is 60° . By the parallelogram rule, the magnitude of their resultant is:

$$F_R = \sqrt{F^2 + F^2 + 2F^2 \cos 60^\circ}$$

Using $\cos 60^\circ = 0.5$, the expression under the square root collapses to $3F^2$:

$$F_R = \sqrt{3F^2} = F\sqrt{3}$$

Substitute the value of F :

$$F_R = (0.40\text{N})\sqrt{3} = 0.69\text{N}$$

The magnitude is $F_R = 0.69\text{N}$.

By symmetry, the direction of the resultant force on the charge at A lies along the perpendicular bisector of the side BC, pointing outward through A away from the centroid of the triangle.

Making Sense of the Answer: The resultant of two equal forces separated by 60° is larger than either alone by a factor of $\sqrt{3}$, not by a factor of two. The reason is that the two contributions do not point in the same direction: they pull at sixty degrees apart, so part of each vector cancels the other's perpendicular component, and only a fraction of each survives in the resultant. The same geometric factor of $\sqrt{3}$ appears at every vertex of every equilateral triangle of like charges, no matter what the charge or the side length, and is worth committing to memory. The result also illustrates the principle of superposition in its purest

form: the force on one charge from a collection is the vector sum of the pair-force contributions, each computed as if the other source were the only one in the universe.

Think Like a Physicist: Symmetric arrangements of **equal** charges always reduce the vector-addition step to a single geometric factor. For three equal like charges at the vertices of an equilateral triangle, the factor on each vertex is $\sqrt{3}$. For four equal like charges at the corners of a square, the factor on each corner is the sum of three pair forces whose components partially cancel by symmetry. The rule of thumb: sketch the arrangement, identify the symmetry axis through the chosen charge, project each pair force onto that axis, and the magnitudes parallel to the axis add while the components perpendicular to it cancel. The arithmetic that remains is one Coulomb force times a geometric factor that depends only on how many charges are in the arrangement and where they sit.

HOT Example 6

In a hydrogen atom, an electron orbits a proton at an average distance of about 5.3×10^{-11} m. The electron has charge $-e$ and mass $m_e = 9.11 \times 10^{-31}$ kg; the proton has charge $+e$ and mass $m_p = 1.67 \times 10^{-27}$ kg; the elementary charge is $e = 1.6 \times 10^{-19}$ C. Calculate (a) the magnitude of the electric force between the electron and the proton, (b) the magnitude of the gravitational force between them, and (c) the ratio of the two forces. Take $G = 6.67 \times 10^{-11}$ Nm²/kg².

Solution

(a) The electric force is attractive (the charges are opposite), and its magnitude follows directly from Coulomb's law:

$$F_E = \frac{kQ_1Q_2}{r^2} = \frac{ke^2}{r^2}$$

Substitute the given values:

$$F_E = (9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times \frac{(1.6 \times 10^{-19} \text{ C})^2}{(5.3 \times 10^{-11} \text{ m})^2} = 8.2 \times 10^{-8} \text{ N}$$

The electric force has magnitude $F_E = 8.2 \times 10^{-8}$ N, directed from the electron towards the proton.

(b) The gravitational force is also attractive, and its magnitude follows from Newton's law of gravitation:

$$F_G = \frac{Gm_em_p}{r^2}$$

Substituting:

$$F_G = (6.67 \times 10^{-11} \text{ Nm}^2/\text{kg}^2) \times \frac{(9.11 \times 10^{-31} \text{ kg})(1.67 \times 10^{-27} \text{ kg})}{(5.3 \times 10^{-11} \text{ m})^2} = 3.6 \times 10^{-47} \text{ N}$$

(c) The ratio of the electric force to the gravitational force is:

$$\frac{F_E}{F_G} = \frac{8.2 \times 10^{-8} \text{ N}}{3.6 \times 10^{-47} \text{ N}} = 2.3 \times 10^{39}$$

The electric force between the electron and the proton in a hydrogen atom is about 2.3×10^{39} times stronger than the gravitational force between them.

Making Sense of the Answer: The number 10^{39} has a way of disappearing into the page when written down. Make it concrete by saying it aloud: a trillion trillion trillion billion. The electric force pulling the electron onto the proton in a hydrogen atom is a trillion-trillion-trillion-billion times stronger than the gravitational pull between them. If gravity had been the binding force of the atom, the orbit of the electron would have to be unimaginably slow and unimaginably wide to be stable, and the universe of chemistry that builds every plant, animal, and human being you have ever met would simply not exist. The electron clings to the proton because of electricity, and gravity is, on that scale, a rumour. Yet gravity, summed over every atom of an entire planet, produces the force that holds your shoes to the floor of the classroom. The two forces operate at very different scales and very different strengths.

Think Like a Physicist: The ratio $\frac{F_E}{F_G}$ computed in this example is independent of the separation r . Both forces fall off as $\frac{1}{r^2}$, so the r^2 in the denominator cancels out, and the ratio depends only on the constants k ,

G, the two masses, and the elementary charge *e*. This is a powerful style of argument: whenever two inverse-square forces are compared, the geometry drops out and only the source quantities and the natural constants remain.

Three more worked examples now lie behind us, and Coulomb's law has done what every great formula in physics does: it has taken a difference of forty orders of magnitude and packed it into a number you can multiply on a single line. The next section asks what would happen if we removed the second charge from these problems and asked, instead, what the first charge alone is doing to the empty space around it. That question takes us out of force-on-a-charge and into the idea of the electric field, which the rest of this chapter is going to live inside.

ELECTRIC FIELD OF A POINT CHARGE

Coulomb's law tells us the force between two charges. It is, in a sense, a complete description of electrostatics: every phenomenon we shall meet from here to the end of the chapter can be derived from it, in principle, by laborious vector addition over many charges. In practice, however, that vector addition is so tedious that a second, more economical idea has, for two hundred years, been the way working physicists think about electric phenomena. The idea is to forget the second charge for a moment and ask what the first charge, sitting alone, is doing to the empty space around it. The answer, in one sentence, is that *the first charge sets up an electric field everywhere around it, and any other charge that ventures into that field feels a force*. The job of this section is to make the idea of the electric field precise, derive its form for the simplest source, learn to draw it as a picture, and apply it to the most important geometry of this entire chapter: the uniform field between two parallel plates.

From force to field

Imagine that a single source charge *Q* has been fixed in place somewhere in space, and we want to know what it is doing at a particular point *P* nearby. The way to find out is to place a small positive test charge *q* at *P* and measure the force it feels. The force *F* on the test charge is, by Coulomb's law, proportional to *q*: doubling the test charge doubles the force, halving it halves the force, and so on. But the ratio of force to test charge, $\frac{F}{q}$, is independent of how much test charge we used. The ratio depends only on the source *Q* and on where the test point *P* is located. We give that ratio its own name and its own symbol.

The **electric field** at a point *P* is defined as *the electric force per unit positive charge experienced by a small positive test charge placed at that point*. Mathematically, if *F* is the force on a small positive test charge *q* at *P*, the electric field at *P* is:

$$\mathbf{E} = \frac{\mathbf{F}}{q}$$

The electric field is a vector: it has the same direction as the force on a positive test charge, and it has the same direction as *F* divided by the positive scalar *q*. The SI unit of electric field is therefore the unit of force divided by the unit of charge, i.e., the newton per coulomb: 1 N/C. (A second unit, the volt per metre, will turn out to be identical to the newton per coulomb when we meet the electric potential in a later section, but for now we shall use only the newton per coulomb.)

Two cautions about the test charge *q* are worth flagging at the start, because students often stumble over them.

- **The test charge must be small.** If we used a test charge so large that its own field disturbed the source *Q* (by attracting or repelling its charges to new positions), the force we measured would no longer represent the field that *Q* produces on its own. In practice we imagine *q* to be vanishingly small.
- **The test charge must be positive by convention.** A negative test charge would feel a force in the opposite direction, and the field defined as *F*/*q* would have the wrong sign. The convention "small and positive" gives the field its own definite direction, independent of the test charge we happen to drop into it.

The field of a point charge

Apply the new definition to the simplest possible source: a single point charge *Q* sitting alone in a vacuum. A small positive test charge *q* placed at a distance *r* from *Q* feels a Coulomb force whose magnitude we already know:

$$F = \frac{kQq}{r^2}$$

Divide this by q as the definition demands:

$$\mathbf{E} = \frac{F}{q} = \frac{kQ}{r^2},$$

and we have the field of a point charge, the cleanest formula in electrostatics. The magnitude of the field at distance r from a point charge Q falls as the square of the distance, exactly as the Coulomb force did before it. The direction is fixed by the sign of the source: *if Q is positive the field at P points radially outward from Q (away from the source); if Q is negative the field points radially inward, towards the source.*

In a uniform material medium of relative permittivity ϵ_r , the constant k is replaced by $\frac{k}{\epsilon_r}$ (because the medium reduces the electric interaction by the factor ϵ_r), so the field of a point charge in a medium is:

$$\mathbf{E} = \frac{kQ}{\epsilon_r r^2}$$

directed in the same radial direction as in a vacuum. The medium weakens the field but does not change its direction.

Electric field lines

The electric field at every point in space is a vector, with a magnitude and a direction. To draw such a vector at every point of a region of space, we use a device invented by Michael Faraday and refined by James Clerk Maxwell: the **electric field line**. A **field line** is an imaginary curve in space whose tangent, at every point along it, gives the direction of the electric field at that point. A small positive test charge released from rest at any point on a field line would, for an instant, start moving along that line. (Only for an instant, because as soon as the test charge begins to move, it accelerates, gathers speed, and its motion no longer follows the field line. But the tangent rule is precise: the field line shows the direction of the force at every point along it.)

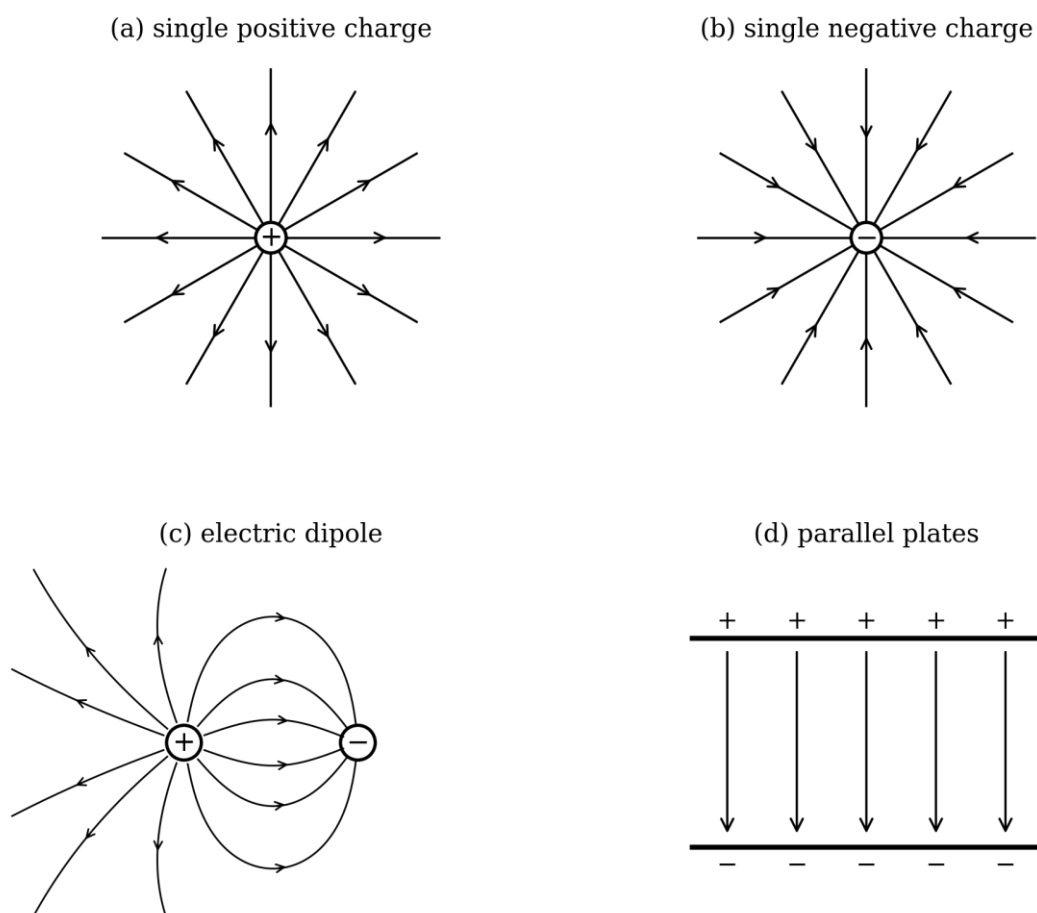


Figure: Patterns of electric field lines for four standard charge arrangements. (a) A single positive point charge: the lines fan out radially in every direction, away from the charge. (b) A single negative point charge: the lines converge radially inward, toward the charge. (c) An electric dipole (one positive charge and one equal negative charge held a fixed distance apart): the lines leave the positive charge, curve through space, and arrive at the negative charge. (d) Two parallel plates carrying opposite uniform charges: the field lines between the plates are straight, parallel, and equally spaced, indicating a uniform field.

Field-line diagrams are not just pretty pictures. They carry real quantitative information, and they obey six rules that examination questions love to test. The six rules follow from the definition of a field line, from the conservation of charge, and from the way charges arrange themselves on conductors. Understand them so that they are remembered without effort.

Rule 1: Field lines begin on positive charges and end on negative charges

A line cannot start in empty space, because there would be no source for the field along it; nor can it end in empty space, for the same reason. Every line must trace a path from a positive charge to a negative one (or, in the limit, to or from infinity).

Rule 2: Two field lines never cross

If they did, the tangent at the crossing point would have two different directions, and the electric field at that point would be undefined.

Rule 3: The lines are denser (closer together) where the field is stronger, and sparser (further apart) where the field is weaker

The density of lines through any small area perpendicular to the field is proportional to the magnitude of E at that area.

Rule 4: A field line meets the surface of a conductor in static equilibrium at right angles

The reason is that the surface of a static conductor is an equipotential surface, and any component of E parallel to it would drive a current along the surface, in contradiction with static equilibrium.

Rule 5: In electrostatics, field lines never form closed loops

A closed loop would mean that a small positive test charge taken around the loop would return to its starting point with a net gain in kinetic energy, in violation of the conservation of energy. (This rule breaks for the magnetic field of a steady current, where closed loops are the rule rather than the exception. Stay tuned for that surprise in a third book in the series.)

Rule 6: The number of lines drawn entering or leaving a charge is, by convention, proportional to the magnitude of the charge

A charge of $+2Q$ is drawn with twice as many lines leaving it as a charge of $+Q$.

A field-line pattern can be made visible in the laboratory with a simple trick due to Faraday himself. Sprinkle small grass-seed or iron-filing particles on a horizontal sheet of paper, then place two charged objects on the paper a few centimetres apart. Each tiny particle, finding itself in the field, polarises (its electrons shift slightly) and becomes a small electric dipole that aligns itself with the local field direction. After a few seconds, the particles trace out the field-line pattern as if it had been drawn by an invisible hand. The figure shows the patterns the experiment reveals for four arrangements that are worth recognising on sight.

The uniform field between two parallel plates

Among the four field-line patterns of the figure in the previous page, one stands apart for both its mathematical simplicity and its industrial importance: the field between two parallel plates carrying opposite uniform charges. The lines between the plates are straight, parallel, and evenly spaced. By rule 3, an evenly spaced bundle of parallel lines means a field of constant magnitude everywhere in the region they fill. The field between the plates is, in a word, uniform.

A **uniform electric field** is one whose magnitude and direction are the same at every point of the region of space it fills. In the gap between two large parallel plates carrying opposite uniform charges, the field is, to an excellent approximation, uniform everywhere away from the plates' edges.

How big is that field? The answer is one of the most useful formulas in the chapter, and we shall derive it properly when Gauss's law arrives in the next section. For the moment, we shall state it and use it. If the two plates are separated by a **perpendicular distance d** and the **potential difference** between them is **V** , then the magnitude of the **uniform electric field** between the plates is:

$$E = \frac{V}{d},$$

directed from the positive plate to the negative plate. The derivation of this result will follow naturally from the definition of potential difference we shall meet shortly, and from the fact that E is constant in the gap. For now, recognise the formula, know the units (volts per metre, which is the same as newtons per coulomb), and notice three things about it.

This result has several important implications.

First, the field depends only on the ratio $\frac{V}{d}$, not on either quantity separately. A potential difference of 100V across a gap of 1cm produces the same E as 1000V across 10cm.

Second, the formula assumes the plates are large enough, and their separation small enough, that the field between them is genuinely uniform. Near the edges of real plates the field bulges outward and the lines are no longer parallel, an effect called fringing that we shall ignore throughout this chapter except in a single advanced remark much later.

Third, $E = V/d$ is the bridge between two ways of describing what is between the plates. Tell someone the plate separation and the voltage across the plates and they can compute the field. Tell them the field and the distance and they can compute the voltage. The two pictures, field-as-vector and voltage-as-scalar, will turn out to describe the same physics from two complementary angles, and the formula $E = V/d$ is the cleanest example of how the two angles meet.

Two worked examples now arrive to check that we can do something with this new field besides admire it from a distance.

BINDER Example 7

A point charge of $+2.0\mu\text{C}$ is placed inside a tank of paraffin oil, whose relative permittivity 5.0. Calculate the magnitude of the electric field at a point P that lies inside the oil at a distance of 10cm from the charge. Take $k = 9.0 \times 10^9 \text{ Nm}^2/\text{C}^2$.

Solution

The field of a point charge inside a dielectric medium of relative permittivity ϵ_r is the vacuum field reduced by the factor ϵ_r :

$$E = \frac{kQ}{\epsilon_r r^2}$$

Substitute the given values, all expressed in SI units:

$$E = \frac{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2)(2.0 \times 10^{-6} \text{ C})}{(5.0)(0.10\text{m})^2} = 3.6 \times 10^5 \text{ N/C}$$

The magnitude of the electric field at P is therefore $3.6 \times 10^5 \text{ N/C}$, directed radially outward from the positive charge.

Making Sense of the Answer: Three hundred sixty thousand newtons per coulomb is, by the standards of school physics, a strong electric field, larger than the field just below a thundercloud about to discharge, and not far below the breakdown field of dry air. A two-microcoulomb charge is the kind of charge a rubbed plastic rod can hold for a few seconds in dry weather, so the field we have just computed is the field a Form Five student could comfortably produce on a laboratory bench. The factor of five in the denominator (the relative permittivity of paraffin) reduces the field by exactly that factor compared with the same charge in air: had the medium been air rather than paraffin, the answer would have been $1.8 \times 10^6 \text{ N/C}$, which is dangerously close to the dry-air breakdown limit we shall meet later.

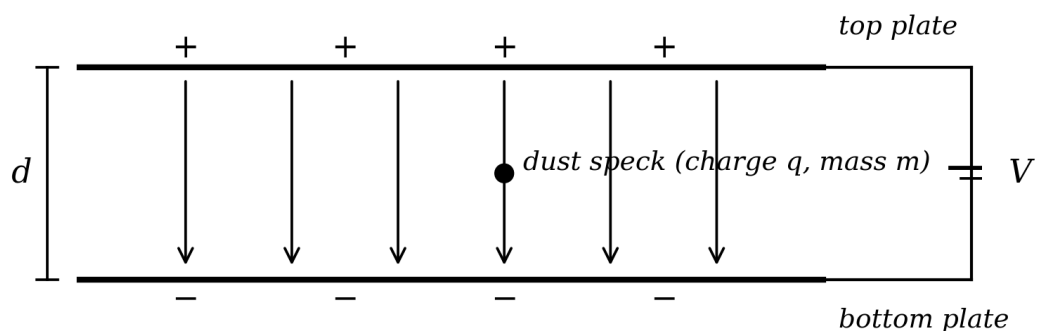
Think Like a Physicist: Whenever a problem asks for an electric field in a medium rather than in vacuum, the only modification to the vacuum formula is the appearance of the relative permittivity ϵ_r in the denominator. No re-derivation is needed; just divide the vacuum result by ϵ_r . The same scaling applies to the Coulomb force between two charges in the same medium, to the field of any collection of point charges in the medium, and to the capacitance of a parallel-plate capacitor with the same dielectric between its plates. One factor of ϵ_r handles them all.

REAL Example 8

In a school physics laboratory, two horizontal metal plates are mounted $d = 4.0\text{cm}$ apart with the upper plate connected to the positive terminal of a $V = 200\text{V}$ battery and the lower plate connected to the negative terminal. A small spherical dust speck of mass $m = 5.0\mu\text{g}$ carries a charge of $q = +1.0\text{nC}$ and finds itself momentarily floating between the plates. Calculate (a) the magnitude of the electric field between the plates, (b) the magnitude of the electric force on the dust speck, and (c) the ratio of the electric force to the weight of the speck. From (c), decide whether the speck will rise, fall, or float. Take $g = 9.8\text{m/s}^2$.

Solution

Consider the following diagram:



a charged parallel-plate capacitor with a small dust speck between the plates

(a) The plates are large enough and close enough together that the field between them is uniform, with magnitude given by the parallel-plate formula:

$$E = \frac{V}{d}$$

Substitute:

$$E = \frac{200\text{V}}{0.040\text{m}} = 5.0 \times 10^3 \text{ V/m}$$

The field between the plates has magnitude $E = 5.0 \times 10^3 \text{ V/m}$ (equivalent to $5.0 \times 10^3 \text{ N/C}$), directed downward from the positive top plate to the negative bottom plate.

(b) The electric force on a charge q sitting in a field E follows from the definition $E = F/q$, rearranged for the force:

$$F = qE$$

Substitute:

$$F = (1.0 \times 10^{-9} \text{ C})(5.0 \times 10^3 \text{ N/C}) = 5.0 \times 10^{-6} \text{ N}$$

The electric force on the dust speck has magnitude $F = 5.0 \times 10^{-6} \text{ N}$, equal to $5.0\mu\text{N}$, directed downward (because the field is downward and the charge is positive).

(c) The weight of the speck is:

$$W = mg = (5.0 \times 10^{-9} \text{ kg})(9.8\text{m/s}^2) = 4.9 \times 10^{-8} \text{ N}$$

And the ratio of the electric force to the weight is:

$$\frac{F}{W} = \frac{5.0 \times 10^{-6} \text{ N}}{4.9 \times 10^{-8} \text{ N}} = 1.02 \times 10^2 = 102$$

The electric force on the speck is about a hundred times its weight, and it points downward in the same direction as the weight. Both forces push the speck downward, so the speck will accelerate downward, not float. It falls to the bottom plate within a fraction of a second.

Making Sense of the Answer: The answer to part (c) is sometimes surprising on first reading. A microscopic dust speck barely visible to the eye weighs almost nothing, yet a hundred times that almost-nothing is still very, very little. The genuine surprise is the ratio of one hundred: an electric field of just five thousand newtons per coulomb (which a Form Five student can produce with two metal plates, a 200V battery, and forty millimetres of air) exerts on a charged speck a force a hundred times stronger than gravity. To make the speck float in equilibrium against gravity, the experimenter would have to reduce its charge by a factor of one hundred (down to 10pC), or weaken the field by the same factor (by turning down the battery to 2V, or moving the plates four metres apart), or arrange for the field to point upward by reversing the polarity of the battery.

Think Like a Physicist: Every uniform-field problem with a small charged object between two parallel plates collapses to the same three equations: $E = V/d$ for the field, $F = qE$ for the electric force, and $W = mg$ for the weight. Once those three numbers are in hand, the question of whether the object rises, falls, or floats reduces to comparing F and W , and to deciding which way F points (which depends on the sign of the charge and the orientation of the plates). The same machinery, with the same three equations, will reappear in the motion of charged particles in fields later in this chapter, in the discussion of the Millikan oil-drop experiment, and in the analysis of charged droplet steering in inkjet printers covered near the end of the chapter.

Two worked examples now lie behind us, and the electric field has earned its keep. We can compute it for a single point charge in any medium, we can draw it as a pattern of lines for any arrangement we can sketch, and we can read off its strength in a parallel-plate gap from a single division. The natural next question is what happens when the charge is no longer a single point sitting on its own, but is smeared along a curve, over a surface, or through a volume. That is exactly the question the next section is about, and the principle of superposition we met in the previous section is going to make the answer beautifully tractable.

ELECTRIC FIELD FROM DISTRIBUTIONS OF CHARGE

A single point charge produces a tidy, radial field that falls as the square of distance from it. Real problems rarely come with a single point charge sitting on its own. They come with two charges at the ends of a thin rod, four charges at the corners of a square, a ring of bracelets-worth of charge, an entire long wire of charge running across a laboratory bench, or, in due course, a charged conducting sphere with all its charge spread across the surface. The job of this section is to take the principle of superposition we met in the previous one and convert it into a tool for finding the field of any reasonable distribution of charges. We will do it three times in three increasingly continuous settings: a finite collection of point charges sitting at the corners of a

square; the continuous distribution of charge around a thin ring; and, as a final case that motivates the deep theorem of the next section, an infinite line of charge.

A collection of point charges

The **principle of superposition for electric fields** states that:

The total electric field at a point due to a collection of source charges is the vector sum of the individual electric fields that each source, taken one at a time, would produce at that point.

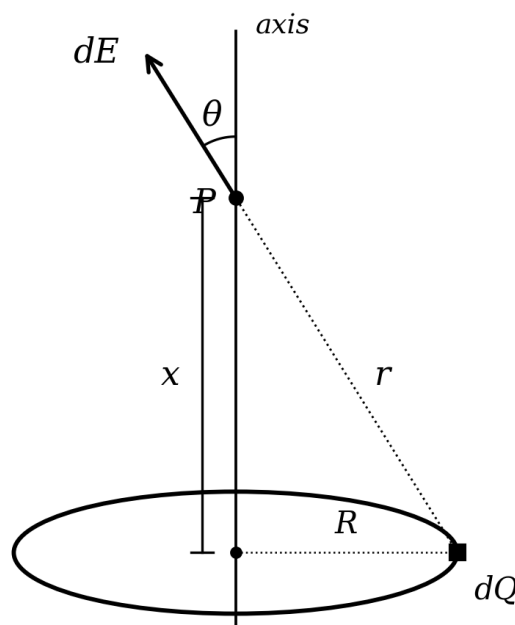
The principle transfers straight from forces to fields, since dividing the total force on a small positive test charge by the test charge simply divides each source-contribution to the force by the same amount. Algebraically, if a small positive test charge q sitting at point P would feel a force F_i from the i th source charge if that source were alone in the universe, then the total electric field at P produced by a whole collection of source charges acting simultaneously is the vector sum of the individual fields:

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 + \cdots$$

Each E_i is computed by treating the corresponding source as if it were alone, using the point-charge formula for the field and pointing the vector along the line from the source to P (radially outward for a positive source, radially inward for a negative one). The arithmetic of the calculation is always the same: write each contribution as a vector, resolve the vectors into components, add the components, recombine into a single resultant vector. In problems with an obvious symmetry, the addition often collapses dramatically because several pairs of components cancel each other, leaving only one or two non-zero terms.

A continuous distribution: the field on the axis of a ring

When the charges are no longer countable on the fingers of one hand but are smeared continuously over a curve, a surface, or a volume, the sum of individual contributions becomes an integral. The simplest such problem with a clean closed-form answer, is the field on the axis of a uniformly charged thin ring. We have a thin ring of radius R , lying flat in a horizontal plane, and the total charge Q is distributed evenly around the ring. We want the electric field at a point P on the axis of the ring, at a perpendicular distance x from the centre.



a uniformly charged ring: geometry for finding E on the axis

Figure: A uniformly charged ring of radius R and total charge Q . The field point P sits on the axis at perpendicular distance x from the ring's centre. A typical charge element dQ on the ring is at distance r from P ; the line dQ -to- P makes an angle θ with the axis. The field dE produced at P by the element points along this line, away from the element for a positive dQ .

Pick any one small piece of the ring, a tiny element of charge dQ . By the point-charge formula, that single element produces a field at P of magnitude:

$$dE = \frac{k dQ}{r^2}$$

Where r is the distance from the element to P. From the figure's geometry, that distance is the hypotenuse of a right triangle whose legs are R and x , so:

$$r = \sqrt{R^2 + x^2}$$

The vector dE points along the line from the element to P, so it has two components: one along the axis (pointing away from the ring towards P), and one perpendicular to the axis (pointing in the plane that contains the element, the axis, and P). If we call the angle between dE and the axis θ , then the axial component of dE is $dE \cos\theta$ and the perpendicular component is $dE \sin\theta$.

Here is where the symmetry of the ring does its job. For every charge element on the ring at the position shown in the figure, there is an exactly equal element on the diametrically opposite side of the ring. That opposite element contributes its own dE , equal in magnitude to the first one but pointing in a direction that is the mirror image of the first across the axis. The axial components of the two add together (both point along the axis in the same direction), while the perpendicular components are equal and opposite, and cancel each other exactly. Pair up every element on the ring with its diametrically opposite partner and the result is the same: all perpendicular components cancel in pairs, and only the axial component survives.

So the field at P is purely axial. Its magnitude is the integral of the axial contribution $dE \cos\theta$ over every element on the ring:

$$E = \oint_{\text{ring}} (dE \cos\theta) = \oint_{\text{ring}} \left(\frac{k dQ}{r^2} \cos\theta \right)$$

To evaluate this integral, parametrise the ring by the angle ϕ that a typical element makes around the centre, with ϕ running from 0 to 2π as we go once around. The arc length of an element subtending an angle $d\phi$ is $R d\phi$. If the linear charge density on the ring is $\lambda = \frac{Q}{2\pi R}$, then the element carries:

$$dQ = \lambda R d\phi$$

From the geometry, every element on the ring is at the same distance $r = \sqrt{R^2 + x^2}$ from P, and every element makes the same angle θ with the axis. From the figure, $\cos\theta = \frac{x}{r}$. Substituting these expressions for dQ and $\cos\theta$ into the integrand:

$$E = \int_0^{2\pi} \left(\frac{k(\lambda R d\phi)}{r^2} \times \frac{x}{r} \right) = \int_0^{2\pi} \left(\frac{k\lambda R x}{r^3} \right) d\phi$$

Now look at the integrand: the factors k , λ , R , x , and r are all constants for a given field point P (they do not depend on ϕ). They pull out of the integral as a common factor, leaving only the integral of $d\phi$:

$$E = \frac{k\lambda R x}{r^3} \int_0^{2\pi} d\phi = \frac{k\lambda R x}{r^3} \times 2\pi = \frac{k(2\pi R\lambda)x}{r^3}$$

But $2\pi R\lambda$ is the total charge on the ring (the linear charge density multiplied by the full circumference), and that is just Q . So our expression collapses to:

$$E = \frac{kQx}{r^3}$$

Substituting $r = \sqrt{R^2 + x^2}$ gives the final result:

$$\mathbf{E}(\mathbf{x}) = \frac{kQx}{(\sqrt{R^2 + x^2})^3} = \frac{kQ\mathbf{x}}{(R^2 + x^2)^{3/2}},$$

directed along the axis, away from the ring (for a positive Q). This is the central formula of the subsection: the magnitude of the electric field at a point on the axis of a uniformly charged ring, at perpendicular distance x from the ring's centre.

Two limiting cases of this formula are worth thinking about before we set numbers into it.

First, at the centre of the ring ($x = 0$), the formula gives $E(0) = 0$. The field is zero at the centre. This may, at first sight, seem surprising: there is charge everywhere around the centre, and a great deal of it close to hand. But each element of charge on the ring has a partner on the opposite side, and the two elements pull a test charge at the centre in exactly opposite directions with exactly equal magnitudes. The pulls cancel pair by pair, and the net field at the centre is zero. Symmetry has done what no amount of arithmetic could: it has predicted the answer without computing a single contribution.

Second, very far from the ring ($x \gg R$), the radius R becomes negligible compared with x , and the formula reduces to $E \approx \frac{kQ}{x^2}$, the field of a point charge Q at distance x . This makes perfect sense: from far enough away, the ring looks like a single point with all of its charge piled at the centre, and the field should reduce to the point-charge form.

Where on the axis is the field largest?

The two limits we have just identified are useful but they do not exhaust the geometry of $E(x)$. The field is zero at $x = 0$, it falls back to zero as $x \rightarrow \infty$, and it is positive in between. By continuity, the field must reach a maximum somewhere on the axis between the centre and infinity. The location of that maximum is fixed by setting the derivative of $E(x)$ to zero:

$$\frac{dE}{dx} = 0$$

Because $E(x)$ is a quotient of two functions of x , the derivative is computed by the quotient rule. Identify the numerator and the denominator:

$$u(x) = kQx, \quad v(x) = (R^2 + x^2)^{3/2}$$

so that $E(x) = u(x)/v(x)$. The quotient rule reads:

$$\frac{dE}{dx} = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2}$$

Compute the two derivatives separately. The numerator derivative is immediate:

$$\frac{du}{dx} = \frac{d}{dx}(kQx) = kQ$$

For the denominator, apply the chain rule. Let $w = R^2 + x^2$, so $v = w^{3/2}$ and $dw/dx = 2x$.

Differentiate $w^{3/2}$ to get $\frac{3}{2}w^{1/2}$, then multiply by $\frac{dw}{dx}$:

$$\frac{dv}{dx} = \frac{3}{2}(R^2 + x^2)^{1/2} \times (2x) = 3x(R^2 + x^2)^{1/2}$$

Substitute these two derivatives into the quotient rule, pulling the common factor kQ out of the numerator:

$$\frac{dE}{dx} = kQ \times \frac{(R^2 + x^2)^{3/2} - 3x^2(R^2 + x^2)^{1/2}}{(R^2 + x^2)^3}$$

Setting this derivative to zero, the denominator is irrelevant (a fraction is zero only when its numerator is), and the numerator must vanish:

$$(R^2 + x^2)^{3/2} = 3x^2(R^2 + x^2)^{1/2}$$

Divide both sides by the common factor $(R^2 + x^2)^{1/2}$ and rearrange:

$$R^2 + x^2 = 3x^2 \Rightarrow R^2 = 2x^2$$

Taking the positive square root gives the location:

$$x = \frac{R}{\sqrt{2}}$$

The field rises from zero at the centre, climbs to its highest value at a distance equal to the radius divided by the square root of two, then falls back to zero in the point-charge limit. The shape of the curve is shown in the following figure.

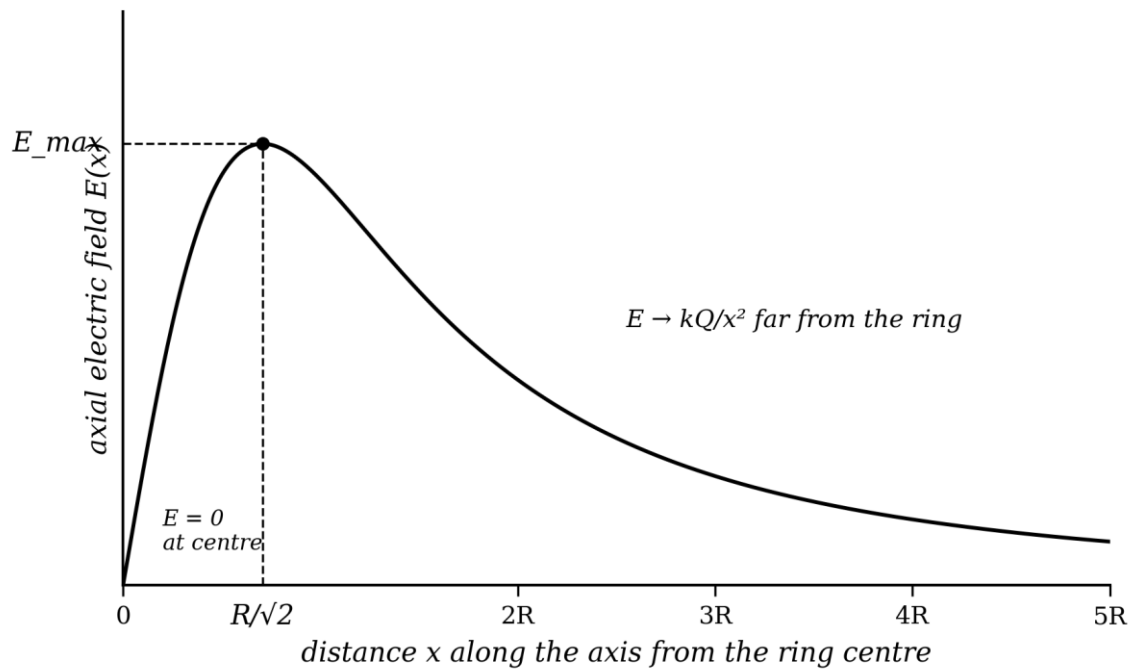


Figure: The axial electric field of a uniformly charged ring as a function of distance x along the axis from the ring centre. The field vanishes at the centre ($x = 0$), rises to a maximum at $x = R/\sqrt{2}$, and decays as $\frac{kQ}{x^2}$ far from the ring, in agreement with the point-charge limit.

Field of a finite line of charge

After the closed loop of the ring, the most natural continuous distribution to study next is a single straight line of charge. Take a thin straight rod of length $2a$, carrying a positive charge spread uniformly along its length with linear charge density λ (we will define it more clearly in the next subsection). Set up the geometry so that the rod lies along the y axis with its midpoint at the origin, and let the field point P sit on the x axis at perpendicular distance x from the midpoint, as the figure shows.

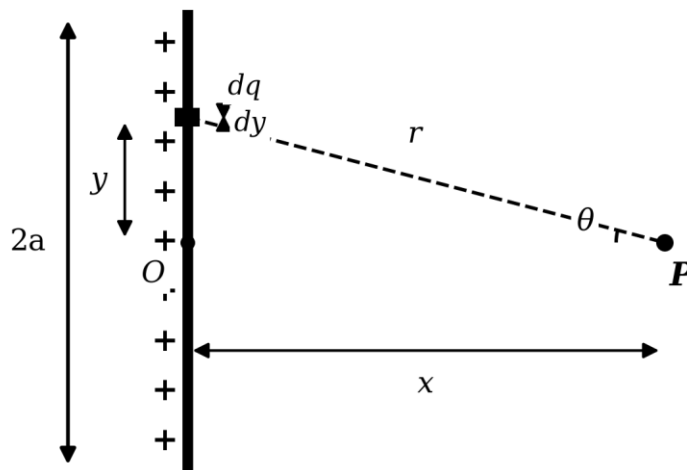


Figure: Geometry for the finite line of charge. A thin straight rod of length $2a$, with linear charge density λ , runs along the y axis. The field point P sits at perpendicular distance x from the midpoint O . A representative element of length dy at height y above the midpoint carries charge $dq = \lambda dy$, sits at distance r from P , and contributes a field dE whose direction is along the line from the element to P .

By the symmetry of the figure, a charge element at height y above the midpoint and the matching element at height $-y$ below it contribute equal magnitudes of dE at P , but with their y components opposite. The two y components cancel in the sum, and only the components along the x axis survive. The x component of the field due to a single element of charge $dq = \lambda dy$ at height y is the Coulomb-law magnitude multiplied by the geometric factor x/r , giving:

$$dE_x = dE \cos\theta = \frac{k dq}{r^2} \times \frac{x}{r} = \frac{k\lambda x dy}{(x^2 + y^2)^{3/2}}$$

With x held fixed and the variable y running over every element of the rod from $-a$ to a , the total field at P is the integral:

$$E_x = k\lambda x \int_{-a}^a \frac{dy}{(x^2 + y^2)^{3/2}}$$

The standard integral evaluates to:

$$\int_{-a}^a \frac{dy}{(x^2 + y^2)^{3/2}} = k\lambda x \left[\frac{y}{x^2 \sqrt{x^2 + y^2}} \right]_{-a}^a = k\lambda x \left(\frac{2a}{x^2 \sqrt{x^2 + a^2}} \right)$$

Therefore the field of the finite line of charge at perpendicular distance x from the midpoint is:

$$E_x = \frac{2k\lambda a}{x\sqrt{x^2 + a^2}}$$

The result behaves correctly at the two natural extremes.

- When the field point sits very close to the rod, x much smaller than a , the square root reduces to a , and the formula collapses to $E = \frac{2k\lambda}{x}$, the field of a uniformly charged wire treated as effectively **infinite**, which is the case the next subsection takes up directly.
- When the field point sits very far from the rod, x much larger than a , the square root is dominated by x , the rod shrinks to a point of total charge $Q = 2\lambda a$, and the formula collapses to $E = \frac{kQ}{x^2}$, the Coulomb field of a point charge.

Between these two extremes the full formula gives the field of any actual finite rod.

An infinite line of charge

One last continuous distribution deserves attention before we close this section, because it both gives a beautiful result and exposes the limits of the brute-force summation method we have been using. Consider an infinite, thin straight wire on which positive charge is distributed uniformly along its length.

The **linear charge density** of a charge distribution along a curve is defined as *the amount of electric charge per unit length of the curve*. We give it the symbol λ , and its SI unit is the coulomb per metre, C/m. If a thin straight piece of wire of length ℓ carries a total charge Q spread uniformly along its length, then:

$$\lambda = \frac{Q}{\ell}$$

The question we want to answer is: *what is the electric field at a point P that sits at a perpendicular distance s from such an infinite line of charge?*

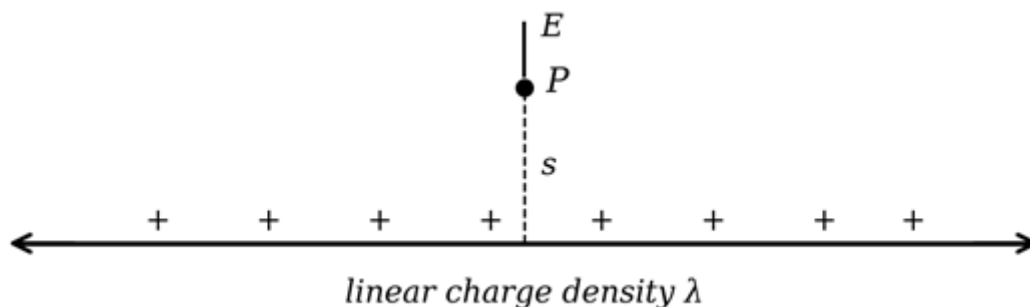


Figure: An infinite straight wire carries a uniform linear charge density λ . The field point P sits at perpendicular distance s from the wire, and by symmetry the field at P points radially outward from the wire, perpendicular to its length.

Set the perpendicular foot of P as the origin and let y be the coordinate of a point on the wire measured from that foot. A small element of the wire of length dy at position y carries charge,

$$dQ = \lambda dy,$$

and lies at a distance,

$$r = \sqrt{s^2 + y^2}$$

from P . The field dE that this element produces at P has magnitude,

$$dE = \frac{k dQ}{r^2} = \frac{k \lambda dy}{s^2 + y^2},$$

and points along the line from the element to P . Let α be the angle this line makes with the perpendicular (the perpendicular itself is the direction of the field we want). The perpendicular component of dE is $dE \cos \alpha$, and the geometry gives $\cos \alpha = \frac{s}{r}$. By symmetry, the components parallel to the wire from the elements at $+y$ and $-y$ cancel exactly, so only the perpendicular component survives in the integral over the whole wire:

$$E = \int_{-\infty}^{+\infty} (dE \cos \alpha) = \int_{-\infty}^{+\infty} \left(\frac{k \lambda dy}{s^2 + y^2} \times \frac{s}{r} \right)$$

Substituting $r = \sqrt{s^2 + y^2}$ into the denominator and using the symmetry of the integrand about $y = 0$ to fold the integral onto the half-line $y = 0$ to $y = +\infty$ (which doubles its value):

$$E = 2k\lambda s \int_0^{+\infty} \frac{dy}{(s^2 + y^2)^{3/2}} = 2k\lambda s \left[\frac{y}{s^2 \sqrt{s^2 + y^2}} \right]_0^{+\infty} = 2k\lambda s \times \frac{1}{s^2} = \frac{2k\lambda}{s}$$

Finally, substituting $k = \frac{1}{4\pi\epsilon_0}$ to write the result in the standard SI form:

$$E = \frac{\lambda}{2\pi\epsilon_0 s},$$

directed radially outward from the wire. Two features of this result deserve underlining.

First, the field falls with the first power of the distance, not the square. The reason is geometric: a line of charge is one-dimensional, so its field spreads out over a circumference (a one-dimensional surface) of length $2\pi s$, rather than over a sphere of area $4\pi s^2$ as a point charge does.

Second, the answer to this problem fell out of an integral that, while standard, took about half a page of calculus to evaluate. It would be nice to have a shortcut, an organising principle that lets us read off the answers to problems with high symmetry without having to set up and evaluate any integrals at all. That shortcut exists, and it is the subject of the next section. The principle is called Gauss's law, and once you have it in hand, you will solve in three lines the problem we have just solved in half a page.

Three worked examples now exercise the machinery just built.

HOT Example 9

Four point charges sit at the corners of a square of side $a = 20\text{cm}$. The two charges at the top corners are each $+Q = +1.0\mu\text{C}$, and the two at the bottom corners are each $-Q = -1.0\mu\text{C}$. Find the magnitude and direction of the electric field at the centre P of the square. Take $k = 9.0 \times 10^9 \text{ Nm}^2/\text{C}^2$.

Solution

Consider the following figure:

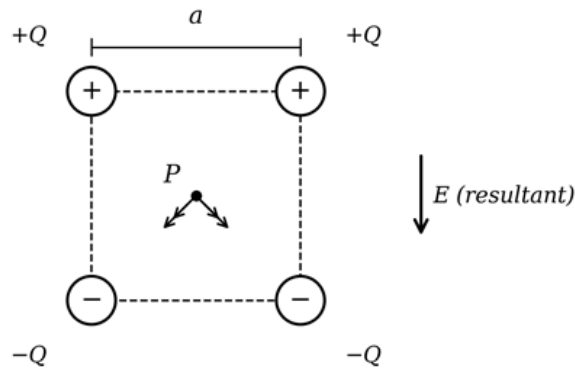


Figure: The four-corner arrangement of HOTS Example 9. The two top charges are $+Q$; the two bottom charges are $-Q$. By symmetry, the horizontal components of the four field contributions at the centre cancel in pairs, and the resultant E points straight down from the $+Q$ row to the $-Q$ row.

The centre of the square is equidistant from all four corners. That distance is half the diagonal of the square:

$$r = \frac{a\sqrt{2}}{2} = \frac{a}{\sqrt{2}}$$

The magnitude of the field at P due to any one of the four charges, taken alone, is given by the point-charge formula:

$$E_1 = \frac{kQ}{r^2} = \frac{kQ}{\frac{a^2}{2}} = \frac{2kQ}{a^2}$$

Each of the four field contributions at P has this magnitude. The directions are fixed by the sign and position of each corner charge. The contribution from each top corner (a positive charge) points away from that corner, i.e., downward and into the square. The contribution from each bottom corner (a negative charge) points towards that corner, i.e., also downward and into the square. Each of the four contributions makes an angle of 45° with the vertical, because the line from any corner to the centre runs along the half-diagonal at this angle.

Now resolve each contribution into horizontal and vertical components. The horizontal components are: from the top-left (positive) charge, rightward (positive x); from the top-right (positive) charge, leftward (negative x); from the bottom-left (negative) charge, leftward (negative x); from the bottom-right (negative) charge, rightward (positive x). Two rightward and two leftward, all of equal magnitude: they cancel in pairs, and the horizontal component of the total field is zero.

The vertical components are all downward: from the two top positive charges, the field at P points away (downward into the square); from the two bottom negative charges, the field at P points towards them (also downward). All four vertical components add. Each has magnitude $E_1 \times \cos 45^\circ$, and the total vertical field is the sum of all four:

$$E = 4 \times E_1 \times \cos 45^\circ = 4 \times \frac{2kQ}{a^2} \times \frac{1}{\sqrt{2}} = \frac{8kQ}{a^2\sqrt{2}} = \frac{4\sqrt{2}kQ}{a^2}$$

Substitute the given values:

$$E = \frac{(4)(1.414)(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2)(1.0 \times 10^{-6} \text{ C})}{(0.20\text{m})^2} = 1.27 \times 10^6 \text{ N/C}$$

The electric field at the centre of the square has magnitude approximately $1.3 \times 10^6 \text{ N/C}$, and it points straight down, from the row of positive charges towards the row of negative charges.

Making Sense of the Answer: A field of just over one million newtons per coulomb sounds enormous on first reading. It is, in fact, the field of a small commercial Van de Graaff generator: the kind of device that can lift the hair of a child standing on the insulating mat and make a fluorescent tube glow when held near. The four-corner arrangement is the smallest collection of point charges that produces a non-zero, easily

computable field at its geometric centre, and the answer $4\sqrt{2} \text{ kQ/a}^2$ is the cleanest combination of the elementary constants and the geometry.

Think Like a Physicist: Before substituting any numbers into a discrete-charge problem, sketch the arrangement on paper, draw the field vector from every source charge to the field point, and inspect the picture for symmetry. In this problem the four contributions to the horizontal component of the field cancelled in pairs by inspection, and the four vertical components added. Symmetry replaced a four-vector addition with the multiplication of one scalar by four. The same collapse occurs in every regular-polygon arrangement of charges, and recognising it before you start computing is the single largest time-saver in this entire chapter.

REAL Example 10

In the school physics laboratory, Mr. Akilikubwa sets a thin wire ring of radius $R = 10\text{cm}$ on the bench and rubs it carefully with a piece of silk until it carries a uniform positive charge of about $Q = 50\text{nC}$. He then asks the class: **what is the electric field at the centre of the ring?** Kipanga and Kipute disagree.

Kipanga: "The field at the centre must be enormous. There is charge everywhere round the ring, only ten centimetres away in every direction. Adding up so many close-by contributions has to give a big answer."

Kipute: "Hold on. There is charge everywhere round the ring, but for every bit of charge on one side of the centre there is an equal bit on the opposite side. The two bits pull a small test charge at the centre in exactly opposite directions, and the pulls have exactly the same size because the two bits are the same distance away. They cancel. Every bit cancels its partner. The total field at the centre is zero."

Kipanga: "Zero? Surrounded by all that charge, the field is zero? That sounds wrong. Kipute, you are over-relying on symmetry."

Mr. Akilikubwa: "Let us settle the matter with the formula we derived a moment ago. The axial field of a uniformly charged ring at distance x from the centre is"

$$E(x) = \frac{kQx}{(R^2 + x^2)^{3/2}}$$

Mr. Akilikubwa: "The centre of the ring is the special case $x = 0$. Substitute that value into the formula and read off the answer."

Kipute carries out the substitution at the chalkboard:

$$E(0) = \frac{kQ \times (0)}{(R^2 + 0)^{3/2}} = \frac{0}{R^3} = 0$$

Kipute: "Zero. Just as I said. The numerator has an x , and when x is zero the whole thing is zero, no matter how big Q or how small R is."

Kipanga: "All right, the formula says zero. But what does the formula know that I do not?"

Mr. Akilikubwa: "It knows that the directions of the contributions matter, not just their sizes. A bit of charge on the right of the ring pushes a test charge at the centre to the left. A bit on the left pushes it to the right. The two pushes are equal in size, and they point in opposite directions, so they add to zero, not to twice one of them. Kipanga, your instinct that the answer should be big was based on adding up only the sizes of the contributions, forgetting their directions. The formula folds direction in for you automatically, through the cosine of the angle the contribution makes with the axis. At the centre, the cosine is zero for every element on the ring, and the axial component of each contribution is zero. There is your answer."

Kipute writes the final answer on the board:

$$E(0) = 0$$

Mr. Akilikubwa: "And the lesson: when a charge distribution is symmetric, trust the symmetry. It will save you the trouble of the calculation, and it will give you the answer with no arithmetic at all."

Making Sense of the Answer: Kipanga's intuition was perfectly reasonable, and it belongs to a long tradition of students for whom a lot of charge close-by should produce a large field. The tradition is half right: a lot of charge close-by does produce a lot of field, contribution by contribution. The other half is that the contributions are vectors, and vectors add by cancellation as easily as they add by reinforcement. The ring at its centre is the cleanest example of perfect cancellation in all of electrostatics: every contribution

has an exact opposite, and the sum collapses to zero. The same principle reappears in capacitor problems, in dipole problems, and in the metallic-cage problems further on in this part of the book.

Think Like a Physicist: Whenever a problem asks for the field at the geometric centre of a symmetric distribution of like charges, the answer is almost certainly zero, and you can write it down without setting up any integrals. The same answer holds at the centre of a uniformly charged sphere, at the centre of a uniformly charged spherical shell, at the centre of a circular disc of uniform charge, at the centre of a regular polygon of equal point charges, and at the centre of a uniformly charged cube. Symmetry is the cheapest shortcut in physics: it costs nothing to use, and it replaces an integral with the single word "zero".

HOT Example 11

A thin ring of radius 10cm carries a uniformly distributed positive charge of 50nC. Find (a) the location along the axis where the electric field is greatest, and (b) the magnitude of the field at that point.

Solution

(a) The axial field of the ring is:

$$E(x) = \frac{kQx}{(R^2 + x^2)^{3/2}}$$

The field is largest at the axial distance:

$$x = \frac{R}{\sqrt{2}}$$

With $R = 0.10\text{m}$, this places the maximum at:

$$x = \frac{0.10\text{m}}{\sqrt{2}} = 0.0707\text{m} = 7.1\text{cm}$$

(b) Substituting $x = R/\sqrt{2}$ into the expression for $E(x)$:

$$E_{\max} = \frac{kQ \times \frac{R}{\sqrt{2}}}{\left(R^2 + \frac{R^2}{2}\right)^{3/2}} = \frac{\frac{kQR}{\sqrt{2}}}{\frac{3\sqrt{3}}{2\sqrt{2}}R^3} = \frac{2kQ}{3\sqrt{3}R^2}$$

Substituting numerical values, with $k = 9.0 \times 10^9 \text{ Nm}^2/\text{C}^2$, $Q = 50 \times 10^{-9} \text{ C}$, and $R = 0.10\text{m}$:

$$E_{\max} = \frac{2 (9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) (50 \times 10^{-9} \text{ C})}{3\sqrt{3}(0.10\text{m})^2} = 1.73 \times 10^4 \text{ N/C}$$

The maximum axial field has magnitude approximately $1.7 \times 10^4 \text{ N/C}$, located on the axis at $x = 7.1\text{cm}$ from the centre of the ring.

Making Sense of the Answer: A field of seventeen thousand newtons per coulomb is in the normal range for charged objects of laboratory size. The location of the maximum, at $R/\sqrt{2}$, is one of the small geometric facts that physics rewards students for remembering.

Think Like a Physicist: When a function vanishes at both ends of an interval and is positive in between, it must have a maximum, and you locate it by setting the derivative to zero.

Three worked examples now lie behind us, and the superposition principle has done its job on three increasingly continuous geometries. The next section gives us a powerful new tool, Gauss's law, that will reduce the infinite-line problem to three lines instead of a page.

GAUSS'S LAW AND FOUR CANONICAL APPLICATIONS

Last lesson finished with three integrals stacked one on top of the other. A ring of charge took us through a circle. An infinite line took us through the entire real number line. A square arrangement asked for four vector additions and a Pythagorean theorem before it agreed to give up its answer. The work was honest, but it was work. A clever person, watching the same problems for the third time, begins to suspect that there must be a shortcut hiding underneath. There is. Its name is **Gauss's law**, and the rest of this section is the story of how it cuts a class of hard electrostatic problems down to a single line of algebra.

Kipanga, Kipute, and Mr. Akilikubwa are once again on the veranda. Kipanga is balancing a textbook on his head, partly out of mischief and partly because his cousin once told him that physicists develop better posture than other people. Kipute is sketching a sphere on her notebook. Mr. Akilikubwa is, as always, watching them both with the patient amusement of someone who has seen this moment many times before.

Kipute: Sir, the ring example showed me how to do the integral, but I will not be honest if I claim the calculation felt natural. Is there a way to get the same answer without all the trigonometry?

Mr. Akilikubwa: For a ring on its axis, no. The geometry asks for an integral and you must do the integral. But for a charged sphere, an infinite plane, or an infinite line, there is a beautiful counting argument that delivers the answer in one line. The counting argument is called Gauss's law. Before we can state it, however, we need one new word: flux.

Kipanga: Flux. That sounds like a disease.

Mr. Akilikubwa: It is in fact the Latin word for flow. We will use it to count field lines crossing a surface.

Electric flux

Imagine a flat surface of area A placed inside a uniform electric field of magnitude E . Suppose first that the surface is held perpendicular to the field, with the field lines piercing it head-on. The number of field lines crossing the surface is proportional to the product EA . That product is called the **electric flux** through the surface, written with the Greek capital letter Φ (phi).

$$\Phi = EA$$

The **electric flux** through a flat surface placed perpendicular to a uniform electric field is the product of the field magnitude and the surface area. It is a **scalar quantity**, and it measures the number of field lines that cross the surface. The SI unit of flux follows directly from the formula: newton metre-squared per coulomb, Nm^2/C .

Now tilt the surface so that its normal makes an angle θ with the field. The same number of field lines still cross the surface, but if we project the tilted surface onto a plane perpendicular to the field, we get a smaller projected area equal to $A \cos \theta$. The flux therefore picks up a cosine factor:

$$\Phi = EA \cos \theta$$

The figure shows the geometry: the flat surface of area A , the unit normal $\hat{n}$ at the centre of the surface, and the angle θ between the field and the normal. When $\theta = 0^\circ$ the surface is perpendicular to the field and the flux is at its maximum value EA . When $\theta = 90^\circ$ the surface is parallel to the field, no field lines cross it, and the flux is zero.

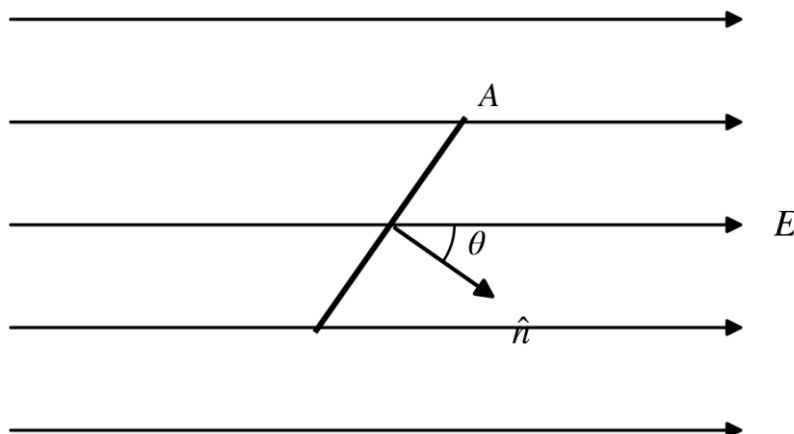


Figure: Electric flux through a flat surface of area A tilted so that its unit normal $\hat{n}$ makes angle θ with the uniform electric field E .

A real surface need not be flat, and a real field need not be uniform. To handle the general case we break the surface into many tiny patches, each so small that the surface is essentially flat across the patch and the field is essentially uniform across it. Let dA denote the area of one such patch, and let its outward normal make angle θ with the local field. The patch contributes a small flux $d\Phi = E dA \cos \theta$. Summing the contributions of all the patches is, in the limit of infinitely small patches, a surface integral:

$$\Phi = \int E dA \cos \theta$$

If the surface is closed (an imaginary balloon with no hole in it), the integral is written with a small circle through the integral sign to remind us of the closure, and the convention is that the normal always points outward:

$$\Phi = \oint E dA \cos \theta$$

The electric flux through a surface is the surface integral of the component of the electric field along the outward normal. **For a closed surface**, the integral counts the net number of field lines leaving the surface. Lines that enter and then leave the surface contribute zero to the total, because their entry counts negatively and their exit counts positively, and the two cancel. This last property is the heart of the law we are about to state.

A word about field lines and space: A point charge sends its field lines out into space isotropically, meaning equally in every direction. The total angle of space around any point is 4π steradians (the three-dimensional analogue of the 2π radians around a point in a plane). The field lines therefore spread their entire share through that total solid angle. Now imagine wrapping any closed surface, of any shape, around the point charge. Every line that leaves the charge must, before it escapes to infinity, pierce that surface exactly once on its way out. The shape of the surface does not matter, only whether the charge is enclosed by it or not. This piece of pure geometry is what makes the law we are about to state independent of the shape of the surface we choose.

Gauss's law

The German mathematician Carl Friedrich Gauss, working in the early 1800s on problems in geodesy and electromagnetism, proved a beautiful result that is now one of the four equations of classical electromagnetism. In the language of electric flux it reads:

$$\oint \mathbf{E} dA \cos \theta = \frac{Q_{\text{enc}}}{\epsilon_0}$$

In words, **Gauss's law** states that: *The net electric flux through any closed surface is equal to the total charge enclosed by the surface divided by the permittivity of free space.* The closed surface is called a **Gaussian surface**. It is imaginary; you draw it in space wherever it is convenient. You may choose any shape that makes the integral easy to evaluate, and the right-hand side will still equal $\frac{Q_{\text{enc}}}{\epsilon_0}$ no matter what shape you pick, because the shape contributes only its enclosed charge to the answer.

Kipanga: *A surface I get to choose. Anywhere. Any shape. That feels too good to be true.*

Mr. Akilikubwa: *It is true, and like every gift in physics it comes with one condition. The right-hand side is always $\frac{Q_{\text{enc}}}{\epsilon_0}$, regardless of shape; but the left-hand side, the integral, is only easy to evaluate when the shape you choose matches the symmetry of the charge distribution. For a single point charge you should choose a sphere centred on the charge. For an infinite line of charge you should choose a cylinder coaxial with the line. For an infinite plane you should choose a small pill-box piercing the plane. The art of using Gauss's law is choosing the right Gaussian surface.*

From Coulomb's law to Gauss's law

Gauss's law is not an extra postulate; it is a consequence of Coulomb's law and the geometry of three-dimensional space. To see this, place a single point charge Q at the centre of an imaginary sphere of radius r , and compute the flux of its electric field through that sphere. The geometry is shown in the figure.

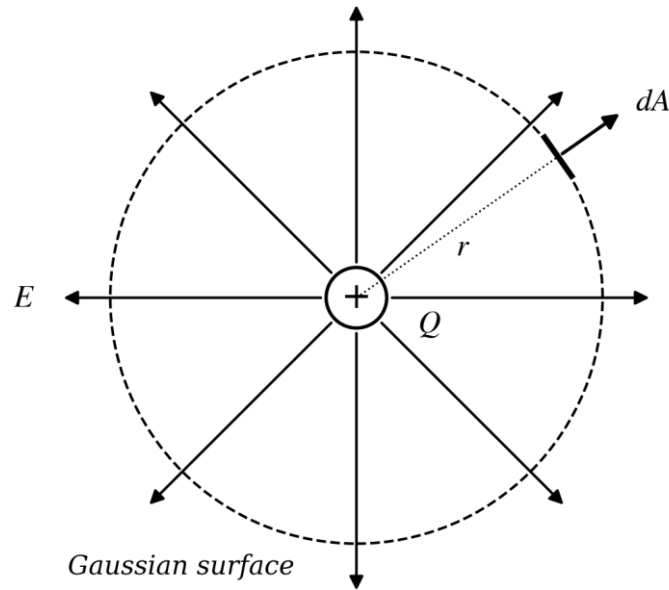


Figure: A Gaussian sphere of radius r around a point charge Q . The radial field E is parallel to the outward area vector dA at every patch of the sphere, so $\cos \theta = 1$ everywhere on the surface.

By the symmetry of the situation, the electric field on the sphere has the same magnitude everywhere on the surface, and it points radially outward at every point. Coulomb's law gives its magnitude:

$$E = \frac{kQ}{r^2}$$

The outward normal dA at every patch of the sphere is parallel to the local field, so the angle θ between them is zero everywhere, and $\cos \theta = 1$. The flux integral therefore simplifies to a constant E multiplied by the surface area of the sphere:

$$\Phi = \oint E \, dA \cos \theta = E \oint dA = E \times 4\pi r^2$$

Substitute the Coulomb expression for E :

$$\Phi = \frac{kQ}{r^2} \times 4\pi r^2 = 4\pi kQ$$

The radius r has dropped out. Now use $k = \frac{1}{4\pi\epsilon_0}$ to replace k :

$$\Phi = 4\pi \times \frac{Q}{4\pi\epsilon_0} = \frac{Q}{\epsilon_0}$$

The factor 4π inside the permittivity exactly cancels the 4π from the sphere's surface area, and we are left with the enclosed charge divided by ϵ_0 , which is the right-hand side of Gauss's law. The geometric reason for the cancellation is the one already given in the solid-angle paragraph: the field lines of a point charge spread through 4π steradians, and ϵ_0 was historically defined with a 4π built into it so that the cancellation would happen cleanly.

From Gauss's law back to Coulomb's law

The argument runs equally well in reverse. Take Gauss's law as the starting point, place a single point charge Q alone in space, and surround it with an imaginary sphere of radius r centred on the charge. By symmetry the field on the sphere has the same magnitude E everywhere on the sphere, and points radially. The flux integral simplifies as before to $E \times 4\pi r^2$, and Gauss's law sets this equal to $\frac{Q}{\epsilon_0}$:

$$E \times 4\pi r^2 = \frac{Q}{\epsilon_0}$$

Make E the subject:

$$E = \frac{Q}{4\pi\epsilon_0 r^2} = \frac{kQ}{r^2},$$

which is Coulomb's law for the field of a point charge. The two laws are therefore equivalent: knowing one of them, you can derive the other. They look different because they encode the same physics from different starting points. Coulomb's law is a statement about pairs of charges and the force between them; Gauss's law is a statement about flux through closed surfaces. The next several pages will show why the flux version, when the geometry is symmetric, is the easier tool to reach for.

The flux recipe is laid out. One worked example arrives now to test the cosine factor before the four canonical applications of Gauss's law begin to fall, one by one, in the rest of the section.

BINDER Example 12

A flat circular surface of radius 20cm is placed inside a uniform electric field of magnitude 500N/C. The unit normal of the surface makes an angle of 60° with the field. Calculate the electric flux through the surface.

Solution

Let A be the area of the flat surface, E the magnitude of the uniform field, and θ the angle between the normal of the surface and the field. The geometry of the problem is shown in the figure.

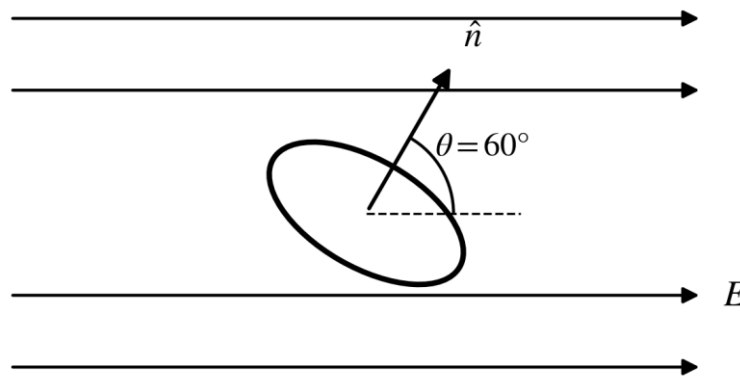


Figure: A flat circular surface tilted so that its unit normal $\hat{n}$ makes an angle of 60° with the uniform electric field E .

The area of a circular surface of radius r is $A = \pi r^2$. The electric flux through a flat surface in a uniform field is given by:

$$\Phi = EA \cos \theta = E\pi r^2 \cos \theta$$

Substitute the data:

$$\Phi = (500\text{N/C}) \times \pi \times (0.20\text{m})^2 \times \cos 60^\circ = 31.4 \text{ Nm}^2/\text{C}$$

The electric flux through the surface is $\Phi = 31.4 \text{ Nm}^2/\text{C}$

Making Sense of the Answer: A quick sanity check. If the surface were perpendicular to the field ($\theta = 0^\circ$), the flux would be $EA = (500\text{N/C}) \times \pi \times 0.040\text{m}^2 = 20\pi \text{ Nm}^2/\text{C} = 62.8 \text{ Nm}^2/\text{C}$. If it were parallel to the field ($\theta = 90^\circ$), no field lines would cross it and the flux would be zero. Our answer of $31.4 \text{ Nm}^2/\text{C}$ is exactly half of the maximum, which matches $\cos 60^\circ = 0.5$ to within the precision of the arithmetic. The cosine is doing the geometric work for us.

Think Like a Physicist: Flux is a scalar that counts the net number of field lines crossing a surface, weighted by how perpendicular the surface is to those lines. The cosine factor is not a physical factor (nothing changes about the field or the surface), it is a geometric factor that accounts for the projection of the surface onto a plane perpendicular to the field. Once you have this idea in your hand, you will recognise

it again in fluid flow, heat flow, and even in optics, because the same picture underlies every conservation law that talks about something flowing across a boundary.

The flux recipe and Gauss's law are now in hand. Four canonical applications will occupy the next several pages. In each case the trick is the same: pick a Gaussian surface whose shape matches the symmetry of the charge distribution, observe that the field is constant on that surface, and read the magnitude of the field straight off the integral. The first application is the most fundamental one of all, because the charge distributions used in many later examples are built out of spherical pieces: a uniformly charged sphere.

1. A uniformly charged sphere

The uniformly charged sphere may be insulating or conducting.

A uniformly charged insulating sphere

Consider a solid insulator (for example a ball of plastic, or a hard rubber sphere) of radius R , carrying a total charge Q spread evenly throughout its volume. By **volume charge density** we mean the *charge per unit volume*:

$$\rho = \frac{Q}{\frac{4}{3}\pi R^3} = \frac{3Q}{4\pi R^3}$$

For a uniformly charged sphere of radius R and total charge Q , the volume charge density is the same number ρ everywhere inside the sphere, and zero outside. To find the electric field at a point at distance r from the centre, draw an imaginary spherical Gaussian surface of radius r centred on the charged sphere. By the spherical symmetry of the problem, the magnitude of the field is the same everywhere on the Gaussian surface and points radially. The flux integral therefore simplifies as in the point-charge derivation:

$$\Phi = E \times 4\pi r^2 = \frac{Q_{\text{enc}}}{\epsilon_0}$$

Two cases follow, depending on whether the Gaussian surface lies outside the charged sphere or inside it. The geometry is shown in the figure.

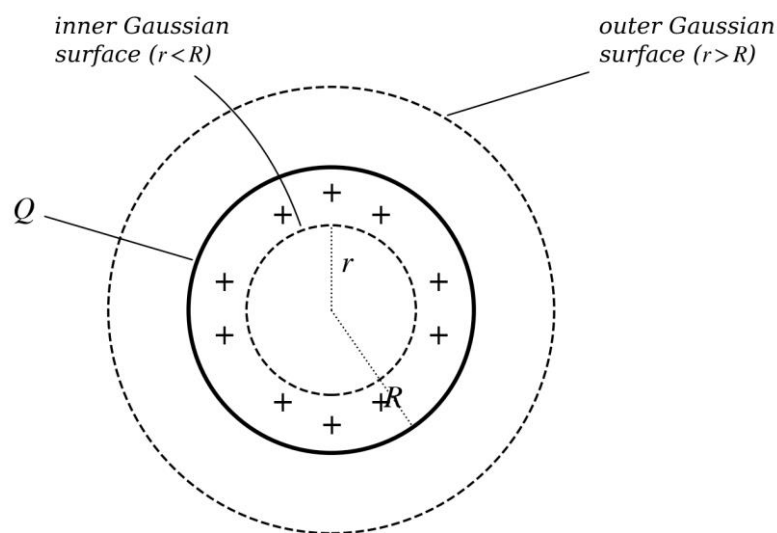


Figure: A uniformly charged insulating sphere of radius R and total charge Q , with an inner Gaussian surface ($r < R$) and an outer Gaussian surface ($r > R$). The + signs indicate that the charge is spread throughout the volume of the sphere.

Outside the sphere ($r > R$): When the Gaussian surface lies outside the charged sphere, it encloses the entire charge Q . Gauss's law gives:

$$E \times 4\pi r^2 = \frac{Q}{\epsilon_0}$$

Make E the subject:

$$E = \frac{Q}{4\pi\epsilon_0 r^2} = \frac{kQ}{r^2}$$

Outside the sphere, the field is exactly the same as if the entire charge Q were concentrated at the centre. This is the spherical analogue of Newton's shell theorem for gravitation, which we met in **Foundations of Physics**: a spherically symmetric shell or ball of mass attracts an external test mass as though all of its mass sat at its centre.

Inside the sphere ($r < R$): When the Gaussian surface lies inside the charged sphere, it encloses only the portion of the charge contained in the smaller sphere of radius r. That charge equals the volume charge density ρ multiplied by the volume of the inner sphere:

$$Q_{\text{enc}} = \rho \times \frac{4}{3}\pi r^3 = \frac{3Q}{4\pi R^3} \times \frac{4}{3}\pi r^3 = Q \frac{r^3}{R^3}$$

The factors of $4\pi/3$ have cancelled, leaving a clean cube ratio. Substitute this enclosed charge into Gauss's law:

$$E \times 4\pi r^2 = \frac{Q \frac{r^3}{R^3}}{\epsilon_0}$$

Make E the subject and simplify:

$$\mathbf{E} = \frac{Qr}{4\pi\epsilon_0 R^3} = \frac{kQr}{R^3}$$

Inside the sphere, the field grows **linearly** with the distance from the centre. It is zero at the centre ($r = 0$), and rises to its peak value $\frac{kQ}{R^2}$ at the surface ($r = R$). Outside the sphere, the field then falls as $\frac{1}{r^2}$. The same value at the surface is approached from both sides, so the field is continuous there. (We will see in a moment that the conducting case is different: it has a jump discontinuity at the surface.)

A uniformly charged conducting sphere

Replace the insulator with a metal sphere of the same radius R and the same total charge Q. A metal contains a sea of free electrons that can move across the bulk of the material. In electrostatic equilibrium (charges no longer moving), the electric field inside the metal must be zero, because if it were not zero the free electrons would feel a force and continue to drift, contradicting equilibrium. The verbal statement of this foundational result is short and worth noting cleanly:

Inside a conductor in electrostatic equilibrium, the electric field is zero everywhere. As a consequence, any net charge placed on an isolated conductor resides entirely on its outer surface; there can be no net charge in the interior bulk.

The figure shows the geometry: the same Gaussian surfaces as before, but now the charge Q lives on the rim of the sphere instead of being distributed through its volume.

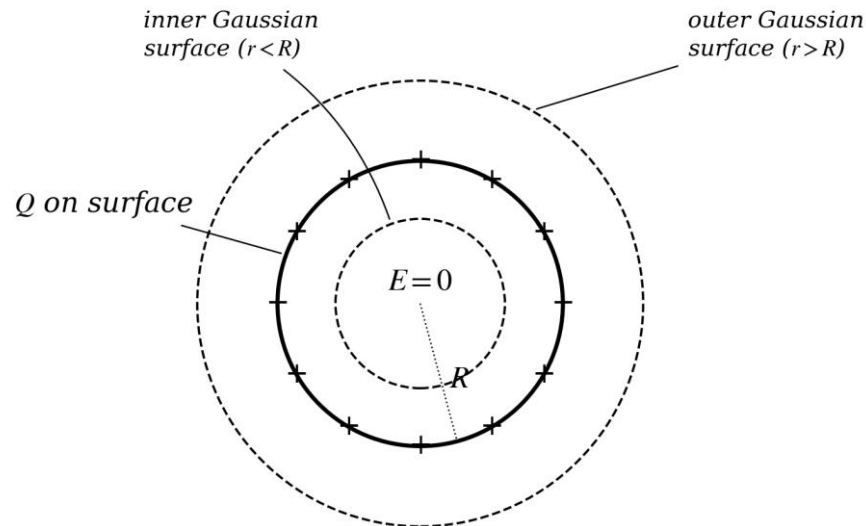


Figure: A uniformly charged conducting sphere of radius R with total charge Q . The charge resides entirely on the outer surface (shown by the $+$ signs along the rim), and the electric field is zero everywhere inside.

Inside the conductor ($r < R$): The inner Gaussian surface encloses no charge at all, because the conductor's charge sits on the outer surface. Gauss's law therefore reads:

$$E \times 4\pi r^2 = \frac{0}{\epsilon_0} = 0,$$

which gives $E = 0$ everywhere inside the conductor. This is the Gauss's-law derivation of the same result we just stated verbally: the interior of a conductor in electrostatic equilibrium is electric field-free.

Outside the conductor ($r > R$): The outer Gaussian surface encloses the full charge Q (all of which sits on the conductor's outer surface). Gauss's law reads exactly as for the insulating sphere:

$$E \times 4\pi r^2 = \frac{Q}{\epsilon_0},$$

and the answer is the same:

$$E = \frac{kQ}{r^2}$$

Outside the conductor, the field is identical to that of either a point charge Q at the centre or the uniformly charged insulating sphere we studied a moment ago. From a distance, you cannot tell a metal sphere with charge Q from a plastic sphere with charge Q from a point charge Q : all three look the same to the field.

Insulating and conducting spheres, side by side

Outside the sphere, the two cases give the same answer. The difference lives inside. The figure plots the field magnitude as a function of distance r from the centre, for both spheres on the same axes.

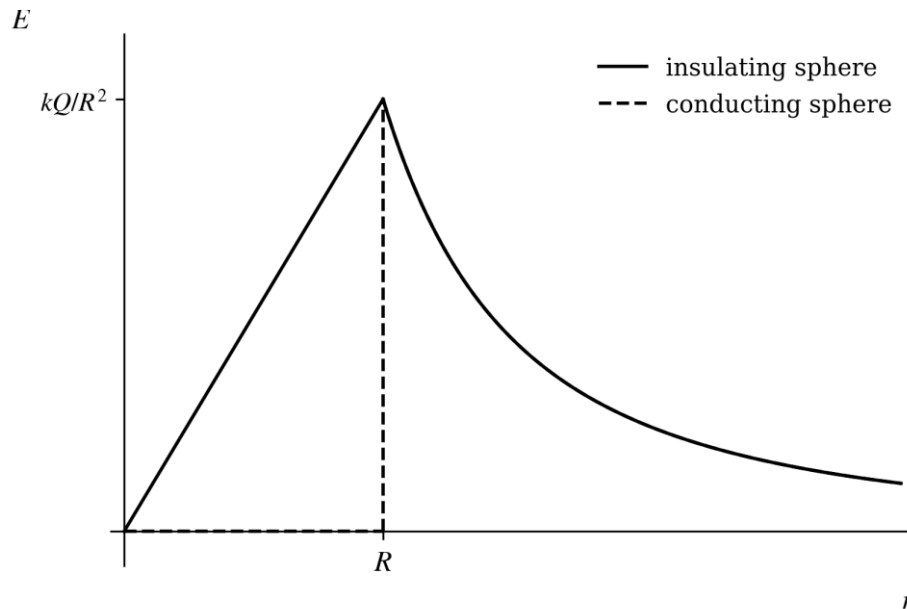


Figure: Electric field magnitude E as a function of distance r from the centre, for a uniformly charged insulating sphere (solid line) and a conducting sphere of the same radius and charge (dashed line). Outside the sphere both curves coincide; inside, the insulating case grows linearly while the conducting case stays at zero.

Two observations are worth committing to memory.

- The two curves are identical for $r > R$, so from outside the sphere we cannot detect by an electric measurement alone whether the body is an insulator or a conductor.
- The conducting case has a step discontinuity at the surface: the field jumps from zero just inside to $\frac{kQ}{R^2}$ just outside. The insulating case is smooth across the surface, because in that case the charge is spread through the volume rather than concentrated on a thin shell.

Kipute: Sir, in the conducting case, where is the charge? You said it sits on the surface, but a surface has no thickness. So how much charge sits in a tiny patch of the surface?

Mr. Akilikubwa: A very good question. The natural quantity is the surface charge density, written with the Greek letter σ (sigma). For a conducting sphere of total charge Q and radius R , the charge is distributed evenly over the spherical surface of area $4\pi R^2$, so $\sigma = \frac{Q}{4\pi R^2}$. The field just outside the surface can then be written as $E = \frac{\sigma}{\epsilon_0}$. We will derive this last result formally in the next application, when we look at an infinite plane sheet.

The insulating-sphere result has two faces: the friendly $\frac{kQ}{r^2}$ on the outside, and the surprising linear-in- r face on the inside. The next example tests both faces in a single calculation, so that the contrast between them is visible in one place.

BINDER Example 13

An insulating sphere of radius 6.0cm carries a total charge of 3.0 μ C uniformly distributed throughout its volume. Calculate the magnitude of the electric field at a point 3.0cm from the centre, and at a point 9.0cm from the centre. Take the Coulomb constant to be $9.0 \times 10^9 \text{ Nm}^2/\text{C}^2$.

Solution

Let Q be the total charge of the sphere, R its radius, and r the distance from the centre to the field point. The geometry is shown in the figure: one field point lies inside the sphere ($r < R$), the other lies outside ($r > R$).

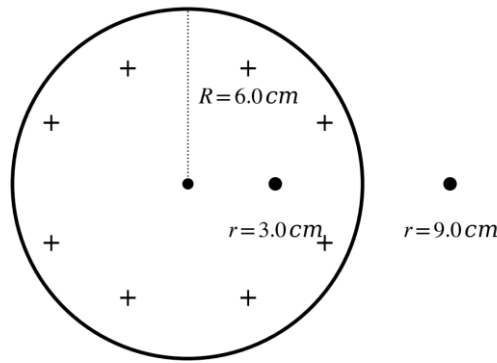


Figure: Example 13: an insulating sphere of radius 6.0 cm carrying a total charge $3.0 \mu\text{C}$. The field is wanted at the two indicated points.

Field at the inside point: Here $r = 3.0\text{cm} = 0.030\text{m}$ is less than R , so the Gauss's-law result for the interior of a uniformly charged sphere applies:

$$E_{\text{in}} = \frac{kQr}{R^3}$$

Substitute the numerical values:

$$E_{\text{in}} = \frac{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times (3.0 \times 10^{-6} \text{ C}) \times (0.030 \text{ m})}{(0.060 \text{ m})^3} = 3.75 \times 10^6 \text{ N/C}$$

Field at the outside point: Here $r = 9.0\text{cm} = 0.090\text{m}$ is greater than R , so the outside-the-sphere result applies, and the field is identical to that of a point charge Q at the centre:

$$E_{\text{out}} = \frac{kQ}{r^2}$$

Substitute the data:

$$E_{\text{out}} = \frac{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times (3.0 \times 10^{-6} \text{ C})}{(0.090 \text{ m})^2} = 3.33 \times 10^6 \text{ N/C}$$

The field at the inside point is $3.8 \times 10^6 \text{ N/C}$, and the field at the outside point is $3.3 \times 10^6 \text{ N/C}$.

The inside field is the larger of the two.

Making Sense of the Answer: A moment of surprise is fair here. Naively, you might expect that getting closer to the centre always means a stronger field, because point-charge intuition suggests that the field should blow up as $1/r^2$ near the source. For a uniformly charged sphere, however, the field reaches its maximum at the surface, then falls off in both directions: linearly toward zero as you move toward the centre, and as $1/r^2$ as you move outward. The two evaluation points in this example sit at the same fractional distance on either side of the surface, $r/R = 0.5$ inside and $r/R = 1.5$ outside. The inside field is kQ/R^2 multiplied by $r/R = 0.5$, and the outside field is the same surface value multiplied by $(R/r)^2 = 1/2.25 \approx 0.44$. The inside number, 0.5, wins by a small margin over the outside number, 0.44, which is why the inside field came out slightly larger than the outside one.

Think Like a Physicist: The inside-versus-outside structure of the uniformly charged sphere is the spherical electrical twin of the Earth-as-ball gravitational problem from **Foundations of Physics**: inside a uniformly dense planet, the gravitational field grows linearly from zero at the centre to its peak value at the surface, then falls as $1/r^2$ outside. The two are not coincidences; they are two instances of the same theorem, which Newton proved for gravity in the seventeenth century and which Gauss proved in greater generality two centuries later. The inverse-square law and the geometry of three-dimensional space conspire to give a beautiful result that the field outside a spherically symmetric distribution depends only on the total amount enclosed, never on how that amount is arranged.

A uniformly charged hollow conducting sphere

The two spheres treated so far were solid all the way through. The case in which the metal is hollowed out, leaving an empty cavity inside an otherwise solid metal shell, is no less common, and far from being a curiosity it turns out to lie behind one of the most useful protective devices in all of electrical engineering. Take a hollow conducting sphere with inner radius R_1 and outer radius R_2 , and let it carry a total positive charge Q deposited on it from outside, the whole arrangement left in electrostatic equilibrium.

The very first question is where the charge sits. The argument is the same as for the solid conductor and is worth repeating because it now does extra work. Inside the metal of any conductor in equilibrium the electric field must be zero, for if it were not, the free electrons of the metal would feel a force and the conductor would not yet be in equilibrium. Gauss's law applied to any closed surface lying wholly inside the metal then says that the charge enclosed must be zero too, since zero field gives zero flux. The total charge Q therefore cannot reside anywhere within the body of the metal; it can only sit on a surface. The metal of a hollow conductor has two surfaces, an inner one of radius R_1 facing the cavity and an outer one of radius R_2 facing the world. For a hollow conductor with no charge inside the cavity, all the charge migrates to the outer surface, and the inner surface carries none. This is the picture we work with.

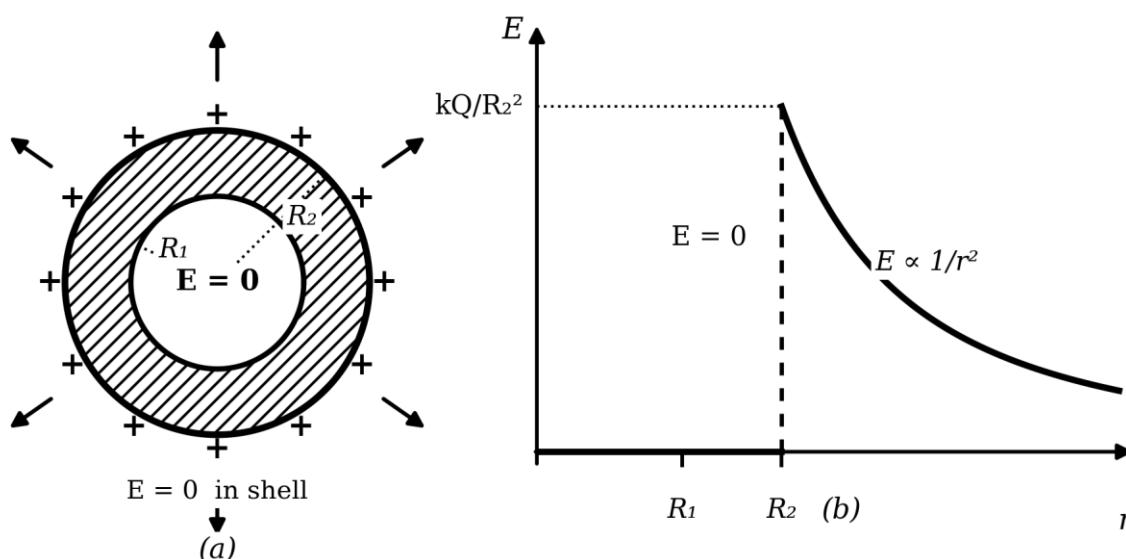


Figure: A uniformly charged hollow conducting sphere. Panel (a): the conductor has inner radius R_1 and outer radius R_2 . Charge sits entirely on the outer surface. The field is zero both in the cavity and within the metal of the shell. Outside the sphere the field is the same as if all the charge were concentrated at the centre. Panel (b): the field strength against distance. Zero everywhere inside, the field rises abruptly to its peak value kQ/R_2^2 at the outer surface, then falls as the inverse square of distance.

With the charge sat down where it can be reached, Gauss's law is applied three times in turn, one Gaussian sphere for each of the three regions a radial coordinate can land in.

First region: Take first the region outside the conductor, $r > R_2$. Draw a Gaussian sphere of radius r concentric with the conductor; it encloses the whole of the charge, Q . By the spherical symmetry of the problem the field is radial and has the same magnitude at every point of the Gaussian sphere, so the flux integral simplifies at once and Gauss's law reads:

$$E \times 4\pi r^2 = \frac{Q}{\epsilon_0}$$

Solving for the field strength:

$$E = \frac{Q}{4\pi\epsilon_0 r^2} = \frac{kQ}{r^2}$$

Outside the hollow conductor the field is exactly the field of a point charge Q sitting at the centre. The hollowness makes no difference to the world beyond the outer surface, which sees only a total charge and the distance to it.

Second region: Take next the region within the metal of the shell itself, $R_1 < r < R_2$. A Gaussian sphere of radius r drawn in this range lies entirely within the body of the metal, where the field is zero by the

argument that opened the subsection. Zero field gives zero flux, and the calculation is complete in a single line: $\mathbf{E} = \mathbf{0}$ everywhere within the metal.

Third region: Take finally the region inside the cavity, $r < R_1$. A Gaussian sphere of radius r drawn here encloses no charge whatever, because all of Q sits on the outer surface and lies outside this Gaussian. By Gauss's law the total flux through it is zero, and by the spherical symmetry the field on it is radial and of one fixed magnitude. The two facts can be reconciled only by setting that one magnitude to zero, and so the field inside the cavity is zero as well:

$$\mathbf{E} = \mathbf{0} \quad \text{for} \quad r < R_1$$

Gathering the three answers together: the electric field is zero everywhere inside the hollow conductor, both within the metal of the shell and within the cavity it encloses. At the outer surface the field jumps abruptly to its peak value kQ/R_2^2 , and from there outward it falls off as the inverse square of distance, exactly as if the whole charge were squeezed to a point at the centre. The two panels of the figure record this behaviour in picture and in graph.

The most striking thing about the result is the second region. A region of empty space, surrounded on all sides by a charged metal shell that may itself sit at a high potential, has zero field inside. The cavity is electrically calm regardless of what the shell is doing. This is the principle of the **Faraday cage**, and it is *why a person sitting inside a closed metal car during a lightning storm is safe even if the car itself is struck*. The metal body of the car becomes a charged hollow conductor for a brief instant, but the cavity inside it, where the occupants sit, holds its field-free peace. The very same result was leaned on in the previous section when the contrast between $E = 0$ and $V \neq 0$ was drawn, the cavity being the showcase example of a region where the field has vanished while the potential has not. Gauss's law has now done the work of placing that vanished field on a firm derivation.

A uniformly charged hollow insulating sphere

The fourth and final sphere of this section combines the hollow geometry of the third with the volume-filled charge distribution of the first. Take a hollow insulating sphere of inner radius R_1 and outer radius R_2 , the shell material between the two radii carrying a uniform volume charge density ρ , the cavity inside the inner radius empty. Because the material is an insulator, the charge has no freedom to migrate; it stays where it was deposited and so really does fill the shell volume uniformly, instead of running to the outer surface as it did for the hollow conductor. The total charge of the sphere is then the volume charge density multiplied by the volume of the shell:

$$Q = \rho \times \frac{4}{3}\pi (R_2^3 - R_1^3)$$

Again, Gauss's law is applied three times in turn, one Gaussian sphere for each of the three radial regions.

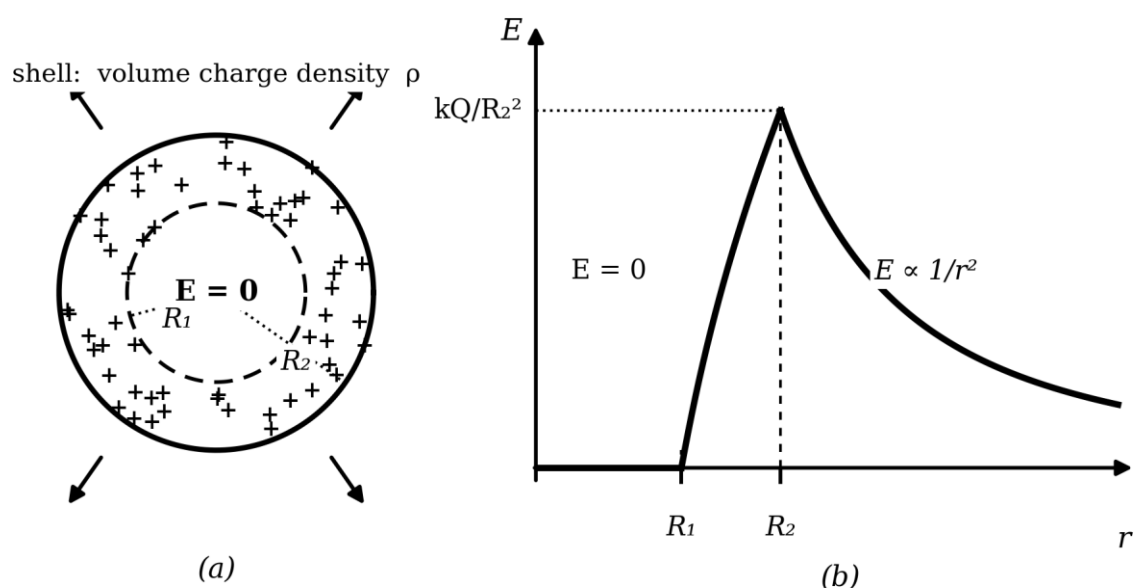


Figure: A uniformly charged hollow insulating sphere. Panel (a): the insulating shell of inner radius R_1 and outer radius R_2 carries a uniform volume charge density ρ throughout its bulk. The cavity is empty. Panel (b): the field strength against distance. Zero inside the cavity, the field rises continuously through the shell from zero at the inner surface to its peak value kQ/R_2^2 at the outer surface, then falls as the inverse square of distance outside.

First region: Take first the region inside the cavity, $r < R_1$. A Gaussian sphere of radius r drawn here lies wholly inside the empty cavity and encloses no charge, since the charge is distributed entirely through the shell material lying outside this Gaussian. By Gauss's law the total flux is zero, and by the spherical symmetry the field on the Gaussian is radial and of one fixed magnitude. The two facts can only be reconciled by setting that magnitude to zero:

$$\mathbf{E} = \mathbf{0} \quad \text{for} \quad r < R_1$$

The cavity is field-free for the same reason as in the hollow conductor of the previous treatment.

Second region: Take next the region within the shell of charge itself, $R_1 < r < R_2$. A Gaussian sphere of radius r drawn in this range encloses only part of the charged shell, namely the portion of the shell that lies between R_1 and the radius r of the Gaussian itself. The volume of this enclosed portion is the volume of a spherical shell stretching from R_1 to r :

$$V_{\text{enc}} = \frac{4}{3}\pi (r^3 - R_1^3)$$

The charge inside the Gaussian is the density multiplied by this volume, and using the relation between Q and ρ above, the result can be written either as a density expression or as a simple fraction of Q :

$$Q_{\text{enc}} = \rho \times \frac{4}{3}\pi (r^3 - R_1^3) = Q \times \frac{(r^3 - R_1^3)}{(R_2^3 - R_1^3)}$$

The ratio on the right is the fraction of the total charge that has been swept up by the Gaussian sphere of radius r , $\left(\frac{Q_{\text{enc}}}{Q} = \frac{r^3 - R_1^3}{R_2^3 - R_1^3}\right)$; the remaining charge sits in the still-unenclosed shell layer between r and R_2 .

Gauss's law now reads:

$$E \times 4\pi r^2 = \frac{Q_{\text{enc}}}{\epsilon_0}$$

Substituting Q_{enc} and solving for the field strength:

$$E = \frac{Q (r^3 - R_1^3)}{4\pi\epsilon_0 r^2 (R_2^3 - R_1^3)} = \frac{kQ (r^3 - R_1^3)}{r^2 (R_2^3 - R_1^3)}$$

This is the field inside the charged shell, and the three-cube structure of the result is the signature of this case. It does its checking at the two boundaries of the shell. At the inner boundary, $r = R_1$, the numerator factor $(r^3 - R_1^3)$ vanishes and the formula returns $E = 0$, matching the field already found in the cavity. At the outer boundary, $r = R_2$, the same factor becomes $(R_2^3 - R_1^3)$, which is exactly the factor sitting in the denominator and cancels with it, leaving $E = kQ/R_2^2$, the field of a point charge Q at the outer-surface distance. The shell formula has joined onto the cavity formula on its inside and joins onto the outside formula on its outside, the whole field varying continuously from one region to the next.

Third region: Take finally the region outside the sphere, $r > R_2$. A Gaussian sphere of radius r drawn here encloses the whole of the charge, Q , and Gauss's law gives the by-now-familiar point-charge field:

$$\mathbf{E} = \frac{kQ}{r^2}$$

Gathering the three answers together: inside the cavity the field is zero, exactly as for the hollow conductor. Within the shell the field rises smoothly from zero at the inner surface R_1 to the peak value kQ/R_2^2 at the outer surface R_2 , the rise governed by the $(r^3 - R_1^3)/r^2$ combination so that the field is small near R_1 and approaches its maximum near R_2 . Outside the shell the field falls as the inverse square of distance, exactly as for a point charge. Panel (b) of the figure draws this whole behaviour as a single continuous curve.

With this fourth case, every spherical charge distribution that an examination is likely to set has now been worked through: the solid insulating sphere, the solid conducting sphere, the hollow conducting sphere, and the hollow insulating sphere.

The sphere is behind us. The next canonical surface is the one a sphere can never become, however large you make it: perfectly, endlessly flat. It is also the surface that produces the single most useful field in all of electrostatics, so it is worth meeting slowly.

2. An infinite plane sheet of charge

Imagine an enormous flat sheet, far wider than anything you intend to do near it: think of a great charged wall, or one face of the metal plates inside a capacitor, with charge spread evenly across it. We measure how thickly the charge is spread with the **surface charge density**, written σ , defined as the *charge carried on each unit of area of the sheet*. Its unit is the coulomb per square metre, C/m^2 . A sheet with a large σ is a crowded sheet; a sheet with a small σ is a sparse one.

Before writing a single equation, stand at a point a short distance in front of this sheet and ask what the field there can possibly look like. The sheet is so vast that, from where you stand, it has no edges and no centre, and it looks exactly the same in every direction you might slide along it. A field cannot do something that the situation gives it no reason to do, and two firm conclusions follow from that thought alone.

First, the field must point straight out of the sheet, at right angles to it. Suppose for a moment that it leaned a little to one side. That sideways lean would have to choose one particular direction along the sheet, and yet every direction along the sheet is identical, so nothing could make the field prefer one over another. The only honest sideways component is no sideways component at all. The field is purely perpendicular to the sheet.

Second, the field has the same strength at every point that lies the same distance from the sheet, and it points directly away from the sheet on both faces if the sheet is positively charged, and directly toward it on both faces if the sheet is negative. A sheet with no special place on it cannot produce a strong field at one spot and a weak field at another the same distance away.

That is the picture. Now we choose a Gaussian surface that respects it, and Gauss's law hands us the magnitude of the field almost for free.

The natural surface to choose is a short, *closed cylinder*, often called a **pillbox**, pushed through the sheet so that it straddles it. Its two flat circular ends, each of area A , sit parallel to the sheet, one end poking out of each face, while its curved side wall runs perpendicular to the sheet. The figure shows the arrangement.

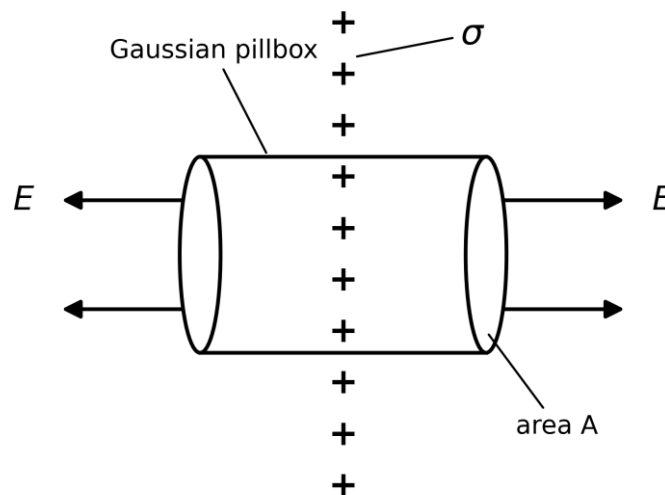


Figure: A short closed cylinder, or pillbox, pushed through an infinite plane sheet of charge so that it straddles the sheet. The two flat caps, each of area A , lie parallel to the sheet, one on each side; the curved wall is perpendicular to it. The electric field E leaves both caps at right angles to the sheet, and the pillbox encloses a patch of the sheet of area A .

Now we add up the electric flux passing out through this closed surface, piece by piece. Begin with the curved side wall. The field there points straight out of the sheet, while the wall itself runs perpendicular to the sheet, so the field merely skims along the wall without ever poking through it. The angle between the field and the wall's outward normal is a full right angle, and $\cos 90^\circ = 0$, so the curved wall contributes no flux whatsoever.

That leaves the two flat end caps. On each cap the field is perpendicular to the surface and points straight out through it, exactly along the cap's outward normal, so $\cos 0^\circ = 1$ and the flux through one cap is simply the field strength multiplied by the cap's area, EA . There are two caps, one on each side, and each contributes the same amount, so the total flux out of the whole pillbox is:

$$\Phi = EA + EA = 2EA$$

Next we need the charge sealed inside the pillbox. The cylinder punches a small disc out of the sheet, a patch of area A , and whatever charge sits on that patch is the charge enclosed. Since σ is the charge on each unit of area, the charge on a patch of area A is $Q_{\text{enc}} = \sigma A$. No other charge lies inside the pillbox; the rest of the sheet is outside it.

Gauss's law states that the total flux out of a closed surface equals the charge it encloses divided by ϵ_0 . Setting our two results equal to each other:

$$2EA = \frac{\sigma A}{\epsilon_0}$$

Look closely at what happens next. The area A appears on both sides of the equation and cancels completely. Making E the subject, the field of an infinite plane sheet of charge stands revealed:

$$E = \frac{\sigma}{2\epsilon_0}$$

Since $\frac{1}{\epsilon_0} = 4\pi k$, the result may equally be written as:

$$E = 2\pi k\sigma,$$

whichever constant is the more convenient for the numbers in front of you.

Two features of this small formula deserve far more attention than its size suggests.

The first is that the area A vanished along the way. This is a relief rather than a surprise. The pillbox was an imaginary surface that we invented purely for the calculation, and a real, physical field cannot possibly depend on the size of an imaginary box. Had A survived into the answer, that would have been a sure sign of a mistake.

The second feature is genuinely startling. Search the formula $E = \frac{\sigma}{2\epsilon_0}$ for the distance from the sheet. It is not there. There is no r , no symbol of any kind for how far the chosen point lies from the sheet. Step one centimetre away from an infinite charged sheet, or step one kilometre away, and provided the sheet still looks infinite from where you stand, the field is exactly the same strength. The field of an infinite sheet does not fade with distance at all.

It is worth seeing why, because the reason is pure geometry. The field of a point charge weakens as $\frac{1}{r^2}$ because its field lines spray outward in every direction and must spread themselves ever more thinly over the surface of an ever-larger sphere. The field of a long straight line of charge weakens more gently, as $\frac{1}{r}$, because its lines spread only over the curved surface of an ever-larger cylinder. An infinite sheet, however, sends its field lines straight out, perpendicular and perfectly parallel to one another, and parallel lines never spread. With nothing to dilute them, the field they represent never weakens. A point charge spreads its lines over a sphere, a line spreads them over a cylinder, and a plane does not spread them at all.

What we have built, then, is a perfectly **uniform field**: the same strength and the same direction at every point on a given side of the sheet. Notice also the 2 sitting in the denominator. The field is $E = \frac{\sigma}{2\epsilon_0}$ and not $\frac{\sigma}{\epsilon_0}$ because the sheet fires its field out of both of its faces at once. The charge on the enclosed patch must share its flux between the two caps, and each side receives exactly half. That modest 2 is worth holding on to: the very next surface we build will remove it, and a great deal of useful physics lives in that difference.

A formula with no distance in it has a faintly suspicious look about it, rather like a train timetable printed with no times. One honest set of numbers will now be fed into it, and the suspicion will turn out to be the entire lesson.

BINDER Example 14

A large flat sheet carries a uniform surface charge density of $4.0\mu\text{C}/\text{m}^2$. Find the electric field at a point 2.0cm from the sheet, and then state the field at a point 50cm from the sheet. Take $\epsilon_0 = 8.854 \times 10^{-12}\text{C}^2/(\text{Nm}^2)$.

Solution

The field of an infinite sheet of charge does not depend on the distance from the sheet, so a single calculation will answer both parts of the question at once. The field of a uniformly charged sheet is:

$$E = \frac{\sigma}{2\epsilon_0}$$

Express every quantity in SI units. The surface charge density is $\sigma = 4.0\mu\text{C}/\text{m}^2 = 4.0 \times 10^{-6}\text{C}/\text{m}^2$. Substituting this value, together with ϵ_0 :

$$E = \frac{4.0 \times 10^{-6}\text{C}/\text{m}^2}{2 \times 8.854 \times 10^{-12}\text{C}^2/(\text{N} \cdot \text{m}^2)} = 2.3 \times 10^5\text{N/C}$$

The field at the point 2.0cm from the sheet is therefore $2.3 \times 10^5\text{N/C}$.

The field at the point 50cm from the sheet is the very same value, $2.3 \times 10^5\text{N/C}$, because the formula holds no distance for the change in position to act upon.

Making Sense of the Answer: *Two different distances were quoted in the question and a single field came back for both of them, which is not an oversight but the whole point of the example. It is worth pausing to feel how strange that is. Had the charge been a point charge instead of a sheet, moving from 2.0cm out to 50cm, a twenty-five-fold increase in distance, would have weakened the field by a factor of twenty-five squared, that is by 625, dimming it almost to nothing. The sheet does no such thing. Its field lines leave the surface perpendicular and travel side by side without ever spreading, so there is nothing to thin the field out, and the needle of a field meter would not so much as twitch as you walked away. One honest caution belongs here for the real world: a genuine sheet has edges, and the result $E = \frac{\sigma}{2\epsilon_0}$ is trustworthy only while the distance from the sheet stays small compared with the width of the sheet itself.*

Think Like a Physicist: *When a formula for a field has no distance in it, treat that empty space as a message rather than as a missing term. The absence is telling you that the geometry has arranged the field lines so that they never spread. This is a habit worth building: read a formula for what it leaves out just as carefully as for what it puts in. Here the three great charge shapes line up into one clean lesson. A point charge spreads its lines over a sphere, and its field falls as $\frac{1}{r^2}$. A long straight line spreads its lines over a cylinder, and its field falls as $\frac{1}{r}$. An infinite plane sends its lines out flat and parallel, they spread over nothing, and its field falls off not at all. Sphere, cylinder, plane: each shape writes its own signature into the way the field fades, and a physicist learns to read that signature straight off the formula.*

A field that flatly refuses to weaken with distance is an odd creature to be holding at the end of a worked example. Odd, but about to become extraordinarily useful. Keep this stubborn, uniform field firmly in mind, because the next surface we build does something clever with it: it brings in a second sheet, sets the two facing each other, and traps the field neatly in the gap between them.

3. Two parallel sheets of charge

Imagine two large flat sheets, each one an infinite sheet of the kind studied in the previous section, held parallel to each other with a narrow gap between them, like the two facing walls of a thin corridor. One sheet carries a positive surface charge density, which we write as $+\sigma$; the other carries an equal and opposite negative density, $-\sigma$. This is not an artificial arrangement dreamt up to make a sum come out neatly. It is exactly the layout of charge inside a capacitor, the device on which every camera flash and every radio quietly depends, and which the next chapter will study in full.

Before any algebra, guess what the field of this pair ought to look like. The positive sheet pushes its field away from itself; the negative sheet pulls field toward itself. In the gap between the two, both of these urges

point the same way, from the positive sheet across to the negative one, so we should expect a strong field there. In the regions outside the pair, the push of one sheet and the pull of the other are set against each other, so we should expect them to fight to a standstill and leave little or nothing behind. A strong, orderly field sealed inside the gap, and near-emptiness outside: that is the guess. Now we make it exact.

The pleasant surprise is that no new Gaussian surface is needed; the hard work was done already. A single infinite sheet, we found, produces a field of magnitude $\frac{\sigma}{2\epsilon_0}$ on each side of itself, pointing directly away from the sheet if it is positive and directly toward the sheet if it is negative. Our arrangement is simply two such sheets, and the field at any point is the vector sum of what each sheet produces there on its own. Adding the separate fields of separate sources is called the **principle of superposition**, and it turns a fresh problem into a short piece of arithmetic.

Set the positive sheet on the left and the negative sheet on the right, and agree to call the direction running from the positive sheet toward the negative one, that is, rightward, the positive direction. Each sheet on its own contributes a field of size $\frac{\sigma}{2\epsilon_0}$; all we must do is work out, in each region, which way those two contributions point, and then add them. The figure shows the arrangement and the field it produces.

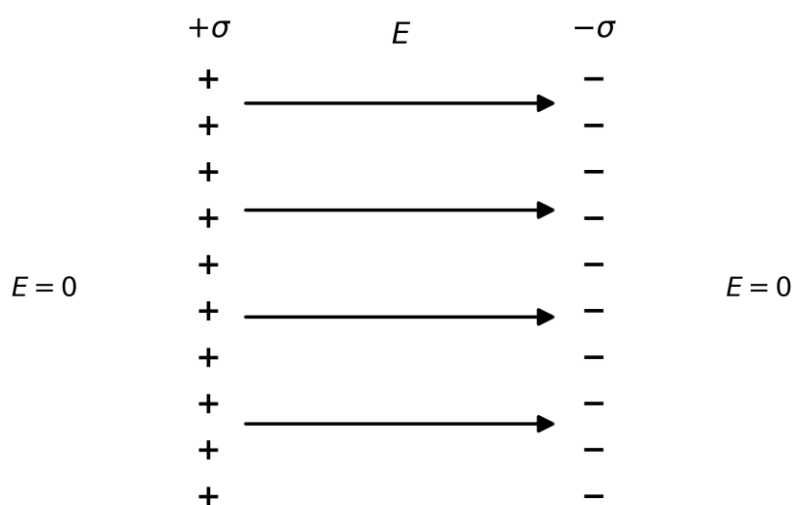


Figure: Two parallel sheets carrying equal and opposite surface charge densities, $+\sigma$ and $-\sigma$. In the gap between the sheets the two single-sheet fields point the same way and reinforce, giving a uniform field E directed from the positive sheet to the negative one. In the regions outside the pair the two fields oppose and cancel, leaving $E = 0$.

Begin **in the gap between the plates**. Here the field of the positive sheet points away from it, which is to the right. The field of the negative sheet points toward it, which is also to the right. The two contributions point the same way, so they add:

$$E = \frac{\sigma}{2\epsilon_0} + \frac{\sigma}{2\epsilon_0} = \frac{\sigma}{\epsilon_0}$$

The two equal halves combine into a single whole. The electric field in the gap between two parallel, oppositely charged sheets is therefore:

$$E = \frac{\sigma}{\epsilon_0}$$

Now step outside the pair, to the **left of the positive sheet**. The field of the positive sheet still points away from it, which out here means to the left. The field of the negative sheet still points toward it, which out here means to the right. The two contributions are equal in size but now point in opposite directions, so they cancel exactly:

$$E = \frac{\sigma}{2\epsilon_0} - \frac{\sigma}{2\epsilon_0} = 0$$

The same cancellation occurs in the region to the right of the negative sheet, by an argument word for word the same. The electric field everywhere outside the pair of sheets is zero. The whole field of the arrangement has been swept into the gap and sealed there, exactly as the opening guess suggested.

It is worth pausing on the 2 that has just disappeared. In the previous section a single sheet gave a field $\frac{\sigma}{2\epsilon_0}$, and the 2 in its denominator was there because that lone sheet had to share its flux between its two faces. Here, between the plates, that sharing works in our favour. Each sheet sends one half-strength field, $\frac{\sigma}{2\epsilon_0}$, into the gap, and because the two half-fields point the same way they combine into a whole, $\frac{\sigma}{\epsilon_0}$. The field between the plates is exactly twice the field of either sheet alone.

Notice what kind of field this is. Each single-sheet field was uniform, the same at every point on its side, and the sum of two uniform fields is uniform as well. The field in the gap has the same strength $\frac{\sigma}{\epsilon_0}$ and the same direction at every point between the plates. It does not sag in the middle nor strengthen near a plate. This steady, even field is the one assumed, usually without a word of comment, in every parallel-plate problem a student meets, and it is the reason a charge released between the plates moves with a constant acceleration.

Notice too where the field is, and where it is not. It lies entirely inside the gap and is exactly zero everywhere outside. The pair of sheets has gathered its whole field into the narrow space between them and sealed it there. A capacitor is little more than this arrangement fitted with a pair of leads, and the field we have just found, $E = \frac{\sigma}{\epsilon_0}$, is the field inside every capacitor ever built. The next chapter takes this single result and, in no more than a few lines, turns it into the rule for how much charge such a device can store, $C = \frac{\epsilon_0 A}{d}$. A great deal of what that chapter does rests on the short derivation just completed.

The companion case: two sheets carrying the same sign

The arrangement just settled was the capacitor: two sheets carrying equal and opposite charge. It is worth to note what happens if the two sheets carry charge of the **same** sign instead, say $+\sigma$ on each, since the result is the exact mirror image of the capacitor's and explains in one stroke a piece of physics already met in this chapter.

Superposition does the whole calculation in three lines. Each sheet on its own contributes a field of magnitude $\frac{\sigma}{2\epsilon_0}$ on each of its sides, directed away from the positive sheet. Setting the two sheets side by side with the rightward direction as positive, the contributions in the three regions add as follows:

$$\text{Left of both: } E = -\frac{\sigma}{2\epsilon_0} + -\frac{\sigma}{2\epsilon_0} = -\frac{\sigma}{\epsilon_0}$$

$$\text{Between: } E = +\frac{\sigma}{2\epsilon_0} + -\frac{\sigma}{2\epsilon_0} = 0$$

$$\text{Right of both: } E = +\frac{\sigma}{2\epsilon_0} + +\frac{\sigma}{2\epsilon_0} = +\frac{\sigma}{\epsilon_0}$$

Between the two like-charged sheets the two single-sheet fields point opposite ways and cancel exactly; outside the pair they point the same way and reinforce. The picture is the mirror image of the capacitor: the field is now expelled from the gap and dumped to either side, and the gap itself sits at zero. Same-sign sheets seal **nothing** between them.

This little result is the planar twin of the hollow-conductor result met under Gauss's law. There, charges spread uniformly over the outer surface of a hollow metal sphere produced zero field everywhere inside the cavity; here, charges of the same sign spread over two parallel sheets produce zero field everywhere between them. The two cases are the same physics in different geometries: when charges of one sign are arranged symmetrically around a region, their fields meet in the middle pointing opposite ways and cancel. This is the deep reason the Faraday cage works, and the present arrangement is the simplest possible setting in which to see it.

A single sheet handed us a field that ignored distance altogether. Two sheets, set face to face, strike an even better bargain: a strong and tidy field in the gap, and nothing whatsoever outside. One honest set of numbers will show how generous the bargain is.

BINDER Example 15

Two large parallel plates are held a short distance apart, one carrying a uniform positive surface charge density and the other an equal and opposite negative density, each of magnitude $8.0\mu\text{C}/\text{m}^2$. (a) Find the electric field in the gap between the plates. (b) State the electric field in the region outside the plates. Take $\epsilon_0 = 8.854 \times 10^{-12}\text{C}^2/(\text{N}\cdot\text{m}^2)$.

Solution

(a) The field in the gap between two parallel, oppositely charged sheets is uniform and has the magnitude:

$$E = \frac{\sigma}{\epsilon_0}$$

Express the surface charge density in SI units: $\sigma = 8.0\mu\text{C}/\text{m}^2 = 8.0 \times 10^{-6}\text{C}/\text{m}^2$. Substituting this value together with ϵ_0 :

$$E = \frac{8.0 \times 10^{-6}\text{C}/\text{m}^2}{8.854 \times 10^{-12}\text{C}^2/(\text{N}\cdot\text{m}^2)} = 9.0 \times 10^5\text{N}/\text{C}$$

The field in the gap is therefore $9.0 \times 10^5\text{N}/\text{C}$, uniform throughout the gap and directed from the positive plate toward the negative plate.

(b) Outside the plates the field of the positive sheet and the field of the negative sheet are equal in size and opposite in direction, so they cancel. The electric field everywhere outside the pair of plates is **zero**.

Making Sense of the Answer: It is worth comparing this gap field with the field of a single sheet carrying the same σ . One sheet on its own would give $\frac{\sigma}{2\epsilon_0}$, which works out to $4.5 \times 10^5\text{N}/\text{C}$; the pair gives $9.0 \times 10^5\text{N}/\text{C}$, exactly twice as much. The second sheet has doubled the field in the gap, and it has done so without adding any charge at all to the point where the field is measured. The zero outside is every bit as important as the large value inside. A field meter carried right up to the back of either plate would read nothing, because the two sheets cancel each other there. The honest caution from the previous section still holds: real plates have edges, and the tidy results, $\frac{\sigma}{\epsilon_0}$ inside and zero outside, hold well in the interior of a wide pair of plates but blur where the edges are.

Think Like a Physicist: The whole of this section was settled without a single new integral and without drawing one fresh Gaussian surface. We already owned the field of one sheet, so we simply placed two of them and added. When a new problem is plainly built out of pieces you have already solved, do not start again from the beginning; identify the pieces and add their effects. That habit is the principle of superposition, and it is one of the most powerful labour-saving devices in all of physics. Notice also the shape of the result: the two fields reinforce in one region and cancel in another, purely because of the directions in which they point. Fields adding where they agree and cancelling where they disagree is a pattern you will meet again and again across physics, and the arithmetic never grows harder than it was here.

So the humble pair of sheets turns out to be quietly magnificent. You have just derived, in a handful of honest lines, the exact field inside every capacitor on Earth. The next chapter does something almost unfair with that result: it takes this one formula and, in three short lines, extracts from it the rule for how much electrical charge a capacitor can hoard. Keep $E = \frac{\sigma}{\epsilon_0}$ somewhere safe. One canonical surface still remains, and it is the thinnest of them all: not a sheet at all, but a single endless line of charge.

4. An infinite line of charge

Imagine an endlessly long, straight, thin line, a charged thread, or a very long charged wire seen from close up, carrying charge spread evenly along its whole length. We measure how thickly the charge is spread with the **linear charge density**, written λ , defined earlier as the *charge carried on each unit length of the line*. Its unit is the coulomb per metre, C/m.

Stand at a point a perpendicular distance x from this line and ask, as we have asked twice before, what the field there is able to be. The line looks identical whether you face along it one way or the other, so the field can carry no component running along the line; any such component would have to choose a direction the line itself flatly refuses to distinguish. The line also looks identical from every direction around it. Two conclusions follow at once. First, the field points straight out from the line, radially, at right angles to it. Second, its strength depends only on x , the perpendicular distance from the line, and is the same at every point that shares that distance.

This is cylindrical symmetry, and it calls for a cylindrical Gaussian surface. We take an imaginary cylinder of radius x and length L , sharing its axis with the line of charge, so that the line runs straight down its centre. The figure shows the arrangement.

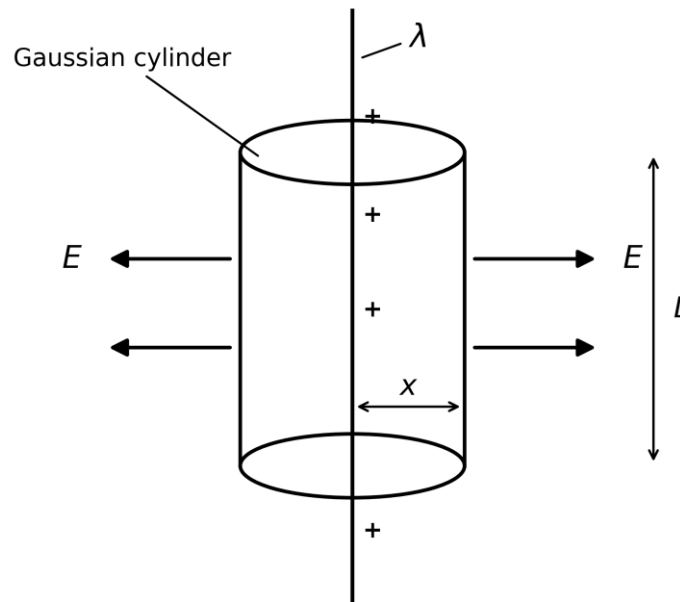


Figure: A cylindrical Gaussian surface of radius x and length L , sharing its axis with an infinite line of charge of linear density λ . The field E points radially outward and pierces only the curved wall; it skims along the two flat end caps and so sends no flux through them.

Add up the flux through this closed cylinder in two parts: its two flat end caps, and its curved side wall. Take the end caps first. On either cap the outward-pointing normal runs parallel to the line itself, while the field there points radially, square across the line. The field thus skims along the face of the cap without ever passing through it; the angle between the field and the cap's normal is a full right angle, $\cos 90^\circ = 0$, and the two end caps together contribute no flux at all.

The curved side wall is another matter entirely. Everywhere on it the field points radially outward, and so does the wall's own outward normal; the two are exactly aligned, $\cos 0^\circ = 1$. Better still, every point of the curved wall lies at the same perpendicular distance x from the line, so by our symmetry the field has the same strength E at every point of it. The flux through the curved wall is therefore E multiplied by the wall's area. A cylinder of radius x and length L has a curved wall of area $2\pi xL$, its circumference $2\pi x$ multiplied by its length L . The total flux out of the whole closed surface is then:

$$\Phi = E(2\pi xL)$$

The charge enclosed is whatever length of line lies sealed inside the cylinder. The cylinder is L long and the line runs straight down its axis, so it encloses a length L of charged line. With λ the charge on each unit of length, the enclosed charge is $Q_{\text{enc}} = \lambda L$.

Gauss's law now sets the total flux equal to the enclosed charge divided by ϵ_0 :

$$E(2\pi xL) = \frac{\lambda L}{\epsilon_0}$$

The length L stands on both sides of the equation and cancels, and well it should, for L was a feature of an imaginary cylinder, never of the real line. Making E the subject delivers the field of an infinite line of charge:

$$E = \frac{\lambda}{2\pi\epsilon_0 x}$$

Since $\frac{1}{2\pi\epsilon_0} = 2k$, the same result is just as often written as:

$$E = \frac{2k\lambda}{x},$$

whichever constant sits more comfortably with the numbers in front of you.

Look closely at how this field fades. It falls off as $\frac{1}{x}$: double your distance from the line and the field drops to one half. This sits neatly between the two fields found earlier in this section. The field of a point charge spreads its lines outward over the surface of a sphere, whose area grows as the square of the distance, so that field falls as $\frac{1}{r^2}$. The field of an infinite sheet sends its lines straight out, parallel, and never lets them spread, so that field does not fall away at all. A line lies exactly in between the two. Its lines spread over the curved wall of a cylinder, whose area grows only in proportion to x , not to its square, so the field falls as $\frac{1}{x}$, gently. Sphere, cylinder, plane; an inverse square, an inverse first power, and no falloff at all. The geometry of the source is written directly into the way its field dies away.

There is one more thing worth saying, and it is really the whole point of this section. Earlier in this chapter we found this very same field, $E = \frac{\lambda}{2\pi\epsilon_0 x}$, by the long road. We chopped the line into countless tiny elements, wrote down the small field contributed by each, and summed them all with an integral. It worked, but it was heavy labour. Here, Gauss's law has reproduced the identical result in a single line of algebra, almost the moment the right cylinder was drawn. That is the power the law has been promising from the start. Choosing a Gaussian surface that matches the symmetry of the charge, a sphere for a point or a ball, a pillbox for a sheet, a cylinder for a line, turns a hard integral into one short multiplication. Picking the surface is the real physics; the algebra that follows is light work.

Four formulae for four shapes are now in hand. The line's formula is the only one of them that fades with distance in a gentle, well-mannered way, and a single calculation will show exactly how gentle that gentleness is.

BINDER Example 16

A long straight wire carries a uniform linear charge density of $5.0 \times 10^{-8} \text{C/m}$. (a) Find the electric field at a point 0.10m from the wire. (b) Hence state the field at a point 0.20m from the wire. Take $\epsilon_0 = 8.854 \times 10^{-12} \text{C}^2/(\text{Nm}^2)$.

Solution

(a) The field of an infinite line of charge, at a perpendicular distance x from it, is:

$$E = \frac{\lambda}{2\pi\epsilon_0 x}$$

Substitute λ , ϵ_0 and $x = 0.10\text{m}$, every quantity in SI units:

$$E = \frac{5.0 \times 10^{-8} \text{C/m}}{2\pi \times 8.854 \times 10^{-12} \text{C}^2/(\text{N} \cdot \text{m}^2) \times 0.10\text{m}} = 9.0 \times 10^3 \text{N/C}$$

The field at the point 0.10m from the wire is $9.0 \times 10^3 \text{N/C}$, directed radially away from the wire.

(b) The field of a line of charge falls off as $\frac{1}{x}$, so doubling the distance halves the field. Without a fresh substitution, the field at 0.20m is simply half the value just found, $4.5 \times 10^3 \text{N/C}$, again directed radially outward.

Making Sense of the Answer: The two answers differ by exactly a factor of two for a doubling of distance, and that is the unmistakable signature of a $\frac{1}{x}$ field. It is worth holding the three signatures of this section side by side. Move twice as far from a point charge and its field drops to a quarter, since an inverse-square law falls away fast. Move twice as far from an infinite sheet and its field does not change in the slightest. Move twice as far from a line and the field drops to one half: the line is the moderate middle case of the three. Part (b) also makes a quiet point about method. Once the shape of a law is known, a second value can often be read straight off the first, with no need to repeat the whole calculation. The familiar real-world caution still belongs here: the formula treats the line as infinitely long, so it is trustworthy only for points whose distance from the wire is small compared with the wire's own length, comfortably clear of its ends.

Think Like a Physicist: This whole section asked one skill of you, over and over: match the imaginary surface to the symmetry of the real charge. A point or a sphere has spherical symmetry, so a spherical surface was the one to draw. A sheet has flat symmetry, so a flat-faced pillbox was drawn. A line has cylindrical symmetry, so a cylinder was drawn. In every single case, once the surface fitted the symmetry,

the field stepped outside the flux as a plain constant and Gauss's law quietly did the rest. So when you meet a new charge distribution, do not reach for a formula first; reach for its symmetry. Ask what the field is forbidden to do, let that forbidding pin down both its direction and the surface that respects it, and the law will hand you the magnitude. Symmetry first, surface second, algebra last: that is the entire method in six words.

And there it is: four shapes, four fields, one law. Gauss's law has now been aimed in turn at a sphere, a sheet, a facing pair of sheets and a line, and each time it has paid out a clean formula almost for the asking. One short calculation still waits to close the section, and it is a faintly dangerous one. It asks how much charge you may pile onto an object before the very air around it loses patience, breaks down, and answers you back with a spark.

That spark deserves a proper explanation, and, pleasingly, it calls for no new physics at all. The first of the four applications just completed, the charged sphere, joined to a single new fact about air, is enough to predict exactly when it happens.

Dielectric breakdown: when the air gives way

Air, in its ordinary state, is an insulator: it does not conduct electricity, which is why charge can sit undisturbed on objects and why the wires inside a wall do not simply discharge into the room. But this refusal to conduct has a limit. Drive a strong enough electric field through air and the field begins to tear electrons loose from the air molecules. Those freed electrons, hurled along by the same field, crash into further molecules and knock loose more electrons still, which are themselves hurled onward to free yet more. In a fraction of a second the air along that path has turned from an insulator into a conductor, and charge floods across it. We see and hear that flood as a spark. Lightning is the very same process, run on the scale of a thundercloud.

Every insulator has a ceiling: the strongest electric field it can withstand before it breaks down in this way and begins to conduct. That ceiling is called the **dielectric strength** of the material. For dry air at ordinary pressure the dielectric strength is about $3.0 \times 10^6 \text{ V/m}$, three million volts per metre. The volt per metre is simply another name for the unit of electric field; a field of one V/m is the same as a field of one N/C. Keep the field in air below that figure and the air insulates in silence; let the field once reach it, and the air gives way.

Now the question that closes this section. Take a metal sphere of radius R carrying a charge Q , standing in open air. The very first of the four applications told us the field such a sphere produces: outside the sphere, the field is exactly that of a point charge Q placed at the centre. That field is strongest right at the sphere's own surface, a distance R from the centre, where its magnitude is $\frac{kQ}{R^2}$; further out it only weakens, falling away as $\frac{1}{r^2}$. As charge is added to the sphere, Q climbs and the surface field climbs with it. The sphere cannot be charged without limit. The moment its surface field reaches the dielectric strength of air, the air at the surface breaks down, and any further charge simply sprays off into the surroundings as **sparks**. The greatest charge the sphere can hold is therefore the charge that lifts its surface field exactly to that ceiling.

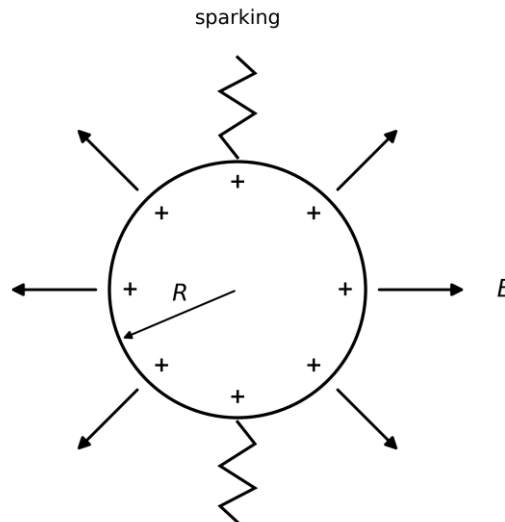


Figure: A charged conducting sphere of radius R . The field E is strongest right at the surface and weakens outward. When the surface field reaches the dielectric strength of the surrounding air, the air breaks down and the sphere sparks; this sets a ceiling on the charge it can hold.

Anyone who has tugged off a wool jumper on a dry day, and heard it crackle in the dark, has already run this experiment in miniature. Here it is again, this time with the numbers filled in.

BINDER Example 17

A metal sphere of radius 10cm stands in dry air, whose dielectric strength is $3.0 \times 10^6 \text{V/m}$. Find the greatest charge the sphere can hold before the air at its surface breaks down and it begins to spark. Take $k = 8.99 \times 10^9 \text{Nm}^2/\text{C}^2$.

Solution

The field just outside a charged sphere is the field of a point charge sitting at its centre. At the surface, a distance R from that centre, its magnitude is:

$$E = \frac{kQ}{R^2}$$

The air breaks down when this surface field reaches the dielectric strength. Setting the surface field equal to that ceiling and making Q the subject:

$$Q = \frac{ER^2}{k}$$

Substitute the dielectric strength for E , with $R = 10\text{cm} = 0.10\text{m}$, keeping every quantity in SI units:

$$Q = \frac{3.0 \times 10^6 \text{N/C} \times (0.10\text{m})^2}{8.99 \times 10^9 \text{N} \cdot \text{m}^2/\text{C}^2} = 3.3 \times 10^{-6} \text{C}$$

The greatest charge the sphere can hold is therefore about $3.3 \times 10^{-6} \text{C}$, a little over three microcoulombs.

Making Sense of the Answer: Three microcoulombs is a startlingly small charge, a few millionths of a single coulomb, and yet it is the most that a sphere the size of a large grapefruit can hold in open air. The result is really a statement about the size of the coulomb itself: the coulomb is an enormous unit, and a whole coulomb of static charge resting on a small object in air is something nature simply will not permit. It is the reason the charges met all through electrostatics are almost always quoted in microcoulombs or nanocoulombs. The size of the sphere matters as well. Because the surface field is $\frac{kQ}{R^2}$, a larger sphere can take on more charge before that field reaches the ceiling, while a small sphere, or worse a sharp point with a very small radius of curvature, reaches the ceiling at a far smaller charge. That is precisely why high-

voltage equipment is built with smooth, well-rounded surfaces, and why a sharp metal spike is the first thing to spark, which is the very principle a lightning conductor is built upon.

Think Like a Physicist: This problem arrived dressed as something new, the dramatic business of sparking air, and yet it was solved with nothing new at all. The whole calculation rested on a single result derived earlier in this section, the field outside a charged sphere, joined to one new fact about air. When a problem looks unfamiliar, the first move is to search it for the familiar piece hidden inside. Notice also the shape of the method. The question asked for a maximum, the most charge the sphere could hold, and the route to it was to take the formula for the relevant quantity, set it equal to its limiting value, and solve. A great many physical questions, the most, the least, the longest, the fastest, the moment just before something gives way, are answered by exactly that move: name the limit, set the formula equal to it, and solve.

With that, Gauss's law has given, for the moment, all that it has to give. This section set out from a single idea, that counting the field lines crossing a closed surface counts the charge sealed within it, and from that one idea it has produced the field of a sphere, of a sheet, of a facing pair of sheets and of a line, and finally the precise charge at which the surrounding air revolts. The electric field has now been mapped thoroughly. The next section turns to a different and equally powerful question: **not how hard the field pushes at each point, but how much energy is stored in the pushing**, the idea of electric potential.

ELECTRIC POTENTIAL AND POTENTIAL DIFFERENCE

Until now the electric field has held the stage: a vector at every point of space, telling any charge placed there which way it will be pushed and how hard. That is one full and correct description of the electrical world. But there is a second description, and for a great many problems it is far the more convenient. It speaks not of forces and directions but of energy, and its central idea is **electric potential**.

To reach that idea, think first about what happens when a charge is moved. Carry a small positive charge from one point in a field to another. If you move it against the field, you must push, and in pushing you do work on it; that work does not vanish but is stored, ready to be released, exactly as the work done in lifting a stone is stored as gravitational potential energy. If instead you let the charge run along the field, the field itself does the work, and the charge speeds up, gaining kinetic energy as it goes. Either way, energy is bound up in a charge's position within a field. This is the thread we now follow.

Take two points in a field, label them A and B, and let an external agent carry a small positive test charge q slowly from B to A, slowly enough that it never picks up speed. Call the work the agent does on that journey W_{AB} . Now comes the key step: divide that work by the size of the charge moved. The quantity $\frac{W}{q}$ does not depend on how large or small a test charge was chosen, for doubling the charge doubles the work and leaves the ratio untouched. It is therefore a property of the two points and the field between them, and of nothing else. We give it a name, the **potential difference** between A and B and is expressed as the work done per unit charge in carrying a positive charge from B to A:

$$V_{AB} = \frac{W_{AB}}{q}$$

In words: **The potential difference between two points is the work done per unit charge in carrying a positive charge from one of the points to the other.**

Because it is an energy divided by a charge, the unit of potential difference is the joule per coulomb. This unit is met so often that it is given its own name, the **volt**, after Alessandro Volta: one volt is one joule per coulomb, $1V = 1J/C$. To say that the potential difference between two points is one volt is to say that one joule of work is needed to carry each coulomb of charge between them, no more and no less.

That definition is the practical one, but it leaves the field itself out of sight. The two are easily joined. To carry the charge against the field, the agent must, at every step of the way, push with a force equal and opposite to the force qE (from $E = \frac{F}{q}$, $F = qE$) that the field exerts on it. Adding up the agent's work, step by tiny step, along the path from B to A, gives the potential difference directly:

$$V_{AB} = - \int_B^A \frac{qE \cdot dr}{q} = - \int_B^A E \cdot dr$$

The integral simply sums the field's effect along the path; the dot is a reminder that only the part of the field lying along each small step does any work, and the minus sign records that moving a positive charge against

the field calls for positive work from the agent. The field, put the other way round, points in the direction in which the potential falls.

One feature of this integral is too important to pass over. Its value depends only on the two endpoints, B and A, and not in the least on the route taken between them. A charge carried from B to A along a short straight path, and the same charge carried along a long and winding one, call for exactly the same work per unit charge. The figure shows two such routes. The reason is that the electrostatic force, like gravity, is a **conservative force**: the work it does around any closed loop is zero. This is what makes potential difference a sensible idea at all. Were the work to depend on the path, $\frac{W}{q}$ would not be a property of the two points, and could not be quoted as a single definite number.

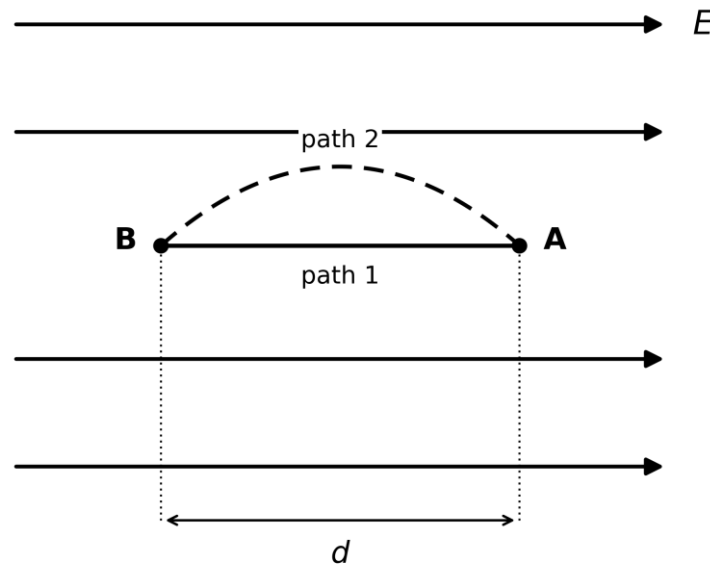


Figure: Two different routes, a straight one and a winding one, carrying a charge between the same two points B and A in a uniform field E . The work done per unit charge is the same along both paths; it depends only on the endpoints, which is what allows a single potential difference to be assigned to the pair.

Potential difference compares two points. It is often useful to speak instead of the potential at a single point, and for that we need one fixed point of reference to measure everything against, just as heights are quoted from sea level. The natural choice is a point infinitely far from all charges, where the field has faded to nothing; there the potential is taken to be zero. This results to the new term, **electric potential**.

The **electric potential** at a point is then *the work done per unit charge in bringing a positive charge from infinity to that point*.

The potential at a point near a positive charge is itself positive, since an agent must do positive work to push a positive test charge inward against the repulsion; the potential near a negative charge is negative. A potential difference is then simply the difference of two such potentials, $V_{AB} = V_A - V_B$.

And here is the great convenience promised at the outset. Electric potential is a **scalar**. At each point of space it is a single number, a count of joules per coulomb, with no direction attached to it. The field needed three numbers at every point, a strength and a direction; the potential needs only one. Problems that are awkward to handle with vectors often become straightforward once recast in terms of this single scalar, as the next section will show in full.

One case is worth working out at once, for it is the one met most often. Suppose the field is uniform, the same in strength and direction at every point, as it is in the gap between the charged parallel plates of the previous section. Carry a charge a distance d directly along such a field. The field strength E never changes along the way, so the sum that the integral performs is no harder than a single multiplication, and the potential difference between the two points is simply:

$$V = Ed$$

The potential difference across a uniform field is the field strength multiplied by the distance moved along it. Travelling along the field, in the direction it points, the potential falls; travelling against the field, it rises. A uniform field of E volts per metre shifts the potential by E volts for every metre crossed along it. This single relation, $V = Ed$, together with the definition V_{AB} as work per unit charge, will carry us through most of what follows.

Definitions stay slippery until a number is poured into them. Here is a uniform field, a marked-off distance and a charge to be carried across it; two short questions will turn the definitions solid.

BINDER Example 18

A uniform electric field of magnitude $2.0 \times 10^3 \text{ V/m}$ fills the region between two points A and B that lie a distance 5.0cm apart, measured along the field. (a) Find the potential difference between A and B. (b) Find the work that must be done to carry a charge of $3.0 \mu\text{C}$ from one of the points to the other.

Solution

(a) The two points lie along a uniform field, so the potential difference between them is the field strength multiplied by the distance separating them, $V = Ed$.

Substituting the given values in SI units, with $d = 5.0\text{cm} = 0.050\text{m}$:

$$V = (2.0 \times 10^3 \text{ V/m})(0.050\text{m}) = 100\text{V}$$

The potential difference between A and B is 100V.

(b) The work needed to carry a charge across a potential difference follows straight from the definition of that difference as work per unit charge, $W = qV$. With $q = 3.0 \times 10^{-6}\text{C}$ and the potential difference just found:

$$W = qV = (3.0 \times 10^{-6}\text{C})(100\text{V}) = 3.0 \times 10^{-4}\text{J}$$

The work that must be done is $3.0 \times 10^{-4}\text{J}$, or 0.30mJ.

Making Sense of the Answer: Notice that part (b) never once needed the field, the distance, or the geometry of the situation. As soon as the potential difference between the two points was known, the work followed from $W = qV$ alone, by the very definition of potential difference as work per unit charge. Notice the size of the answer too. Carrying three microcoulombs through a hundred volts costs only three ten-thousandths of a joule, less than the work of lifting a grain of rice a finger's width. Everyday electrostatic energies are small, even when the voltages, hundreds or thousands of volts, sound large, simply because the charges in question are so tiny. One point of care: a potential difference has a sense as well as a size. The potential is the higher at the point reached by moving against the field. Carrying a positive charge from the low-potential point to the high one demands positive work from you; carry it the other way and the field does the work for you, handing the charge energy it can spend.

Think Like a Physicist: Two descriptions of electricity now sit side by side, and it is worth being clear about how they differ. The electric field is a vector: at every point it carries both a strength and a direction, and to combine fields means to add vectors, with all the care that arrows demand. The electric potential is a scalar: at every point it is one plain number. Both describe the same physics, and the relation between them, the line integral met above, or in a uniform field the tidy $V = Ed$, lets you pass freely from one to the other. The skill is to choose the easier language for the question in hand. If a problem asks which way a charge will move, the field, with its built-in direction, answers at once. If a problem asks how much energy a charge will gain, or how much work a process will cost, the potential, a single number needing no vector arithmetic, is almost always the quicker road. A physicist keeps both maps of the one territory, and reads whichever makes the present journey shorter.

We have now defined the potential and learned to count it out in a uniform field. But uniform fields are the gentle exception. Most fields, the field of a point charge first among them, vary from place to place. The next task is to find the potential of a single point charge, and to discover, with it, a small miracle of convenience: where electric fields must be added as vectors, electric potentials, being mere numbers, may simply be added up.

The potential of a point charge

We ended the last section with a promise: that potentials, unlike fields, may be combined as plain numbers. To redeem it we need first the potential of the simplest source there is, a single point charge, and then the rule for adding several of them together.

Take a single point charge Q sitting alone in space, and ask for the potential it produces at a point a distance r away. That potential, as the last section established, is the work done per unit charge in carrying a small positive test charge to the point from infinitely far off, where the potential is counted as zero. The field of a point charge is already known to us: it points radially outward, with magnitude $E = \frac{kQ}{r^2}$. Because the test charge may be carried straight in along a radius, the field lies squarely along every step of the path, and the line integral of the field becomes an ordinary integral over the distance from the charge:

$$V = - \int_{\infty}^r E \cdot dr = - \int_{\infty}^r \left(\frac{kQ}{r^2} \right) dr$$

The constant k and the charge Q may both be taken outside the integral, since neither alters as the test charge moves:

$$V = -kQ \int_{\infty}^r \left(\frac{1}{r^2} \right) dr = -kQ \left[-\frac{1}{r} \right]_{\infty}^r = -kQ \left(-\frac{1}{r} - 0 \right) = -kQ \left(-\frac{1}{r} \right)$$

The potential of a point charge is therefore:

$$V = \frac{kQ}{r}$$

Read this formula closely, for it is among the most useful in the whole subject. The potential falls off as $\frac{1}{r}$, more gently than the field's $\frac{1}{r^2}$: halve your distance from the charge and the potential doubles, where the field would quadruple.

The sign of the potential follows the sign of the charge. Around a positive charge the potential is everywhere positive, sloping upward as the charge is approached, a hill; around a negative charge it is everywhere negative, a valley. And, decisively, $V = \frac{kQ}{r}$ carries no direction with it. It is a pure number at each point, a count of joules per coulomb.

Since k is the same constant that appears in Coulomb's law, the potential is just as often written as:

$$V = \frac{Q}{4\pi\epsilon_0 r}$$

Now to the promise. Suppose several point charges are present together, Q_1 , Q_2 , Q_3 and so on, lying at distances r_1 , r_2 , r_3 from some point of interest. *What is the potential there?* Because potential is energy per unit charge, and energies simply add, the potential of the whole group is the ordinary sum of the potentials each charge would produce alone:

$$V = \frac{kQ_1}{r_1} + \frac{kQ_2}{r_2} + \frac{kQ_3}{r_3} + \dots$$

This is the **superposition of potentials**, and it is worth pausing on how much lighter it is than the superposition of fields. To combine fields, every field's direction had to be tracked, the vectors resolved into components, the components added separately and the whole reassembled, with an angle to compute at the end. To combine potentials there are no directions and no components: each charge offers a single number, $\frac{kQ}{r}$, and the numbers are simply added. The one care needed is with signs, *for a negative charge contributes a negative potential, and that minus sign must be carried faithfully into the sum*. A positive and a negative charge at equal distances from a point contribute potentials that cancel each other exactly, leaving nothing.

An example will show the method at work, and the figure sets the scene: four equal positive charges, one pinned at each corner of a square, with the potential to be found at the dead centre.

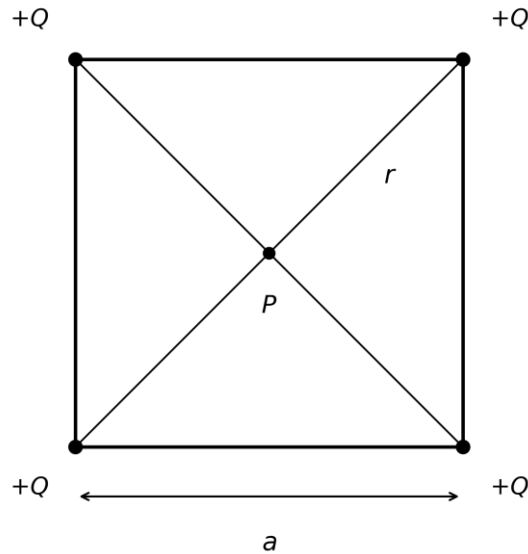


Figure: Four equal positive charges, one at each corner of a square of side a . The centre P lies the same distance r from every corner, so each charge contributes an equal share to the potential there.

The potential difference between two finite points

The result $V = \frac{kQ}{r}$ just settled is the **absolute** potential at a point distance r from a point charge Q , taken with the conventional zero of potential at infinity. In practice, however, an absolute potential is not the natural quantity. The thing that does the physical work, that drives charges through a wire and accelerates them in a field, is the **difference** of potential between two points. It is the difference that the equation $V_{AB} = V_A - V_B$, met earlier, defines, and it is the difference that a voltmeter actually measures. The absolute potential is a special case of the difference, recovered when one of the two points is sent to infinity.

For a point charge Q at the origin and two field points A and B at radial distances r_A and r_B from it, the potential difference follows directly from $V = kQ/r$ applied at each point:

$$V_{AB} = V_A - V_B = \frac{kQ}{r_A} - \frac{kQ}{r_B} = kQ \left(\frac{1}{r_A} - \frac{1}{r_B} \right)$$

This is the work per unit positive charge that an external agent must do to carry a charge from B to A . It is finite for any two finite radial distances. Sending B to infinity, the reciprocal $\frac{1}{r_B}$ vanishes ($\frac{1}{\infty} \approx 0$) and the equation collapses to the absolute potential $V_A = \frac{kQ}{r_A}$.

The absolute potential is therefore just a difference in which one of the two points has been pushed where its contribution can be safely ignored. The version with two finite points is the one that survives in distributions for which infinity is not a sensible reference, the infinite line of charge being the most important example.

Four charges, one tidy square, and a question that the scalar nature of potential turns from a chore into barely three lines of arithmetic.

BINDER Example 19

Four equal point charges, each of $8.0 \times 10^{-9}\text{C}$, are fixed one at each corner of a square of side 0.40m . Find the electric potential at the centre of the square. Take $k = 8.99 \times 10^9 \text{Nm}^2/\text{C}^2$.

Solution

The centre of a square lies the same distance from all four of its corners. That distance is half the length of the diagonal, $r = \frac{a}{\sqrt{2}}$, which for a side of 0.40m comes to $r = \frac{0.40\text{m}}{\sqrt{2}} = 0.28\text{m}$.

Electric potential is a scalar, so the potential at the centre is simply the sum of the four potentials, one from each corner charge. The four charges are equal and all lie the same distance r from the centre, so the four contributions are identical, and the total is four times any one of them:

$$V = 4 \frac{kQ}{r}$$

Substituting:

$$V = \frac{4 \times 8.99 \times 10^9 \text{Nm}^2/\text{C}^2 \times 8.0 \times 10^{-9} \text{C}}{0.28 \text{m}} = 1.0 \times 10^3 \text{V}$$

The electric potential at the centre of the square is about $1.0 \times 10^3 \text{V}$, that is one kilovolt.

Making Sense of the Answer: It is worth setting this result beside what the electric field does at the very same point. The four charges are equal and sit symmetrically about the centre, so the four field vectors there are equal in length and point in four opposing directions; they cancel in perfect pairs, and the field at the centre is exactly zero. The potential at that same point is a thousand volts. Field nothing, potential everything, and no contradiction in it. The field is a vector, and four vectors set symmetrically can sum to nothing; the potential is a scalar, and four positive numbers can only pile up.

Think Like a Physicist: Let the scalar nature of potential shape how you plan a problem. Whenever several charges are in play, ask first what the question truly wants. If it asks for an energy, a work done, or a speed gained, it wants the potential, and the potential is a single number got by plain addition. If it insists on the field, on the direction a force will point, then the vectors cannot be dodged. But here is the move worth keeping: even when the field is what you finally want, it is often easier to find the potential everywhere first, as one scalar map laid over space, and then read the field off from the way that potential changes from place to place. The field is the slope of the potential, a relationship glimpsed in the opening section and soon to be made exact. Scalar first, vector afterwards, is for a great many problems the shorter road.

Electric potential of a finite line of charge

Every potential met so far in this section has been built up from point charges. The same recipe extends bodily to a continuous distribution: cut the distribution into infinitesimal pieces, each effectively a point charge of charge dq , compute the potential each piece would contribute at the field point using the point-charge formula, and add the contributions by integral. As the cleanest illustration of the method, take a thin straight rod of length $2a$, uniformly charged with linear density λ , and let P sit at perpendicular distance x from the midpoint of the rod, as the figure shows.

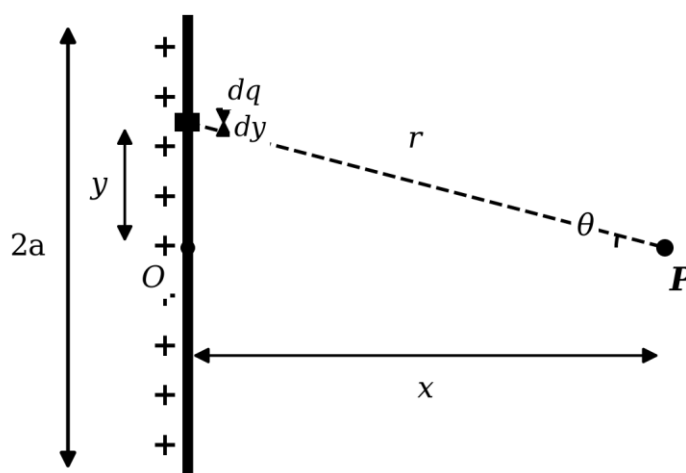


Figure: Geometry for the finite line of charge. The same geometry used earlier. The element of length dy at height y carries charge $dq = \lambda dy$ and sits at distance r from P . The potential it contributes at P is $dV = k dq / r$, a scalar that adds simply to the contributions of every other element along the rod.

Each element of length dy along the rod carries the small charge $dq = \lambda dy$, and lies at distance $r = \sqrt{x^2 + y^2}$ from P . The potential this single element would contribute at P , treating it as a tiny point charge, is the point-charge formula applied to its charge:

$$dV = \frac{k dq}{r} = \frac{k\lambda dy}{\sqrt{x^2 + y^2}}$$

Adding the contributions of every element along the rod, with y running from $-a$ to a :

$$V = k\lambda \int_{-a}^a \frac{dy}{\sqrt{x^2 + y^2}} = k\lambda \left[\ln \left(y + \sqrt{x^2 + y^2} \right) \right]_{-a}^a = k\lambda \ln \left(\frac{\sqrt{x^2 + a^2} + a}{\sqrt{x^2 + a^2} - a} \right)$$

Substituting $k = \frac{\lambda}{4\pi\epsilon_0}$:

$$V = \frac{\lambda}{4\pi\epsilon_0} \ln \left(\frac{\sqrt{x^2 + a^2} + a}{\sqrt{x^2 + a^2} - a} \right)$$

This is the electric potential at perpendicular distance x from the midpoint of a uniformly charged rod of length $2a$, taken with the usual reference of zero potential at infinity. Two short remarks close the case.

- The result depends only on the perpendicular distance from the midpoint, not on the orientation of the rod in space; the potential is a scalar, and the contributions of every element add by simple arithmetic, free of any worry about directions.
- The formula degenerates in an instructive way when the rod is allowed to grow infinitely long, a tending to infinity. The argument of the logarithm grows without bound, and the potential at any finite point P diverges.

From the finite line to its infinite-line difference

When the finite rod is allowed to grow infinitely long, a tending to infinity, the absolute potential of the previous derivation diverges, as already noted. The remedy is based on the fact that: the natural quantity is no longer the absolute potential but the difference of potentials between two points, and that difference is finite even when each of its terms, taken on its own, is not. The line of charge is the showcase example.

For an infinite straight line carrying linear charge density λ , the field at perpendicular distance r from the line is, as derived earlier,

$$E = \frac{\lambda}{2\pi\epsilon_0 r}$$

Take two field points at perpendicular distances r_a and r_b from the line. The potential difference between them follows from the general relation $V = -\int \vec{E} \cdot d\vec{r}$, integrated radially from the second point to the first:

$$V_a - V_b = -\int_{r_b}^{r_a} \frac{\lambda}{2\pi\epsilon_0 r} dr = -\frac{\lambda}{2\pi\epsilon_0} [\ln r]_{r_b}^{r_a} = \frac{\lambda}{2\pi\epsilon_0} \ln \left(\frac{r_b}{r_a} \right)$$

The result is perfectly finite for any two finite radii. The logarithm grows only slowly and never diverges away from the special value $r = 0$ of the line itself. The divergence of the absolute potential, which looked like a failure of the physics when it first appeared, was nothing of the sort; it was simply a sign that infinity was the wrong reference for this distribution, and any other finite reference will do. The line of charge is the prototype for every distribution of infinite extent.

The conducting cylinder, as a corollary

A natural choice of reference for the line of charge is the surface of the line itself, when the line is realised in practice as a long conducting cylinder of radius R , uniformly charged. Setting the potential at the cylinder's surface to zero, $V(R) = 0$, and applying the difference formula with $r_b = R$ and $r_a = r$ for any $r > R$:

$$V(r) = \frac{\lambda}{2\pi\epsilon_0} \ln \left(\frac{R}{r} \right)$$

This is the potential at perpendicular distance r from the axis of a long charged conducting cylinder, taken with the cylinder's own surface as the zero of potential. The formula is negative for any $r > R$ (the logarithm of a fraction less than one is negative), which is the right sign: the potential falls as one moves away from a positively charged conductor, and falling below the zero set at the surface means becoming negative. Inside the conductor itself, the field is zero throughout, so the potential is constant; it must therefore take everywhere within the metal the value it had on the surface, which is the zero we chose.

Electric potential of a charged conducting sphere

One further continuous distribution is worth its turn before the dipole, and a familiar one. The conducting sphere whose field was settled in the previous section is also a continuous distribution of charge, and its potential, by the relation $V = -\int \vec{E} \cdot d\vec{r}$ that we have just been working with, follows from the field with no fresh effort. The result is worth seeing because it places side by side, in a single picture, the most useful contrast in all of electrostatics: a field that jumps abruptly at a surface and a potential that does not.

Take a solid conducting sphere of radius R , carrying a total positive charge Q , left in electrostatic equilibrium. The field, recalled from the previous section, is the field of a point charge Q sitting at the centre outside the sphere, and is zero inside the conductor:

$$E = \frac{kQ}{r^2} \quad \text{for } r > R, \quad E = 0 \quad \text{for } r \leq R$$

The potential at any point in space, taken with the conventional zero at infinity, follows from integrating this field along a path that runs in from infinity to the point in question. Two regions need separate attention.

For a point outside the sphere, at distance r from the centre with $r > R$, the integration is exactly the integration that was done for a point charge. The field along the inward path is kQ/r'^2 , and

$$V(r) = -\int_{\infty}^r \frac{kQ}{r'^2} dr' = \frac{kQ}{r}$$

From outside, then, a conducting sphere looks like a point charge of total Q at its centre, just as far as the potential is concerned as it does as far as the field is concerned. At the surface itself, $r = R$, this gives:

$$V(R) = \frac{kQ}{R}$$

For a point inside the conductor, $r < R$, the field is zero throughout. The change in potential between any two points inside the conductor is therefore zero as well, since

$$dV = -\vec{E} \cdot d\vec{r} = 0 \quad \text{inside.}$$

The potential cannot change anywhere within the body of the conductor, and must take everywhere the value it had at the surface. The whole interior of the conductor sits at the constant potential.

$$V = \frac{kQ}{R} \quad \text{for } r \leq R$$

Drawing the two regions together: the potential of a charged conducting sphere is constant at the value $\frac{kQ}{R}$ throughout its whole interior, falls smoothly as kQ/r outside, and joins continuously from one regime to the other at the surface $r = R$. The figure draws the picture, and sets the potential alongside the field for direct comparison.

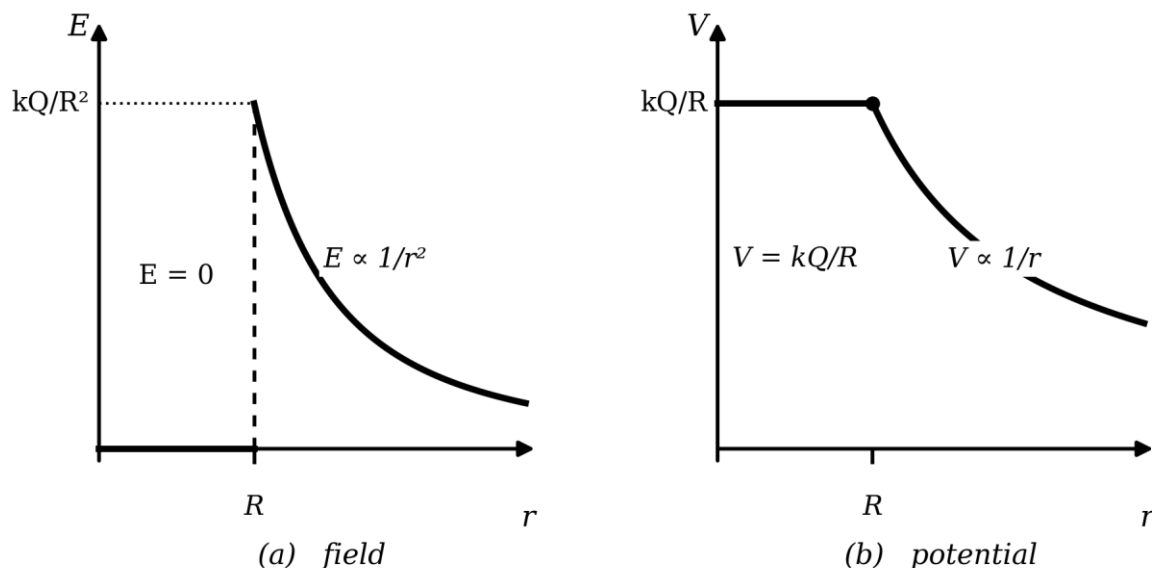


Figure: Electric field and electric potential of a charged conducting sphere. Panel (a): the field is zero inside the conductor, jumps abruptly to its peak value kQ/R^2 at the surface, and falls as the inverse square outside. Panel (b): the potential is constant at kQ/R throughout the interior, joins continuously to the curve kQ/r at the surface, and falls as the inverse first power outside. The field is discontinuous at the surface; the potential is not.

The two graphs are worth looking at together, because they tell the same story in two different scripts.

- Outside the sphere both quantities fall away with distance, the field as the inverse square and the potential as the inverse first power, the slower fall reflecting the fact that the potential is the running sum of the field's effect inward from infinity.
- At the surface itself the two part company in a way worth noticing. The field is discontinuous there: it leaps from zero just inside to its peak value kQ/R^2 just outside, a jump that records the thin layer of charge mathematically present on the idealised conductor's surface. The potential, by contrast, is perfectly continuous at the surface: the constant $\frac{kQ}{R}$ carried out from the interior matches exactly the value the outside formula kQ/r delivers at $r = R$, and the curve sails through the surface without a kink. The contrast is precisely the one that the next section will name: the field is the slope of the potential, and a potential with a finite step in its slope answers with a finite jump in the field.

This is also the formal underpinning of the puzzle that runs through the next section. Inside a hollow charged conductor, met earlier in the catalogue of Gauss applications, the field is zero but the potential is emphatically not. The conducting cavity sits at the constant surface value kQ/R_2 , a region of zero field at non-zero potential that finds its cleanest statement in the present formula and its cleanest picture in the present figure. Field and potential, the chapter is beginning to say, are independent enough that either can vanish without the other doing the same, and a careful student keeps both in view at every step.

So far the charges added have all carried the same sign, and four positive numbers can do nothing but pile up. The next arrangement is the most important of all the simple ones, and the first to set a positive charge against a negative one: a single pair, plus and minus, held a small fixed distance apart. It is called the electric dipole, and superposing its two potentials, one positive and one negative, yields a result whose shape is surprising, widely useful, and the perfect illustration of how a field and a potential can part ways.

The electric dipole

The previous section closed on a promise of the most instructive of all the simple charge arrangements. This section makes good on it.

A positive charge and a negative charge, equal in magnitude and held a small fixed distance apart: this modest pair is called the **electric dipole**. Modest it may look, but it is one of the most important arrangements in all of physics. Every water molecule is a tiny dipole, its oxygen end slightly negative and its hydrogen end slightly positive, and that single fact is why water dissolves so much and behaves so strangely. Radio

aerials are driven dipoles. An atom placed in an electric field stretches into a dipole. The arrangement repays close study.

Two numbers fix a dipole: the **size of the charge, q** , and the **separation between the charges, d** . The two are bundled into a single quantity, the **dipole moment**, written **$\mathbf{p}$** and defined as *the product of either charge and the separation vector between the charges in an electric dipole*.

Mathematically:

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

The dipole moment is a vector: it points along the line joining the charges, *from the negative charge towards the positive*, and its size qd measures how strong the dipole is. Everything a dipole does at a distance, it does through its dipole moment $\mathbf{p}$.

Two questions now press: *what potential does a dipole set up in the space around it, and what field?* We shall answer both, at two telling places: along the dipole's **axis**, the line through both charges, and along its **perpendicular bisector**, the line that cuts the dipole in half at right angles. The figure marks them.

A point A sits on the axis and a point B on the bisector, each a distance r from the centre of the dipole. Throughout, we take r to be large compared with the small separation d , which is the case of real interest.

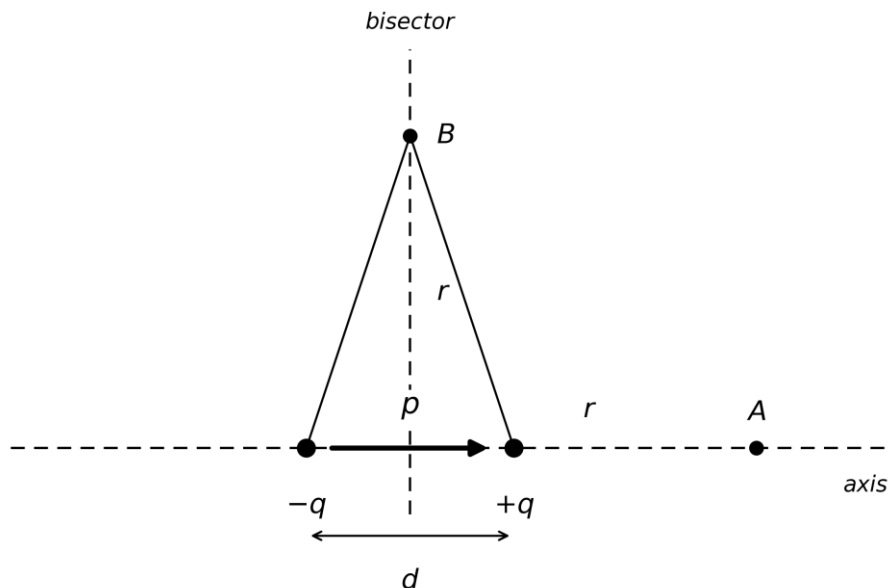


Figure: An electric dipole: charges $+q$ and $-q$ a distance d apart, with dipole moment $\mathbf{p}$ pointing from the negative charge to the positive. The point A lies on the axis and B on the perpendicular bisector, each a distance r from the centre.

Begin with the potential at A, on the axis. The point stands a distance $r - \frac{d}{2}$ from the nearer charge, the positive one, and $r + \frac{d}{2}$ from the farther charge, the negative one. Potential is a scalar, so the potential at A is the plain sum of the two contributions, each carrying the sign of its charge:

$$V = \frac{kq}{\left(r - \frac{d}{2}\right)} - \frac{kq}{\left(r + \frac{d}{2}\right)} = \frac{kqd}{\left(r - \frac{d}{2}\right)\left(r + \frac{d}{2}\right)} = \frac{kqd}{r^2 - \frac{d^2}{4}}$$

Now the case of interest. When A is far from the dipole compared with its tiny separation, r much greater than d , the term $\frac{d^2}{4}$ is negligible beside r^2 and may be dropped. Writing the dipole moment $\mathbf{p} = q\mathbf{d}$ in place of qd , the potential on the axis of a dipole is:

$$V = \frac{kqd}{r^2} = \frac{k\mathbf{p}}{r^2}$$

Notice the inverse-square falloff. A single point charge's potential dies away as $\frac{1}{r}$; a dipole's, on its axis, dies faster, as $\frac{1}{r^2}$. The reason is the dipole's neutrality: from far off its two charges very nearly cancel, and only their slight separation keeps any potential alive at all.

Now the bisector, and the point B. Here something different happens. B lies on the line that cuts the dipole symmetrically in half, so it is exactly the same distance from the positive charge as from the negative one. Call that common distance s . The positive charge raises the potential at B by $\frac{kq}{s}$; the negative charge lowers it by precisely the same $\frac{kq}{s}$; and the two cancel without remainder:

$$V = \frac{kq}{s} - \frac{kq}{s} = 0$$

This zero is exact, and no approximation was made. It holds not at one special point but at every point along the whole perpendicular bisector, near the dipole or far from it. The bisector is a line of zero potential.

So much for the potential. We turn now to the field, and we shall let the bisector spring its surprise in a moment.

First the axis. The field at A is found by the same superposition, but now of the two fields rather than the two potentials, and a field is a vector. Both fields here lie along the axis: the positive charge pushes outward, away from the dipole, with magnitude $\frac{kq}{(r-\frac{d}{2})^2}$; the negative charge pulls inward, back towards itself, with magnitude $\frac{kq}{(r+\frac{d}{2})^2}$. Taking the outward direction as positive, the net field is their difference:

$$E = \frac{kq}{\left(r - \frac{d}{2}\right)^2} - \frac{kq}{\left(r + \frac{d}{2}\right)^2} = \frac{kq \left(\left(r + \frac{d}{2}\right)^2 - \left(r - \frac{d}{2}\right)^2 \right)}{\left(r - \frac{d}{2}\right)^2 \left(r + \frac{d}{2}\right)^2} = \frac{2kqrd}{\left(r^2 - \frac{d^2}{4}\right)^2}$$

Once more taking r much larger than d , the denominator becomes simply r^4 , and with $p = qd$ the field on the axis of a dipole is:

$$E = \frac{2kqd}{r^3} = \frac{2kp}{r^3}$$

The falloff is faster still, as $\frac{1}{r^3}$, and the field points the same way as the dipole moment p .

And now the bisector. The potential there was zero; *is the field zero too?* It is not. At B each charge produces a field of magnitude $\frac{kq}{s^2}$, the positive charge's pointing away from it and the negative charge's pointing towards it. Resolve each of these two fields into a part along the axis and a part along the bisector. The parts along the bisector are equal in size and opposite in direction, one up and one down, and cancel exactly. The parts along the axis are equal in size and point the same way, and so they add. Each axial part is the field magnitude $\frac{kq}{s^2}$ times the cosine of its angle to the axis, which the geometry fixes at $\frac{d}{2s}$. Two equal axial parts

therefore give, with $s = \sqrt{r^2 + \frac{d^2}{4}}$:

$$E = 2 \times \frac{kq}{s^2} \times \left(\frac{d}{2s}\right) = \frac{kqd}{s^3}$$

With r much greater than d the distance s becomes r , and with $p = qd$ the field on the bisector of a dipole is:

$$E = \frac{kqd}{r^3} = \frac{kp}{r^3}$$

It points along the axis but antiparallel to the dipole moment; it is therefore opposite to the field on the axis, and exactly half that size at the same distance.

Set the two bisector results side by side, for together they carry the lesson this section has been driving towards.

Along the perpendicular bisector of a dipole the potential is zero, exactly and everywhere. Along that very same line the electric field is $\frac{kp}{r^3}$, and most certainly not zero. A student who argues that the potential here is zero and so the field here must be zero has made a real and tempting mistake. Potential is energy per unit charge; field is force per unit charge; they are different quantities, answering different questions, and a place may be empty of the one while full of the other. The bisector of a dipole is the plainest proof of it in all of electrostatics: zero potential, and a living field, on the very same line.

A short calculation will fix the zero-potential result in the mind, by checking it at the one point easiest to reach: the dipole's own midpoint.

BINDER Example 20

Two point charges, $+3.0 \times 10^{-9}\text{C}$ and $-3.0 \times 10^{-9}\text{C}$, are fixed 20cm apart. Find the electric potential at the midpoint of the line joining them. Take $k = 8.99 \times 10^9 \text{N} \cdot \text{m}^2/\text{C}^2$.

Solution

The midpoint lies halfway between the charges, a distance of 10cm, or 0.10m, from each. Potential is a scalar, so the potential at the midpoint is the sum of the two contributions, one from each charge, each carrying the sign of its charge:

$$V = \frac{k(+q)}{r} + \frac{k(-q)}{r}$$

Both charges lie the same distance from the midpoint and are equal in magnitude, so the two contributions are equal in size; but one charge is positive and the other negative, so the two are opposite in sign. Substituting $3.0 \times 10^{-9}\text{C}$ and $r = 0.10\text{m}$:

$$V = \frac{k(3.0 \times 10^{-9}\text{C})}{0.10\text{m}} - \frac{k(3.0 \times 10^{-9}\text{C})}{0.10\text{m}} = 0$$

The potential at the midpoint is zero.

Making Sense of the Answer: *The zero is worth dwelling on. It did not come about because the charges are far away or weak; each charge, on its own, sets up a healthy 270 volts at the midpoint. It came about because the two contributions are equal and opposite, and a scalar sum lets them cancel completely. And the midpoint is nothing special: it is simply the most convenient point on the perpendicular bisector at which to test. Every other point on that bisector would have given zero by the identical argument, the two charges always equidistant and their potentials always equal and opposite. This single calculation stands for a whole line of zeros. Yet, as the section showed, the electric field marches straight through that line of zero potential without itself vanishing at all.*

Think Like a Physicist: *Carry one habit away from the dipole: keep the sign of every charge with it, always, into every potential sum. The whole result here, the clean cancellation to zero, lives or dies on the minus sign of the negative charge. Drop it, treat both contributions as positive, and the midpoint potential comes out as 540 volts instead of nothing: a wrong answer produced by a careless hand rather than a wrong method. Potential is a scalar, which makes it easy; but a scalar still carries a sign, and the sign is not optional. When charges of both kinds are present, write each contribution as the charge, complete with its sign, divided by its distance and multiplied by k , and let the arithmetic take care of the rest.*

The dipole has shown how potentials combine, signs and all, and how sharply a dipole's reach falls away with distance. One large question about potential now remains, and it is the one that ties this whole section back to energy. We have spoken throughout of the work done in moving a charge, and of energy carried per unit charge; it is time to speak of the energy itself: the electric potential energy stored when charges are assembled against their will, and released when they are let fly apart. That is the business of the section's final part.

Electric potential energy

The section's final task is the one its very first definition pointed towards. Potential was introduced as energy carried per unit charge; it remains now to speak of that energy itself.

Every charge that sits in an electric field holds a store of energy by virtue of where it sits, and that store is its **electric potential energy**. The potential introduced at the start of this section delivers it at once. Potential was defined as energy per unit charge, which is to say that the potential V at a point is the potential energy

belonging to each coulomb of charge placed there. A charge q is q coulombs, so its potential energy is q times the energy carried by each:

$$U = qV$$

The unit is the joule, as a charge in coulombs multiplied by a potential in volts is bound to give. The sign looks after itself. A positive charge at a point of positive potential has positive energy, the signature of a charge pushed into place against a repulsion, which would spring back the moment it were freed. A positive charge at a point of negative potential has negative energy, the signature of a charge that is bound, one that could be dragged away to infinity only if work were done on it from outside.

The most important case is also among the simplest. It is not one charge sitting in another's field, but two charges together, and the energy that belongs to the pair of them. Take two point charges, Q_1 and Q_2 , a distance r apart, and ask how much energy is stored in that arrangement. The honest way to count it is to ask how much work was needed to build the arrangement at all, starting with the two charges infinitely far apart.

Bring in Q_1 first. Empty space resists it not at all, so placing the first charge costs no work. With Q_1 in position it has spread a potential through the space around it, and at the distance r where Q_2 is to go that potential has the value $\frac{kQ_1}{r}$. Carrying Q_2 in to that spot costs work equal to Q_2 multiplied by the potential already waiting there, and that work is the potential energy of the pair:

$$U = Q_2 \times \frac{kQ_1}{r} = \frac{kQ_1 Q_2}{r}$$

The result makes no distinction between the two charges. Bringing Q_1 in to the potential of Q_2 instead would have produced the identical $\frac{kQ_1 Q_2}{r}$; the energy belongs to the pair, not to either charge by itself. Its sign tells the pair's whole story.

Two like charges, both positive or both negative, give a positive product $Q_1 Q_2$ and so a positive energy: work was spent forcing them together against their mutual repulsion, and that work is stored, ready to be handed back the instant they are released. Two unlike charges give a negative product and a negative energy: the pair is bound, and energy would have to be supplied from outside before they could be parted to infinity. Positive energy is a coiled spring; negative energy is a knot.

Three charges, or four, or more, bring no new principle. The energy of any collection of point charges is the sum of the energies of all the separate pairs that can be drawn within it. For three charges (Q_1, Q_2, Q_3) there are three pairs, and the total energy is the three pair-energies added together:

$$U = \frac{kQ_1 Q_2}{r_{12}} + \frac{kQ_1 Q_3}{r_{13}} + \frac{kQ_2 Q_3}{r_{23}}$$

Each term is one pair's $\frac{kQ_i Q_j}{r}$, carrying its own two charges and its own separation. The figure shows the three pairs of a triangle of charges. Four charges would furnish six pairs, five charges ten, and the recipe never alters: find every distinct pair, price it with $\frac{kQ_i Q_j}{r}$, and add.

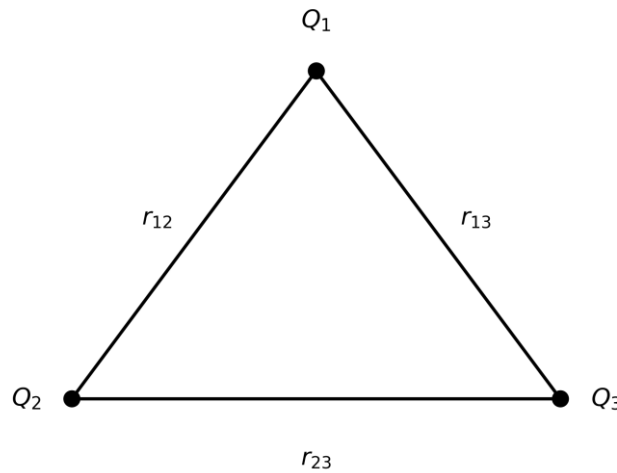


Figure: Three point charges form three pairs. The energy of the system is the sum of the three pair-energies, one for each side of the triangle.

One caution attaches to this counting, and it is the source of a factor of one-half that often appears alongside the formula. The pair-by-pair sum just described counts each pair exactly once, and is correct as it stands. A tempting alternative is to visit each charge in turn, find the potential at its location due to all the others, multiply by the charge itself to obtain its qV , and add those products. That alternative counts every pair twice, once at each of the two charges that form it. To undo the double-counting the whole sum must be halved, and that halving is the one-half. The pair-by-pair count needs no such repair, and is the safer of the two.

A last and very practical question closes the section: *how much energy does it take to charge a conductor, to gather a total charge Q onto an isolated metal sphere of radius R ?*

The charge cannot arrive all at once. It must be carried on in small instalments, and each instalment proves harder to add than the one before. A sphere, seen from its surface or from beyond, behaves exactly as a point charge at its centre would, so while it carries a charge q its surface rests at the potential $\frac{kq}{R}$. Adding the next small instalment dq means pushing it onto a sphere already held at that potential, and the work for that one small step is the potential multiplied by the small charge, $dW = V dq$. The total work is the sum of every such step as the charge builds from nothing to its final value Q , which is an integral:

$$W = \int_0^Q V dq = \int_0^Q \left(\frac{kq}{R} \right) dq$$

The constant k and the radius R both come outside the integral, and the integral of q is $\frac{q^2}{2}$:

$$W = \left(\frac{k}{R} \right) \int_0^Q q dq = \left(\frac{k}{R} \right) \left[\frac{q^2}{2} \right]_0^Q$$

Evaluating the bracket between its limits, $\frac{q^2}{2}$ is $\frac{Q^2}{2}$ at the upper limit and zero at the lower, and the energy stored in charging the conductor is:

$$W = \frac{kQ^2}{2R}$$

Read what the integral has done. The stored energy grows not as the charge Q but as its square. Double the charge carried by the sphere and four times the energy is stored, not twice. The reason lies in that climbing potential: the first faint instalments of charge are added almost for nothing, onto a sphere of low potential, while the final instalments must be forced onto a sphere whose potential has risen high. The average cost across the whole process comes to exactly half the final potential, and that one-half is why the energy is $\frac{kQ^2}{2R}$.

and not $\frac{kQ^2}{R}$. The same one-half can be seen directly: the result is one-half of the final charge Q multiplied by the final surface potential $\frac{kQ}{R}$.

A short and concrete calculation will set the energy of a pair to work.

BINDER Example 21

Two point charges, $+5.0 \times 10^{-9}\text{C}$ and $+2.0 \times 10^{-9}\text{C}$, are held 30cm apart. Find the electric potential energy of the pair. Take $k = 8.99 \times 10^9 \text{Nm}^2/\text{C}^2$.

Solution

The potential energy of a pair of point charges is $\frac{kQ_1Q_2}{r}$, each charge entered complete with its sign and r the separation between them:

$$U = \frac{kQ_1Q_2}{r}$$

Both charges are positive, and the separation of 30cm is 0.30m. Substituting, with every quantity in SI units:

$$U = \frac{8.99 \times 10^9 \text{Nm}^2/\text{C}^2 \times 5.0 \times 10^{-9}\text{C} \times 2.0 \times 10^{-9}\text{C}}{0.30\text{m}} = 3.0 \times 10^{-7}\text{J}$$

The potential energy of the pair is about $3.0 \times 10^{-7}\text{J}$.

Making Sense of the Answer: *The energy is positive, and that positive sign is the physics of the answer. Both charges are positive, so they repel; to have brought them as close together as 30 centimetres, work had to be spent against that repulsion, and the result is precisely that spent work, now held in the arrangement. Release the charges and the store is repaid in full: they fly apart, and the energy reappears as kinetic energy. Had one charge been negative, the product of the two charges would have been negative, the energy negative with it, and the meaning turned around: a bound pair, holding together, needing energy supplied before it could ever be separated.*

Think Like a Physicist: *Two habits from this section are worth carrying onward. The first is that potential energy, like potential itself, is a scalar, and it is counted pair by pair. When a problem sets several charges before you, do not hunt for a single clever formula; find every pair, price each with the pair formula, and add the prices, signs and all. The second is to watch for the square. The energy stored in an arrangement of charge very often grows as the square of that charge, because the act of building the charge up raises the very potential it must be built against. Whenever doubling a charge is found to quadruple an energy, the integral met in this section is quietly at work behind the result.*

With this the section on electric potential and potential energy is complete. Potential difference, the potential of a single point charge, the potential and field of a dipole, and the energy stored in assembled charge have all been set out and worked through. What follows is a gathering of further worked examples, the real-world cases and the harder ones, drawing on everything the section has built.

Further worked examples on electric potential and potential difference

The four parts behind us built the tools; the examples that follow are where the tools are made to work. The physics does not change in them. The questions simply grow more demanding, and one or two will look, at first sight, as though nature has broken its own rules. It has not.

REAL Example 22

A bird settles on a single bare conductor of a high-voltage transmission line and perches there, both feet gripping the one wire, entirely unharmed, even though that wire may be carrying electricity at many thousands of volts. A person who took hold of the same wire while standing on the ground would be killed at once. Explain, in terms of electric potential, why the bird is safe.

Solution

Whether a current is driven through a body depends not on the potential of that body but on the potential difference between its parts. The potential difference between two points is the work needed, per unit charge, to carry charge from one to the other; where there is no potential difference, no work is done on the charge, and nothing is there to drive a current.

The bird grips the wire with two feet set only a few centimetres apart. The wire is an excellent conductor, and along so short a length of good conductor there is almost no potential difference at all: the bird's two feet rest at the same potential, whatever great value that potential may hold relative to the distant ground. With no potential difference between its feet, no work is done on the charges in the bird's body, no current flows through it, and the bird feels nothing.

The bird is safe not because the wire is at a low potential, for it is not, but because both of its feet touch the wire at one and the same potential. The person on the ground stands in a wholly different case. One hand grips the wire at a potential of thousands of volts while the feet rest on the earth at a potential near zero. The body then bridges an enormous potential difference, a large current is driven through it from hand to foot, and the outcome is fatal. In electricity the danger lives in the potential difference, never in the potential alone.

REAL Example 23

The safety advice for anyone caught in the open during a fierce thunderstorm, or who comes upon a power line lying broken on the ground, is to crouch low with the two feet held tightly together, and on no account to walk normally or to run. Using the idea of electric potential, explain why feet held together are safer than feet set apart.

Solution

When a large current pours into the earth at a single place, where lightning strikes the ground or where a live wire falls upon it, that point of entry is raised to a very high potential. The earth is not a perfect conductor, so the potential does not stay high everywhere. It falls away steeply with distance from the point of entry, leaving the ground around that point carrying a slope of potential: high close to the entry, lower and lower further off.

A person standing on that sloping ground with the feet apart has each foot resting at a different potential, since each foot stands a different distance from the point of entry. The difference between those two potentials appears across the person's legs, and if it is large it drives a dangerous current up one leg and down the other.

Crouching with the feet pressed together takes the danger away. Two feet in nearly the same place rest at nearly the same potential; the potential difference between them falls almost to zero, and with no potential difference to drive it, little current passes through the body. The warning never to run carries the same physics: a long running stride sets the two feet far apart on the slope of potential, across a large potential difference.

This is the bird on the wire seen from the other side. The bird is safe because both its feet share a single potential, while the person is endangered because their feet, set apart on sloping ground, do not.

BINDER Example 24

Inside a uranium-238 nucleus two protons may sit a distance of about $6.0 \times 10^{-15}\text{m}$ apart. Treating each proton as a point charge of $+1.6 \times 10^{-19}\text{C}$, calculate the electric potential energy of this pair of protons. Take $k = 8.99 \times 10^9 \text{Nm}^2/\text{C}^2$.

Solution

The electric potential energy of a pair of point charges is the product of the constant k with the two charges, divided by their separation:

$$U = \frac{kQ_1Q_2}{r}$$

Both charges are protons, so each is $1.6 \times 10^{-19}\text{C}$, and the separation is $6.0 \times 10^{-15}\text{m}$. Substituting:

$$U = \frac{8.99 \times 10^9 \text{N} \cdot \text{m}^2/\text{C}^2 \times (1.6 \times 10^{-19}\text{C})^2}{6.0 \times 10^{-15}\text{m}} = 3.8 \times 10^{-14}\text{J}$$

The electric potential energy of the pair of protons is about $3.8 \times 10^{-14}\text{J}$.

HOT Example 25

Four equal positive point charges, each of magnitude q , are carried in from infinity and fixed one at each corner of a square of side a . Find the total work that must be done to assemble them.

Solution

The total work done is equal to the electric potential energy of the finished arrangement.

The potential energy of a system of point charges is the sum of $\frac{kQ_1Q_2}{r}$ taken over every distinct pair of charges. Four charges form six distinct pairs. As the figure shows, four of these pairs lie along the sides of the square, separated by the side length a , and the remaining two lie along the diagonals, separated by the diagonal length $a\sqrt{2}$.

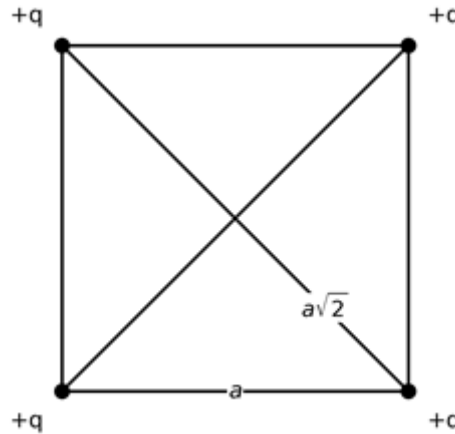


Figure: Four charges at the corners of a square form six pairs: four pairs lie along the sides of the square and two along its diagonals.

Each of the four side-pairs contributes $\frac{kq^2}{a}$, the two charges of a pair being both positive and equal.

Each of the two diagonal-pairs contributes $\frac{kq^2}{a\sqrt{2}}$.

The total energy is the sum of all six pairs:

$$U = 4 \times \left(\frac{kq^2}{a} \right) + 2 \times \left(\frac{kq^2}{a\sqrt{2}} \right)$$

The two diagonal terms simplify, since 2 divided by $\sqrt{2}$ is itself $\sqrt{2}$:

$$U = (4 + \sqrt{2}) \left(\frac{kq^2}{a} \right) \approx 5.41 \left(\frac{kq^2}{a} \right)$$

HOT Example 26

A speck of dust of mass $5.0\mu\text{g}$ carries a charge of $2.0\mu\text{C}$. It is released from rest and moves freely, acted on by no force but the electric one, between two points whose potential difference is 2700V . Find the speed the dust speck reaches.

Solution

As the charged speck moves through the potential difference, the electric force does work on it. The work done on a charge q carried through a potential difference ΔV is $q\Delta V$, and since no other force acts, all of that work becomes kinetic energy. The speck starts from rest, so its whole kinetic energy at the end is $\frac{1}{2}mv^2$:

$$q\Delta V = \frac{1}{2}mv^2$$

The speed v is wanted, so it is made the subject:

$$v = \sqrt{\frac{2q\Delta V}{m}}$$

Now the numbers, each in SI units: the charge $2.0\mu\text{C}$ is $2.0 \times 10^{-6}\text{C}$, the mass $5.0\mu\text{g}$ is $5.0 \times 10^{-9}\text{kg}$, and the potential difference is 2700V :

$$v = \sqrt{\frac{2 \times 2.0 \times 10^{-6} \text{C} \times 2700 \text{V}}{5.0 \times 10^{-9} \text{kg}}} = 1.5 \times 10^3 \text{m/s}$$

The dust speck reaches a speed of about $1.5 \times 10^3 \text{m/s}$.

HOT Example 27

A charge of $+3.0 \text{nC}$ and a charge of -6.0nC are fixed 24cm apart. Find every point on the straight line through the two charges at which the total electric potential is zero.

Solution

The potential at any point is the sum of the contributions of the two charges, each given by $\frac{kQ}{r}$ and each carrying the sign of its own charge. The positive charge raises the potential; the negative charge lowers it. The total is zero only where the positive contribution of the $+3.0 \text{nC}$ charge is exactly matched by the negative contribution of the -6.0nC charge.

Writing r_1 for the distance from the $+3.0 \text{nC}$ charge and r_2 for the distance from the -6.0nC charge, this balance is:

$$\frac{k(3.0 \text{nC})}{r_1} = \frac{k(6.0 \text{nC})}{r_2}$$

The constant k and the unit nanocoulomb stand on both sides and cancel, and the balance reduces to a plain statement about the two distances:

$$r_2 = 2r_1$$

Wherever the potential vanishes, the point must lie twice as far from the -6.0nC charge as from the $+3.0 \text{nC}$ charge. Two places on the line meet this condition. The first lies between the charges. There the two distances together make up the whole separation, so $r_1 + r_2 = 24 \text{cm}$; with $r_2 = 2r_1$ this becomes $3r_1 = 24 \text{cm}$, giving $r_1 = 8.0 \text{cm}$. **One point of zero potential lies between the charges, 8.0cm from the $+3.0 \text{nC}$ charge.**

The second lies outside the pair, beyond the $+3.0 \text{nC}$ charge on the side away from the -6.0nC charge. There the greater distance exceeds the smaller by the whole separation, so $r_2 - r_1 = 24 \text{cm}$; with $r_2 = 2r_1$ this becomes $r_1 = 24 \text{cm}$. **The second point of zero potential lies 24cm from the $+3.0 \text{nC}$ charge, on its far side.**

Beyond the -6.0nC charge there is no third point: that charge is the larger and would always be the nearer, so its negative potential can never be cancelled there.

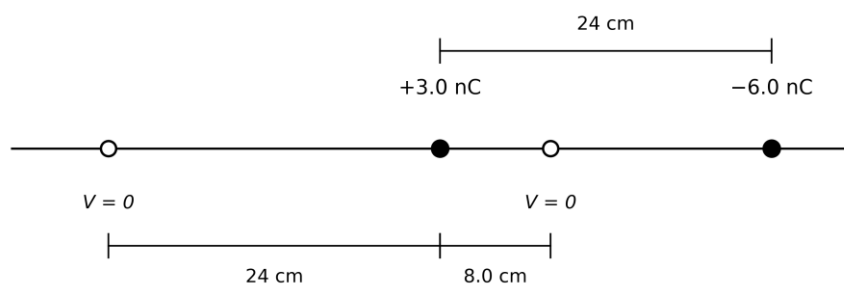


Figure: The two points of zero potential on the line of the charges: one between the charges, the other beyond the positive charge.

Make Sense of the Answer: Two points of zero potential, and it is worth ending on what they are not. At neither of them is the electric field zero. Between the charges the field of the positive charge and the field of the negative charge point the same way, and two vectors in the same direction cannot cancel. At the outer point the two fields do oppose each other, but the point stands at unequal distances from two unequal charges, so their strengths differ and again they fail to cancel. Potential is a scalar: signed numbers cancel by simple addition, and zeros of potential are common. The field is a vector: it vanishes only where strength and direction conspire, at some quite different place. The lesson the dipole first taught holds to the end of the section. A point where the potential is zero is not a point where nothing is going on.

From a bird on a wire to two protons inside a nucleus, these examples leaned on one small set of ideas: the potential of a point charge, the potential difference that alone drives a current, the energy bound up in a pair of charges and in a system of them, and the conservation of energy as a charge crosses a potential difference. Each question gave way to one of them. With this the section on electric potential and potential difference is complete, its formulas derived, its meaning drawn out, and its reach tested against the world.

The next section takes that potential one step further: it joins the points of equal potential into surfaces of their own, and shows that the electric field can be read straight from the way the potential changes from place to place.

EQUIPOTENTIAL SURFACES AND THE FIELD-POTENTIAL RELATIONSHIP

Anyone who has ever used a map of a mountainous place has already met equipotential surfaces, although under another name. On such a map the hills are drawn with thin closed lines that loop around one another, and every one of those lines is a contour: it joins together all the points on the ground that stand at the same height above the sea. Walk along a single contour line and a curious thing happens, you go neither up nor down. You may walk north for a while, then swing east, then curl back towards the south, but as long as your feet keep to the one line your height above the sea does not change by a single metre. To gain height you must do the very thing the contour forbids, you must step across it and onto the next line.

The map carries a second message, and it is hidden in the spacing of the lines. Where the contours are drawn far apart the ground rises gently, a long and easy walk between one marked height and the next. Where the contours crowd together into a tight band the ground is steep, the same gain in height now squeezed into a few hard strides. A walker who has learned to read a map never stops to measure a slope; the walker simply looks at how closely the lines run and knows at once where the climbing will be punishing and where it will be kind. The steepness of the hill is written in the spacing of its contours.

The electric potential, built up point by point in the previous section, possesses contour lines of exactly this kind. Gather together all the points in a field that happen to share one single value of potential, and they do not form a scattered dust of separate points; they join up into a continuous surface. That surface is called an **equipotential surface**, and it is the electrical twin of the contour line on the mountain map. Carry a charge along it and, as the pages ahead will show, no work is done, just as the walker on the contour gains no height. Watch how the equipotentials of a field are spaced and the strength of the field can be read straight off them, just as the walker reads the steepness of the hill from the spacing of the lines. This section follows that one picture wherever it leads, and it leads a surprisingly long way.

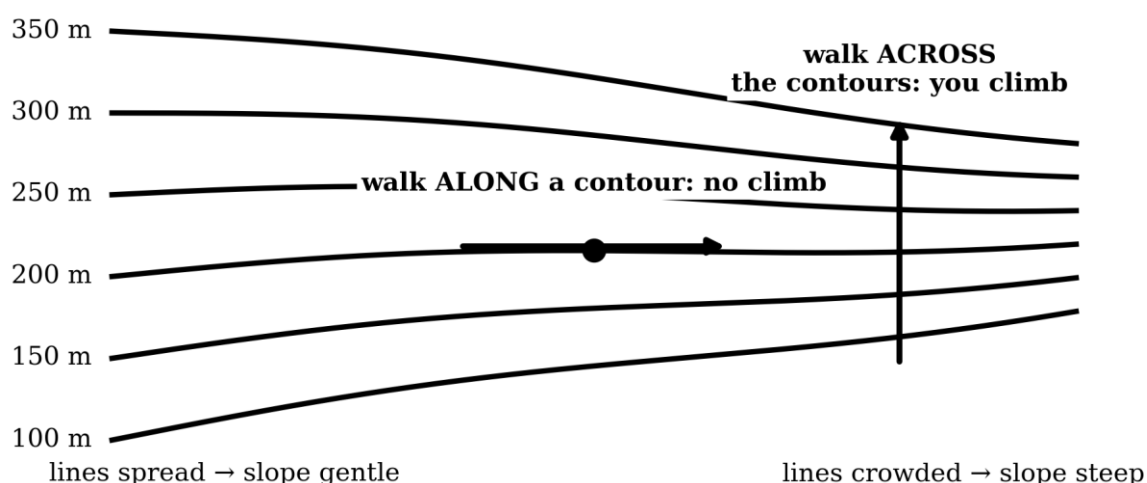


Figure: The contour-map analogy. On a map of a hill each line joins points that stand at the same height above the sea. Walking along a single line involves no climbing at all, while the spacing of the lines reveals how steep the ground is: lines crowded close together mark steep ground, lines spread far apart mark gentle ground.

Here, then, is the idea stated plainly. An **equipotential surface** is a surface on which the electric potential has the same value at every point. Move a test charge anywhere over such a surface and the potential it sits in never alters. Different surfaces carry different values, one for twelve volts, another for twenty, another for

a hundred, and the whole family of them fills the space around a charge in the same way the contour lines fill a map, each surface a sheet of constant potential nested inside or outside its neighbours.

The shape of these surfaces is not left to guesswork; it follows at once from the formula for the potential. Around a single point charge the potential is $V = \frac{kQ}{r}$, a quantity that depends on nothing whatever except the distance from the charge. Every point lying the same distance from the charge therefore carries the same potential. And the set of all points at one fixed distance from a central point is, by its very definition, a sphere. The equipotential surfaces of a point charge are thus a family of spheres, all centred on the charge, one nested inside the next like the skins of an onion. On the flat of a page they can only be drawn as circles, but the reader should always picture the full spheres standing behind the circles.

Between two large parallel charged plates the picture is different but just as simple. There the field is uniform, and the potential climbs at a perfectly steady rate as one crosses the gap from the lower plate to the higher one. Every point the same distance from a plate sits at the same potential, and the set of all points at one fixed distance from a flat plate is itself a flat plane. The **equipotentials of the uniform field** are therefore a *stack of flat planes, all parallel to the plates and evenly spaced through the gap*. The two cases, the nest of spheres and the stack of planes, are drawn together in the figure.

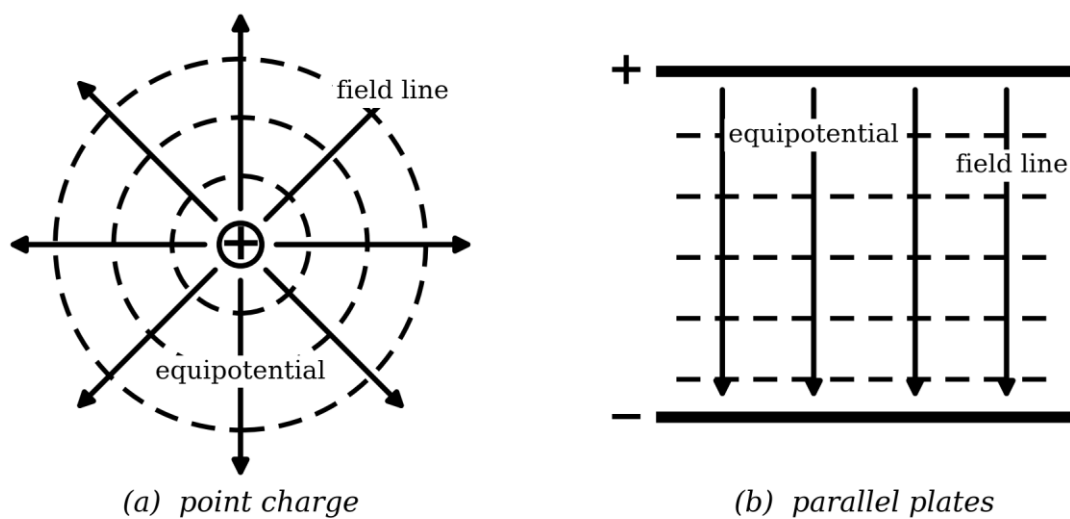


Figure: Equipotential surfaces. Around a point charge the equipotentials are concentric spheres, drawn here as circles, and the field lines run straight out along the radii. Between parallel plates the equipotentials are flat parallel planes evenly spaced through the gap. In both cases the field crosses every equipotential at right angles.

Equipotential surfaces are useful because a small handful of plain facts about them can be stated once and then leaned upon for the rest of electrostatics. Five such facts deserve to be set out carefully, each together with the reasoning that makes it true, for a fact whose reason is understood is a fact that can never afterwards be forgotten.

The first property is that *the electric field is everywhere at right angles to the equipotential surface*. To see why, suppose for a moment that it were not so, and that the field at some point on a surface had a part of itself lying flat along the surface. That flat part would be a force, and it would push a charge resting there sideways, along the surface. Let the charge move a little way in that direction, and the field, pushing as the charge moves, would do work upon it. But the charge has stayed on the equipotential, where the potential does not change, and the work done on a charge is the charge multiplied by the change of potential it passes through. With no change of potential there can be no work. The field cannot both do work and do no work on the same small journey. The only escape from the contradiction is to deny the thing that was supposed: the field has no part lying along the surface at all. Whatever the field does, it does straight out of the surface, exactly perpendicular to it. On the diagrams this shows as field lines piercing the equipotentials cleanly at ninety degrees, never grazing along them.

The second property is the one just used to prove the first, now promoted to a headline in its own right: *no work is done in moving a charge from one point of an equipotential surface to another point of the same*

surface. The reasoning is a single line. The work done by an outside agent in carrying a charge between two points is the charge multiplied by the difference of potential between those points, the relationship $W = q\Delta V$. Both points lie on the one equipotential, so they are at the same potential; the difference of potential between them is therefore zero, and the work, being the charge times that difference, is zero along with it. It does not matter how long the path is, how sharply it twists, or whether the charge is large or small, positive or negative. If the journey begins and ends on the same equipotential, the work is zero. This is the electrical echo of the walker who returns to the same contour having spent no net effort against gravity.

The third property restores the message that was hidden in the spacing of the map. *Where equipotential surfaces are crowded close together the field is strong, and where they are spread far apart the field is weak.* Imagine two neighbouring equipotentials whose potentials differ by some fixed step, say ten volts. If the two surfaces lie close together, those ten volts are used up over a very short distance, so the potential is changing quickly as one moves through the region. If the surfaces lie far apart, the same ten volts are spread thinly over a long distance, and the potential changes only slowly. As the next part of this section will shortly make exact, a potential that changes quickly with distance means a strong field, and a potential that changes slowly means a weak one. Crowded equipotentials, strong field; widely spaced equipotentials, weak field. The map told the truth about hills, and it tells the same truth about fields.

The fourth property is a short one. *Two different equipotential surfaces can never cross one another.* The argument is the very one that forbade two field lines from crossing, back in the discussion of the point charge. If two equipotentials of different value did meet at some point, that single point would lie on both surfaces at once, and so would be obliged to hold both values of potential at the same time. But the potential at a point in space is one single definite number; a place cannot be at twelve volts and at twenty volts together. The crossing is impossible, and the equipotentials, like well-mannered contour lines, thread their way among one another without ever touching.

The fifth property ties the whole idea back to conductors. *The surface of a conductor sitting in electrostatic equilibrium is an equipotential surface,* and more than that, the entire solid body of the conductor is at one and the same potential. The reason was half built already in the work on Gauss's law. Inside a conductor in equilibrium the electric field is zero, for if any field remained the free charges would still feel a force and would still be drifting, and the conductor would not yet be in equilibrium at all. With the field zero everywhere inside, no work is needed to carry a charge from any one interior point to any other, and where no work is needed to move between two points those two points must be at the same potential. Every point of the metal, the surface and the depths alike, therefore shares one common value. A conductor in equilibrium is an equipotential body wrapped in an equipotential skin, and this is precisely why the field lines met its surface at right angles in the earlier figures: that surface is an equipotential, and the first property demands it.

So far the equipotential surfaces have been read like a map, by eye. It is time to make the reading exact, and to do that the section returns once more to the steepness of the hill. The steepness at any spot is the answer to a precise question: *as I take one short step across the contours, by how much does my height change?* A small change of height for the step means gentle ground; a large change means a cliff. The electric field, it now turns out, is the answer to exactly the same question asked of the potential: *as I take one short step through the field, by how much does the potential change?* The field is the steepness of the potential landscape, and the time has come to write that down as an equation.

In the previous section the potential was built out of the field by adding up the effect of the field all along a path, an operation of accumulation, an integral. The reverse operation, travelling from the potential back to the field, must therefore be an operation of rate of change. For a field that points along a single direction, that relationship is:

$$\mathbf{E} = -\frac{dV}{dr}$$

The quantity $\frac{dV}{dr}$ in that equation is the rate at which the potential changes with distance, and it carries a name of its own, the **potential gradient**. The equation says, in a few symbols, everything the map said in pictures: *the electric field at a point is equal to the negative of the potential gradient there.*

Two parts of that statement repay a careful look, the size and the sign. Take the size first. The equation says the field is large wherever the potential is changing rapidly with distance, and small wherever the potential changes only sluggishly. That is the third property over again, no longer dressed as a picture of crowded lines but written out as a piece of arithmetic: closely spaced equipotentials simply mean a large potential

gradient, and a large potential gradient means a strong field. The strength of the field and the steepness of the potential are not two separate facts but one single fact seen twice.

Now the sign. The minus sign standing in front is not decoration; it fixes the direction in which the field points. It declares that the field points the way the potential decreases, from high potential towards low, downhill on the landscape of potential. Set a positive charge free in a field and it rolls, exactly as a ball set down on a hillside rolls, from high potential to low. The field is the arrow that shows which way is downhill, and the minus sign is the small but essential piece of bookkeeping that keeps that arrow pointing the correct way.

A check on units is worth a moment, because it heals an apparent split in the language of the subject. The potential gradient is a potential divided by a distance, so its unit is the volt per metre, written V/m. Yet the electric field was first introduced, chapters and sections ago, as a force per unit charge, measured in newtons per coulomb, written N/C. The two units look unrelated, and are in fact identical. A volt is a joule per coulomb, and a joule is a newton metre, so a volt per metre is a newton metre per coulomb, divided again by a metre, and the metre cancels to leave simply a newton per coulomb. One volt per metre is one newton per coulomb, $1\text{V/m} = 1\text{N/C}$. The force picture and the energy picture have been describing the very same field, in two accents, all along.

The relationship runs in both directions. Just as the field can be found from the potential by taking a gradient, the potential can be recovered from the field by the accumulation met in the previous section,

$$V_B - V_A = -\int \mathbf{E} \cdot d\mathbf{r},$$

the field summed all along a path from one point to the other. The words all along a path are the important ones, and they were earned in the previous section: because the electrostatic force is conservative, the sum comes out the same whatever route is chosen between the two points. The two equations, $\mathbf{E} = -\frac{dV}{dr}$ and its inverse, form a matched pair, each one undoing the work of the other. Together they hand the student a freedom well worth having: a problem may be attacked from the field end or from the potential end, whichever proves the easier, and the pair of equations then carries the answer safely across to the other description.

The simplest case of all is the uniform field between parallel plates, and it is worth seeing how the grand relationship collapses neatly onto the modest formula already in hand. Between the plates the potential does not curve; it changes at a perfectly steady rate from the one plate to the other. A steady rate of change means the potential gradient is just the whole potential difference divided by the whole separation, and the field is then the same everywhere, with magnitude:

$$E = \frac{V}{d}$$

That small equation, first met when parallel plates were introduced, now stands revealed for what it truly is: not a separate rule to be memorised on its own, but the gentlest possible example of $\mathbf{E} = -\frac{dV}{dr}$, a straight-line potential giving a constant gradient giving a constant field.

One question, more than any other, decides whether a student has truly grasped the relationship between field and potential, and it can be put in two halves.

First: *if the electric field is zero everywhere throughout some region of space, must the potential be zero there too?*

Second, the mirror of the first: *if the potential is zero at some point, must the field be zero there?*

The honest answer to both halves is no, and the contour map explains the no before a single charge has been drawn.

Think once more of the hillside. A wide flat plateau has zero slope, the ground level across the whole of it, and yet that plateau may sit two thousand metres above the sea. Flatness is not lowness; a piece of ground can be perfectly level and still be very high. That settles the first half. And a particular spot may happen to lie exactly at sea level, its height reading zero, while the ground there plunges steeply down towards a cove; the height is zero but the slope is not. That settles the second half. Height and slope are simply different things, and either one of them may vanish while the other stands.

Translated into electricity, the potential is the height and the field is the slope. Suppose the field is zero throughout a region. The equation for the field then says the gradient of the potential is zero there, which

means the potential does not change from place to place; it is constant. But constant is a long way from zero. The cleanest illustration is the hollow charged conductor met earlier under Gauss's law. Everywhere inside the hollow the field is exactly zero, and yet the potential inside is not zero at all: it holds, unchanged, the same value the charged surface itself has, carried into the interior precisely because there is no field in there to alter it. A region of zero field is a region of flat, constant potential, and that constant may be a thousand volts as easily as nothing.

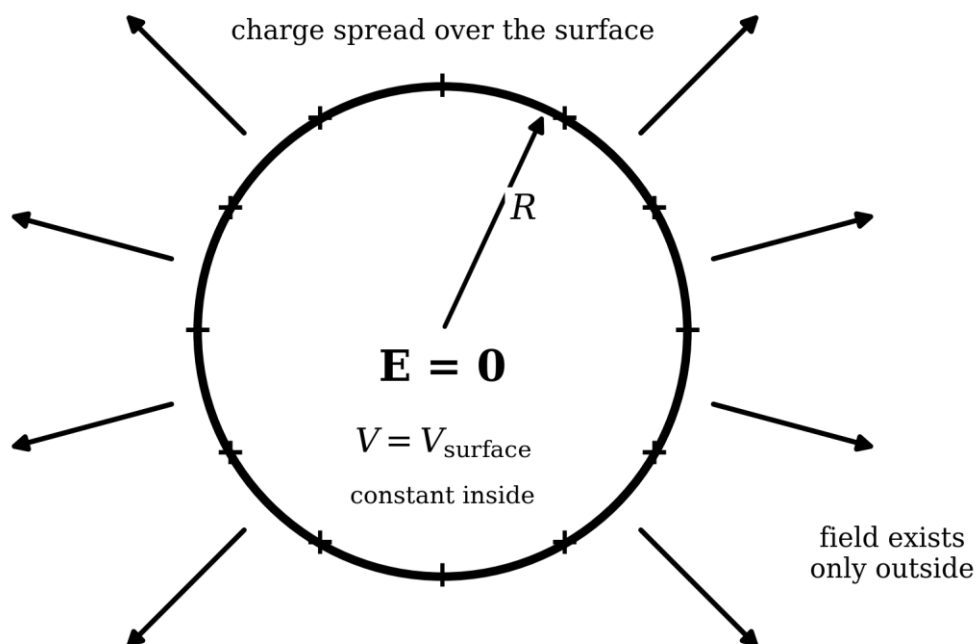


Figure: A hollow charged conductor. The charge spreads itself over the outer surface. Inside the hollow the field is zero at every point, yet the potential is not zero: it holds the constant value of the charged surface all the way through the interior. Field lines exist only outside the conductor.

The other half is settled by an object already familiar, the electric dipole of the previous section. At every point on the plane lying halfway between its two equal and opposite charges, the potential is zero, the positive contribution and the negative contribution being equal numbers of opposite sign that cancel exactly in the sum. Yet the field on that plane is not zero; it is strong and definite, because the field is a vector sum, and there the two field vectors of the two charges point so as to add together rather than to cancel. A point of zero potential is merely a place where signed numbers have happened to total zero; it makes no promise whatever about the field.

The lesson behind the two halves is one single lesson, and it is the very heart of this section. The field is obtained from the potential by differences and by rates of change, never from the bare value the potential happens to have at one point. The field equation has eyes only for how the potential changes; the actual figure the potential has reached is invisible to it. That is why a whole region sitting at one constant potential, however high that potential may be, contains no field, and why a lone zero of potential, with changing values pressing in all around it, can sit calmly in the middle of a powerful field.

Equipotential surfaces have now done their work. They have sorted the cluttered space around a charge into a clean set of level sheets, they have shown the field to be nothing grander than the steepness of those sheets, and they have kept the two ideas of field and potential firmly distinct even while binding them together with a single equation. Everything in the chapter so far, the field, the potential and the equipotential, has been a description of the empty space that charges drape around themselves. The next section at last lets something move. It places a charged particle into a uniform field and follows it as the field takes hold, accelerates it, bends its flight into a graceful curve, and, in one final and elegant trick, holds it perfectly still in the air against the steady pull of gravity.

Worked examples now, and the equipotential idea proves an unusually generous host: it has a habit of pressing an answer into the candidate's hand before the pen has properly warmed up. Four examples follow. Approach them in good spirits, for not a single one of them bites.

BINDER Example 28

A certain equipotential surface in an electric field has a potential of 12V at every point on it. A charge of +3.0C is carried from one point on this surface to another point on the same surface, by a slow and thoroughly winding route. Calculate the work done by the external agent in carrying the charge.

Solution

The work done in carrying a charge between two points is the charge multiplied by the difference of potential between those points:

$$W = q\Delta V$$

Both points lie on the same equipotential surface, so each of them is at a potential of 12V. The difference of potential between them is therefore:

$$\Delta V = 12V - 12V = 0V$$

Substituting this difference of potential, together with the charge, into the equation for the work:

$$W = q\Delta V = (3.0C)(0V) = 0J$$

The work done by the external agent is zero.

Making Sense of the Answer: *The long winding route was a deliberate distraction and changed nothing whatever. Work done on a charge depends only on the potentials of the starting point and the finishing point, never on the journey taken between them, and here those two potentials were equal. Had the charge instead been carried to a point on a different equipotential, a surface at twenty volts say, the answer would not have been zero. It is the sameness of the two potentials, and not the motion itself, that drives the work down to nothing.*

Think Like a Physicist: *When a question places the start and the finish of a journey on the same equipotential surface, write zero for the work at once, and read the rest of the question only out of politeness. The length of the path, the number of its bends and the size of the charge are all decoration laid over a single fact: equal potentials, no work.*

BINDER Example 29

Two large parallel metal plates are mounted horizontally, one above the other, separated by a gap of 4.0cm. The upper plate is held at a potential of 200V and the lower plate at 0V.

- Find the magnitude of the uniform electric field in the gap between the plates.
- State the potential of the equipotential surface lying exactly halfway between the plates, and find how far above the lower plate the 50V equipotential surface lies.

Solution

(a) Between parallel plates the field is uniform, which means the potential changes at a perfectly steady rate across the gap and the potential gradient is simply the whole potential difference divided by the whole separation:

$$E = \frac{\Delta V}{d}$$

The potential difference between the plates is $200V - 0V = 200V$, and the separation is 4.0cm, which is 0.040m. Substituting these values:

$$E = \frac{200V}{0.040m} = 5.0 \times 10^3 V/m$$

The field between the plates has magnitude $5.0 \times 10^3 V/m$, which is the same thing as $5.0 \times 10^3 N/C$.

(b) Because the field is uniform, the potential rises in step with distance: equal climbs of distance bring equal climbs of potential. Halfway across the gap the potential has therefore reached exactly half of its full range, so the equipotential surface there carries a potential of:

$$\frac{200V + 0V}{2} = 100V$$

To locate the 50V surface, measure a height upward from the lower plate. The potential at that height, rising steadily from zero at the lower plate at a rate equal to the field strength, is the field strength multiplied by the height, $V = Ex$, so the height is the potential divided by the field strength:

$$x = \frac{V}{E} = \frac{50\text{V}}{5.0 \times 10^3 \text{V/m}} = 0.010\text{m}$$

The 50V equipotential surface lies 0.010m, that is to say 1.0cm, above the lower plate.

Making Sense of the Answer: *The equipotentials between the plates are evenly spaced, and that even spacing is the whole signature of a uniform field. The 50V surface sits a quarter of the way up the gap for the plainest of reasons: 50V is one quarter of the 200V total, and in a uniform field equal steps of potential always occupy equal steps of distance. The numbers have simply shared the gap out in proportion.*

Think Like a Physicist: *In a uniform field the potential behaves like a ruler laid flat across the gap, its markings spaced evenly from the one plate to the other. Locating a particular equipotential is then not really a piece of physics at all but a piece of proportion: find what fraction the wanted potential is of the total, and the surface sits that very same fraction of the way across the gap.*

REAL Example 30

During a storm a high-voltage TANESCO line is brought down, and one end of it comes to rest on the open ground of a field. The point where the wire touches the earth is now held at a high potential, and the potential of the ground falls away steadily with distance from that point, so that the earth around the fallen wire is covered with a set of equipotential surfaces shaped like rings centred on the point of contact. A person caught in that field is endangered not by touching the wire but by the ground itself.

- Using the idea of a potential gradient, explain why a person who walks away with long, normal strides is in greater danger than one who shuffles away with the feet kept close together.
- Explain why, once the danger is understood, the safest way to cross such ground is to move with the two feet always touching or kept together.

Solution

(a) What threatens a person standing on that ground is not the potential of the ground itself but the difference of potential between the two feet, for it is that difference, acting as a voltage applied across the body, that drives a current up one leg and down the other. The potential of the ground changes with distance from the point of contact, so two feet planted at two different distances stand on two different equipotential rings and are held at two different potentials. A long stride sets the feet far apart. The potential gradient, the rate at which the potential changes with distance, then acts across that long gap between the feet, and a gradient multiplied by a long distance gives a large difference of potential, and so a large and dangerous current through the body. A short shuffling step sets the feet almost together. The same gradient now acts across only a tiny distance, the two feet sit on very nearly the same ring at very nearly the same potential, the difference between them is small, and only a small and survivable current can flow. Cutting the distance between the feet cuts the voltage across the body in the very same proportion.

(b) The principle behind the danger points straight at the remedy. Since the hazard is the difference of potential between the feet, and since that difference grows with the distance between the feet, safety lies in never letting the feet be far apart. Ordinary walking is dangerous because at every stride the feet are spread wide across many rings, and running is worse still, for it throws the feet far apart and lands them hard on widely separated rings. The safe method is the exact opposite of a stride: keep the two feet touching one another and move by sliding them forward together in small shuffles, or by hopping with the feet kept firmly together, so that at every instant both feet rest on practically the same equipotential ring and practically no difference of potential is ever allowed to appear across the body.

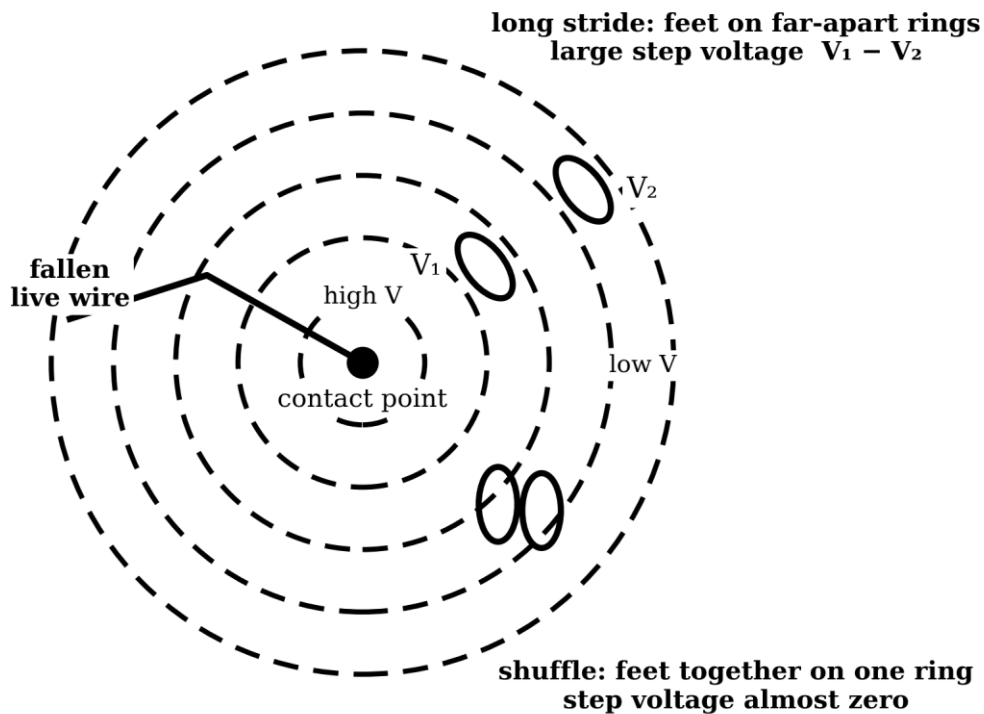


Figure: Step potential near a fallen conductor. Seen from above, the ground around the point of contact is covered with equipotential rings. Two feet placed a long stride apart rest on widely separated rings, and a large step voltage appears between them. Feet kept together rest on a single ring, and almost no step voltage appears.

Making Sense of the Answer: The danger is carried by the difference of potential, and not by the potential itself, the very same lesson taught by the bird that perches unharmed on a single high-voltage line. Here that lesson is turned into a rule for survival. The equipotential rings lying on the ground cannot be seen, but they are every bit as real as the contour lines on a map, and a person's safety is decided entirely by how many of those rings lie trapped between the two feet.

Think Like a Physicist: A hazard that announces itself as a high voltage is very often, when it is looked at properly, a hazard about a voltage difference taken across a small span. Reduce the span and the difference shrinks along with it. Reading the danger through the idea of a potential gradient turns a frightening and confusing situation into a simple matter of geometry: keep the two sampling points, here the two feet, close together, and the gradient is left with only a short distance over which to build a difference.

HOT Example 31

Two students are arguing.

The first says: If the electric field is zero everywhere inside some region of space, then the electric potential must be zero everywhere in that region as well.

The second disagrees: A zero field only means the potential is the same at every point of the region, and that same value may be anything at all, a large value included.

To settle the matter they consider a hollow metal sphere of radius 0.10m carrying a charge of +8.0nC spread over its surface. Decide which student is correct, and support the decision by finding both the electric field and the electric potential at the centre of the sphere.

Solution

Two separate facts about a charged hollow conducting sphere are needed, and both were established earlier in this chapter with the help of Gauss's law.

Take the field first. A Gaussian surface drawn anywhere inside the hollow sphere encloses no charge whatever, since all of the charge sits on the outer surface of the metal. Gauss's law then gives zero for the field at every interior point, the centre included:

$$E_{\text{centre}} = 0$$

Now take the potential. Outside the sphere the potential is that of a point charge, $V = \frac{kQ}{r}$.

Following it inward, it rises steadily as the distance shrinks until, at the surface, where the distance equals the radius of the sphere, it reaches the value $\frac{kQ}{R}$.

Inside the surface the field is zero, so there is no gradient anywhere in the interior to change the potential any further; it therefore holds that same surface value, unchanged, all the way to the centre:

$$V_{\text{centre}} = \frac{kQ}{R}$$

Substituting:

$$V_{\text{centre}} = \frac{(9.0 \times 10^9 \text{Nm}^2/\text{C}^2)(8.0 \times 10^{-9} \text{C})}{0.10 \text{m}} = 7.2 \times 10^2 \text{V}$$

At the centre of the sphere the field is zero while the potential is 720V. The field has vanished; the potential plainly has not. The first student is wrong and the second student is right.

Making Sense of the Answer: *The result is the hollow-conductor fact and the dipole fact pressed together into one single lesson. A field of zero does not empty the potential out; it only flattens it, fixing it at one constant value throughout the region. Field and potential are independent enough that either one of them may stand at zero while the other does not, because the field measures the gradient of the potential, and a gradient of zero is the mark of a constant, not the mark of a nothing.*

Think Like a Physicist: *When a question sets a field against a potential, refuse to let the one quietly stand in for the other. Ask two separate questions. For the field, the only question is whether the potential is changing at the point; if it is not, the field is zero there. For the value of the potential, the question is what the potential has added up to at that point, which is a quite separate matter. The hollow charged sphere is the cleanest place in all of electrostatics to keep those two questions apart, and a student who can keep them apart there can keep them apart anywhere.*

Four examples, and not a single casualty among them. Equipotential surfaces, for all the quiet labour they save a person, turn out to be about the most peaceable and best-mannered objects in the whole of electrostatics.

MOTION OF A CHARGED PARTICLE IN A UNIFORM ELECTRIC FIELD

Every page of this chapter so far has drawn the electric field as a kind of landscape, a fixed pattern of arrows and level surfaces hanging in empty space, waiting. Nothing in it has actually moved. That changes now. This section takes a single charged particle, sets it down inside a uniform field, and simply watches what the field does to it. And the most useful thing to carry into that watching is a comparison the reader already owns. A uniform electric field does to a charge exactly what the uniform gravity near the ground does to a thrown stone. Every result about projectiles, every parabola traced by a thrown ball, every trick of splitting a motion into a horizontal part and a vertical part, carries straight across. The physics of this section is, in a real sense, already familiar to the reader; only the names have changed.

The natural place to stage that comparison is the region between two parallel charged plates, because there, and almost nowhere else so cleanly, the field is uniform. The field E has the same size and the same direction at every point of the gap, so a charge placed anywhere in that gap meets one and the same field. As the next few paragraphs show, that single fact of uniformity is what makes the motion so simple, and so very like the flight of a stone.

Begin with the force. A charge Q sitting in a field E feels an electric force, and the size of that force was settled in the very first pages of the chapter, in the definition of the field itself:

$$F = QE$$

In a uniform field E is constant, so this force is constant too, the same in size and the same in direction wherever the charge happens to lie in the gap. A constant force is the key that unlocks everything else in the section.

For a constant force acting on a particle of mass m , Newton's second law gives a constant acceleration. Dividing the force by the mass:

$$\mathbf{a} = \frac{\mathbf{F}}{m} = \frac{Q\mathbf{E}}{m}$$

The acceleration of a charged particle in a **uniform** field is **constant**. That one sentence is the whole engine of the section. A constant acceleration is precisely the condition under which all the equations of uniformly accelerated motion, the ones already learned for objects falling under gravity, apply without the smallest change.

The charge in the field is, mathematically, a projectile, and the quantity $\frac{QE}{m}$ simply takes the part that g took for falling bodies.

Two remarks before the motion is set going.

First, the direction. A positive charge is pushed along the field and accelerates in the direction of $\mathbf{E}$; a negative charge, an electron above all, is pushed against the field and accelerates in the direction opposite to $\mathbf{E}$.

Second, a point of bookkeeping. The particle has mass, so the Earth pulls on it with a weight mg as well. Whether that weight matters depends entirely on what the particle is. For an electron or an ion the mass is so fantastically small that the weight is utterly swamped by the electric force, thousands upon thousands of times smaller, and it is dropped without apology. For a larger object, a charged drop of oil say, the weight is comparable with the electric force and cannot be ignored at all. This section meets both cases, and the difference between them is simply the difference between an electron and an oil drop.

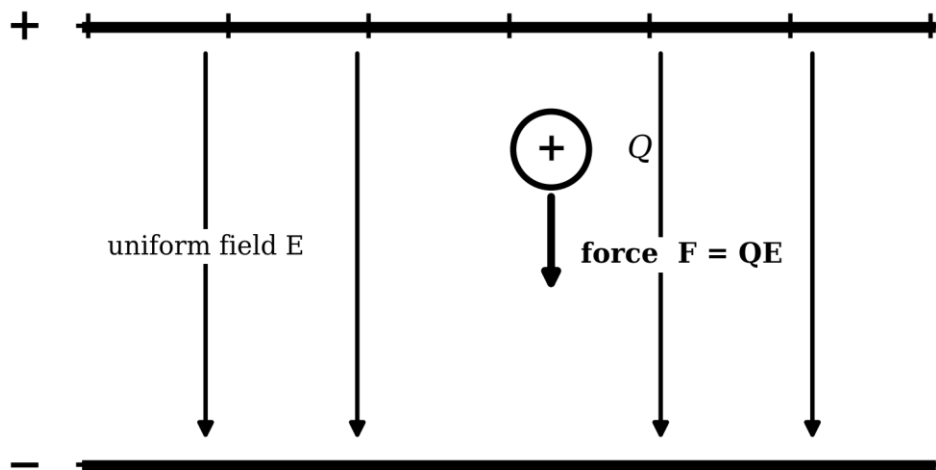


Figure: A charge in a uniform field. Between parallel plates the field is uniform, so the electric force on a charge has the same size and direction everywhere in the gap. A constant force produces a constant acceleration, exactly as a constant gravitational field does for a falling body.

The simplest motion is the one in which the charge is released from rest, or is already travelling straight along a field line. With the force directly along the motion, the charge runs in a straight line and simply speeds up, gaining velocity at the steady rate $\mathbf{a} = \frac{Q\mathbf{E}}{m}$, exactly as a stone dropped from the hand gathers speed at the rate g . There is nothing fresh to derive: the familiar equations of straight-line accelerated motion deliver the velocity gained and the distance covered. The energy side of it was settled already in the work on potential, where the push of the field across a potential difference was shown to hand the charge a kinetic energy equal to Q times the potential difference crossed. Speeding a charge up in a straight line is the principle of the electron gun that launches the beam in every cathode-ray tube. It is motion along the field, and it holds no surprises.

The interesting motion is the one across the field. Let a charged particle be fired into the uniform field so that it enters moving at right angles to the field, travelling parallel to the plates rather than along the gap between them. This is the exact counterpart of a stone thrown horizontally from the top of a cliff, and it is handled in the exact same way: split the motion into two independent parts and treat each part entirely on its own.

Set an axis x along the direction the particle was fired, and an axis y along the field. Along x there is no force at all, because the field, and so the force QE , points entirely along y . With no force along x , the particle simply keeps whatever speed u it entered with, and after a time t it has travelled a distance:

$$x = ut$$

Along y the story is the story of the falling stone. The particle starts with no velocity in the y direction, and the constant force gives it the constant acceleration $a = \frac{QE}{m}$. Starting from rest and accelerating steadily for the same time t , it is carried sideways a distance:

$$y = \frac{1}{2}at^2$$

Two motions, but one clock. The time t is the single thread running through both, and eliminating it joins the two motions into the shape of the path.

From the first equation the time is $t = \frac{x}{u}$. Substituting that into the second equation, and then replacing the acceleration by $\frac{QE}{m}$, gives

$$y = \frac{1}{2}a\left(\frac{x}{u}\right)^2 = \frac{QEx^2}{2mu^2}$$

The crucial thing in that result is not the cluster of constants but the shape. The deflection y is proportional to x^2 , and a relationship in which one coordinate grows as the square of the other is a parabola. The charged particle, crossing the uniform field, curves along a parabola, just as surely as the stone thrown from the cliff traces one under gravity.

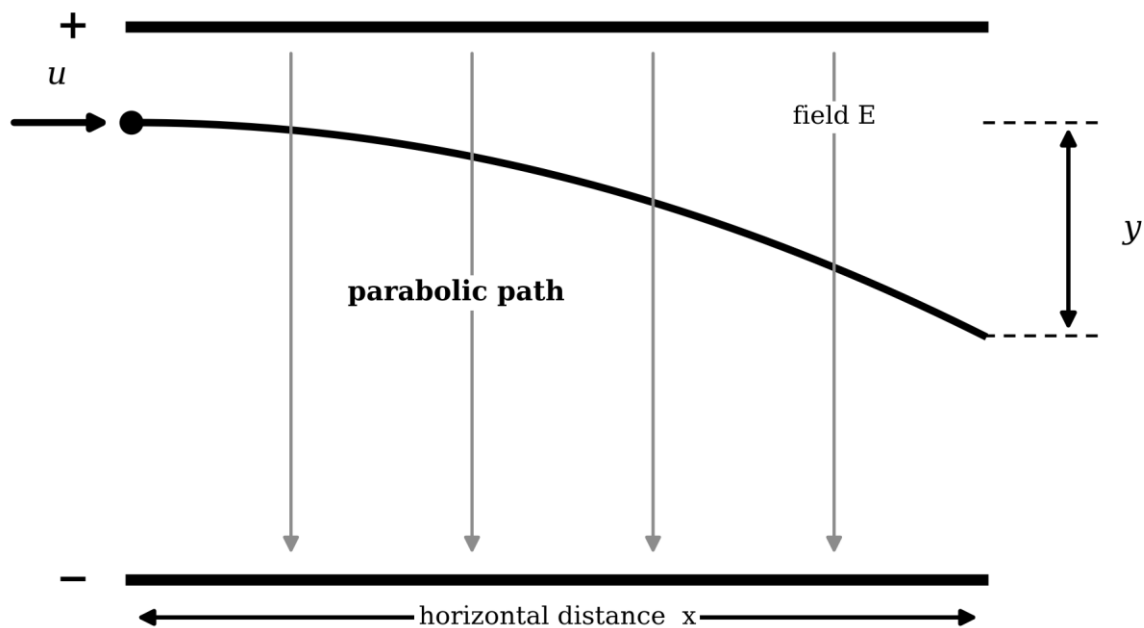


Figure: The parabolic path. A particle fired into the field at right angles to it keeps a steady speed along its original direction while being steadily accelerated along the field. The two motions combine into a parabola, the same curve a stone traces when thrown horizontally.

This parabola is not a curiosity. It is the working principle of a whole family of instruments. Steer a beam of electrons between a pair of charged plates and the size of the deflection can be read straight off the parabola; vary the voltage on the plates and the beam can be aimed at any chosen spot on a screen. That is the cathode-ray oscilloscope tracing its glowing line, the older television painting its picture, and the inkjet printer flicking charged droplets of ink to their appointed places on the page. Every one of them is this paragraph's parabola, put to work.

The last motion is the most surprising of all, because in it nothing moves. Take the charged oil drop mentioned earlier, the case in which the weight is not negligible, and set it between two horizontal plates

with the field running vertically. Two forces now act on the drop: Its **weight mg** pulling it straight down, and the **electric force QE** which, if the field is pointed correctly, pushes it straight up.

Adjust the strength of the field, by adjusting the voltage on the plates, until the upward electric force exactly matches the downward weight. At that one setting the two forces cancel, the total force on the drop is zero, and the drop hangs in the air, perfectly still, suspended on nothing but the balance of electricity against gravity.

The condition for that **motionless balance** is simply that the two forces are equal:

$$QE = mg$$

It is a short equation, but one of the most quietly powerful in all of physics, because of what can be done with it. Every quantity in it can be measured. The weight side, mg , can be found by other means; the field E is known from the plate voltage and the plate spacing. That leaves the charge Q as the only unknown, and the equation hands it straight over.

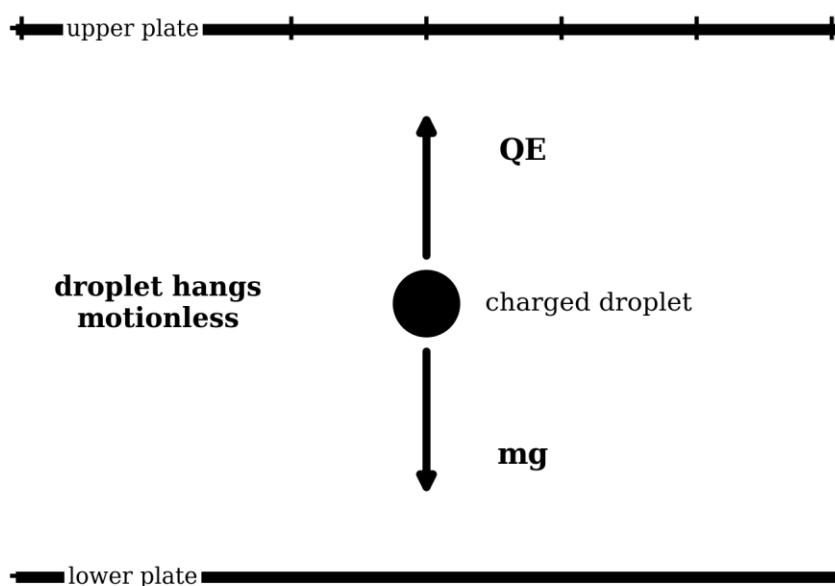


Figure: The balancing act. A charged droplet between horizontal plates feels its weight downward and the electric force upward. When the field is adjusted so the two forces are equal the drop hangs motionless, and the balance condition becomes a way of measuring the charge.

This is precisely the experiment Robert Millikan performed (**Millikan oil-drop experiment**), beginning in 1909, with tiny drops of oil. He sprayed a fine mist of oil between two horizontal plates, picked out a single charged droplet under a microscope, and adjusted the voltage until that droplet hung motionless. From $QE = mg$ he then calculated the charge on the drop. Repeating the measurement on droplet after droplet, he found something remarkable: the charge was never just any value, but always a whole-number multiple of one smallest basic amount. That basic amount was the charge on a single electron, the elementary charge e , the very quantisation of charge met at the start of this chapter.

Four worked examples follow, and they share a single secret: every one of them is a projectile problem wearing a disguise. A reader who once survived the mechanics of thrown stones already has, tucked away, everything needed to finish all four. The disguises are rather good, but they are only disguises.

BINDER Example 32

A uniform electric field of strength $5.0 \times 10^3 \text{ N/C}$ is set up between two parallel plates. A small charged particle, of charge $2.0 \times 10^{-6} \text{ C}$ and mass $4.0 \times 10^{-8} \text{ kg}$, is placed in the field.

- Find the electric force acting on the particle.
- Find the acceleration this force produces.

Solution

(a) The electric force on a charge in a field is the charge multiplied by the field strength:

$$F = QE$$

Substituting the charge and the field strength:

$$F = QE = (2.0 \times 10^{-6} \text{C})(5.0 \times 10^3 \text{N/C}) = 1.0 \times 10^{-2} \text{N}$$

(b) The acceleration follows from Newton's second law, the force divided by the mass:

$$a = \frac{F}{m} = \frac{1.0 \times 10^{-2} \text{N}}{4.0 \times 10^{-8} \text{kg}} = 2.5 \times 10^5 \text{m/s}^2$$

The particle accelerates at $2.5 \times 10^5 \text{m/s}^2$ in the direction of the field.

Making Sense of the Answer: The acceleration is enormous, hundreds of thousands of metres per second in every second, and yet the force that produces it, a hundredth of a newton, would not trouble a housefly. It is the smallness of the mass that magnifies a gentle force into a violent acceleration. It is worth noticing, too, that the weight of this particle, its mass times g , is only about $4 \times 10^{-7} \text{N}$, some twenty-five thousand times smaller than the electric force. For a particle this light, gravity is rightly left out of the account.

Think Like a Physicist: A problem of this kind is two steps, never one. The first step, $F = QE$, belongs to electricity; the second step, $a = \frac{F}{m}$, belongs to mechanics and has nothing electrical in it at all. Keeping the two steps mentally apart, the field producing a force and then the force producing an acceleration, is what keeps a charged-particle problem from ever turning into a muddle.

REAL Example 33

In an experiment modelled on Millikan's, a tiny charged oil drop of mass $3.3 \times 10^{-15} \text{kg}$ is held perfectly motionless in the uniform field between two horizontal plates. The field has strength $2.0 \times 10^5 \text{N/C}$ and is directed so that the electric force on the drop is upward.

(a) State the condition for the drop to hang motionless, and use it to find the charge on the drop.

(b) Taking the elementary charge as $1.6 \times 10^{-19} \text{C}$, find how many excess electrons the drop carries.

Solution

(a) For the drop to hang motionless the total force on it must be zero, which means the upward electric force must exactly balance the downward weight:

$$QE = mg$$

Making the charge the subject and substituting the values, with g taken as 9.8m/s^2 :

$$Q = \frac{mg}{E} = \frac{(3.3 \times 10^{-15} \text{kg})(9.8 \text{m/s}^2)}{2.0 \times 10^5 \text{N/C}} = 1.6 \times 10^{-19} \text{C}$$

(b) The number of excess electrons is the charge on the drop divided by the charge on a single electron:

$$n = \frac{Q}{e} = \frac{1.6 \times 10^{-19} \text{C}}{1.6 \times 10^{-19} \text{C}} = 1$$

The drop carries a single excess electron.

Making Sense of the Answer: The charge came out equal, to the precision of the measurement, to the elementary charge itself, so the drop carries exactly one electron more than its protons can balance. This is the very heart of Millikan's discovery. Drop after drop returned a charge that was always a whole number of these elementary units and never a fraction of one, and that stubborn whole-number pattern was the proof that electric charge comes in indivisible grains.

Think Like a Physicist: The balance condition turns a stationary object into a measuring instrument. Nothing moves, and yet because the drop is motionless a physicist may write down at once that the forces upon it are equal, and an equation with a single unknown is an equation already half solved. Whenever an object is described as hanging, floating, or suspended, the first line of the answer should record that the forces upon it sum to zero.

REAL Example 34

Kipanga and Kipute are watching a demonstration in which charged particles are sent into a uniform electric field. Kipanga argues that a positive ion and an electron, sent into the very same field, must come away with accelerations of equal magnitude, since the field is identical for both and the field is what does the pushing. Kipute insists that the two accelerations are nothing alike. Decide which student is right, and explain the reasoning.

Solution

Both students have hold of a piece of the truth, which is what makes the disagreement worth settling with care. Begin with the force. The electric force on a particle is QE . An electron and a singly charged positive ion carry the same magnitude of charge, the elementary charge e , and they are placed in the same field E , so the force on each of them genuinely does have the same magnitude. To that extent Kipanga is right, and it is presumably this equal force that the argument has in mind. But a force is not an acceleration. The acceleration is the force divided by the mass, $a = \frac{QE}{m}$, and here the two particles part company entirely. An electron is among the lightest objects in all of nature; even the lightest possible ion, a lone proton, is already about 1.8×10^3 times more massive, and a typical ion thousands of times more massive still. Dividing the same force by these wildly different masses produces wildly different accelerations. Set side by side, the ratio of the ion's acceleration to the electron's is:

$$\frac{a_{\text{ion}}}{a_e} = \frac{m_e}{m_{\text{ion}}},$$

the inverse ratio of their masses, which is a tiny fraction. The electron, for the same force, is flung forward with an acceleration tens of thousands of times greater than the ion's. Kipute is right: equal force, but utterly unequal acceleration. Kipanga's slip was to let the word force quietly turn into the word acceleration somewhere in the middle of the argument.

Making Sense of the Answer: *The resolution is the equation $a = \frac{QE}{m}$ read with care. The field fixes the force; the mass then fixes how much motion that force can produce. Two particles in one field feel matched forces and answer with mismatched accelerations, in exact inverse proportion to their masses. This is the reason that, in every device which steers particles with fields, it is electrons that are steered: being the least massive, they are the most obedient to the field.*

Think Like a Physicist: *Whenever two objects are compared in the same field, separate the two questions the comparison hides inside itself. Is the force the same? That depends only on the charge and the field. Is the acceleration the same? That depends also on the mass. A great many confident wrong answers come from settling the first question and then forgetting that the second one was ever asked.*

HOT Example 35

An electron is fired horizontally, at a speed of $3.0 \times 10^6 \text{ m/s}$, into a uniform electric field of strength 200 N/C directed vertically downward. The field fills a region 10 cm long, measured along the electron's original direction of travel. Taking the electron's charge as $1.6 \times 10^{-19} \text{ C}$ and its mass as $9.11 \times 10^{-31} \text{ kg}$, find the vertical deflection of the electron by the time it leaves the field region.

Solution

The motion splits into two independent parts, one along the original direction of travel and one along the field.

Along the direction of travel there is no force, so the electron keeps its entry speed. So:

$$t = \frac{x}{u} = \frac{0.10 \text{ m}}{3.0 \times 10^6 \text{ m/s}} = 3.33 \times 10^{-8} \text{ s}$$

Along the field the electron starts from rest and accelerates. The acceleration is the electric force divided by the mass:

$$a = \frac{QE}{m} = \frac{(1.6 \times 10^{-19} \text{ C})(200 \text{ N/C})}{9.11 \times 10^{-31} \text{ kg}} = 3.51 \times 10^{13} \text{ m/s}^2$$

The vertical deflection is the distance the electron is carried, from rest, under this acceleration in the time just found:

$$y = \frac{1}{2}at^2 = \frac{1}{2}(3.51 \times 10^{13} \text{m/s}^2)(3.33 \times 10^{-8} \text{s})^2 = 1.95 \times 10^{-2} \text{m}$$

The electron leaves the field deflected by $1.95 \times 10^{-2} \text{m}$, which is very nearly 2.0cm.

Making Sense of the Answer: *The deflection is small, around two centimetres, and yet it is built from an acceleration of staggering size, $3.51 \times 10^{13} \text{m/s}^2$. What keeps the deflection so modest is the brevity of the crossing: the electron is inside the field for only about thirty nanoseconds, and even a colossal acceleration can move a particle only so far in so short a time. The same answer follows in a single line from the parabola formula $y = \frac{QEx^2}{2mu^2}$, which is just this two-step calculation compressed into one.*

Think Like a Physicist: *The whole method rests on refusing to think about the curved path directly. The parabola looks complicated; its two shadows, one cast along the direction of travel and one along the field, are each as simple as anything in mechanics, a steady speed and a steady acceleration from rest. Solve the two simple shadows on their own, let the shared time link them at the end, and the complicated curve has been conquered without ever being faced head on.*

With that, the moving charge has been followed through its three motions: straight down the field with a steady acceleration, across the field along a parabola, and held in a dead balance against gravity. The field has been mapped, the potential has been measured, the equipotentials have been drawn, and now a charge has been set loose in the field and watched. The theory of the electrostatic field is, with this section, complete. What remains is to see it earn its living.

The next section leaves the tidy world of plates and test charges and walks out into the ordinary one, into the machines and the hazards and the everyday wonders, from the lightning conductor on a roof to the photocopier humming in an office, that are nothing more than the electrostatics of this chapter quietly at work.

APPLICATIONS OF ELECTROSTATICS

Every idea in this chapter so far has been built and tested in the clean, quiet world of the textbook, with its point charges and its perfect plates. It is time to open the door and let the chapter out into the ordinary world, where it turns out to have been working all along, unremarked, inside machines and weather and the dust of a cement works. Seven applications follow, each a short look at one familiar thing through the eyes of this chapter. Not one of them needs an idea that has not already been met.

1. The Van de Graaff generator

This is the machine that makes a person's hair stand on end at a science fair, and it is electrostatics made visible. A moving rubber belt carries charge, sprayed onto it at the bottom, up into a large hollow metal sphere, where it is collected. Because the charge on a conductor always travels to the outer surface, met earlier in this chapter, the charge spreads over the outside of the sphere and the potential of the dome climbs higher and higher, reaching hundreds of thousands of volts.

The spraying of charge onto the belt relies on the point-discharge action, the same crowding of field at a sharp point that the lightning conductor will use below. A research version of exactly this machine, used as a particle accelerator, runs in physics laboratories all over the world. The principle, dome, belt, and sharp points, never changes.

2. The electrostatic precipitator

Inside the chimney of a cement works or a coal power station, a high-voltage wire runs down the centre of the gas duct. The strong field around the wire charges every speck of ash and smoke that drifts past it. Once charged, each particle feels the force $F = QE$ and is driven sideways onto earthed collecting plates, where it sticks and is later shaken off. Clean gas leaves the top of the chimney.

This is electrostatics doing a public service. The cement industry is a major one in Tanzania, and a works such as the Wazo Hill plant at Dar es Salaam would, without precipitators, blanket the surrounding district in fine grey dust. The whole device is nothing more than the force on a charge in a field, scaled up and put to work cleaning the air.

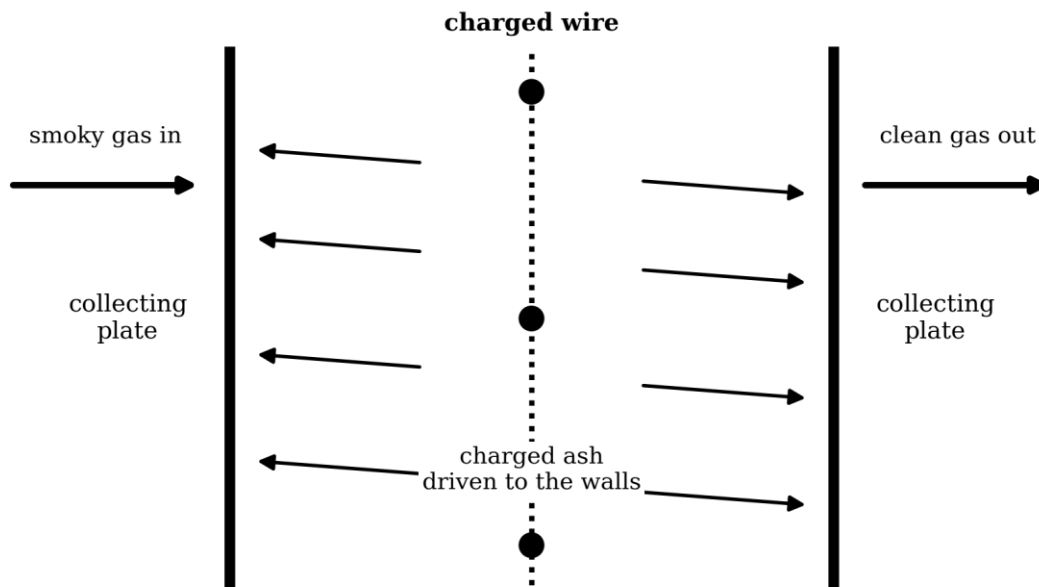


Figure: The electrostatic precipitator. Smoky gas passes a charged wire that gives every ash particle a charge. The charged particles are then pulled sideways onto the collecting plates, and clean gas leaves the chimney.

3. The photocopier and the laser printer

Both machines paint with charge. A metal drum is first given a uniform electrostatic charge in the dark. Light, reflected from the white parts of the page or written by a laser, lands on the drum and quietly removes the charge wherever it falls, leaving a hidden pattern of charge in the shape of the dark text and pictures alone.

Fine particles of toner, themselves charged, are then dusted over the drum. They cling by electrostatic attraction only to the charged regions, and so the invisible pattern becomes a visible one. The toner is pressed onto paper and fixed with heat. Every copy that slides out of the machine is a picture drawn entirely by the attraction and repulsion of charge.

4. The lightning conductor

A lightning conductor is a metal rod, sharply pointed at the top, fixed to the highest part of a building and joined by a thick copper strap to a metal plate buried in the earth. Its sharp point is the heart of it. Field crowds intensely at any sharp point on a conductor, a fact established earlier in the chapter, and the air there is ionised so strongly that charge leaks steadily and silently away into the sky, easing the build-up that would otherwise grow into a strike.

Should a strike come even so, the conductor offers it a broad, easy, metal path straight to the ground, and the dangerous current passes down the strap instead of through the brick and timber of the building. This is also why a tall coconut tree is struck in preference to a short mango tree beside it: greater height means a shorter air gap to the cloud, and a field that reaches the breakdown value of air sooner.

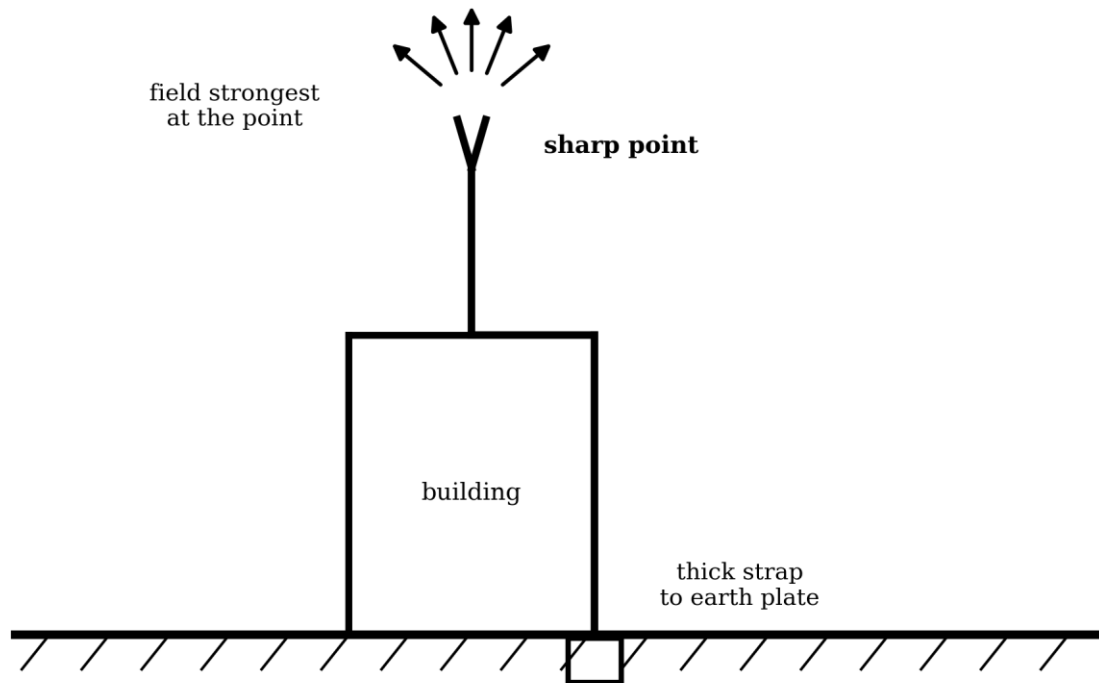


Figure: The lightning conductor. A sharp point at the top of the conductor concentrates the electric field, so charge leaks quietly into the air there. A thick strap carries any strike safely down to an earth plate.

5. The inkjet printer

An inkjet printer fires a fine stream of tiny ink droplets, and gives each droplet a controlled charge as it is formed. The charged droplets then fly between a pair of charged deflecting plates, where the uniform field bends each one along the parabola derived in the previous section. By setting the charge on a droplet, the machine chooses how far that droplet is deflected, and so steers it to its exact place on the page. It is the parabola of a charged particle in a uniform field, printing a letter.

6. Electrostatic spray painting

When a vehicle body is spray-painted in a factory, the paint leaving the spray gun is given an electric charge, and the metal body is earthed. Every charged droplet of paint is then attracted to the metal, and even droplets that would otherwise sail past are pulled back and curve around onto the far side of the panel. The result is an even coat, painted edges and hidden corners included, with very little paint wasted. The attraction between a charge and an earthed conductor, here, simply paints a car.

7. The static hazard in grain handling

Not every application is a benefit; the same physics also names a danger. On a farm or in a processing plant, dry grain pouring along a conveyor belt or down a chute rubs constantly against the surfaces it touches, and that friction separates charge exactly as a comb does through hair. Charge builds up on the insulated metal of the equipment, the potential climbs, and if it climbs far enough a spark jumps. In the dust-laden air above stored grain that spark can start a fire or an explosion.

The cure is the chapter in reverse: earth every piece of metal equipment with a conducting strap, so that charge drains harmlessly away as fast as friction deposits it and the potential is never allowed to build. Understanding electrostatics, here, is a matter of safety in an industry on which many Tanzanian livelihoods depend.

Seven machines and hazards, and a single chapter beneath all of them. The force on a charge, the crowding of field at a point, the parabola in a uniform field, the charge that flees to the surface of a conductor, the breakdown of air: each idea, built carefully in its turn, has been seen here earning its keep. The theory was never the destination. This was.

MISCELLANEOUS WORKED EXAMPLES ON ELECTROSTATICS

Example 36

- (a) A glass rod that has been rubbed with silk is brought close to small pieces of dry uncharged paper. Explain why the pieces of paper are attracted to the rod even though they carry no net charge.
- (b) Three small charged objects A, B and C are fixed at the vertices of an equilateral triangle of side 20cm. The charge at A is $+5.0\mu\text{C}$, at B is $-3.0\mu\text{C}$ and at C is $+4.0\mu\text{C}$. Calculate the magnitude and direction of the resultant electrostatic force on the object at A. Take $k = 9.0 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2$.

Solution

(a) The positive rod polarises the neutral paper: electrons within the paper shift slightly toward the surface facing the rod, leaving the surface farther from the rod slightly positive. The two induced surfaces carry equal and opposite charge, but the negative one is closer to the rod than the positive one. The attraction between the rod and the nearer negative surface therefore exceeds the repulsion between the rod and the farther positive surface, and the net force on the paper is one of attraction.

(b) Place A at the origin, with B on the positive x-axis at (0.20m,0) and C above the x-axis at (0.10m, 0.10 $\sqrt{3}$ m).

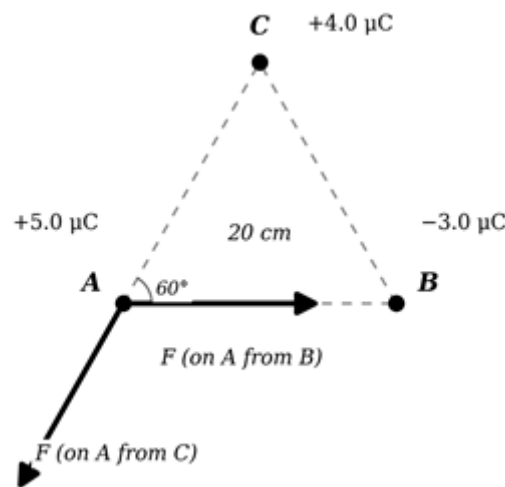


Figure: Equilateral triangle ABC of side 20 cm with the charges marked at each vertex. The two Coulomb forces on the charge at A are shown: the force from B is attractive and points from A toward B; the force from C is repulsive and points from C through A and beyond.

Force on A due to B (attractive, directed from A toward B):

$$F_{AB} = k \frac{|Q_A Q_B|}{r^2} = \frac{(9.0 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2) \times (5.0 \times 10^{-6} \text{ C}) \times (3.0 \times 10^{-6} \text{ C})}{(0.20 \text{ m})^2} = 3.375 \text{ N}$$

Components: $F_{AB,x} = +3.375 \text{ N}$, $F_{AB,y} = 0 \text{ N}$.

Force on A due to C (repulsive, directed from C toward A and beyond):

$$F_{AC} = k \frac{|Q_A Q_C|}{r^2} = \frac{(9.0 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2) \times (5.0 \times 10^{-6} \text{ C}) \times (4.0 \times 10^{-6} \text{ C})}{(0.20 \text{ m})^2} = 4.5 \text{ N}$$

The interior angle of the equilateral triangle at A is 60° . The line CA extended past A therefore lies in the third quadrant, making 60° below the negative x-axis. The repulsive force F_{AC} acts along this line, so its components are:

$$F_{AC,x} = -F_{AC} \cos 60^\circ = -(4.5 \text{ N}) \times (0.500) = -2.25 \text{ N}$$

$$F_{AC,y} = -F_{AC} \sin 60^\circ = -(4.5 \text{ N}) \times (0.866) = -3.897 \text{ N}$$

Resultant components on A:

$$\Sigma F_x = 3.375 \text{ N} - 2.25 \text{ N} = 1.125 \text{ N}, \quad \Sigma F_y = 0 \text{ N} - 3.897 \text{ N} = -3.897 \text{ N}$$

Magnitude:

$$F = \sqrt{(\Sigma F_x)^2 + (\Sigma F_y)^2} = \sqrt{(1.125\text{N})^2 + (3.897\text{N})^2} = 4.06\text{N}$$

Direction:

$$\theta = \tan^{-1}\left(\frac{|\Sigma F_y|}{|\Sigma F_x|}\right) = \tan^{-1}\left(\frac{3.897}{1.125}\right) = 73.9^\circ$$

The resultant force on A has magnitude 4.06N, directed at 73.9° below the line AB, on the side of AB opposite to C.

Example 37

- (a) A student claims that the electric field on the axis of a uniformly charged ring is greatest at the centre of the ring, “because the charges are closest there”. Identify the error in the student’s reasoning, and state the actual location on the axis at which the field is greatest.
- (b) A thin straight rod of length 1.0m carries a uniform positive linear charge density $\lambda = 4.0\text{nC/m}$. Determine the electric field strength at a point P located perpendicular to the midpoint of the rod, 0.30m from it. Compare your result with the value obtained if the rod were treated as effectively infinite at the same perpendicular distance.

Solution

(a) The error is that the student weighs each charge element only by its distance, ignoring the directions of the field contributions. At the centre of the ring every element’s contribution to the field at the centre is radial; by symmetry these contributions cancel in opposite pairs across the centre, and the net field at the centre is zero. The field on the axis is zero at the centre, rises to a maximum at $x = R/\sqrt{2}$ from the centre (where R is the radius of the ring) (and decays as kQ/x^2 far from the ring).

(b) Half-length $a = 0.50\text{m}$; perpendicular distance $x = 0.30\text{m}$. Then:

$$\sqrt{x^2 + a^2} = \sqrt{(0.30\text{m})^2 + (0.50\text{m})^2} = 0.583\text{m}$$

Using the finite-line field formula at the perpendicular bisector of the rod:

$$E = \frac{2k\lambda a}{x\sqrt{x^2 + a^2}}$$

$$E = \frac{2 \times (9.0 \times 10^9 \text{N} \cdot \text{m}^2/\text{C}^2) \times (4.0 \times 10^{-9} \text{C/m}) \times (0.50\text{m})}{(0.30\text{m}) \times (0.583\text{m})} \approx 206\text{N/C}$$

For the infinite-line approximation:

$$E_\infty = \frac{2k\lambda}{x} = \frac{2 \times (9.0 \times 10^9 \text{Nm}^2/\text{C}^2) \times (4.0 \times 10^{-9} \text{C/m})}{(0.30\text{m})} = 240\text{N/C}$$

The finite-line value is about $206/240 \approx 86\%$ of the infinite-line value, because the two ends of the actual rod contribute less than an extension to infinity would.

Example 38

- (a) Explain why the electric potential at every point on the perpendicular bisector of an electric dipole is zero, while the electric field at the same points is not.
- (b) A thin straight rod of length 1.0m carries a uniform positive linear charge density $\lambda = 4.0\text{nC/m}$. Determine the electric potential at a point P located perpendicular to the midpoint of the rod, 0.30m from it, taking the reference of zero potential at infinity.

Solution

(a) The electric potential at a point is the scalar sum of the potential contributions from each charge. The two charges of the dipole, equal in magnitude and opposite in sign, sit equidistant from any point on the perpendicular bisector; their potential contributions are equal and opposite, and cancel to zero. The electric field is a vector sum. The two contributions at a point on the bisector are vectors of equal magnitude pointing in different directions (one away from the positive charge, one toward the negative one). They do not cancel; they combine into a non-zero resultant lying parallel to the dipole axis.

(b) Half-length $a = 0.50\text{m}$; perpendicular distance $x = 0.30\text{m}$. $\sqrt{x^2 + a^2} = 0.583\text{m}$.

Using the finite-line potential formula at the perpendicular bisector:

$$V = k\lambda \ln \left(\frac{\sqrt{x^2 + a^2} + a}{\sqrt{x^2 + a^2} - a} \right)$$

$$V = (9.0 \times 10^9 \text{Nm}^2/\text{C}^2) \times (4.0 \times 10^{-9} \text{C/m}) \times \ln \left(\frac{0.583\text{m} + 0.500\text{m}}{0.583\text{m} - 0.500\text{m}} \right) = 92.5\text{V}$$

The electric potential at P is approximately 92.5V, taking the reference at infinity.

Example 39

(a) Two identical conducting spheres of mass m are suspended from a common point by insulating threads of equal length L , and each carries a charge Q . Show that at small-angle equilibrium the centre-to-centre separation of the spheres satisfies

$$d^3 = \frac{Q^2 L}{2\pi\epsilon_0 mg}$$

(b) A hollow conducting sphere of inner radius 6.0cm and outer radius 10cm carries a total charge of +12nC in electrostatic equilibrium. Determine the magnitude of the electric field at each of the following points: (i) the centre of the cavity, (ii) a point 3.0cm from the centre, (iii) a point 8.0cm from the centre, (iv) a point 15cm from the centre.

Solution

(a) Each sphere experiences three forces in equilibrium: weight mg downward, tension T along the thread, and Coulomb repulsion $F = kQ^2/d^2$ horizontal. Resolving the tension along the thread:

$$T \cos\theta = mg, \quad T \sin\theta = \frac{kQ^2}{d^2}$$

Dividing:

$$\tan\theta = \frac{kQ^2}{d^2 mg}$$

For small angles, $\tan\theta \approx \sin\theta \approx d/(2L)$. Therefore:

$$\frac{d}{2L} = \frac{kQ^2}{d^2 mg}$$

$$d^3 = \frac{2L \times kQ^2}{mg} = \frac{2Q^2 L}{4\pi\epsilon_0 mg} = \frac{Q^2 L}{2\pi\epsilon_0 mg}$$

which is the required result.

(b) Let $R_1 = 0.060\text{m}$, $R_2 = 0.100\text{m}$, $Q = 12 \times 10^{-9}\text{C}$.

(i) The centre of the cavity sits at $r = 0 < R_1$; the Gaussian sphere of radius r encloses no charge, so:

$$E = 0$$

(ii) At $r = 0.030\text{m}$, still in the cavity, $r < R_1$, so $E = 0$.

(iii) At $r = 0.080\text{m}$, the Gaussian sphere lies entirely within the metal of the shell. The field inside a conductor in electrostatic equilibrium is zero, so $E = 0$.

(iv) At $r = 0.150\text{m} > R_2$, the Gaussian sphere encloses the whole of the charge Q :

$$E = \frac{kQ}{r^2} = \frac{(9.0 \times 10^9 \text{Nm}^2/\text{C}^2) \times (12 \times 10^{-9} \text{C})}{(0.150\text{m})^2} = 4800 \text{N/C}$$

Example 40

(a) Briefly explain how Millikan determined the elementary charge using charged oil drops.

- (b) A hollow insulating sphere of inner radius 4.0cm and outer radius 10cm is uniformly charged through its shell volume with a total charge of +8.0nC. Determine the magnitude of the electric field at: (i) a point 3.0cm from the centre, (ii) a point 7.0cm from the centre, (iii) the outer surface of the sphere, (iv) a point 15cm from the centre.

Solution

(a) Millikan sprayed a fine mist of oil drops between two horizontal parallel plates. He selected an individual charged drop under a microscope and adjusted the potential difference between the plates until the upward electric force on its charge balanced its downward weight. From $QE = mg$, with E known from the plate voltage and spacing and m found from the drop's terminal velocity in the absence of the field, he calculated the charge Q on each drop. Repeated measurements yielded charges that were always whole-number multiples of one smallest amount, the elementary charge $e \approx 1.6 \times 10^{-19}\text{C}$.

(b) Let $R_1 = 0.040\text{m}$, $R_2 = 0.100\text{m}$, $Q = 8.0 \times 10^{-9}\text{C}$.

Then $R_1^3 = 6.4 \times 10^{-5}\text{m}^3$, $R_2^3 = 1.0 \times 10^{-3}\text{m}^3$, $R_2^3 - R_1^3 = 9.36 \times 10^{-4}\text{m}^3$.

(i) At $r = 0.030\text{m} < R_1$: the Gaussian sphere lies inside the empty cavity and encloses no charge, so:

$$E = 0$$

(ii) At $r = 0.070\text{m}$ the point lies within the charged shell, $R_1 < r < R_2$. Apply the shell-region formula:

$$E = \frac{kQ(r^3 - R_1^3)}{r^2(R_2^3 - R_1^3)}$$

With $r^3 = 3.43 \times 10^{-4}\text{m}^3$, $r^3 - R_1^3 = 2.79 \times 10^{-4}\text{m}^3$, $r^2 = 4.9 \times 10^{-3}\text{m}^2$:

$$E = \frac{(9.0 \times 10^9 \text{Nm}^2/\text{C}^2) \times (8.0 \times 10^{-9}\text{C}) \times (2.79 \times 10^{-4}\text{m}^3)}{(4.9 \times 10^{-3}\text{m}^2) \times (9.36 \times 10^{-4}\text{m}^3)} \approx 4380\text{N/C}$$

(iii) At the outer surface, $r = R_2 = 0.100\text{m}$, the Gaussian encloses the whole charge, so:

$$E = \frac{kQ}{R_2^2} = \frac{(9.0 \times 10^9 \text{N} \cdot \text{m}^2/\text{C}^2) \times (8.0 \times 10^{-9}\text{C})}{(0.100\text{m})^2} = 7200\text{N/C}$$

(iv) At $r = 0.150\text{m} > R_2$:

$$E = \frac{kQ}{r^2} = \frac{(9.0 \times 10^9 \text{Nm}^2/\text{C}^2) \times (8.0 \times 10^{-9}\text{C})}{(0.150\text{m})^2} = 3200\text{N/C}$$

Example 41

- (a) Explain what is meant by the statement that electric charge is quantised, and give the numerical value of the smallest free charge in nature.
- (b) An isolated solid conducting sphere of radius 2.0cm carries a total positive charge of +8.0nC. Determine: (i) the number of electrons that have been removed from the sphere, (ii) the electric potential at the centre of the sphere, (iii) the electric potential at its surface, (iv) the electric potential at a point 5.0cm from the centre, (v) the electric potential at a point 20cm from the centre.

Solution

(a) Electric charge is quantised in the sense that the charge on any object is always an integer multiple of one smallest indivisible amount, namely the charge on a single electron (or, equivalently, on a single proton, with opposite sign). This smallest free charge in nature is the elementary charge $e = 1.6 \times 10^{-19}\text{C}$.

(b) Let $R = 0.020\text{m}$ and $Q = 8.0 \times 10^{-9}\text{C}$.

(i) Number of electrons removed:

$$n = \frac{Q}{e} = \frac{8.0 \times 10^{-9}\text{C}}{1.6 \times 10^{-19}\text{C}} = 5.0 \times 10^{10} \text{electrons}$$

(ii) Inside a conductor in equilibrium the field is zero, so the potential is constant throughout the interior and equals its value at the surface. At the centre:

$$V = \frac{kQ}{R} = \frac{(9.0 \times 10^9 \text{Nm}^2/\text{C}^2) \times (8.0 \times 10^{-9} \text{C})}{0.020\text{m}} = 3600\text{V}$$

(iii) The surface sits at the same constant interior value, $V = 3600\text{V}$.

(iv) At $r = 0.050\text{m} > R$, the conducting sphere looks from outside like a point charge Q at its centre:

$$V = \frac{kQ}{r} = \frac{(9.0 \times 10^9 \text{Nm}^2/\text{C}^2) \times (8.0 \times 10^{-9} \text{C})}{0.050\text{m}} = 1440\text{V}$$

(v) At $r = 0.20\text{m}$:

$$V = \frac{kQ}{r} = \frac{(9.0 \times 10^9 \text{Nm}^2/\text{C}^2) \times (8.0 \times 10^{-9} \text{C})}{0.20\text{m}} = 360\text{V}$$

Example 42

- (a) Explain conservation of charge as a named principle, and give an example of a process in which it must hold.
- (b) Two large parallel conducting plates are held a short distance apart in air. Each plate carries the same sign of charge, with a uniform surface charge density $\sigma = 4.0\mu\text{C}/\text{m}^2$. Determine the magnitude and direction of the electric field in: (i) the region to the left of both plates, (ii) the gap between the plates, (iii) the region to the right of both plates. Take $\epsilon_0 = 8.854 \times 10^{-12} \text{C}^2/(\text{N} \cdot \text{m}^2)$.

Solution

(a) Conservation of electric charge is a named principle of physics which states that the total electric charge of an isolated system is constant in time. This means charge may be transferred from one body to another or separated within a body, but it can be neither created nor destroyed. Example: when a glass rod is rubbed with silk, an equal magnitude of positive charge appears on the glass and negative charge on the silk; the total charge of the glass-and-silk system remains zero throughout the process.

(b) Each sheet alone produces a uniform field of magnitude $\frac{\sigma}{2\epsilon_0}$ on each of its sides, directed away from the sheet (positive charge). With the rightward direction taken as positive, superpose the two contributions in each region.

(i) To the left of both plates, both single-sheet fields point to the left:

$$E_{\text{left}} = -\frac{\sigma}{2\epsilon_0} - \frac{\sigma}{2\epsilon_0} = -\frac{\sigma}{\epsilon_0} = -\frac{(4.0 \times 10^{-6} \text{C}/\text{m}^2)}{(8.854 \times 10^{-12} \text{C}^2/(\text{Nm}^2))} = -4.5 \times 10^5 \text{N/C}$$

(ii) In the gap, the field of the left sheet points to the right (into the gap) and the field of the right sheet points to the left (also into the gap). The two contributions are equal in magnitude but opposite in direction:

$$E_{\text{gap}} = +\frac{\sigma}{2\epsilon_0} - \frac{\sigma}{2\epsilon_0} = 0$$

(iii) To the right of both plates, both single-sheet fields point to the right:

$$E_{\text{right}} = +\frac{\sigma}{2\epsilon_0} + \frac{\sigma}{2\epsilon_0} = +\frac{\sigma}{\epsilon_0} = +\frac{(4.0 \times 10^{-6} \text{C}/\text{m}^2)}{(8.854 \times 10^{-12} \text{C}^2/(\text{Nm}^2))} = +4.5 \times 10^5 \text{N/C}$$

The field is $4.5 \times 10^5 \text{N/C}$ directed away from the pair on each side, and zero in the gap between the plates.

Example 43

- (a) Explain the three methods of charging a neutral conductor, naming each and stating in one sentence what happens to the charge in each case.
- (b) A long straight wire of negligible thickness carries a uniform positive linear charge density of $5.0\text{nC}/\text{m}$. Treating the wire as effectively infinite, determine: (i) the magnitude of the electric field at a perpendicular distance of 10cm from the axis, (ii) the potential difference between two points at perpendicular distances of 5.0cm and 20cm respectively from the axis, stating which point is at the higher potential.

Solution

(a) The methods are:

1. Charging by friction: rubbing two suitable insulators together transfers electrons from one to the other, leaving the two with equal and opposite net charges.

2. Charging by contact: a charged body touched briefly to a neutral conductor shares its charge with the conductor, and on separation each carries part of the original charge.

3. Charging by induction: a charged body brought near a neutral conductor without touching it polarises the conductor; momentarily earthing the far end drains the unwanted sign, and on disconnection the conductor is left with a net charge opposite in sign to the inducing body.

(b) (i) Field of the infinite line at perpendicular distance r :

$$E = \frac{\lambda}{2\pi\epsilon_0 r} = \frac{2k\lambda}{r} = \frac{2 \times (9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times (5.0 \times 10^{-9} \text{ C/m})}{(0.10 \text{ m})} = 900 \text{ N/C}$$

directed radially outward from the wire.

(ii) Potential difference between a point at distance r_a and one at distance r_b :

$$V_a - V_b = \frac{\lambda}{2\pi\epsilon_0} \ln\left(\frac{r_b}{r_a}\right) = 2k\lambda \ln\left(\frac{r_b}{r_a}\right)$$

With $r_a = 0.050 \text{ m}$ and $r_b = 0.20 \text{ m}$:

$$V_a - V_b = 2 \times \left(9.0 \times \frac{10^9 \text{ Nm}^2}{\text{C}^2}\right) \times \left(5.0 \times \frac{10^{-9} \text{ C}}{\text{m}}\right) \times \ln\left(\frac{0.20 \text{ m}}{0.050 \text{ m}}\right) = 124.7 \text{ V}$$

The point at 5.0cm sits at the higher potential, since the potential of a positively charged line falls as one moves outward.

Example 44

- (a) Explain why electric field lines never cross each other and never form closed loops in electrostatics.
 (b) Four point charges are fixed at the corners of a square ABCD of side 1.0m, arranged so that A is at the top-left, B at the top-right, C at the bottom-right and D at the bottom-left. The charges are: at A, $+2.0 \mu\text{C}$; at B, $-3.0 \mu\text{C}$; at C, $+1.0 \mu\text{C}$; at D, $+4.0 \mu\text{C}$. Determine the magnitude and direction of the resultant electric field at the centre O of the square.

Solution

(a) Field lines never cross because the electric field at any point has one definite direction; if two lines crossed at a point, the field there would have two directions at once, which is impossible. Field lines never form closed loops in electrostatics because the electrostatic field is conservative; the work done by the field on a test charge taken around any closed path is zero, which forbids a continuous loop of field along which the work would accumulate.

(b) Each corner of the square lies on a diagonal that passes through the centre O. The distance from any corner to O is half the diagonal:

$$r = \frac{1.0 \text{ m}}{\sqrt{2}} \quad r^2 = \frac{(1.0 \text{ m})^2}{2} = 0.50 \text{ m}^2$$

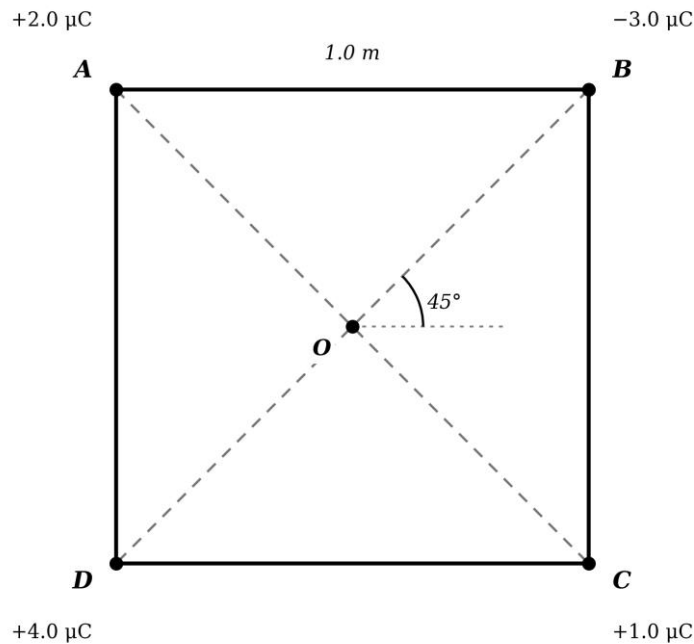


Figure: Square ABCD with charges at each corner and centre O. Both diagonals (dashed) pass through O. Each diagonal makes 45° with the sides of the square, so every field vector at O lies along a diagonal and resolves into components by $\cos 45^\circ$ and $\sin 45^\circ$.

Field magnitudes at O due to each charge (using $E = k|Q|/r^2$):

$$E_A = \frac{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times (2.0 \times 10^{-6} \text{ C})}{(0.50 \text{ m})^2} = 3.6 \times 10^4 \text{ N/C}$$

$$E_B = \frac{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times (3.0 \times 10^{-6} \text{ C})}{(0.50 \text{ m})^2} = 5.4 \times 10^4 \text{ N/C}$$

$$E_C = \frac{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times (1.0 \times 10^{-6} \text{ C})}{(0.50 \text{ m})^2} = 1.8 \times 10^4 \text{ N/C}$$

$$E_D = \frac{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times (4.0 \times 10^{-6} \text{ C})}{(0.50 \text{ m})^2} = 7.2 \times 10^4 \text{ N/C}$$

Take +x along the line from A to B (parallel to side AB) and +y along the line from D to A (parallel to side AD). Each field vector at O lies along a diagonal and therefore makes 45° with both axes.

Field due to A. A is positive, so E_A at O points away from A, toward the lower-right corner. Its x-component is positive (rightward), its y-component negative (downward):

$$E_{A,x} = +E_A \cos 45^\circ = +(3.6 \times 10^4 \text{ N/C}) \times (0.7071) = +2.55 \times 10^4 \text{ N/C}$$

$$E_{A,y} = -E_A \sin 45^\circ = -(3.6 \times 10^4 \text{ N/C}) \times (0.7071) = -2.55 \times 10^4 \text{ N/C}$$

Field due to B. B is negative, so E_B at O points toward B (upper-right). Both components positive:

$$E_{B,x} = +E_B \cos 45^\circ = +(5.4 \times 10^4 \text{ N/C}) \times (0.7071) = +3.82 \times 10^4 \text{ N/C}$$

$$E_{B,y} = +E_B \sin 45^\circ = +(5.4 \times 10^4 \text{ N/C}) \times (0.7071) = +3.82 \times 10^4 \text{ N/C}$$

Field due to C. C is positive, so E_C at O points away from C, toward the upper-left. x-component negative, y-component positive:

$$E_{C,x} = -E_C \cos 45^\circ = -(1.8 \times 10^4 \text{ N/C}) \times (0.7071) = -1.27 \times 10^4 \text{ N/C}$$

$$E_{C,y} = +E_C \sin 45^\circ = +(1.8 \times 10^4 \text{ N/C}) \times (0.7071) = +1.27 \times 10^4 \text{ N/C}$$

Field due to D. D is positive, so E_D at O points away from D, toward the upper-right. Both components positive:

$$E_{D,x} = +E_D \cos 45^\circ = +(7.2 \times 10^4 \text{ N/C}) \times (0.7071) = +5.09 \times 10^4 \text{ N/C}$$

$$E_{D,y} = +E_D \sin 45^\circ = +(7.2 \times 10^4 \text{ N/C}) \times (0.7071) = +5.09 \times 10^4 \text{ N/C}$$

Summing components:

$$\Sigma E_x = (2.55 + 3.82 - 1.27 + 5.09) \times 10^4 \text{ N/C} = +10.18 \times 10^4 \text{ N/C}$$

$$\Sigma E_y = (-2.55 + 3.82 + 1.27 + 5.09) \times 10^4 \text{ N/C} = +7.64 \times 10^4 \text{ N/C}$$

Magnitude of the resultant:

$$E = \sqrt{(10.18 \times 10^4)^2 + (7.64 \times 10^4)^2} \text{ N/C} = 1.27 \times 10^5 \text{ N/C}$$

Direction (angle above the +x axis):

$$\theta = \tan^{-1} \left(\frac{7.64}{10.18} \right) = 36.87^\circ$$

The resultant field at O has magnitude $1.27 \times 10^5 \text{ N/C}$, directed at 36.87° above the line AB, on the side of corner B.

Example 45

- (a) A bird perched on a single high-voltage overhead transmission line is unharmed; a person who simultaneously touches the same line and a metal post is electrocuted. Account for this difference.
- (b) Three point charges are placed at the vertices A, B and C of an equilateral triangle of side 60 cm. The charge at A is $+5.0 \text{ nC}$, at B is -3.0 nC , and at C is $+4.0 \text{ nC}$. Calculate the electric potential at the midpoint M of side AB, taking the reference of zero potential at infinity.

Solution

(a) A bird perched on a single overhead wire rests both feet on essentially the same point of the conductor. The conductor is an equipotential, so the potential difference between the bird's feet is zero, no current flows through the bird and it is unharmed. A person standing on the ground and simultaneously touching the live wire and an earthed metal post is held at two very different potentials: the wire at tens of thousands of volts and the ground at zero. The large potential difference drives a substantial current through the person's body, with fatal consequences.

(b) M is equidistant from A and B, at a distance equal to half the side of the triangle:

$$r_{AM} = r_{BM} = \frac{0.60 \text{ m}}{2} = 0.30 \text{ m}$$

Vertex C lies at the apex; its distance to M is the altitude of the equilateral triangle from C:

$$r_{CM} = (0.60 \text{ m}) \times \frac{\sqrt{3}}{2} = (0.30 \text{ m}) \times \sqrt{3} = 0.520 \text{ m}$$

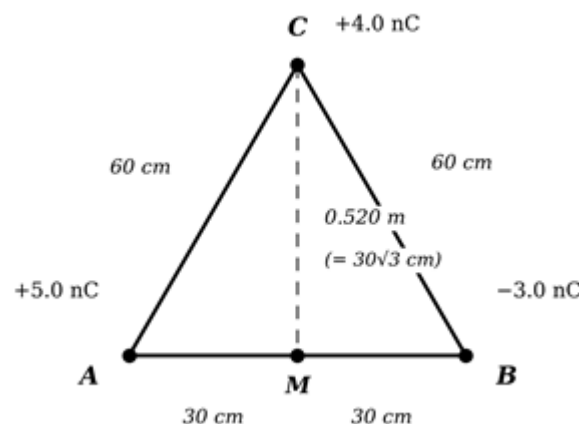


Figure: Equilateral triangle ABC of side 60 cm. The charges at the vertices are $+5.0 \text{ nC}$, -3.0 nC and $+4.0 \text{ nC}$ at A, B and C respectively. M is the midpoint of side AB; $AM = BM = 30 \text{ cm}$. The altitude from C to M has length $30\sqrt{3} \text{ cm} = 0.520 \text{ m}$.

Potential at M is the scalar sum of the three contributions:

$$V = k \left(\frac{Q_A}{r_{AM}} + \frac{Q_B}{r_{BM}} + \frac{Q_C}{r_{CM}} \right)$$

$$V = (9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times \left(\frac{(+5.0 \times 10^{-9} \text{ C})}{0.30 \text{ m}} + \frac{(-3.0 \times 10^{-9} \text{ C})}{0.30 \text{ m}} + \frac{(+4.0 \times 10^{-9} \text{ C})}{0.520 \text{ m}} \right) = 129.2 \text{ V}$$

The potential at the midpoint of AB is approximately 129V.

Example 46

- (a) Cling-film stretched off the roll sometimes pulls itself onto your hand from a distance of a centimetre. Explain this observation.
- (b) Two point charges $Q_1 = +5.0 \text{ nC}$ and $Q_2 = -5.0 \text{ nC}$ are fixed at two points 4.0cm apart on a straight line. A small particle of mass $8.0 \mu\text{g}$ carrying a charge of $+3.0 \mu\text{C}$ is released from rest at a point a on the same line, 1.0cm from Q_1 . Calculate the speed of the particle when it reaches point b, located on the same line 1.0cm from Q_2 . Neglect gravity.

Solution

(a) Pulling the cling-film off the roll involves friction between the film and the surface of the roll, which separates charge: the film comes away with a net charge of one sign and the roll keeps an equal and opposite charge. Brought close to the neutral hand, the charged film polarises the hand by induction, drawing opposite charge toward the surface of the hand nearer the film. The film's charge and the hand's nearer induced charge are of opposite sign, so they attract across the gap and pull the film onto the hand even before contact.

(b) The particle starts at point a with $r_{1a} = 0.010 \text{ m}$ and $r_{2a} = 0.030 \text{ m}$, and ends at point b with $r_{1b} = 0.030 \text{ m}$ and $r_{2b} = 0.010 \text{ m}$.

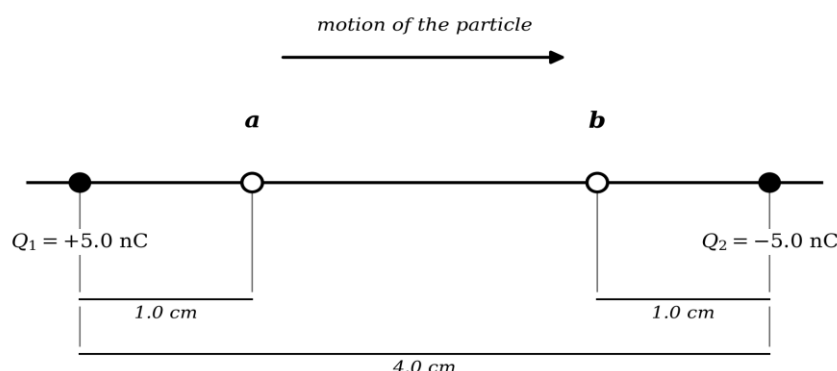


Figure: The two fixed charges sit at the ends of a 4.0 cm line; Q_1 is at the left, Q_2 at the right. The moving particle starts at point a (1.0 cm from Q_1) and ends at point b (1.0 cm from Q_2). Open circles distinguish the trajectory endpoints from the fixed charges.

Potential at point a:

$$V_a = k \left(\frac{Q_1}{r_{1a}} + \frac{Q_2}{r_{2a}} \right)$$

$$V_a = (9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times \left(\frac{(+5.0 \times 10^{-9} \text{ C})}{0.010 \text{ m}} + \frac{(-5.0 \times 10^{-9} \text{ C})}{0.030 \text{ m}} \right) = +3000 \text{ V}$$

Potential at point b (by symmetry of the geometry):

$$V_b = (9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times \left(\frac{(+5.0 \times 10^{-9} \text{ C})}{0.030 \text{ m}} + \frac{(-5.0 \times 10^{-9} \text{ C})}{0.010 \text{ m}} \right) = -3000 \text{ V}$$

Work done by the electric field on the particle from a to b:

$$W = q(V_a - V_b) = (3.0 \times 10^{-6} \text{ C}) \times (3000 \text{ V} - (-3000 \text{ V})) = 0.018 \text{ J}$$

By the work-energy theorem, this work equals the kinetic energy gained, since the particle starts at rest:

$$\frac{1}{2}mv^2 = W = 0.018\text{J}$$

$$v = \sqrt{\frac{2W}{m}} = \sqrt{\frac{2 \times (0.018\text{J})}{(8.0 \times 10^{-9}\text{kg})}} = 2.1 \times 10^3\text{m/s}$$

The particle reaches point b with a speed of approximately $2.1 \times 10^3\text{m/s}$.

Example 47

- (a) A petrol tanker truck is required by law to drag a metal chain along the road as it moves. Explain why the chain is essential.
- (b) An electron, initially at rest, is accelerated through a potential difference of 300V in vacuum. Determine: (i) the kinetic energy acquired by the electron, (ii) its speed at the end of the acceleration. Take the electron mass $m_e = 9.11 \times 10^{-31}\text{kg}$ and the elementary charge $e = 1.6 \times 10^{-19}\text{C}$.

Solution

(a) As the tanker travels, friction between the tyres and the road, and between the moving petrol and the metal walls of the tank, separates charge and leaves the tanker with a net accumulated charge. If this charge is allowed to build up unchecked, the tanker reaches a potential at which a spark can jump between it and a nearby earthed object, such as a petrol pump or the ground, at the moment of stopping. Petrol vapour around the tanker is highly flammable, and a spark would ignite it. The trailing metal chain provides a continuous conducting path from the tanker to the ground, draining the accumulated charge to earth as fast as friction deposits it, so the tanker stays essentially at ground potential and no spark can form.

- (b) (i) Work done by the field on the electron as it is accelerated through $\Delta V = 300\text{V}$:

$$W = e\Delta V = (1.6 \times 10^{-19}\text{C}) \times (300\text{V}) = 4.8 \times 10^{-17}\text{J}$$

Since the electron starts at rest, all this work appears as kinetic energy:

$$\text{KE} = W = 4.8 \times 10^{-17}\text{J}$$

- (ii) From $\text{KE} = \frac{1}{2}m_ev^2$:

$$v = \sqrt{\frac{2\text{KE}}{m_e}} = \sqrt{\frac{2 \times (4.8 \times 10^{-17}\text{J})}{(9.11 \times 10^{-31}\text{kg})}} = 1.0 \times 10^7\text{m/s}$$

The electron acquires a kinetic energy of $4.8 \times 10^{-17}\text{J}$ and reaches a speed of approximately $1.0 \times 10^7\text{m/s}$ (about 3.3% of the speed of light).

Example 48

- (a) Explain why the factor $\frac{1}{4\pi\epsilon_0}$ appears in Coulomb's law in SI units.
- (b) Two equal positive point charges, each $Q = 8.0\text{nC}$, are fixed in vacuum at the points $x = -a$ and $x = +a$ on a horizontal line, with $a = 4.0\text{cm}$ and the origin O at the midpoint between them. A small bead of mass $m = 1.0\text{g}$, carrying a positive charge $q = +8.0\text{nC}$, is threaded on a smooth insulating wire that lies along the x-axis between the two fixed charges and can slide along the wire without friction. (i) Show that the bead is in equilibrium at O. (ii) Show that for small displacements x along the wire from O, the bead executes simple harmonic motion, and derive an expression for the period of small oscillations. (iii) Calculate the period numerically.

Solution

- (a) The factor $\frac{1}{4\pi\epsilon_0}$ appears in Coulomb's law in SI units because the SI system is built upon Gauss's law as the defining relation between charge and field. By Gauss's law, the total electric flux from a point charge Q through any closed surface that encloses it is Q/ϵ_0 ; spreading this flux uniformly over a sphere of surface area $4\pi r^2$ gives $E = Q/(4\pi\epsilon_0 r^2)$. The 4π is the geometric factor for the surface area of a sphere, and the $1/\epsilon_0$ is the dimensional factor that converts a charge in coulombs into a field in V/m. In Gaussian (CGS) units the constant is absorbed into the definition of the unit of charge and Coulomb's law reduces to $F = \frac{q_1 q_2}{r^2}$; the explicit $\frac{1}{4\pi\epsilon_0}$ of SI is the cost of having defined the coulomb as the fundamental unit of charge.

(b) (i) Place $+x$ along the wire, with O at the origin. The bead at position x (with $|x| < a$) feels two repulsive Coulomb forces: from the charge at $-a$, a force in the $+x$ direction with magnitude $kQq/(x+a)^2$; from the charge at $+a$, a force in the $-x$ direction with magnitude $kQq/(a-x)^2$. The net force, with $+x$ positive, is:

$$F(x) = kQq \left(\frac{1}{(x+a)^2} - \frac{1}{(a-x)^2} \right)$$

At $x = 0$, $F(0) = kQq \left[\frac{1}{a^2} - \frac{1}{a^2} \right] = 0$, so the midpoint is an equilibrium position.

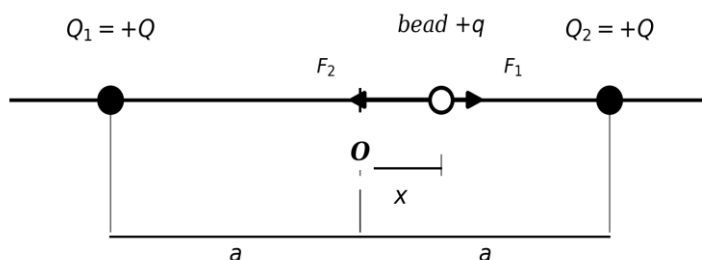


Figure: The bead $+q$ sits on a smooth insulating wire between the two fixed charges $Q_1 = Q_2 = +Q$. At a small displacement x to the right of O , the rightward force F_1 from the more distant charge Q_1 at $-a$ is smaller than the leftward force F_2 from the nearer charge Q_2 at $+a$, so the net force on the bead is restoring (toward O).

(ii) For small x (with $|x| \ll a$), expand each denominator to first order in x/a :

$$\frac{1}{(x+a)^2} \approx \frac{1}{a^2} \left(1 - \frac{2x}{a} \right), \quad \frac{1}{(a-x)^2} \approx \frac{1}{a^2} \left(1 + \frac{2x}{a} \right)$$

Substituting:

$$F(x) = \frac{kQq}{a^2} \left(1 - \frac{2x}{a} - 1 - \frac{2x}{a} \right) = -\frac{4kQq}{a^3} x$$

This has the form $F = -k_{\text{eff}} x$, with effective spring constant $k_{\text{eff}} = \frac{4kQq}{a^3}$.

By Newton's second law, $F = ma = m \frac{d^2x}{dt^2} = -k_{\text{eff}} x$, which is the equation of simple harmonic motion with angular frequency:

$$\omega^2 = \frac{k_{\text{eff}}}{m} = \frac{4kQq}{ma^3}$$

Period of small oscillations:

$$T = \frac{2\pi}{\omega} = 2\pi \sqrt{\frac{ma^3}{4kQq}} = \pi \sqrt{\frac{ma^3}{kQq}}$$

(iii) Substitute the numerical values:

$$T = \pi \sqrt{\frac{(1.0 \times 10^{-3} \text{ kg}) \times (0.040 \text{ m})^3}{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times (8.0 \times 10^{-9} \text{ C}) \times (8.0 \times 10^{-9} \text{ C})}} = 1.05 \text{ s}$$

The bead oscillates about O with a period of approximately 1.05s.

Example 49

- Explain why two independent natural light sources cannot produce a sustained spark-pattern in the way that two charged spheres can produce a sustained Coulomb force.
- An electron travelling horizontally at $v = 2.0 \times 10^7 \text{ m/s}$ enters the region between two horizontal parallel metal plates. The plates are $L = 4.0 \text{ cm}$ long and $d = 2.0 \text{ cm}$ apart, and a constant potential

difference of 200V is maintained between them, with the upper plate positive. The electron enters midway between the plates, parallel to them, and the field outside the plate region is negligible. Take $m_e = 9.11 \times 10^{-31}\text{kg}$ and $e = 1.6 \times 10^{-19}\text{C}$. Determine: (i) the magnitude and direction of the electric field between the plates, (ii) the magnitude and direction of the electric force on the electron, (iii) the magnitude of its acceleration, (iv) the time spent between the plates, (v) the vertical deflection of the electron at the moment it emerges from the plates, and verify that it does not strike either plate during transit, (vi) the angle its velocity makes with the horizontal as it emerges.

Solution

(a) Light from two ordinary independent sources consists of wave-trains whose phases vary randomly and very rapidly with respect to one another, much faster than the eye or any practical detector can resolve. At each instant the two waves produce some interference pattern, but the pattern shifts position with every phase fluctuation and is washed out in the time-average; no sustained pattern is observed unless the sources are phase-locked, i.e., coherent. The electrostatic interaction between two charged spheres is not a wave phenomenon at all: it is a static configuration of field vectors in space, fixed entirely by the positions and magnitudes of the charges, with no phase relationship to vary in time. The contrast is the basic difference between an electrostatic field, which is well-defined at every point in space, and an oscillating wave field, where steady patterns require coherence between sources.

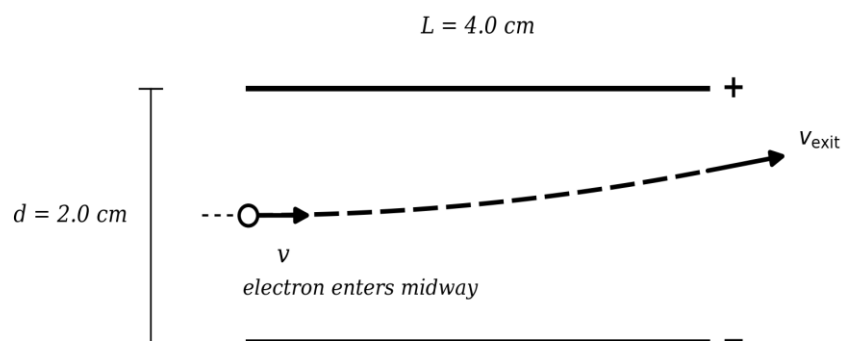


Figure: The electron enters midway between the plates moving horizontally with speed v . The uniform electric field in the gap points downward (from $+$ to $-$), so the force on the negatively charged electron is upward. The trajectory inside the plates is a parabolic curve toward the upper (positive) plate. On exit, the velocity is inclined at an angle θ above the horizontal.

(b) (i) Field strength between the plates:

$$E = \frac{\Delta V}{d} = \frac{(200\text{V})}{(0.020\text{m})} = 1.0 \times 10^4 \text{V/m}$$

directed from the upper (+) plate to the lower (-), i.e., vertically downward.

(ii) The electron carries charge $-e$; the electric force is $\vec{F} = (-e)\vec{E}$, opposite to E , i.e., vertically upward (toward the positive plate). Magnitude:

$$F = eE = (1.6 \times 10^{-19}\text{C}) \times (1.0 \times 10^4 \text{V/m}) = 1.6 \times 10^{-15}\text{N}$$

(iii) Acceleration:

$$a = \frac{F}{m_e} = \frac{(1.6 \times 10^{-15}\text{N})}{(9.11 \times 10^{-31}\text{kg})} = 1.76 \times 10^{15} \text{m/s}^2$$

directed vertically upward.

(iv) Horizontal motion is uniform at v ; time to traverse the length of the plates:

$$t = \frac{L}{v} = \frac{(0.040\text{m})}{(2.0 \times 10^7 \text{m/s})} = 2.0 \times 10^{-9} \text{s}$$

(v) Vertical motion is uniform acceleration from rest in the vertical direction (entry parallel to the plates, $v_{y,0} = 0$):

$$y = \frac{1}{2}at^2 = \frac{1}{2} \times (1.76 \times 10^{15} \text{m/s}^2) \times (2.0 \times 10^{-9} \text{s})^2 = 3.5 \times 10^{-3} \text{m}$$

That is, $y = 3.5 \text{mm}$.

The electron entered midway between plates that are $d = 2.0 \text{cm}$ apart, so the clearance to either plate was $d/2 = 10 \text{mm}$. Since $y < d/2$, the electron emerges from the plates without striking either.

(vi) Vertical velocity at exit:

$$v_y = at = (1.76 \times 10^{15} \text{m/s}^2) \times (2.0 \times 10^{-9} \text{s}) = 3.51 \times 10^6 \text{m/s}$$

Angle with the horizontal:

$$\theta = \tan^{-1} \left(\frac{v_y}{v} \right) = \tan^{-1} \left(\frac{(3.51 \times 10^6 \text{m/s})}{(2.0 \times 10^7 \text{m/s})} \right) = 9.96^\circ \approx 10^\circ$$

The electron emerges from the plates deflected 3.5mm upward, with its velocity inclined at 10° above the horizontal.

Example 50

- (a) Photocopiers and laser printers use electrostatic attraction of toner particles to a charged drum. Explain in two sentences how the toner forms a pattern that becomes the image on the paper.
- (b) Three point charges $+q$, $+q$, and $-2q$ are fixed at the vertices A, B, C of an equilateral triangle of side $a = 30 \text{cm}$, with the $-2q$ charge at vertex C. Take $q = 5.0 \text{nC}$. (i) Determine the total electric potential energy of this configuration. (ii) A fourth charge $+q$ is now brought from infinity to the centroid G of the triangle. State, with reasoning, the work that must be done by an external agent to achieve this.

Solution

(a) A photoconductive drum is first charged uniformly, then exposed to a light pattern representing the image: where light strikes the drum the surface becomes conducting and the charge there leaks to ground, leaving the unexposed regions still charged. Charged toner particles of opposite sign are attracted only to those still-charged regions and transfer the pattern to the drum, from where the toner is transferred to paper and fused with heat.

(b) Consider the following diagram:

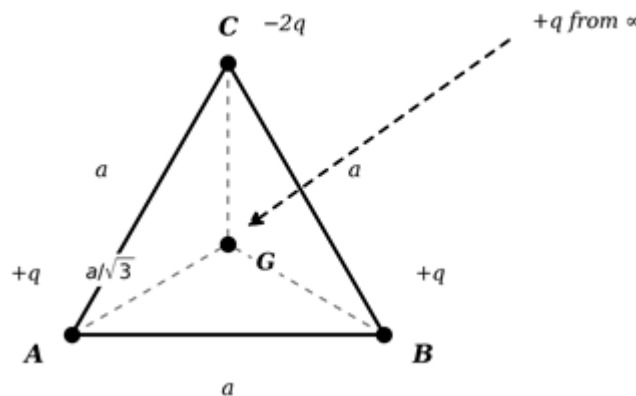


Figure: Equilateral triangle ABC of side a . The charges at the vertices are $+q$, $+q$ and $-2q$. The centroid G of the triangle is equidistant from all three vertices, with $r_G = a/\sqrt{3}$. The dashed arrow indicates the fourth charge being brought from infinity to G.

(i) The total electric potential energy of a system of point charges is the scalar sum of the pair-wise potential energies $U_{ij} = kQ_iQ_j/r_{ij}$ over all distinct pairs. For three charges on an equilateral triangle of side a , all three pair separations equal a :

$$U_{AB} = \frac{k(+q)(+q)}{a} = +\frac{kq^2}{a}$$

$$U_{AC} = \frac{k(+q)(-2q)}{a} = -\frac{2kq^2}{a}$$

$$U_{BC} = \frac{k(+q)(-2q)}{a} = -\frac{2kq^2}{a}$$

Sum:

$$U = U_{AB} + U_{AC} + U_{BC} = \frac{kq^2}{a} - \frac{2kq^2}{a} - \frac{2kq^2}{a} = -\frac{3kq^2}{a}$$

Numerically:

$$U = -3 \times \frac{(9.0 \times 10^9 \text{ Nm}^2/\text{C}^2) \times (5.0 \times 10^{-9} \text{ C})^2}{(0.30 \text{ m})} = -2.25 \times 10^{-6} \text{ J} = -2.25 \mu\text{J}$$

The negative value tells us the system is bound: an external agent would have to do positive work of $2.25 \mu\text{J}$ against the net electrostatic forces to disassemble the configuration to infinite mutual separation.

(ii) The work an external agent must do in bringing a charge q' from infinity to a point P is $W = q'V_P$, where V_P is the electric potential at P due to all other charges already in place. The centroid G of an equilateral triangle is equidistant from all three vertices, at distance:

$$r_G = \frac{a}{\sqrt{3}} = \frac{0.30 \text{ m}}{\sqrt{3}} = 0.173 \text{ m}$$

Potential at G due to the three vertex charges:

$$V_G = \frac{k}{r_G}((+q) + (+q) + (-2q)) = \frac{k}{r_G} \times 0 = 0$$

The contributions from the two $+q$ charges and the one $-2q$ charge cancel exactly at the centroid because the three distances are equal and the charges sum to zero. Therefore:

$$W = qV_G = q \times 0 = 0 \text{ J}$$

No work is done by the external agent in bringing the fourth charge from infinity to the centroid.

The result is counter-intuitive (three nearby charges and yet zero work), but is a clean consequence of the equidistance of G from the three vertices and the fact that the three charges sum to zero.

From the basic fact that charge exists only in discrete quanta, through Coulomb's law and the field it generates, through Gauss's law applied to lines, sheets, and spheres, to the electric potential and the work-energy bridge $W = q\Delta V$, the chapter has built electrostatics step by step, and the fifteen miscellaneous examples have exercised those ideas in combination. With the topic now viewed as a whole, we are well prepared to move on to the Digging Deeper Exercise on the next page.

DIGGING DEEPER EXERCISE 10

EXERCISE 10A: BINDER QUESTIONS

Question 1

When two or more point charges are present, the resultant electric field at a point in space is found by vector addition of the contributions of the individual charges, but the resultant electric potential at the same point is found by scalar (arithmetic) addition. Explain why the two quantities are combined by different rules.

Question 2

Explain in your own words what an electric field is, and how its strength at a point is operationally measured.

Question 3

On the surface of an irregularly shaped charged conductor in electrostatic equilibrium, the surface charge density is greater at regions of high curvature, such as sharp points and edges, than at regions of low curvature, such as broad flat surfaces. Account for this concentration of charge.

Question 4

Explain why the electric field inside an empty cavity of a hollow charged conductor in electrostatic equilibrium is zero, even though the surface of the conductor carries a non-zero charge.

Question 5

A person who has walked across a dry carpet sometimes draws a small visible spark when reaching for a metal doorknob. Explain how this is consistent with the idea that air has a finite dielectric strength.

Question 6

Provide the statement of Gauss's law for electric fields, and explain in one or two sentences why the law is particularly useful for finding the field of a charge distribution that possesses a high degree of symmetry.

Question 7

Explain what is meant by an equipotential surface, and explain three of its properties.

Question 8

Explain why no work is done by the electric field in moving a test charge along an equipotential surface, by giving two independent arguments (one in terms of the potential difference, and one in terms of the directions of the field and of the displacement).

Question 9

The relationship between the electric field strength and the electric potential along a chosen direction is $E = -\frac{dV}{dr}$. Explain this relationship in words, and explain what the negative sign means physically.

Question 10

A conductor connected to the Earth by a wire is said to be earthed, and is assigned a potential of zero. Explain (i) why the Earth is treated as the universal reference of zero potential, and (ii) what happens physically to the charge on a previously charged conductor at the moment it is earthed.

EXERCISE 10B: REAL QUESTIONS

Question 11

Kipanga: “I noticed that the lightning during last night’s storm struck the tall coconut tree behind our house in Mtwara, not the shorter mango tree right next to it.”

Mr. Akilikubwa: “You have noticed something every lightning conductor in the world is designed around.”

Explain why a lightning bolt is far more likely to strike a tall pointed object than a shorter one standing close beside it.

Question 12

Kipute pulls a wool jumper over her head on a dry day and her hair stands out as straight wisps that refuse to lie flat.

Kipanga: *“The hairs are pushing each other apart.”*

Mr. Akilikubwa: *“You are half right.”*

Explain Kipanga’s correct point, and the second mechanism that Mr. Akilikubwa is adding.

Question 13

A lightning conductor on a tall building has two unmistakable features: a sharp metal point projecting above the roof, and a thick copper strap leading from the point down to a heavy metal plate buried deep in moist soil at the building’s foot. Explain the purpose of each of the two features.

Question 14

A car parked in the open during a thunderstorm is sometimes struck by lightning, yet the people sitting inside it are unharmed. Explain why the metal body of the car protects them.

Question 15

An inkjet printer forms each letter on the page by firing a stream of tiny ink droplets between a pair of charged metal plates. Each droplet is given a controlled amount of charge as it leaves the nozzle, and the print head varies that charge from droplet to droplet. Explain how varying the charge on each droplet determines where on the page it lands.

Question 16

At the Tanga Cement factory, the hot flue gas leaving the kiln passes through an electrostatic precipitator before reaching the chimney. Inside the precipitator, a very high voltage is applied between a central wire and the surrounding earthed collecting plates, and the dust and ash particles are removed from the gas. Explain how the precipitator removes the dust.

Question 17

At an airport, the operating manual for a fuel truck instructs the operator to clip a heavy bonding cable from the truck to the body of the aircraft before opening the fuel valve, and to remove the cable only after the valve has been closed. Explain why the bonding cable is essential, even when the operator is standing on dry tarmac in rubber boots.

Question 18

When lightning strikes the ground close to the foot of a tall isolated tree, a person sheltering beneath that tree may be killed by the strike even though the lightning did not pass through the body. Explain why the ground itself becomes dangerous near the point at which the lightning enters it.

Question 19

A plastic comb rubbed briskly on the hair retains its charge for many minutes and continues to attract small pieces of paper for a long time. A clean metal rod, held in the bare hand and rubbed in exactly the same way on the same hair, cannot be made to attract paper at all. Explain the difference.

Question 20

In a school physics demonstration in Mwanza, the teacher holds one end of a long fluorescent tube near the highly charged metal dome of a Van de Graaff generator without making contact, and without connecting the tube to any electrical supply. The tube glows along its length with a faint flickering light. Explain how the tube can glow without being plugged in.

EXERCISE 10C: HOT QUESTIONS

Question 21

A point charge of $+4.0\mu\text{C}$ is placed at the origin of an x-axis, and a second point charge of $-2.0\mu\text{C}$ is placed at $x = 0.30\text{m}$. Determine the position on the x-axis at which the net electric field due to the two charges is zero.

Question 22

Four point charges of magnitudes $+2.0\mu\text{C}$, $-3.0\mu\text{C}$, $+2.0\mu\text{C}$ and $-3.0\mu\text{C}$ are placed in that cyclic order at the four corners of a square of side 0.10m . Find the magnitude of the net electric field at the centre of the square.

Question 23

In the right-angled triangle ABC, the right angle is at vertex B, with $AB = 3.0\text{cm}$ and $BC = 4.0\text{cm}$. A point charge of $+6.0\text{nC}$ is fixed at A and a point charge of $+8.0\text{nC}$ is fixed at C. Find the magnitude and direction of the resultant electric field at vertex B.

Question 24

Identical point charges of magnitude $+Q$ are placed along the x-axis at the positions $x = 1.0\text{m}$, 2.0m , 4.0m , 8.0m , and so on, with each successive charge at twice the distance of its predecessor from the origin. Find the magnitude of the electric field at the origin in terms of kQ .

Question 25

A uniformly charged conducting sphere of diameter 0.60m carries a uniform surface charge density of $\sigma = 50\mu\text{C}/\text{m}^2$. Determine (i) the magnitude of the electric field just outside the surface of the sphere, (ii) the total charge on the sphere, and (iii) the total electric flux through any closed surface that completely encloses the sphere.

Question 26

A student calculates the electric field at a point inside a uniformly charged insulating solid sphere of total charge Q and radius R , at a perpendicular distance r from the centre (with $r < R$), by writing $E = kQ/r^2$. Identify the conceptual error in the student's reasoning, and derive the correct expression for the magnitude of the electric field inside the sphere at radius r .

Question 27

Two large parallel sheets carry uniform surface charge densities of $+\sigma$ and $-\sigma$ respectively, with $\sigma = 6.0\mu\text{C}/\text{m}^2$. Find the magnitude of the electric field (i) in the region between the two sheets, and (ii) in the region outside the sheets.

Question 28

Three point charges of magnitudes Q , $+q$ and $+q$ are placed at the vertices of an equilateral triangle of side L . Given that the total electrostatic potential energy of the three-charge system is zero, find Q in terms of q .

Question 29

Two thin concentric spherical shells of radii r and R (with $r < R$) each carry the same uniform surface charge density σ . Find the electric potential at the common centre of the two shells, taking the potential to be zero at infinity.

Question 30

Two equal positive charges, each of magnitude Q , are fixed on the x-axis at the positions $x = +a$ and $x = -a$. A small particle of charge $+q$ and mass m is placed at the midpoint between them, $x = 0$. The particle is given a small displacement x along the x-axis and released. (a) Show that the resulting motion is simple harmonic. (b) Find the angular frequency ω in terms of Q , q , m , a , and Coulomb's constant k .

Question 31

A student computes the work done by the electrostatic field in moving a charge of 3.0C along a closed loop that lies entirely on an equipotential surface of value 12V . The student writes $W = qV = (3.0\text{C})(12\text{V}) = 36\text{J}$. Give the correct value of the work done, and identify the conceptual error in the student's reasoning.

Question 32

Two horizontal parallel plates are placed 8.0mm apart and are maintained at a potential difference of 980V , with the upper plate at the higher potential. A small negatively charged oil drop of mass $1.0 \times 10^{-14}\text{kg}$ is observed to be at rest between the plates. Calculate the number of excess electrons on the drop. Take $g = 9.8\text{m/s}^2$ and $e = 1.6 \times 10^{-19}\text{C}$.

Question 33

Two identical small pith balls, each of mass m and carrying the same charge Q , are suspended by silk threads of equal length L from a common point. In air the two threads come to rest making an angle of 30° between them. When the entire system is then immersed in a non-conducting oil of relative permittivity K , the threads come to rest making an angle of 18° between them. Assuming small-angle equilibrium throughout, and that the buoyancy on the balls in the oil is negligible, find K .

Question 34

An electrostatic precipitator at a Tanzanian cement factory must remove ash particles of diameter $1.0\mu\text{m}$ from the flue gas. Each particle is assumed to acquire a charge of $100e$ by collision with ions close to the central wire, and must then be driven sideways across a 15cm gap onto an earthed collecting plate during the 2.0s residence time of the gas inside the precipitator. Take the viscosity of the flue gas as that of air, $\eta = 1.8 \times 10^{-5}\text{Pa}\cdot\text{s}$. Estimate the minimum applied electric field that will deliver every particle onto the plate within the residence time, and judge whether the chosen field is realistic in light of the dielectric strength of dry air, $3.0 \times 10^6\text{V/m}$.

Question 35

A laboratory Van de Graaff generator has a spherical metal dome of diameter 12cm and is rated to operate at a potential of 350kV above ground. Calculate the electric field at the surface of the dome at the rated potential, and determine whether the dome can hold this potential in ordinary dry air without breakdown. If the rated potential cannot be held, suggest two design modifications that would allow the dome to operate at 350kV safely. Take the dielectric strength of dry air as $3.0 \times 10^6\text{V/m}$.

Chapter 11

CAPACITANCE

INTRODUCTION

The old black-and-white television in the Miono physics prep room had been unplugged for two days. The lab technician, his hands full of soldering iron and replacement parts, glanced up from his workbench at the corner of the room. He was waiting for a new picture tube. The cabinet of the old set sat with its back panel removed, exposed to the morning light.

Kipanga, never one to leave equipment unexplored, stepped over to look inside. His curiosity rarely consulted his sense of caution. He reached his hand in past the chassis, intent on examining the picture tube. The technician, glancing up just a moment too late, opened his mouth to call out.

SNAP!

Kipanga yanked his hand back, sucking at the side of his thumb. A small red mark had risen on it. He stared at the back of the television, then at his hand, then at the technician, with all the affront of someone who has been personally insulted by an inanimate object.

Kipanga: *It bit me! The television bit me, and it is not even plugged in!*

Mr. Akilikubwa, who had been writing up notes at the doorway of the prep room, looked up with the slow, settled expression of a man who has seen this exact scene before, many times, and resigned himself to its repetition.

Mr. Akilikubwa (without rising from his notebook): *Disconnected is not the same as discharged, Kipanga. Inside an old television there is a part called a capacitor, and in fact there are many such parts, including one very large one connected to the picture tube. It is charged to thousands of volts when the set is in use. When the set is unplugged, that capacitor holds onto its charge for hours, even days, sometimes weeks. The technician knows this. That is why he never reaches inside before discharging it himself.*

Kipute (from across the lab, laughing): *Sir, this is the same boy who, three weeks ago, was told not to lick the terminals of a battery to test whether they were still good.*

Kipanga: *I learn by experience. The book only tells you the answer. The shock tells you the reason.*

Mr. Akilikubwa rose, walked over to the bench, and picked up a long screwdriver with a thick insulated rubber handle. He touched the metal tip of the screwdriver to the anode cap on top of the picture tube and then immediately to the chassis of the set. A small, brief snap followed; it was smaller than the one that had stung Kipanga's hand. He held the tip in place for a moment longer, withdrew it, and repeated the action a second time. The second snap was barely audible.

Mr. Akilikubwa: *Now it is safe to work on. The first shorting drains the bulk of the stored charge. The second is for anything that might have re-built itself from the residual fields. A technician who skips this step is the kind of technician one only meets once.*

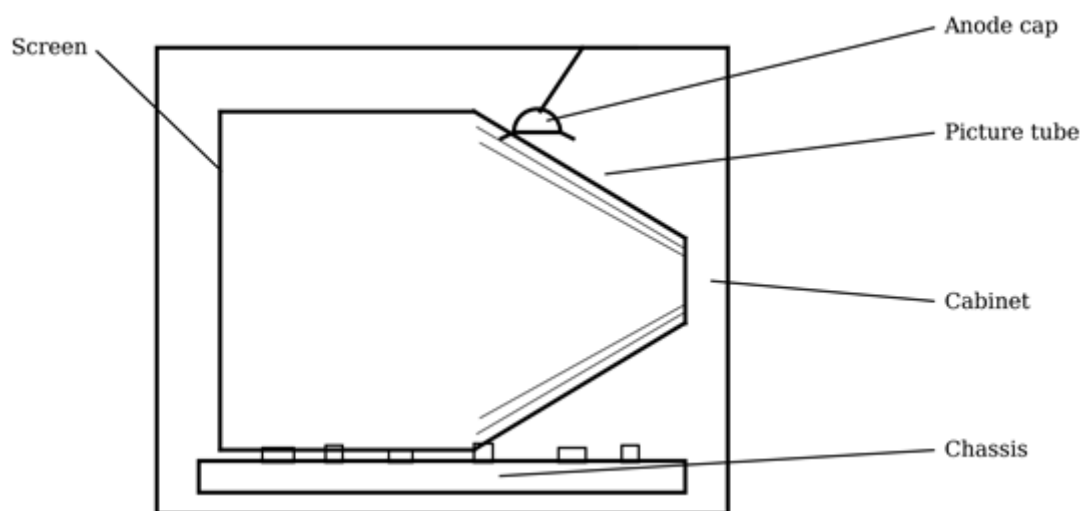


Figure: *Cut-away view of an old CRT television set. The picture tube, the chassis, and the anode cap (which holds the high voltage) are labelled. A capacitor inside the set keeps the anode cap charged for a long time after the set is unplugged.*

The capacitor is one of the three foundation components of any circuit that does something more interesting than light a single bulb. The other two are the resistor and the inductor; the resistor has been with us since O-level, and the inductor is the subject of Chapter 13. Capacitors are the components that store electric charge for later use. They make the sudden flash of a camera possible; they smooth out the trembling voltage from a power supply; they tune a radio to one station among many; they fire the high-voltage spark plug of a petrol engine; and, as Kipanga has just discovered, they hold their charge long after their source has been removed.

In Chapter 10 we calculated the electric field between two large parallel plates carrying opposite surface charges, and found it to have the simple value $E = \frac{\sigma}{\epsilon_0}$. In this chapter we use that result to derive a formula for how much charge a capacitor of given geometry holds for a given applied voltage. Then we extend the result to combinations of capacitors, to capacitors with insulating filler material between their plates, to the energy they store, to the force their charged plates exert on one another, and to the curve of charge versus time as a capacitor fills and empties through a resistor. We close with a set of applications that use capacitors in serious work: the camera flash, the defibrillator that revives a stopped heart, the soil-moisture sensor on a farm in Mbeya, the power-factor correction bank at a substation in Tanga, and several more.

One thing should already be clear from Kipanga's small adventure with the television: a capacitor remembers. It remembers the charge put on it long after the source has been removed. To work safely with capacitors, and to use them intelligently, we need to understand exactly how they remember, what determines how much they remember, and how they release what they remember when finally given the chance. The next eight sections of this chapter answer those questions, in the order they are best learned. Disconnected, as Mr. Akilikubwa reminded us, is not the same as discharged.

THE CAPACITOR

Open a portable radio, the controller of a desk fan, or the power-supply box of any old computer. Among the resistors, the tiny black integrated-circuit chips, and the loops of copper wire, you will see something tubular, sometimes blue, sometimes black, sometimes silver, with two wire legs sticking out of one end and the markings "10 μ F 25V" or "470 μ F 16V" printed on its side. That is an electrolytic capacitor. Pry open an old desk radio and you will find a small beige ceramic disc capacitor with a code such as "104" or "103" stamped on it. Look inside a digital voltmeter and you will see a flat rectangular box capacitor of the polyester or film type. Inside a camera, the same shape but larger, charged to several hundred volts, waits patiently to fire the flash.

These are all capacitors. They look different on the outside because of the materials and the geometries inside, but the underlying physics is the same in every case. **A capacitor is a device for storing electric charge on a pair of conductors held very close together but never allowed to touch.** The conductors are usually flat plates. The thin gap between them is filled with an *insulating material* called a **dielectric**, which we will discuss in detail later in the chapter. For now we will simply take the gap to be air or vacuum.

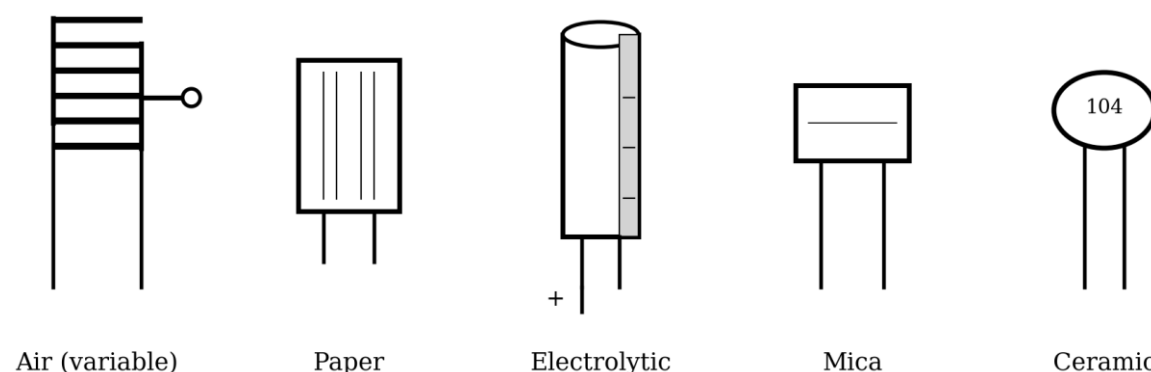


Figure: Five common capacitor types. Each is named after the material that fills the gap between its plates.

To get a feel for what 'storing charge' looks like in practice, look at the simple charging circuit in the figure. A battery of voltage V is connected through a switch S to a capacitor with plates labelled M and N . While the switch is open, both plates are electrically neutral. The moment the switch is closed, the battery starts pushing electrons from plate M through the external wire, around through the battery, and onto plate N , where they begin to accumulate. Plate N becomes more and more negative; plate M , which has just lost the same number of electrons, becomes equally positive.

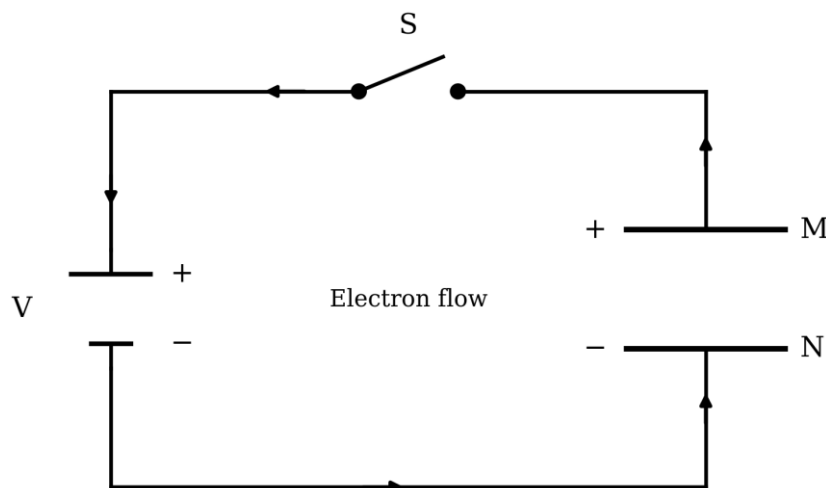


Figure: **Charging a capacitor.** Electrons leave plate M, flow through the external circuit, and accumulate on plate N. The electrons never cross the gap between the plates; the dielectric prevents that.

The electrons do not at any stage cross the gap between the plates. They cannot because the dielectric is an insulator. They only go around the outside, through the battery, depositing themselves on plate N from outside while peeling off plate M from outside. By the time the capacitor is fully charged, plate N carries a charge of magnitude Q with negative sign, and plate M carries the same magnitude Q with positive sign. The total charge on the device is $+Q + (-Q) = 0$. The capacitor as a whole is electrically neutral, even when it is "holding a charge of Q ". When we say the capacitor holds charge Q , we mean either plate alone carries that magnitude; the total is always, at every instant, zero. The plates have simply taken sides.

How much charge piles up depends on two things only: how hard the battery pushes (its voltage V) **and** how generously the capacitor is built to receive the charge. The first is set by the supply; the second is set by the capacitor. Experiment shows that for a given capacitor, the charge Q piled on either plate is strictly proportional to the applied voltage V : double V and Q doubles, halve V and Q halves. The constant of proportionality is called the **capacitance** of the capacitor, written **C**:

$$C = \frac{Q}{V}$$

This is the defining equation of capacitance. Read it carefully: it says that **capacitance** is the amount of charge stored per unit of voltage applied. A capacitor of large capacitance stores a lot of charge for each volt applied to it; a capacitor of small capacitance stores very little.

The SI unit of capacitance is the **farad (F)**, in honour of Michael Faraday, the English physicist who first studied capacitors and dielectrics in serious detail in the 1830s. By definition, a capacitor of 1F stores 1C of charge when 1V is applied across its plates:

$$1\text{F} = \frac{1\text{C}}{1\text{V}}$$

By any reasonable engineering standard, the farad is an absurdly large unit. A typical desk-radio capacitor lives in the picofarad range; six trillionths (10^{-12}) of a farad. The tubular capacitors in your power supply live in the microfarad range; millionths (10^{-6}) of a farad. Only one class of specialised device, the so-called supercapacitor or ultracapacitor, reaches whole farads, and it pays for the privilege with a low working voltage and a noticeably bulky physical size. The multiples you will routinely meet in this chapter, and in every examination paper from here onwards, are:

$$1\text{mF} = 10^{-3}\text{F}$$

$$1\mu\text{F} = 10^{-6}\text{F}$$

$$1\text{nF} = 10^{-9}\text{F}$$

$$1\text{pF} = 10^{-12}\text{F}$$

The microfarad and the picofarad do almost all the practical work. A 1F capacitor on a circuit board would be a curiosity, not a component.

A brief word now on the five common capacitor types and where each is at home in a circuit.

- 1) **Air capacitors** use *ordinary air* as their dielectric. They are still common in the variable-tuning sections of older AM radios; turning the tuning knob slides one set of plates partway out from between another set, changing the effective overlap area and therefore the capacitance.
- 2) **Paper capacitors** use *waxed or oil-impregnated paper* as the dielectric. They were once the standard in radio receivers, motor-starting circuits, and old electrical machinery; modern equipment has largely replaced them with film capacitors, but they remain common in equipment that comes back to the bench for repair.
- 3) **Electrolytic capacitors** achieve very large capacitance values in small packages by using a *very thin chemical layer* (the *electrolyte*) as the dielectric between two metallic foils. They appear wherever a circuit needs to store a large amount of charge: power-supply smoothing, motor-control circuits, audio amplifier coupling. They are polarised. Each electrolytic capacitor has a definite positive lead (often the longer leg, sometimes marked with stripes on the body) and a definite negative lead. Connecting an electrolytic capacitor the wrong way around results in a quiet hiss, a sharp burning smell, and a permanent change in the capacitor's circuit behaviour, in the sense that it stops being a capacitor at all.
- 4) **Mica capacitors** use thin *sheets of natural mica* as the dielectric between metal plates. Mica has a remarkably stable dielectric constant and a very high dielectric strength, so mica capacitors are prized for accuracy at high frequencies, particularly in radio oscillator circuits and precision measurement instruments.
- 5) **Ceramic capacitors** use a *thin disc of ceramic*, usually a *titanate or oxide compound*, as the dielectric. They are by far the cheapest and most numerous capacitors on a modern circuit board, performing routine bypass and decoupling duties wherever a little capacitance is needed. The "**104**" or "**103**" stamped on a ceramic capacitor reads as *10 followed by 4 zeros (or 3 zeros) in picofarads, giving 100nF or 10nF respectively.*

In circuit diagrams, capacitors are represented by a small standard set of symbols, shown in the following figure.

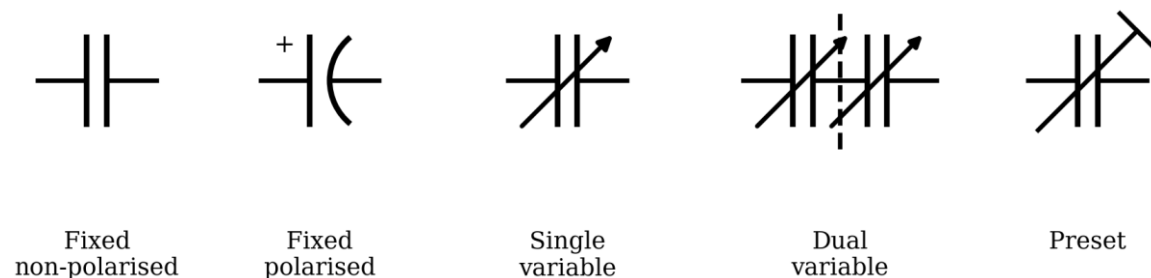


Figure: Standard circuit symbols for fixed (non-polarised, polarised) and variable (single, dual, preset) capacitors.

One small geometric matter is worth dispatching before we move on. The two plates of a capacitor **need not** be the same size. If we built a capacitor with one plate a small coin and the other plate a large sheet of metal, the small coin would simply carry a higher surface charge density (charge per unit area) than the large sheet, but the total charges on the two plates would remain exactly equal in magnitude and opposite in sign. This is not a coincidence of geometry; it is the consequence of charge conservation. Every electron that arrived on the negative plate had to leave the positive plate. The battery cannot manufacture electrons out of thin air; it can only push electrons around a closed loop. Whatever number of electrons arrives at one plate must have departed the other.

What follow are three short worked examples in which nothing exciting happens. They are here because every formula in physics needs an easy outing before it is asked to do real work.

BINDER Example 1

A capacitor of capacitance $5\mu\text{F}$ carries a charge of $60\mu\text{C}$ on either plate. Find the potential difference between the plates.

Solution

From the defining equation of capacitance:

$$C = \frac{Q}{V} \quad \Rightarrow \quad V = \frac{Q}{C}$$

Substituting the given values:

$$V = \frac{60 \times 10^{-6} \text{C}}{5 \times 10^{-6} \text{F}} = 12 \text{V}$$

The potential difference between the plates is 12V.

Making Sense of the Answer: Twelve volts is a perfectly ordinary working voltage for a microfarad-class capacitor; nothing here is unusual or extreme. The microfarad and the microcoulomb are companion units in the same way that the volt and the ampere are companions in everyday electrical work; both prefixes mean "one millionth", and when they appear together the result is a clean voltage of order ten volts. The $5\mu\text{F}$ – $60\mu\text{C}$ – 12V triple sits comfortably in the middle of the working range of any small microfarad capacitor.

Think Like a Physicist: Watch the prefixes cancel. Sixty divided by five is twelve, and the "micro" on top divides the "micro" on the bottom, leaving a pure voltage in volts. Whenever the charge and the capacitance are quoted in matching prefixes (microcoulombs with microfarads, nanocoulombs with nanofarads, picocoulombs with picofarads), the prefixes cancel and the answer is in straight volts. Noticing this saves several mental steps.

BINDER Example 2

A capacitor of capacitance 90pF is connected across a 12V battery. Find the number of electrons transferred from one plate to the other during charging.

Solution

The charge that accumulates on either plate when the capacitor is fully charged is:

$$Q = CV = (90 \times 10^{-12} \text{F})(12 \text{V}) = 1.08 \times 10^{-9} \text{C}$$

Each electron carries a charge of magnitude $e = 1.6 \times 10^{-19} \text{C}$. The number of electrons N transferred from one plate to the other is therefore the total charge divided by the charge per electron:

$$N = \frac{Q}{e} = \frac{1.08 \times 10^{-9} \text{C}}{1.6 \times 10^{-19} \text{C}} = 6.75 \times 10^9$$

The number is 6.75×10^9 electrons (about 6.75 billion electrons move from one plate, through the external circuit, and onto the other plate during the charging process).

Making Sense of the Answer: Six and three-quarter billion electrons sounds like an unimaginable number, but in fact it is a comparatively small number of electrons by physics standards. One coulomb of charge contains about 6.25×10^{18} electrons; what we transferred here is only about one part in a billion of a coulomb. The reason capacitors store such "small" numbers of electrons in absolute terms is that even a tiny imbalance of charge between two close plates produces a substantial voltage. Twelve volts across 90pF is the comfortable operating range of any pocket radio; the same twelve volts across a much larger pair of plates might store a microcoulomb, ten thousand times what we just calculated, yet still only a millionth of a coulomb. Electric forces are strong, and small charge differences make large voltages when the plates are close.

Think Like a Physicist: Any "how many electrons" question hides a two-step calculation: first find the total charge using the capacitance formula, then divide by the magnitude of the electronic charge. The two steps are independent of each other and are best written on two separate lines, since trying to crush them into one over-stuffed equation invites algebra errors. Notice also the scale of the answer: a few billion electrons, not a few million or a few trillion. Once a reader has done two or three of these calculations, the order of magnitude: billions, for a small picofarad-class capacitor at battery voltages becomes a useful sanity check on every future answer.

REAL Example 3

In the school prep room, Kipute opens a small drawer of spare components and pulls out a ceramic disc capacitor with the marking 104 stamped on its face. She wants to know how much charge the capacitor will hold when she connects it across a 5V supply. Find (a) the charge stored on either plate when the 5V supply is connected; (b) the number of electrons transferred from one plate to the other during charging.

Solution

(a) The marking 104 is read as 10 followed by 4 zeros, with the result in picofarads. The first two digits give the leading value, which is 10, and the third digit gives the number of trailing zeros, which is 4. The value of the capacitance is therefore:

$$C = 10 \times 10^4 \text{ pF} = 10^5 \text{ pF} = 10^5 \times 10^{-12} \text{ F} = 10^{-7} \text{ F}$$

From the defining equation of capacitance, the charge stored on either plate when 5V is applied is:

$$Q = CV = (10^{-7} \text{ F})(5\text{V}) = 5 \times 10^{-7} \text{ C} = 0.5 \mu\text{C}$$

(b) Each electron carries a charge of magnitude $e = 1.6 \times 10^{-19} \text{ C}$. The number of electrons transferred from one plate to the other during charging is therefore:

$$N = \frac{Q}{e} = \frac{5 \times 10^{-7} \text{ C}}{1.6 \times 10^{-19} \text{ C}} = 3.13 \times 10^{12} \text{ electrons}$$

Making Sense of the Answer: *Half a microcoulomb is more than the 1.08 nanocoulombs of the previous example by a factor of about five hundred. The two examples differ in capacitance by a factor of about a thousand (100nF versus 90pF) and in voltage by about a half (5V versus 12V), so the ratio of charges should be roughly $1000 \div 2 = 500$, which is what we got. The arithmetic checks against the physics, and the habit of estimating ratios like this catches half the algebra mistakes a student can make on a capacitor question.*

Think Like a Physicist: *A three-digit code stamped on a ceramic capacitor is read in the same way the colour bands on an old resistor are read: the first two digits give the leading value, and the third digit gives the power-of-ten multiplier. The result is always in picofarads, the smallest of the common engineering units for capacitance. Knowing the convention turns a component on the bench into a usable value in a few seconds, with no datasheet required.*

That is what the formula $C = \frac{Q}{V}$ looks like with its training wheels still on. From here on it works without them.

Across the corridor from the prep room, while the technician was still waiting for the replacement television tube, Mr. Akilikubwa had set up a small display on the bench: a 1000-microfarad electrolytic capacitor, a 9V battery, a low-current LED, and a 220Ω resistor.

Mr. Akilikubwa: *Kipanga, since you have already met a charged capacitor today, you may as well be the person who charges this one. Touch the leads of this capacitor briefly across the two terminals of the battery, just a second is enough, and then place those same leads across the LED with this resistor in series. Mind which way around the LED goes. It is polarised. The longer leg is the positive lead.*

Kipanga touches the capacitor leads to the battery for about a second, then transfers them, with the polarity carefully observed, to the LED-and-resistor combination. The LED lights up brightly, dims visibly over two or three seconds, then fades to dark.

Kipanga: *Sir, it glowed and then it faded.*

Mr. Akilikubwa: *It did. It glowed because the capacitor was driving current through the LED. It faded because the capacitor was running out of charge to drive that current with. By the end of the fade, the voltage across the capacitor had fallen to a value too small to light the LED.*

Kipute: *And the resistor controlled how slowly it faded?*

Mr. Akilikubwa: *It did. A larger resistor would have made the fade slower. A smaller resistor would have made it faster, and probably also burnt out the LED in the first rush of current before any fading could begin. The capacitor itself plays a role too. Replace this 1000-microfarad capacitor with a smaller 100-microfarad one, and the fade is ten times faster. Replace it with a 10 000-microfarad capacitor, and the LED holds its glow for tens of seconds.*

Kipanga: *So the capacitor is a kind of small battery?*

Mr. Akilikubwa: *It behaves like one during the discharge. But unlike a battery, a capacitor stores no chemical energy. It stores only the energy of the electric field between its plates. When the field drains away, so does the energy. A battery, once exhausted, can sometimes be revived by chemistry; a capacitor, once discharged, is simply empty until something charges it again. Later in this chapter we will calculate exactly*

how long the fade takes for any given combination of capacitor and resistor, using an exponential law and a single number called the time constant. For now you have the picture. A capacitor fills when connected to a battery and empties when connected across a load. The speed of either process depends on the resistance of the path.

THE PARALLEL-PLATE CAPACITOR

In the previous section we wrote down the defining equation $C = \frac{Q}{V}$, but that equation by itself tells us only what capacitance is. It does not tell us what determines the value of C for any particular device. To predict that value, we have to look inside the capacitor and account for the geometry of the plates and the material that fills the gap between them. We start with the simplest geometry of all: two flat parallel plates separated by a uniform gap.

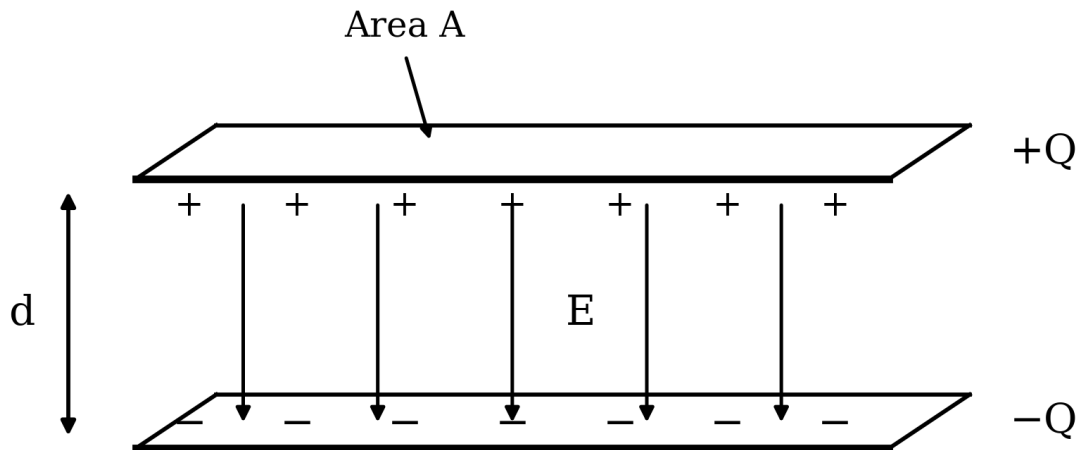


Figure: A **parallel-plate capacitor**. Each plate has area A and the gap between them is d . The top plate carries charge $+Q$, spread evenly over its area; the bottom plate carries an equal and opposite charge $-Q$. The field between the plates is uniform and points from the positive plate to the negative one.

Let each plate have area A , and let the uniform gap between them be d . Suppose the capacitor has been connected to a battery long enough to be fully charged. The top plate carries a positive charge Q distributed evenly over its surface, and the bottom plate carries the same magnitude of charge with the opposite sign, also distributed evenly. The surface charge density on each plate is therefore:

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

In the previous chapter we showed that the electric field between two parallel plates of opposite surface charge has the simple uniform value $E = \frac{\sigma}{\epsilon_0}$ in the region between the plates and is essentially zero outside, provided the plate dimensions are much larger than the gap d . This proviso is met by almost every real capacitor; the gap is always small compared with the plate dimensions. Substituting our expression for σ :

$$E = \frac{\sigma}{\epsilon_0} = \frac{Q}{A\epsilon_0}$$

This is the magnitude of the uniform electric field everywhere in the gap. The potential difference between the two plates is the work done per unit charge to move a test charge from one plate to the other across the gap. Because the field is uniform and the displacement is parallel to it, the potential difference is simply the field strength multiplied by the gap:

$$V = Ed = \frac{Qd}{A\epsilon_0}$$

Both Q and V depend on how much charge the battery has placed on the plates; the bigger the battery, the bigger both numbers. But their ratio is set entirely by the geometry of the capacitor, with no reference to the battery:

$$C = \frac{Q}{V} = \frac{A\epsilon_0}{d}$$

That is the central result for a parallel-plate capacitor. With three symbols on the right-hand side, it tells us three things at once.

1) Large plates make large capacitance

Doubling the plate area A doubles the capacitance, with no change to the voltage rating. This is the most direct way of increasing the capacitance of a parallel-plate device: make the plates bigger. The cost is bulk, since the device itself grows in the same proportion.

2) Close plates make large capacitance

Halving the gap d also doubles the capacitance, this time with no change in the plate area. In some ways this is an even more useful relation, because it works without enlarging the device. But it cannot be pushed without limit: the smaller the gap, the more easily the dielectric between the plates breaks down under the applied field, and the smaller the working voltage at which the capacitor can be operated safely.

3) The constant ϵ_0 is the property of empty space

No engineering choice changes its value. But if we replace the air or vacuum gap with an insulating material called a dielectric, the effective permittivity rises above ϵ_0 , and the capacitance rises in the same proportion. This is the third method of increasing capacitance, and the most powerful in practice. The whole industry of large-value capacitors is built on the development of dielectrics of high permittivity. We return to this idea in detail in the section on dielectrics later in the chapter.

For now, with the gap full of air or vacuum, the formula stands as $C = \frac{A\epsilon_0}{d}$, and ϵ_0 takes its SI value:

$$\epsilon_0 = 8.854 \times 10^{-12} \text{F/m}$$

There is one practical limit on how thin the gap d can be made. At very small separations the electric field $E = \frac{V}{d}$ in the gap becomes so large that the air between the plates ionises, and the capacitor discharges itself across the gap as a brief spark.

The field at which this happens is called the dielectric strength of the medium, and for dry air at ordinary pressure and temperature it has the value $E_{\text{breakdown}} \approx 3 \times 10^6 \text{ V/m}$. This is the same field that we met at the end of Chapter 10, when we examined the lightning bolt between a charged cloud base and the ground: lightning is the breakdown of the air column when the field between cloud and ground exceeds this threshold. Inside an air-gap capacitor the same physics sets the maximum voltage the device can withstand for a given gap: $V_{\text{max}} = E_{\text{breakdown}} \times d$. Push the applied voltage past V_{max} and the capacitor flashes over. We will return to dielectric strengths of other materials in a later section, when we examine what happens when an insulating slab is placed between the plates.

It is also worth pausing on the geometry of the field between the plates. Inside the central region (the part of the gap well away from the plate edges) the field is straight, parallel, and uniform, as we have already seen. The surfaces of constant potential in this region, called the equipotential surfaces, are flat planes parallel to the plates themselves, evenly spaced through the gap; the field lines cross every equipotential at right angles. Near the edges of the plates, however, the field is no longer perfectly uniform. *The lines bulge outward into the space just beyond the plates*, an effect called **edge fringing** or the **fringe field**. The equipotentials curve to follow them. For plates whose linear dimensions are large compared with the gap d , the fringing region is a small fraction of the total field volume, and the uniform-field approximation used in the derivation of $C = \frac{\epsilon_0 A}{d}$ is excellent. For very small plates with a comparatively large gap, fringing becomes a noticeable correction and more careful methods are required.

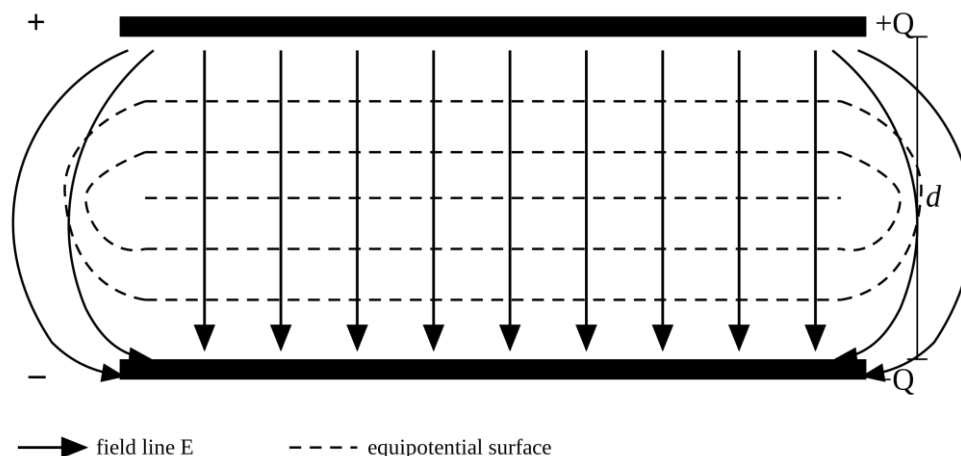


Figure: Equipotential surfaces and field lines for a parallel-plate capacitor. The field lines (solid arrows) are straight and evenly spaced in the central region between the plates and curve outward at the edges in the fringing region. The equipotentials (dashed) are flat planes in the central region and curve outward near the edges. Field lines always cross equipotentials at right angles.

Multi-plate capacitors

A real capacitor can multiply its effective plate area without growing in overall size by stacking many plates and connecting them alternately to the two leads. Suppose we have N plates stacked with a uniform gap d between each adjacent pair, and we wire the odd-numbered plates (1, 3, 5, and so on) to one lead and the even-numbered plates (2, 4, 6, and so on) to the other lead.

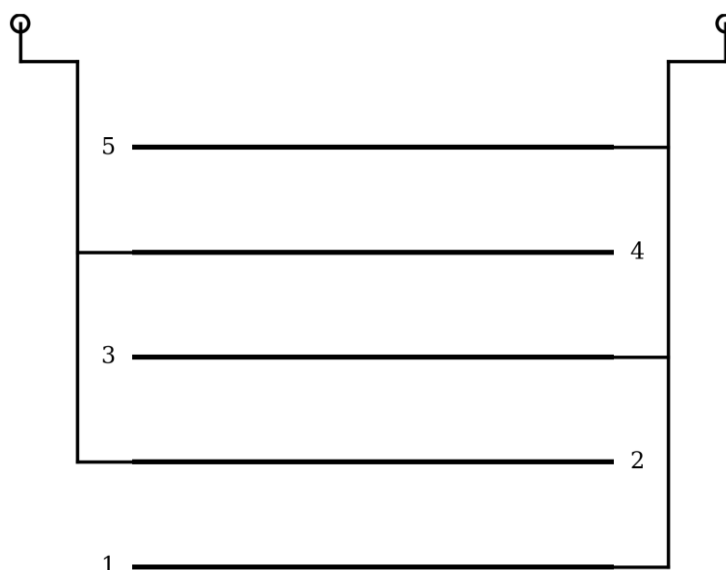


Figure: A multi-plate capacitor with five plates. Odd-numbered plates are wired to one terminal, even-numbered plates to the other. Each adjacent pair of plates forms one parallel-plate capacitor, and the four such pairs are all wired in parallel.

Each adjacent pair of plates now behaves as its own small parallel-plate capacitor with plate area A and gap d . The number of such pairs is one less than the number of plates, since between N plates there are $N - 1$ gaps. (For the five-plate stack drawn here, the four adjacent pairs are 1–2, 2–3, 3–4, and 4–5.) All of these pairs share the same two leads, and so they are wired in parallel. As we will see in the next section, capacitors wired in parallel have their capacitances added directly. The total capacitance of the multi-plate stack is therefore:

$$C = (N - 1) \frac{A\epsilon_0}{d}$$

Each additional plate adds one more pair, and so one more effective capacitor, to the parallel sum.

Most **air-variable capacitors** in old radios are built on exactly this principle. A stack of about 10 to 30 thin metal plates is split in two: half of them are fixed in the cabinet, and half are mounted on a rotating shaft. Turning the tuning knob slides the rotating set in or out of the fixed set, changing the effective overlap area A smoothly between almost zero (plates fully apart) and the full geometrical area (plates fully meshed). The capacitance then varies smoothly between a few picofarads and a few hundred picofarads, which is exactly the range needed to tune a radio across the AM broadcast band.

Five worked examples follow.

BINDER Example 4

A multi-plate mica capacitor is built from 7 thin plates, each of area 50cm^2 , stacked with a uniform gap of 0.2mm between consecutive plates. The gaps are filled with mica, whose dielectric constant is $\epsilon_r = 6$. Find the capacitance of the device.

Solution

With $N = 7$ plates in the stack, the number of adjacent pairs (and so the number of capacitor gaps in parallel) is one less:

$$N - 1 = 7 - 1 = 6$$

Each plate has area $A = 50\text{cm}^2 = 5 \times 10^{-3} \text{m}^2$. The gap is $d = 0.2\text{mm} = 2 \times 10^{-4} \text{m}$. With mica in the gap, the effective permittivity (ϵ) is $\epsilon_r\epsilon_0$, so the multi-plate formula becomes:

$$C = \frac{(N - 1) \epsilon_r \epsilon_0 A}{d}$$

Substituting:

$$C = \frac{6 \times 6 \times (8.854 \times 10^{-12} \text{F/m}) \times (5 \times 10^{-3} \text{m}^2)}{2 \times 10^{-4} \text{m}} = 7.97 \times 10^{-9} \text{F} = 7.97 \text{nF}$$

The capacitance of the seven-plate mica stack is 7.97nF .

Making Sense of the Answer: A few nanofarads is exactly the order of magnitude one finds for the high-stability mica capacitors used in radio-frequency tuned circuits, where dielectric losses must be tiny and the capacitance must not drift with temperature. The two design moves that pushed our number out of the bare-air picofarad range and into the nanofarad range were the multi-plate stacking (six gaps in parallel rather than one) and the use of a high- ϵ_r material in the gap (mica multiplies the bare-air value by six). Each of these moves contributes a factor of six in the present example, giving the device a capacitance thirty-six times larger than a single-gap air capacitor of the same plate area would have at the same plate separation.

Think Like a Physicist: Multi-plate formulae and dielectric formulae multiply together cleanly, because their underlying physical effects are independent: stacking plates adds capacitor gaps in parallel, while a dielectric in each gap raises that gap's per-pair capacitance. The composite formula $C = \frac{(N - 1) \epsilon_r \epsilon_0 A}{d}$ is best read as the single-pair vacuum capacitance $\frac{\epsilon_0 A}{d}$, multiplied by $(N - 1)$ for parallel-gap stacking, multiplied by ϵ_r for the dielectric fill. Identifying these three factors separately on the worksheet makes both the arithmetic and the physical meaning easier to track.

BINDER Example 5

A parallel-plate capacitor consists of two square metal plates of side 10cm separated by an air gap of 1mm . (a) Calculate the capacitance of the device. (b) Find the maximum voltage that can be applied across the plates before air breakdown, given a dielectric strength of $3 \times 10^6 \text{V/m}$ for dry air. (c) Find the maximum charge the device can hold at this voltage.

Solution

(a) Convert each given length to metres, since the SI unit of ϵ_0 is built on the metre. The plates are squares of side 10cm , so each has area:

$$A = (0.10\text{m})^2 = 1 \times 10^{-2}\text{m}^2$$

The gap, converted to metres, is:

$$d = 1\text{mm} = 1 \times 10^{-3}\text{m}$$

With $\epsilon_0 = 8.854 \times 10^{-12}\text{F/m}$, the capacitance is:

$$C = \frac{A\epsilon_0}{d} = \frac{(1 \times 10^{-2}\text{m}^2)(8.854 \times 10^{-12}\text{F/m})}{1 \times 10^{-3}\text{m}} = 8.854 \times 10^{-11}\text{F} \approx 88.5\text{pF}$$

The capacitance of the device is about 88.5pF.

(b) The maximum voltage that can be applied across the gap before the air ionises is the dielectric strength multiplied by the gap:

$$V_{\text{max}} = E_{\text{breakdown}} \times d = (3 \times 10^6 \text{ V/m})(1 \times 10^{-3} \text{ m}) = 3 \times 10^3 \text{ V} = 3000\text{V}$$

The capacitor can safely sustain potential differences up to about 3kV before the air between its plates ionises and the device discharges across the gap in a spark.

(c) The maximum charge the capacitor can hold at this maximum voltage follows from the defining equation of capacitance:

$$Q_{\text{max}} = C V_{\text{max}} = (88.5 \times 10^{-12} \text{ F})(3000 \text{ V}) = 2.66 \times 10^{-7} \text{ C} \approx 0.27\mu\text{C}$$

The maximum charge the device can store before discharging through the air is about 0.27 μC .

Making Sense of the Answer: About ninety picofarads for 10cm by 10cm plates a millimetre apart is exactly the order of magnitude one finds in the variable tuning capacitors of old AM radios. Small, tens to hundreds of picofarads, is the natural scale of a hand-built parallel-plate device. To climb into the microfarad range one must take one of three paths: enlarge the plates beyond what a hobbyist can build, shrink the gap to a few micrometres at the price of low working voltage, or fill the gap with a dielectric of high permittivity. The first path leads to bulky equipment, the second to fragile equipment, and the third to the electrolytic capacitor in every power supply on this planet.

Think Like a Physicist: The constant ϵ_0 has the awkward value $8.854 \times 10^{-12} \text{ F/m}$, and that minus-twelve hides in every parallel-plate calculation, waiting to be dropped. A useful habit is to write ϵ_0 in scientific notation at the very top of the working, before substituting anything else, and then carry the powers of ten through the algebra by adding and subtracting their exponents in a clearly visible column. Forgetting a single 10^{-12} gives an answer that is wrong by a factor of a trillion, and on examination scripts this is the single commonest mistake on parallel-plate questions.

REAL Example 6

In the school prep room, **Kipute** opens a small capacitance meter and clips its leads across an air-gap parallel-plate capacitor whose plates are exposed. She breathes gently on the gap between the plates and watches the meter reading: it shifts only by a few per cent. She then slides a thin sheet of dry paper between the plates (without touching either plate) and watches the reading again: it now changes by a factor of about three. Explain qualitatively why the two interventions produce such different effects.

Solution

Both interventions are doing the same physical thing: altering the dielectric properties of the material that fills the gap. But they do it on very different scales.

Air at ordinary temperature and humidity is almost a perfect vacuum from the point of view of the electric field. Its dielectric constant is $\epsilon_r \approx 1.0006$ in dry air, and rises only slightly with humidity. When Kipute breathes on the gap, she increases the water-vapour content of the air slightly, which raises ϵ_r by a fraction of a percent. The capacitor sees a tiny corresponding rise in its capacitance, and the meter responds by a few per cent. Even a wet breath barely lifts ϵ_r above one.

Paper is a very different beast with greater value of ϵ_r . Even dry paper has $\epsilon_r \approx 3$, a value three times the dielectric constant of air. Slipping the paper between the plates replaces most of the air in the gap with a much more easily polarised material; the resulting capacitance is therefore about three times the empty-gap value. The meter responds by a factor of three.

The qualitative lesson is that what counts is not what material is touching the plates, but what material fills the gap between the plates. That is where the field stores its energy and the dielectric does its work. Humid air remains, electrically, almost vacuum. Paper is fundamentally different.

Making Sense of the Answer: A factor of three in capacitance from a single sheet of paper is the simplest demonstration that the medium between the plates matters as much as the geometry. Breathing on the gap, by contrast, is a stress-test of how close to vacuum ordinary air really is.

Think Like a Physicist: Whenever a question contrasts two experimental interventions and asks you to explain why one matters more than the other, the cleanest answer names the underlying parameter that is being varied (here, the dielectric constant ϵ_r of the material in the gap) and the magnitude of the variation in each case (fractional per-cent for humidity in air, factor of three for paper). The qualitative answer is usually a one-line ratio between the two relevant ϵ_r values, with the rest of the paragraph spent justifying the ratio.

REAL Example 7

The variable tuning capacitor of an old AM radio receiver is built from a stack of 21 thin semicircular brass plates of radius 2.5cm, separated by air gaps of 0.5mm. The odd-numbered plates are fixed to the cabinet, and the even-numbered plates are mounted on a tuning shaft so that they rotate in and out of the fixed set. Calculate the maximum capacitance of the device, achieved when the moving plates are fully meshed with the fixed ones.

Solution

With 21 plates in the stack, the number of adjacent pairs is one less:

$$N - 1 = 21 - 1 = 20$$

Each plate is a semicircle of radius $r = 2.5\text{cm} = 2.5 \times 10^{-2}\text{m}$. The area of one semicircle is:

$$A = \frac{1}{2}\pi r^2 = \frac{1}{2}\pi(2.5 \times 10^{-2}\text{m})^2 = 9.817 \times 10^{-4}\text{m}^2$$

The gap, converted to metres, is $d = 0.5\text{mm} = 5 \times 10^{-4}\text{m}$. The total capacitance of the meshed stack is:

$$C = (N - 1) \frac{A\epsilon_0}{d}$$

$$C = 20 \times \frac{(9.817 \times 10^{-4}\text{m}^2)(8.854 \times 10^{-12}\text{F/m})}{5 \times 10^{-4}\text{m}} = 3.48 \times 10^{-10}\text{F} = 348\text{pF}$$

The maximum capacitance of the tuner is about 348pF.

Making Sense of the Answer: The medium-wave AM broadcast band runs from about 530kHz to 1600kHz. An AM radio tunes across this band by varying the capacitance of a parallel LC circuit; the inductor is a fixed coil of two or three hundred microhenrys, and the variable capacitor has to swing from about 20pF (when the moving plates are fully open) to 350 or 400pF (when fully meshed). Our calculated maximum of 348pF lands exactly where it should. The reader who builds an AM radio at home is therefore building a capacitor of this rough size and shape, with their own hands, and the result of their effort is the ability to select a station out of a few dozen contenders by rotating a single knob.

Think Like a Physicist: A multi-plate capacitor's capacitance is the per-pair capacitance multiplied by the number of pairs, not by the number of plates. With 21 plates the number of pairs is 20, with the simple rule that the number of gaps is one less than the number of plates. Forgetting this off-by-one is the single commonest mistake on multi-plate questions. A student who recognises a multi-plate problem on sight should write the quantity $(N - 1)$ on the working sheet before doing anything else, just as a student who recognises a circular wire problem on sight should write down $\pi d^2/4$ before doing anything else.

HOT Example 8

A parallel-plate capacitor consists of two circular metal plates of radius 5cm separated by an air gap of 2mm. The capacitor is connected to a 50V supply and allowed to charge fully. The supply is then disconnected. After the supply has been removed, the gap between the plates is reduced to 0.5mm by a mechanical adjustment. Find (a) the original capacitance and the original charge on each plate; (b) the new capacitance after the gap has been reduced; (c) the new voltage between the plates.

Solution

Before any algebra, write down the central physical fact about this problem: the supply has been disconnected before the geometry change. With the supply gone, there is no path for charge to flow on or off the plates, and so the total charge Q on each plate is held fixed throughout the gap change. The capacitance C changes because the geometry changes; the voltage V then adjusts to keep the relation $Q = CV$ consistent.

(a) The plate area is:

$$A = \pi r^2 = \pi(5 \times 10^{-2} \text{m})^2 = 7.854 \times 10^{-3} \text{m}^2$$

With the original gap $d_{\text{old}} = 2 \text{mm} = 2 \times 10^{-3} \text{m}$, the original capacitance is:

$$C_{\text{old}} = \frac{A\epsilon_0}{d_{\text{old}}} = \frac{(7.854 \times 10^{-3} \text{m}^2)(8.854 \times 10^{-12} \text{F/m})}{2 \times 10^{-3} \text{m}} = 3.48 \times 10^{-11} \text{F} = 34.8 \text{pF}$$

The original charge on each plate is:

$$Q = C_{\text{old}} V_{\text{old}} = (34.8 \times 10^{-12} \text{F})(50 \text{V}) = 1.74 \times 10^{-9} \text{C} = 1.74 \text{nC}$$

(b) The new gap is $d_{\text{new}} = 0.5 \text{mm} = 5 \times 10^{-4} \text{m}$, which is one quarter of the original gap. Since the capacitance is inversely proportional to the gap (with everything else held fixed), reducing the gap by a factor of four multiplies the capacitance by four:

$$C_{\text{new}} = 4C_{\text{old}} = 4 \times 34.8 \text{pF} = 139.2 \text{pF}$$

(c) The charge Q is held fixed (the supply was disconnected before the gap was changed). The new voltage is therefore:

$$V_{\text{new}} = \frac{Q}{C_{\text{new}}} = \frac{1.74 \times 10^{-9} \text{C}}{139.2 \times 10^{-12} \text{F}} = 12.5 \text{V}$$

After the gap is reduced to 0.5mm, the capacitance rises to about 139pF and the voltage between the plates falls to about 12.5V, while the charge of 1.74nC remains undisturbed on each plate.

Making Sense of the Answer: Reducing the gap by a factor of four has increased the capacitance by a factor of four (because C is inversely proportional to d), and with the charge held fixed by the disconnected supply, the voltage must drop by a factor of four to keep the relation $Q = CV$ satisfied. Fifty volts divided by four is twelve and a half volts, which is what we obtained. The total charge of 1.74nC sat undisturbed on the plates throughout the gap change; the geometry change only altered the ease with which that charge could be held at a given voltage. The problem is, at heart, the relation $Q = CV$ restated three times in three different roles.

Think Like a Physicist: Every "what happens after the supply is disconnected" question is asking which quantities are held fixed by the disconnection (always: the charge on each plate) and which quantities are free to change (always: the voltage, if the capacitance changes; and the capacitance, if the geometry changes). The first move on any such question is to write the words " **Q is held fixed**" at the top of the working before any algebra begins. The opposite pattern, in which the supply stays connected throughout the change, holds the voltage fixed instead and lets the charge change. Knowing which of the three quantities (Q , V , C) is the constant of the problem decides the entire solution; getting it wrong sends every subsequent line of working in the wrong direction.

We now know how to compute the capacitance of a single parallel-plate device of any geometry, and the capacitance of a stack of such devices acting as one. The next section asks the inevitable follow-up question: *what happens when several finished capacitors are wired together in a circuit?*

COMBINATIONS OF CAPACITORS

In a perfect engineering world, every capacitor on the market would be available in exactly the value a designer wants. In the real world, capacitors come in standard sizes (1μF, 2.2μF, 4.7μF, 10μF, and so on), and the designer who needs, say, 4.5μF or 13.6μF must combine two or more standard parts to obtain the value required.

There are two ways to combine capacitors in a circuit. They can be wired in parallel, in which case both plates of each capacitor are connected across the same pair of nodes; or they can be wired in series, in which case the capacitors form a single chain between two nodes. The two arrangements produce very different total capacitances, and the way they share charge and voltage is different again. We deal with each in turn, and then with the related question of what happens when two capacitors that have been charged separately are afterwards connected together, the so-called charge-sharing problem.

Capacitors in parallel

When two or more capacitors are wired in parallel, both plates of each capacitor are connected across the same pair of nodes in the circuit. The same pair of wires links every capacitor's top plate, and another pair links every capacitor's bottom plate. Every capacitor in the combination therefore experiences the same potential difference V across its terminals: the voltage of the supply.

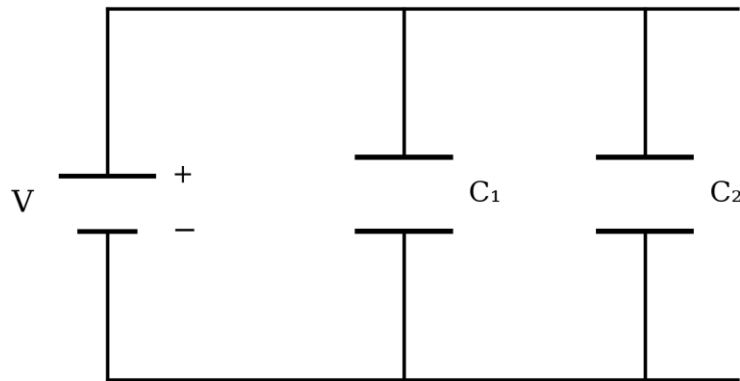


Figure: Two capacitors in parallel across a battery. Each capacitor sees the full battery voltage V across its plates, but each stores its own charge in proportion to its own capacitance.

The charge each capacitor accumulates is proportional to its own capacitance:

$$Q_1 = C_1 V, \quad Q_2 = C_2 V, \quad Q_3 = C_3 V, \quad \dots$$

The total charge stored across the combination is the sum of the charges on each individual capacitor:

$$Q_{\text{total}} = Q_1 + Q_2 + Q_3 + \dots = (C_1 + C_2 + C_3 + \dots)V$$

Defining the **equivalent capacitance** C_{eq} as the single capacitance that, when connected to the same supply, would store the same total charge:

$$Q_{\text{total}} = C_{\text{eq}} V$$

Comparing the two expressions for Q_{total} :

$$C_{\text{eq}} V = (C_1 + C_2 + C_3 + \dots)V$$

$$C_{\text{eq}} = C_1 + C_2 + C_3 + \dots$$

In parallel, capacitances add directly. This is the opposite of resistors, where parallel resistances combine reciprocally. The asymmetry is not a coincidence: *capacitors store charge while resistors pass it*. More capacitors in parallel provide more charge storage, in the same way that more buckets standing side by side hold more water; more resistors in parallel provide less resistance, in the same way that more lanes on a road create less obstruction to traffic flow.

Capacitors in series

When two or more capacitors are wired in series, they form a single chain between two nodes. The right plate of one capacitor is connected to the left plate of the next, and so on along the chain, with no other path for charge to enter or leave the middle wires.

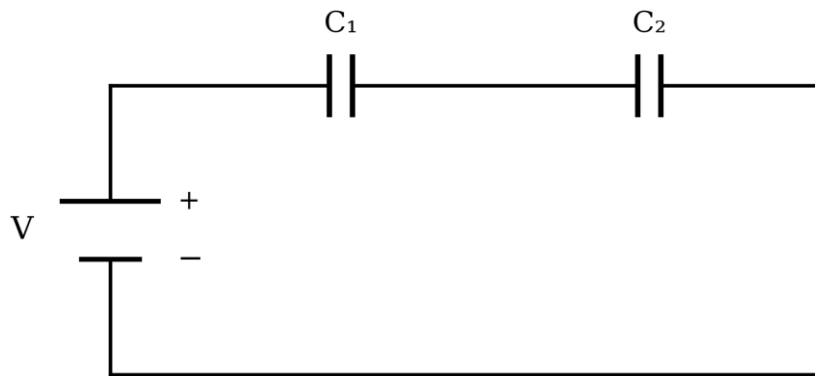


Figure: Two capacitors in series across a battery. The middle wire between the two capacitors is isolated from the rest of the circuit, and the charge on every plate in the chain has the same magnitude.

This isolation of the middle wires is the key to understanding the series combination. When the supply pushes charge $+Q$ onto the leftmost plate of the chain, the right plate of that same capacitor must carry $-Q$. The bottom-plate charge cannot come from the supply, because the supply is not connected to the middle wires. Instead, electrons from the next capacitor's left plate are drawn through the middle wire to settle on the first capacitor's right plate. The next capacitor's left plate, having lost those electrons, becomes $+Q$. The same argument repeats along the entire chain. Every capacitor in the series chain therefore carries the same magnitude of charge Q on its plates. This is the central fact about series-connected capacitors.

The voltage across each individual capacitor in the chain is set by its own capacitance:

$$V_1 = \frac{Q}{C_1}, \quad V_2 = \frac{Q}{C_2}, \quad V_3 = \frac{Q}{C_3}, \quad \dots$$

The total voltage across the chain is the sum of these individual voltages:

$$V = V_1 + V_2 + V_3 + \dots = Q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \dots\right)$$

Defining the **equivalent capacitance** C_{eq} as the single capacitor that would carry the same charge Q at the same total voltage V :

$$V = \frac{Q}{C_{\text{eq}}}$$

Comparing the two expressions for V :

$$\begin{aligned} \frac{Q}{C_{\text{eq}}} &= Q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \dots\right) \\ \frac{1}{C_{\text{eq}}} &= \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \dots \end{aligned}$$

In series, the reciprocals of the capacitances add, never the capacitances themselves.

For the common special case of just two capacitors in series, the reciprocal-sum formula rearranges into a more convenient form. Starting from:

$$\frac{1}{C_{\text{eq}}} = \frac{1}{C_1} + \frac{1}{C_2} = \frac{C_1 + C_2}{C_1 C_2}$$

Inverting both sides:

$$C_{\text{eq}} = \frac{C_1 C_2}{C_1 + C_2}$$

This is the "product over sum" rule, valid for two capacitors in series and **only** two.

It is worth recording the closed form for three capacitors in series as well. Starting from:

$$\frac{1}{C_{\text{eq}}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} = \frac{C_2 C_3 + C_1 C_3 + C_1 C_2}{C_1 C_2 C_3}$$

Inverting both sides:

$$C_{eq} = \frac{C_1 C_2 C_3}{C_1 C_2 + C_1 C_3 + C_2 C_3}$$

For four or more capacitors the closed form becomes unwieldy and the reciprocal sum must be evaluated directly, or the capacitors combined two at a time.

Before leaving the series formula, one observation is worth pinning down.

In a series chain, every capacitor carries the same charge Q (charge conservation through the chain), so the voltage across each capacitor is $V_i = \frac{Q}{C_i}$. The capacitor with the smallest C therefore carries the largest V . *Series capacitors divide the supply voltage in inverse proportion to their capacitances.*

By the same logic, in a parallel combination every capacitor sees the same V across its plates, so the charge on each is $Q_i = C_i V$. The capacitor with the largest C therefore carries the largest Q . *Parallel capacitors divide the supplied charge in direct proportion to their capacitances.*

These two observations are the engine of every exam-style potential-divider problem and are worth remembering in this compact form: ***in series, smallest C takes the largest V ; in parallel, largest C takes the largest Q .***

Charge sharing

A peculiar but instructive scenario arises when two capacitors are charged separately and then connected together. Suppose a capacitor C_1 has been charged to a voltage V_1 by one source, and a second capacitor C_2 has been charged separately to a different voltage V_2 by another source (or by the same source set to a different output). Both sources are then disconnected, and the two capacitors are wired together so that their positive plates are joined by one wire and their negative plates by another. The two capacitors are now in parallel, but each is initially at a different voltage.

Charge cannot remain in this disequilibrium. Electrons flow from the higher-voltage capacitor, through the connecting wire, and onto the lower-voltage one until both come to a common voltage. *What is that common voltage?*

The key idea is conservation of charge. Before the connection, the total charge held on the (originally) positive plates is:

$$Q_{total} = Q_1 + Q_2 = C_1 V_1 + C_2 V_2$$

After the connection, charge has redistributed between the two capacitors, but the total charge on the connected side is unchanged: nothing has been added to the system, and nothing has left it. The two capacitors are now in parallel with a common voltage V_{final} :

$$Q_{total} = (C_1 + C_2) V_{final}$$

Solving for the final voltage:

$$C_1 V_1 + C_2 V_2 = (C_1 + C_2) V_{final}$$

$$V_{final} = \frac{C_1 V_1 + C_2 V_2}{C_1 + C_2}$$

This is a weighted average of V_1 and V_2 , with the weights being the two capacitances. The result always lies between the two initial voltages, and is closer to whichever capacitor is the larger of the two: the bigger capacitor dominates the final voltage, in the same way that a large reservoir dominates the level of a small tank joined to it by a pipe.

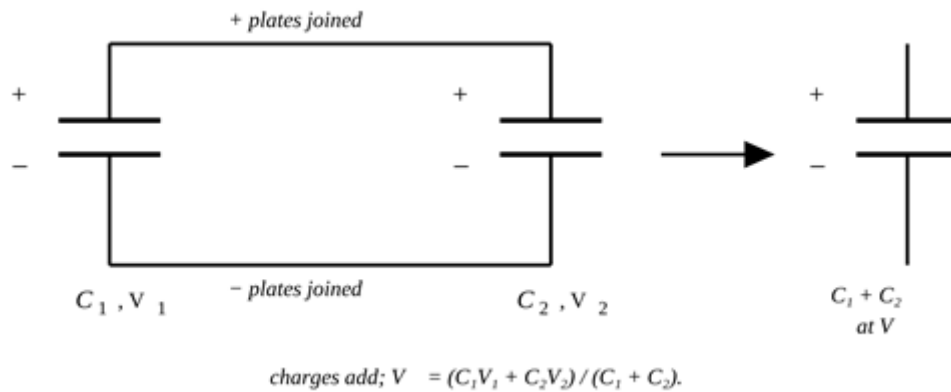


Figure: Charge sharing. Like plates joined: the positive plate of each capacitor is wired to the positive plate of the other, so charges of the same sign add.

A warning: The stated formula assumes the two capacitors are connected positive plate to positive plate, and negative to negative. If they are instead connected with opposite polarity (positive to negative), the charges on the joined plates partly cancel, and the formula takes the modified form:

$$V_{\text{final}} = \frac{|C_1V_1 - C_2V_2|}{C_1 + C_2}$$

The polarity of the connection therefore matters.

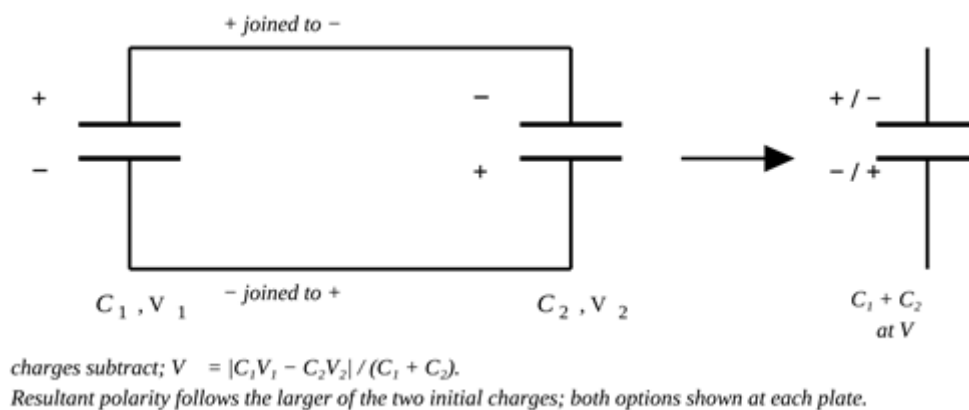


Figure: Charge sharing. Unlike plates joined: the positive plate of one capacitor is wired to the negative plate of the other, so charges of opposite sign partially cancel.

Five worked examples now follow.

BINDER Example 9

A $2\mu\text{F}$ capacitor and a $3\mu\text{F}$ capacitor are wired in parallel across a 10V battery. Find (a) the equivalent capacitance of the combination; (b) the charge stored on each capacitor; (c) the total charge supplied by the battery.

Solution

(a) Parallel capacitances add directly:

$$C_{\text{eq}} = C_1 + C_2 = 2\mu\text{F} + 3\mu\text{F} = 5\mu\text{F}$$

(b) In a parallel combination, every capacitor sees the full battery voltage $V = 10\text{V}$. The charge on each is:

$$Q_1 = C_1V = (2\mu\text{F})(10\text{V}) = 20\mu\text{C}$$

$$Q_2 = C_2V = (3\mu\text{F})(10\text{V}) = 30\mu\text{C}$$

(c) The total charge supplied by the battery is the sum of the charges stored on the two capacitors:

$$Q_{\text{total}} = Q_1 + Q_2 = 20\mu\text{C} + 30\mu\text{C} = 50\mu\text{C}$$

Alternatively, applying the parallel formula directly: $Q_{\text{total}} = C_{\text{eq}}V = (5\mu\text{F})(10\text{V}) = 50\mu\text{C}$, in agreement.

Making Sense of the Answer: The arithmetic is undramatic, but two features are worth noting. First, every capacitor in the parallel combination carries the same voltage, in the same way that every lamp wired in parallel across a supply glows at the supply voltage. Second, the larger capacitor stores more charge, in proportion to its capacitance: the $3\mu\text{F}$ capacitor stores $30\mu\text{C}$, the $2\mu\text{F}$ stores $20\mu\text{C}$, and the ratio of charges is the ratio of capacitances.

Think Like a Physicist: A good cross-check on any parallel-capacitor question is to compute the total charge in two ways: once as the sum of the individual charges (which uses every individual capacitance), and once as (equivalent capacitance)(voltage) (which uses only the equivalent). The two answers must agree. If they do not, an arithmetic mistake is somewhere in the working, and the cross-check has caught it before any mark has been lost. This double-bookkeeping habit takes ten extra seconds and saves marks that would otherwise leak out of the script unnoticed.

BINDER Example 10

A $4\mu\text{F}$ capacitor and a $6\mu\text{F}$ capacitor are wired in series across a 10V battery. Find (a) the equivalent capacitance of the combination; (b) the charge stored on each capacitor; (c) the voltage across each capacitor.

Solution

(a) Using the product-over-sum formula for two capacitors in series:

$$C_{\text{eq}} = \frac{C_1 C_2}{C_1 + C_2} = \frac{(4\mu\text{F})(6\mu\text{F})}{4\mu\text{F} + 6\mu\text{F}} = 2.4\mu\text{F}$$

(b) In a series combination, every capacitor in the chain carries the same charge, equal to the charge that the equivalent capacitor would carry at the supply voltage:

$$Q = C_{\text{eq}}V = (2.4\mu\text{F})(10\text{V}) = 24\mu\text{C}$$

Both capacitors carry the same charge $Q = 24\mu\text{C}$ on their plates.

(c) The voltage across each capacitor is set by its own capacitance:

$$V_1 = \frac{Q}{C_1} = \frac{24\mu\text{C}}{4\mu\text{F}} = 6\text{V}$$

$$V_2 = \frac{Q}{C_2} = \frac{24\mu\text{C}}{6\mu\text{F}} = 4\text{V}$$

As a cross-check, $V_1 + V_2 = 6\text{V} + 4\text{V} = 10\text{V}$, which is the supply voltage. The chain has divided the supply voltage between its two members in inverse proportion to their capacitances.

Making Sense of the Answer: Three lessons are buried in this short calculation. First, the equivalent capacitance of a series combination is always smaller than the smallest individual capacitance ($2.4\mu\text{F}$ here is smaller than $4\mu\text{F}$, and very much smaller than $6\mu\text{F}$). Series wiring weakens a capacitor; parallel wiring strengthens it. Second, the voltage is shared between the two capacitors in inverse proportion to their capacitances: the smaller capacitor holds more of the voltage (6V across $4\mu\text{F}$, 4V across $6\mu\text{F}$), because for the same charge, a smaller capacitance needs a larger voltage. Third, the two voltages must add to the supply voltage, which is the most reliable cross-check available on any examination problem.

Think Like a Physicist: The two-capacitor product-over-sum rule ($C_1 C_2 / (C_1 + C_2)$) is fast and reliable. Use it whenever exactly two capacitors are in series. For three or more, do not look for a three-capacitor shortcut; there is no clean one. Either compute the reciprocal sum in full or combine the capacitors two at a time, applying the two-capacitor formula iteratively. Trying to invent a three-capacitor shortcut on the fly during an examination is a guaranteed source of arithmetic mistakes.

REAL Example 11

In the circuit on a hobbyist's bench, three capacitors are wired together as follows: a $4\mu\text{F}$ capacitor C_1 sits in parallel with an $8\mu\text{F}$ capacitor C_2 , and this parallel pair is wired in series with a $6\mu\text{F}$ capacitor C_3 . The whole arrangement is then connected across a 12V supply. Find (a) the equivalent capacitance of the entire circuit; (b) the charge on each capacitor; (c) the voltage across each capacitor.

Solution

(a) First combine C_1 and C_2 , which are in parallel:

$$C_{12} = C_1 + C_2 = 4\mu\text{F} + 8\mu\text{F} = 12\mu\text{F}$$

Now combine C_{12} in series with C_3 using the product-over-sum rule:

$$C_{\text{eq}} = \frac{C_{12}C_3}{C_{12} + C_3} = \frac{(12\mu\text{F})(6\mu\text{F})}{12\mu\text{F} + 6\mu\text{F}} = 4\mu\text{F}$$

(b) In the series part of the circuit, the same charge passes through C_{12} and C_3 . This charge is equal to the charge on the equivalent capacitor at the supply voltage:

$$Q_{\text{series}} = C_{\text{eq}}V = (4\mu\text{F})(12\text{V}) = 48\mu\text{C}$$

The charge on C_3 is therefore $Q_3 = 48\mu\text{C}$. The same magnitude of charge sits across the parallel pair, but it splits between C_1 and C_2 in proportion to their individual capacitances. The voltage across the parallel pair is found first:

$$V_{12} = \frac{Q_{\text{series}}}{C_{12}} = \frac{48\mu\text{C}}{12\mu\text{F}} = 4\text{V}$$

Now the charges on the two parallel capacitors:

$$Q_1 = C_1V_{12} = (4\mu\text{F})(4\text{V}) = 16\mu\text{C}$$

$$Q_2 = C_2V_{12} = (8\mu\text{F})(4\text{V}) = 32\mu\text{C}$$

Cross-check: $Q_1 + Q_2 = 16\mu\text{C} + 32\mu\text{C} = 48\mu\text{C}$, which equals the series charge as it must.

(c) The voltage across C_3 is:

$$V_3 = \frac{Q_{\text{series}}}{C_3} = \frac{48\mu\text{C}}{6\mu\text{F}} = 8\text{V}$$

Both C_1 and C_2 carry the same voltage as the parallel pair, $V_{12} = 4\text{V}$.

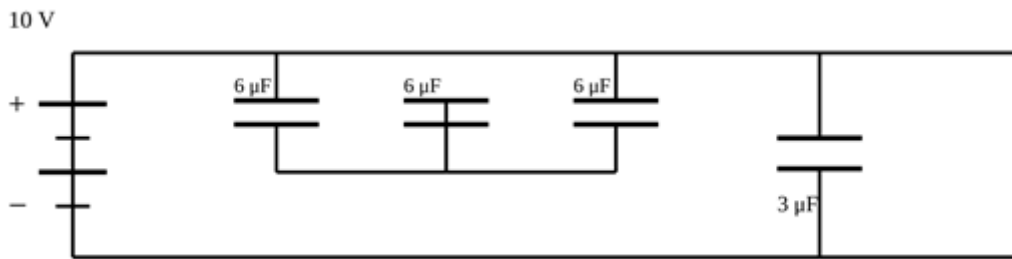
Cross-check on the series voltage rule: $V_{12} + V_3 = 4\text{V} + 8\text{V} = 12\text{V}$, the supply voltage.

Making Sense of the Answer: Three quick lessons appear in this answer. First, the equivalent capacitance of the whole circuit ($4\mu\text{F}$) is exactly equal to the individual capacitance of C_1 . This is a coincidence of the chosen numbers; do not read any general meaning into it. Second, the smaller member of the series part (the $6\mu\text{F}$ capacitor C_3) takes the larger share of the supply voltage (8V out of 12V), in exact accordance with the rule that voltage in series divides in inverse proportion to capacitance. Third, the parallel pair shares its $48\mu\text{C}$ in the ratio of its two capacitances, with the larger $8\mu\text{F}$ capacitor taking the larger $32\mu\text{C}$ share. Every cross-check in this problem passed; if any one of them had failed, an arithmetic mistake would have been hiding in the working.

Think Like a Physicist: A compound capacitor problem is a sequence of two-capacitor problems, each handled with the appropriate parallel or series rule. The discipline is to combine in stages: identify the innermost group, reduce it to a single equivalent, then move outward. At each stage, write down the new combined capacitance and carry it forward as a single labelled quantity (here, C_{12}) rather than substituting its numerical value into the next equation. Carrying symbolic labels prevents arithmetic errors from accumulating through the stages, and it also makes it easy to see which individual capacitor is responsible for which physical effect at the end.

HOT Example 12

Use the following figure to find: (a) the equivalent capacitance of the whole combination; (b) the charge stored on the $3\mu\text{F}$ capacitor; and (c) the charge on each of the $6\mu\text{F}$ capacitors.

**Solution**

(a) The three $6\mu\text{F}$ capacitors are in series. Using the reciprocal-sum formula:

$$\frac{1}{C_{\text{series}}} = \frac{1}{6\mu\text{F}} + \frac{1}{6\mu\text{F}} + \frac{1}{6\mu\text{F}} = \frac{3}{6\mu\text{F}} = \frac{1}{2\mu\text{F}}$$

So $C_{\text{series}} = 2\mu\text{F}$.

The series stack is now in parallel with the $3\mu\text{F}$ capacitor, and parallel capacitances add directly:

$$C_{\text{eq}} = C_{\text{series}} + 3\mu\text{F} = 2\mu\text{F} + 3\mu\text{F} = 5\mu\text{F}$$

(b) The $3\mu\text{F}$ capacitor sees the full supply voltage of 10V , so its charge is:

$$Q_{3\mu\text{F}} = CV = (3\mu\text{F})(10\text{V}) = 30\mu\text{C}$$

(c) The series stack also sees 10V across its terminals as a whole, so the charge on the equivalent $2\mu\text{F}$ capacitor is:

$$Q_{\text{series}} = (2\mu\text{F})(10\text{V}) = 20\mu\text{C}$$

Every capacitor inside the series chain carries this same charge, since the chain is a single closed loop with no branching: each of the three $6\mu\text{F}$ capacitors carries $20\mu\text{C}$ on its plates.

Making Sense of the Answer: Three identical $6\mu\text{F}$ capacitors in series behave as a single $2\mu\text{F}$ capacitor: the series chain divides each individual capacitance by the number of capacitors in the chain when those capacitors are equal. Adding a parallel $3\mu\text{F}$ lifts the combination to $5\mu\text{F}$. The pattern recurs constantly in practical circuit design: series stacking lets a designer build a high-voltage capacitor from a chain of lower-voltage units, at the cost of a smaller total capacitance, and a parallel branch can then be used to top up the total to a desired value.

Think Like a Physicist: A compound capacitor network is always disentangled by combining innermost groups first. Identify any sub-network whose capacitors are all in pure series or all in pure parallel, replace that sub-network with its single equivalent, redraw the circuit, and repeat. Within three or four redraws even an intimidating-looking network reduces to a single equivalent capacitor across the supply. The discipline of redrawing at every stage is the only thing that keeps the bookkeeping clean on harder problems.

HOT Example 13

A $4\mu\text{F}$ capacitor is charged to a voltage of 12V by one source. Separately, a $6\mu\text{F}$ capacitor is charged to a voltage of 18V by another source. Both sources are then disconnected. The two capacitors are wired together so that their positive plates are joined by one wire and their negative plates by another. Find (a) the final common voltage across the two capacitors; (b) the new charge on each capacitor; (c) the magnitude of charge that flowed from one capacitor to the other during the redistribution.

Solution

(a) Before the connection, the charge on each capacitor is:

$$Q_{1,\text{initial}} = C_1 V_1 = (4\mu\text{F})(12\text{V}) = 48\mu\text{C}$$

$$Q_{2,\text{initial}} = C_2 V_2 = (6\mu\text{F})(18\text{V}) = 108\mu\text{C}$$

After the connection, charge has redistributed but the total charge on the connected (positive) side is unchanged:

$$Q_{\text{total}} = Q_{1,\text{initial}} + Q_{2,\text{initial}} = 48\mu\text{C} + 108\mu\text{C} = 156\mu\text{C}$$

The two capacitors are now in **parallel**; their equivalent capacitance is:

$$C_{eq} = C_1 + C_2 = 4\mu\text{F} + 6\mu\text{F} = 10\mu\text{F}$$

The common final voltage on the two capacitors is:

$$V_{final} = \frac{Q_{total}}{C_{eq}} = \frac{156\mu\text{C}}{10\mu\text{F}} = 15.6\text{V}$$

(b) The new charge on each capacitor at this common voltage is:

$$Q_{1,new} = C_1 V_{final} = (4\mu\text{F})(15.6\text{V}) = 62.4\mu\text{C}$$

$$Q_{2,new} = C_2 V_{final} = (6\mu\text{F})(15.6\text{V}) = 93.6\mu\text{C}$$

Cross-check on charge conservation: $Q_{1,new} + Q_{2,new} = 62.4\mu\text{C} + 93.6\mu\text{C} = 156\mu\text{C}$, which equals Q_{total} as it must.

(c) The charge that flowed from C_2 to C_1 during the redistribution is the difference between the initial and final charges on C_1 :

$$\Delta Q = Q_{1,new} - Q_{1,initial} = 62.4\mu\text{C} - 48\mu\text{C} = 14.4\mu\text{C}$$

An identical magnitude of charge ($14.4\mu\text{C}$) flowed off C_2 , since $Q_{2,initial} - Q_{2,new} = 108\mu\text{C} - 93.6\mu\text{C} = 14.4\mu\text{C}$, in agreement with charge conservation.

Making Sense of the Answer: The final voltage of 15.6V sits between the two initial voltages of 12V and 18V, as expected: charge can only redistribute, never appear from nothing or vanish into nothing. The final voltage is closer to 18V than to 12V (15.6V is about 60% of the way from 12V to 18V), and the reason is that the larger capacitor was the one that started at 18V: the $6\mu\text{F}$ capacitor carries more weight in the weighted average that determines V_{final} . The $4\mu\text{F}$ capacitor's voltage rose by 3.6V, gaining $14.4\mu\text{C}$ of charge; the $6\mu\text{F}$ capacitor's voltage fell by 2.4V, losing the same $14.4\mu\text{C}$. The smaller capacitor's voltage rose more sharply because the same transferred charge has a larger effect on a smaller capacitance.

Think Like a Physicist: Every charge-sharing problem reduces to a single principle: conservation of charge between the moment before the connection and the moment after. Begin every such problem by writing down the total charge on the connected side before the connection. Then equate it to (sum of capacitances)(final voltage) after the connection. The whole problem is then one equation in one unknown. A common trap, set by question setters with mischievous intentions, is to specify the connection as positive plate to negative plate, which makes the initial charges partly cancel. Always read the question carefully on this point: the polarity of the connection is part of the physics, not a stylistic detail. For the reversed-polarity case mentioned in the derivation, the total charge after the connection is:

$$Q_{total, reversed} = |C_1 V_1 - C_2 V_2|$$

and the final voltage is correspondingly smaller, sometimes dramatically so.

Combinations of capacitors are now in hand. The next section asks *what is being stored when a capacitor holds a charge, and how much of it*.

ENERGY STORED IN A CAPACITOR

A capacitor that holds a charge Q at a voltage V holds something else as well: **energy**. The work done by the source to push that charge onto the plates does not vanish at the end of charging; it is stored in the electric field between the plates, ready to be released the moment a path is provided. The camera flash, the defibrillator paddle, the spark plug of a petrol engine, and the smoothing capacitor of every laptop charger: all of these are devices that exist specifically to take energy slowly from a low-voltage supply and release it suddenly when the user demands it.

To find how much energy a capacitor stores, imagine the capacitor partway through being charged. At some intermediate instant the plates carry a charge q (less than the final charge Q), and the voltage at this instant is:

$$v = \frac{q}{C}$$

The source now pushes a tiny extra charge dq onto the positive plate, against the voltage v that the already-accumulated charge is producing. The work done by the source in this tiny step is the voltage times the charge moved:

$$dW = vdq = \frac{q}{C} dq$$

The total work done to charge the capacitor from empty ($q = 0$) to fully charged ($q = Q$) is the integral of this expression from 0 to Q :

$$W = \int_0^Q \frac{q}{C} dq = \left[\frac{q^2}{2C} \right]_0^Q = \frac{Q^2}{2C}$$

All of this work has been stored in the capacitor as the energy of the electric field between its plates. We call this stored energy U :

$$U = \frac{Q^2}{2C}$$

The same result can be obtained by a picture. The voltage during the charging process grows in direct proportion to the charge already on the plates: $v = q/C$ is a straight line through the origin. If we plot v against q , as charge climbs from 0 to Q and voltage climbs from 0 to V , the work done is the area under this straight line. That area is a triangle of base Q and height V .

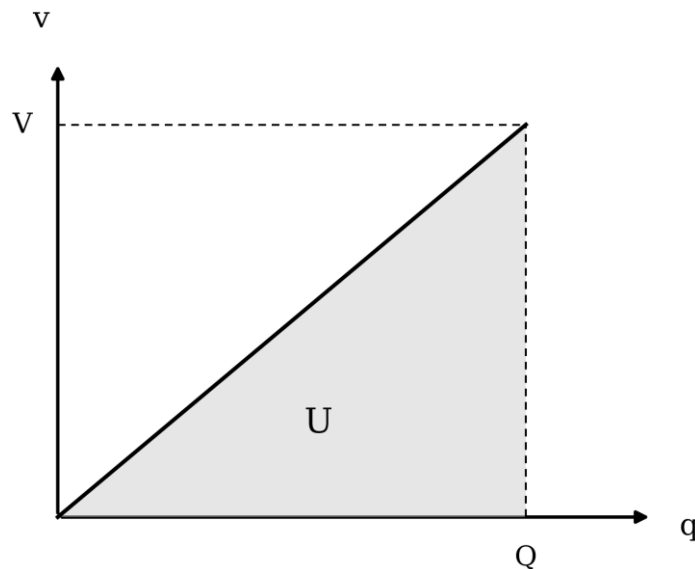


Figure: During charging, the voltage v grows linearly with the charge q . The shaded triangle has area $\frac{1}{2}QV$, and this area is the energy stored in the capacitor when it is fully charged.

The area of a triangle of base Q and height V is:

$$U = \frac{1}{2}QV$$

Two further forms follow at once from the relation $Q = CV$. Substituting $Q = CV$ in the formula above:

$$U = \frac{1}{2}QV = \frac{1}{2}(CV)V = \frac{1}{2}CV^2$$

Or, substituting $V = Q/C$ in the same expression:

$$U = \frac{1}{2}QV = \frac{1}{2}Q \times \frac{Q}{C} = \frac{Q^2}{2C}$$

The energy stored in a capacitor can therefore be written in three equivalent forms, depending on which two of the three quantities (Q , V , C) are most naturally available in the problem at hand:

$$U = \frac{1}{2}QV = \frac{1}{2}CV^2 = \frac{Q^2}{2C}$$

The first form is the most general; it relies on knowing both Q and V . The second is most useful when the capacitor is connected to a fixed voltage supply, in which case V is given directly. The third is most useful when the supply has been disconnected and the charge Q is held fixed by isolation. All three forms describe the same physical quantity; substituting $Q = CV$ takes any one of them to either of the other two.

It is illuminating to ask not how much energy a particular capacitor stores, but how much energy is stored per unit volume of the field between its plates. For a parallel-plate capacitor of plate area A , gap d , and field $E = V/d$ between the plates, the energy is:

$$U = \frac{1}{2} CV^2 = \frac{1}{2} \left(\frac{A\epsilon_0}{d} \right) \times (Ed)^2 = \frac{1}{2} \epsilon_0 E^2 (Ad)$$

The product Ad is the volume of space between the plates, in which the field E resides. Dividing the energy by this volume gives *the energy stored per unit volume of the field*, called the **energy density** and written u :

$$u = \frac{\text{Total energy stored}}{\text{Volume of electric field}} = \frac{U}{Ad} = \frac{1}{2} \epsilon_0 E^2$$

This is a general result, not restricted to parallel-plate capacitors. Wherever an electric field E exists, the energy stored in that field, per unit volume of space, is $\frac{1}{2} \epsilon_0 E^2$. This formula gives a meaning to the field itself as a storehouse of energy, independent of whatever distribution of charges produced it.

Force between the plates

Two oppositely-charged plates attract each other. *How strong is the attraction?* The cleanest way to find out is to use the energy formula. Consider a parallel-plate capacitor with charge Q on each plate, isolated from its supply so that no charge can come or go. The stored energy at gap d is:

$$U(d) = \frac{Q^2}{2C} = \frac{Q^2 d}{2A\epsilon_0}$$

Now imagine pulling the plates apart by a tiny extra distance dd , doing positive mechanical work against the attractive force. The work done in this small move is $dW = F dd$, and since no charge can enter or leave the isolated capacitor, this work goes entirely into increasing the stored energy:

$$F dd = dU = \frac{Q^2}{2A\epsilon_0} dd$$

Dividing both sides by dd , the magnitude of the attractive force between the plates is:

$$F = \frac{Q^2}{2A\epsilon_0}$$

This can be rewritten in terms of the field E between the plates. Using $Q = A\epsilon_0 E$ (which follows from $Q = CV$ with $C = A\epsilon_0/d$ and $V = Ed$):

$$F = \frac{A^2 \epsilon_0^2 E^2}{2A\epsilon_0} = \frac{1}{2} \epsilon_0 E^2 A$$

The force per unit area is therefore $\frac{F}{A} = \frac{1}{2} \epsilon_0 E^2$, which is exactly the energy density u of the field between the plates. Force per area equals energy per volume; a deep coincidence rooted in the principle of virtual work, and one that the reader will meet again whenever a system stores energy in a field that can be deformed.

The factor of $\frac{1}{2}$ in $F = \frac{Q^2}{2\epsilon_0 A}$ has a simple physical origin worth pausing over. The total field in the gap of a parallel-plate capacitor has magnitude $\frac{\sigma}{\epsilon_0}$, where $\sigma = \frac{Q}{A}$ is the surface charge density. But this total field is the sum of equal contributions from the two plates: each plate alone produces a field of magnitude $\frac{\sigma}{2\epsilon_0}$, and the two contributions add in the gap (where they point in the same direction) to give the total $\frac{\sigma}{\epsilon_0}$.

When we ask for the force on one plate, we must use only the field produced by the other plate, since a plate cannot exert a force on itself. The field acting on either plate is therefore $\frac{\sigma}{2\epsilon_0}$, and the force on the plate of charge Q is $F = Q \times \frac{\sigma}{2\epsilon_0} = \frac{Q^2}{2A\epsilon_0}$. Hence the factor of $\frac{1}{2}$ in the answer is the geometric record of the fact

that exactly half of the gap field is produced by the very plate one is computing the force on. We can cross-check the energy-method result against this direct one: both routes give the same F , and the agreement is the consistency between virtual-work bookkeeping and direct field-charge calculation.

Seven worked examples follow. None of them is harder than it looks.

BINDER Example 14

A capacitor of capacitance $5\mu\text{F}$ is charged to a potential difference of 12V . Calculate the energy stored in three different ways: using $U = \frac{1}{2} CV^2$; using $U = \frac{1}{2} QV$; and using $U = \frac{Q^2}{2C}$. Show that all three forms yield the same numerical answer.

Solution

First compute the charge on the plates, which is needed for two of the three formulae:

$$Q = CV = (5\mu\text{F})(12\text{V}) = 60\mu\text{C}$$

Now compute the energy three ways. Using the second form:

$$U = \frac{1}{2} CV^2 = \frac{1}{2} (5\mu\text{F})(12\text{V})^2 = \frac{1}{2} (5\mu\text{F})(144\text{V}^2) = 360\mu\text{J}$$

Using the first form:

$$U = \frac{1}{2} QV = \frac{1}{2} (60\mu\text{C})(12\text{V}) = 360\mu\text{J}$$

Using the third form:

$$U = \frac{Q^2}{2C} = \frac{(60\mu\text{C})^2}{2 \times 5\mu\text{F}} = \frac{3600\mu\text{C}^2}{10\mu\text{F}} = 360\mu\text{J}$$

All three forms agree: the energy stored is $360\mu\text{J}$ (roughly one third of a millijoule).

Making Sense of the Answer: A capacitor of $5\mu\text{F}$ at 12V stores about a third of a millijoule, which is in the same ballpark as the kinetic energy of an ant walking at a brisk pace. The stored energy is small in everyday terms, but if it is released through a wire of small resistance in a fraction of a microsecond, the average power dissipated during the discharge can climb into the watt range. Capacitors are not about how much energy is stored; they are about how quickly that energy can be released.

Think Like a Physicist: The three forms $U = \frac{1}{2} QV = \frac{1}{2} CV^2 = \frac{Q^2}{2C}$ are the same equation wearing three different costumes, and learning to recognise which costume the question is asking for is half the work of solving energy problems. If the question gives C and V , use the middle form. If the question disconnects the supply and asks what happens next, use the third form, because Q is the quantity held fixed and V is the one that changes. Choosing the right form before substituting saves arithmetic and prevents the mistake of using a V that is no longer correct.

BINDER Example 15

A $100\mu\text{F}$ capacitor is charged to a potential difference of 200V . The capacitor is then short-circuited through a low-resistance wire and the discharge lasts a total time of 5ms . Find (a) the charge stored on the capacitor before the discharge; (b) the energy stored before the discharge; (c) the average current during the discharge; and (d) the average power delivered to the wire during the discharge.

Solution

(a) The charge stored on the capacitor before the discharge is:

$$Q = CV = (100 \times 10^{-6} \text{ F})(200\text{V}) = 2 \times 10^{-2} \text{ C} = 20\text{mC}$$

(b) The stored energy is:

$$U = \frac{1}{2} CV^2 = \frac{1}{2} \times (100 \times 10^{-6} \text{ F}) \times (200\text{V})^2 = 2\text{J}$$

(c) The average current during the discharge is the total charge that flows divided by the time:

$$I_{\text{avg}} = \frac{Q}{t} = \frac{2 \times 10^{-2} \text{ C}}{5 \times 10^{-3} \text{ s}} = 4 \text{ A}$$

(d) The average power delivered is the total energy divided by the time:

$$P_{\text{avg}} = \frac{U}{t} = \frac{2 \text{ J}}{5 \times 10^{-3} \text{ s}} = 400 \text{ W}$$

Making Sense of the Answer: Two joules sounds modest, but delivered in five milliseconds it amounts to 400 W of average power and four amps of average current, both numbers comparable to those of a small electric kettle, but only for a five-millisecond pulse. The peak current at the start of the discharge is much larger than the four-amp average (because the discharge follows an exponential decay rather than a uniform rate), but the average is the right number to compare with the kettle. A 100 μF capacitor at 200 V is a very ordinary component in a photographic flash unit.

Think Like a Physicist: Two of the four quantities asked for here are pure averages (current and power), and average is in both cases the total quantity divided by the time. Average is almost never the same as instantaneous in a capacitor discharge, where current and power both fall exponentially to zero. The student who blurs the two and writes $I_{\text{max}} = 4 \text{ A}$ or $P_{\text{max}} = 400 \text{ W}$ has misread the question. When a question asks for an average over a discharge, the answer is total energy (or total charge) divided by total time. When a question asks for an initial or peak value, the answer is the current or power at $t = 0$, which is set by the initial voltage and the circuit resistance, not by the time over which the discharge completes.

REAL Example 16

A capacitor of capacitance 100 μF is charged to a voltage of 10V by a battery, which is then disconnected. The charged capacitor is now connected by a pair of wires to a second, identical uncharged capacitor of capacitance 100 μF . The two capacitors quickly come to a common voltage. Find (a) the common final voltage; (b) the energy stored on the system before the connection; (c) the energy stored on the system after the connection; (d) account for the difference.

Solution

Mr. Akilikubwa stood at the blackboard with the problem written out, and pointed his chalk in turn at Kipanga and Kipute.

Mr. Akilikubwa: Four parts, three workings. Kipanga, take part (a). Apply charge conservation and find the final voltage on the combination.

Kipanga went to the board and wrote out the initial charge on the first capacitor:

$$Q_{\text{initial}} = C_1 V_{1,\text{initial}} = (100 \mu\text{F})(10\text{V}) = 1000 \mu\text{C}$$

After the connection, the two capacitors are wired in parallel; their equivalent capacitance is:

$$C_{\text{eq}} = C_1 + C_2 = 100 \mu\text{F} + 100 \mu\text{F} = 200 \mu\text{F}$$

Charge cannot disappear, so the same 1000 μC of total charge is now spread across the parallel combination. The common final voltage is:

$$V_{\text{final}} = \frac{Q_{\text{initial}}}{C_{\text{eq}}} = \frac{1000 \mu\text{C}}{200 \mu\text{F}} = 5 \text{ V}$$

Mr. Akilikubwa: Good. The connection has cut the voltage in half. Kipute, take part (b). The energy before.

Kipute moved to the board. Before the connection, all the energy is stored on the first capacitor; the second is empty.

$$U_{\text{before}} = \frac{1}{2} C_1 V_{1,\text{initial}}^2 = \frac{1}{2} (100 \mu\text{F})(10\text{V})^2 = 5000 \mu\text{J} = 5 \text{ mJ}$$

Mr. Akilikubwa: And part (c). After the connection.

Both capacitors now hold the same voltage 5V, so the total stored energy is the sum of the energies of the two:

$$U_{\text{after}} = \frac{1}{2} C_1 V_{\text{final}}^2 + \frac{1}{2} C_2 V_{\text{final}}^2 = \frac{1}{2} (100 \mu\text{F})(5\text{V})^2 + \frac{1}{2} (100 \mu\text{F})(5\text{V})^2 = 2500 \mu\text{J} = 2.5 \text{ mJ}$$

Kipanga: Sir, the answer is exactly half the original energy. We started with 5mJ, we now have 2.5mJ. Where is the other half?

Mr. Akilikubwa: That is part (d) of the question, and the part the question really cared about. Kipute, you tried this experiment yesterday with the two capacitors on the bench and the LED. What did you observe at the moment of connection?

Kipute: There was a small spark at the contact, and a faint click. The wire warmed slightly under my finger.

Mr. Akilikubwa: The wire **warmed**, the spark sparked, and a little **sound** was made. That is where the missing 2.5mJ went. The current during the equalisation was enormous (large charge moving in a short interval) and even the small resistance of the wires was enough to dissipate a significant fraction of the stored energy as **heat**. The remainder went into the snap of the spark and the faint **sound wave** you heard as a click.

Kipanga: But **what if the wires had been perfect conductors with no resistance at all?** Then no heat could be dissipated, and yet the missing energy is still half of the original.

Mr. Akilikubwa: An excellent question, and the answer is that perfect conductors would not save the energy either. A circuit made of two capacitors and a perfect-conducting loop is an oscillator: charge would slosh back and forth between the two capacitors forever, but in doing so it would broadcast **electromagnetic waves** into the surrounding space, as any oscillating charge does. The lost energy is then radiated away as **electromagnetic radiation**. Whether by heat, by sound, by spark, or by radiation, exactly half the original energy must leave the system during a direct capacitor-to-capacitor connection. There is no way to avoid it.

Kipute: So if we wanted to transfer energy from one capacitor to another without losing it, a direct wire is not the way.

Mr. Akilikubwa: Correct. You would have to put an inductor or a transformer between the two capacitors and let the energy pass through the magnetic field of the inductor on the way. An inductor can store energy without dissipating it (in the ideal case), and the energy can be transferred from one capacitor to the other through the magnetic energy of the inductor with no loss in principle. This is exactly what every camera flash circuit and every defibrillator charger does internally: it puts an inductor in the path between the small battery and the large storage capacitor, precisely to avoid the half-energy loss that a direct connection would suffer.

Making Sense of the Answer: The half-energy theorem of capacitor coupling is a feature, not a bug. It explains why every device that needs to take energy from a small supply and pile it into a storage capacitor uses an inductor or a transformer between the two: a direct wire connection would lose half the energy in the very act of transferring it. The numerical result we obtained (2.5mJ out of 5mJ) is exact, not approximate. It depends only on the fact that the two capacitors were **equal**; for unequal capacitors the lost fraction is also computable and is somewhere between 0 and 0.5.

Think Like a Physicist: Whenever an apparent paradox arises in physics, the resolution is almost always found in an assumption that you did not realise you were making. Here, the assumption was that the connecting wires were perfectly idealised: zero resistance, zero inductance, zero radiation. As soon as any one of those three properties is real, and at least one of them always is, the energy bookkeeping closes. The half-energy theorem is a special case of a far more general truth: nature does not let energy disappear, but she is sometimes shy about telling you where it has gone.

HOT Example 17

A 5 μ F capacitor is charged to a potential difference of 100V and disconnected from the supply. It is then connected, with like plates joined together, to an uncharged 3 μ F capacitor. Find (a) the common final voltage on the combination; (b) the charge on each capacitor after the connection; (c) the initial and final stored energies; and (d) the energy lost in making the connection.

Solution

(a) Before the connection, the 5 μ F capacitor carries:

$$Q_1 = C_1 V_1 = (5\mu\text{F})(100\text{V}) = 500\mu\text{C}$$

The 3 μ F capacitor carries no charge. The total charge on the joined positive side is therefore 500 μ C. After the connection, the two capacitors are in **parallel**; their equivalent capacitance is:

$$C_{eq} = C_1 + C_2 = 5\mu\text{F} + 3\mu\text{F} = 8\mu\text{F}$$

The common final voltage is:

$$V_f = \frac{Q_{total}}{C_{eq}} = \frac{500\mu\text{C}}{8\mu\text{F}} = 62.5\text{V}$$

(b) The new charges on the two capacitors at this common voltage:

$$Q_{1,new} = C_1 V_f = (5\mu\text{F})(62.5\text{V}) = 312.5\mu\text{C}$$

$$Q_{2,new} = C_2 V_f = (3\mu\text{F})(62.5\text{V}) = 187.5\mu\text{C}$$

Cross-check on charge conservation: $Q_{1,new} + Q_{2,new} = 312.5\mu\text{C} + 187.5\mu\text{C} = 500\mu\text{C}$, equal to the initial charge, as it must be.

(c) Initial and final stored energies:

$$U_{initial} = \frac{1}{2} C_1 V_1^2 = \frac{1}{2} \times (5 \times 10^{-6}\text{F}) \times (100\text{V})^2 = 2.5 \times 10^{-2}\text{J} = 25\text{mJ}$$

$$U_{final} = \frac{1}{2} C_{eq} V_f^2 = \frac{1}{2} \times (8 \times 10^{-6}\text{F}) \times (62.5\text{V})^2 = 1.56 \times 10^{-2}\text{J} = 15.6\text{mJ}$$

(d) The energy lost in making the connection is the difference:

$$\Delta U_{lost} = U_{initial} - U_{final} = 25\text{mJ} - 15.6\text{mJ} = 9.4\text{mJ}$$

About 9.4 millijoules of the original 25mJ has been dissipated as heat in the connecting wire and as a brief electromagnetic burst.

Making Sense of the Answer: Roughly 38 per cent of the original stored energy disappears in the very act of connecting two capacitors at different voltages. The loss is large not because the wires were bad, but because any real (or even ideal) connection forces a sudden current through whatever resistance and inductance the loop happens to have, and the work done by that current against itself is exactly the missing energy. This is why energy is never transferred between capacitors through a direct wire in any serious circuit; an inductor or a transformer is always placed in the path so that the energy can pass through a magnetic field and arrive at its destination intact.

Think Like a Physicist: Every charge-sharing problem reduces to a single principle: total charge on the joined positive side is unchanged by the connection. Write Q_{total} at the very top of the working, then divide by the **parallel** equivalent capacitance to get V_f . From V_f every other quantity (new charges, new energies, energy lost) follows in one line. Make energy bookkeeping a separate paragraph from the voltage and charge calculation: students who try to compute initial and final energies in the same paragraph almost always lose a factor somewhere.

REAL Example 18

The photographic flash unit on an old film camera uses a capacitor of capacitance $800\mu\text{F}$ charged to a voltage of 300V . When the camera button is pressed, the capacitor discharges through the flash bulb in 1.0ms . Calculate (a) the energy stored in the capacitor; (b) the average power dissipated by the flash bulb during the discharge; (c) compare this average power to that of an ordinary 100W incandescent lamp.

Solution

(a) The energy stored in the capacitor is:

$$U = \frac{1}{2} CV^2 = \frac{1}{2} (800\mu\text{F})(300\text{V})^2 = \frac{1}{2} (800 \times 10^{-6}\text{F})(90000\text{V}^2) = 36\text{J}$$

The capacitor stores 36J of energy.

(b) The discharge lasts 1.0ms . The average power dissipated by the flash bulb during this brief interval is:

$$P = \frac{U}{t} = \frac{36\text{J}}{1.0 \times 10^{-3}\text{s}} = 36000\text{W} = 36\text{kW}$$

(c) Calculating their ratio:

$$\frac{P_{\text{flash}}}{P_{\text{bulb}}} = \frac{36000\text{W}}{100\text{W}} = 360$$

The flash bulb operates at 360 times the power of an ordinary 100W incandescent lamp for the brief duration of the discharge.

Making Sense of the Answer: The same energy 36J delivered slowly by a 100W lamp would take 0.36s to be released, an interval over which the human eye sees an ordinary glow. Released instead in 1.0ms, the same energy produces a flash so intense that the shutter of a camera, which is open for a comparable duration, can record a complete photograph in the time between the start and the end of the discharge. The capacitor has not made any more energy; it has simply concentrated the same energy into a much shorter time, and time-concentration of energy is the entire purpose of the capacitor in a flash circuit.

Think Like a Physicist: Whenever a problem asks for the average power dissipated over a discharge interval, the working has two clean steps: first compute the total energy released, and second divide it by the duration. Doing these in the opposite order or trying to compute instantaneous power from current and voltage requires an exponential discharge curve, and is far more work for the same answer. The energy-and-time approach works regardless of the shape of the discharge curve, because U and t are total quantities that do not care about the micro-structure of how the discharge happens. Conservation of energy is a friend on every such problem.

HOT Example 19

Two capacitors of capacitance $1\mu\text{F}$ and $5\mu\text{F}$ respectively are charged separately, each to a potential difference of 100V. The positive plate of the $1\mu\text{F}$ capacitor is then wired to the negative plate of the $5\mu\text{F}$ capacitor, and the negative plate of the $1\mu\text{F}$ capacitor to the positive plate of the $5\mu\text{F}$ capacitor. Find (a) the final common voltage on the combination, with its correct sign; (b) the energy lost in the connection.

Solution

(a) Before the connection, each capacitor carries:

$$Q_1 = C_1 V_1 = (1\mu\text{F})(100\text{V}) = 100\mu\text{C}$$

$$Q_2 = C_2 V_2 = (5\mu\text{F})(100\text{V}) = 500\mu\text{C}$$

Because unlike plates are joined together, the charges of opposite sign partially cancel on the joined side. The net charge on the joined positive side is the magnitude of the difference:

$$Q_{\text{total}} = |Q_2 - Q_1| = |500\mu\text{C} - 100\mu\text{C}| = 400\mu\text{C}$$

The sign of the resulting field is determined by the larger of the two initial charges (here Q_2 , the charge on the $5\mu\text{F}$ capacitor), so the joined plate that was originally the positive plate of the $5\mu\text{F}$ capacitor remains positive in the final state.

The two capacitors are now in parallel with $C_{\text{eq}} = 1\mu\text{F} + 5\mu\text{F} = 6\mu\text{F}$, and the common final voltage is:

$$V_f = \frac{Q_{\text{total}}}{C_{\text{eq}}} = \frac{400\mu\text{C}}{6\mu\text{F}} = 66.7\text{ V}$$

(b) The initial total stored energy is:

$$U_{\text{initial}} = \frac{1}{2} C_1 V_1^2 + \frac{1}{2} C_2 V_2^2 = \frac{1}{2} (1 \times 10^{-6}\text{F})(100\text{V})^2 + \frac{1}{2} (5 \times 10^{-6}\text{F})(100\text{V})^2 = 0.03\text{J} = 30\text{mJ}$$

The final stored energy is:

$$U_{\text{final}} = \frac{1}{2} C_{\text{eq}} V_f^2 = \frac{1}{2} \times (6 \times 10^{-6}\text{F}) \times (66.7\text{V})^2 = 0.0133\text{J} = 13.3\text{ mJ}$$

Energy lost in the connection:

$$\Delta U_{\text{lost}} = U_{\text{initial}} - U_{\text{final}} = 30\text{mJ} - 13.3\text{mJ} = 16.7\text{mJ}$$

Making Sense of the Answer: More than half of the original stored energy (about 56 per cent) is dissipated during the unlike-plates connection, substantially more than in the like-plates case of the previous example. The reason is geometric. Unlike-plates connection forces a much larger redistribution of charge through the wire than the like-plates connection of the previous example, where charge merely redistributes between two

capacitors of like polarity. Larger transient current means more heat dissipated, full stop. The lesson for designers is that unlike-plates accidents are far more damaging than like-plates errors.

Think Like a Physicist: Sign conventions are the natural hazard of unlike-plates problems. The safest habit is to compute the magnitudes of all charges first, then ask which initial charge is the larger of the two: that capacitor's polarity determines the sign of the final state, and the magnitude of the difference is the conserved net charge. Trying to carry signed charges through the arithmetic without first clarifying which way they will end up almost always produces sign errors that are difficult to track back to their source.

HOT Example 20

A parallel-plate capacitor consists of two square metal plates of side 10cm separated by an air gap of 1mm. The capacitor is connected to a 500V supply and allowed to charge fully. Calculate the magnitude of the attractive force between the plates.

Solution

The plate area, converted to metres, is:

$$A = (0.10\text{m})^2 = 0.01\text{m}^2 = 10^{-2}\text{m}^2$$

The gap, converted to metres, is:

$$d = 1\text{mm} = 10^{-3}\text{m}$$

The field between the plates is $E = V/d$:

$$E = \frac{V}{d} = \frac{500\text{V}}{10^{-3}\text{m}} = 5 \times 10^5\text{V/m}$$

The magnitude of the attractive force between the plates is $F = \frac{1}{2} \epsilon_0 E^2 A$:

$$F = \frac{1}{2} (8.854 \times 10^{-12}\text{F/m}) (5 \times 10^5\text{V/m})^2 (10^{-2}\text{m}^2) = 1.1 \times 10^{-2}\text{N}$$

The attractive force between the plates is approximately $1.1 \times 10^{-2}\text{N}$ (about 11mN).

Making Sense of the Answer: A force of 11mN is small but easily measurable; a sensitive postal balance would detect it without difficulty. The force scales as E^2 , so doubling the voltage at the same gap would quadruple the force, and halving the gap at the same voltage would also quadruple it (because $E = V/d$ doubles when d halves). Modern electrostatic loudspeakers, electrostatic micro-electromechanical sensors, and the inkjet printer heads that put writing on paper all exploit this scaling: tiny gaps and modest voltages give very large forces per unit area, in places where mechanical actuators of the same size could not deliver anything useful.

Think Like a Physicist: Two formulae for the inter-plate force exist, and they are equivalent: $F = \frac{Q^2}{2A\epsilon_0}$, which is best when the supply has been disconnected and the charge is held fixed; and $F = \frac{1}{2} \epsilon_0 E^2 A$, which is best when the supply is connected and the voltage (and so the field) is held fixed. Picking the right form for the constraint of the problem saves a substitution and often a sign error. As with the energy formula, the discipline of choosing the form before substituting is more than half the battle.

The energy and the force a capacitor exerts on its own plates are now in hand. The next section turns to a question we have so far avoided: what difference does it make to all of the above if the gap between the plates is filled with an insulating material instead of air or vacuum?

DIELECTRICS

So far the gap between the plates has been air, or, if we have wished to be careful about the small ionisation correction, vacuum. In every real capacitor the gap is filled with a solid or liquid insulator: paper, mica, ceramic, polyester, polypropylene, electrolytic gel. The insulator does three jobs at once.

- It increases the capacitance by a factor that depends on the material.
- It raises the voltage at which the capacitor breaks down, often by a large factor.
- It gives the plates something to lean on, so a capacitor the size of a fingernail can have its plates only ten micrometres apart without the smallest mechanical knock putting them in contact.

The insulating filler is called a **dielectric**. The factor by which the capacitance increases when the dielectric fills the gap is called the **dielectric constant** of the material, written ϵ_r . It is dimensionless. Vacuum has $\epsilon_r = 1$ by definition; air has $\epsilon_r = 1.0006$, indistinguishable from vacuum in almost every practical calculation. Paper has ϵ_r between 2 and 4 depending on the grade; mica is around 5 to 7; common ceramic capacitor materials run from 6 up to several thousand; water has $\epsilon_r \approx 81$; certain titanate ceramics used in modern surface-mount components can reach $\epsilon_r \approx 10\,000$.

How a dielectric works

Before applying any formula, consider the following figure.

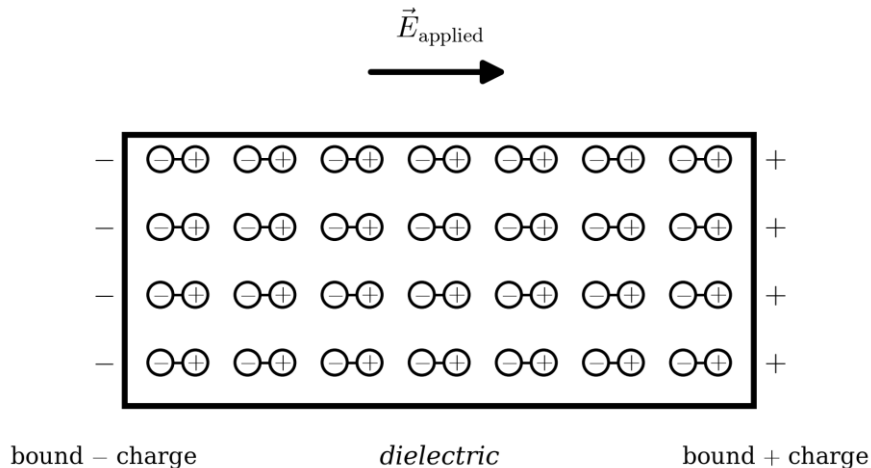


Figure: Polar molecules inside a dielectric align with the applied field $\vec{E}_{\text{applied}}$. The aligned dipoles cancel inside the bulk of the material, leaving uncompensated bound charge only on the two faces of the slab: negative bound charge on the face nearest the positive plate, positive bound charge on the face nearest the negative plate.

Every molecule contains electrons and nuclei. In the absence of an external field, these charges are arranged so that the molecule has no net dipole moment; or, if it is a polar molecule like water, its many dipoles point in random directions because of thermal motion, and the average dipole is zero. Place the molecule in an electric field, and one of two things happens. If the molecule is non-polar, the electron cloud is pulled slightly against the field direction relative to the positive nuclei, producing a small induced dipole. If the molecule is polar, the existing permanent dipole is partly turned into alignment with the field, against the randomising effect of thermal motion.

Either way, the population of aligned dipoles produces a thin layer of bound positive charge on the face of the dielectric pointing towards the negative plate, and a matching layer of bound negative charge on the face pointing towards the positive plate. These bound charges produce their own electric field, in the opposite direction to the applied field. The net electric field inside the dielectric is therefore **less** than the field that would be there if the dielectric were absent.

Quantitatively, the net electric field inside the dielectric is reduced by exactly the factor ϵ_r :

$$\mathbf{E}_{\text{in dielectric}} = \frac{\mathbf{E}_{\text{vacuum}}}{\epsilon_r}$$

For a parallel-plate capacitor with a fixed charge Q on the plates, the field outside the dielectric is set by:

$$\sigma = \frac{Q}{A},$$

and is independent of what fills the gap.

The field inside the dielectric is reduced by ϵ_r . The voltage across the plates, which is the line integral of the field across the gap, is therefore reduced by ϵ_r . Since $C = Q/V$, and Q is the same, C is multiplied by ϵ_r :

$$C = \epsilon_r C_0$$

Where C_0 is the capacitance of the same geometry with vacuum (or air) in the gap. The whole of this section reduces, in slogan form, to this one equation. Everything that follows is a matter of unpacking what ϵ_r means when the dielectric fills only part of the gap, or when more than one dielectric is present.

Two examples follow. The first puts the slogan formula to work; the second uses it to solve a small manufacturing problem.

BINDER Example 21

An air-gap parallel-plate capacitor has capacitance 50pF. A slab of mica with dielectric constant $\epsilon_r = 6$ is cut to exactly fit the gap and slid into place, filling the space between the plates completely. Find the new capacitance.

Solution

The capacitance with vacuum (or air) in the gap is the value given: $C_0 = 50 \text{ pF}$

With the gap completely filled by a dielectric of constant ϵ_r , the new capacitance is:

$$C = \epsilon_r C_0 = 6 \times 50 \text{ pF} = 300 \text{ pF}$$

The new capacitance is 300pF.

Making Sense of the Answer: *The mica did exactly what the slogan promised: it multiplied the capacitance by exactly ϵ_r . Nothing more is needed to predict the new value once ϵ_r is known. A factor of 6 from a slip of mica thinner than a postage stamp is a remarkable efficiency, and is the reason mica capacitors held the high-precision market for half a century before ceramic took over.*

Think Like a Physicist: *The formula $C = \epsilon_r C_0$ does not depend on the geometry. It works for any capacitor whose gap is uniformly filled with the same dielectric: parallel plates, concentric cylinders, concentric spheres, anything. The geometric details are hidden inside C_0 ; the material details are hidden inside ϵ_r ; and the two factors enter the answer multiplicatively.*

REAL Example 22

A small factory in Tabora makes rolled paper capacitors for export. The paper they use is 40mm wide, 0.015mm thick, and has a dielectric constant of 2.5. Two long aluminium foil strips are laid against the paper, a second sheet of paper goes on top, and the four-layer sandwich is rolled into a small cylinder. The capacitor is meant to have a capacitance of 2.0μF.

- Find the length of paper, and so of foil, that the factory must roll up to make a single capacitor of this rating.
- Explain one practical reason why the rolled construction works, and one practical reason why the design might fail in service.

Solution

Treat the unrolled sandwich as a single parallel-plate capacitor of plate area $A = Lw$, where L is the length of the foil and $w = 40 \text{ mm} = 0.040 \text{ m}$ is the width, with the paper as the dielectric of thickness $d = 0.015 \text{ mm} = 1.5 \times 10^{-5} \text{ m}$. The capacitance is:

$$C = \frac{\epsilon_r \epsilon_0 A}{d} = \frac{\epsilon_r \epsilon_0 Lw}{d}$$

Solving for L :

$$L = \frac{Cd}{\epsilon_r \epsilon_0 w} = \frac{(2.0 \times 10^{-6} \text{ F})(1.5 \times 10^{-5} \text{ m})}{(2.5)(8.854 \times 10^{-12} \text{ F/m})(0.040 \text{ m})} = 33.9 \text{ m}$$

A length of about 33.9m of paper, and the same of each foil strip, is required for one 2.0μF capacitor.

- Practical observations.

Why the construction works: The rolled-strip design delivers a very large plate area in a very small package. Thirty-four metres of foil is the area that the formula demands for a 2μF capacitor at this gap and dielectric, but a 34m-long flat sheet would not fit on a workbench. Rolling the sandwich into a cylinder a few centimetres long preserves the foil-to-foil area exactly (the formula does not care how the area is shaped) while shrinking the package down to something that fits comfortably onto a circuit board. The rolled construction is therefore a geometric trick that buys area density without changing the electrical parameters.

Why the design might fail in service: Paper is hygroscopic; it absorbs water vapour from the surrounding air. Once moisture has soaked into the paper between the foils, the dielectric strength of the paper drops sharply, because dissolved ions in the absorbed water provide a conduction path between the plates. In a humid coastal climate such as that of Tanga or Dar es Salaam, an unsealed paper capacitor can lose its rated voltage in months. Manufacturers solve this by impregnating the paper with oil or wax and sealing the rolled cylinder inside a metal can; an exported capacitor that arrives without such sealing will eventually short out in service.

Making Sense of the Answer: *Thirty-four metres of paper to make a single small component sounds extravagant until one remembers that the rolled cylinder is perhaps three centimetres long. The roll squeezes that long strip into a package the size of a thumb. The trick of the rolled construction is exactly this: trading a long thin sheet for a stubby cylinder, with no change in the area or in the gap that the formula sees.*

Think Like a Physicist: *Notice how nothing in the calculation depended on the radius of the rolled cylinder, on how tightly the cylinder was wound, or on how many turns the paper made. All that matters is the total area of foil facing foil with paper between them, and that is Lw , the area of the unrolled sandwich. When a problem looks geometrically intimidating, ask yourself whether the answer cares about the intimidating part of the geometry. Very often, it does not.*

Partially-filled gap

The mica slab in the last example was cut to exactly fit the gap. Suppose now it is too thin: it fills part of the gap but leaves a small air space between the slab and one of the plates. Or suppose, for safety reasons, the designer wants a sliver of air on either side of the slab so that the slab is not in electrical contact with either plate. What is the capacitance of such a partially-filled gap?

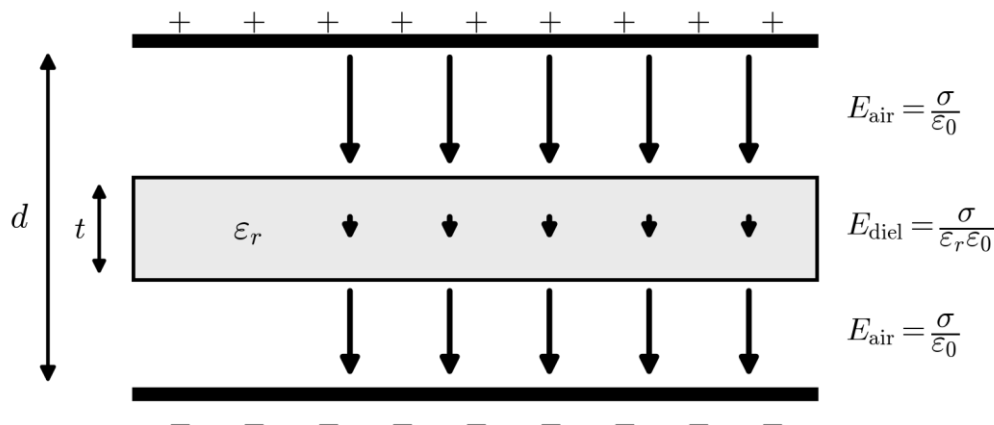


Figure: Partially-filled gap. Total separation d , slab of thickness t with dielectric constant ϵ_r somewhere in the gap, leaving $(d - t)$ of air. The field in the air region is the full vacuum value σ/ϵ_0 ; the field inside the dielectric is reduced to $\sigma/(\epsilon_r \epsilon_0)$. The arrows inside the slab are drawn shorter to indicate the smaller field strength there.

Let the total gap be d and the slab thickness be t , with $t < d$. The plates carry surface charge density $\sigma = Q/A$. The field outside the dielectric is the full vacuum value, σ/ϵ_0 ; the field inside the dielectric is reduced by ϵ_r to $\sigma/(\epsilon_r \epsilon_0)$. The potential difference across the whole gap is the sum of the two contributions:

$$V = E_{\text{air}}(d - t) + E_{\text{diel}} t = \frac{\sigma}{\epsilon_0} (d - t) + \frac{\sigma}{\epsilon_r \epsilon_0} t$$

Factor out σ/ϵ_0 :

$$V = \frac{\sigma}{\epsilon_0} \left((d - t) + \frac{t}{\epsilon_r} \right) = \frac{Q}{\epsilon_0 A} \left((d - t) + \frac{t}{\epsilon_r} \right)$$

The capacitance is Q/V :

$$C = \frac{\epsilon_0 A}{(d - t) + \frac{t}{\epsilon_r}}$$

Two limit checks.

- When $t \rightarrow 0$, the slab vanishes and the bracket reduces to d , giving $C = \frac{\epsilon_0 A}{d}$, the full air-gap capacitance, as it must.
- When $t \rightarrow d$, the slab fills the gap and the bracket reduces to d/ϵ_r , giving $C = \frac{\epsilon_r \epsilon_0 A}{d}$, the full dielectric result. Both limits recover the cases we already understand.

A second way to arrive at the same formula is to notice that the partially-filled gap is two capacitors in series: an air capacitor of thickness $(d - t)$ stacked on top of a dielectric capacitor of thickness t . The series combination gives the same answer, with a little less arithmetic if the student is comfortable with series-capacitor algebra. Either route is acceptable.

Two examples follow.

HOT Example 23

A parallel-plate capacitor has plates of area 200cm^2 separated by an air gap of 0.5mm . A slab of dielectric constant 4 and thickness 0.3mm is inserted, with the remaining 0.2mm left as air. Find the new capacitance.

Solution

Convert all lengths and the area to SI base units:

$$A = 200\text{cm}^2 = 2.00 \times 10^{-2}\text{m}^2 \quad d = 0.5\text{mm} = 5 \times 10^{-4}\text{m} \quad t = 0.3\text{mm} = 3 \times 10^{-4}\text{m} \quad \epsilon_r = 4$$

Use the partial-fill formula:

$$C = \frac{\epsilon_0 A}{(d - t) + \frac{t}{\epsilon_r}} = \frac{(8.854 \times 10^{-12}\text{F/m})(2.00 \times 10^{-2}\text{m}^2)}{(5 \times 10^{-4}\text{m} - 3 \times 10^{-4}\text{m}) + \left(\frac{3 \times 10^{-4}\text{m}}{4}\right)} = 6.44 \times 10^{-10}\text{F} = 644\text{pF}$$

The partially-filled capacitor stores about 644pF .

Alternative solution (as two capacitors in series)

The partial-fill geometry is two capacitors stacked between the same pair of plates: an air-only layer of thickness $(d - t) = 0.2\text{mm}$ on one side of the slab, and a fully dielectric-filled layer of thickness $t = 0.3\text{mm}$ on the other. The same charge Q passes through both layers, so they are in series.

The air layer alone has capacitance:

$$C_{\text{air}} = \frac{\epsilon_0 A}{d - t} = \frac{(8.854 \times 10^{-12}\text{F/m})(2.00 \times 10^{-2}\text{m}^2)}{2 \times 10^{-4}\text{m}} = 8.854 \times 10^{-10}\text{F} = 885.4\text{pF}$$

The dielectric layer alone has capacitance:

$$C_{\text{diel}} = \frac{\epsilon_r \epsilon_0 A}{t} = \frac{(4)(8.854 \times 10^{-12}\text{F/m})(2.00 \times 10^{-2}\text{m}^2)}{3 \times 10^{-4}\text{m}} = 2.361 \times 10^{-9}\text{F} = 2361\text{pF}$$

The product-over-sum form of the two-capacitor series rule gives:

$$C = \frac{C_{\text{air}} C_{\text{diel}}}{C_{\text{air}} + C_{\text{diel}}} = \frac{(885.4\text{pF})(2361\text{pF})}{885.4\text{pF} + 2361\text{pF}} = 644\text{pF}$$

Either route is acceptable in an examination.

Making Sense of the Answer: Compare with the two endpoints. The same gap filled entirely with air would give $C_0 = \epsilon_0 A/d \approx 354\text{pF}$; filled entirely with the $\epsilon_r = 4$ slab it would give $4C_0 \approx 1416\text{pF}$. The actual partial-fill answer of 644pF sits between these two, closer to the air endpoint than to the full-slab endpoint. This is the correct intuition: 0.3mm of slab in a 0.5mm gap is more than half the gap, but the air strip is still close enough to the full-air answer that the boost from the slab is only about 82 per cent of the maximum.

Think Like a Physicist: A slab in a gap is two capacitors in series, an air one and a dielectric one. Series capacitors are dominated by the smaller of the two: whichever capacitor has the smaller C controls the total. The air strip here has $C_{\text{air}} = \epsilon_0 A/(d - t) \approx 885\text{pF}$, and the slab itself has $C_{\text{slab}} = \epsilon_r \epsilon_0 A/t \approx 2360\text{pF}$.

The air strip is the smaller and therefore dominates. To raise C further, the designer would have to thin the air strip, which is exactly why real commercial capacitors leave no air at all and fill the entire gap with dielectric.

HOT Example 24

A parallel-plate capacitor is charged from a battery, and the battery is then disconnected. A slab of dielectric of thickness 3mm is then introduced into the gap. To keep the potential difference across the plates the same as it was before the slab was inserted, the plate separation must be increased by 2.4mm. Find the dielectric constant of the slab.

Solution

After the battery is disconnected, the charge Q on the plates is fixed. The potential difference $V = Q/C$ therefore stays the same if and only if the capacitance C stays the same. The condition of the problem is therefore $C_{\text{new}} = C_{\text{old}}$.

Before the slab is inserted, the capacitor is a pure air gap of width d , with capacitance:

$$C_{\text{old}} = \frac{\epsilon_0 A}{d}$$

After the slab is inserted and the gap widened by $\Delta d = 2.4$ mm, the new total gap is $(d + \Delta d)$, of which thickness $t = 3$ mm is slab and the rest is air. Apply the partial-fill formula:

$$C_{\text{new}} = \frac{\epsilon_0 A}{(d + \Delta d - t) + \frac{t}{\epsilon_r}}$$

Setting $C_{\text{new}} = C_{\text{old}}$, the $\epsilon_0 A$ factors match and the numerators equate, so the denominators must equate:

$$d = (d + \Delta d - t) + \frac{t}{\epsilon_r}$$

The d terms cancel:

$$0 = \Delta d - t + \frac{t}{\epsilon_r}$$

Solve for ϵ_r :

$$\epsilon_r = \frac{t}{t - \Delta d} = \frac{3\text{mm}}{3\text{mm} - 2.4\text{mm}} = 5$$

The dielectric constant of the slab is 5.

Making Sense of the Answer: $\epsilon_r = 5$ is in the ballpark of mica, glass, or a stiff polymer; not the highest- K materials, but well above paper. The arithmetic is striking in how little of it survives. The answer comes from a single relation, $t(1 - 1/\epsilon_r) = \Delta d$, which has the simple interpretation that the slab replaces a layer of air of thickness Δd , and the dielectric constant is set by how thick a slab is required to replace a given thickness of air.

Think Like a Physicist: Two-line solutions are not always easier than three-line ones, but they always teach more. The clean relation $t(1 - 1/\epsilon_r) = \Delta d$ works for any slab and any compensation distance, and is worth keeping. It says that a thick slab with high ϵ_r is hardly any electrically denser than a thin slab with the same effect, because the ratio $\Delta d/t$ saturates at 1 as $\epsilon_r \rightarrow \infty$. A perfect conductor in the gap (where $\epsilon_r \rightarrow \infty$) replaces its own thickness exactly, which is the result we will meet again later in this chapter for the conducting slab.

Two dielectrics in one gap

If a single slab fills the gap, the answer is $C = \epsilon_r \epsilon_0 A/d$. If the gap is only partly filled, the answer is the formula derived just now. But sometimes a designer wants two different materials between the same pair of plates, either stacked one on top of the other across the full plate area, or sitting side by side, each covering its own piece of the plate area. Each arrangement gives a different formula, and which formula applies depends entirely on whether the two slabs share the same charge or the same voltage.

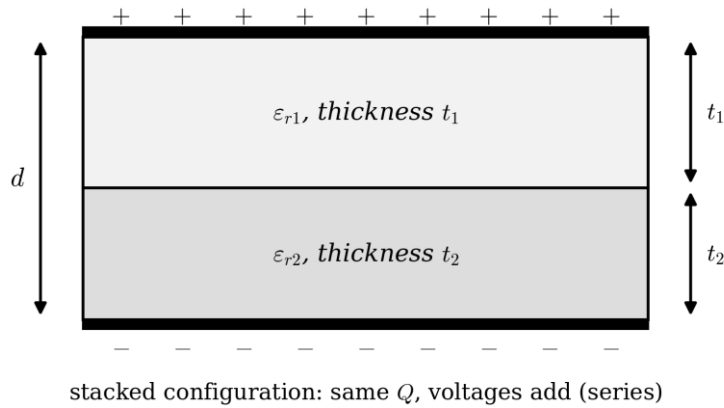


Figure: Stacked configuration. Two slabs of dielectric constants ϵ_{r1} and ϵ_{r2} , of thicknesses t_1 and t_2 , both spanning the full plate area A , stacked to fill the gap $d = t_1 + t_2$. The same charge Q passes through both slabs, so they are in series.

In the stacked configuration, both slabs span the full area A but each occupies a different fraction of the gap. The same charge Q flows through both regions, since charge cannot accumulate at the slab-to-slab interface; whatever leaves one plate must arrive at the other, passing through both materials in succession. The voltage across the pair is the sum of the voltages across each slab. This is the algebra of capacitors in series:

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} = \frac{t_1}{\epsilon_{r1}\epsilon_0 A} + \frac{t_2}{\epsilon_{r2}\epsilon_0 A} = \frac{\epsilon_{r2}t_1 + \epsilon_{r1}t_2}{\epsilon_{r1}\epsilon_{r2}\epsilon_0 A}$$

$$C = \frac{\epsilon_{r1}\epsilon_{r2}\epsilon_0 A}{\epsilon_{r2}t_1 + \epsilon_{r1}t_2} = \frac{\epsilon_0 A}{\frac{t_1}{\epsilon_{r1}} + \frac{t_2}{\epsilon_{r2}}}$$

For three or more layers the formula extends in the obvious way: the denominator gains another t/ϵ_r term for each new layer. In the symmetric case of two equal layers each of thickness $d/2$:

$$C_{\text{stacked, equal}} = \frac{2\epsilon_0 A}{d} \left(\frac{\epsilon_{r1}\epsilon_{r2}}{\epsilon_{r1} + \epsilon_{r2}} \right)$$

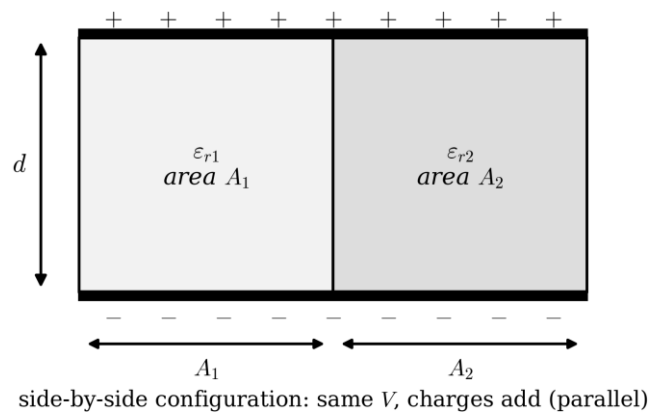


Figure: Side-by-side configuration. Two slabs of dielectric constants ϵ_{r1} and ϵ_{r2} , each spanning the full gap d but each covering a different piece of the plate area, A_1 and A_2 , with $A_1 + A_2 = A$. Both halves see the same voltage V , so they are in parallel.

In the side-by-side configuration, both slabs span the full gap d but each covers a different fraction of the plate area. Both halves see the same voltage V across the gap; the total charge is the sum of the charges on each half. This is the algebra of capacitors in parallel:

$$C = C_1 + C_2 = \frac{\epsilon_{r1}\epsilon_0 A_1}{d} + \frac{\epsilon_{r2}\epsilon_0 A_2}{d} = \frac{\epsilon_0}{d} (\epsilon_{r1}A_1 + \epsilon_{r2}A_2)$$

For two slabs covering equal halves of the plate, $A_1 = A_2 = \frac{A}{2}$:

$$C_{\text{side-by-side, equal}} = \frac{\epsilon_0 A}{2d} (\epsilon_{r1} + \epsilon_{r2})$$

The two configurations give different formulae for the same pair of materials. The stacked formula uses the harmonic mean of the dielectric constants, weighted by thickness; the side-by-side formula uses the arithmetic mean, weighted by area. The arithmetic mean is always larger than the harmonic mean, so for the same materials the side-by-side configuration always gives the larger capacitance, a result worth checking by direct substitution after every calculation.

These two configurations are nothing more than the series and parallel combinations from earlier in the chapter, dressed in different clothes. The dielectric slabs are themselves the capacitors. Once the student has spotted which configuration applies, the rest is bookkeeping. One example to demonstrate.

HOT Example 25

A parallel-plate capacitor has plates of area 100cm^2 separated by a gap of 4mm . The gap is completely filled by two dielectric slabs of equal thickness 2mm each. The upper slab has dielectric constant $\epsilon_{r1} = 2$ and the lower slab has $\epsilon_{r2} = 6$. Each slab spans the full plate area. Find the capacitance.

Solution

The two slabs are stacked one above the other, each spanning the full plate area. This is the stacked, or series, configuration. Use the symmetric form of the stacked formula, since $t_1 = t_2 = d/2$:

$$C = \frac{2\epsilon_0 A}{d} \left(\frac{\epsilon_{r1}\epsilon_{r2}}{\epsilon_{r1} + \epsilon_{r2}} \right)$$

$$C = \frac{2(8.854 \times 10^{-12} \text{ F/m})(1.00 \times 10^{-2} \text{ m}^2)}{4 \times 10^{-3} \text{ m}} \left(\frac{2 \times 6}{2 + 6} \right) = 66.4 \text{ pF}$$

The stacked two-dielectric capacitor stores about 66.4pF .

Making Sense of the Answer: Compare with the two materials acting alone in the full gap. A full $\epsilon_r = 2$ fill would give $2C_0 = 44.3 \text{ pF}$; a full $\epsilon_r = 6$ fill would give $6C_0 = 132.8 \text{ pF}$. The stacked answer of 66.4pF sits between these but well below the arithmetic mean of 88.5pF , because series combinations are dominated by the smaller capacitor, which here is the $\epsilon_r = 2$ slab. If the same two slabs were arranged side by side instead, the answer would be $\epsilon_0 A(2 + 6)/(2d) = 88.5 \text{ pF}$, exactly the arithmetic-mean prediction and larger than the stacked result.

Think Like a Physicist: Faced with a two-dielectric problem, the very first question to ask is not “what formula?” but “what is conserved across the slab boundary?”. If the same charge has to pass through both slabs (stacked), the formula is series. If the same voltage sits across both slabs (side by side), the formula is parallel. The geometry tells you which conservation applies, and the conservation tells you the formula. The harmonic-versus-arithmetic mean distinction follows automatically.

The capacitor and its materials are now in hand. The next part of this section asks what happens when the dielectric is inserted into a **charged** capacitor, with the battery either connected or disconnected, and the surprising contrast between those two cases.

Battery connected versus battery disconnected

When a dielectric is inserted into a charged capacitor, the result depends decisively on whether the battery is still connected to the plates or has been removed. The two cases give different final voltages, different final stored charges, different final field strengths, and different final stored energies. Examination problems frequently put this contrast at the centre of a question, and students who try to answer them without first sorting out which case applies almost always get the signs wrong. The contrast is worth pinning down once, carefully, and then using as a checklist.

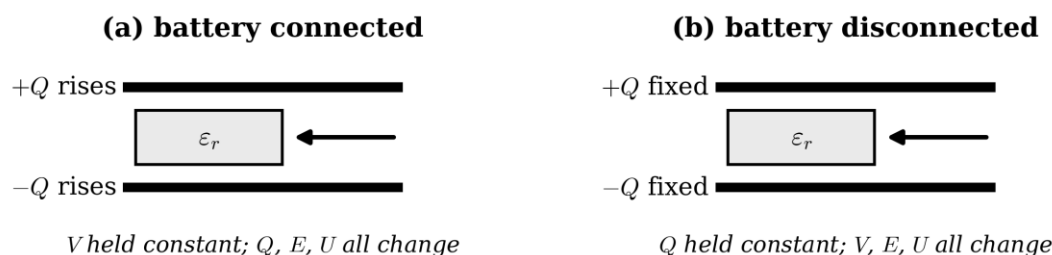


Figure: A dielectric slab is being inserted into a parallel-plate capacitor. In (a) the battery is connected throughout the insertion, so the voltage across the plates is fixed by the battery and the charge on the plates is whatever the battery has to supply to satisfy the new capacitance. In (b) the battery has been disconnected before the slab is introduced, so the charge on the plates is locked in and the voltage adjusts to whatever the new capacitance requires.

Take both cases starting from the same initial state: an air-gap capacitor of capacitance C_0 charged to a voltage V_0 , holding a charge $Q_0 = C_0 V_0$, with a field $E_0 = V_0/d$ between the plates and a stored energy $U_0 = 1/2 C_0 V_0^2$. A slab of dielectric constant ϵ_r is now inserted to fill the gap completely.

Case 1: battery connected

The battery enforces $V = V_0$ at all times. The capacitance rises to $\epsilon_r C_0$, so the charge on the plates must rise to $Q = (\epsilon_r C_0) V_0 = \epsilon_r Q_0$.

The field: Since V is unchanged and d is unchanged, the electric field stays at E_0 .

The stored energy: $1/2 CV^2$, rises to $1/2 (\epsilon_r C_0) V_0^2 = \epsilon_r U_0$. The extra charge $\epsilon_r Q_0 - Q_0 = (\epsilon_r - 1) Q_0$ is supplied by the battery, which does work $(\epsilon_r - 1) Q_0 V_0$ in the process. Of this work, the capacitor stores an extra $(\epsilon_r - 1) U_0$; the rest, also $(\epsilon_r - 1) U_0$, is done as mechanical work on the dielectric drawing it into the gap, against the fringing-field forces at the edges of the plates. The energy bookkeeping balances cleanly.

Case 2: battery disconnected

The charge $Q = Q_0$ is locked in. The capacitance still rises to $\epsilon_r C_0$, but with Q fixed and C larger, the voltage must drop. From $V = \frac{Q}{C}$, the new voltage is $V = \frac{Q_0}{\epsilon_r C_0} = \frac{V_0}{\epsilon_r}$.

The field: Since the gap is unchanged but V has dropped, the electric field falls to $E = \frac{V}{d} = \frac{E_0}{\epsilon_r}$.

The stored energy: $\frac{Q^2}{2C}$, drops to $\frac{Q_0^2}{2\epsilon_r C_0} = \frac{U_0}{\epsilon_r}$. The energy lost, $U_0 \left(1 - \frac{1}{\epsilon_r}\right)$, is also done as mechanical work on the dielectric drawing it into the gap; it comes out of the capacitor's stored energy rather than out of a battery, because no battery is present.

Both cases lose energy from the field-and-battery system. Both lose it for the same physical reason: *the dielectric is drawn into the gap and the system does work on it*. The difference between the two cases is purely accounting. In case 1, the battery refills the energy account; in case 2, the account starts smaller because nothing refills it.

The five-row summary, with every quantity tracked side by side, is the table every student should keep in mind while solving a problem of this type.

Quantity	Initial (air)	Battery connected	Battery disconnected
Capacitance	C_0	$\epsilon_r C_0$	$\epsilon_r C_0$
Charge	Q_0	$\epsilon_r Q_0$	Q_0
Voltage	V_0	V_0	$\frac{V_0}{\epsilon_r}$

Electric field	E_0	E_0	$\frac{E_0}{\epsilon_r}$
Stored energy	U_0	$\epsilon_r U_0$	$\frac{U_0}{\epsilon_r}$

One worked example follows, with numbers.

HOT Example 26

An air-gap parallel-plate capacitor has capacitance 8pF, and is charged from a battery to a potential difference of 200V. The dielectric in the gap is then replaced by a mica slab of dielectric constant $\epsilon_r = 4$. Compute the capacitance, charge, voltage, and stored energy under two scenarios: (a) the battery is connected throughout the replacement; (b) the battery has been disconnected before the replacement.

Solution

Compute the initial quantities first; both cases share them.

The initial capacitance is the value given:

$$C_0 = 8\text{pF} = 8 \times 10^{-12}\text{F}$$

The initial charge:

$$Q_0 = C_0 V_0 = (8 \times 10^{-12}\text{F})(200\text{V}) = 1.6 \times 10^{-9}\text{C} = 1.6\text{nC}$$

The initial stored energy:

$$U_0 = 1/2 C_0 V_0^2 = 1/2 (8 \times 10^{-12}\text{F})(200\text{V})^2 = 1.6 \times 10^{-7}\text{J} = 0.16\mu\text{J}$$

The new capacitance is the same in both cases:

$$C = \epsilon_r C_0 = (4)(8\text{pF}) = 32\text{pF}$$

(a) battery connected:

V is held at $V_0 = 200\text{V}$. The new charge:

$$Q = C V_0 = (32 \times 10^{-12}\text{F})(200\text{V}) = 6.4 \times 10^{-9}\text{C} = 6.4\text{nC}$$

The new stored energy:

$$U = 1/2 C V_0^2 = 1/2 (32 \times 10^{-12}\text{F})(200\text{V})^2 = 6.4 \times 10^{-7}\text{J} = 0.64\mu\text{J}$$

The charge has risen by a factor of $\epsilon_r = 4$; the energy has risen by a factor of $\epsilon_r = 4$.

(b) battery disconnected: Q is held at $Q_0 = 1.6\text{nC}$. The new voltage:

$$V = \frac{Q_0}{C} = \frac{1.6 \times 10^{-9}\text{C}}{32 \times 10^{-12}\text{F}} = 50\text{V}$$

The new stored energy:

$$U = \frac{Q_0^2}{2C} = \frac{(1.6 \times 10^{-9}\text{C})^2}{2(32 \times 10^{-12}\text{F})} = 4 \times 10^{-8}\text{J} = 0.04\mu\text{J}$$

The voltage has dropped by a factor of $\epsilon_r = 4$; the energy has dropped by a factor of $\epsilon_r = 4$.

Both columns of the comparison are now filled with numbers.

(a) ends at 6.4nC, 200V, 0.64μJ;

(b) ends at 1.6nC, 50V, 0.04μJ.

The two final energies (0.64μJ and 0.04μJ) have geometric mean $\sqrt{(0.64)(0.04)}\mu\text{J} = 0.16\mu\text{J}$, which is the initial energy, as expected.

Making Sense of the Answer: The two cases differ by a factor of sixteen in the stored energy. In (a) the capacitor ends up holding sixteen times the energy that (b) ends up holding. The reason is that (a) had a battery doing extra work for the duration of the insertion, while (b) had no such top-up; the battery in (a)

supplied four times as much energy as case (b) had at the start, and a quarter of that supply went into stored field while the rest went into mechanical work on the slab.

Think Like a Physicist: Always begin a battery-connected-vs-disconnected problem by writing down the five-row table for the geometry, even if some of the rows are not asked. Filling the table forces the student to commit to a single column and stops the typical error of mixing rows from both columns. The geometric-mean check on the two final energies costs almost no time and catches arithmetic slips in real time. Examination question that asks for “the energy lost” almost always expect the student to compute the difference between the initial energy and the final energy of the appropriate column, not the difference between the two columns.

Dielectric constant, dielectric strength, water, and temperature

So far the dielectric has been treated as if it were specified by one number, the dielectric constant ϵ_r . Two further properties matter in practice. The first is the maximum field the material can withstand before it breaks down and conducts; this is called the **dielectric strength** of the material, measured in volts per metre. The second is how the dielectric constant changes with temperature; for most polar materials, ϵ_r drops as the temperature rises, sometimes by a substantial fraction.

Dielectric strength is a different quantity from dielectric constant, and the two are not correlated. A high- ϵ_r material may have a low dielectric strength, in which case it cannot hold a large voltage even though it would, in principle, store a lot of charge for the voltage it can hold. A low- ϵ_r material with a high dielectric strength may end up storing more energy than the high- ϵ_r material, because the maximum field, and so the maximum energy density, depends on the square of the field. The following short table gives representative values to fix the orders of magnitude.

Material	Dielectric constant ϵ_r	Dielectric strength
Vacuum	1 (exact)	∞
Dry air	1.0006	3MV/m
Paper	2 to 4	16MV/m
Polyethylene	2.3	50MV/m
Polystyrene	2.6	25MV/m
Mica	5 to 7	100MV/m
Glass	5 to 10	14MV/m
Water (deionised)	80	low (depends on ions)
Titanate ceramic	up to 10000	varies, often modest

Water is a striking entry in this list. Its dielectric constant is enormous, around 80 at room temperature, far higher than any common solid. It would seem to be the perfect dielectric. In practice, water is almost never used as a capacitor dielectric, for three reasons.

- Water dissolves ions readily; even highly purified water has enough dissolved ionic content to conduct measurably, which short-circuits any capacitor.
- The dielectric strength of water depends sharply on the ionic content, and trace contamination drops it by orders of magnitude.
- The dielectric constant of water is strongly temperature-dependent: ϵ_r drops from about 88 at 0°C to about 56 at 100°C, a fall of more than a third over the range of liquid water. A capacitor whose capacitance changed by a third over the operating temperature would be useless for any precision application.

The temperature dependence of polar liquids has a simple physical reason. At higher temperatures, the random thermal motion of the molecules competes with the aligning effect of the applied field. The hotter

the liquid, the more thoroughly thermal motion scrambles the dipole orientation, and the smaller the net polarisation produced by the field. Since the polarisation is what gives a material its ϵ_r above 1, a smaller polarisation means a smaller ϵ_r . Non-polar dielectrics (paper, polyethylene, mica) show much weaker temperature dependence because their ϵ_r comes from induced dipoles rather than the alignment of permanent ones, and induced dipoles are largely insensitive to temperature.

Two examples follow. The first is a quick check on the figure-of-merit for energy storage; the second is a real-world drift problem with a soil moisture sensor in Morogoro.

BINDER Example 27

A capacitor designer must choose between two candidate dielectric materials for a high-energy storage capacitor. Material A has dielectric constant $\epsilon_r = 100$ and dielectric strength 5MV/m. Material B has $\epsilon_r = 5$ and dielectric strength 200MV/m. For a fixed gap and fixed plate area, which material stores more energy at its breakdown limit, and by what factor?

Solution

The maximum energy density in a dielectric capacitor is the energy density of the field, evaluated at the breakdown field:

$$u_{\max} = 1/2 \epsilon_r \epsilon_0 E_{\text{strength}}^2$$

For Material A:

$$u_A = 1/2 (100)(8.854 \times 10^{-12} \text{F/m})(5 \times 10^6 \text{V/m})^2 = 1.107 \times 10^4 \text{J/m}^3$$

For Material B:

$$u_B = 1/2 (5)(8.854 \times 10^{-12} \text{F/m})(2 \times 10^8 \text{V/m})^2 = 8.854 \times 10^5 \text{J/m}^3$$

The ratio of maximum stored energies, for the same gap and area, is:

$$\frac{u_B}{u_A} = \frac{8.854 \times 10^5 \text{J/m}^3}{1.107 \times 10^4 \text{J/m}^3} = 80$$

Material B stores 80 times more energy at the breakdown limit than Material A does, despite Material A having a dielectric constant twenty times larger.

Making Sense of the Answer: The figure of merit for energy storage is not ϵ_r alone but the product $\epsilon_r E_{\text{strength}}^2$. Material B's strength is 40 times larger than Material A's, so its E^2 contribution is $40^2 = 1600$ times larger; Material A's ϵ_r is 20 times larger, which means B is down by a factor of 20 in the ϵ_r contribution. The net advantage for B is $1600/20 = 80$, and that is exactly the factor by which B beats A in energy density. This is why modern high-energy-density capacitors are built around polymer films with modest ϵ_r but very high dielectric strength, rather than around the high- ϵ_r titanate ceramics that dominate signal-coupling applications.

Think Like a Physicist: Whenever a quantity is bounded by something that enters squared in the figure of merit, the bound dominates. Maximum voltage on a capacitor, maximum stress on a beam, maximum current in a resistor: in every case the limit on one factor that appears squared in the answer matters more than a much larger first-power factor that does not face such a limit. The lesson generalises: hunt for the squared variable first when comparing designs.

REAL Example 28

In a soil-moisture monitoring station at Sokoine University of Agriculture (SUA) in Morogoro, a water-dielectric capacitor sensor is calibrated at 25°C, where the dielectric constant of the water in the sensor is 78. During the dry-season midday peak the sensor temperature rises to 45°C, where the dielectric constant of the water has dropped to 71. The sensor electronics convert the measured capacitance directly into a soil moisture reading, with no temperature compensation. Find the percentage error in the moisture reading caused purely by the temperature drift, and state the sign of the error.

Solution

The capacitance of the sensor is proportional to the dielectric constant of the water in it, with all geometric factors held fixed. The fractional change in capacitance, from the calibration temperature to the operating temperature, is:

$$\frac{\Delta C}{C_{25}} = \frac{C_{45} - C_{25}}{C_{25}} = \frac{\epsilon_{r,45} - \epsilon_{r,25}}{\epsilon_{r,25}} = \frac{71 - 78}{78} = \frac{-7}{78} = -0.0897 \text{ (negative)}$$

The percentage error is approximately:

$$\left| \frac{\Delta C}{C_{25}} \right| \times 100\% = 8.97\% \quad (\text{reading low})$$

The capacitance has fallen by about 9 per cent. Since the electronics read off moisture content as proportional to capacitance, the moisture reading falls by the same 9 per cent. The error is **negative**: the sensor reports the soil as drier than it actually is.

Making Sense of the Answer: *A nine per cent error is small enough to miss in a single reading but large enough to skew a season-long irrigation schedule. The error has a definite sign: it always tells the operator the soil is drier than it really is, on a hot afternoon. An operator who responds to such readings by ordering extra irrigation will over-water on hot days and under-water on cool days, a pattern that disrupts both the soil moisture and the irrigation budget. The fix is either a temperature-compensated electronics design, in which the reading is corrected using a separate temperature sensor, or a non-polar dielectric like a polymer film whose ϵ_r changes much less with temperature.*

Think Like a Physicist: *Whenever a transducer relies on a material property that has a temperature dependence, the temperature dependence enters the calibration. A device that reads correctly in the laboratory at one temperature need not read correctly in the field at a different temperature. The systematic, signed nature of the error (always reading low when it gets hotter) is a strong clue to its origin: random errors do not have a definite sign; systematic errors do.*

Capacitor partly immersed in a dielectric liquid

Before introducing any formula, let us first examine an important physical situation involving a dielectric liquid and a charged capacitor.

Imagine a vertical parallel-plate capacitor lowered partially into a tank containing a dielectric liquid. The lower portion of the plates lies below the liquid surface, while the upper portion remains above it. The plates are connected to a voltage source, so a constant potential difference is maintained across them.

One might expect the liquid level between the plates to remain the same as the level in the surrounding tank. Surprisingly, this does not happen. Instead, the liquid climbs upward between the plates and settles at a height h above the external liquid surface.

This rise is not caused by capillary action, but by electrostatic forces. The electric field between the plates polarises the dielectric liquid, and the fringing electric field near the boundary pulls the liquid into the region where the electric field exists. In effect, the capacitor attracts the dielectric into the gap in much the same way that it would pull in a solid dielectric slab. The liquid rises until the upward electric force is exactly balanced by the downward gravitational force of the raised liquid column, establishing equilibrium.

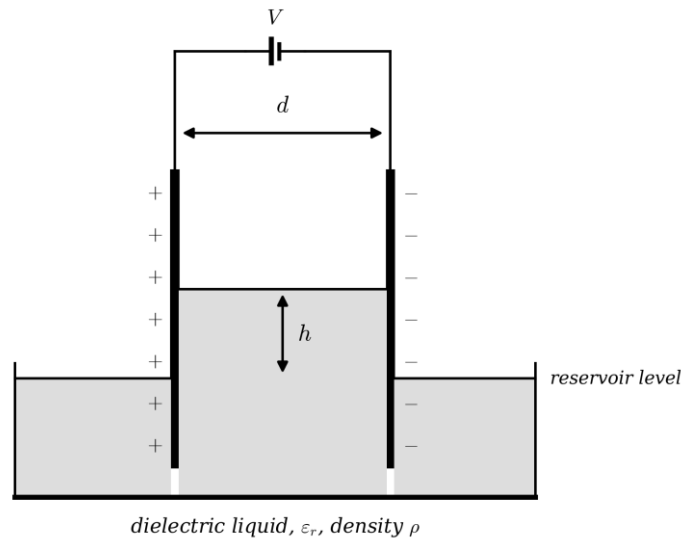


Figure: A parallel-plate capacitor with vertical plates is dipped partway into a tank of dielectric liquid of dielectric constant ϵ_r and density ρ , with the plates connected to a battery of voltage V . The liquid rises inside the gap to a height h above the reservoir level outside the plates, and stays there in equilibrium.

The height h can be found by an energy argument. Treat the gap as two capacitors in parallel, side by side: the dielectric-filled lower portion of height h and the air-filled upper portion of height $(H - h)$, where H is the immersed plate height. With the plates connected to a battery of voltage V , the total capacitance is:

$$C(h) = \frac{\epsilon_0 L}{d} ((H - h) + \epsilon_r h) = \frac{\epsilon_0 L}{d} (H + (\epsilon_r - 1)h)$$

Where L is the plate width and d is the gap. The voltage V is constant; the charge Q on the plates and the stored energy U change with h as the liquid rises.

By the energy method with a connected battery, the upward force on the dielectric liquid as it is drawn into the gap is:

$$F_{\text{up}} = \frac{1}{2} V^2 \frac{dC}{dh} = \frac{1}{2} V^2 \frac{\epsilon_0 L (\epsilon_r - 1)}{d}$$

This force acts on the column of liquid above the reservoir level, lifting it. The lifted column has height h , width L , depth d (the gap), and density ρ , so its weight is:

$$W = \rho L h d g$$

In equilibrium the electric lifting force equals the weight of the lifted column:

$$\frac{1}{2} V^2 \frac{\epsilon_0 L (\epsilon_r - 1)}{d} = \rho L h d g$$

Solving for h , with L cancelling from both sides:

$$h = \frac{(\epsilon_r - 1) \epsilon_0 V^2}{2 \rho g d^2} = \frac{(\epsilon_r - 1) \epsilon_0 E^2}{2 \rho g},$$

where the right-hand form uses $E = V/d$ for the electric field in the gap.

The rise is proportional to $(\epsilon_r - 1)$, to the square of the field, and is inversely proportional to the density and to gravity. It does not depend on the plate width L , which is a good consistency check: a wider capacitor lifts a wider column with the same height. It also does not depend on the depth of immersion below the reservoir, provided the upper edge of the plates is above the rise level so the column has somewhere to go.

HOT Example 29

Two vertical parallel plates of width 5cm and separation 2mm are connected to a 5kV supply and partly immersed in a tank of transformer oil (dielectric constant 2.2, density 880kg/m³). Find the height to which the oil rises inside the gap, above the reservoir level outside the plates. Take $g = 9.81\text{m/s}^2$.

Solution

The field between the plates is the voltage divided by the gap:

$$E = \frac{V}{d} = \frac{5000\text{V}}{2 \times 10^{-3}\text{m}} = 2.5 \times 10^6 \text{V/m}$$

Apply the height formula derived above:

$$h = \frac{(\epsilon_r - 1)\epsilon_0 E^2}{2\rho g} = \frac{(2.2 - 1)(8.854 \times 10^{-12}\text{F/m})(2.5 \times 10^6\text{V/m})^2}{2(880\text{kg/m}^3)(9.81\text{m/s}^2)} = 3.85 \times 10^{-3}\text{m} = 3.85\text{mm}$$

The transformer oil rises about 3.85mm inside the gap, above the level of the surrounding bath.

Making Sense of the Answer: A rise of nearly four millimetres is easily visible to the eye and easily measurable with a ruler held against the outside of the capacitor. The plate width 5cm did not enter the answer, as the derivation showed; only the field, the dielectric constant, the density, and gravity.

Think Like a Physicist: The dimensions of the rise formula are worth checking. The numerator $(\epsilon_r - 1)\epsilon_0 E^2$ has units of energy per volume, since $1/2 \epsilon_0 E^2$ is the energy density of the field. The denominator ρg has units of weight per volume (kilogram per cubic metre times metre per second squared, which is newton per cubic metre, or equivalently joules per metre to the fourth). The ratio J/m^3 divided by $\text{N/m}^3 = \text{J/m}^4$ comes out cleanly in metres, as h must. Carrying units through every algebraic step, as we have done here, is the cheapest insurance against the kind of slip that drops a factor of g and gives an answer in inverse seconds.

Conducting slab inserted in the gap

What happens if the slab inserted into the gap is a conductor rather than a dielectric? A perfect conductor cannot sustain an electric field inside itself; any field would drive its free electrons until they piled up on the surfaces and exactly cancelled the field in the bulk. So a conducting slab of thickness t in a parallel-plate gap has zero field inside it: all the voltage drop between the plates occurs in the air on either side of the slab.

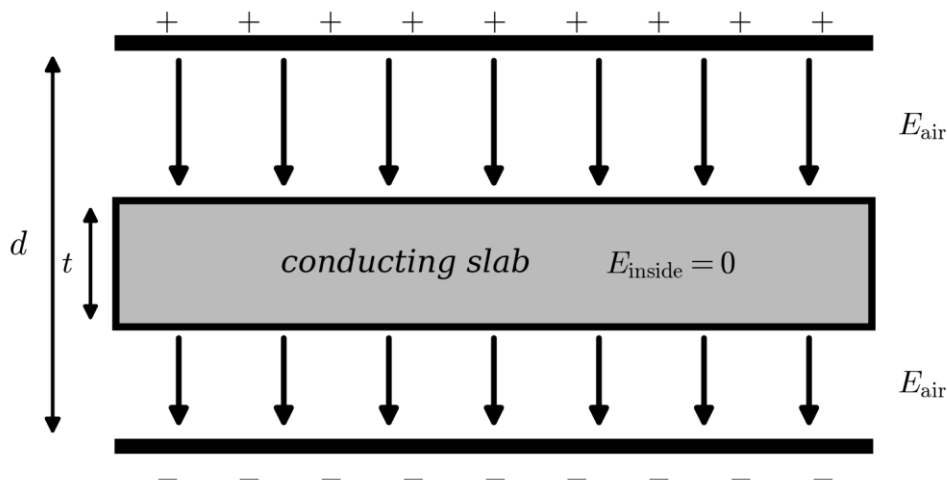


Figure: A conducting slab of thickness t inserted into a parallel-plate gap of width d . The field inside the slab is zero. The field outside the slab, in the two air regions, is the full vacuum value σ/ϵ_0 . The effective electrical gap is $d - t$, independent of where in the gap the slab sits.

The effective electrical gap is therefore $d - t$, regardless of where the slab sits in the gap. The new capacitance is the vacuum capacitance with the reduced gap:

$$C = \frac{\epsilon_0 A}{d - t}$$

A conducting slab can be thought of as the limit of a dielectric slab with $\epsilon_r \rightarrow \infty$. In the partial-fill formula, as ϵ_r becomes very large, the term t/ϵ_r tends to zero, and the denominator reduces to $(d - t)$, recovering the conducting-slab answer. The contrast between a dielectric and a conductor is exact in the limit: a dielectric

reduces the field inside itself by a factor $1/\epsilon_r$; a conductor reduces it by a factor 0. Both raise the capacitance, but the conductor is the more efficient choice geometrically.

In practice a conducting slab is rarely used to fill part of the gap of a capacitor, because the slightest contact between the slab and either plate would short the capacitor. But the conducting-slab result is the cleanest baseline for partial-fill problems and the limiting case that calibrates the dielectric formulas.

HOT Example 30

An air-gap parallel-plate capacitor has plates separated by 5mm and capacitance 100pF. A conducting slab of thickness 3mm is inserted into the gap, without touching either plate. Find the new capacitance. Then find the thickness of a dielectric slab of $\epsilon_r = 4$ that would produce exactly the same new capacitance.

Solution

The original (air-only) capacitance, the gap, and so the product $\epsilon_0 A$, follow from the given:

$$\epsilon_0 A = C_0 d = (100 \times 10^{-12} \text{F})(5 \times 10^{-3} \text{m}) = 5 \times 10^{-13} \text{Fm}$$

After the conducting slab is inserted, the effective gap is reduced to $d - t$:

$$d - t = 5\text{mm} - 3\text{mm} = 2\text{mm} = 2 \times 10^{-3} \text{m}$$

The new capacitance:

$$C = \frac{\epsilon_0 A}{d - t} = \frac{5 \times 10^{-13} \text{Fm}}{2 \times 10^{-3} \text{m}} = 2.5 \times 10^{-10} \text{F} = 250 \text{pF}$$

Or, using the ratio form, $C/C_0 = d/(d - t) = 5/2 = 2.5$, so $C = 2.5C_0 = 250 \text{pF}$.

For the dielectric slab equivalent, set the partial-fill formula equal to 250pF and solve for t_d :

$$\frac{\epsilon_0 A}{(d - t_d) + \frac{t_d}{\epsilon_r}} = 250 \text{pF}$$

The denominator must equal $\epsilon_0 A / (250 \text{pF}) = 2 \text{mm}$:

$$(d - t_d) + \frac{t_d}{4} = 2 \text{mm}$$

With $d = 5 \text{mm}$:

$$5 - t_d + \frac{t_d}{4} = 2 \quad \Rightarrow \quad 5 - \frac{3t_d}{4} = 2 \quad \Rightarrow \quad t_d = 4 \text{mm}$$

A dielectric of $\epsilon_r = 4$ must be 4mm thick to produce the same 250pF as a 3mm conducting slab. The dielectric needs to be thicker to make up for not being a perfect conductor.

Making Sense of the Answer: A 3mm conductor and a 4mm dielectric of $\epsilon_r = 4$ are electrically equivalent for this geometry, both raising the capacitance from 100pF to 250pF. The conductor wins by a millimetre of thickness because it reduces the field inside itself by the factor $1/\infty$ rather than by the factor $1/4$. The factor-of-four dielectric makes up three quarters of the thickness it occupies, electrically speaking, but not all of it; the remaining quarter is “wasted” gap.

Think Like a Physicist: The conducting-slab result is the cleanest test of the partial-fill formula. If a student plugs $\epsilon_r \rightarrow \infty$ into the partial-fill formula and does not recover the conducting-slab answer, the student has the formula wrong. Limiting cases are the cheapest, fastest, and most reliable formula checks available in physics. They cost almost no time and save a great deal of grief when an answer turns out to disagree with intuition.

The dielectric is now thoroughly mapped: where it sits in the gap, how much of it sits there, in what configuration, with what battery condition, and at what temperature, all addressed and tabulated. But the capacitor and its insulators have, so far, all been in steady states. Nothing in this section has asked the question of how the charge gets onto the plates in the first place, or how long it takes. The next section watches the charge accumulate over time, through the resistance of the charging circuit, and exposes the second timescale that decides what a capacitor will actually do in a real device.

CHARGING AND DISCHARGING THROUGH A RESISTOR

The capacitor in the camera flash mentioned at the start of this chapter does its work in two distinct stages. Two small dry-cell batteries cannot deliver the 2000V that a flash bulb needs; what they can do is take their time, half a second or so, to push a small current through a high-resistance charging coil and into a $50\mu\text{F}$ capacitor, which sits and waits with its voltage rising slowly towards the supply value. When the shutter button is pressed, the capacitor is suddenly connected to the flash bulb through a near-zero resistance, and the stored energy comes out of the capacitor in under a millisecond, lighting the bulb so brightly that for one frame of motion the room is brighter than midday.

The capacitor here is acting as a bucket. The batteries fill the bucket slowly, through the narrow neck of the charging resistor. The flash bulb empties the bucket quickly, through the wide outlet of the low-resistance discharge path. The capacitor itself does not care which mode it is in; it simply holds whatever charge has been pushed onto it and lets whatever current the rest of the circuit allows flow into or out of it. What this section does is put numbers on the timing of both processes: *how long it takes to fill the bucket, and how long it takes to empty it.*

Charging through a resistor

The circuit for charging a capacitor through a resistor is shown below. A battery of EMF V drives current through a resistor R and into a capacitor C . Before the switch is closed, the capacitor is uncharged. The instant the switch is closed, current begins to flow, charge begins to build up on the plates, and the voltage across the capacitor begins to rise from zero.

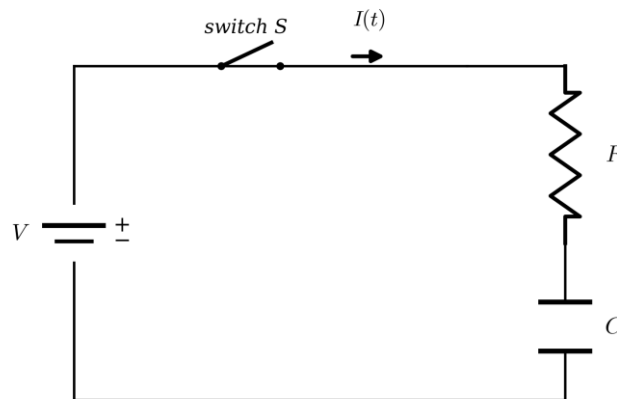


Figure: Charging circuit. A battery of EMF V is connected in series with a switch S , a resistor R , and an initially uncharged capacitor C . When S is closed, current $I(t)$ flows and the capacitor charges towards a final voltage equal to the battery EMF.

To find $Q(t)$, the charge on the capacitor at time t after the switch is closed, apply **Kirchhoff's voltage law** around the loop. The sum of the potential drops around the closed loop must equal the EMF:

$$V = IR + \frac{Q}{C}$$

The first term is the voltage drop across the resistor, the second is the voltage across the capacitor. The current I is the rate at which charge arrives on the capacitor plates: $I = \frac{dQ}{dt}$. Substituting:

$$V = R \frac{dQ}{dt} + \frac{Q}{C}$$

Rearrange to put Q and $\frac{dQ}{dt}$ on one side:

$$R \frac{dQ}{dt} = V - \frac{Q}{C} = \frac{CV - Q}{C}$$

Separate the variables, with Q and dQ on the left and t and dt on the right:

$$\frac{dQ}{CV - Q} = \frac{dt}{RC}$$

Integrate both sides.

$$\begin{aligned} \int_0^Q \frac{dQ}{CV - Q} &= \int_0^t \frac{dt}{RC} \\ [-\ln(CV - Q)]_0^Q &= \left[\frac{t}{RC} \right]_0^t \\ \ln \frac{CV}{CV - Q} &= \frac{t}{RC} \end{aligned}$$

Exponentiating both sides:

$$\frac{CV}{CV - Q} = e^{t/RC} \Rightarrow CV - Q = CVe^{-t/RC}$$

Solving for $Q(t)$:

$$Q = Q(t) = CV(1 - e^{-t/RC})$$

This is the charging law. The charge starts at zero when $t = 0$, rises towards the asymptotic value:

$$Q_{\max} = CV \text{ as } t \rightarrow \infty,$$

and never actually reaches it; the approach is exponential. The current $I(t)$ is the derivative of $Q(t)$:

$$I(t) = \frac{dQ}{dt} = CV \left(\frac{1}{RC} e^{-t/RC} \right) = \frac{V}{R} e^{-t/RC}$$

The current starts at the value $I_0 = \frac{V}{R}$ at $t = 0$, which is the value it would have if the capacitor were not there at all, and decays exponentially to zero as the capacitor fills and starts to oppose further charge flow.

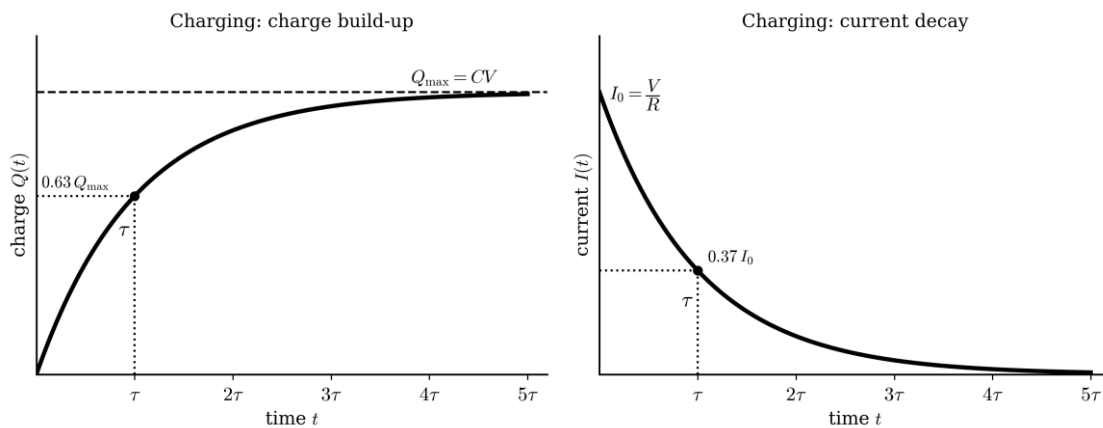


Figure: Charging curves. Left: the charge $Q(t)$ rises from zero towards its asymptotic value $Q_{\max} = CV$, reaching 63% of the asymptote at time $t = \tau$. Right: the current $I(t)$ starts at its maximum $I_0 = V/R$ and decays exponentially, falling to 37% of I_0 at $t = \tau$.

The time constant

The combination RC that appears in both exponentials has units of time and is called the **time constant** of the circuit, written τ :

$$\tau = RC$$

A quick dimensional check confirms that the product of resistance and capacitance is indeed a time:

$$[R] \times [C] = \frac{[V]}{[i]} \times \frac{[Q]}{[V]} = \frac{[Q]}{[i]} = \frac{[I][T]}{[I]} = [T]$$

The time constant is the natural timescale of the circuit. After one time constant, the charge has reached $1 - e^{-1} = 0.632$ of its maximum value, often quoted as 63%. After two time constants, the charge has reached $1 - e^{-2} = 0.865$ of its maximum. After five time constants, the charge has reached 0.993, within a per cent of its asymptotic value; for almost any practical purpose, the capacitor is fully charged after 5τ .

Concisely: **Time constant** of an electrical circuit is the time taken for a changing quantity, such as charge, current, or voltage, to reach about 63% of its final value during growth, or to fall to about 37% of its initial value during decay.

The resistance R does not change the final value $Q_{\max} = CV$ that the capacitor approaches; that is set by the battery EMF and the capacitance alone. What R changes is the speed at which $Q_{\max}$ is approached. A large R gives a long time constant and a slow approach; a small R gives a short time constant and a fast approach. This is the meaning of the bucket metaphor with which this section opened: the bucket size is set by C and V , while the rate of filling is set by R .

BINDER Example 31

A capacitor of capacitance $10\mu\text{F}$ is to be charged from a 12V supply through a series resistor of $1.0\text{M}\Omega$. Find the time constant of the circuit, and the time required for the capacitor to reach 99% of the supply voltage.

Solution

The time constant is the product of resistance and capacitance:

$$\tau = RC = (1.0 \times 10^6 \Omega)(10 \times 10^{-6} \text{F}) = 10\text{s}$$

The charge at time t obeys the charging law:

$$\frac{Q(t)}{CV} = \frac{Q(t)}{Q_{\max}} = 1 - e^{-t/\tau}$$

Set this fraction equal to 0.99 and solve for t :

$$1 - e^{-t/\tau} = 0.99 \Rightarrow e^{-t/\tau} = 0.01$$

Take the natural logarithm of both sides:

$$-\frac{t}{\tau} = \ln(0.01) = -\ln(100) = -4.605$$

So:

$$t = 4.605\tau = 4.605 \times 10\text{s} = 46.05\text{s} \approx 46\text{s}$$

The capacitor reaches 99% of the supply voltage in about 46s.

Making Sense of the Answer: Forty-six seconds is just under five time constants, which is the rule-of-thumb point at which a capacitor is “fully” charged for engineering purposes. Each successive time constant gets the capacitor closer to $Q_{\max}$, but by smaller and smaller increments: 63% after 1τ , 86% after 2τ , 95% after 3τ , 98% after 4τ , 99% after 4.6τ . The exponential approach means that “all the way” is never reached in finite time, but the time to reach any sensible fraction of the way is set by the time constant alone.

Think Like a Physicist: Whenever an exponential approach to a fixed value is asked about, the same algebra applies regardless of the physical setting. The same time-constant arithmetic will return in the next chapter when current grows through an inductor, and again in the chapter on radioactivity when a sample decays away. The trick “set the relevant ratio equal to the target fraction and take a logarithm” is one of the most reusable in physics. The harder skill is recognising which ratio belongs in the exponential: for a quantity that grows towards a limit, it is the deficit (the un-arrived-at remainder); for a quantity that decays towards zero, it is the quantity itself.

HOT Example 32

A capacitor of capacitance $1.0\mu\text{F}$ is connected in series with a resistor of $10\text{M}\Omega$ and a 12.0V battery. Before the switch is closed, the capacitor is uncharged. At time $t = 46\text{s}$ after the switch is closed, find: (a) the fraction of the final charge that the capacitor has acquired; (b) the fraction of the initial current that is still flowing.

Solution

The time constant of the circuit is:

$$\tau = RC = (10 \times 10^6 \Omega)(1.0 \times 10^{-6} \text{F}) = 10 \text{s}$$

At $t = 46 \text{s}$, the dimensionless time variable is:

$$\frac{t}{\tau} = \frac{46 \text{s}}{10 \text{s}} = 4.6$$

The exponential factor at this time:

$$e^{-t/\tau} = e^{-4.6} = 0.01005$$

(a) Fraction of the final charge:

$$\frac{Q(t)}{Q_{\max}} = 1 - e^{-t/\tau} = 1 - 0.01005 = 0.98995 \approx 0.99$$

The capacitor has acquired about 99% of its final charge.

(b) Fraction of the initial current still flowing:

$$\frac{I(t)}{I_0} = e^{-t/\tau} = 0.01005 \approx 0.01$$

The current has fallen to about 1% of its initial value.

Making Sense of the Answer: The two results are mirror images. The charge has come 99% of the way; the current has fallen 99% of the way. Their sum is exactly one, every time, because $(1 - e^{-t/\tau}) + e^{-t/\tau} = 1$ is an identity. As the capacitor fills, the back-voltage it presents to the battery rises towards V and the net driving voltage falls towards zero, taking the current with it. The current and the deficit charge ($Q_{\max} - Q$) track each other exactly.

Think Like a Physicist: The complementary relation between $Q(t)/Q_{\max}$ and $I(t)/I_0$ is more than an arithmetic curiosity. It says that the same exponential factor controls everything in the charging problem: the deficit charge, the instantaneous current, and through them the instantaneous power, the rate of energy storage, and the rate of resistor heating. Once $e^{-t/\tau}$ is computed for a given moment, every time-dependent quantity in the circuit can be read off without doing more exponentials. Spotting a single underlying scalar that controls many quantities is a habit worth developing; it cuts exam-paper arithmetic.

Discharging through a resistor

Once a capacitor has been charged, it can be allowed to discharge through a resistor by disconnecting the battery and closing a switch that puts the capacitor directly across the resistor. The capacitor's own voltage now drives the current, which flows out of the positive plate, through the resistor, and back to the negative plate, gradually neutralising the charge.

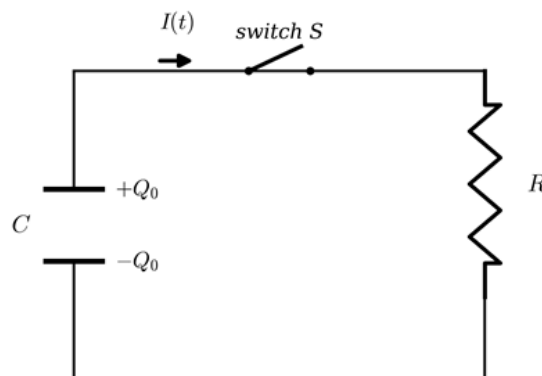


Figure: Discharging circuit. A capacitor C , initially charged to Q_0 , is connected through a switch S across a resistor R . When S is closed, the capacitor's stored charge drives a current $I(t)$ around the loop, decreasing Q to zero as $t \rightarrow \infty$.

Kirchhoff's voltage law for the discharge loop says the voltage across the capacitor equals the voltage drop across the resistor:

$$\frac{Q}{C} = IR$$

The current flows out of the capacitor and so reduces Q , which means $I = -\frac{dQ}{dt}$ (the minus sign because Q is decreasing).

Substituting:

$$\frac{Q}{C} = -R \frac{dQ}{dt}$$

Rearrange and separate variables:

$$\frac{dQ}{Q} = -\frac{dt}{RC}$$

Integrate both sides, from $Q = Q_0$ at $t = 0$ to Q at time t on the left, and from 0 to t on the right:

$$\int_{Q_0}^Q \frac{dQ}{Q} = \int_0^t -\frac{dt}{RC}$$

$$[\ln Q]_{Q_0}^Q = \left[-\frac{t}{RC}\right]_0^t$$

$$\ln \frac{Q}{Q_0} = -\frac{t}{RC}$$

Exponentiating:

$$Q = Q(t) = Q_0 e^{-t/RC}$$

The charge decays exponentially from Q_0 to zero, with the same time constant $\tau = RC$ as the charging case. The current is:

$$I(t) = -\frac{dQ}{dt} = \frac{Q_0}{RC} e^{-t/RC} = I_0 e^{-t/\tau}$$

where $I_0 = \frac{Q_0}{RC} = \frac{V_0}{R}$ is the initial current, the same value one would get by treating the capacitor at $t = 0$ as a battery of voltage $V_0 = \frac{Q_0}{C}$ driving current through R . After one time constant, both the charge and the current have fallen to $1/e \approx 37\%$ of their initial values. After 5τ , both have fallen to less than 1% of their initial values; the capacitor is effectively discharged.

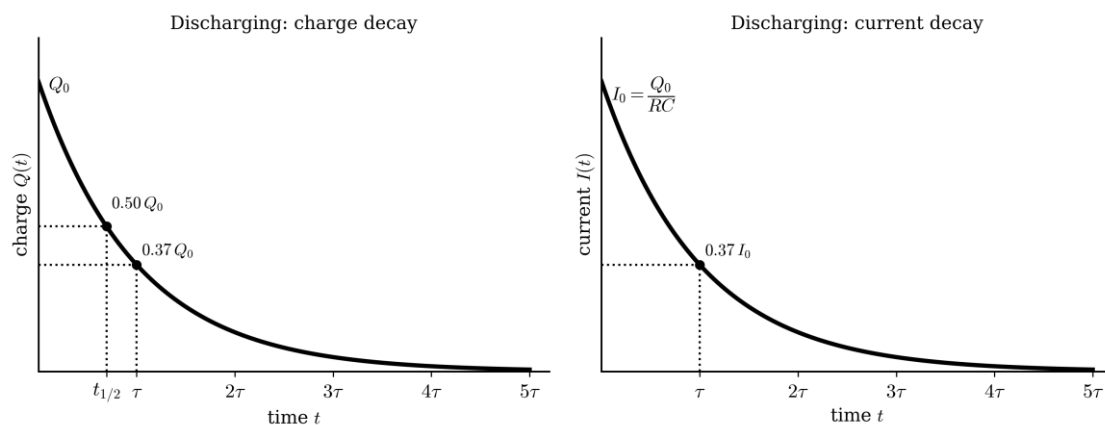


Figure: Discharging curves. Left: the charge $Q(t)$ decays from Q_0 towards zero, falling to half its initial value at $t = t_{1/2} = \tau \ln 2$ and to 37% at $t = \tau$. Right: the current $I(t)$ decays from its initial value $I_0 = Q_0/(RC)$ along the same exponential.

Half-discharge time

A useful single-number summary of how fast a capacitor discharges is the time required for the charge to fall to half its initial value. Set $Q(t) = \frac{Q_0}{2}$ in the discharge law:

$$\frac{Q_0}{2} = Q_0 e^{-t_{1/2}/\tau} \Rightarrow e^{-t_{1/2}/\tau} = \frac{1}{2}$$

Take the natural logarithm:

$$\frac{t_{1/2}}{\tau} = \ln 2 \approx 0.693$$

So:

$$t_{1/2} = \tau \ln 2 = 0.693\tau$$

The half-discharge time is fixed at 0.693 time constants, independent of the initial charge. This is the same exponential half-life that appears for radioactive decay; the underlying mathematics of any exponential decay is the same regardless of whether the decaying quantity is the charge on a capacitor, the number of undecayed nuclei in a radioactive sample, the concentration of a reactant in a first-order chemical reaction, the height of foam on a cup of cappuccino, or the height of froth on a glass of milk gradually collapsing with time.

The same half-time formula applies to charging, by symmetry: the deficit charge ($Q_{\max} - Q$) during charging halves every $\tau \ln 2$ seconds, because the deficit follows an exponential decay law:

$$(Q_{\max} - Q) = Q_{\max} e^{-t/\tau}$$

BINDER Example 33

A capacitor of $50\mu\text{F}$ is initially charged to 200V and then discharged through a $20\text{k}\Omega$ resistor. Find the time constant, the half-discharge time, and the time at which the voltage across the capacitor has fallen to 25V .

Solution

The time constant:

$$\tau = RC = (20 \times 10^3 \Omega)(50 \times 10^{-6} \text{F}) = 1.0\text{s}$$

The half-discharge time:

$$t_{1/2} = \tau \ln 2 = (1.0\text{s})(0.693) = 0.693\text{s}$$

For the voltage to fall to 25V from 200V , note that voltage and charge fall together (since $V = \frac{Q}{C}$ and C is fixed). The required fraction is $\frac{25\text{V}}{200\text{V}} = \frac{1}{8} = 0.125$.

$$\frac{V(t)}{V_0} = e^{-t/\tau} = 0.125$$

Take the natural logarithm:

$$-\frac{t}{\tau} = \ln(0.125) = -2.079$$

So:

$$t = 2.079\tau = (2.079)(1.0\text{s}) = 2.08\text{s}$$

Or by the half-time trick: $200\text{V} \rightarrow 100\text{V} \rightarrow 50\text{V} \rightarrow 25\text{V}$ is three halvings, which takes $3t_{1/2} = 3(0.693\text{s}) = 2.08\text{s}$. Both routes give the same answer.

Making Sense of the Answer: The half-time route is the quicker check because the target voltage is exactly one-eighth of the initial value, which is 2^{-3} . Whenever the target is a power-of-two fraction of the initial value, the answer is an integer number of half-times. For target fractions that are not powers of two, the natural-logarithm route is the general method, but the half-time view still gives a quick estimate: at $t_{1/2}$ we are at 50%, at $2t_{1/2}$ at 25%, at $3t_{1/2}$ at 12.5%, and so on.

Think Like a Physicist: The exponential law gives the discharging capacitor a memoryless quality: from any moment onwards, the time required to fall to half of the present value is always $t_{1/2} = \tau \ln 2$, regardless of how much charge has already drained away. The first half-discharge takes 0.693τ , the second half-discharge from the halfway point also takes 0.693τ , and so on indefinitely. Memoryless decay is a hallmark

of any first-order linear system, and recognising it lets the problem be solved by simple halving arguments instead of by the full exponential formula.

REAL Example 34

A clinical defibrillator in a Dar es Salaam regional hospital uses a $100\mu\text{F}$ capacitor that is charged from a high-voltage supply and then discharged through the chest of a patient between two paddles. The chest tissue between the paddles has an electrical resistance of about $1.0\text{k}\Omega$. For one rescue, the capacitor is charged to 50V before the discharge. Find: (a) the initial discharge current; (b) the time constant of the discharge; (c) the half-discharge time; (d) the total energy delivered to the chest tissue; (e) the rise in temperature of a 1.0g packet of tissue between the paddles, if its specific heat capacity is the same as water, $4180\text{J}/(\text{kgK})$.

Solution

(a) Initial discharge current:

At $t = 0$ the capacitor acts as a battery of voltage $V_0 = 50\text{V}$ across the resistance $R = 1.0\text{k}\Omega$:

$$I_0 = \frac{V_0}{R} = \frac{50\text{V}}{1.0 \times 10^3\Omega} = 0.050\text{A} = 50\text{mA}$$

(b) Time constant:

$$\tau = RC = (1.0 \times 10^3\Omega)(100 \times 10^{-6}\text{F}) = 0.10\text{s} = 100\text{ms}$$

(c) Half-discharge time:

$$t_{1/2} = \tau \ln 2 = (0.10\text{s})(0.693) = 0.0693\text{s} \approx 69\text{ms}$$

(d) Total energy delivered:

All of the initial stored energy on the capacitor is dissipated in the resistor (there is nowhere else for it to go):

$$U_0 = \frac{1}{2} CV_0^2 = \frac{1}{2} (100 \times 10^{-6}\text{F})(50\text{V})^2 = 0.125\text{J}$$

(e) Temperature rise:

The heat absorbed by the 1.0g packet equals $mc\Delta T$, set equal to the energy delivered:

$$\Delta T = \frac{U_0}{mc} = \frac{0.125\text{J}}{(1.0 \times 10^{-3}\text{kg})(4180\text{J}/(\text{kgK}))} = 0.0299\text{K} \approx 30\text{mK}$$

The chest tissue between the paddles rises by about 30mK , or three hundredths of a kelvin, far too small to feel.

Making Sense of the Answer: A temperature rise of three hundredths of a kelvin from a defibrillator-style discharge sounds suspiciously small, and it is. The reason is that the 50V used in this teaching example is far below the voltage of a real defibrillator. Clinical units charge the capacitor to between 1500V and 3000V , where the stored energy is 30^2 to 60^2 times larger, so the energy delivered jumps from 0.125J to around 100J to 400J , in line with the specifications of clinical equipment. The current also rises by the same factor of 30 to 60 , from 50mA here to 1.5 to 3 amperes in the real machine, which is enough to deliver the shock that restarts a heart in fibrillation. The teaching numbers are scaled down purely for arithmetic simplicity; the physical formulas and time scales transfer directly to the real device.

Think Like a Physicist: A defibrillator is a deliberately fast discharge into a deliberately low resistance. The product $RC = 100\text{ms}$ here is short on a human timescale but exactly right for the heart: long enough to deliver substantial energy in one shock, short enough to leave the cardiac tissue unharmed between shocks. Designers of medical equipment do not pick R and C independently; they pick the product to match the biological time constant they need to interact with. Same trick is used in camera flashes (matched to the time the shutter is open), in radar pulse generators (matched to the round-trip travel time to the target), and in switching power supplies (matched to the switching frequency). RC is not an accident; it is a design knob.

Energy bookkeeping during charging

A small but striking question hides inside the charging process. The battery does a known amount of work moving charge around the circuit. The capacitor stores a known amount of energy at the end. The resistor heats up by a known amount during the charging. The question is: *do these three numbers add up correctly*,

and what is the ratio between them? The answer, the same in every charging problem regardless of the value of R , is surprising enough to be worth deriving carefully.

The total work done by the battery from the moment the switch is closed until the capacitor is fully charged is the time integral of the instantaneous power supplied by the battery, $P_{\text{battery}} = VI(t)$:

$$W_{\text{battery}} = \int_0^{\infty} VI(t)dt = V \int_0^{\infty} \frac{V}{R} e^{-t/RC} dt$$

The integral evaluates to RC , giving:

$$W_{\text{battery}} = \frac{V^2}{R} \times RC = CV^2$$

The energy ultimately stored on the capacitor, from the standard formula:

$$U_{\text{cap}} = 1/2 CV^2$$

The difference between battery work and stored energy must have been dissipated in the resistor as heat:

$$W_R = W_{\text{battery}} - U_{\text{cap}} = CV^2 - 1/2 CV^2 = 1/2 CV^2$$

So the battery does CV^2 of work, exactly half of which ends up stored on the capacitor and exactly half of which is dissipated as heat in the resistor.

This 50-50 split is independent of the value of R . A small resistor charges the capacitor quickly with a large current; a large resistor charges slowly with a small current. Either way, exactly half of the battery's energy is dissipated as heat. This is the **half-energy theorem of capacitor charging**: *it is impossible to charge a capacitor from a battery without losing exactly half of the supplied energy as heat in the resistance of the circuit, no matter how small that resistance is.*

The result can be verified by direct integration of the resistor power. The instantaneous power dissipated in the resistor is $P_R(t) = I^2(t)R$, and the total heat is:

$$W_R = \int_0^{\infty} I^2(t)Rdt = R \int_0^{\infty} \left(\frac{V}{R}\right)^2 e^{-2t/RC} dt = \frac{V^2}{R} \int_0^{\infty} e^{-2t/RC} dt$$

The integral evaluates to $\frac{RC}{2}$:

$$W_R = \frac{V^2}{R} \times \frac{RC}{2} = 1/2 CV^2$$

The same $1/2 CV^2$, independent of R as advertised. Even in the limit $R \rightarrow 0$, the current spikes to infinity for an infinitesimal time and the integral still evaluates to exactly $1/2 CV^2$. There is no resistance small enough to avoid the half-energy loss.

HOT Example 35

A $100\mu\text{F}$ capacitor is charged from a 50V battery through a 200Ω series resistor. Compute the total work done by the battery, the energy ultimately stored on the capacitor, and the heat dissipated in the resistor, and verify the half-energy theorem.

Solution

The total work done by the battery:

$$W_{\text{battery}} = CV^2 = (100 \times 10^{-6}\text{F})(50\text{V})^2 = 0.25\text{J}$$

The energy ultimately stored on the capacitor:

$$U_{\text{cap}} = 1/2 CV^2 = 1/2 (0.25\text{J}) = 0.125\text{J}$$

The heat dissipated in the resistor, by the half-energy theorem:

$$W_R = 1/2 CV^2 = 0.125\text{J}$$

Verification by direct integration of $P_R = I^2 R$. The time constant is:

$$\tau = RC = (200\Omega)(100 \times 10^{-6}\text{F}) = 0.020\text{s}$$

The initial current:

$$I_0 = \frac{V}{R} = \frac{50V}{200\Omega} = 0.25A$$

The heat integral:

$$W_R = \int_0^\infty I_0^2 e^{-2t/\tau} R dt = I_0^2 R \times \frac{\tau}{2} = (0.25A)^2 (200\Omega) (0.010s) = 0.125J$$

The direct integration gives the same 0.125J that the half-energy theorem predicts. Energy bookkeeping closes: $W_{\text{battery}} = U_{\text{cap}} + W_R$, with $0.25J = 0.125J + 0.125J$, exactly.

Making Sense of the Answer: The 200Ω resistor heated up to a noticeable degree during the $\approx 100ms$ of charging. The peak power dissipated in the resistor at $t = 0$ was $I_0^2 R = (0.25)^2 (200) = 12.5W$, a substantial wattage for a brief instant. By the time the capacitor was charged, the resistor had absorbed 0.125J of heat. For a typical wire-wound resistor of mass 1g and specific heat capacity similar to copper ($385J/(kg \cdot K)$), this is a temperature rise of $0.125/(10^{-3} \times 385) = 0.32K$, just detectable to the touch.

Think Like a Physicist: The R -independence of the half-energy theorem is the surprising part. Intuition might say: surely a smaller R wastes less energy, because $I^2 R$ would be smaller. The catch is that a smaller R also lets the current $I_0 = V/R$ be larger, by exactly the same factor. The two effects cancel out. Whenever a result turns out to be independent of a parameter that the algebra contained, there is usually a cancellation worth understanding, and the cancellation is often a hint of a deeper conservation principle. In this case, the principle is that work must be done to push charge against an increasing back-voltage, and that work is set by the geometry of the Q - V relationship alone.

Heat dissipated during partial discharge

The last quantity to handle is the heat that comes out of a capacitor during a partial discharge: from time $t = 0$ to time $t = T$, how much energy has been dissipated in the resistor by the discharging capacitor? Once T is known the answer is a short calculation. The energy left on the capacitor at time T is:

$$U(T) = \frac{Q^2(T)}{2C} = \frac{Q_0^2}{2C} e^{-2T/\tau} = U_0 e^{-2T/\tau}$$

The energy dissipated is the difference between the initial and the remaining stored energies:

$$W_R(T) = U_0 - U(T) = U_0 - U_0 e^{-2T/\tau} = U_0 (1 - e^{-2T/\tau}) = 1/2 CV_0^2 (1 - e^{-2T/\tau})$$

Two limit checks confirm the formula. As $T \rightarrow \infty$, the exponential goes to zero and the heat reaches $1/2 CV_0^2$, which is the entire initial stored energy as expected; the capacitor is then fully discharged. As $T \rightarrow 0$, the exponential goes to one and the heat is zero; nothing has discharged yet, so nothing has been dissipated. The formula passes both checks.

HOT Example 36

A capacitor of $200\mu F$ is charged to 100V and then discharged through a $1.0k\Omega$ resistor. Find the heat dissipated in the resistor during the first 100ms of the discharge.

Solution

The time constant:

$$\tau = RC = (1.0 \times 10^3 \Omega)(200 \times 10^{-6} F) = 0.20s$$

The initial stored energy on the capacitor:

$$U_0 = 1/2 CV_0^2 = 1/2 (200 \times 10^{-6} F)(100V)^2 = 1.0J$$

At $T = 100ms = 0.10s$, the exponent in the partial-discharge formula is:

$$\frac{2T}{\tau} = \frac{2(0.10s)}{0.20s} = 1.0$$

The exponential factor:

$$e^{-2T/\tau} = e^{-1} = 0.3679$$

The heat dissipated from $t = 0$ to $t = T$:

$$W_R(T) = U_0(1 - e^{-2T/\tau}) = (1.0\text{J})(1 - 0.3679) = 0.632\text{J}$$

About 0.63J of heat has been dissipated in the resistor during the first 100ms of the discharge, with the remaining 0.37J still stored on the capacitor.

Making Sense of the Answer: In one time constant of charging or discharging, the capacitor moves from 100% to 37% of its initial state in voltage and charge, but it moves from 100% to $37^2\% = 14\%$ of its initial energy, because energy is quadratic in voltage. Equivalently, 63% of the charge has flowed but 86% of the energy has come out. The exponential decay is faster in energy than in charge by a factor of two in the exponent, because $U \propto Q^2$.

Think Like a Physicist: The factor of two in the energy exponent traces back to the squaring in $U = Q^2/(2C)$. Any quantity that depends quadratically on an exponentially-decaying variable will decay exponentially with twice the rate of the original variable. The same idea explains why intensity in optics, power in alternating-current circuits, and kinetic energy in damped motion all decay twice as fast (in exponent) as the underlying field, current, or velocity. Squaring a decaying exponential is a habit-forming move in physics, and the doubling of the rate is the bookkeeping that follows.

The charge dynamics of a capacitor are now complete. We have a formula for the charge as a function of time, for the current at any moment, for the time constant that sets the speed of approach, for the half-time at which the charge falls to half its value, and for the heat dissipated during any window of the charging or discharging process. The next section returns from the time-dependent picture to a steady-state geometric question: *what is the capacitance of an isolated conducting sphere, or of two concentric spheres, or of two coaxial cylinders?* These geometries close the chapter and prepare the ground for Chapter 13, where the capacitor reappears as a reactive element in alternating-current circuits.

CAPACITANCE OF CONDUCTORS

Every capacitor seen so far has had two clearly defined plates with an air or dielectric gap between them. But it is possible to assign a capacitance even to a single isolated conductor, with no second plate anywhere nearby. The trick is to recognise that the “second plate” is, in such cases, the surface of the universe itself, sitting at zero potential infinitely far away. With that convention, any charged conductor has a definite capacitance, set by its geometry and the medium around it. This section works out the capacitance of three important geometries (the isolated sphere, the concentric spherical capacitor, and the coaxial cable) and uses them to settle a few questions about how large or small a real capacitor might reasonably be.

1. Isolated conducting sphere

Consider a metal sphere of radius R held in vacuum, far from any other conductor, and carrying a total charge Q . From Chapter 10, the potential at the surface of such a sphere, measured against a reference at infinity, is:

$$V = \frac{Q}{4\pi\epsilon_0 R}$$

The potential is the same everywhere on the surface, because the surface is an equipotential of a conductor in equilibrium. Inside the conductor, the potential is the same as at the surface, because there is no electric field inside a conductor in equilibrium; the potential is constant throughout the bulk and on the surface. Outside the conductor, the potential falls as $1/r$, exactly as if all the charge were concentrated at the centre.

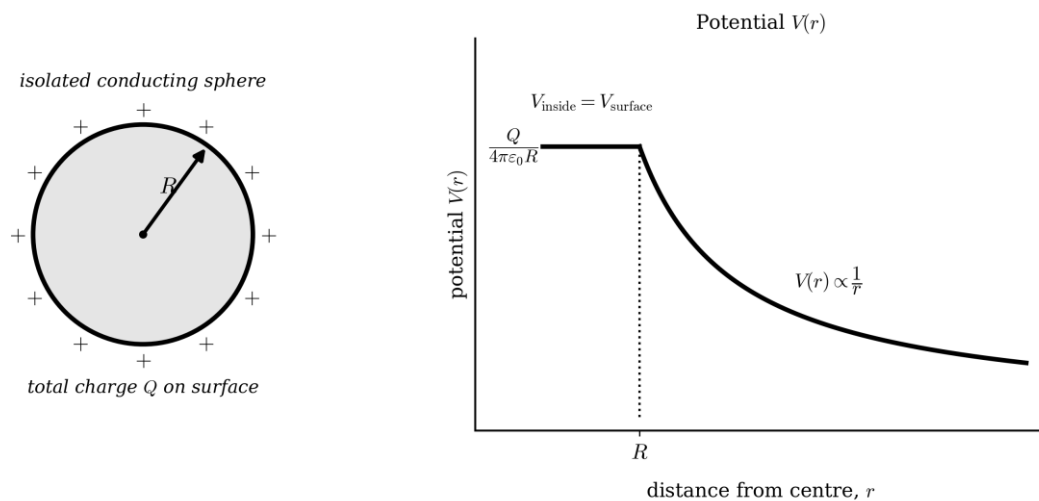


Figure: An isolated conducting sphere of radius R holds total charge Q on its surface. The potential $V(r)$ is constant inside the conductor at the surface value $Q/(4\pi\epsilon_0 R)$, and falls as $1/r$ outside.

By the definition $C = \frac{Q}{V}$, the capacitance of the isolated sphere is:

$$C = \frac{Q}{V} = \frac{Q}{\frac{Q}{4\pi\epsilon_0 R}} = 4\pi\epsilon_0 R$$

The capacitance depends only on the radius R and on ϵ_0 . It is independent of the charge Q that the sphere happens to be carrying, and independent of any property of the conductor other than its size. A larger sphere can hold more charge at the same potential, and so has larger capacitance, by a factor exactly equal to the ratio of radii.

If the sphere sits in a dielectric medium of constant ϵ_r rather than in vacuum, the surface potential drops by the factor ϵ_r for the same charge, so the capacitance rises by the same factor:

$$C = 4\pi\epsilon_r\epsilon_0 R$$

This is the same multiplicative rule that applied to the parallel-plate geometry: replace vacuum with a dielectric of constant ϵ_r and the capacitance is multiplied by ϵ_r regardless of geometry.

BINDER Example 37

A small isolated metal sphere of radius 12cm is held in vacuum, far from any other conductor. Find its capacitance.

Solution

Apply the isolated-sphere formula with $R = 0.12\text{m}$:

$$C = 4\pi\epsilon_0 R = 4\pi(8.854 \times 10^{-12}\text{F/m})(0.12\text{m}) = 1.335 \times 10^{-11}\text{F} = 13.35\text{pF}$$

The capacitance of the isolated 12cm sphere is 13.35pF.

Making Sense of the Answer: A sphere the size of a small football has a capacitance of only about 13pF. This is enormously less than a $1\mu\text{F}$ electrolytic capacitor the size of a fingernail, which is to say: making capacitance from an isolated charged conductor is wildly inefficient compared with making it from two close-spaced plates with a thin dielectric between them. The reason is geometric. The isolated-sphere “second plate” is at infinity, so the average distance between the plates is the radius itself, twelve centimetres in this case, and capacitance scales inversely with that gap. A parallel-plate capacitor with a $10\mu\text{m}$ dielectric gap shrinks the gap by a factor of ten thousand, gaining four orders of magnitude in capacitance per unit area at the cost of needing a second plate.

Think Like a Physicist: The remarkable fact about the isolated-sphere formula $C = 4\pi\epsilon_0 R$ is that it has no V in it, no Q in it, no current in it; only the geometry and the constant of free space. Capacitance is a property of the conductor system, not of the state it happens to be in. The same lesson applies to all the geometries that follow: the formula always emerges from the geometry alone, with ϵ_0 and ϵ_r as the only

material inputs. When a quantity depends only on geometry, it is robust against the kinds of accidents that can change the state of the system without changing what the system is.

2. Concentric spherical capacitor

Now bring up a second conductor: a thin metallic shell of radius b , concentric with the first sphere, with the first sphere now of radius a , where $b > a$. The inner sphere carries charge $+Q$; the outer shell carries charge $-Q$, drawn from the earth to which it is connected. The region $a < r < b$ between the spheres holds an electric field, and integrating that field from a to b gives the voltage across the device.

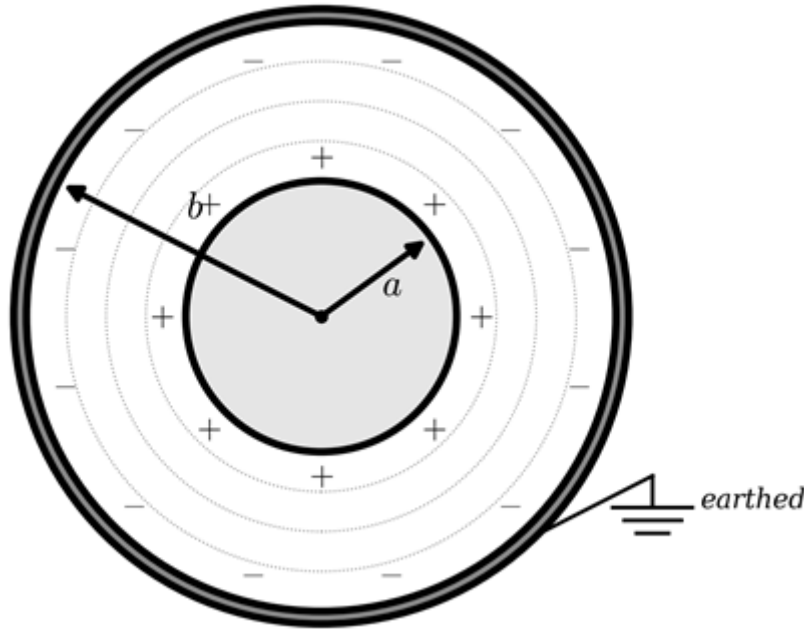


Figure: A concentric spherical capacitor. The inner sphere has radius a and charge $+Q$; the outer shell has inner radius b and is earthed, drawing charge $-Q$ from the ground. The field in the annular gap $a < r < b$ falls as $1/r^2$. Outside the outer shell, the field is zero by Gauss's law.

By Gauss's law applied to a spherical surface of radius r inside the gap, the field strength is:

$$E(r) = \frac{Q}{4\pi\epsilon_0 r^2} \quad \text{for } a < r < b$$

The potential difference between the inner and outer surfaces is the line integral of $E(r)$ from a to b :

$$V = \int_a^b E(r) dr = \frac{Q}{4\pi\epsilon_0} \int_a^b \frac{dr}{r^2} = \frac{Q}{4\pi\epsilon_0} \left[-\frac{1}{r} \right]_a^b = \frac{Q}{4\pi\epsilon_0} \left(\frac{1}{a} - \frac{1}{b} \right)$$

Combine the two fractions:

$$V = \frac{Q}{4\pi\epsilon_0} \left(\frac{b-a}{ab} \right)$$

The capacitance is $\frac{Q}{V}$:

$$C = \frac{Q}{V} = \frac{4\pi\epsilon_0 ab}{b-a}$$

Two limit checks confirm the formula sits between the two geometries we already understand.

Limit one: $b \rightarrow \infty$. As the outer shell is moved out to infinity, the formula reduces to $C \rightarrow 4\pi\epsilon_0 a$, the isolated-sphere result for the inner sphere alone. The outer shell, far enough away, plays no role; the inner sphere “sees” only the universe at infinity.

Limit two: $b \rightarrow a$. As the outer shell is brought very close to the inner sphere, the gap $b - a$ becomes small while the product ab stays roughly a^2 . The formula becomes $C \rightarrow 4\pi\epsilon_0 a^2 / (b - a)$. Identify $4\pi a^2$ as the

surface area of the inner sphere, so this can be written $C \rightarrow \epsilon_0 A / (b - a)$, exactly the parallel-plate formula with area $A = 4\pi a^2$ and gap $d = b - a$. The two-sphere geometry has collapsed to a parallel-plate geometry, as it should once the gap is much smaller than the radii.

Both limits work. The concentric spherical capacitor sits, formally, between the parallel-plate capacitor (gap much smaller than the curvature) and the isolated sphere (no outer plate at all), and it interpolates smoothly between them.

HOT Example 38

Treat the Earth as an isolated conducting sphere of radius $R_E = 6.4 \times 10^6 \text{ m}$ held in vacuum.

- Find the capacitance of the Earth, treated as an isolated sphere.
- Now treat the Earth together with the stratosphere (modelled as a thin conducting layer at altitude $h = 50 \text{ km}$ above the Earth's surface) as a concentric spherical capacitor. Find this capacitance.
- Compare the two values and explain which is larger and why.

Solution

- Apply the isolated-sphere formula with $R = 6.4 \times 10^6 \text{ m}$:

$$C_E = 4\pi\epsilon_0 R_E = 4\pi(8.854 \times 10^{-12} \text{ F/m})(6.4 \times 10^6 \text{ m}) = 712 \times 10^{-6} \text{ F} = 712 \mu\text{F}$$

The Earth, as an isolated sphere, has a capacitance of about $712 \mu\text{F}$.

- Inner radius $a = R_E = 6.4 \times 10^6 \text{ m}$;

Outer radius $b = R_E + h = 6.4 \times 10^6 \text{ m} + 5.0 \times 10^4 \text{ m} = 6.45 \times 10^6 \text{ m}$. The gap is $b - a = 5.0 \times 10^4 \text{ m}$.

Apply the concentric-sphere formula:

$$C_{ES} = \frac{4\pi\epsilon_0 ab}{b - a} = \frac{4\pi(8.854 \times 10^{-12} \text{ F/m})(6.4 \times 10^6 \text{ m})(6.45 \times 10^6 \text{ m})}{5.0 \times 10^4 \text{ m}} = 0.0919 \text{ F} = 91.9 \text{ mF}$$

- The ratio of the two values:

$$\frac{C_{ES}}{C_E} = \frac{91.9 \times 10^{-3} \text{ F}}{712 \times 10^{-6} \text{ F}} = 129$$

The Earth-stratosphere capacitor is about 129 times larger than the isolated-Earth capacitance. The reason follows from the concentric-sphere formula: when the gap $b - a$ is small compared to ab , the capacitor is in the parallel-plate-like regime, and $C \approx \epsilon_0 \cdot 4\pi a^2 / (b - a)$. The stratosphere is only 50 km above the surface, while the Earth's radius is 6400 km, so the gap is just under 1% of the radius. In that thin-gap limit, the concentric formula gives a much larger capacitance than the isolated-sphere formula, by roughly the ratio of the radius to the gap, which is $6400 \text{ km} / 50 \text{ km} = 128$, in excellent agreement with the numerical answer.

Making Sense of the Answer: *The Earth-and-stratosphere system has a capacitance of nearly 0.1 farads, an enormous value by any laboratory standard but unsurprising given the four-million-square-metre plate area. The atmospheric electrical circuit really does behave as a giant leaky capacitor: thunderstorms charge it from below, while a slow leakage current of about 1,500 A worldwide closes the loop through the fair-weather atmosphere. The fact that this circuit can be sized using nothing more than the spherical-capacitor formula is a striking demonstration of how the simple geometries learned in this section apply to systems orders of magnitude larger than any device in the laboratory.*

Think Like a Physicist: *The crossover from the “isolated sphere” formula to the “parallel plate” formula happens whenever the outer shell is brought close enough to the inner sphere that the gap $b - a$ becomes much smaller than the radius a . This is exactly what happens to a beach ball when a metal sheet is brought up close to its surface; the ball, considered as an isolated conductor, has tiny capacitance, but bringing a near-by plate up to it raises the capacitance by the ratio of radius to gap. The lesson is that “isolated conductor” capacitance is a special case of two-conductor capacitance with the second conductor at infinity, and the second conductor matters whenever it is closer than the radius.*

REAL Example 39

Find the radius an isolated conducting sphere in vacuum would need to have a capacitance of exactly 1F. Compare with the radius of the Earth, and state whether such a capacitor would be a practical engineering object.

Solution

Rearrange the isolated-sphere formula to solve for the radius:

$$C = 4\pi\epsilon_0 R \Rightarrow R = \frac{C}{4\pi\epsilon_0}$$

Substitute $C = 1.0\text{F}$:

$$R = \frac{1.0\text{F}}{4\pi(8.854 \times 10^{-12}\text{F/m})} = 8.99 \times 10^9\text{m}$$

The radius needed is $8.99 \times 10^9\text{m}$, or roughly nine billion metres.

Compare with the Earth's radius:

$$\frac{R_{1\text{F}}}{R_E} = \frac{8.99 \times 10^9\text{m}}{6.4 \times 10^6\text{m}} \approx 1400$$

So a 1F isolated sphere would have to be about 1400 times the radius of the Earth. Such an object is not practical by any stretch of the imagination, and the one-farad rating, as a unit of capacitance, is fittingly enormous; capacitances of one farad and above exist in the laboratory only as electrolytic and supercapacitor constructions that achieve the rating through extreme plate area in a small volume, not through the use of any isolated sphere.

Making Sense of the Answer: *The one-farad isolated sphere would not even fit comfortably in the solar system. A sphere of radius $9 \times 10^9\text{m}$ has a diameter of $1.8 \times 10^{10}\text{m}$, larger than the radius of Jupiter's orbit. The size of the farad as a unit of capacitance is therefore set by the inefficiency of the isolated-conductor geometry: a unit defined by $C = Q/V$ with Q in coulombs and V in volts has values, for any reasonable charged object, that are very small fractions of a farad. The microfarad, nanofarad, and picofarad are the units that actually appear on real components, and even the millifarad is on the large side for most circuits.*

Think Like a Physicist: *Whenever a unit comes out to "an impractically large value for any natural object", there is usually a story about which geometries can or cannot achieve the unit efficiently. A 1F capacitance is unreachable in a single isolated sphere of any natural size, but easily reached in an electrolytic capacitor the size of a thumb, because the parallel-plate geometry trades a large gap for a large area at small volume. The farad is large by the standards of isolated geometry but small by the standards of parallel-plate geometry, and the unit makes more sense when the latter is the reference.*

3. Coaxial (cylindrical) capacitor

The third geometry is the coaxial cable: two long concentric cylinders of radii a and b , with $b > a$ and length L , the inner cylinder carrying charge $+Q$ and the outer shell carrying $-Q$. This is the standard form of every transmission line: the inner conductor of a television antenna cable, the signal core of an oscilloscope lead, the centre conductor of any coaxial radio-frequency cable.

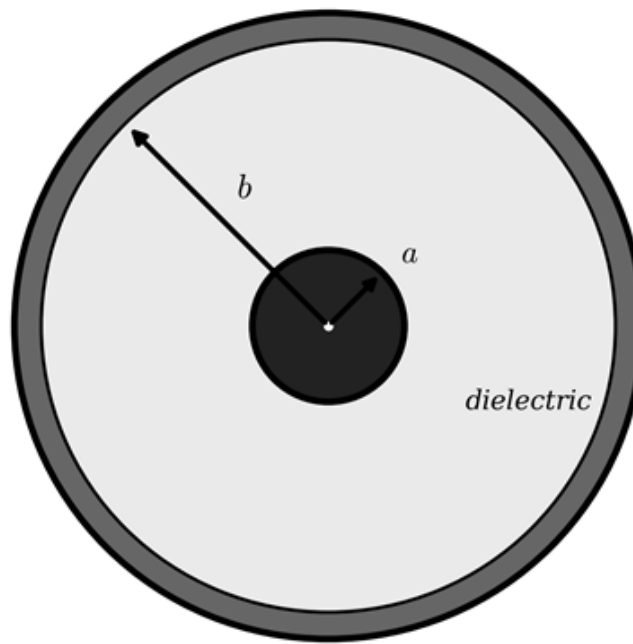


Figure: Cross-section of a coaxial cable. The inner conductor of radius a carries charge $+Q$; the outer conductor of inner radius b carries the equal and opposite charge $-Q$. The dielectric between them, typically air, polyethylene, or foam, supports the field $E(r)$ that runs radially from inner to outer conductor.

By Gauss's law applied to a cylindrical surface of radius r and length L , with the linear charge density $\lambda = \frac{Q}{L}$ on the inner conductor, the field in the gap is:

$$E(r) = \frac{\lambda}{2\pi\epsilon_0 r} \quad \text{for } a < r < b$$

The field falls as $1/r$, more slowly than the spherical $1/r^2$, because a cylinder spreads its flux over a circumference that grows linearly with r , while a sphere spreads its flux over a surface that grows quadratically.

Integrate the field from a to b to find the potential difference:

$$V = \int_a^b E(r) dr = \frac{\lambda}{2\pi\epsilon_0} \int_a^b \frac{dr}{r} = \frac{\lambda}{2\pi\epsilon_0} \ln \frac{b}{a}$$

Substitute $\lambda = \frac{Q}{L}$ to get V in terms of the total charge Q :

$$V = \frac{Q}{2\pi\epsilon_0 L} \ln \frac{b}{a}$$

The capacitance follows from $C = Q/V$:

$$C = \frac{2\pi\epsilon_0 L}{\ln(b/a)}$$

This is the coaxial-cable capacitance formula. If the gap between the conductors is filled with a dielectric of constant ϵ_r , the formula is multiplied by ϵ_r in the usual way:

$$C = \frac{2\pi\epsilon_r\epsilon_0 L}{\ln(b/a)}$$

Two features of the formula deserve comment. First, the capacitance depends on the logarithm of the radius ratio, not on the ratio itself; doubling the outer radius adds only $\ln 2 = 0.69$ to the denominator, modest compared with the strong dependence on length L . Second, the formula gives capacitance proportional to length, so the natural quoted figure for a cable is capacitance per unit length, C/L , in picofarads per metre. A typical RG-58 television cable, with $\epsilon_r \approx 2.3$, runs at about 100pF/m; a 30-metre run from the rooftop antenna to the receiver thus carries 3nF of stored capacitance that the signal has to drive through.

BINDER Example 40

A standard 50-ohm coaxial cable used in laboratory instrumentation has an inner conductor radius of 0.45mm and an outer conductor inner radius of 1.5mm. The gap between them is filled with polyethylene of dielectric constant $\epsilon_r = 2.3$. Find the capacitance per metre of this cable.

Solution

Apply the coaxial formula divided by L:

$$\frac{C}{L} = \frac{2\pi\epsilon_r\epsilon_0}{\ln(b/a)}$$

$$\frac{C}{L} = \frac{2\pi(2.3)(8.854 \times 10^{-12}\text{F/m})}{\ln\left(\frac{1.5\text{mm}}{0.45\text{mm}}\right)} = 1.063 \times 10^{-10}\text{F/m} \approx 106\text{pF/m}$$

The cable has a capacitance per metre of about 106pF/m (in line with the manufacturer's specification of around 100pF/m for cables of this type).

Making Sense of the Answer: A standard television cable from the roof to the set, running about 20m, has a total cable capacitance of $20 \times 106 = 2120\text{pF} \approx 2.1\text{nF}$. This capacitance loads the antenna at high signal frequencies, where the cable's impedance drops to ohms and the antenna sees a heavy load. Cable specifications quote both the capacitance per metre and the characteristic impedance precisely because both numbers control the behaviour of the system at signal frequencies. The 50-ohm impedance and the 100pF/m capacitance are linked: changing one without changing the other would require changing the dielectric constant or the radius ratio, which would in turn change both.

Think Like a Physicist: The coaxial formula's $\ln(b/a)$ in the denominator reflects the radial geometry. In any flux problem where the field source is a line rather than a point, the field falls as $1/r$ rather than $1/r^2$, and the integral of $1/r$ over radius is a logarithm. The logarithm is a slowly varying function: doubling the radius ratio from 3 to 6 raises $\ln(b/a)$ from 1.10 to 1.79, a change of 63%, much less than the doubling of the geometry. Cable designers exploit this insensitivity to keep cable capacitance roughly constant even when the geometry has to flex for mechanical or manufacturing reasons.

The three classical geometries (isolated sphere, concentric spheres, coaxial cylinders) now have their capacitances in closed form. They are the building blocks of every textbook capacitance calculation, and between them they cover the geometric range from “infinite gap” (isolated sphere) through “comparable gap and curvature” (concentric spheres) to “long cylinder with translation symmetry” (coaxial cable). The next section steps back from these analytical geometries to survey the practical applications of the capacitor: where it lives in everyday devices, what role it plays in each, and why the eight applications that close this chapter explain so much of the modern world.

APPLICATIONS OF CAPACITORS

Capacitors appear in almost every electronic device that exists, and in many that pass for non-electronic. The eight applications surveyed below cover the most important roles the capacitor plays in everyday equipment, from the camera flash that opened this chapter to the soil-moisture sensor on a maize plot. Each application is presented in two short paragraphs: what the capacitor is physically doing, and a few numbers or anchors that give the application its scale. None of the applications introduces new physics; everything here follows from the formulas already developed.

1. Flash photography

A camera flash unit charges a 50μF to 200μF capacitor from two AA batteries through a step-up oscillator to several hundred volts, taking about half a second. When the shutter is pressed, the capacitor is dumped through a xenon flash bulb of resistance a few ohms in under a millisecond, releasing all of its stored energy as light. The capacitor is the bucket from the opening of this chapter: filled slowly from a small source, emptied quickly into a low-resistance load.

A typical compact flash designed for 50J of light energy at a peak voltage of 300V requires a capacitor of $C = 2U/V^2 = 2(50)/(300)^2\text{F} \approx 1.1\text{mF}$. The discharge time constant through a 10Ω flash bulb is $\tau = RC \approx 11\text{ms}$, with the light pulse lasting roughly 1ms to 5ms depending on the design. The high-pitched whine heard after pressing the flash button on an older camera is the oscillator stepping the battery voltage up to refill the bucket for the next shot.

2. Defibrillators

A clinical defibrillator stores around 200J of energy on a $100\mu\text{F}$ capacitor charged to about 2kV, and discharges it through paddles pressed against a patient's chest in 5ms to 10ms. The brief, intense current pulse depolarises the entire heart muscle simultaneously, allowing the sinoatrial node to restart a coordinated rhythm. The same capacitor formulas that apply to flash photography apply here, but at much higher voltage and energy.

The chest tissue between the paddles has a resistance of about $1\text{k}\Omega$, so the discharge time constant is $\tau = RC = 100\text{ms}$, intentionally short enough to deliver most of the energy in one decisive pulse and long enough not to be dangerously fast. This is the application worked through quantitatively in Example 34 of the previous section; the present paragraph simply explains why the numbers chosen there matter for the human heart.

3. DC blocking and AC coupling

A capacitor placed in series with a signal path blocks any DC component while passing AC, because a steady voltage produces no current through a capacitor in equilibrium but a varying voltage does. This is the standard way to feed an AC signal from one amplifier stage to the next without disturbing the DC bias of either; the capacitor passes the music or speech, the DC bias stays where the designer set it. The capacitor “couples” the AC signal across without “loading” the DC.

The capacitance is chosen so that the reactance $X_C = 1/(\omega C)$ is small compared with the load resistance at the lowest signal frequency of interest. For audio coupling from a 20Hz to 20kHz amplifier into a $10\text{k}\Omega$ stage, a $1\mu\text{F}$ capacitor gives $X_C = 1/(2\pi(20)(10^{-6})) \approx 8\text{k}\Omega$ at 20Hz, comparable to the load and starting to roll off the bass. Chapter 13 takes up the reactance formula and what it does in alternating-current circuits in full.

4. Power-supply smoothing

A bridge rectifier converts 50Hz mains AC to a pulsating DC output at 100Hz (because both halves of the AC cycle now point the same way). A large capacitor placed in parallel with the load smooths these pulses by discharging slowly through the load between rectifier peaks, holding the load voltage almost constant. The capacitor is the reservoir; the rectifier refills it 100 times per second; the load drains it continuously.

The required capacitance grows with the load current: for a 1A load running off a 230V rectified supply with a ripple of 1% allowed, the capacitor must hold the voltage within 2.3V during the 10ms between peaks, giving $C = I\Delta t/\Delta V = (1)(0.010)/(2.3) \approx 4,400\mu\text{F}$. Large smoothing capacitors of $10,000\mu\text{F}$ at 400V working voltage are standard inside the cabinets of the Tanga substation and every other industrial DC supply in Tanzania.

5. Tuning circuits in AM radios

An inductor L placed in parallel with a variable capacitor C forms a resonant circuit that responds most strongly at the frequency $f_0 = 1/(2\pi\sqrt{LC})$. A radio receiver uses this resonance to pick out one broadcast station from the dozens overlapping at its antenna: turning the tuning knob rotates one set of plates of the variable capacitor partly out of the gap between the fixed plates, changing the effective area A and so changing C , which shifts f_0 to the frequency of a different station.

A typical AM tuning capacitor varies from 20pF to 400pF , paired with an inductor of around $250\mu\text{H}$, giving a tuning range from $f_0 \approx 510\text{kHz}$ to $f_0 \approx 2.3\text{MHz}$. The AM broadcast band that a Radio Free Africa receiver picks up runs from 540kHz to 1610kHz , comfortably inside this range. The same trick is used in FM radios with smaller capacitors and inductors to reach the much higher FM frequencies.

6. Touch screens

Modern phone and tablet screens are tiled with a grid of tiny capacitors, each formed between two transparent conducting layers separated by a thin insulator. When a finger comes close to one of these capacitors, the finger acts as a dielectric of $\epsilon_r \approx 50$ (because the human body is mostly salty water), raising the capacitance of just that one element by a measurable amount. The screen electronics scan the grid many times per second, detect which capacitor has changed, and compute the location of the touch from the position of the affected element.

Typical touch-screen elements have base capacitance around 5pF to 10pF and respond to a finger by changing by 0.5pF to 2pF , well above the noise floor of the detection electronics. A 5mm wide grid on a

10cm screen requires $20 \times 20 = 400$ capacitors and the same number of detection channels; modern phones have grids ten times finer than that. The capacitor formulas of this chapter, applied to elements only a few square millimetres in area, are what make capacitive touch interfaces commercially possible.

7. Electrostatic precipitators

A large industrial smoke duct can be cleaned by passing the flue gases between charged plates that act as the two plates of a capacitor on the kilowatt scale. The dust and ash particles in the gas pick up negative charge from a corona-discharge wire held at 30kV to 80kV, then drift across the duct under the field of the capacitor and stick to a positively charged collector plate. The collected dust is dislodged periodically by mechanical rapping and falls into hoppers below.

A typical precipitator at a cement works at Tanga Cement factory removes about 99.5% of the particulate matter from $100,000\text{m}^3/\text{h}$ of flue gas. The collector plates form a parallel-plate capacitor several metres on a side; the energy stored on the plates is modest, but the field strength of around 5kV/cm is enough to drive even the smallest charged dust grains across the gap in seconds. The same physics is used in coal-fired power stations and, on a much smaller scale, in domestic air purifiers.

8. Soil-moisture sensor for a Tanzanian maize plot

A pair of metal plates, each $5\text{cm} \times 5\text{cm}$ with a 2mm gap, is pushed vertically into the soil of a maize field outside Mbeya. The soil is the dielectric, and its dielectric constant rises sharply with water content: from about $\epsilon_r = 4$ in dry soil to about $\epsilon_r = 18$ in soil at field capacity. The capacitance of the sensor element therefore tracks soil moisture directly, and a small electronic module reads the capacitance and reports it as a percentage on a calibrated dial.

For the sensor described, the capacitance in dry soil is $C = \epsilon_r \epsilon_0 A/d = (4)(8.854 \times 10^{-12})(2.5 \times 10^{-3})/(2 \times 10^{-3}) \approx 44\text{pF}$, rising to about 200pF in fully saturated soil. The agricultural extension officer carries a hand-held unit and calls on each cooperative member once a week during the dry season to take a reading. Earlier in this chapter (Example 28) the temperature dependence of this same sensor was worked through in detail; the sensor is therefore the practical setting for both a capacitor formula and a dielectric-temperature correction.

The dead television in the prep room is now, retroactively, explained. The anode cap inside the picture tube was the storage plate of a high-voltage capacitor, several thousand volts above chassis ground; the glass envelope was the dielectric; and the chassis itself was the other plate. The slow leakage path that would eventually have drained the capacitor was through the (very small) conductivity of the glass and the surrounding air, with a time constant of weeks rather than seconds. This is why a technician shorts the anode cap to the chassis with a screwdriver before opening any television, even one unplugged for days. Disconnected, in this context, really is not the same as discharged.

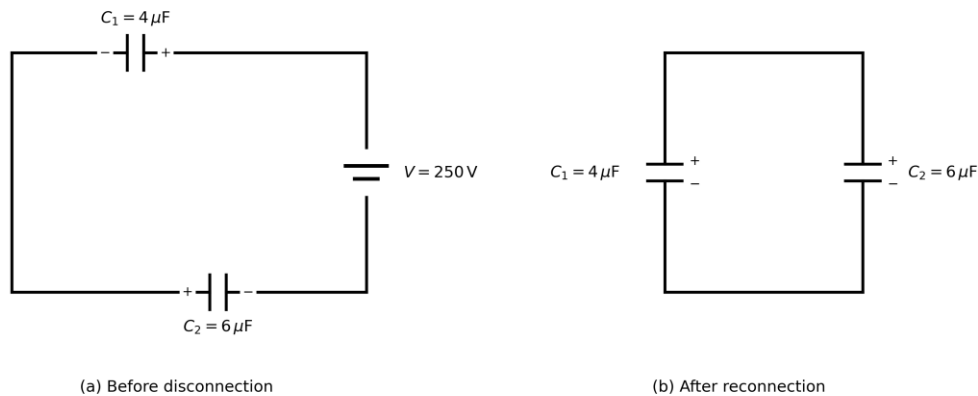
Each of the eight applications above either fills its bucket slowly and empties it quickly (camera flash, defibrillator), or fills it fast and empties it slowly (smoothing capacitor), or holds a moderate charge that responds to small environmental changes (touch screen, soil-moisture sensor). Every one of these behaviours follows from a small set of formulas: the geometric capacitance $C = \epsilon_0 A/d$, the energy storage $U = 1/2 CV^2$, the time constant $\tau = RC$, and the dielectric multiplier $C = \epsilon_r C_0$. The applications differ in geometry and in time scale but not in physics. Two chapters from now, the capacitor will reappear as a circuit element in alternating-current circuits, where the role of resistance is played by the reactance $X_C = 1/(\omega C)$; everything X_C depends on, the reader now knows how to compute.

The chapter body is now complete. What follows is the Miscellaneous Worked Examples section, in which ten additional problems are worked through end to end.

MISCELLANEOUS WORKED EXAMPLES ON CAPACITANCE

Example 41

- State the conditions that must be met for a capacitor to be considered fully charged in a DC circuit, and explain briefly why these conditions are never strictly satisfied in finite time.
- Two capacitors connected as shown in the figure are first charged from the supply and are then disconnected and reconnected with like terminals joined. Calculate the new common potential difference and the new charge on each capacitor.

**Solution**

(a) A capacitor in a DC circuit is fully charged when:

- (i) no current flows in the circuit, and
- (ii) the potential difference across the capacitor equals the EMF of the source.

The charge during charging follows $Q(t) = CV(1 - e^{-t/RC})$, which approaches CV only as $t \rightarrow \infty$. For any finite time t , the factor $e^{-t/RC}$ is strictly positive, so $Q(t)$ is strictly less than CV and a small current is still flowing.

(b) **Series** combination across the 250V supply. The equivalent capacitance is:

$$C_s = \frac{C_1 C_2}{C_1 + C_2} = \frac{(4\mu\text{F})(6\mu\text{F})}{4\mu\text{F} + 6\mu\text{F}} = 2.4\mu\text{F}$$

In a series chain each capacitor carries the same charge, equal to the charge supplied by the battery:

$$Q = C_s V = (2.4\mu\text{F})(250\text{V}) = 600\mu\text{C}$$

Reconnection with **like** terminals joined. After disconnection each capacitor still carries $600\mu\text{C}$, and when the positive plates are joined to one common rail and the negative plates to another, the total charge on the joined positive rail is:

$$Q_{\text{total}} = Q_1 + Q_2 = 600\mu\text{C} + 600\mu\text{C} = 1200\mu\text{C}$$

The capacitors now sit in **parallel**, so the combined capacitance of the parallel pair is:

$$C_p = C_1 + C_2 = 4\mu\text{F} + 6\mu\text{F} = 10\mu\text{F}$$

The common potential difference across the parallel combination is the total charge divided by the combined capacitance:

$$V' = \frac{Q_{\text{total}}}{C_p} = \frac{1200\mu\text{C}}{10\mu\text{F}} = 120\text{V}$$

The new charges on each capacitor are therefore:

$$Q'_1 = C_1 V' = (4\mu\text{F})(120\text{V}) = 480\mu\text{C}$$

$$Q'_2 = C_2 V' = (6\mu\text{F})(120\text{V}) = 720\mu\text{C}$$

The new common potential difference is 120V. The new charges are $480\mu\text{C}$ on the $4\mu\text{F}$ capacitor and $720\mu\text{C}$ on the $6\mu\text{F}$ capacitor.

Example 42

- (a) Commercial paper capacitors are often manufactured by rolling two long strips of metal foil with the paper dielectric folded as a zig-zag between them, rather than laying the foil and the paper out flat. State two reasons why the zig-zag construction is preferred over flat-plate construction.

- (b) Two capacitors of $2.0\mu\text{F}$ and $4.0\mu\text{F}$ are connected to form a potential divider, with the $4.0\mu\text{F}$ capacitor observed to have 5.0V across it. The input supply voltage is 15V . Find the two possible values of energy stored in the $2.0\mu\text{F}$ capacitor, and explain in physical terms why two values are possible.

Solution

(a) (i) A long zig-zag of metal foil packs a much greater effective plate area into a small physical volume than a flat-plate stack of the same outer dimensions, so the capacitance per unit volume is much higher; this is the dominant reason.

(ii) The zig-zag pattern breaks the single long dielectric interface of a flat-plate stack into many shorter segments, so a single pinhole or microscopic defect in the paper is less likely to span the gap between the foils and trigger premature breakdown.

(b) The two possibilities arise because the polarity of the 5.0V across the $4.0\mu\text{F}$ capacitor is not specified relative to the 15V supply. The two possible values of the potential difference across the $2.0\mu\text{F}$ capacitor are:

$$V_{2,(a)} = 15\text{V} - 5\text{V} = 10\text{V} \quad \text{or} \quad V_{2,(b)} = 15\text{V} + 5\text{V} = 20\text{V}$$

The corresponding energies stored in the $2.0\mu\text{F}$ capacitor are:

$$U_{(a)} = \frac{1}{2} CV^2 = \frac{1}{2} (2.0 \times 10^{-6}\text{F})(10\text{V})^2 = 1.0 \times 10^{-4}\text{J} = 100\mu\text{J}$$

$$U_{(b)} = \frac{1}{2} CV^2 = \frac{1}{2} (2.0 \times 10^{-6}\text{F})(20\text{V})^2 = 4.0 \times 10^{-4}\text{J} = 400\mu\text{J}$$

The two possible energy values are $100\mu\text{J}$ and $400\mu\text{J}$.

Physically, if the 5V drop across the $4.0\mu\text{F}$ capacitor opposes the supply, the $2.0\mu\text{F}$ must supply the remaining 10V to make up the loop; if the same 5V aids the supply, the $2.0\mu\text{F}$ must drop the full 20V to satisfy Kirchhoff's voltage rule.

Example 43

- (a) Explain briefly why the half-life of a discharging capacitor depends only on the product RC , and not on the initial charge stored or the voltage that originally charged the capacitor.
- (b) A capacitor of capacitance $8.0\mu\text{F}$ is charged through negligible resistance to a potential difference of 18V , and is then discharged through a resistor of resistance $2.5\text{M}\Omega$. Find (i) the time constant of the discharge circuit; (ii) the charge remaining on the capacitor 8.0s after the discharge begins; (iii) the potential difference across the capacitor at the same instant; (iv) the current in the resistor at the same instant.

Solution

(a) The discharge of a capacitor through a resistor follows $Q(t) = Q_0 e^{-t/RC}$, where Q_0 is the initial charge. The half-life is the time at which $Q(t)/Q_0 = 1/2$, which gives $e^{-t_{1/2}/RC} = 1/2$, and hence $t_{1/2} = RC \ln 2$. The initial charge Q_0 appears identically on both sides of the ratio and cancels, and the original charging voltage enters Q_0 only through $Q_0 = CV_0$ and similarly cancels. The half-life therefore depends only on the product RC .

(b) (i) The time constant is the product of resistance and capacitance:

$$\tau = RC = (2.5 \times 10^6\Omega)(8.0 \times 10^{-6}\text{F}) = 20\text{s}$$

(ii) The initial charge is:

$$Q_0 = C V_0 = (8.0\mu\text{F})(18\text{V}) = 144\mu\text{C}$$

The charge remaining at $t = 8.0\text{s}$ is:

$$Q(8) = Q_0 e^{-t/\tau} = (144\mu\text{C}) e^{-8.0/20} = (144\mu\text{C})(0.6703) = 96.5\mu\text{C}$$

(iii) The potential difference across the capacitor at the same instant follows from $V = Q/C$:

$$V(8) = \frac{Q(8)}{C} = \frac{96.5\mu\text{C}}{8.0\mu\text{F}} = 12.1\text{V}$$

(iv) The current in the resistor at the same instant follows from Ohm's law applied to the resistor:

$$I(8) = \frac{V(8)}{R} = \frac{12.1V}{2.5 \times 10^6 \Omega} = 4.83 \mu A$$

Example 44

- (a) Show by direct argument that the force per unit area between the plates of a parallel-plate capacitor is numerically equal to the energy density of the field between the plates.
- (b) A parallel-plate capacitor has plate area 0.05m^2 and plate separation 1.0mm , and is charged to a potential difference of 200V . (i) Find the force of attraction between the plates while the capacitor remains connected to the battery. (ii) The plates are then drawn apart slowly, with the battery still connected, until the separation is 2.0mm . Find the new force. (iii) Find the work done by the external agent in increasing the separation.

Solution

(a) The total electric field between the plates of a parallel-plate capacitor has magnitude $E = \sigma/\epsilon_0$, where $\sigma = Q/A$ is the surface charge density on either plate. This total field is the sum of the field $\sigma/(2\epsilon_0)$ contributed by each plate separately, and a plate cannot exert a force on itself, so only the field due to the **other** plate acts on the charge σA of the first. The force per unit area on one plate is therefore the surface charge density times the field due to the other plate alone:

$$\frac{F}{A} = \sigma \times \frac{\sigma}{2\epsilon_0} = \frac{\sigma^2}{2\epsilon_0}$$

Substituting $\sigma = \epsilon_0 E$ recasts this in terms of the electric field:

$$\frac{F}{A} = \frac{(\epsilon_0 E)^2}{2\epsilon_0} = \frac{1}{2} \epsilon_0 E^2$$

The right-hand side is exactly the energy density $u = \frac{1}{2} \epsilon_0 E^2$ of the field between the plates.

Hence $F/A = u$: force per unit area equals energy per unit volume.

- (b) Plate area $A = 0.05\text{m}^2$, initial separation $d_0 = 1.0\text{mm} = 1.0 \times 10^{-3}\text{m}$, final separation $d_f = 2.0\text{mm} = 2.0 \times 10^{-3}\text{m}$, supply voltage $V = 200\text{V}$.

(i) The field while the capacitor sits at the initial separation is:

$$E_0 = \frac{V}{d_0} = \frac{200V}{1.0 \times 10^{-3}\text{m}} = 2.0 \times 10^5 \text{V/m}$$

The force of attraction follows from $F = \frac{1}{2} \epsilon_0 E^2 A$:

$$F_0 = \frac{1}{2} (8.854 \times 10^{-12} \text{F/m}) (2.0 \times 10^5 \text{V/m})^2 (0.05\text{m}^2) = 8.85 \times 10^{-3} \text{N}$$

(ii) With the battery still connected, the voltage is held at 200V . The new field at the doubled separation is:

$$E_f = \frac{V}{d_f} = \frac{200V}{2.0 \times 10^{-3}\text{m}} = 1.0 \times 10^5 \text{V/m}$$

The new force, with E halved, is therefore:

$$F_f = \frac{1}{2} (8.854 \times 10^{-12} \text{F/m}) (1.0 \times 10^5 \text{V/m})^2 (0.05\text{m}^2) = 2.21 \times 10^{-3} \text{N}$$

(iii) Because the battery holds V constant, the attractive force depends on the separation as $F(d) = \frac{\epsilon_0 A V^2}{2 d^2}$, and the external agent must do positive work against this force in pulling the plates apart. The work is:

$$W_{\text{ext}} = \int_{d_0}^{d_f} F(d) dd = \int_{d_0}^{d_f} \frac{\epsilon_0 A V^2}{2 d^2} dd = \frac{\epsilon_0 A V^2}{2} \int_{d_0}^{d_f} \frac{dd}{d^2} = \frac{\epsilon_0 A V^2}{2} \left(\frac{1}{d_0} - \frac{1}{d_f} \right)$$

Substituting the numerical values:

$$W_{\text{ext}} = \frac{(8.854 \times 10^{-12})(0.05)(200)^2}{2} \left(\frac{1}{1.0 \times 10^{-3}} - \frac{1}{2.0 \times 10^{-3}} \right) = 4.43 \times 10^{-6} \text{ J}$$

Example 45

- (a) For a capacitor that must operate at high voltage without dielectric breakdown, explain why the dielectric strength of the material matters more than its dielectric constant. Use this to explain why mica capacitors are preferred over electrolytic capacitors for radio-frequency tuning circuits.
- (b) A parallel-plate capacitor of capacitance 5.0 nF is to be designed to operate at a working voltage of 600 V. The available dielectric is mica with relative permittivity $\epsilon_r = 6$ and dielectric strength $E_b = 100 \text{ MV/m}$. A safety factor of 2 on the field is required. Find (i) the minimum plate separation that prevents breakdown at the working voltage; (ii) the required plate area; (iii) state one practical reason why the design might still fail in service even with the safety factor applied.

Solution

(a) The voltage at which a capacitor breaks down is $V_{\text{max}} = E_b d$, where E_b is the dielectric strength of the filling material and d is the plate separation. A higher relative permittivity ϵ_r increases the capacitance and hence the charge stored at a given voltage, but it does not change V_{max} . For a high-voltage capacitor the limiting design factor is therefore the dielectric strength, not the dielectric constant. Mica has a high dielectric strength of about 100 MV/m and very low dielectric loss at radio frequencies, so it can hold off the peak voltages of an RF resonant circuit without breakdown and without dissipating the signal as heat. An electrolytic capacitor has a low dielectric strength (set by the thin oxide layer) and a high dielectric loss at radio frequencies; the same RF voltage would either break down its oxide or convert most of the signal energy into heat. Mica is therefore preferred.

- (b) (i) With a safety factor of 2 applied to the field, the design field at the working voltage is:

$$E_{\text{design}} = \frac{E_b}{2} = 50 \text{ MV/m} = 5.0 \times 10^7 \text{ V/m}$$

The minimum plate separation is:

$$d_{\text{min}} = \frac{V_{\text{working}}}{E_{\text{design}}} = \frac{600 \text{ V}}{5.0 \times 10^7 \text{ V/m}} = 1.2 \times 10^{-5} \text{ m} = 12 \mu\text{m}$$

- (ii) The required plate area follows from $C = \epsilon_r \epsilon_0 A/d$:

$$A = \frac{C d_{\text{min}}}{\epsilon_r \epsilon_0} = \frac{(5.0 \times 10^{-9} \text{ F})(1.2 \times 10^{-5} \text{ m})}{(6)(8.854 \times 10^{-12} \text{ F/m})} = 1.13 \times 10^{-3} \text{ m}^2$$

(iii) The data-sheet figure of 100 MV/m is the dielectric strength of clean, dry, defect-free mica. In service the mica sheet may carry microscopic voids, edge moisture absorption, or surface contamination, any of which can reduce the local breakdown field well below the data-sheet value and trigger premature failure of the device. A safety factor on the average field does not protect against a local field concentration at a defect that the bulk calculation does not see.

Example 46

- (a) Earthing one plate of a charged parallel-plate capacitor does not change the capacitor's capacitance. Explain why.
- (b) A parallel-plate capacitor of plate area 250 cm^2 with plates 2.0 mm apart is fully charged from a 150 V battery. With the battery still connected, the plates are then drawn apart slowly to a new separation of 4.0 mm. (i) State, with brief justification, how each of the field E , the charge Q on each plate, and the stored energy U changes during the process. (ii) Compute E , Q , and U before and after the separation, and account in one sentence for where the lost capacitor energy went.

Solution

(a) Capacitance depends only on the geometry of the two-conductor system (the plate area A and the gap d) and on the dielectric between the plates. Earthing one plate fixes its potential at zero relative to earth, but it does not change A , d , or the dielectric, so the ratio $C = Q/V$ between the plates is unchanged.

(b) (i) With the battery connected the potential difference V is held fixed at 150 V. The field $E = V/d$ halves as d doubles. The capacitance $C = \epsilon_0 A/d$ halves as d doubles. The charge $Q = CV$ halves because C halves and V is fixed. The stored energy $U = 1/2 CV^2$ halves for the same reason.

(ii) $A = 250\text{cm}^2 = 0.025\text{m}^2$, $d_0 = 2.0\text{mm} = 2.0 \times 10^{-3}\text{m}$, $d_f = 4.0\text{mm} = 4.0 \times 10^{-3}\text{m}$, $V = 150\text{V}$.

Initial configuration:

$$E_0 = \frac{V}{d_0} = \frac{150\text{V}}{2.0 \times 10^{-3}\text{m}} = 7.5 \times 10^4\text{V/m}$$

$$C_0 = \frac{\epsilon_0 A}{d_0} = \frac{(8.854 \times 10^{-12}\text{F/m})(0.025\text{m}^2)}{2.0 \times 10^{-3}\text{m}} = 1.107 \times 10^{-10}\text{F} = 110.7\text{pF}$$

$$Q_0 = C_0 V = (110.7\text{pF})(150\text{V}) = 16.6\text{nC}$$

$$U_0 = \frac{1}{2} C_0 V^2 = \frac{1}{2} (110.7 \times 10^{-12}\text{F})(150\text{V})^2 = 1.25\mu\text{J}$$

Final configuration, with $d_f = 2d_0$:

$$E_f = \frac{V}{d_f} = 3.75 \times 10^4\text{V/m}, \quad C_f = \frac{C_0}{2} = 55.4\text{pF}$$

$$Q_f = C_f V = 8.30\text{nC}, \quad U_f = \frac{1}{2} C_f V^2 = 0.623\mu\text{J}$$

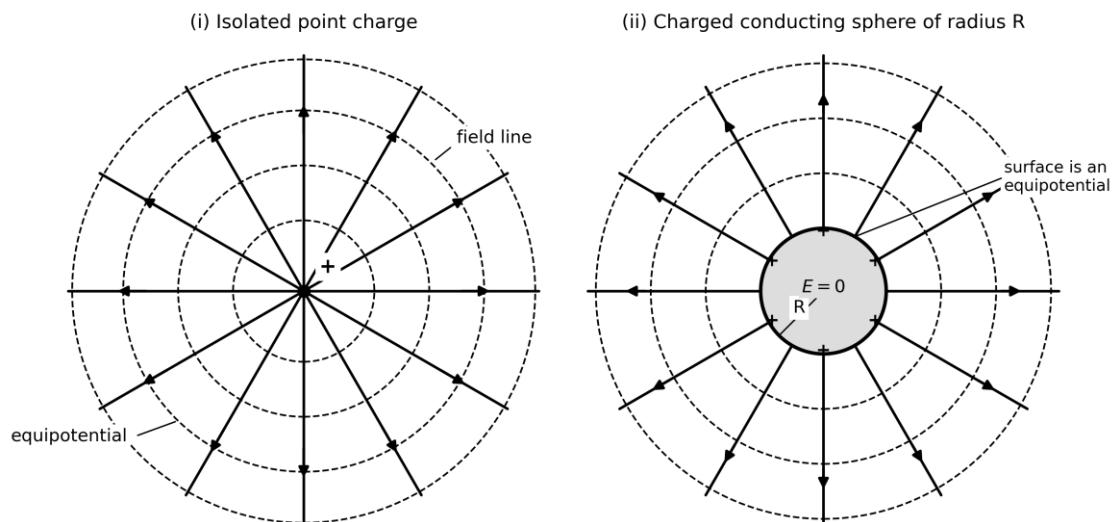
The capacitor returned exactly half of its initial stored energy ($0.623\mu\text{J}$) to the battery as charge flowed off the plates, and the external agent did an additional $0.623\mu\text{J}$ of mechanical work in pulling the plates apart against the residual attraction; both contributions ended up in the battery, which therefore gained a net $1.25\mu\text{J}$ in total.

Example 47

- (a) Sketch the electric field lines and the equipotential surfaces for (i) an isolated positive point charge in vacuum, and (ii) a charged conducting sphere of radius R , both inside and outside the sphere. State in each case why the equipotential surfaces are everywhere perpendicular to the field lines.
- (b) A parallel-plate capacitor has plate area 4m^2 and the gap between the plates is completely filled by three layers of dielectric material stacked perpendicular to the plates. The three layers have relative permittivities 3, 6, and 9, and thicknesses 1.0mm, 3.0mm, and 0.6mm respectively. Calculate the capacitance of the capacitor.

Solution

- (a) The sketches are shown in the figure below.



The equipotential surfaces are everywhere perpendicular to the field lines because V is constant on an equipotential surface; any small displacement $d\vec{s}$ along the surface gives $dV = -\vec{E} \cdot d\vec{s} = 0$, so $\vec{E}$ has no component along the surface and must lie wholly perpendicular to it. For the point charge, the equipotentials are concentric spheres centred on the charge, and the field lines are radial straight lines piercing them at right angles. For the conducting sphere, the entire interior is at one potential and the field inside is zero; on the

outside, the field and equipotentials look identical to those of a point charge of the same total charge placed at the centre, with the sphere's own surface as the innermost equipotential.

(b) The three dielectric layers, all with the same plate area A and stacked perpendicular to the plates, act as three capacitors in series. The combined capacitance follows from the parallel-plate formula applied to a stack:

$$C = \frac{\epsilon_0 A}{\frac{d_1}{\epsilon_{r1}} + \frac{d_2}{\epsilon_{r2}} + \frac{d_3}{\epsilon_{r3}}}$$

$$C = \frac{(8.854 \times 10^{-12} \text{F/m})(4\text{m}^2)}{\left(\frac{1.0}{3} + \frac{3.0}{6} + \frac{0.6}{9}\right) \times 10^{-3} \text{m}} = 3.94 \times 10^{-8} \text{F} = 39.4 \text{nF}$$

The capacitance of the three-layer arrangement is 39.4nF.

Example 48

- (a) Liquid dielectrics whose molecules are polar, such as water and propylene carbonate, have relative permittivities that decrease as the temperature is raised. Explain the molecular reason for this behaviour, and contrast it with the behaviour of non-polar dielectrics such as paraffin oil.
- (b) A capacitor of unknown capacitance C is fully charged from a 200V battery. The battery is then disconnected and the capacitor is discharged through a small coil of resistance wire embedded in a thermally insulated block of specific heat capacity $250 \text{J kg}^{-1} \text{K}^{-1}$ and mass 0.10kg. The temperature of the block rises by 0.40K as a result of the discharge. Calculate C .

Solution

(a) A polar liquid owes its high relative permittivity to the alignment of its permanent molecular dipoles with the applied field; the larger the aligned fraction of dipoles, the larger the induced opposing field inside the medium, and the larger the relative permittivity. Thermal motion opposes this alignment: the higher the temperature, the more vigorous the random reorientation of molecules and the smaller the fraction of dipoles that remain aligned at any instant, so ϵ_r falls. A non-polar liquid like paraffin oil has no permanent molecular dipoles; its (much smaller) ϵ_r comes from induced dipoles whose magnitudes depend only on the local field strength and not on the temperature, so ϵ_r is essentially independent of T .

(b) The energy initially stored in the capacitor is delivered entirely to the resistance-wire coil during the discharge, and the coil delivers it as heat to the thermally insulated block. The thermal energy gained by the block is therefore equal to the initial stored energy of the capacitor:

$$\frac{1}{2} C V^2 = m c \Delta T$$

Substituting the given quantities:

$$\frac{1}{2} C (200\text{V})^2 = (0.10\text{kg})(250 \text{J kg}^{-1} \text{K}^{-1})(0.40\text{K})$$

$$C = \frac{2 \times 10\text{J}}{40000\text{V}^2} = 5.0 \times 10^{-4} \text{F} = 500\mu\text{F}$$

The capacitance is 500 μ F.

Example 49

- (a) An air capacitor of negligible self-capacitance is connected across an electrostatic voltmeter and the system is charged to 80V. A glass slab of dielectric constant 5 is then inserted to fill the gap of the capacitor completely while the voltmeter remains connected. The voltmeter reading is observed to fall to 24V. Explain why the voltmeter reading does not fall to the value $(80/5)\text{V} = 16\text{V}$ expected from the dielectric constant alone, and identify the source of the discrepancy.
- (b) Two capacitors of capacitance C_1 and $C_2 = 12\mu\text{F}$ are connected in series and charged through negligible resistance from a battery. The battery is then disconnected and the series combination is discharged through a resistor of resistance $R = 20\text{M}\Omega$. The measured time constant of the discharge is $\tau = 60\text{s}$. Find C_1 .

Solution

(a) The electrostatic voltmeter has its own (small but non-zero) capacitance C_V , which sits in parallel with the air capacitor C_0 , and the initial charge on the system is $Q = (C_0 + C_V)(80V)$, not $C_0(80V)$. When the glass slab is inserted, the air capacitor becomes $5C_0$, but the voltmeter is not affected, so the new parallel combination has capacitance $5C_0 + C_V$. Charge is conserved during insertion, so the new reading is $V' = (C_0 + C_V)(80V)/(5C_0 + C_V)$, which exceeds $80V/5 = 16V$ because the voltmeter contributes a capacitance that the dielectric does not enhance. The discrepancy comes from the parasitic capacitance of the voltmeter itself.

(b) The two capacitors in series behave electrically as a single capacitor of capacitance:

$$C_s = \frac{C_1 C_2}{C_1 + C_2}$$

The time constant for discharging through R is $\tau = R C_s$, so:

$$C_s = \frac{\tau}{R} = \frac{60s}{20 \times 10^6 \Omega} = 3.0 \times 10^{-6} F = 3.0 \mu F$$

Substituting back into the series formula and solving for C_1 :

$$\frac{1}{C_1} = \frac{1}{C_s} - \frac{1}{C_2} = \frac{1}{3.0 \mu F} - \frac{1}{12 \mu F} = \frac{4 - 1}{12 \mu F} = \frac{1}{4.0 \mu F}$$

$$C_1 = 4.0 \mu F$$

The unknown capacitance is $4.0 \mu F$.

Example 50

- (a) Explain how a projected-capacitance touch screen on a mobile phone detects the location of a finger.
- (b) Two identical small conducting spheres, each of mass $5.0g$ and carrying a charge of $60 \mu C$, are suspended from a common point by light insulating threads of equal length $30cm$. In air, the threads hang at 36° to each other. The arrangement is now fully immersed in a transparent insulating liquid of density $800 kg m^{-3}$, and the angle between the threads is observed to remain unchanged. The density of the sphere material is $2400 kg m^{-3}$. Determine the relative permittivity of the liquid.

Solution

(a) A projected-capacitance touch screen is built as a fine grid of tiny capacitor cells, each formed by two transparent electrodes separated by a thin dielectric. When a finger approaches a cell, the finger sits in the cell's stray field and acts locally as a high-permittivity dielectric (the human body is mostly water, ϵ_r of order 50), replacing part of the air (ϵ_r close to 1) above the cell. That cell's capacitance therefore rises by tens of per cent. A controller chip scans every cell hundreds of times a second, identifies the cell whose capacitance has changed, and reports the grid coordinates of that cell as the touch location.

(b) Consider the following figure:

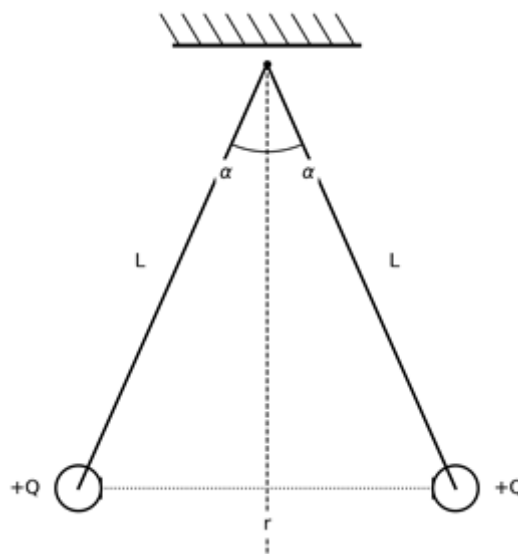


Figure: Geometry of the suspension in air. Each thread of length L makes the half-angle α with the vertical, the spheres are separated by r , and each sphere carries charge $+Q$.

Set the half-angle to α , so that $2\alpha = 36^\circ$ and $\alpha = 18^\circ$. The geometry is the same in air and in the liquid; only the forces acting on each sphere change.

In air, three forces act on each sphere: the tension T along the thread, the weight mg downward, and the horizontal Coulomb repulsion F_C from the other sphere. Resolving the equilibrium horizontally and vertically:

$$T \sin \alpha = F_C, \quad T \cos \alpha = mg \quad \Rightarrow \quad F_C = mg \tan \alpha$$

The Coulomb force in air is given by:

$$F_C = \frac{kQ^2}{r^2} \quad (\text{with } r = 2L \sin \alpha)$$

In the liquid the angle is unchanged, so the separation r between the spheres is unchanged. But two things have changed:

- The Coulomb force is reduced by the relative permittivity of the liquid (from Chapter 10, the force between two charges in a dielectric medium is:

$$F = \frac{kQ_1 Q_2}{\epsilon_r r^2} = \frac{kQ^2}{\epsilon_r r^2} = \frac{F_C}{\epsilon_r}$$

- The effective weight of each sphere is reduced by the buoyant upthrust of the liquid. The buoyant force on each sphere is $V_{\text{sphere}} \rho_L g$, where $V_{\text{sphere}} = m/\rho_s$. The apparent weight in the liquid is therefore:

$$m'g = mg - \frac{m}{\rho_s} \rho_L g = mg \left(1 - \frac{\rho_L}{\rho_s}\right)$$

Equilibrium in the liquid gives:

$$\frac{F_C}{\epsilon_r} = mg \left(1 - \frac{\rho_L}{\rho_s}\right) \tan \alpha$$

Dividing this by the equilibrium equation in air, $F_C = mg \tan \alpha$, makes the unknowns m , Q , L , and α all cancel cleanly, leaving the densities alone:

$$\frac{1}{\epsilon_r} = 1 - \frac{\rho_L}{\rho_s} \quad \Rightarrow \quad \epsilon_r = \frac{\rho_s}{\rho_s - \rho_L}$$

Substituting the given densities:

$$\epsilon_r = \frac{2400 \text{ kg m}^{-3}}{2400 \text{ kg m}^{-3} - 800 \text{ kg m}^{-3}} = \frac{2400}{1600} = 1.50$$

The relative permittivity of the liquid is 1.50.

Notice that the answer is independent of the charge, the mass, the thread length, and the angle of suspension: only the density ratio of the sphere material to the liquid enters.

Ten miscellaneous worked examples now lie behind us. The Digging Deeper Exercise now follows.

DIGGING DEEPER EXERCISE 11

EXERCISE 11A: BINDER QUESTIONS

Question 1

A capacitor stores charge, yet the net charge on a fully charged capacitor is zero. Reconcile these two statements.

Question 2

Identify four factors that determine the capacitance of a parallel-plate capacitor, and indicate in each case whether the capacitance is directly or inversely proportional to that factor.

Question 3

Two capacitors of capacitances C_1 and C_2 are connected first in series and then in parallel across the same battery. In which configuration is the total charge delivered by the battery greater? Justify your answer.

Question 4

On the same pair of axes, sketch the variation of charge Q with time t for (i) a capacitor charging through a resistor from a constant-voltage source, and (ii) the same capacitor discharging through the same resistor. Mark the time constant τ on both curves.

Question 5

Give two reasons why mica is preferred over paper as the dielectric for a high-frequency tuning capacitor.

Question 6

Explain how an electrostatic precipitator cleans smoke from a chimney.

Question 7

Distinguish between the dielectric constant of a material and its dielectric strength. Explain why water is rarely used as a capacitor dielectric in spite of its high dielectric constant.

Question 8

Show by dimensional analysis that the product RC has the units of time.

Question 9

When a capacitor is being charged through a resistor by a constant-voltage source, the current in the resistor decreases steadily with time. Explain why the current does not stop suddenly at the instant the capacitor reaches its maximum charge.

Question 10

In an RC charging circuit, the series resistor R affects the time taken to reach the steady-state charge on the capacitor, but not the value of that steady-state charge. Identify the quantity that does determine the steady-state charge, and explain physically why R does not.

Question 11

Sketch the electric field lines and the equipotential surfaces between the two oppositely charged plates of a parallel-plate capacitor, including the regions near the edges of the plates.

Question 12

When the two plates of a parallel-plate capacitor have unequal areas, do they carry charges of equal magnitude after the capacitor has been connected to a battery? Justify your answer by appealing to conservation of charge.

EXERCISE 11B: REAL QUESTIONS**Question 13**

The owner of an old film camera complains that the flash takes longer to recharge between shots as the batteries weaken. Explain why this is so.

Question 14

A simple washing-machine timer uses an RC circuit to switch off the motor after a fixed interval. If the time constant required is 10s and the resistor available is 100k Ω , estimate the required capacitance.

Question 15

Turning the tuning knob of an old AM radio rotates the moving plates of a variable capacitor partly out of the fixed plates. Explain how this changes the station received.

Question 16

A piezoelectric lighter produces a spark when the trigger is pressed. The spark lasts only a fraction of a second even if the trigger is held down. By treating the piezoelectric element as a small capacitor, explain why the spark is not sustained.

Question 17

Electrolytic capacitors are marked with a polarity and the manufacturer warns the user not to connect them with the polarity reversed. Explain in physical terms what happens if the polarity is reversed.

Question 18

A small torch designed for use in emergency power cuts contains a capacitor in parallel with the bulb, which continues to glow faintly for about 30s after the batteries die. Given that the bulb has resistance 100 Ω and operates at 3V, estimate the capacitance of the torch's capacitor.

Question 19

A parallel-plate soil-moisture sensor placed in a maize plot near Mbeya reads 4nF in dry conditions and 15nF a day after heavy rain. Explain why the reading changes by such a large factor.

Question 20

When the switch of a torch with a capacitor across the bulb is opened, the bulb continues to glow for nearly a second before going dark. Explain this observation.

Question 21

Lightning between a cloud and the ground is often described as the discharge of a giant capacitor. Identify the two plates of this capacitor and the dielectric between them, and explain why the breakdown occurs only when the electric field exceeds approximately $3 \times 10^6 \text{V/m}$ for dry air.

EXERCISE 11C: HOT QUESTIONS**Question 22**

Three capacitors of capacitance 2 μF , 4 μF , and 6 μF are connected first in series and then in parallel across the same 240V DC supply. In each case find the equivalent capacitance, the total charge delivered by the supply, and the total energy stored.

Question 23

Three 4 μF capacitors are arranged in a network with two of them in parallel and the combination in series with the third. The network is connected across a 120V DC supply. Find (i) the equivalent capacitance; (ii) the charge on each capacitor; (iii) the potential difference across each capacitor; (iv) the total energy stored.

Question 24

A $2\mu\text{F}$ capacitor is connected to a 20V supply through a $3\text{M}\Omega$ series resistor and an open switch. Find the time taken after the switch is closed for the energy stored in the capacitor to reach 50 per cent of its maximum final value.

Question 25

A $50\mu\text{F}$ capacitor in series with a $4\text{M}\Omega$ resistor discharges through the resistor. Find the half-life of the discharge, and the number of half-lives needed for the charge to fall below 1 per cent of its initial value.

Question 26

Two square plates of side 10cm separated by 1.0mm carry charges of magnitude 8nC each, with opposite signs. Find the force of attraction between the plates and compare it with the weight of 1g of water.

Question 27

A $2\mu\text{F}$ capacitor carries a charge of 10mC and is discharged through a $5\text{M}\Omega$ resistor. Find (i) the charge remaining on the capacitor after 2s ; (ii) the heat dissipated in the resistor during those 2s .

Question 28

A parallel-plate capacitor of plate area A and gap d has the gap filled by two dielectric slabs of relative permittivities $K_1 = 2$ and $K_2 = 4$, each of thickness $d/2$, stacked perpendicular to the plates. Show that the equivalent capacitance is:

$$C_{\text{eq}} = \frac{2\varepsilon_0 A}{d} \left(\frac{K_1 K_2}{K_1 + K_2} \right)$$

Question 29

A capacitor of capacitance $20\mu\text{F}$ charged to 500V is short-circuited through a coil of resistance 2Ω . Find (i) the peak current immediately after the short-circuit; (ii) the time constant of the discharge; (iii) the total energy dissipated as heat in the coil.

Question 30

Two concentric thin metallic spherical shells have radii 5cm and 10cm . The inner shell carries a charge $+Q$ and the outer shell is earthed. Show that the system behaves as a spherical capacitor, find its capacitance, and compare it with the capacitance of an isolated sphere of radius 5cm .

Question 31

A coaxial cable consists of an inner cylindrical conductor of radius 1mm and an outer cylindrical conductor of radius 5mm , separated by an insulating dielectric of relative permittivity 2. The cable is 100m long. Find its capacitance.

Question 32

A parallel-plate capacitor with air between its plates has a capacitance of 8pF . Find the new capacitance if the distance between the plates is reduced by half and the space between them is filled with a substance of dielectric constant 6.

Question 33

A Van de Graaff generator has its shell electrode held at a potential of $2.5 \times 10^6\text{V}$. The dielectric strength of the gas surrounding the electrode is $5 \times 10^7\text{V/m}$. Find the minimum radius of the spherical shell that will prevent breakdown.

Question 34

A $100\mu\text{F}$ capacitor is charged by a 12V battery and then disconnected. It is then discharged through a thin copper wire of mass 2g and specific heat capacity $380\text{J kg}^{-1}\text{K}^{-1}$. Assuming all the stored energy goes into heating the wire, find the rise in temperature of the wire.

Question 35

A simple parallel-plate soil-moisture sensor for a Tanzanian maize plot has plates of side 5cm and gap 2mm. Calculate (i) the capacitance in air; (ii) the capacitance in dry soil of relative permittivity 4; (iii) the capacitance in saturated soil of relative permittivity 18; (iv) suggest a simple circuit arrangement that converts the capacitance reading into a meter deflection a farmer can read in the field.

Question 36

In an RC discharging circuit, the time taken for the charge on the capacitor to fall to half its initial value is found to be 0.50s. Determine (i) the time constant of the circuit; (ii) the time taken for the charge on the capacitor to fall from 8.0mC to 1.0mC.

Question 37

A parallel-plate capacitor with plates of area 250cm^2 and separation 6.0mm is charged to a potential difference of 240V from a battery, after which the battery is disconnected. A slab of relative permittivity 5 and thickness 4.0mm is then inserted into the gap, leaving the remaining 2.0mm as air. Find (i) the new charge on each plate; (ii) the new potential difference between the plates; (iii) the new energy stored.

Question 38

A capacitor of capacitance $3.0\mu\text{F}$ is charged from one source to a potential difference of 80V, and a separate capacitor of capacitance $5.0\mu\text{F}$ is charged from another source to a potential difference of 200V. Both sources are then disconnected, and the capacitors are wired together with their like-polarity plates joined. Find (i) the energy stored in each capacitor before the connection; (ii) the total energy stored after the connection; (iii) the energy lost in making the connection, and account in one sentence for the loss.

ANSWERS TO DIGGING DEEPER EXERCISES

EXERCISE 10A

1. Electric field is a vector quantity, defined as the force per unit positive charge that a test charge would experience at the point. Forces are themselves vectors; the force a test charge experiences from several source charges is the vector sum of the individual forces, so the field, which is the force per unit charge, also adds as a vector. Electric potential is a scalar quantity, defined as the work done per unit positive charge in bringing a test charge from infinity to the point. Work is a scalar, the contributions to the total work from each source charge depend only on the radial distance from that source and not on direction, and scalar quantities add by ordinary arithmetic. The two quantities are combined by different rules because the underlying physical quantities they record (force and work) are themselves of different kinds, vector and scalar respectively.

2. An electric field is a region of space in which a stationary electric charge experiences a force of electrical origin. The field at a point is a vector quantity whose magnitude is the force per unit positive charge that a small positive test charge would experience at the point, and whose direction is the direction of that force. Operationally, the strength of the field at a point is measured by placing a small positive test charge q at the point, measuring the electrostatic force F on it, and computing $E = F/q$. The test charge is required to be small enough that its own field does not appreciably disturb the source charges responsible for the field being measured.

3. Consider two conducting spheres of different radii joined by a long thin wire, so that they share the same potential. The potential at the surface of each sphere is $V = kQ/R$; for a given V the charge held by a sphere is therefore proportional to its radius, $Q \propto R$. The surface area of a sphere is $4\pi R^2$, so the surface charge density is $\sigma = Q/(4\pi R^2) \propto 1/R$: the smaller the radius, the higher the surface charge density. An irregular conductor in electrostatic equilibrium behaves locally like a small portion of a sphere whose radius equals the radius of curvature at that point. Regions of high curvature (sharp points and edges) have small effective radius and therefore concentrate more charge per unit area than regions of low curvature, which behave like large-radius spheres.

4. In electrostatic equilibrium all excess charge on a conductor resides on its outer surface, and no charge is present in the bulk or in an empty cavity within it. The conducting material that bounds the cavity is at a single potential (the conductor is itself an equipotential), and the cavity is enclosed by a surface on which the potential is uniform. Apply Gauss's law to any closed surface drawn entirely within the empty cavity: the surface encloses no charge, so the total flux through it is zero. Combined with the symmetry imposed by the equipotential boundary, this forces the field at every point of the cavity to be zero. Equivalently, a non-zero field inside the cavity would drive a steady current in the surrounding conductor, in contradiction with the equilibrium condition.

5. Air, like any insulator, can withstand only a limited electric field before its molecules ionise and the air becomes momentarily conducting; this limiting field is the dielectric strength of air, approximately 3.0×10^6 V/m at ordinary conditions. Friction between the soles of the shoes and the dry carpet transfers electrons by the usual triboelectric mechanism, and the person walks away with a net charge that raises their potential to several thousand volts. As the hand approaches the metal doorknob, the gap between the fingertip and the knob shrinks, and the field across the gap rises sharply. When the field exceeds the dielectric strength of the intervening air, the air ionises, the gap becomes momentarily conducting, and the accumulated charge crosses the gap as a visible spark with an audible click. The spark is therefore the very mechanism by which the dielectric strength of air is exceeded and then restored.

6. Gauss's law states that the total electric flux through any closed surface in space is equal to the net charge enclosed by the surface, divided by the permittivity of free space: $\Phi = \oint \mathbf{E} \cdot d\mathbf{A} = \frac{Q_{\text{enc}}}{\epsilon_0}$. The law is particularly useful for charge distributions of high symmetry (spherical, cylindrical, or planar) because for such distributions a Gaussian surface can be chosen on which the field magnitude is constant and either parallel or perpendicular to every element of the surface. The integral on the left then reduces to a simple product $E \times A$ and the field is found by a single line of algebra, without any need to evaluate Coulomb's-law integrals over the source distribution.

7. An equipotential surface is a surface in space all of whose points lie at the same electric potential. Three of its properties are: (i) no work is done by the electric field in moving a test charge from one point of the surface to any other, because the two points are at the same potential and the work done equals the charge times the potential difference; (ii) the electric field at every point of an equipotential surface is perpendicular to the surface, because any component of the field along the surface would do work as the charge moved, contradicting the first property; (iii) no two distinct equipotential surfaces ever intersect, because at the point of intersection the potential would simultaneously have two different values, which is impossible.

8. Argument by potential difference: the work done by the electric field in moving a charge q from a point at potential V_1 to a point at potential V_2 is $W = q(V_1 - V_2)$. On an equipotential surface every point has the same potential by definition, so $V_1 = V_2$, the potential difference is zero, and the work done is zero. Argument by directions: the electric field at every point of an equipotential surface is perpendicular to the surface (otherwise it would have a component along the surface that could do work, contradicting the first argument). The displacement of a charge along the surface lies in the surface itself, so the displacement is perpendicular

to the field. The work done is $W = \vec{F} \cdot \vec{s} = Fs \cos\theta$, and with the angle between force and displacement equal to 90° the cosine is zero, so the work is zero. The two arguments give the same answer because they describe the same physical situation viewed from different angles.

9. The electric field component along a chosen direction equals the negative of the rate of change of the electric potential with distance in that direction. Equivalently, the electric field points in the direction in which the potential decreases most rapidly, and its strength equals the magnitude of that rate of decrease. The negative sign carries a physical meaning: a positive test charge released in an electric field naturally accelerates from regions of higher potential toward regions of lower potential (as a mass rolls downhill on a gravitational slope). For the calculated field to point in this natural direction (from high V to low V), the rate of change dV/dr must be reversed in sign, since dV/dr is itself positive when V increases with distance. The negative sign is therefore the mathematical statement that fields point downhill, from high potential to low.

10. (i) The Earth is treated as the universal reference of zero potential because it is an effectively infinite reservoir of charge: any quantity of charge that flows to or from it produces a change in its own potential too small to measure. Like a vast lake compared with a teacup, the Earth's level does not rise or fall when charge is added or removed in ordinary quantities, so by convention this stable level is taken as the zero of potential, and the potential of any other body is reported relative to it.

(ii) At the moment a charged conductor is earthed (connected to the Earth by a conducting wire), the conductor and the Earth form a single conductor that must reach a single potential. If the conductor was at a positive potential, electrons flow from the Earth to the conductor (so the conventional current flows from conductor to Earth) and continue until the conductor is at the same potential as the Earth, that is, zero. If the conductor was at a negative potential, excess electrons flow from the conductor to the Earth (current flows from Earth to conductor) and continue until the conductor is again at zero potential. In both cases the conductor ends with no net charge.

EXERCISE 10B

11. The thundercloud above carries a large electric charge of one sign and induces a strong electric field in the air below it. A tall pointed object pushes a conducting path upward into this field, and at its sharp top the field is concentrated to values far higher than over the open ground around it. Air ionises and breaks down at a fixed field strength of about $3 \times 10^6 \text{ V/m}$, so it is at the top of the tallest object that this critical field is reached first. The discharge therefore begins at, and travels back down through, that tall pointed conductor. A shorter neighbouring tree never raises the field at its tip to the breakdown level, and is bypassed.

12. Kipanga is right that each hair, brought into rubbing contact with the wool jumper, has gained charge of the same sign as every other hair (one sign of charge has gone onto the hairs, and the opposite sign onto the wool). Like charges on the individual hairs repel one another, and so the hairs splay apart. Mr. Akilikubwa is adding that the body itself has acquired the same sign of charge through the rubbing and behaves as a charged conductor. Electric field lines around such a conductor point radially outward from its surface, and each charged hair, free to swing about its root, aligns itself along these outward field lines. It therefore stands up away from the head, not merely away from the neighbouring hairs.

13. The sharp upper point concentrates the electric field around the conductor to its highest value just above the tip, so this is the spot at which the surrounding air breaks down first. The point therefore takes any lightning strike in preference to the rest of the building, and, more importantly on every other day, it leaks small currents of charge quietly into the air through the same point-discharge mechanism, partially discharging the cloud above before any full strike can develop. The thick copper strap and the buried plate together provide a path of extremely low resistance leading any actual lightning current straight down through the strap to the earth, where it spreads harmlessly into the soil. Without the earth connection, the lightning current would force its way through the building's ordinary structure, with potentially fatal results.

14. A car body is essentially a hollow metal shell, a Faraday cage. When charge is delivered to its outer skin by a lightning strike, that charge moves entirely to the outer surface of the metal almost instantaneously, and the electric field inside the cavity of the shell is zero. The lightning current likewise spreads over the outside of the car and flows down to the ground through the tyres or through the air, never crossing the empty interior. The occupants, sitting in this field-free space, experience no electric force from the strike and are unhurt, provided they are not in contact with the metal frame at the moment the lightning is flowing.

15. The two charged plates set up an almost uniform electric field across the gap, and a charged droplet flying through that field experiences a steady sideways electric force equal to its charge multiplied by the field strength. While the droplet is between the plates it follows a parabolic path, just as a body in a uniform gravitational field would. Doubling the charge on the droplet doubles the sideways force and doubles the sideways deflection by the time the droplet leaves the field; halving the charge halves the deflection. By choosing the charge on each successive droplet, the print head selects exactly where on the page that droplet strikes, painting the shape of each letter one droplet at a time.

16. The very high voltage on the central wire ionises the air close to its surface, producing a stream of free electrons and gas ions. As dust and ash particles in the flue gas drift past this region they collide with the ions, and so each particle acquires a net electric charge of the same sign as that of the central wire. The strong electric field between the central wire and the surrounding earthed plates then exerts a Coulomb force on each charged particle, pulling it sideways across the gas stream and onto a collecting plate. The cleaned gas continues on up the chimney while the collected dust is shaken or rinsed from the plates and removed.

17. As fuel flows from the truck through the hose into the aircraft, friction between the moving fuel and the inner walls of the hose generates electric charge: the fuel acquires charge of one sign while the truck acquires the opposite. If the truck and the aircraft are not connected, a potential difference quickly grows between them as the truck loses charge and the aircraft gains it. Once that potential difference reaches the breakdown field of the air in the gap between the two metal bodies, a spark crosses the gap, and in the petrol-vapour atmosphere of a refuelling site that single spark is enough to start a fire. The bonding cable holds the truck and the aircraft at the same potential throughout the operation, so no potential difference ever develops between them and no spark can occur. The operator's rubber boots cannot perform this function, since they are insulators and would simply leave the two metal bodies free to drift to different potentials.

18. The lightning current entering the ground at the foot of the tree spreads radially outward through the soil, and the resistance of the soil makes the electric potential fall off steeply with distance from the strike point. Two points on the ground a small distance apart can therefore differ in potential by hundreds or thousands of volts during the brief flow of the strike. A person standing with feet a short distance apart has one foot at a higher potential than the other; current then flows from one foot, up one leg, across the body, down the other leg, and into the other foot. The current crossing the heart is what makes this step potential lethal, and explains why sheltering beneath an isolated tall tree during a thunderstorm is dangerous even when one is not struck directly.

19. Both the comb and the metal rod gain charge by friction with the hair just as readily; the question is what happens to that charge once it is on the surface. The plastic comb is an insulator, so any charge deposited on its surface stays in place, because electrons cannot flow through the plastic. The metal rod is a conductor in direct contact with the hand and so with the body of the experimenter, which is itself a fair conductor connected to the ground through the floor. Charge placed on the metal rod therefore flows along the rod, into the hand, and away through the body to the ground as quickly as the rubbing generates it. The rod is left uncharged a fraction of a second after each stroke, with nothing left with which to attract the paper.

20. The Van de Graaff dome at hundreds of kilovolts sets up a very strong electric field in the space around it. When the fluorescent tube is brought into that field, the field penetrates the thin glass and acts on the small amount of mercury vapour and noble gas inside, accelerating any free electrons that happen to be present. The accelerated electrons collide with the gas atoms and excite them, and as the atoms return to their lower energy states they emit ultraviolet light; the phosphor coating on the inside wall of the tube then converts this ultraviolet light into the visible glow that is seen. The glow does not require a current to flow through the tube from any external supply, only a strong electric field reaching into it from a charge in the neighbourhood.

EXERCISE 10C

21. The net field vanishes at the single point $x \approx 1.02\text{m}$, beyond the $-2.0\mu\text{C}$ charge on the side away from the origin.

22. By the cyclic ordering, the two $+2.0\mu\text{C}$ charges occupy diagonally opposite corners of the square, and the two $-3.0\mu\text{C}$ charges occupy the other diagonal pair. The field at the centre due to each charge lies along the diagonal joining that charge to the centre; the two contributions from the $+2.0\mu\text{C}$ pair are therefore equal in magnitude and exactly antiparallel, and cancel, and the same is true of the two contributions from the $-3.0\mu\text{C}$ pair. The net electric field at the centre of the square is therefore, $\mathbf{E} = \mathbf{0}$

23. $7.5 \times 10^4 \text{ N/C}$, 36.9° with the line BA extended past B, lying in the quadrant facing away from both A and C.

24. $E = \frac{4kQ}{3}$

25. (i) $5.65 \times 10^6 \text{ N/C}$ (ii) $56.5\mu\text{C}$ (iii) $6.39 \times 10^6 \text{ Nm}^2/\text{C}$

26. The student's formula $E = kQ/r^2$ treats the whole charge Q as if it were concentrated at the centre of the sphere. This is valid only for points outside the sphere, where the spherically symmetric distribution can be replaced by a single point charge at its centre. For a point inside the sphere, the field at radius r is determined only by the charge that lies inside that radius.

Apply Gauss's law to a spherical Gaussian surface of radius r centred on the sphere's centre. For a uniform volume charge density $\rho = Q/(\frac{4}{3}\pi R^3)$, the charge enclosed within the Gaussian surface is

$$Q_{\text{enc}} = \rho \times \frac{4}{3}\pi r^3 = Q\left(\frac{r}{R}\right)^3$$

so Gauss's law gives

$$E \times 4\pi r^2 = \frac{Q_{\text{enc}}}{\epsilon_0} = \frac{Q}{\epsilon_0} \left(\frac{r}{R}\right)^3$$

$$E = \frac{Qr}{4\pi\epsilon_0 R^3} = \frac{kQr}{R^3}$$

The correct field grows linearly with r from zero at the centre, reaching kQ/R^2 at the surface ($r = R$), where it joins continuously with the external kQ/r^2 result. The student's formula gives the wrong dependence on r , and would diverge as r approaches zero, which is unphysical for a smooth volume distribution.

27. (i) $6.78 \times 10^5 \text{ N/C}$ (ii) 0

28. The total electrostatic potential energy of a system of point charges is the sum over all distinct pairs of $kQ_i Q_j / r_{ij}$. For the three vertices of an equilateral triangle of side L , every pair-separation equals L , and the three distinct pairs are $(Q, +q)$, $(Q, +q)$ and $(+q, +q)$:

$$U_{\text{total}} = \frac{kQq}{L} + \frac{kQq}{L} + \frac{kq^2}{L} = \frac{k}{L}(2Qq + q^2)$$

Setting $U_{\text{total}} = 0$ and dividing through by kq/L (which is non-zero):

$$2Q + q = 0 \Rightarrow Q = -\frac{q}{2}$$

The unknown charge must be negative and half the magnitude of either of the other two, so that the attractive interactions of Q with the two $+q$ charges exactly cancel the repulsive interaction between those two $+q$ charges.

29. The total charge on the inner shell is $Q_1 = \sigma \times 4\pi r^2$ and on the outer shell is $Q_2 = \sigma \times 4\pi R^2$. For a thin spherical shell of charge Q and radius R , the potential at any point inside the shell (including at the centre) is the constant surface value kQ/R . By the superposition of potentials, the potential at the common centre is the sum of the contributions of the two shells, each evaluated at its own radius:

$$V_{\text{centre}} = \frac{kQ_1}{r} + \frac{kQ_2}{R} = \frac{k(\sigma 4\pi r^2)}{r} + \frac{k(\sigma 4\pi R^2)}{R}$$

$$V_{\text{centre}} = 4\pi k\sigma(r + R) = \frac{\sigma(r + R)}{\epsilon_0}$$

The contribution of each shell to the potential at the centre is proportional to its radius, not to its enclosed volume, because the field inside a uniformly charged shell is zero and the potential there is constant at its surface value.

30. (a) Let the displaced position of the test charge be x , with $|x| \ll a$. The $+Q$ at $x = -a$ lies to the left of the test charge at a distance $(a + x)$, and pushes the test charge in the $+x$ direction; the $+Q$ at $x = +a$ lies to the right at a distance $(a - x)$ and pushes the test charge in the $-x$ direction. Taking $+x$ as positive, the net Coulomb force on the test charge is

$$F = \frac{kQq}{(a + x)^2} - \frac{kQq}{(a - x)^2}$$

For small displacement, expand each term to first order in x/a using the binomial series:

$$\frac{1}{(a \pm x)^2} \approx \frac{1}{a^2} \left(1 \mp \frac{2x}{a} \right)$$

so that

$$F \approx \frac{kQq}{a^2} \left(\left(1 - \frac{2x}{a} \right) - \left(1 + \frac{2x}{a} \right) \right) = -\frac{4kQq}{a^3} x$$

The net force is directed opposite to the displacement and proportional to it, which is the condition for simple harmonic motion.

(b) By Newton's second law:

$$m \frac{d^2 x}{dt^2} = -\frac{4kQq}{a^3} x$$

Comparing with the standard SHM form $\frac{d^2 x}{dt^2} = -\omega^2 x$:

$$\omega^2 = \frac{4kQq}{ma^3} \Rightarrow \omega = 2 \sqrt{\frac{kQq}{ma^3}}$$

31. The correct value of the work done is $W = 0$.

The student's error is to apply $W = qV$ as if V were the potential at a single point on the path. The relevant quantity is the potential difference between the start point A and the end point B of the path:

$$W = q(V_A - V_B)$$

For any closed loop, A and B are the same point, so $V_A = V_B$ and the potential difference $V_A - V_B = 0$. The work done is therefore zero, regardless of the value of the equipotential on which the loop happens to lie. The 12V value never enters the calculation.

The same conclusion follows from the conservative nature of the electrostatic force: the line integral of E around any closed loop in an electrostatic field is identically zero.

32. For the drop to remain at rest, the upward electric force on its negative charge must exactly balance its downward weight:

$$qE = mg$$

Between parallel plates the field is uniform, with $E = V/d$, so

$$q = \frac{mgd}{V} = \frac{(1.0 \times 10^{-14} \text{ kg})(9.8 \text{ m/s}^2)(8.0 \times 10^{-3} \text{ m})}{980 \text{ V}} = 8.0 \times 10^{-19} \text{ C}$$

The number of excess electrons is therefore

$$n = \frac{q}{e} = \frac{8.0 \times 10^{-19} \text{ C}}{1.6 \times 10^{-19} \text{ C}} = 5$$

The drop carries 5 excess electrons. The fact that q comes out as an integer multiple of e is the experimental signature that charge is quantised, which is exactly what Millikan was looking for.

33. The small-angle limit gives the centre-to-centre separation d of the two balls at equilibrium in vacuum or air as

$$d^3 = \frac{Q^2 L}{2\pi\epsilon_0 mg}$$

In a non-conducting medium of relative permittivity K the Coulomb force between the balls is reduced by a factor K , which is equivalent to replacing ϵ_0 by $K\epsilon_0$, so

$$d_{\text{oil}}^3 = \frac{Q^2 L}{2\pi K\epsilon_0 mg} = \frac{d_{\text{air}}^3}{K}$$

For small angle θ measured from the vertical, $d = 2L\sin\theta \approx 2L\theta$, so the ratio of separations is the ratio of half-angles:

$$K = \left(\frac{d_{\text{air}}}{d_{\text{oil}}}\right)^3 = \left(\frac{\theta_{\text{air}}}{\theta_{\text{oil}}}\right)^3 = \left(\frac{15^\circ}{9^\circ}\right)^3 = \frac{125}{27} \approx 4.6$$

The relative permittivity of the oil is approximately 4.6.

34. Assumptions: every ash particle reaches terminal drift velocity almost immediately (the time to reach terminal velocity is negligible compared with 2.0s); buoyancy and gravity on the micrometre particle are small compared with the electric driving force; the particle remains a uniform sphere of radius $r = 0.5\mu\text{m}$ throughout its drift; the applied field is uniform between the wire and the plate.

At terminal velocity the electric force balances the Stokes drag:

$$qE = 6\pi\eta rv \Rightarrow v = \frac{qE}{6\pi\eta r}$$

To cross the 15cm gap during the 2.0s residence time, the drift speed must be

$$v_{\text{min}} = \frac{0.15 \text{ m}}{2.0 \text{ s}} = 0.075 \text{ m/s}$$

so the minimum applied field is

$$E_{\text{min}} = \frac{6\pi\eta rv_{\text{min}}}{q} = \frac{6\pi (1.8 \times 10^{-5})(0.5 \times 10^{-6})(0.075)}{(100)(1.6 \times 10^{-19})} \approx 8.0 \times 10^5 \text{ V/m}$$

The required field of about $8 \times 10^5 \text{ V/m}$ is roughly one quarter of the breakdown field of dry air ($3 \times 10^6 \text{ V/m}$), so it lies safely below the breakdown limit. The chosen field is therefore realistic for a precipitator operating in ordinary flue gas conditions.

35. For an isolated spherical conductor of radius R at a potential V above ground, the field at the surface is $E = V/R$. With $R = 0.06\text{m}$ and $V = 350\text{kV}$:

$$E_{\text{surface}} = \frac{V}{R} = \frac{350 \times 10^3 \text{ V}}{0.060 \text{ m}} = 5.83 \times 10^6 \text{ V/m}$$

This is about 1.9 times the dielectric strength of dry air ($3.0 \times 10^6 \text{ V/m}$). The air immediately above the dome will therefore ionise and a continuous corona discharge (or sparking to any nearby earthed object) will drain charge from the dome as fast as the belt delivers it. The dome cannot reach, and cannot hold, the rated 350kV in ordinary dry air.

Either of the following design modifications will allow the rating to be met:

- (i) Increase the radius of the dome. The minimum radius required to keep E_{surface} below the breakdown field is $R_{\text{min}} = V/E_{\text{breakdown}} = 350 \times 10^3 / 3.0 \times 10^6 = 0.117 \text{ m}$, so a dome of diameter at least 24cm (twice the present 12cm) would hold 350kV safely.
- (ii) Enclose the generator in a chamber of higher-dielectric-strength gas (for example sulphur hexafluoride, SF_6 , whose breakdown field at atmospheric pressure is about three times that of air), or operate the generator at elevated pressure, since the breakdown field rises with gas pressure.

EXERCISE 11A

1. On a charged capacitor, the two plates carry charges of equal magnitude but opposite sign: one plate holds $+Q$ and the other plate holds $-Q$, so the net charge on the device as a whole is $+Q + (-Q) = 0$. The phrase “the capacitor stores charge” refers not to a net charge on the device but to the magnitude Q held separately on each plate, since it is the separation of equal and opposite charges across the gap, and not any net charge, that the capacitor is designed to maintain. The energy stored is contained in the electric field between the plates, set up by the equal and opposite surface-charge densities on the two facing plates. The two statements are therefore consistent: the device stores $+Q$ on one plate and $-Q$ on the other, with zero net charge as a whole, but with energy and a field between.

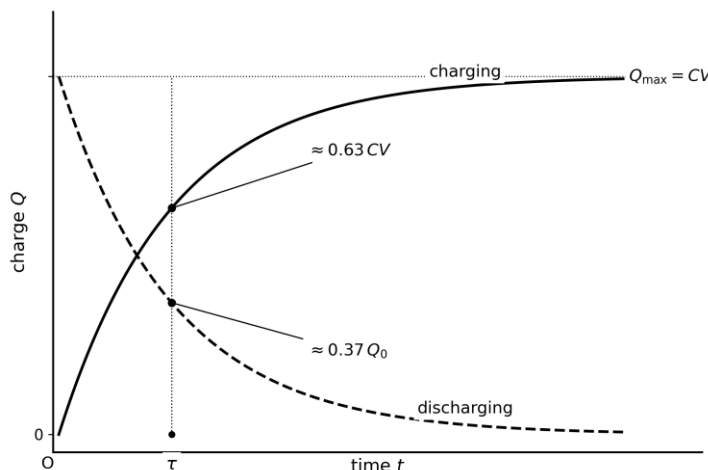
2. The capacitance of a parallel-plate capacitor is $C = \epsilon_r \epsilon_0 A/d$, so four factors determine it:

- (i) the relative permittivity ϵ_r of the dielectric, with C directly proportional to ϵ_r ;
- (ii) the area A of overlap of the two plates, with C directly proportional to A ;
- (iii) the separation d between the plates, with C inversely proportional to d ; and
- (iv) the permittivity of free space ϵ_0 , a universal constant entering as a multiplicative factor.

Note: The capacitance does **not** depend on the charge Q on the plates or the potential difference V across them, because $C = Q/V$ is fixed by the geometry and the dielectric alone and is the same whatever values Q and V take.

3. In the parallel configuration each capacitor sees the full battery voltage V , so the total charge delivered by the battery is $Q_p = (C_1 + C_2)V$. In the series configuration the equivalent capacitance is $C_s = C_1 C_2 / (C_1 + C_2)$, and the battery delivers $Q_s = C_s V$ to the chain. Since $C_1 + C_2 > C_1 C_2 / (C_1 + C_2)$ for any positive C_1 and C_2 , the **parallel** configuration draws the **greater** total charge from the battery at the same voltage.

4. The charging curve follows $Q(t) = CV(1 - e^{-t/\tau})$, rising from zero and approaching the asymptotic value $Q_{\text{max}} = CV$ as t becomes large. The discharging curve follows $Q(t) = Q_0 e^{-t/\tau}$, falling from Q_0 and decaying to zero. Both curves share the same time constant $\tau = RC$. At $t = \tau$ the charging curve has reached approximately 0.63 CV (that is, $CV(1 - 1/e)$), and the discharging curve has fallen to approximately 0.37 Q_0 (that is, Q_0/e).



5. Mica is preferred for two reasons. First, mica has a dielectric strength of about 100MV/m, several times larger than the value of about 16MV/m for impregnated paper, so a mica capacitor can hold off the peak RF voltage in a tuned circuit without breakdown,

whereas a paper capacitor of the same gap would arc over. Second, mica has very low dielectric loss at radio frequencies, while paper's dielectric loss rises sharply with frequency; a paper capacitor in an RF tuning circuit would dissipate a significant fraction of the signal energy as heat, broadening the resonance peak and degrading the Q-factor of the tuned circuit.

6. An electrostatic precipitator is built as a thin central wire held at very high electric potential (typically several tens of kilovolts negative) inside an enclosure of earthed collector plates. The intense electric field close to the surface of the wire ionises the gas locally, producing a stream of free electrons that attach themselves to passing smoke particles and charge each particle up. The strong field between the wire and the collector plates then exerts a Coulomb force on each charged particle, driving it sideways out of the gas stream and onto a plate, where it sticks and is later shaken off mechanically. The cleaned gas continues to the chimney while the collected dust is removed from the plates and disposed of.

7. The dielectric constant, or relative permittivity ϵ_r , is the dimensionless ratio of the capacitance with the given material filling the gap to the capacitance of the same capacitor in vacuum, and measures how strongly the material polarises in the applied field. The dielectric strength is the maximum electric field, in V/m, that the material can sustain before electrical breakdown turns it from an insulator into a conductor.

Pure water has an exceptionally high ϵ_r of about 81 at room temperature, but its dielectric strength is very low because even trace amounts of dissolved ions allow it to conduct electrolytically as soon as a modest field is applied. A water-filled capacitor would therefore self-discharge through ionic conduction long before any useful charge could be stored, so the high ϵ_r is of no practical benefit.

8. From $V = IR$ we have $R = \frac{V}{I}$. From $Q = CV$ we have $C = \frac{Q}{V}$.

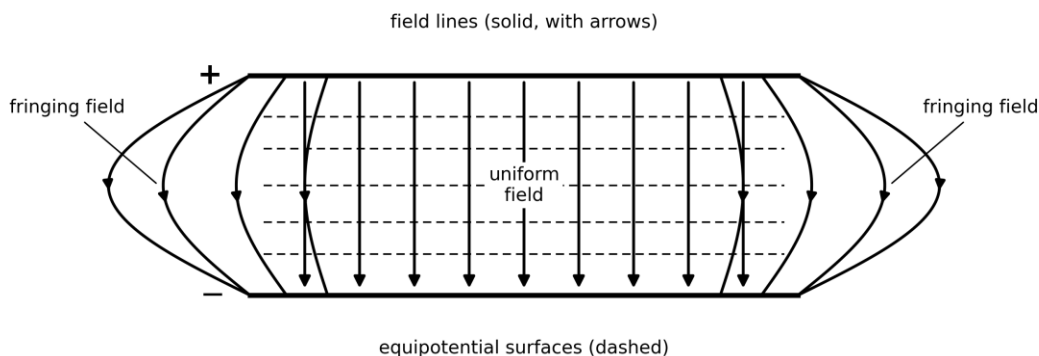
$$[RC] = [R] \times [C] = \frac{[V]}{[I]} \times \frac{[Q]}{[V]} = \frac{[Q]}{[I]} = \frac{[I][T]}{[I]} = [T]$$

Since RC has dimensions of time, it must also have units of time.

9. During charging, the growing potential difference Q/C across the capacitor opposes the EMF V_0 of the source, so the net driving voltage in the loop is $V_0 - Q/C$ and the current in the resistor is $I(t) = (V_0 - Q/C)/R$. As Q approaches its asymptotic maximum CV_0 , the numerator $V_0 - Q/C$ smoothly approaches zero, and so does the current. The current cannot snap off because the approach of Q to CV_0 is exponential and asymptotic, not a sudden hit against a ceiling: the same exponential law that drove the early rapid charging governs the slow tail at the end, and the current dies away smoothly as $e^{-t/RC}$.

10. The steady-state charge is determined by the capacitance C and the EMF V_0 of the source together, as $Q_{\max} = CV_0$. The series resistance R does not appear in this expression because, at the end of charging, the current has fallen to zero and no voltage is dropped across R ; the entire EMF then sits across the capacitor, requiring $Q = CV_0$ regardless of R . What R does control is the time constant $\tau = RC$, which sets how fast $Q_{\max}$ is approached. A larger R simply means a longer time constant; the capacitor still reaches the same $Q_{\max}$ eventually, only more slowly.

11. In the central region between the plates, away from the edges, the field lines run straight from the + plate to the - plate, parallel to one another and uniformly spaced, and the field there is uniform. The equipotential surfaces in this central region are flat planes parallel to the plates, intersecting the field lines at right angles. Near the edges of the plates the field lines bulge outwards beyond the plate boundaries (the fringing field), and the equipotential surfaces curve outwards to follow them, while remaining perpendicular to the field at every point.



12. Yes, the two plates carry charges of equal magnitude after being connected to the battery, despite their unequal areas. Charge is conserved in the isolated capacitor-battery loop: any positive charge that flows from one terminal of the battery onto one plate

must be matched by the same amount of positive charge flowing from the other plate into the other terminal of the battery, so the two plates finish with charges $+Q$ and $-Q$ regardless of the plate areas. The surface charge densities $\sigma_1 = Q/A_1$ and $\sigma_2 = Q/A_2$ differ because the areas A_1 and A_2 differ, with σ larger on the smaller plate; the totals on the two plates are nevertheless equal in magnitude and opposite in sign, as conservation of charge demands.

EXERCISE 11B

13. A camera flash is a capacitor charged through the camera's circuit by the batteries. The total resistance of the charging loop is the internal resistance r of the batteries plus the fixed circuit resistance R_{circ} , so $R_{\text{tot}} = r + R_{\text{circ}}$. As the cells weaken, their internal resistance r grows steadily, so R_{tot} grows and the time constant $\tau = R_{\text{tot}} C$ grows with it. The capacitor still reaches its full charge eventually, because $Q_{\text{max}} = CV$ is set by C and the EMF, not by R ; but it does so more slowly. The owner therefore notices a longer wait between flashes as the cells weaken, since the flash circuit cannot fire until the capacitor has reached close to its full charge.

14. The time constant for a series RC circuit is $\tau = RC$, so the required capacitance is:

$$C = \frac{\tau}{R} = \frac{10\text{s}}{100 \times 10^3 \Omega} = 1.0 \times 10^{-4} \text{F} = 100 \mu\text{F}$$

The required capacitance is $100 \mu\text{F}$.

15. Rotating the moving plates of the variable capacitor partly out of the fixed plates reduces the effective area of overlap A between the two sets of plates, and the capacitance $C = \epsilon_0 A/d$ therefore decreases. The radio's resonant frequency is $f_0 = 1/(2\pi\sqrt{LC})$, which is inversely proportional to $\sqrt{C}$. As C decreases, f_0 rises, and the radio tunes from a station of lower carrier frequency to one of higher carrier frequency.

16. When the trigger is pressed, the sudden compression of the piezoelectric element produces a brief charge pulse on its faces; the element behaves at that instant as a small capacitor whose plates have been charged from rest to a high voltage. The stored charge dumps almost instantly through the spark gap as a single high-current discharge, producing the spark. Once the element has discharged through the gap, it carries no further charge while the trigger is held down; no current flows and no further spark is produced. Only when the trigger is released and re-pressed is the element compressed again, generating a fresh charge pulse and a fresh single spark.

17. An electrolytic capacitor's dielectric is a thin oxide layer formed and maintained on the surface of one of the aluminium foils by the chemistry of the electrolyte under the correct polarity. The oxide layer is what gives the device its very high capacitance per unit volume. Reverse polarity reverses the chemistry: the oxide dissolves rather than forms, the dielectric layer rapidly thins until it can no longer block conduction, and the capacitor then behaves as a short circuit at the applied voltage. A large current flows, boils the electrolyte, and can rupture the can, sometimes violently. The polarity marking is therefore a safety warning, not merely an electrical convenience.

18. Power dissipated by the bulb:

$$P = \frac{V^2}{R} = \frac{(3\text{V})^2}{100 \Omega} = 0.090\text{W} = 90\text{mW}$$

Energy required for 30s of glow:

$$U = P \times t = (0.090\text{W})(30\text{s}) = 2.7\text{J}$$

The capacitor must store at least this much energy. From $U = 1/2 C V^2$:

$$C = \frac{2U}{V^2} = \frac{2 \times 2.7\text{J}}{(3\text{V})^2} = 0.6\text{F}$$

The capacitance is about 0.6F .

19. Water has a relative permittivity ϵ_r of about 81 at room temperature, far larger than that of dry soil components (ϵ_r about 4). When water fills the spaces between soil particles in the sensor's gap, the effective ϵ_r of the soil-water mixture rises sharply, from about 4 in dry conditions to about 15 to 20 when saturated. The capacitance $C = \epsilon_r \epsilon_0 A/d$ scales linearly with this effective ϵ_r , so a factor-of-four increase in capacitance (from 4nF to 15nF) reflects a factor-of-four increase in the effective dielectric constant, exactly the change expected when dry soil is wetted to saturation by the previous day's rain.

20. With the switch open, the battery is disconnected from the rest of the circuit, and the only available path for the capacitor to discharge is through the bulb itself. The capacitor therefore discharges as a simple RC circuit with R equal to the bulb's resistance, the current decaying exponentially as $I(t) = (V_0/R) e^{-t/RC}$. The bulb is lit by this decaying current and dims smoothly as the current falls. The afterglow lasts of the order of one or two time constants $\tau = RC$; for a torch designed with τ of about half a second, the bulb dims noticeably within a second and then goes dark.

21. The two plates of the giant capacitor are the lower charge layer of the storm cloud, typically carrying a strong negative charge at its base, and the surface of the Earth below, which is induced positively by the cloud's charge. The dielectric is the column of air between them. Air, like any insulator, has a dielectric strength: a maximum electric field above which the air molecules ionise and the gas conducts as a plasma. For dry air this critical field is approximately $3 \times 10^6 \text{V/m}$. As long as the field between cloud and ground remains below this value, the air holds off the stored charge and the cloud-Earth system behaves as a charged capacitor; once the field exceeds the dielectric strength of air, breakdown occurs along a conducting channel, and the stored energy releases as the lightning stroke. Humid air has a lower breakdown field, which is part of the reason why lightning is most common in stormy, humid conditions.

EXERCISE 11C

22. Series: $1.09 \mu\text{F}$, $262 \mu\text{C}$, 31.4mJ . Parallel: $12 \mu\text{F}$, 2.88mC , 0.346J .

23. (i) $2.67 \mu\text{F}$ (ii) $320 \mu\text{C}$ (iii) 80V (iv) 19.2mJ

24. 7.37s .

25. Half-life: 138.6s , 7 half-lives are required.

26. $3.6 \times 10^{-4} \text{N}$, the plate force is about **27** times **smaller** than the weight of 1g of water.

27. (i) 8.19mC (ii) 8.2J

28. The stack is two capacitors in series, the upper of capacitance $C_1 = 2K_1\epsilon_0 A/d$ and the lower $C_2 = 2K_2\epsilon_0 A/d$, each with full plate area A and thickness $d/2$. In series:

$$\frac{1}{C_{\text{eq}}} = \frac{1}{C_1} + \frac{1}{C_2} = \frac{d}{2\epsilon_0 A} \left(\frac{1}{K_1} + \frac{1}{K_2} \right) = \frac{d}{2\epsilon_0 A} \left(\frac{K_1 + K_2}{K_1 K_2} \right)$$

Hence:

$$C_{\text{eq}} = \frac{2\epsilon_0 A}{d} \left(\frac{K_1 K_2}{K_1 + K_2} \right)$$

29. (i) $I_{\text{peak}} = 250 \text{A}$ (ii) $\tau = 40 \mu\text{s}$ (iii) The entire stored energy is dissipated as heat in the coil: 2.5J .

30. Earthing the outer shell forces its potential to zero, and the inner shell at potential V_{in} relative to it has:

$$V_{\text{in}} = \frac{kQ}{a} - \frac{kQ}{b} = kQ \left(\frac{b-a}{ab} \right)$$

So $C = \frac{Q}{V_{\text{in}}} = \frac{ab}{k(b-a)} = \frac{4\pi\epsilon_0 ab}{b-a}$, confirming a capacitor relationship $Q = CV$.

With $a = 0.05 \text{m}$, $b = 0.10 \text{m}$:

$$C = \frac{0.05 \times 0.10}{(9 \times 10^9)(0.05)} = 11.1 \text{pF}$$

An isolated sphere of radius $a = 0.05 \text{m}$ has $C_{\text{iso}} = 4\pi\epsilon_0 a = a/k = 5.56 \text{pF}$. The double-shell capacitor has twice the capacitance of the isolated inner sphere alone.

31. 6.9nF .

32. 96pF .

33. 5cm

34. $9.5 \times 10^{-3} \text{K}$

35. (i) 11.1pF (ii) 44.3pF (iii) 200pF (iv) Place the sensor as the timing capacitor of an RC or LC oscillator. As soil moisture rises, the sensor's capacitance rises and the oscillator frequency falls. A frequency-to-voltage converter (a simple monostable triggered by the oscillator, with the output averaged through a low-pass filter) then drives a moving-coil meter that has been calibrated against soil samples of known moisture content, giving the farmer a direct percentage reading.

36. (i) 0.721s (ii) 1.50s

37. (i) 8.85nC (ii) 112V (iii) $0.495 \mu\text{J}$

38. (i) 109.6mJ (ii) 96.1mJ (iii) 13.5mJ . The energy lost is dissipated as heat in the connecting wires during the brief transient redistribution of charge between the two capacitors.