Electromagnetism I: Electrostatics

The material here is covered at the right level in chapters 1–3 of Purcell. For a separate introduction to vector calculus, see the resources mentioned in the syllabus, or chapter 1 of Griffiths. Electrostatics is covered in more mathematical detail in chapter 2 of Griffiths. For interesting general discussion, see chapters II-1 through II-5 of the Feynman lectures. There is a total of **80** points.

1 Coulomb's Law and Gauss's Law

We'll begin with some basic problems which can be solved with symmetry arguments.

Idea 1

Gauss's law is written in integral form as

$$\oint \mathbf{E} \cdot d\mathbf{S} = \frac{Q}{\epsilon_0}.$$

In practice, you will only apply this form to situations with high symmetry, where

$$E = \begin{cases} Q/4\pi\epsilon_0 r^2 & \text{spherical symmetry,} \\ \lambda/2\pi\epsilon_0 r & \text{cylindrical symmetry,} \\ \sigma/2\epsilon_0 & \text{infinite plane.} \end{cases}$$

Example 1

Consider a spherical shell of uniform surface charge density σ. A small hole is cut out of the surface of the shell. What is the electric field at the center of this hole?

Solution

We use the principle of superposition. First, consider the entire spherical shell, without a hole. By Gauss's law and spherical symmetry, the radial electric field at a point P infinitesimally outside the sphere is σ/ϵ_0, while the electric field at a nearby point P' infinitesimally inside is zero.

This field is the superposition of the fields of the charges near P and P', and charges from the entire rest of the sphere. Consider the effect of a small piece of the surface, near P and P'. From the perspective of these points, this piece looks like an infinite plane, so its radial electric field is $\sigma/2\epsilon_0$ at P, and $-\sigma/2\epsilon_0$ at P'. Therefore, the entire rest of the sphere must contribute a radial electric field of $\sigma/2\epsilon_0$, at both P and P'. Therefore, when one cuts out a hole, this is the only contribution that remains, so the field is just $\sigma/2\epsilon_0$.

[2] **Problem 1** (Griffiths 2.18). Some questions about uniformly charged spheres.

(a) Consider a sphere of radius R and uniform charge density ρ. Find the electric field everywhere.

(b) Now two spheres, each of radius R and carrying uniform charge densities ρ and $-\rho$, are placed so that they partially overlap. Call the vector from the positive center to the negative center $\mathbf{d}$. Find the electric field in the overlap region.

[2] **Problem 2.** Consider a cube with a corner at the origin, and sides parallel to the x, y, and z axes. If a charge q is placed at $(\epsilon, \epsilon, \epsilon)$ for some tiny ϵ, what's the flux through each face of the cube?

[2] **Problem 3** (BAUPC)**.** In both parts below, take the potential to be zero at infinity.

(a) Consider a solid sphere of uniform charge density. Find the ratio of the electrostatic potential at the surface to that at the center.

(b) Consider a solid cube of uniform charge density. Find the ratio of the electrostatic potential at a corner to that at the center. (Hint: use symmetry.)

Idea 2

If you follow an electric field line, the potential monotonically decreases along it.

[2] **Problem 4.** Two questions about electrostatic equilibrium.

(a) Prove that when a system of point charges is in equilibrium (i.e. the net force on *each* of the charges due to the others vanishes), the total potential energy of the system is zero.

(b) Show that for a positive point charge in the electric fields of fixed, positive point charges, there is a path along which the charge can be moved to infinity without ever needing positive external work, i.e. a path along which the potential only decreases.

Idea 3

Gauss's law is written in differential form as

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}.$$

The divergence of a vector field $\mathbf{F} = F_x \hat{\mathbf{x}} + F_y \hat{\mathbf{y}} + F_z \hat{\mathbf{z}}$ is

$$\nabla \cdot \mathbf{F} = \partial_x F_x + \partial_y F_y + \partial_z F_z$$

in Cartesian coordinates, where ∂_x stands for $\partial/\partial x$, and so on.

Example 2

Show that the two forms of Gauss's law are equivalent.

Solution

To do this, we need to establish the geometric meaning of the divergence. For simplicity we consider two dimensions; the proof for three dimensions is similar. Consider a small rectangle prism with one corner at the origin, with axes aligned with the Cartesian coordinate axes and side lengths Δx and Δy. To apply Gauss's law in integral form, we need to compute the flux through each side. The flux going out the top side is

$$\int_0^{\Delta x} E_y(x, \Delta y)\, dx$$

while the flux going out the bottom side is

$$-\int_0^{\Delta x} E_y(x,0)\,dx.$$

The sum of these two terms is

$$\int_0^{\Delta x} (E_y(x,\Delta y) - E_y(x,0))\,dx \approx \Delta y \int_0^{\Delta x} (\partial_y E_y)|_{(x,0)}\,dx$$

where we applied a tangent line approximation, and the subscript indicates where the function $\partial_y E_y$ is evaluated. Higher-order terms in the Taylor series would be proportional to higher powers of Δy, which is small, so we can ignore them.

The integrand is still a function of x, but we can Taylor expand it about the origin as

$$(\partial_y E_y)|_{(x,0,0)} = (\partial_y E_y)|_{(0,0,0)} + \Delta x(\ldots) + \ldots.$$

These extra terms are again higher-order in Δx and Δy, so we ignore them. The net flux through the top and bottom faces is hence, to lowest order,

$$\Delta y \int_0^{\Delta x} (\partial_y E_y)|_{(0,0,0)}\,dx = \Delta x \Delta y\,(\partial_y E_y)|_{(0,0,0)}.$$

By similar reasoning, pairing up the left and right faces gives

$$\text{flux} = \Delta x \Delta y\,(\partial_x E_x + \partial_y E_y)|_{(0,0,0)} = \Delta x \Delta y\,(\nabla \cdot \mathbf{E})|_{(0,0,0)}.$$

Thus the divergence is the outgoing flux per unit area, or volume in three dimensions.

This shows us why the two forms of Gauss's law area equivalent. For example, starting from the differential form, the left-hand side is the flux per volume, while the right-hand side is the charge per volume, divided by ϵ_0. Integrating both sides over some volume relates the total flux to the total charge divided by ϵ_0, which is Gauss's law in integral form.

If the above derivation was a bit abstract, we can also show the idea using specific examples.

Example 3

Suppose the region $0 < x < d$ has charge density $-\rho$, and the region $-d < x < 0$ has charge density ρ. Find the electric field everywhere.

Solution

By translational symmetry, the field always points along $\hat{\mathbf{x}}$ and only depends on x, $\mathbf{E}(\mathbf{r}) = E(x)\,\hat{\mathbf{x}}$. By applying the integral form of Gauss's law to a rectangular prism, with one side at x_l and another at x_r, we have

$$E(x_r) - E(x_l) = \frac{1}{\epsilon_0} \int_{x_l}^{x_r} \rho(x)\, dx, \quad E(x) = \frac{1}{\epsilon_0} \int_0^x \rho(x)\, dx + E_0.$$

Since the divergence of $\mathbf{E}(\mathbf{r})$ is just $\partial E(x)/\partial x$, this clearly satisfies the differential form of Gauss's law. To fix the undetermined constant E_0, we could demand the field be zero on both sides of the charge distribution, motivated by symmetry. Then we have

$$E(x) = \frac{\rho}{\epsilon_0} \times \begin{cases} d - x & 0 < x < d, \\ d + x & -d < x < 0, \\ 0 & \text{elsewhere.} \end{cases}$$

Example 4

Find the electric field of a spherically symmetric charge density $\rho(r)$.

Solution

By spherical symmetry, the field always points radially and only depends on r, $\mathbf{E}(\mathbf{r}) = E(r)\,\hat{\mathbf{r}}$. By applying the integral form of Gauss's law to a sphere of radius r,

$$4\pi r^2 E(r) = \frac{1}{\epsilon_0} \int_0^r dr'\, 4\pi r'^2 \rho(r'), \quad E(r) = \frac{1}{\epsilon_0} \frac{1}{r^2} \int_0^r dr'\, r'^2 \rho(r').$$

Let's check that this indeed satisfies the differential form of Gauss's law, using the divergence in spherical coordinates. For any vector field $\mathbf{F} = F_r \hat{\mathbf{r}} + F_\theta \hat{\boldsymbol{\theta}} + F_\varphi \hat{\boldsymbol{\varphi}}$, the divergence is

$$\nabla \cdot \mathbf{F} = \frac{1}{r^2} \frac{\partial (r^2 F_r)}{\partial r} + \frac{1}{r \sin\theta} \frac{\partial}{\partial \theta} (F_\theta \sin\theta) + \frac{1}{r \sin\theta} \frac{\partial F_\varphi}{\partial \varphi}.$$

This looks complicated, but things turn out simple because $\mathbf{E}$ only has a radial component, $E_r = E(r)$, which gives

$$\nabla \cdot \mathbf{E} = \frac{1}{r^2} \frac{\partial (r^2 E(r))}{\partial r} = \frac{1}{r^2 \epsilon_0} \frac{\partial}{\partial r} \int_0^r dr'\, r'^2 \rho(r') = \frac{r^2 \rho(r)}{r^2 \epsilon_0} = \frac{\rho(r)}{\epsilon_0}$$

just as desired.

[3] **Problem 5.** Consider a vector field expressed in polar coordinates, $\mathbf{F} = F_r \hat{\mathbf{r}} + F_\theta \hat{\boldsymbol{\theta}}$ where $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ are unit vectors in the radial and tangential directions. Gauss's law in differential form still works in these coordinates, but the form of the divergence is different.

By considering the flux per unit area out of a small region bounded by r and $r + dr$, and θ and $\theta + d\theta$, and applying Gauss's law in integral form, find what the divergence in polar coordinates must be for Gauss's law in differential form to hold. (Optional: try generalizing to spherical coordinates.)

[4] **Problem 6.** This problem is quite subtle, but will enhance your understanding of electromagnetism. Suppose that all of space is filled with uniform charge density ρ.

(a) Show that $\mathbf{E} = (\rho/\epsilon_0) x \hat{\mathbf{x}}$ obeys the differential form of Gauss's law.

(b) Show that $\mathbf{E} = (\rho/3\epsilon_0)r\hat{\mathbf{r}}$ also obeys Gauss's law.

(c) Argue that by symmetry, $\mathbf{E} = 0$. Show that this does not obey Gauss's law.

(d) ⋆ What's going on? Which, if any, is the actual field? If you think there's more than one possible field, how could that be consistent with Coulomb's law, which gives the answer explicitly? For that matter, what does Coulomb's law say about this setup, anyway?

Idea 4

A tricky, occasionally useful idea is to use Newton's third law: it may be easier to calculate the force of A on B than the force of B on A.

Example 5: Purcell 1.28

Consider a point charge q. Draw any imaginary sphere of radius R around the charge. Show that the average of the electric field over the surface of the sphere is zero.

Solution

Imagine placing a uniform surface charge σ on the sphere. Then the average of the point charge's electric field over the sphere times $4\pi R^2\sigma$ is the total force of the point charge on the charged sphere. But this is equal in magnitude to the force of the charged sphere on the point charge, which must be zero by the shell theorem. Thus the average field over the sphere has to vanish.

Example 6

Consider two spherical balls of charge q and radii a_i, with their centers separated by a distance $r > a_1 + a_2$. What is the net force of the first on the second?

Solution

It might seem obvious that the answer is $q^2/4\pi\epsilon_0 r^2$, with no dependence on a_1 and a_2. In fact, if you've done any orbital mechanics, you've almost certainly assumed that the force between two spherical bodies (such as the Earth and Sun) is Gm_1m_2/r^2, which is equivalent.

This has a simple but slightly tricky proof. By the shell theorem, we can set $a_1 = 0$, replacing the first ball with a point charge, because this produces the same field at the second ball. But the force on the second ball depends on the electric field at every point on it, which seems to require doing an integral. To avoid this, we use Newton's third law, which tells us it's equivalent to compute the force on the first ball. To compute *that*, we may set $a_2 = 0$ by the shell theorem again. This reduces us to the case of two point charges, giving the answer.

[3] **Problem 7** (Purcell 1.28)**.** Some extensions of the previous example.

(a) Show that if the charge q is instead outside the sphere, a distance $r > R$ from its center, the average electric field over the surface of the sphere is the same as the electric field at the center of the sphere.

(b) Show that for any overall neutral charge distribution contained within a sphere of radius R, the average electric field over the interior of the sphere is $-\mathbf{p}/4\pi\epsilon_0 R^3$ where $\mathbf{p}$ is the total dipole moment.

[3] **Problem 8.** There are two point charges, $q_1 > 0$ and $q_2 < 0$, in empty space. An electric field line leaves q_1 at an angle α from the line connecting the two charges. Determine whether this field line hits q_2, and if so, at what angle β from the line connecting the two charges. (Hint: this can be done without solving any differential equations.)

Idea 5

The integral $\int d\mathbf{S}$ over a surface with a fixed boundary is independent of the surface.

We proved this in a mechanical way in **M2**. If you want to see a proof using vector calculus, see problem 1.62 of Griffiths.

[3] **Problem 9.** A hemispherical shell of radius R has uniform charge density σ and is centered at the origin. Find the electric field at the origin. (Hint: combine the previous two ideas.)

[3] **Problem 10.** A point charge q is placed a distance $a/2$ above the center of a square of charge density σ and side length a. Find the force of the square on the point charge.

[4] **Problem 11** (Griffiths 2.47, PPP 113, MPPP 140)**.** Consider a uniformly charged spherical shell of radius R and total charge Q.

(a) Find the net electrostatic force that the southern hemisphere exerts on the northern hemisphere.

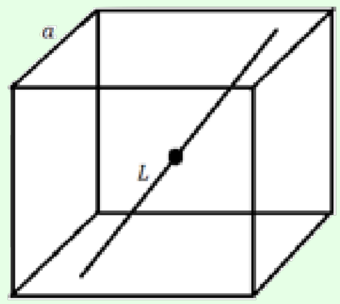
(b) Generalize part (a) to the case where the sphere is split into two parts by a plane whose minimum distance to the sphere's center is h.

(c) Generalize part (a) to the case where the two hemispherical shells have uniform charge density, opposite orientation, and the same center, but have different total charges q and Q, and different radii r and R, where $r < R$.

Hint: see example 10, and use superposition and symmetry when applicable.

Example 7: IdPhO 2020.1A

A point charge of mass m and charge $-q$ is placed at the center of a cube with side length a, whose volume has uniform charge density ρ. The point charge is allowed to slide along a straight line, which has an arbitrary orientation, so that the distance along the line from the center to one of the cube's faces is L.



Find the angular frequency of small oscillations.

Solution

The official solution goes as follows: consider displacing the point charge by some small amount $\Delta \mathbf{r}$. The cube of charge can then be decomposed into (1) a slightly smaller cube of charge centered around the point charge's new position, and (2) three thin plates of charge on the faces opposite to the charge's motion. By symmetry, (1) contributes nothing, and we know what (2) contributes from the answer to problem 10. The result is a restoring force proportional to $-\Delta \mathbf{r}$, whose magnitude has no dependence on the orientation of $\Delta \mathbf{r}$, so the oscillation frequency doesn't depend on L. Once you know this, you can orient the line any way you want, so the problem is simple to finish.

Personally, I don't like this problem because the intended solution requires knowing the answer to problem 10, which itself is pretty tricky. That is, the difficulty of the problem depends mostly on whether you've seen that tough, but standard problem elsewhere. However, I'm including it as an example because there's another way to solve it, which is a bit more advanced, but quite illustrative.

Since this is a question about small oscillations, it suffices to expand the potential energy to second order about the center of the cube. The most general possible expression is

$$V(x, y, z) = a + b_1 x + b_2 y + b_3 z + c_1 x^2 + c_2 y^2 + c_3 z^2 + c_4 xy + c_5 yz + c_6 xz + O(r^3).$$

The constant a doesn't matter, so we can just ignore it. And since $\mathbf{E}$ vanishes at the center, the linear terms b_i are all zero as well. Because the x, y, and z axes are all equivalent by cubical symmetry (e.g. we can rotate them into each other, while keeping the cube the same),

$$c = c_1 = c_2 = c_3, \quad c' = c_4 = c_5 = c_6.$$

Thus, our complicated Taylor series boils all the way down to

$$V(x, y, z) = c(x^2 + y^2 + z^2) + c'(xy + yz + xz) + O(r^3)$$

without even having to do any work! Finally, notice that the cube is symmetric under reflections $x \to -x$, $y \to -y$, or $z \to -z$. These reflections keep the c term the same, but flip the c' term. Therefore, we must have $c' = 0$, so

$$V(r) = cr^2 + O(r^3)$$

which is remarkably simple. The potential near the origin is spherically symmetric, even though the setup as a whole isn't! It's not automatic: it wouldn't be this simple if we had a slightly more complex shape. This "accidental" spherical symmetry is a consequence of the combination of cubical symmetry and the simplicity of Taylor series.

Therefore, to finish the problem we only need to find the coefficient c. While there are simpler ways to do this, I'll do it in a way that introduces some useful facts. Combining the definition of V and Gauss's law, we have

$$\nabla \cdot (\nabla V) = -\nabla \cdot \mathbf{E} = -\frac{\rho}{\epsilon_0}.$$

This is a standard and fundamental result in electrostatics, called Poisson's equation, which we will see again later. The divergence of a gradient is also called a Laplacian, and written as

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -\frac{\rho}{\epsilon_0}.$$

Using this, we can easily compute the value of c, giving

$$V(\mathbf{r}) = -\frac{\rho r^2}{6\epsilon_0} + O(r^3).$$

Therefore, for a displacement $\Delta\mathbf{r}$ in *any* direction, the restoring force is $\rho q r/3\epsilon_0$ in the opposite direction, which means

$$\omega = \sqrt{\frac{\rho q}{3\epsilon_0 m}}$$

independent of the orientation of the line.

Remark

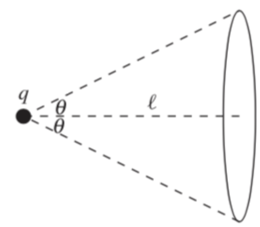
Accidental symmetry is important in modern physics. For example, protons are stable because of an accidental symmetry in the Standard Model, which ensures that baryon number is conserved. That explains why we often expect proton decay to occur in extensions of the Standard Model, such as grand unified theories, as explained in this nice article.

2 Continuous Charge Distributions

Idea 6

In almost all cases in Olympiad physics, there will be sufficient symmetry to reduce any multiple integral to a single integral. Remember that when using Gauss's law, the Gaussian surface may be freely deformed as long as it doesn't pass through any charges.

[2] **Problem 12** (Purcell 1.15). A point charge q is located at the origin. Compute the electric flux that passes through a circle a distance ℓ from q, subtending an angle 2θ as shown below.



[3] **Problem 13** (Purcell 1.8)**.** A ring with radius R has uniform positive charge density λ. A particle with positive charge q and mass m is initially located in the center of the ring and given a tiny kick. If the particle is constrained to move in the plane of the ring, show that it exhibits simple harmonic motion and find the angular frequency.

[3] **Problem 14** (Purcell 1.12)**.** Consider the setup of problem 9. If the hemisphere is centered at the origin and lies entirely above the xy plane, find the electric field at an arbitrary point on the z-axis. (This is a bit complicated, and is representative of the most difficult kinds of integrals you might have to set up in an Olympiad. For a useful table of integrals, see Appendix K of Purcell.)

[3] **Problem 15.** USAPhO 2018, problem B1.

Idea 7: Electric Dipoles

The dipole moment of two charges q and $-q$ separated by $\mathbf{d}$ is $\mathbf{p} = q\mathbf{d}$. More generally, the dipole moment of a charge configuration is defined as

$$\mathbf{p} = \int \rho(\mathbf{r})\mathbf{r}\, d^3\mathbf{r}.$$

For an overall neutral charge configuration, the leading contribution to its electric potential far away is the dipole potential,

$$\phi(r, \theta) = \frac{p\cos\theta}{4\pi\epsilon_0 r^2}$$

where θ is the angle of $\mathbf{r}$ to $\mathbf{p}$.

Remark

Here's a trick to remember the dipole potential. Let $\phi_0(\mathbf{r}) = k/r$ be the potential for a unit charge at the origin. An ideal point dipole of dipole moment p consists of charges $\pm p/d$ separated by d, in the limit $d \to 0$. So the potential is

$$p \lim_{\mathbf{d}\to 0} \frac{\phi_0(\mathbf{r}) - \phi_0(\mathbf{r}+\mathbf{d})}{d}.$$

But this is precisely the (negative) derivative, so you can get the dipole potential by differentiating the ordinary potential! Indeed, for a dipole aligned along the $\hat{\mathbf{z}}$ axis,

$$-\frac{\partial}{\partial z}\frac{kp}{r} = \frac{kp}{r^2}\frac{\partial r}{\partial z} = \frac{kp}{r^2}\frac{z}{r} = \frac{kp\cos\theta}{r^2}$$

which matches the above result. You can use the same trick for quadrupoles and higher multipoles, which we'll see in **E8**.

[3] **Problem 16.** In this problem we'll derive essential results about dipoles, which will be used later.

(a) Using the binomial theorem, derive the dipole potential given above, for a dipole made of a pair of point charges $\pm q$ separated by distance d, oriented along the z-axis.

(b) Differentiate this result to find the dipole field,

$$\mathbf{E}(\mathbf{r}) = \frac{p}{4\pi\epsilon_0 r^3}(2\cos\theta\,\hat{\mathbf{r}} + \sin\theta\,\hat{\boldsymbol{\theta}})$$

where the expression above is in spherical coordinates. (Hint: feel free to use the expression for the gradient in spherical coordinates.)

(c) Show that this may also be written as

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0 r^3}(3(\mathbf{p}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}} - \mathbf{p}).$$

You don't need to memorize these expressions, but it's useful to remember what a dipole field looks like, the fact that its magnitude is roughly $p/4\pi\epsilon_0 r^3$, and the fact that the numeric prefactor is 2 along the dipole's axis and 1 perpendicular to it.

[3] **Problem 17.** USAPhO 2002, problem B2.

[3] **Problem 18.** USAPhO 2009, problem B2. This essential problem introduces useful facts about dipole-dipole interactions.

Idea 8

The potential energy of a set of point charges is

$$U = \frac{1}{4\pi\epsilon_0}\sum_{i\neq j}\frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|} = \frac{1}{2}\sum_i q_i V(\mathbf{r}_i).$$

We sum over $i \neq j$ to avoid computing the energy of a single point charge due to its interaction with itself, which would be infinite. For a continuous distribution of charge, we don't have this problem, and instead find

$$U = \frac{1}{2}\int \rho(\mathbf{r})V(\mathbf{r})\,d^3\mathbf{r} = \frac{\epsilon_0}{2}\int |\mathbf{E}(\mathbf{r})|^2\,d^3\mathbf{r}.$$

Unlike the other quantities we've considered, energy doesn't obey the superposition principle.

[3] **Problem 19.** In this problem we'll apply the above results to balls of charge.

(a) Compute the potential energy of a uniformly charged ball of total charge Q and radius R.

(b) Show that the potential energy of two point charges of charge $Q/2$ separated by radius R is lower than the result of part (a).

(c) Hence it appears that it is energetically favorable to compress a ball of charge into two point charges. Is this correct?

[3] **Problem 20.** An insulating circular disc of radius R has uniform surface charge density σ.

(a) Find the electric potential on the rim of the disc.

(b) Find the total electric potential energy stored in the disc.

[3] **Problem 21.** Consider a uniformly charged ball of total charge Q and radius R. Decompose this ball into two parts, A and B, where B is a ball of radius $R/2$ whose center is a distance $R/2$ of the ball's center, and A is everything else. Find the potential energy due to the interaction of A and B, i.e. the work necessary to bring in B from infinity, against the field of A.

[2] **Problem 22** (PPP 149)**.** A distant planet is at a very high electric potential compared with Earth, say 10^6 V higher. A metal space ship is sent from Earth for the purpose of making a landing on the planet. Is the mission dangerous? What happens when the astronauts open the door on the space ship and step onto the surface of the planet?

Example 8

Since Newton's law of gravity is so similar to Coulomb's law, the results we've seen so far should have analogues in Newtonian gravity. What are they? For example, what's the gravitational Gauss's law?

Solution

The fundamental results to compare are

$$F = -\frac{Gm_1m_2}{r^2}, \quad F = \frac{q_1q_2}{4\pi\epsilon_0 r^2}$$

where the minus sign indicates that the gravitational force is attractive, while the electrostatic force between like charges is repulsive. Then we can transform a question involving (only positive) electric charges to one involving masses if we map

$$q \to m, \quad \frac{1}{4\pi\epsilon_0} \to -G, \quad \mathbf{E} \to \mathbf{g}.$$

Thus, while electrostatics is described by

$$\nabla \times \mathbf{E} = 0, \quad \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad \oint \mathbf{E} \cdot d\mathbf{S} = \frac{Q}{\epsilon_0},$$

the gravitational field is described by

$$\nabla \times \mathbf{g} = 0, \quad \nabla \cdot \mathbf{g} = -4\pi G\rho_m, \quad \oint \mathbf{g} \cdot d\mathbf{S} = -4\pi GM$$

where ρ_m is the mass density. Similarly, the potential energy can be written in two ways,

$$U = \frac{1}{2}\int \rho_m(\mathbf{r})\phi(\mathbf{r})\, d^3\mathbf{r} = -\frac{1}{8\pi G}\int |\mathbf{g}(\mathbf{r})|^2\, d^3\mathbf{r}$$

where $\phi(\mathbf{x})$ is the gravitational potential. This result was first written down by Maxwell.

Remark

Here's a philosophical question: is potential energy "real"? You likely think the answer is obvious, but about half of your friends probably think the opposite answer is obviously correct! In fact, in the 1700s, there was a lively debate over whether the ideas of kinetic energy and momentum, which at the time were given various other names, were worthwhile. Which one of the two was the *true* measure of motion? In our modern language, proponents of energy pointed out that the momentum always vanished in the center of mass frame, which made it "trivial", while supporters of momentum replied that kinetic energy was clearly not conserved in even the simplest of cases, like inelastic collisions.

In the 1800s, thermodynamics was developed, allowing the energy seemingly lost in inelastic collisions to be accounted for as internal energy. But there still remained the problem that kinetic energy was lost in simple situations, such as when balls are thrown upward. By the mid-1800s, the modern language that "kinetic energy is converted to potential energy" was finally standardized, but it was still common to read in textbooks that potential energy was fake, a mathematical trick used to patch up energy conservation. After all, potential energy has some suspicious qualities. If a ball has lots of potential energy, you can't see or feel it, or even know it's there by considering the ball alone. It doesn't seem to be located anywhere in space, and its amount is arbitrary, as a constant can always be added.

In the late 1800s, a revolution on physics answered some of these questions. Maxwell and his successors recast electromagnetism as a theory of fields, and showed that the dynamics of charges and currents were best understood by allowing the fields themselves to carry energy and momentum. We'll cover this in detail in **E7**, but for now, it implies that electrostatic potential energy is fundamentally stored in the field, with a density of $\epsilon_0 E^2/2$. This implies that its location and total amount are directly measurable.

Maxwell believed that the dynamics of fields emerged from the microscopic motions and elastic deformations of an all-pervading ether, in the same way that, say, a fluid's velocity field emerges from the average motion of fluid molecules. This makes it manifestly positive, so he was disturbed to find that the energy density of a gravitational field is *negative*!

A few decades later, the arrival of special relativity answered some questions and reopened others. On one hand, it demolished Maxwell's vision of the ether. On the other hand, it finally answered the question of whether all kinds of potential energy are "real", and it got rid of the freedom to add arbitrary constants. That's because in special relativity, the total energy of a system at rest is related to its mass by $E = mc^2$, and the mass is directly measurable. This finally puts thermal energy, elastic potential energy, and field energy on an equal footing.

Here's the most modern view of energy conservation. All particles and their interactions are fundamentally described by relativistic quantum fields. A famous result called Noether's theorem implies that whenever such a theory is time-translationally symmetric, there is a conserved quantity which we call the energy. (The distinction between kinetic and potential energy becomes irrelevant; it's all just energy.) The density of energy in space can be computed from the state of the fields, but it doesn't need to be explained, as

Maxwell imagined, by the internal motion of whatever the fields are made of. The fields are fundamental: they aren't made of anything; instead, they make up everything!

What happens when we throw gravity into the mix? As we'll discuss further in **R3**, it turns out that at nonrelativistic velocities, the dynamics of gravitating particles can be described by "gravitoelectromagnetism", a theory closely analogous to electromagnetism, where moving masses also source "gravitomagnetic" fields $\mathbf{B}_g$, which result in $m\mathbf{v} \times \mathbf{B}_g$ forces. But the situation gets much more subtle when we upgrade to full general relativity. Here, the notion of a gravitational field disappears completely, and is replaced by the curvature of spacetime, making it hard to define an energy density for it at all. For an accessible overview of the debate, see this paper. Ultimately, though, it doesn't matter that much, since it doesn't impair our ability to use either Newtonian gravity or general relativity.

Example 9

For an infinite line of linear charge density λ, find the potential $V(r)$ by dimensional analysis.

Solution

This example illustrates a famous subtlety of dimensional analysis. The only quantities in the problem with dimensions are λ, ϵ_0, and r. To get the electrical units to balance, we have

$$V(r) = \frac{\lambda}{2\pi\epsilon_0} f(r)$$

where $f(r)$ is a dimensionless function. But there are *no* nontrivial dimensionless functions of a dimensionful quantity r. The only possibilities are that $f(r)$ is a dimensionless constant, or that $f(r)$ is infinite. In the first case, the electric field would vanish, which can't be right. In the second case, it is unclear how to calculate the electric field at all.

In fact, the electric potential *is* infinite, if you insist on the usual convention of setting $V(\infty) = 0$. In that case, we have

$$V(r) = \int_r^\infty \frac{\lambda}{2\pi\epsilon_0} \frac{dr}{r} = \infty$$

independent of r. But this is useless; to get a finite result we can actually work with, we need to subtract off an infinite constant from the potential. Equivalently, we need to set the potential to be zero at some finite distance $r = r_0$. This process is known as renormalization, and it is extremely important in modern physics. After renormalization, we have

$$V(r) = \frac{\lambda}{2\pi\epsilon_0} \log \frac{r_0}{r}$$

which is perfectly consistent with dimensional analysis.

Notice that in the process of renormalization, a new dimensionful quantity r_0 appeared out of nowhere. This phenomenon is known as dimensional transmutation. Of course, physical predictions don't depend on this new scale (e.g. the electric field is independent of r_0), but you can't write down quantities like the potential without it.

3 Conductors

Idea 9

In electrostatic conditions, $\mathbf{E} = 0$ inside a conductor, which implies the conductor has constant electric potential V. This further implies that $\mathbf{E}$ is always perpendicular to a conductor's surface. By Gauss's law, the conductor has $\rho = 0$ everywhere inside, so all charge resides on the surface.

Example 10

Consider a point on the surface of a conductor with surface charge density σ. Show that the outward pressure on the charges at this point is $\sigma^2/2\epsilon_0$.

Solution

Gauss's law tells us that the difference of the electric fields right inside and outside the conductor at this point is

$$E_{\text{out}} - E_{\text{in}} = \frac{\sigma}{\epsilon_0}$$

by drawing a pillbox-shaped Gaussian surface. But we also know that $E_{\text{in}} = 0$ since we're dealing with a conductor, so $E_{\text{out}} = \sigma/\epsilon_0$.

Let's think about how this electric field is made. If there were no charges around except for the ones at this surface, then the interior and exterior fields would have been $\pm\sigma/2\epsilon_0$. This means that all of the other charges, that lie elsewhere on the surface of the conductor, must provide a field $\sigma/2\epsilon_0$ here, so that E_{in} cancels out.

The pressure on the charges at this point on the surface is equal to the product of the surface charge density with the field due to the *rest* of the charges, since the charges at this point can't exert an overall force on themselves, so

$$P = \sigma \left(\frac{\sigma}{2\epsilon_0} \right) = \frac{\sigma^2}{2\epsilon_0}$$

as required. Equivalently, we can conclude that $P = \epsilon_0 E_{\text{out}}^2/2$.

Example 11

Is the charge density at the surface of a charged conductor usually greater at regions of higher or lower curvature?

Solution

We can't answer this question directly, because it is essentially impossible to find the charge distribution of an irregularly shaped conductor. However, we can get some insight by considering the limiting case of a conductor made of two spheres of radii R_1 and R_2, connected by a very long rod.

For the potential to be the same at both spheres, we must have $Q_1/R_1 = Q_2/R_2$, so the charge is proportional to the radius, and the charge density is inversely proportional to the radius. Thus, there's generally higher charge density at sharper points of the conductor.

[1] **Problem 23.** Show that any surface of charge density σ with electric fields $\mathbf{E}_1$ and $\mathbf{E}_2$ immediately on its two sides experiences a force $\sigma(\mathbf{E}_1 + \mathbf{E}_2)/2$ per unit area. (This is a generalization of the example above, where one side was inside a conductor.)

[2] **Problem 24.** Is it possible for a single solid, isolated conductor with a positive total charge to have a negative surface charge density at any point on it? If not, prove it. If so, sketch an example.

Idea 10: Existence and Uniqueness

In a system of conductors where the total charge or potential of each conductor is specified, there exists a unique charge configuration that satisfies those boundary conditions.

This is very useful because in many cases, it is difficult to directly derive the charge distributions or fields. Instead, sometimes one can simply insightfully guess an answer; then it must be the correct answer by uniqueness. For further discussion, see section 2.5 of Griffiths.

Example 12

Consider a conductor with nonzero net charge, and an empty cavity inside. Show that the electric field is zero in the cavity.

Solution

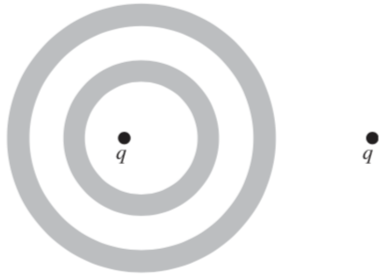
Let's consider a second conductor with the same net charge and the same shape, but without the cavity. By the existence and uniqueness theorem, we know there exists some charge configuration on the second conductor's surface which satisfies the boundary conditions, namely that the electric field vanishes everywhere inside the conductor. In particular, that means the field is zero where the cavity of the original conductor would have been.

Now consider the original conductor again. If we give this conductor precisely the same surface charge distribution, then this will again solve the boundary conditions, and it'll have no field in the cavity. But by the existence and uniqueness theorem, the charge distribution is unique, so this is the only possible answer: the field *must* be zero in the cavity.

If this is your first time seeing this, it can sound like a fast-talking swindle (which is why I made it an example rather than a problem!). It looks like we used no effort and got a strong conclusion out. Of course, that's because all the work is done by the uniqueness theorem.

[1] **Problem 25.** Consider a spherical conducting shell with an arbitrary charge distribution inside, with net charge Q. Find the electric field outside the shell.

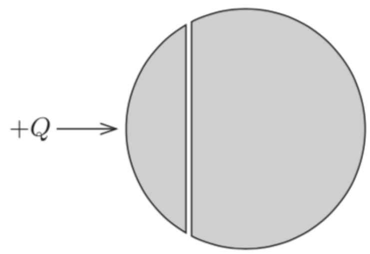
[2] **Problem 26** (Purcell 3.33)**.** The shaded regions represent two neutral conducting spherical shells.



Carefully sketch the electric field. What changes if the two shells are connected by a wire?

[3] **Problem 27.** USAPhO 2014, problem A4.

[4] **Problem 28** (MPPP 150)**.** A solid metal sphere of radius R is divided into two parts by a planar cut, so that the surface area of the curved part of the smaller piece is πR^2. The cut surfaces are coated with a negligibly thin insulating layer, and the two parts are put together again, so that the original shape of the sphere is restored. Initially the sphere is electrically neutral.



The smaller part of the sphere is now given a small positive electric charge Q, while the larger part of the sphere remains neutral. Find the charge distribution throughout the sphere, and the electrostatic interaction force between the two pieces of the sphere.

[3] **Problem 29.** In this problem we'll work through a heuristic proof of a version of the uniqueness theorem. In particular, we will show that for a system of conductors in empty space, specifying the total charge on each conductor alone specifies the entire surface charge distribution.

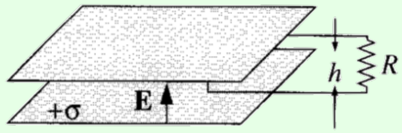
(a) Suppose for the sake of contradiction that two different charge distributions can exist, and consider their difference, which has zero total charge on each conductor. Argue that at least one conductor must have electric field lines both originating from and terminating on it.

(b) Show that at least one of these field lines must originate from or terminate on another one of the conductors.

(c) By generalizing this reasoning, prove the desired result. (Hint: consider the conductors with nonzero surface charges that have the highest and lowest potentials.)

Example 13: Griffiths 7.6

A wire loop of height h and resistance R has one end placed inside a parallel plate capacitor with electric field $\mathbf{E}$, as shown.

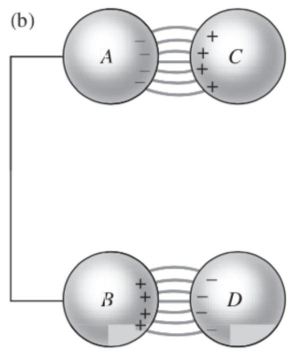


The other end of the loop is far away, where the field is negligible. Find the emf in the loop.

Solution

This is a trick question: if the answer were nonzero, the current would run forever, yielding a perpetual motion machine. Electrostatic fields always produce zero total emf along any loop. The $\sigma h/\epsilon_0$ voltage drop inside the capacitor is canceled out by the voltage drop due to the fringe fields, which are small, but accumulate over a long distance. The point of this example is that, while we can ignore fringe fields for some calculations, they are often essential to get a consistent overall picture. We'll revisit the subtleties of fringe fields in **E2**.

[2] **Problem 30** (Purcell 3.2). Spheres A and B are connected by a wire; the total charge is zero. Two oppositely charged spheres C and D are brought nearby, as shown.



The spheres C and D induce charges of opposite sign on A and B. Now suppose C and D are connected by a wire. Then the charge distribution should not change, because the charges on C and D are being held in place by the attraction of the opposite charge density. Is this correct?

Electromagnetism I: Electrostatics

The material here is covered at the right level in chapters 1–3 of Purcell. For a separate introduction to vector calculus, see the resources mentioned in the syllabus, or chapter 1 of Griffiths. Electrostatics is covered in more mathematical detail in chapter 2 of Griffiths. For interesting general discussion, see chapters II-1 through II-5 of the Feynman lectures. There is a total of **80** points.

1 Coulomb's Law and Gauss's Law

We'll begin with some basic problems which can be solved with symmetry arguments.

Idea 1

Gauss's law is written in integral form as

$$\oint \mathbf{E} \cdot d\mathbf{S} = \frac{Q}{\epsilon_0}.$$

In practice, you will only apply this form to situations with high symmetry, where

$$E = \begin{cases} Q/4\pi\epsilon_0 r^2 & \text{spherical symmetry,} \\ \lambda/2\pi\epsilon_0 r & \text{cylindrical symmetry,} \\ \sigma/2\epsilon_0 & \text{infinite plane.} \end{cases}$$

Example 1

Consider a spherical shell of uniform surface charge density σ. A small hole is cut out of the surface of the shell. What is the electric field at the center of this hole?

Solution

We use the principle of superposition. First, consider the entire spherical shell, without a hole. By Gauss's law and spherical symmetry, the radial electric field at a point P infinitesimally outside the sphere is σ/ϵ_0, while the electric field at a nearby point P' infinitesimally inside is zero.

This field is the superposition of the fields of the charges near P and P', and charges from the entire rest of the sphere. Consider the effect of a small piece of the surface, near P and P'. From the perspective of these points, this piece looks like an infinite plane, so its radial electric field is $\sigma/2\epsilon_0$ at P, and $-\sigma/2\epsilon_0$ at P'. Therefore, the entire rest of the sphere must contribute a radial electric field of $\sigma/2\epsilon_0$, at both P and P'. Therefore, when one cuts out a hole, this is the only contribution that remains, so the field is just $\sigma/2\epsilon_0$.

[2] **Problem 1** (Griffiths 2.18)**.** Some questions about uniformly charged spheres.

(a) Consider a sphere of radius R and uniform charge density ρ. Find the electric field everywhere.

(b) Now two spheres, each of radius R and carrying uniform charge densities ρ and $-\rho$, are placed so that they partially overlap. Call the vector from the positive center to the negative center $\mathbf{d}$. Find the electric field in the overlap region.

Solution. (a) The field inside a uniform sphere of density ρ and center $\mathbf{a}$ is

$$\mathbf{E} = \frac{\rho}{3\epsilon_0}(\mathbf{r} - \mathbf{a}).$$

Outside the sphere, the field falls off as an inverse square,

$$\mathbf{E} = \frac{\rho}{3\epsilon_0} \frac{R^3}{|\mathbf{r} - \mathbf{a}|^3}(\mathbf{r} - \mathbf{a}).$$

(b) If the two centers are $\mathbf{a}_1$ and $\mathbf{a}_2$, then by superposition,

$$\mathbf{E} = \frac{\rho}{3\epsilon_0}((\mathbf{r} - \mathbf{a}_1) - (\mathbf{r} - \mathbf{a}_2)) = \frac{\rho}{3\epsilon_0}\mathbf{d}$$

which is a constant.

[2] **Problem 2.** Consider a cube with a corner at the origin, and sides parallel to the x, y, and z axes. If a charge q is placed at $(\epsilon, \epsilon, \epsilon)$ for some tiny ϵ, what's the flux through each face of the cube?

Solution. There are three "opposite" faces with the same flux, and three "adjacent" faces with the same flux. Now consider adding seven more cubes, so that the charge is now at the center of a $2 \times 2 \times 2$ cube. The total flux through the outer faces of the cube is q/ϵ_0, and there are 24 unit faces, so the flux out of each "opposite" face is $q/24\epsilon_0$. Now consider the original cube. By Gauss's law the total flux out must be q/ϵ_0, which means the flux out of each "adjacent" face is $7q/24\epsilon_0$.

(Note that if the charge were instead *exactly* at one of the corners, the fluxes through the opposite faces would still be $q/24\epsilon_0$, while the fluxes though the adjacent faces would technically be undefined, since the electric field blows up on the face. But roughly speaking, the flux ought to be zero. Then the total flux out of the cube is only $q/8\epsilon_0$, and that's because the corner cuts out one "octant" of the point charge's field.)

Here's a followup question, proposed by Mike Winer and first solved by Jason Youm. If a charge q is at the corner of a regular tetrahedron, what fraction of its flux goes through the tetrahedron's far face? You can't solve it with the same trick as the cube, but it's possible to get the answer without any explicit integration by cleverly considering the flux through combinations of simpler surfaces, and using a little three-dimensional geometry. The answer is

$$\frac{1}{2} - \frac{3 \arctan \sqrt{2}}{2\pi} \approx 0.044.$$

You can try deriving this for yourself, but it's quite tricky; roughly 4 points by the standards of this problem set. In fact, it turns out that it's possible to generalize these kinds of arguments even further, to solve the more general case where the charge is displaced from a vertex of a cube in an *arbitrary* direction! For a deep dive, see this paper.

[2] **Problem 3** (BAUPC)**.** In both parts below, take the potential to be zero at infinity.

(a) Consider a solid sphere of uniform charge density. Find the ratio of the electrostatic potential at the surface to that at the center.

(b) Consider a solid cube of uniform charge density. Find the ratio of the electrostatic potential at a corner to that at the center. (Hint: use symmetry.)

Solution. (a) Let the uniform charge density be ρ and the sphere have radius R, so the total charge is $Q = \frac{4}{3}\pi\rho R^3$. We can treat the field outside the sphere to be like a point charge, so the potential at the surface relative to infinity is $U_0 = Q/4\pi\epsilon_0 R$.

To go from the surface to the center, we need to go against the field lines and change the potential by $\Delta U = -\int_R^0 E(r)\,dr$. The electric field inside the sphere can be found with Gauss's law inside the sphere: $E = \frac{4}{3}\pi\rho r^3/4\pi\epsilon_0 r^2$.

$$\Delta U = \int_0^R \frac{kQr}{R^3} dr = \frac{1}{2}U_0$$

Thus the potential at the center of the sphere is $U_0 + \frac{1}{2}U_0 = \frac{3}{2}U_0$, so the ratio of the potential at the surface to that at the center is $\frac{2}{3}$.

(b) Being at the center of the cube is like being at the corner of 8 identical cubes with half the length. From $U \sim kQ/r \sim k\rho r^2$, we see that the potential is proportional to the square of the length scale. Let the potential at the corner be U_0. For each cube with half the length, the potential from that cube is $\frac{1}{4}U_0$. With eight of those half cubes at the center, the potential at the center of the cube is $2U_0$. So the ratio of the potentials at corner to center is $\frac{1}{2}$.

Idea 2

If you follow an electric field line, the potential monotonically decreases along it.

[2] **Problem 4.** Two questions about electrostatic equilibrium.

(a) Prove that when a system of point charges is in equilibrium (i.e. the net force on *each* of the charges due to the others vanishes), the total potential energy of the system is zero.

(b) Show that for a positive point charge in the electric fields of fixed, positive point charges, there is a path along which the charge can be moved to infinity without ever needing positive external work, i.e. a path along which the potential only decreases.

Solution. (a) Fix some point O not on any of the charges, and scale the system up about O continuously, to send all the charges to infinity. At all points in time, there are no forces on any of the charges, so no work is done. The final potential energy is zero, so the initial potential energy must also have been zero.

(b) Consider the field line going through the test charge. It can't end on a negative charge, since there are none, so it must end at infinity. Moving the charge along this field line gives the desired path.

Idea 3

Gauss's law is written in differential form as

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}.$$

The divergence of a vector field $\mathbf{F} = F_x\hat{\mathbf{x}} + F_y\hat{\mathbf{y}} + F_z\hat{\mathbf{z}}$ is

$$\nabla \cdot \mathbf{F} = \partial_x F_x + \partial_y F_y + \partial_z F_z$$

in Cartesian coordinates, where ∂_x stands for $\partial/\partial x$, and so on.

Example 2

Show that the two forms of Gauss's law are equivalent.

Solution

To do this, we need to establish the geometric meaning of the divergence. For simplicity we consider two dimensions; the proof for three dimensions is similar. Consider a small rectangle prism with one corner at the origin, with axes aligned with the Cartesian coordinate axes and side lengths Δx and Δy. To apply Gauss's law in integral form, we need to compute the flux through each side. The flux going out the top side is

$$\int_0^{\Delta x} E_y(x, \Delta y)\, dx$$

while the flux going out the bottom side is

$$-\int_0^{\Delta x} E_y(x, 0)\, dx.$$

The sum of these two terms is

$$\int_0^{\Delta x} (E_y(x, \Delta y) - E_y(x, 0))\, dx \approx \Delta y \int_0^{\Delta x} (\partial_y E_y)|_{(x,0)}\, dx$$

where we applied a tangent line approximation, and the subscript indicates where the function $\partial_y E_y$ is evaluated. Higher-order terms in the Taylor series would be proportional to higher powers of Δy, which is small, so we can ignore them.

The integrand is still a function of x, but we can Taylor expand it about the origin as

$$(\partial_y E_y)|_{(x,0,0)} = (\partial_y E_y)|_{(0,0,0)} + \Delta x(\ldots) + \ldots.$$

These extra terms are again higher-order in Δx and Δy, so we ignore them. The net flux through the top and bottom faces is hence, to lowest order,

$$\Delta y \int_0^{\Delta x} (\partial_y E_y)|_{(0,0,0)}\, dx = \Delta x \Delta y\, (\partial_y E_y)|_{(0,0,0)}.$$

By similar reasoning, pairing up the left and right faces gives

$$\text{flux} = \Delta x \Delta y\, (\partial_x E_x + \partial_y E_y)|_{(0,0,0)} = \Delta x \Delta y\, (\nabla \cdot \mathbf{E})|_{(0,0,0)}.$$

Thus the divergence is the outgoing flux per unit area, or volume in three dimensions.

This shows us why the two forms of Gauss's law area equivalent. For example, starting from the differential form, the left-hand side is the flux per volume, while the right-hand side is the charge per volume, divided by ϵ_0. Integrating both sides over some volume relates the total flux to the total charge divided by ϵ_0, which is Gauss's law in integral form.

If the above derivation was a bit abstract, we can also show the idea using specific examples.

Example 3

Suppose the region $0 < x < d$ has charge density $-\rho$, and the region $-d < x < 0$ has charge density ρ. Find the electric field everywhere.

Solution

By translational symmetry, the field always points along $\hat{\mathbf{x}}$ and only depends on x, $\mathbf{E}(\mathbf{r}) = E(x)\,\hat{\mathbf{x}}$. By applying the integral form of Gauss's law to a rectangular prism, with one side at x_l and another at x_r, we have

$$E(x_r) - E(x_l) = \frac{1}{\epsilon_0}\int_{x_l}^{x_r} \rho(x)\,dx, \quad E(x) = \frac{1}{\epsilon_0}\int_0^x \rho(x)\,dx + E_0.$$

Since the divergence of $\mathbf{E}(\mathbf{r})$ is just $\partial E(x)/\partial x$, this clearly satisfies the differential form of Gauss's law. To fix the undetermined constant E_0, we could demand the field be zero on both sides of the charge distribution, motivated by symmetry. Then we have

$$E(x) = \frac{\rho}{\epsilon_0} \times \begin{cases} d - x & 0 < x < d, \\ d + x & -d < x < 0, \\ 0 & \text{elsewhere.} \end{cases}$$

Example 4

Find the electric field of a spherically symmetric charge density $\rho(r)$.

Solution

By spherical symmetry, the field always points radially and only depends on r, $\mathbf{E}(\mathbf{r}) = E(r)\,\hat{\mathbf{r}}$. By applying the integral form of Gauss's law to a sphere of radius r,

$$4\pi r^2 E(r) = \frac{1}{\epsilon_0}\int_0^r dr'\,4\pi r'^2 \rho(r'), \quad E(r) = \frac{1}{\epsilon_0}\frac{1}{r^2}\int_0^r dr'\,r'^2\rho(r').$$

Let's check that this indeed satisfies the differential form of Gauss's law, using the divergence in spherical coordinates. For any vector field $\mathbf{F} = F_r\hat{\mathbf{r}} + F_\theta\hat{\boldsymbol{\theta}} + F_\varphi\hat{\boldsymbol{\varphi}}$, the divergence is

$$\nabla \cdot \mathbf{F} = \frac{1}{r^2}\frac{\partial(r^2 F_r)}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(F_\theta \sin\theta) + \frac{1}{r\sin\theta}\frac{\partial F_\varphi}{\partial\varphi}.$$

This looks complicated, but things turn out simple because $\mathbf{E}$ only has a radial component, $E_r = E(r)$, which gives

$$\nabla \cdot \mathbf{E} = \frac{1}{r^2}\frac{\partial(r^2 E(r))}{\partial r} = \frac{1}{r^2\epsilon_0}\frac{\partial}{\partial r}\int_0^r dr'\,r'^2\rho(r') = \frac{r^2\rho(r)}{r^2\epsilon_0} = \frac{\rho(r)}{\epsilon_0}$$

just as desired.

[3] **Problem 5.** Consider a vector field expressed in polar coordinates, $\mathbf{F} = F_r\hat{\mathbf{r}} + F_\theta\hat{\boldsymbol{\theta}}$ where $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ are unit vectors in the radial and tangential directions. Gauss's law in differential form still works in these coordinates, but the form of the divergence is different.

By considering the flux per unit area out of a small region bounded by r and $r + dr$, and θ and $\theta + d\theta$, and applying Gauss's law in integral form, find what the divergence in polar coordinates must be for Gauss's law in differential form to hold. (Optional: try generalizing to spherical coordinates.)

Solution. By summing up contributions from each of the four sides, and letting (F_r, F_θ) be the vector field at one of the corners, the flux through the region is

$$d\Phi = (F_r + dF_r)((r + dr)d\theta) - F_r(r d\theta) + (F_\theta + dF_\theta)dr - F_\theta dr.$$

In two dimensions, the divergence is the flux per area, $dA = r\,dr\,d\theta$, so

$$\nabla \cdot \mathbf{F} = \frac{d\Phi}{dA} = \frac{1}{r}\frac{\partial(rF_r)}{\partial r} + \frac{1}{r}\frac{\partial F_\theta}{\partial \theta}.$$

[4] **Problem 6.** This problem is quite subtle, but will enhance your understanding of electromagnetism. Suppose that all of space is filled with uniform charge density ρ.

(a) Show that $\mathbf{E} = (\rho/\epsilon_0)x\hat{\mathbf{x}}$ obeys the differential form of Gauss's law.

(b) Show that $\mathbf{E} = (\rho/3\epsilon_0)r\hat{\mathbf{r}}$ also obeys Gauss's law.

(c) Argue that by symmetry, $\mathbf{E} = 0$. Show that this does not obey Gauss's law.

(d) $\star$ What's going on? Which, if any, is the actual field? If you think there's more than one possible field, how could that be consistent with Coulomb's law, which gives the answer explicitly? For that matter, what does Coulomb's law say about this setup, anyway?

Solution. (a) We see that $\nabla \cdot \mathbf{E} = \partial_x((\rho/\epsilon_0)x) = \rho/\epsilon_0$, as desired.

(b) In Cartesian coordinates, this field is

$$\mathbf{E} = \frac{\rho}{3\epsilon_0}(x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}})$$

whose divergence is ρ/ϵ_0, as desired.

(c) This has to hold by symmetry because the electric field can't point in any particular direction, by rotational symmetry. It also can't just point radially, because that breaks translational symmetry; the center is a special point. So the only option is $\mathbf{E} = 0$, but $\nabla \cdot \mathbf{E} = 0$, so Gauss's law is not obeyed.

(d) The issue is *boundary conditions*. Just like any differential equation, the solution for the electric field is not defined without boundary conditions (or initial conditions, as we called them in mechanics). Usually, we get a unique solution by demanding the fields go to zero at infinity. (Though in some cases, this might not be the right physical answer. For example, the electric field of a capacitor in the lab doesn't have to be zero outside, because it might be inside some bigger capacitor.) However, we can't do this here because the charge density extends out to infinity too. By taking different choices of boundary conditions, we can get (a), (b), or many other answers. The symmetry argument in (c) fails, because *any* choice of boundary conditions will break the perfect translational symmetry.

At first glance, it could seem that Coulomb's law could give us a unique answer. Coulomb's law for a point charge is itself derived by implicitly assuming that there are no "extra" fields flying around, just the spherically symmetric field of the point charge itself. This looks very reasonable, so what stops us from just saying that each charge in this problem has such a field, and then integrating over the charges? Well, if you write down the integral, you'll find that it's divergent, analogous to the integral $\int_{-\infty}^{\infty} x\,dx$. By itself, the integral is not even well-defined.

In order to get an answer, you have to "regulate" the integral (i.e. change it in a way that makes it well-defined). One possible regulator, for example, is to just chop off the limits of integration at finite values, like $\int_{-L}^{L} x\,dx$. But that particular regulator is equivalent to just replacing the charge distribution with a finite one centered at the origin! In other words, Coulomb's law also fails to give a unique answer, because it requires a regulator to give a well-defined answer, and there are many possible regulators. If you treat the charge distribution as a giant ball with center at the origin, you get the result of part (b). If you treat it as a thick, huge slab along the yz plane centered at the origin, you get the result of part (a). The symmetry argument fails once again, because all the regulators break translational symmetry. This is a simple example of an "anomalous symmetry", an important idea in theoretical physics.

The exact same problem appears in Newtonian cosmology, where charge density is replaced with mass density, and this problem confused Newton himself, who incorrectly thought that $\mathbf{g} = 0$ by symmetry. In this context, *all* regulators/boundary conditions are unsatisfactory. Of course, we want a rotationally symmetric universe to match experiment, so we have to put that in by hand. But then every solution has a center towards which everything collapses, so to keep the solar system an inertial frame, we'd have to put it at the center of the universe! Surely, this would make Copernicus roll in his grave.

Some of these problems are fixed in general relativity. You still have to postulate rotational symmetry (again, on the basis of experimental data), but once you do that, there are no further problems. That's because in general relativity, acceleration is not absolute in the way it is in Newtonian mechanics. Instead, there is no center; everything just gets closer to everything else. For further discussion and references, see this paper.

Idea 4

A tricky, occasionally useful idea is to use Newton's third law: it may be easier to calculate the force of A on B than the force of B on A.

Example 5: Purcell 1.28

Consider a point charge q. Draw any imaginary sphere of radius R around the charge. Show that the average of the electric field over the surface of the sphere is zero.

Solution

Imagine placing a uniform surface charge σ on the sphere. Then the average of the point charge's electric field over the sphere times $4\pi R^2\sigma$ is the total force of the point charge on the charged sphere. But this is equal in magnitude to the force of the charged sphere on the point charge, which must be zero by the shell theorem. Thus the average field over the sphere has to vanish.

Example 6

Consider two spherical balls of charge q and radii a_i, with their centers separated by a distance $r > a_1 + a_2$. What is the net force of the first on the second?

Solution

It might seem obvious that the answer is $q^2/4\pi\epsilon_0 r^2$, with no dependence on a_1 and a_2. In fact, if you've done any orbital mechanics, you've almost certainly assumed that the force between two spherical bodies (such as the Earth and Sun) is Gm_1m_2/r^2, which is equivalent.

This has a simple but slightly tricky proof. By the shell theorem, we can set $a_1 = 0$, replacing the first ball with a point charge, because this produces the same field at the second ball. But the force on the second ball depends on the electric field at every point on it, which seems to require doing an integral. To avoid this, we use Newton's third law, which tells us it's equivalent to compute the force on the first ball. To compute *that*, we may set $a_2 = 0$ by the shell theorem again. This reduces us to the case of two point charges, giving the answer.

[3] **Problem 7** (Purcell 1.28)**.** Some extensions of the previous example.

(a) Show that if the charge q is instead outside the sphere, a distance $r > R$ from its center, the average electric field over the surface of the sphere is the same as the electric field at the center of the sphere.

(b) Show that for any overall neutral charge distribution contained within a sphere of radius R, the average electric field over the interior of the sphere is $-\mathbf{p}/4\pi\epsilon_0 R^3$ where $\mathbf{p}$ is the total dipole moment.

Solution. The same Newton's third law trick will work for both parts.

(a) Let the desired answer be $\mathbf{E}_{\text{avg}}$ and let the charge q be at $\mathbf{r}$. Now imagine a charge Q is uniformly distributed over the surface of the sphere. The force of the charge q on the distributed charge Q is precisely $\mathbf{F}_{qQ} = Q\mathbf{E}_{\text{avg}}$. But we also know that

$$\mathbf{F}_{qQ} = -\mathbf{F}_{Qq} = -\frac{kQq}{r^2}\hat{\mathbf{r}}$$

by Newton's third law and the shell theorem. Therefore we have

$$\mathbf{E}_{\text{avg}} = -\frac{kq}{r^2}\hat{\mathbf{r}}$$

which is precisely the electric field at the center of the sphere due to q. (Note that $\hat{\mathbf{r}}$ points from the center of the sphere to the charge q.)

(b) Let the desired answer be $\mathbf{E}_{\text{avg}}$. Now imagine a charge Q is uniformly distributed over the volume of the sphere. The force of the charge distribution (with charge density $\rho(\mathbf{x})$) on the distributed charge Q is precisely $\mathbf{F}_{qQ} = Q\mathbf{E}_{\text{avg}}$. But we also know that

$$\mathbf{F}_{qQ} = -\mathbf{F}_{Qq} = -\int \rho(\mathbf{r})\mathbf{E}_Q(\mathbf{r})\, d^3\mathbf{r}$$

where $\mathbf{E}_Q$ is the field due to Q. Now, this field is easy to find, as it is just the field of a uniformly charged sphere, so

$$\mathbf{E}_Q = \frac{kQ}{R^3}\mathbf{r}$$

as shown in problem 1. Putting this in the integral, we have

$$Q\mathbf{E}_{\text{avg}} = -\frac{kQ}{R^3}\int \rho(\mathbf{r})\mathbf{r}\, d^3\mathbf{r}$$

so by the definition of the dipole moment,

$$\mathbf{E}_{\text{avg}} = -\frac{k}{R^3}\int \rho(\mathbf{r})\mathbf{r}\, d^3\mathbf{r} = -\frac{k\mathbf{p}}{R^3}$$

as desired.

[3] **Problem 8.** There are two point charges, $q_1 > 0$ and $q_2 < 0$, in empty space. An electric field line leaves q_1 at an angle α from the line connecting the two charges. Determine whether this field line hits q_2, and if so, at what angle β from the line connecting the two charges. (Hint: this can be done without solving any differential equations.)

Solution. Suppose the field line does hit q_2. Rotate the field line about the line connecting the two charges, to form a Gaussian surface. Because no electric field lines go across this surface, the total charge inside must be zero. Now, this surface envelopes "slices" of each point charge. (If you're not happy with "slicing a point charge", just replace the point charges with tiny uniformly charged spheres; everything stays the same.) The solid angle of the first point charge enveloped is

$$\int d\Omega = \int_0^{2\pi} d\phi \int_0^{\alpha} \sin\theta\, d\theta = 2\pi(1-\cos\alpha)$$

so the amount of charge enclosed is

$$\frac{\Omega}{4\pi}\, q_1 = \frac{1-\cos\alpha}{2}\, q_1 = q_1 \sin^2\frac{\alpha}{2}.$$

Reasoning similarly for the other surface, we have

$$q_1 \sin^2\frac{\alpha}{2} = |q_2| \sin^2\frac{\beta}{2}$$

and the field line hits q_2 if there is a solution for β, i.e. when $|q_1/q_2|\,\sin^2(\alpha/2) < 1$.

Idea 5

The integral $\int d\mathbf{S}$ over a surface with a fixed boundary is independent of the surface.

We proved this in a mechanical way in **M2**. If you want to see a proof using vector calculus, see problem 1.62 of Griffiths.

[3] **Problem 9.** A hemispherical shell of radius R has uniform charge density σ and is centered at the origin. Find the electric field at the origin. (Hint: combine the previous two ideas.)

Solution. Place a point charge q at the origin. To find the magnitude of the field, we will compute the force on the hemisphere divided by q. The force on the hemisphere is

$$\int \frac{q}{4\pi\epsilon_0 R^2}\,\sigma\,d\mathbf{S} = \frac{q\sigma}{4\pi\epsilon_0 R^2}\int d\mathbf{S}.$$

By idea 5, we can replace the surface of integration with a flat disk, so $|\int d\mathbf{S}| = \pi R^2$. Thus, the force is $F = q\sigma/4\epsilon_0$, so the field is

$$E = \frac{\sigma}{4\epsilon_0}.$$

[3] **Problem 10.** A point charge q is placed a distance $a/2$ above the center of a square of charge density σ and side length a. Find the force of the square on the point charge.

Solution. This is a tricky problem, whose solution uses a one-time trick. It's equivalent to find the force of the point charge on the square. Set up coordinates so that the square is in the xy plane, and its center is the origin. Then we have

$$\mathbf{F} = \sigma\int \mathbf{E}\,dS$$

where the surface integral is over the square. On the other hand, we know that $\mathbf{F}$ is along the $\hat{\mathbf{z}}$ direction by symmetry, so

$$F = \mathbf{F}\cdot\hat{\mathbf{z}} = \sigma\int E_z\,dS.$$

Now, since $d\mathbf{S}$ is parallel to $\hat{\mathbf{z}}$, this is in fact the same thing as

$$F = \sigma\int \mathbf{E}\cdot d\mathbf{S}$$

where the integral is just the electric flux through the square! By symmetry, this flux is $q/6\epsilon_0$, so

$$F = \frac{\sigma q}{6\epsilon_0}.$$

[4] **Problem 11** (Griffiths 2.47, PPP 113, MPPP 140)**.** Consider a uniformly charged spherical shell of radius R and total charge Q.

(a) Find the net electrostatic force that the southern hemisphere exerts on the northern hemisphere.

(b) Generalize part (a) to the case where the sphere is split into two parts by a plane whose minimum distance to the sphere's center is h.

(c) Generalize part (a) to the case where the two hemispherical shells have uniform charge density, opposite orientation, and the same center, but have different total charges q and Q, and different radii r and R, where $r < R$.

Hint: see example 10, and use superposition and symmetry when applicable.

Solution. (a) The net force that the northern hemisphere exerts on itself is 0, so it is equivalent to find the force on the north due to the entire sphere. The surface charge density is $\sigma =$

$Q/(4\pi R^2)$. By the result of example 10, the outward pressure on the northern hemisphere is $\sigma^2/2\epsilon_0$. Therefore, the total force is

$$F = \left| \frac{\sigma^2}{2\epsilon_0} \int_N d\mathbf{S} \right| = \frac{\sigma^2}{2\epsilon_0}(\pi R^2) = \frac{Q^2}{32\pi\epsilon_0 R^2}$$

where N refers to the northern hemisphere, and the surface integral was done as in problem 9.

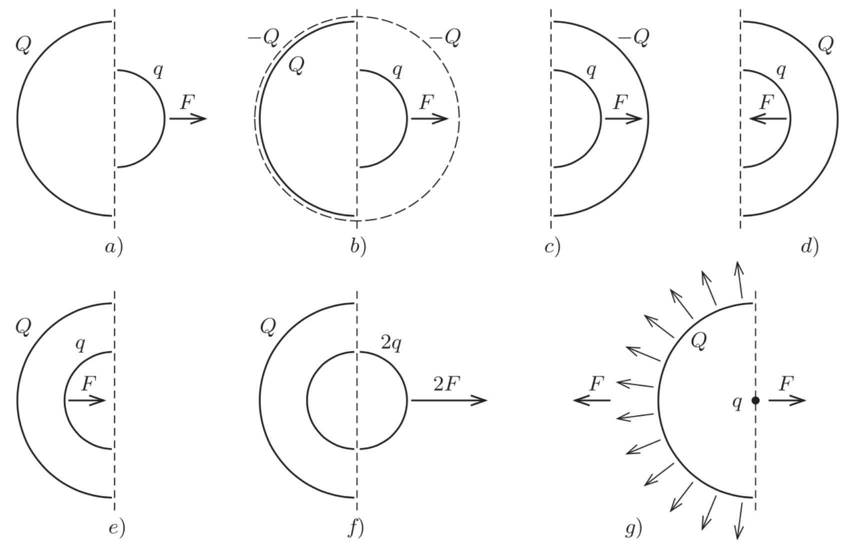
(b) This is exactly the same as in part (a), except that now the integral over the piece is

$$\left| \int d\mathbf{S} \right| = \pi(R^2 - h^2)$$

which gives the result

$$F = (\sigma^2/2\epsilon_0)\pi(R^2 - h^2) = \frac{Q^2}{32\pi\epsilon_0 R^2}(1 - h^2/R^2).$$

(c) This can be solved using an ingenious superposition and symmetry argument.



The force we want to compute is shown in (a). Now consider superposing a uniformly negatively charged sphere with radius just larger than R, as shown in (b). By the shell theorem, this doesn't change the force on the hemisphere of radius r. The result of the superposition is (c). Flipping the charge of one of the hemispheres in (c) flips the force, leading to (d). Finally, reflecting (d) gives (e).

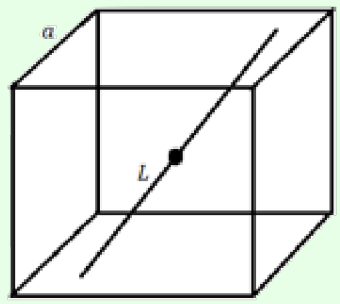
This has all been preamble to the ingenious step: superpose (a) and (e) to get (f), which involves the force on a *complete* sphere of radius r. Using Newton's third law, $2F$ can now be computed by finding the force on the hemisphere. But that is easy because of the shell theorem, which tells us that F is the net force on the hemisphere shown in (g). Using the method of problem 9 again, we conclude

$$F = \frac{q}{4\pi\epsilon_0 R^2}(\pi R^2)\frac{Q}{2\pi R^2} = \frac{Qq}{8\pi\epsilon_0 R^2}$$

which is independent of r! (Setting $r = 0$ and $r = R$ recovers the answers to two previous problems.) By the way, this problem illustrates why you can't trust online solutions. If you google it, you'll find mostly incorrect answers which are in turn copied from an incorrect solution in a JEE book.

Example 7: IdPhO 2020.1A

A point charge of mass m and charge $-q$ is placed at the center of a cube with side length a, whose volume has uniform charge density ρ. The point charge is allowed to slide along a straight line, which has an arbitrary orientation, so that the distance along the line from the center to one of the cube's faces is L.



Find the angular frequency of small oscillations.

Solution

The official solution goes as follows: consider displacing the point charge by some small amount $\Delta\mathbf{r}$. The cube of charge can then be decomposed into (1) a slightly smaller cube of charge centered around the point charge's new position, and (2) three thin plates of charge on the faces opposite to the charge's motion. By symmetry, (1) contributes nothing, and we know what (2) contributes from the answer to problem 10. The result is a restoring force proportional to $-\Delta\mathbf{r}$, whose magnitude has no dependence on the orientation of $\Delta\mathbf{r}$, so the oscillation frequency doesn't depend on L. Once you know this, you can orient the line any way you want, so the problem is simple to finish.

Personally, I don't like this problem because the intended solution requires knowing the answer to problem 10, which itself is pretty tricky. That is, the difficulty of the problem depends mostly on whether you've seen that tough, but standard problem elsewhere. However, I'm including it as an example because there's another way to solve it, which is a bit more advanced, but quite illustrative.

Since this is a question about small oscillations, it suffices to expand the potential energy to second order about the center of the cube. The most general possible expression is

$$V(x,y,z) = a + b_1x + b_2y + b_3z + c_1x^2 + c_2y^2 + c_3z^2 + c_4xy + c_5yz + c_6xz + O(r^3).$$

The constant a doesn't matter, so we can just ignore it. And since $\mathbf{E}$ vanishes at the center, the linear terms b_i are all zero as well. Because the x, y, and z axes are all equivalent by cubical symmetry (e.g. we can rotate them into each other, while keeping the cube the same),

$$c = c_1 = c_2 = c_3, \quad c' = c_4 = c_5 = c_6.$$

Thus, our complicated Taylor series boils all the way down to

$$V(x, y, z) = c(x^2 + y^2 + z^2) + c'(xy + yz + xz) + O(r^3)$$

without even having to do any work! Finally, notice that the cube is symmetric under reflections $x \to -x$, $y \to -y$, or $z \to -z$. These reflections keep the c term the same, but flip the c' term. Therefore, we must have $c' = 0$, so

$$V(r) = cr^2 + O(r^3)$$

which is remarkably simple. The potential near the origin is spherically symmetric, even though the setup as a whole isn't! It's not automatic: it wouldn't be this simple if we had a slightly more complex shape. This "accidental" spherical symmetry is a consequence of the combination of cubical symmetry and the simplicity of Taylor series.

Therefore, to finish the problem we only need to find the coefficient c. While there are simpler ways to do this, I'll do it in a way that introduces some useful facts. Combining the definition of V and Gauss's law, we have

$$\nabla \cdot (\nabla V) = -\nabla \cdot \mathbf{E} = -\frac{\rho}{\epsilon_0}.$$

This is a standard and fundamental result in electrostatics, called Poisson's equation, which we will see again later. The divergence of a gradient is also called a Laplacian, and written as

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -\frac{\rho}{\epsilon_0}.$$

Using this, we can easily compute the value of c, giving

$$V(\mathbf{r}) = -\frac{\rho r^2}{6\epsilon_0} + O(r^3).$$

Therefore, for a displacement $\Delta\mathbf{r}$ in *any* direction, the restoring force is $\rho q r/3\epsilon_0$ in the opposite direction, which means

$$\omega = \sqrt{\frac{\rho q}{3\epsilon_0 m}}$$

independent of the orientation of the line.

Remark

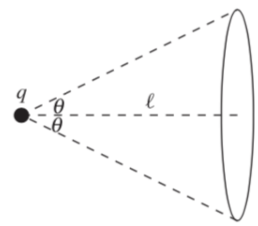
Accidental symmetry is important in modern physics. For example, protons are stable because of an accidental symmetry in the Standard Model, which ensures that baryon number is conserved. That explains why we often expect proton decay to occur in extensions of the Standard Model, such as grand unified theories, as explained in this nice article.

2 Continuous Charge Distributions

Idea 6

In almost all cases in Olympiad physics, there will be sufficient symmetry to reduce any multiple integral to a single integral. Remember that when using Gauss's law, the Gaussian surface may be freely deformed as long as it doesn't pass through any charges.

[2] **Problem 12** (Purcell 1.15). A point charge q is located at the origin. Compute the electric flux that passes through a circle a distance ℓ from q, subtending an angle 2θ as shown below.



Solution. Let $\ell = R\cos\theta$, and deform the disk into a spherical cap with radius R. Then the answer is then just kq/ϵ_0, where k is the ratio of the area of the cap to the total area of the sphere. In spherical coordinates,

$$k = \frac{1}{4\pi}\int_0^\theta 2\pi\sin\theta\,d\theta = \frac{1-\cos\theta}{2}$$

so the answer is

$$\frac{1-\cos\theta}{2}\frac{q}{\epsilon_0}.$$

You can also show this using the original flat Gaussian surface, though that takes more work.

[3] **Problem 13** (Purcell 1.8). A ring with radius R has uniform positive charge density λ. A particle with positive charge q and mass m is initially located in the center of the ring and given a tiny kick. If the particle is constrained to move in the plane of the ring, show that it exhibits simple harmonic motion and find the angular frequency.

Solution. Suppose it is moved by $r \ll R$ in the x direction. Set up polar coordinates with $\theta = 0$ being the positive x axis. By the law of cosines, we have

$$\begin{aligned} U(r) &= 2\int_0^\pi \frac{1}{4\pi\epsilon_0}\frac{q(\lambda R d\theta)}{\sqrt{R^2+r^2-2Rr\cos\theta}} \\ &= \frac{q\lambda}{2\pi\epsilon_0}\int_0^\pi \frac{d\theta}{\sqrt{1+(r^2/R^2)-2(r/R)\cos\theta}}. \end{aligned}$$

Next, we can expand the square root using a Taylor series. If we expand to first order in r/R, then the result will be proportional to the integral of $\cos\theta$, which vanishes. Thus, to get the leading contribution we must expand to second order, giving

$$\begin{aligned} U(r) &= \frac{q\lambda}{2\pi\epsilon_0}\int_0^\pi \left[1 - \frac{1}{2}\frac{r^2}{R^2} + \frac{3}{8}\left(-\frac{2r}{R}\cos\theta\right)^2\right] d\theta \\ &= \frac{q\lambda}{2\pi\epsilon_0}\int_0^\pi \frac{r^2}{2R^2}(3\cos^2\theta - 1)\,d\theta + \text{const} \\ &= \frac{q\lambda r^2}{8\epsilon_0 R^2} + \text{const}. \end{aligned}$$

This is essentially the same calculation as an example in **M6**. We conclude that the effective spring constant is $k = q\lambda/4\epsilon_0 R^2$, so

$$\omega = \sqrt{\frac{q\lambda}{4m\epsilon_0 R^2}}.$$

[3] **Problem 14** (Purcell 1.12)**.** Consider the setup of problem 9. If the hemisphere is centered at the origin and lies entirely above the xy plane, find the electric field at an arbitrary point on the z-axis. (This is a bit complicated, and is representative of the most difficult kinds of integrals you might have to set up in an Olympiad. For a useful table of integrals, see Appendix K of Purcell.)

Solution. Set up spherical coordinates with the hemisphere being the equation of $r = R$ and $\theta \in [0, \pi/2]$. Suppose our location is $(0, 0, z)$. The hemisphere has surface charge σ. We see that the field points in the z-direction by symmetry, so we'll only worry about that piece. The ring at angle θ with width $d\theta$ provides fields at an angle, and some geometry shows that we have to correct by a factor of $\frac{R\cos\theta - z}{r}$ where $r \equiv \sqrt{R^2 + r^2 - 2Rz\cos\theta}$. We then have

$$dE_z = -\frac{\sigma(2\pi R^2 \sin\theta\, d\theta)}{4\pi\epsilon_0 r^2} \cdot \frac{R\cos\theta - z}{r},$$

so

$$E(z) = -\frac{\sigma R^2}{2\epsilon_0} \int_0^{\pi/2} \frac{(R\cos\theta - z)\sin\theta\, d\theta}{(R^2 + r^2 - 2Rz\cos\theta)^{3/2}}.$$

Consulting Appendix K tells us that

$$E(z) = \frac{\sigma R^2}{2\epsilon_0 z^2} \left(\frac{R}{\sqrt{R^2 + z^2}} - \frac{R - z}{\sqrt{(R - z)^2}} \right).$$

Taking some care with the square root, we conclude

$$E(z) = \frac{\sigma R^2}{2\epsilon_0 z^2} \times \begin{cases} \frac{1}{\sqrt{1 + z^2/R^2}} - 1 & z < R \\ \frac{1}{\sqrt{1 + z^2/R^2}} + 1 & z > R \end{cases}.$$

[3] **Problem 15.** USAPhO 2018, problem B1.

Idea 7: Electric Dipoles

The dipole moment of two charges q and $-q$ separated by $\mathbf{d}$ is $\mathbf{p} = q\mathbf{d}$. More generally, the dipole moment of a charge configuration is defined as

$$\mathbf{p} = \int \rho(\mathbf{r})\mathbf{r}\, d^3\mathbf{r}.$$

For an overall neutral charge configuration, the leading contribution to its electric potential far away is the dipole potential,

$$\phi(r, \theta) = \frac{p\cos\theta}{4\pi\epsilon_0 r^2}$$

where θ is the angle of $\mathbf{r}$ to $\mathbf{p}$.

Remark

Here's a trick to remember the dipole potential. Let $\phi_0(\mathbf{r}) = k/r$ be the potential for a unit charge at the origin. An ideal point dipole of dipole moment p consists of charges $\pm p/d$ separated by d, in the limit $d \to 0$. So the potential is

$$p \lim_{\mathbf{d}\to 0} \frac{\phi_0(\mathbf{r}) - \phi_0(\mathbf{r}+\mathbf{d})}{d}.$$

But this is precisely the (negative) derivative, so you can get the dipole potential by differentiating the ordinary potential! Indeed, for a dipole aligned along the $\hat{\mathbf{z}}$ axis,

$$-\frac{\partial}{\partial z}\frac{kp}{r} = \frac{kp}{r^2}\frac{\partial r}{\partial z} = \frac{kp}{r^2}\frac{z}{r} = \frac{kp\cos\theta}{r^2}$$

which matches the above result. You can use the same trick for quadrupoles and higher multipoles, which we'll see in **E8**.

[3] **Problem 16.** In this problem we'll derive essential results about dipoles, which will be used later.

(a) Using the binomial theorem, derive the dipole potential given above, for a dipole made of a pair of point charges $\pm q$ separated by distance d, oriented along the z-axis.

(b) Differentiate this result to find the dipole field,

$$\mathbf{E}(\mathbf{r}) = \frac{p}{4\pi\epsilon_0 r^3}(2\cos\theta\,\hat{\mathbf{r}} + \sin\theta\,\hat{\boldsymbol{\theta}})$$

where the expression above is in spherical coordinates. (Hint: feel free to use the expression for the gradient in spherical coordinates.)

(c) Show that this may also be written as

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0 r^3}(3(\mathbf{p}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}} - \mathbf{p}).$$

You don't need to memorize these expressions, but it's useful to remember what a dipole field looks like, the fact that its magnitude is roughly $p/4\pi\epsilon_0 r^3$, and the fact that the numeric prefactor is 2 along the dipole's axis and 1 perpendicular to it.

Solution. (a) Let the charges be at $(0,0,0)$ and $(0,0,d)$. Then

$$V(r,\theta) = \frac{q}{4\pi\epsilon_0 r}\left(-1 + \frac{1}{\sqrt{1-2(d/r)\cos\theta + (d/r)^2}}\right) \approx \frac{qd\cos\theta}{4\pi\epsilon_0 r^2}.$$

(b) We use the definition $\mathbf{E} = -\nabla V$, along with the gradient in spherical coordinates. Then

$$E_r = -\frac{\partial V}{\partial r} = \frac{p}{4\pi\epsilon_0 r^3}\cdot 2\cos\theta$$

and

$$E_\theta = -\frac{1}{r}\frac{\partial V}{\partial\theta} = \frac{p}{4\pi\epsilon_0 r^3}\cdot\sin\theta,$$

as desired.

(c) We see that $\mathbf{p} \cdot \hat{\mathbf{r}} = p\cos\theta$ and $\mathbf{p} = p\hat{\mathbf{z}} = p(\hat{\mathbf{r}}\cos\theta - \hat{\boldsymbol{\theta}}\sin\theta)$. Thus,

$$3(\mathbf{p}\cdot\hat{\mathbf{r}})\,\hat{\mathbf{r}} - \mathbf{p} = 3p\cos\theta\,\hat{\mathbf{r}} - p(\hat{\mathbf{r}}\cos\theta - \hat{\boldsymbol{\theta}}\,\sin\theta) = p(2\cos\theta\,\hat{\mathbf{r}} + \sin\theta\,\hat{\boldsymbol{\theta}}),$$

as desired.

[3] **Problem 17.** USAPhO 2002, problem B2.

[3] **Problem 18.** USAPhO 2009, problem B2. This essential problem introduces useful facts about dipole-dipole interactions.

Idea 8

The potential energy of a set of point charges is

$$U = \frac{1}{4\pi\epsilon_0}\sum_{i\neq j}\frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|} = \frac{1}{2}\sum_i q_i V(\mathbf{r}_i).$$

We sum over $i \neq j$ to avoid computing the energy of a single point charge due to its interaction with itself, which would be infinite. For a continuous distribution of charge, we don't have this problem, and instead find

$$U = \frac{1}{2}\int \rho(\mathbf{r})V(\mathbf{r})\,d^3\mathbf{r} = \frac{\epsilon_0}{2}\int |\mathbf{E}(\mathbf{r})|^2\,d^3\mathbf{r}.$$

Unlike the other quantities we've considered, energy doesn't obey the superposition principle.

[3] **Problem 19.** In this problem we'll apply the above results to balls of charge.

(a) Compute the potential energy of a uniformly charged ball of total charge Q and radius R.

(b) Show that the potential energy of two point charges of charge $Q/2$ separated by radius R is lower than the result of part (a).

(c) Hence it appears that it is energetically favorable to compress a ball of charge into two point charges. Is this correct?

Solution. (a) We can find the potential by building up the ball by placing charges from infinity. Consider a shell of charge at radius r, and let the charge density by $\rho = Q/(\frac{4}{3}\pi R^3)$. The energy needed to put the shell there is $dU = kQ_{enc}\,dQ/r$, where $Q_{enc} = \frac{4}{3}\rho\pi r^3$ is the charge inside and $dQ = 4\rho\pi r^2\,dr$ is the charge in the shell added to the sphere. Then the energy needed to build the ball, which is the potential energy of the ball, is

$$U_a = \int_0^R kQ\frac{r^3}{R^3}(3Qr^2\,dr/R^3)/r = \frac{3kQ^2}{R^6}\int_0^R r^4 d\,r = \frac{3kQ^2}{5R} = \frac{3Q^2}{20\pi\epsilon_0 R}.$$

(b) From $U = kq_1q_2/r$, we find that for two point charges the potential energy is

$$U_b = \frac{kQ^2}{4R} = \frac{Q^2}{16\pi\epsilon_0 R}$$

which is less than U_a.

(c) It's wrong because in part (b), the energy needed to create the point charges, by squeezing the two halves of the ball down, is not included. Plugging in a radius of zero into part (a), we see that this energy is actually infinite. (Of course, in reality it doesn't take infinite energy to produce electrons, which are point charges. Classical electrodynamics breaks down when describing such a process, which can only be properly understood within relativistic quantum field theory.)

[3] **Problem 20.** An insulating circular disc of radius R has uniform surface charge density σ.

(a) Find the electric potential on the rim of the disc.

(b) Find the total electric potential energy stored in the disc.

Solution. (a) Place the origin at a point on the rim and use polar coordinates. Because the polar equation of a circle is $r = 2R\cos\theta$, we have

$$V = \int_{-\pi/2}^{\pi/2} d\theta \int_0^{2R\cos\theta} \frac{\sigma}{4\pi\epsilon_0}\, dr = \int_{-\pi/2}^{\pi/2} \frac{\sigma R}{2\pi\epsilon_0} \cos\theta\, d\theta = \frac{\sigma R}{\pi\epsilon_0}.$$

(b) Consider building up the ring outward in radius. When we add charges to bring the radius from r to $r + dr$, we do work

$$dW = V\, dq = \frac{\sigma r}{\pi\epsilon_0}(2\pi r\sigma\, dr) = \frac{2\sigma^2 r^2}{\epsilon_0} dr$$

which means

$$W = \int_0^R \frac{2\sigma^2 r^2}{\epsilon_0}\, dr = \frac{2\sigma^2 R^3}{3\epsilon_0}.$$

[3] **Problem 21.** Consider a uniformly charged ball of total charge Q and radius R. Decompose this ball into two parts, A and B, where B is a ball of radius $R/2$ whose center is a distance $R/2$ of the ball's center, and A is everything else. Find the potential energy due to the interaction of A and B, i.e. the work necessary to bring in B from infinity, against the field of A.

Solution. If we tried to compute the potential energy directly, by integrating over A and B, we would get messy integrals. Instead, let's consider bringing in B in three steps:

1. At infinity, compress B into a point charge $Q/8$.
2. Move this point charge to the center of the B-shaped hole in A.
3. Expand the point charge back into the original shape of B.

Our first claim is that the total work needed to do steps (1) and (3) is zero. These two steps are very close to being opposites; the only difference in that in step (3), the expansion takes place within the field of A. By the same reasoning as in problem 1, the field of A within the B-shaped hole is constant, with magnitude

$$E = \frac{k(Q/8)}{(R/2)^2} = \frac{kQ}{2R^2}$$

This constant field does no net work when B is expanded, because the positive work done on one half of B is cancelled by the negative work on the other half.

Therefore, we only have to calculate the work done for step (2), which is easy. Applying superposition and the shell theorem, the work needed to bring the point charge to the point where the surface of A meets the surface of the B-shaped hole is

$$W_1 = \frac{Q}{8}\left(\frac{kQ}{R} - \frac{k(Q/8)}{R/2}\right).$$

Next, moving the point charge from this point to the center of the B-shaped hole takes work

$$W_2 = \frac{R}{2}\frac{Q}{8}E = \frac{kQ^2}{32R}.$$

The total work is

$$W_1 + W_2 = \frac{kQ^2}{8R}.$$

[2] **Problem 22** (PPP 149). A distant planet is at a very high electric potential compared with Earth, say 10^6 V higher. A metal space ship is sent from Earth for the purpose of making a landing on the planet. Is the mission dangerous? What happens when the astronauts open the door on the space ship and step onto the surface of the planet?

Solution. As the space ship approaches the planet, its potential gradually increases from that of the Earth, to that of the distant planet. Meanwhile, all the astronauts inside are doing just fine since the ship acts like a Faraday cage. Once the ship lands, it's already at the same potential as the planet, and when the astronauts step out, nothing happens. In other words, it's electric field that's dangerous, not potential, and the electric fields in this problem are always small.

Another way to see that there's no danger is to replace electric fields with gravitational fields, and thus electric potential with gravitational potential. An elevator in a skyscraper takes you from a low to a very high gravitational potential. But nothing violent happens when you get off!

Example 8

Since Newton's law of gravity is so similar to Coulomb's law, the results we've seen so far should have analogues in Newtonian gravity. What are they? For example, what's the gravitational Gauss's law?

Solution

The fundamental results to compare are

$$F = -\frac{Gm_1m_2}{r^2}, \quad F = \frac{q_1q_2}{4\pi\epsilon_0 r^2}$$

where the minus sign indicates that the gravitational force is attractive, while the electrostatic force between like charges is repulsive. Then we can transform a question involving (only positive) electric charges to one involving masses if we map

$$q \to m, \quad \frac{1}{4\pi\epsilon_0} \to -G, \quad \mathbf{E} \to \mathbf{g}.$$

Thus, while electrostatics is described by

$$\nabla \times \mathbf{E} = 0, \quad \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad \oint \mathbf{E} \cdot d\mathbf{S} = \frac{Q}{\epsilon_0},$$

the gravitational field is described by

$$\nabla \times \mathbf{g} = 0, \quad \nabla \cdot \mathbf{g} = -4\pi G \rho_m, \quad \oint \mathbf{g} \cdot d\mathbf{S} = -4\pi G M$$

where ρ_m is the mass density. Similarly, the potential energy can be written in two ways,

$$U = \frac{1}{2} \int \rho_m(\mathbf{r}) \phi(\mathbf{r}) \, d^3\mathbf{r} = -\frac{1}{8\pi G} \int |\mathbf{g}(\mathbf{r})|^2 \, d^3\mathbf{r}$$

where $\phi(\mathbf{x})$ is the gravitational potential. This result was first written down by Maxwell.

Remark

Here's a philosophical question: is potential energy "real"? You likely think the answer is obvious, but about half of your friends probably think the opposite answer is obviously correct! In fact, in the 1700s, there was a lively debate over whether the ideas of kinetic energy and momentum, which at the time were given various other names, were worthwhile. Which one of the two was the *true* measure of motion? In our modern language, proponents of energy pointed out that the momentum always vanished in the center of mass frame, which made it "trivial", while supporters of momentum replied that kinetic energy was clearly not conserved in even the simplest of cases, like inelastic collisions.

In the 1800s, thermodynamics was developed, allowing the energy seemingly lost in inelastic collisions to be accounted for as internal energy. But there still remained the problem that kinetic energy was lost in simple situations, such as when balls are thrown upward. By the mid-1800s, the modern language that "kinetic energy is converted to potential energy" was finally standardized, but it was still common to read in textbooks that potential energy was fake, a mathematical trick used to patch up energy conservation. After all, potential energy has some suspicious qualities. If a ball has lots of potential energy, you can't see or feel it, or even know it's there by considering the ball alone. It doesn't seem to be located anywhere in space, and its amount is arbitrary, as a constant can always be added.

In the late 1800s, a revolution on physics answered some of these questions. Maxwell and his successors recast electromagnetism as a theory of fields, and showed that the dynamics of charges and currents were best understood by allowing the fields themselves to carry energy and momentum. We'll cover this in detail in **E7**, but for now, it implies that electrostatic potential energy is fundamentally stored in the field, with a density of $\epsilon_0 E^2/2$. This implies that its location and total amount are directly measurable.

Maxwell believed that the dynamics of fields emerged from the microscopic motions and elastic deformations of an all-pervading ether, in the same way that, say, a fluid's velocity field emerges from the average motion of fluid molecules. This makes it manifestly positive, so he was disturbed to find that the energy density of a gravitational field is *negative*!

A few decades later, the arrival of special relativity answered some questions and reopened others. On one hand, it demolished Maxwell's vision of the ether. On the other hand, it finally answered the question of whether all kinds of potential energy are "real", and it got rid of the freedom to add arbitrary constants. That's because in special relativity, the total energy of a system at rest is related to its mass by $E = mc^2$, and the mass is directly measurable. This finally puts thermal energy, elastic potential energy, and field energy on an equal footing.

Here's the most modern view of energy conservation. All particles and their interactions are fundamentally described by relativistic quantum fields. A famous result called Noether's theorem implies that whenever such a theory is time-translationally symmetric, there is a conserved quantity which we call the energy. (The distinction between kinetic and potential energy becomes irrelevant; it's all just energy.) The density of energy in space can be computed from the state of the fields, but it doesn't need to be explained, as Maxwell imagined, by the internal motion of whatever the fields are made of. The fields are fundamental: they aren't made of anything; instead, they make up everything!

What happens when we throw gravity into the mix? As we'll discuss further in **R3**, it turns out that at nonrelativistic velocities, the dynamics of gravitating particles can be described by "gravitoelectromagnetism", a theory closely analogous to electromagnetism, where moving masses also source "gravitomagnetic" fields $\mathbf{B}_g$, which result in $m\mathbf{v} \times \mathbf{B}_g$ forces. But the situation gets much more subtle when we upgrade to full general relativity. Here, the notion of a gravitational field disappears completely, and is replaced by the curvature of spacetime, making it hard to define an energy density for it at all. For an accessible overview of the debate, see this paper. Ultimately, though, it doesn't matter that much, since it doesn't impair our ability to use either Newtonian gravity or general relativity.

Example 9

For an infinite line of linear charge density λ, find the potential $V(r)$ by dimensional analysis.

Solution

This example illustrates a famous subtlety of dimensional analysis. The only quantities in the problem with dimensions are λ, ϵ_0, and r. To get the electrical units to balance, we have

$$V(r) = \frac{\lambda}{2\pi\epsilon_0} f(r)$$

where $f(r)$ is a dimensionless function. But there are *no* nontrivial dimensionless functions of a dimensionful quantity r. The only possibilities are that $f(r)$ is a dimensionless constant, or that $f(r)$ is infinite. In the first case, the electric field would vanish, which can't be right. In the second case, it is unclear how to calculate the electric field at all.

In fact, the electric potential *is* infinite, if you insist on the usual convention of setting $V(\infty) = 0$. In that case, we have

$$V(r) = \int_r^\infty \frac{\lambda}{2\pi\epsilon_0} \frac{dr}{r} = \infty$$

independent of r. But this is useless; to get a finite result we can actually work with, we need to subtract off an infinite constant from the potential. Equivalently, we need to set the potential to be zero at some finite distance $r = r_0$. This process is known as renormalization, and it is extremely important in modern physics. After renormalization, we have

$$V(r) = \frac{\lambda}{2\pi\epsilon_0} \log \frac{r_0}{r}$$

which is perfectly consistent with dimensional analysis.

Notice that in the process of renormalization, a new dimensionful quantity r_0 appeared out of nowhere. This phenomenon is known as dimensional transmutation. Of course, physical predictions don't depend on this new scale (e.g. the electric field is independent of r_0), but you can't write down quantities like the potential without it.

3 Conductors

Idea 9

In electrostatic conditions, $\mathbf{E} = 0$ inside a conductor, which implies the conductor has constant electric potential V. This further implies that $\mathbf{E}$ is always perpendicular to a conductor's surface. By Gauss's law, the conductor has $\rho = 0$ everywhere inside, so all charge resides on the surface.

Example 10

Consider a point on the surface of a conductor with surface charge density σ. Show that the outward pressure on the charges at this point is $\sigma^2/2\epsilon_0$.

Solution

Gauss's law tells us that the difference of the electric fields right inside and outside the conductor at this point is

$$E_{\text{out}} - E_{\text{in}} = \frac{\sigma}{\epsilon_0}$$

by drawing a pillbox-shaped Gaussian surface. But we also know that $E_{\text{in}} = 0$ since we're dealing with a conductor, so $E_{\text{out}} = \sigma/\epsilon_0$.

Let's think about how this electric field is made. If there were no charges around except for the ones at this surface, then the interior and exterior fields would have been $\pm\sigma/2\epsilon_0$. This means that all of the other charges, that lie elsewhere on the surface of the conductor, must provide a field $\sigma/2\epsilon_0$ here, so that E_{in} cancels out.

The pressure on the charges at this point on the surface is equal to the product of the surface charge density with the field due to the *rest* of the charges, since the charges at this point can't exert an overall force on themselves, so

$$P = \sigma\left(\frac{\sigma}{2\epsilon_0}\right) = \frac{\sigma^2}{2\epsilon_0}$$

as required. Equivalently, we can conclude that $P = \epsilon_0 E_{\text{out}}^2/2$.

Example 11

Is the charge density at the surface of a charged conductor usually greater at regions of higher or lower curvature?

Solution

We can't answer this question directly, because it is essentially impossible to find the charge distribution of an irregularly shaped conductor. However, we can get some insight by considering the limiting case of a conductor made of two spheres of radii R_1 and R_2, connected by a very long rod.

For the potential to be the same at both spheres, we must have $Q_1/R_1 = Q_2/R_2$, so the charge is proportional to the radius, and the charge density is inversely proportional to the radius. Thus, there's generally higher charge density at sharper points of the conductor.

[1] **Problem 23.** Show that any surface of charge density σ with electric fields $\mathbf{E}_1$ and $\mathbf{E}_2$ immediately on its two sides experiences a force $\sigma(\mathbf{E}_1 + \mathbf{E}_2)/2$ per unit area. (This is a generalization of the example above, where one side was inside a conductor.)

Solution. Let $\mathbf{E}$ be the field due to all the other charges. Again, we have $\mathbf{E}_1 = \mathbf{E} + \sigma/2\epsilon_0 \hat{\mathbf{n}}$ and $\mathbf{E}_2 = \mathbf{E} - \sigma/2\epsilon_0 \hat{\mathbf{n}}$. Thus, $\mathbf{E} = \frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_2)$, and the force per area is $\sigma\mathbf{E}$.

[2] **Problem 24.** Is it possible for a single solid, isolated conductor with a positive total charge to have a negative surface charge density at any point on it? If not, prove it. If so, sketch an example.

Solution. This can't happen. Note that the surface of the conductor has a constant, positive potential. Now suppose there was a region with negative charge on the conductor, and consider a field line that ends on such a charge. It can't have come from infinity, because the potential at infinity is lower than that of the conductor. And it can't have come from elsewhere on the conductor, because the conductor is an equipotential. This yields a contradiction.

Idea 10: Existence and Uniqueness

In a system of conductors where the total charge or potential of each conductor is specified, there exists a unique charge configuration that satisfies those boundary conditions.

This is very useful because in many cases, it is difficult to directly derive the charge distributions or fields. Instead, sometimes one can simply insightfully guess an answer; then it must be the correct answer by uniqueness. For further discussion, see section 2.5 of Griffiths.

Example 12

Consider a conductor with nonzero net charge, and an empty cavity inside. Show that the electric field is zero in the cavity.

Solution

Let's consider a second conductor with the same net charge and the same shape, but without the cavity. By the existence and uniqueness theorem, we know there exists some charge configuration on the second conductor's surface which satisfies the boundary conditions, namely that the electric field vanishes everywhere inside the conductor. In particular, that means the field is zero where the cavity of the original conductor would have been.

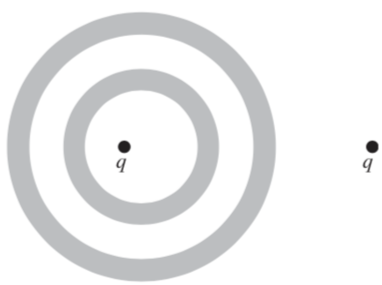
Now consider the original conductor again. If we give this conductor precisely the same surface charge distribution, then this will again solve the boundary conditions, and it'll have no field in the cavity. But by the existence and uniqueness theorem, the charge distribution is unique, so this is the only possible answer: the field *must* be zero in the cavity.

If this is your first time seeing this, it can sound like a fast-talking swindle (which is why I made it an example rather than a problem!). It looks like we used no effort and got a strong conclusion out. Of course, that's because all the work is done by the uniqueness theorem.

[1] **Problem 25.** Consider a spherical conducting shell with an arbitrary charge distribution inside, with net charge Q. Find the electric field outside the shell.

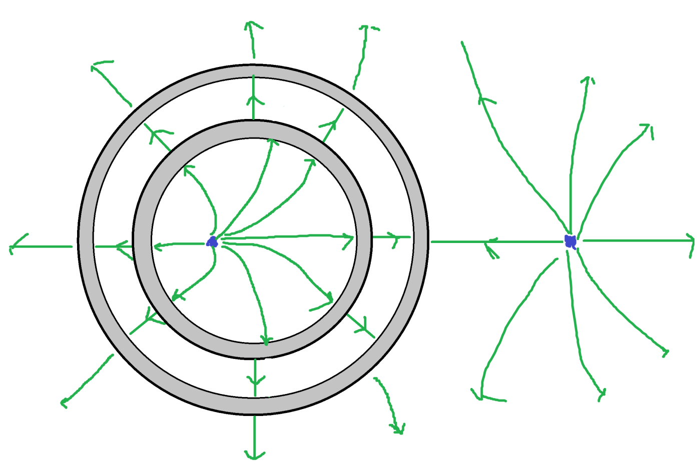
Solution. The shell is an equipotential. The field of a point charge Q at the center of the shell hence satisfies the boundary conditions. By the uniqueness theorem, this is the only solution.

[2] **Problem 26** (Purcell 3.33)**.** The shaded regions represent two neutral conducting spherical shells.



Carefully sketch the electric field. What changes if the two shells are connected by a wire?

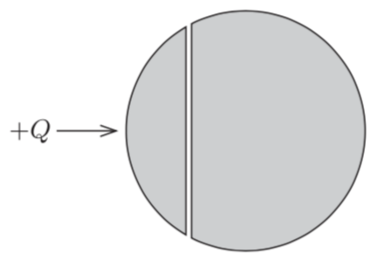
Solution. The results are shown below.



Whether or not there are any field lines coming from the charge outside the shell depends on how close that charge is to the shell. (The entire field configuration in this problem can be found exactly using the method of images, as shown in **E2**.) In the case that the two spheres are connected with a wire, the field in between the two spheres disappears, but nothing else changes.

[3] **Problem 27.** USAPhO 2014, problem A4.

[4] **Problem 28** (MPPP 150)**.** A solid metal sphere of radius R is divided into two parts by a planar cut, so that the surface area of the curved part of the smaller piece is πR^2. The cut surfaces are coated with a negligibly thin insulating layer, and the two parts are put together again, so that the original shape of the sphere is restored. Initially the sphere is electrically neutral.



The smaller part of the sphere is now given a small positive electric charge Q, while the larger part of the sphere remains neutral. Find the charge distribution throughout the sphere, and the electrostatic interaction force between the two pieces of the sphere.

Solution. We know that distributing charge uniformly on the outer surface of the entire sphere will give a valid configuration, in the sense that the field is everywhere perpendicular to the conductors. Similarly, distributing equal and opposite charges uniformly on the two flat faces will give a valid configuration, since it acts like a parallel plate capacitor, making the field vanish everywhere outside.

Neither of these solutions have the right total charge on each piece, but we can fix this by superposing the two. By solving a system of two equations, we find the charge distribution is

- total charge Q distributed uniformly on the sphere,
- charges $\pm(3/4)Q$ distributed uniformly on the flat surfaces.

The two flat faces attract each other and the two curved faces repel each other; there are no other forces by the shell theorem. The pressure on the flat faces is $\sigma^2/2\epsilon_0$. With a little trigonometry, we find the area of the flat faces is $(3/4)\pi R^2$, giving a force

$$F_1 = \frac{3Q^2}{8\pi\epsilon_0 R^2}.$$

As for the repulsive force, using the result of problem 11 we get

$$F_2 = -\frac{3Q^2}{128\pi\epsilon_0 R^2}, \quad F_{\text{tot}} = F_1 + F_2 = \frac{45}{128}\frac{Q^2}{\pi\epsilon_0 R^2}.$$

[3] **Problem 29.** In this problem we'll work through a heuristic proof of a version of the uniqueness theorem. In particular, we will show that for a system of conductors in empty space, specifying the total charge on each conductor alone specifies the entire surface charge distribution.

(a) Suppose for the sake of contradiction that two different charge distributions can exist, and consider their difference, which has zero total charge on each conductor. Argue that at least one conductor must have electric field lines both originating from and terminating on it.

(b) Show that at least one of these field lines must originate from or terminate on another one of the conductors.

(c) By generalizing this reasoning, prove the desired result. (Hint: consider the conductors with nonzero surface charges that have the highest and lowest potentials.)

Solution. (a) Since the overall charge distributions are different, at least one conductor C must have different charge distributions in the two cases. So when we consider the difference, C has areas of both positive and negative surface charge. Field lines come out of the former, and go into the latter.

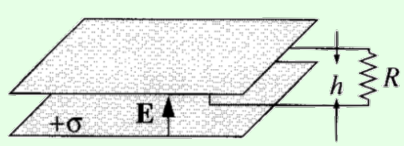
(b) The field lines can connect back to C, because by following the field line, one would prove that C has a higher potential than itself, which is impossible. They also can't all go off to infinity, because we can consider "infinity" to just be a big, far away neutral conductor at zero potential. If lines both came from infinity to C and from C to infinity, then C would again have a higher potential than itself, which is impossible. So some field line must go between C and another conductor C'.

(c) By assumption, at least some of the conductors have nontrivial surface charges on them. So among those conductors, consider the one with the highest potential ϕ_{max}. At we argued in part (a), this conductor has to have both field lines coming from it and going into it. Since potential decreases along field lines, the field lines going into it can't come from any of the other conductors, so they have come from infinity. Since infinity is at zero potential, we have $\phi_{\text{max}} \leq 0$.

Now consider the conductor with the lowest potential ϕ_{min}, which has nontrivial surface charges. Again, at least some field lines have to leave this conductor, but they can't go anywhere except for infinity. Since infinity is at zero potential, $\phi_{\text{min}} \geq 0$. This forces $\phi_{\text{min}} = \phi_{\text{max}} = 0$, so everything must be at zero potential, which means there aren't any electric field lines at all.

Example 13: Griffiths 7.6

A wire loop of height h and resistance R has one end placed inside a parallel plate capacitor with electric field $\mathbf{E}$, as shown.

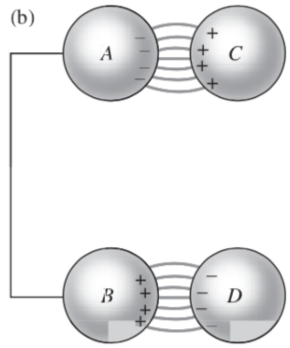


The other end of the loop is far away, where the field is negligible. Find the emf in the loop.

Solution

This is a trick question: if the answer were nonzero, the current would run forever, yielding a perpetual motion machine. Electrostatic fields always produce zero total emf along any loop. The $\sigma h/\epsilon_0$ voltage drop inside the capacitor is canceled out by the voltage drop due to the fringe fields, which are small, but accumulate over a long distance. The point of this example is that, while we can ignore fringe fields for some calculations, they are often essential to get a consistent overall picture. We'll revisit the subtleties of fringe fields in **E2**.

[2] **Problem 30** (Purcell 3.2). Spheres A and B are connected by a wire; the total charge is zero. Two oppositely charged spheres C and D are brought nearby, as shown.



The spheres C and D induce charges of opposite sign on A and B. Now suppose C and D are connected by a wire. Then the charge distribution should not change, because the charges on C and D are being held in place by the attraction of the opposite charge density. Is this correct?

Solution. This isn't correct. To see this rigorously, we can use the uniqueness theorem. After connecting the wires, we have two conductors (A/B, and C/D), each with zero net charge. One possible solution is to have zero charge everywhere. By uniqueness, this is the only possible solution, so anything else cannot have been in equilibrium.

That is rigorous, but it might not be intuitive; after all, it sure looks like the charges on C are stuck where they are. However, though the charges on C are attracted towards A, they also strongly repel each other. It's this repulsion that causes charge on C to start flowing to D when the wire is connected.

Electromagnetism II: Electricity

Chapters 3 and 4 of Purcell cover the material presented here, as does chapter 6 of Wang and Ricardo, volume 2. Image charges are covered in more detail in section 3.2 of Griffiths. For an array of interesting physical examples, see chapters II-6 through II-9 of the Feynman lectures. There is a total of **77** points.

1 The Method of Images

Idea 1

The method of images can be used in some highly symmetric situations to compute the electric field in the vicinity of a conductor. Specifically, consider any configuration of static charges and take any equipotential surface containing some of the charges. Then the resulting field configuration outside that surface is the field configuration we would have if that surface bounded a conductor. This is simply because it has constant potential on the conductor surface, so it must be the right answer by the uniqueness theorem.

[4] **Problem 1.** The simplest application of the method of images is the case of a charge q a distance a from an infinite grounded conducting plane. This problem explores some of its subtleties, assuming you've already read the basic treatment in section 3.4 of Purcell.

(a) Find the force on the charge.

(b) Find the work needed to move the charge out to infinity. (Answer: $q^2/16\pi\epsilon_0 a$.)

(c) Find the total potential energy of the charges on the conducting plane, i.e. the potential energy associated only with their interaction with each other. (Answer: $q^2/16\pi\epsilon_0 a$.)

(d) Now suppose there is another parallel grounded conducting plane on the other end of the charge, a distance b away. How many image charges are needed now? Draw some of them.

(e) A conducting plane forces the electric field to be perpendicular to it. Suppose we modified the plane so that the electric field was instead always *parallel* to it. Find the force on the charge.

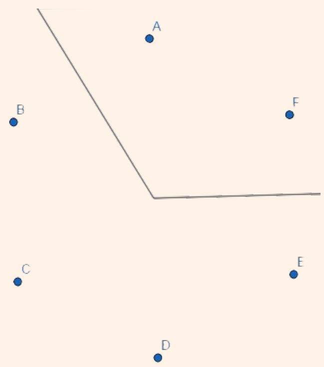
Example 1

Two grounded conducting half-planes intersect, so that in cylindrical coordinates, the equations describing the planes are $\theta = 0$ and $\theta = \theta_p = \pi/2$. A charge q is placed somewhere between the planes. Can the method of images be used to find the force on the charge? What if $\theta_p = 2\pi/3$, or for general θ_p?

Solution

We can solve the first case with three image charges. Suppose the charge q is at (x, y). Then we can reflect in the plane $\theta = 0$, adding a charge $-q$ at $(x, -y)$ to satisfy its boundary condition. Then we can reflect both the real charge and this image charge in the plane $\theta = \pi/2$ to satisfy that plane's boundary condition, adding a charge $-q$ at $(-x, y)$ and a charge q at (x, y).

But when the other plane is at $\theta = 2\pi/3$, there is no configuration of image charges that works. For concreteness, let's suppose the real charge is at point A, on the y-axis.



Reflecting in the $\theta = 0$ plane forces us to have an image charge $-q$ at D, reflecting in the $\theta = 2\pi/3$ plane yields an image charge q at E, and reflecting in the $\theta = 0$ plane again yields a $-q$ charge at F, which is *real* since it's in the same region as A. But this isn't allowed: the point of image charges is to provide an easy way of calculating the effects of screening charges on conducting surfaces on a given set of *real* charges (i.e. the charge at A), so it's not legal to introduce *new* real charges in the process. We would get the same conclusion if we reflected about the planes in a different order – we always need a charge at F. More generally, the method of images works for this problem if and only if $\theta_p = \pi/n$ for integer n.

[4] **Problem 2.** In this problem you'll develop the method of images for spheres.

(a) A point charge $-q$ is located at $x = a$ and a point charge Q is located at $x = A$. Show that the locus of points with $\phi = 0$ is a circle in the xy plane, and hence a spherical shell in space.

(b) Show that the center of the sphere is at the origin provided that

$$a = \frac{r^2}{A}, \qquad -q = -\frac{Qr}{A}$$

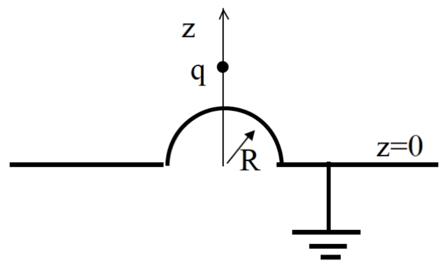
where r is the sphere's radius. These results will be used throughout the problem set.

(c) Now suppose a point charge Q is a distance b from the center of a spherical grounded conducting shell of radius r. Find the force on the charge, considering both the cases $b < r$ and $b > r$.

(d) The case $b < r$ is a bit confusing. On one hand, argue that the total charge is Q plus the image charge, and hence nonzero. On the other hand, argue that the total charge must be zero, by considering an appropriate Gaussian surface. One of these arguments is wrong – which one?

As an aside, the fundamental reason the method of images works for spheres is that electromagnetism has conformal symmetry, a symmetry under any local rescaling of space which preserves angles. (One example of a conformal transformation is inversion in Euclidean geometry.) The setup here is related to the conducting plane by such a transformation.

[2] **Problem 3.** An infinite grounded conducting plane at $z = 0$ is deformed with a hemispherical bump of radius R centered at the origin, as shown. A charge q is placed at $z = a$ as shown.



Can the method of images be used to find the potential in the region with the charge? If so, specify the image charges; if not, explain why not.

[2] **Problem 4** (Purcell 3.50). A point charge q is located a distance $b > r$ from the center of a *nongrounded* conducting spherical shell of radius r, which also has charge q. When b is close to r, the charge is attracted to the shell because it induces negative charge; when b is large the charge is clearly repelled. Find the value of b so that the point charge is in equilibrium. (Hint: you should have to solve a difficult polynomial equation. You can either use a computer or calculator, or use the fact that it contains a factor of $1 - x - x^2$.)

[2] **Problem 5.** Consider a neutral spherical conductor of radius R placed in a uniform external field $\mathbf{E}_0$. By using the method of images, argue that the field created by the sphere is the same as the field of an infinitesimal dipole at the center of the sphere, and find its dipole moment. (Hint: suppose the external field is sourced by two point charges that are very far away.)

[5] **Problem 6** (Purcell 3.45). [A] Consider a point charge q located between two parallel infinite grounded conducting planes. The planes are a distance ℓ apart, and the point charge is a distance b from the left plane. The goal of this problem is to find the total charge induced on each plane.

(a) Argue that the total charge on each plane would not change if we replaced the point charge q with two point charges $q/2$, both a distance b from the left plane. By iterating this process, convert the point charge into a uniformly charged plane, and use this to get the answer.

(b) Alternatively, using image charges, show that the electric field on the inside surface of the left plane, perpendicular to the plane, at a point a distance r from the axis containing all the image charges, satisfies

$$4\pi\epsilon_0 E_\perp = \sum_{n=-\infty}^{\infty} \frac{2q(2n\ell + b)}{((2n\ell + b)^2 + r^2)^{3/2}}.$$

(c) Since $\sigma = -\epsilon_0 E_\perp$, we can integrate both sides to find the total charge on the left plane. However, the integral of each term by itself is simply q, so the series doesn't converge. To get the result, do the following steps in this specific order: group the terms $\pm n$ together, then integrate them only out to a distance $R \gg b$, then sum over the values of $|n|$, then take the limit $R \to \infty$. You should get a finite result that matches that of part (a). As you'll probably see in the process, if you do the steps in any other order, you'll get a nonsensical answer.

Those concerned with mathematical rigor might be bothered by the many choices made in part (c). You might ask, couldn't we have gotten a different result by changing how we did the computation? In fact, by the Riemann rearrangement theorem, we could have gotten almost *any* result. But the way we did it is the physically correct way. It roughly sums the terms "in to out", which respects the fact that real plates are finite. Closely related ideas are used to "cancel infinities" in quantum field theory, in a process known as renormalization. We'll see another example in **X1**.

2 Capacitors

Idea 2

There are multiple definitions of capacitance. For a single, isolated conductor with charge Q, the self-capacitance is defined as

$$Q = C\phi$$

where ϕ is the potential difference between the conductor and infinity. But for a set of two isolated conductors with charges $\pm Q$, you can also define a "mutual" capacitance by

$$Q = C\phi$$

where ϕ is the potential difference between the two conductors. When someone talks about a "capacitance" without specification, such as in idea 4, they probably mean this latter one.

Idea 3

The definitions of C above are only useful when you have only one or two conductors in the problem, respectively. In a situation with more than two, it's very tricky to use the above definitions, because all the conductors will affect each other; even a neutral conductor will have an effect since there will be induced charges on its surface.

Instead, it's better to revert to more general principles. The underlying principle behind capacitance is linearity: the charges are linearly related to the fields. For multiple capacitors, the most general possible linear relation is

$$Q_i = \sum_j C_{ij}\phi_j$$

where conductor i has charge Q_i and potential ϕ_i, the potential is taken to be zero at infinity, and the C_{ij} are called general coefficients of capacitance, or in electrical engineering, the Maxwell capacitance matrix. Similarly, inverting this relation,

$$\phi_i = \sum_j p_{ij} Q_j$$

where the p_{ij} are called coefficients of potential.

In Olympiad physics, you'll almost never want to compute the C_{ij} or p_{ij} explicitly. Instead, the point here is that if you're given the charges and want the potentials, or vice versa, you can build up the answer you want using the principle of superposition, computing all the fields you need, e.g. using Gauss's law.

Remark

General capacitance coefficients are discussed further in section 3.6 of Purcell. One nontrivial fact is that $C_{ij} = C_{ji}$, which is proven by energy conservation in problem 3.64 of Purcell. Capacitance coefficients can be clunky to work with. For example, suppose you want to compute the familiar capacitance of a system of two conductors. By definition, we have

$$Q_1 = C_{11}\phi_1 + C_{12}\phi_2, \quad Q_2 = C_{21}\phi_1 + C_{22}\phi_2.$$

An ordinary two-plate capacitor corresponds to the special case of opposite charges on the plates, so we write $Q = Q_1 = -Q_2$. There is a potential difference V across the plates, so $\phi_1 = \phi_2 + V$, and plugging this in gives

$$Q = (C_{11} + C_{12})\phi_2 + C_{11}V, \quad -Q = (C_{22} + C_{21})\phi_2 + C_{21}V.$$

Eliminating ϕ_2 from the system of equations above, we find the familiar mutual capacitance

$$C = \frac{Q}{V} = \frac{C_{11}C_{22} - C_{12}^2}{C_{11} + C_{22} + 2C_{12}}$$

where we used $C_{12} = C_{21}$. This is quite an inconvenient formula, so as a result we won't consider general capacitance coefficients any further, except briefly for practice in problem 8.

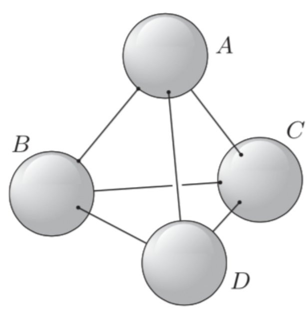
[2] **Problem 7** (Purcell 3.21)**.** Consider a capacitor made of four parallel plates with large area A, evenly spaced with small separation s. The first and third are connected by a wire, as are the second and fourth. What is the capacitance of the system?

[3] **Problem 8.** Consider two concentric spherical metal shells, with radii $a < b$.

(a) Compute their capacitance using Gauss's law.

(b) Compute their capacitance by computing the four capacitance coefficients, verifying that $C_{12} = C_{21}$ along the way, and using the result for C above.

[3] **Problem 9** (MPPP 152)**.** Four identical non-touching metal spheres are positioned at the vertices of a regular tetrahedron, as shown.



A charge $4q$ given to sphere A raises it to a potential V. Sphere A can also be raised to potential V if it and one of the other spheres are each charged with $3q$. What must be the size of equal charges given to A and two other spheres for the potential of A to again be raised to V? What if all four spheres are used?

[3] **Problem 10.** USAPhO 2008, problem A1.

Idea 4

A two-plate capacitor with voltage difference V and mutual capacitance C stores energy

$$U = \frac{1}{2}QV = \frac{1}{2}CV^2.$$

Many circuits have multiple two-plate capacitors. In general, these need to be handled with the capacitance coefficients introduced in idea 3. But in practice, capacitors used in circuits are designed to produce fields confined within themselves, so that different capacitors don't interact with each other. In that case, we can just use mutual capacitance throughout, and C adds in parallel, while $1/C$ adds in series. (But this not work if, e.g. you put one capacitor inside another, in which case you should think about the charges and fields directly.)

[2] **Problem 11** (Purcell 3.24)**.** Some estimates involving capacitance.

(a) Estimate the capacitance of the Earth.

(b) Make a rough estimate of the capacitance of the human body.

(c) By shuffling over a nylon rug on a dry winter day, you can easily charge yourself up to a couple of kilovolts, as shown by the length of the spark when your hand comes too close to a grounded conductor. How much energy would be dissipated in such a spark?

[2] **Problem 12.** The total energy can also be found by integrating the electric field energy,

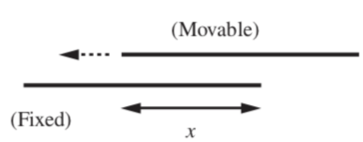
$$U = \frac{\epsilon_0}{2} \int E^2 \, dV.$$

(a) Show that this agrees with $U = CV^2/2$ for a parallel plate capacitor.

(b) Show that this agrees with $U = CV^2/2$ for a capacitor made of concentric spheres.

The general proof is more advanced, but if you're interested, one slick method is given in problem 1.33 of Purcell.

[3] **Problem 13** (Purcell 3.26)**.** A parallel-plate capacitor consists of a fixed plate and a movable plate that is allowed to slide in the direction parallel to the plates. Let x be the distance of overlap.



The separation between the plates is fixed. Let $C(x)$ be the capacitance.

(a) Assume the plates are electrically isolated, so that their charges $\pm Q$ are constant. By differentiating the energy, find the leftward force on the movable plate in terms of Q and $C(x)$.

(b) Now assume the plates are connected to a battery, so that their potential difference ϕ is held constant. Find the leftward force on the movable plate, in terms of ϕ and $C(x)$.

(c) If the movable plate is held in place, the two answers above should be equal because nothing is moving. Verify that this is the case, being careful with signs.

(d) In terms of electric fields, why is there a force on the movable plate? Does the effect invoked in the answer to this part change the conclusion of parts (a) through (c) at all?

Idea 5: Dielectrics

While we're on the subject of capacitors, it's useful to introduce dielectrics. A dielectric is an insulator which polarizes in the presence of an electric field, with positive charges displaced slightly along the field. The resulting electric dipoles distributed throughout the material in turn create a field that tends to weaken the original applied field within the material.

Each part of a dielectric polarizes based on the local electric field, but that electric field depends on the applied field, and the polarization of *every other* piece of the dielectric. Thus, solving for the electric field for a general dielectric geometry is very difficult, and usually not possible in closed form, just like how it's usually not possible to solve for the field of a charged conductor. In Olympiad physics, you will almost always consider highly symmetric situations, where a dielectric simply reduces the applied electric field everywhere inside by a factor of κ, called the dielectric constant. (We'll consider some trickier situations in **E8**.)

Consider a parallel plate capacitor with charge $\pm Q$ on each plate. If a dielectric is inserted with the charge kept the same, then the field inside is reduced by a factor of κ. Thus, the capacitance $C = Q/V$ increases by a factor of κ. Dielectrics may increase the amount of energy that can be stored in a capacitor, which is typically limited by the voltage V_0 where electrical breakdown occurs. So if V_0 stays the same, the maximal stored energy $U = CV_0^2/2$ goes up by a factor of κ.

Plugging in the definition of C, this result implies that the energy density in the capacitor is $\kappa\epsilon_0 E^2$. But we showed in **E1** that the energy density of the electric field is only $\epsilon_0 E^2$. The extra energy is stored in the dielectric material itself: it takes energy to separate positive and negative charges within the dielectric, as if we were stretching many microscopic springs. This potential energy is released when the capacitor is discharged.

3 Tricky Problems

Example 2: PPP 151

A closed body with conducting surface F has self-capacitance C. The surface is now dented so that the new surface F^* is entirely inside F. Prove that the capacitance has decreased.

Solution

The energy stored in the capacitor is $U = Q^2/2C$. Therefore, if we give the capacitor a fixed charge Q, proving that F^* has lower C is equivalent to showing that it takes positive work to dent the foil from F to F^*. It's cleaner to show the other direction, i.e. that starting from F^*, we can get to F while only lowering the energy.

Suppose without loss of generality that F is infinitesimally larger than F^*. (We can break any finite change into infinitesimal stages and repeat this argument.) We can go from F^* to F by just taking each charge on the surface and moving it outward until it hits F.

This lowers the energy because the electric field is always directed outward, as we proved in **E1**.

At this point, the charges lie on F, but they don't have the right distribution, i.e. F is not an equipotential. Now we let the charges spontaneously redistribute themselves so that F is again an equipotential. This again lowers the energy, proving the desired result.

Example 3

Are there charge distributions that aren't spherically symmetric, but which produce an *exact* $\hat{\mathbf{r}}/r^2$ field outside of them?

Solution

If you know a bit about the multipole expansion, this might seem like a daunting question. To make the field exactly $\hat{\mathbf{r}}/r^2$, you need to make sure the charge distribution has no dipole moment, no quadrupole moment, no octupole moment, and so on to infinity, and it seems impossible to satisfy all of these constraints without spherical symmetry. But we have *already* seen an example of such a charge distribution earlier in the problem set!

Recall that when we treated the method of images for spheres, we found that in some situations, the complicated charge densities on conducting spheres were exactly the same as those produced by a fictitious image charge inside the sphere, and generally away from its center. If we place the origin at that image charge, then we have an example of a charge distribution that is perfectly $\hat{\mathbf{r}}/r^2$ far away, but which isn't spherically symmetric. (The general solution is given here.)

[2] **Problem 14.** Consider a set of n conducting, very large parallel plates, placed in zero external electric field. The plates are given charges Q_i. If the left ends of the plates are at locations x_i, and the plates have thickness d_i, what is the total charge on the left end of the leftmost plate, and the right end of the rightmost plate?

[2] **Problem 15** (Purcell 3.9)**.** A conducting spherical shell has charge Q and radius R_1. A larger concentric conducting spherical shell has charge $-Q$ and radius R_2.

(a) If the outer shell is grounded, explain why nothing happens to the charge on it.

(b) If instead the inner shell is grounded, e.g. by connecting it to ground by a very thin wire that passes through a very small hole in the outer shell, find its final charge.

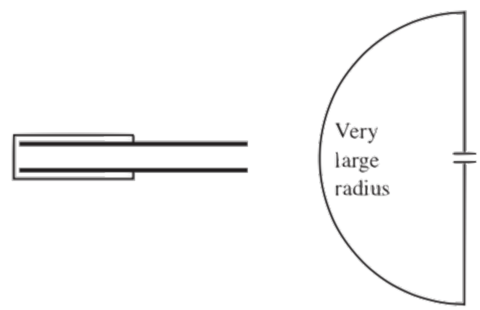
(c) It's not so clear why charge would leave the inner shell in part (b), thinking in terms of forces. A small bit of positive charge will certainly want to hop on the wire and follow the electric field across the gap to the larger shell. But when it gets to the larger shell, it seems like it has no reason to keep going to infinity, because the field is zero outside. And, even worse, the field will point *inward* once some positive charge has moved away from the shells. So it seems like the field will drag back any positive charge that has left. Does charge actually leave the inner shell? If so, what's wrong with the above reasoning?

[2] **Problem 16.** The usual expression for the capacitance of a parallel-plate capacitor is $A\epsilon_0/d$. However, in reality the field within the capacitor is not perfectly uniform, and there are fringe fields outside. Taking these effects into account, is the true capacitance higher or lower than $A\epsilon_0/d$?

[2] **Problem 17** (Purcell 4.16)**.** In a parallel plate capacitor, the quantity $\int \mathbf{E} \cdot d\mathbf{s}$ should be equal to V for any path that connects the two plates.

A charged capacitor can be discharged by attaching a wire to the external surfaces of the plates. No matter how one attaches the wire, $\int \mathbf{E} \cdot d\mathbf{s}$ along the wire should be equal to V. And as we've argued in **E2**, this is sufficient to cause charges to move along the wire, even if the electric field points in the "wrong" direction at some points along the wire, because the wire has negligible capacitance: charges within it move rigidly, each pushing the next one and pulling the previous one.

But it's puzzling how this works for a capacitor, because the electric field is supposed to be essentially zero just outside it. Consider two possible limiting cases for the wire's shape.



In each case, explain qualitatively how $\int \mathbf{E} \cdot d\mathbf{s}$ can be equal to V. In particular, how large are the contributions from the distinct segments of the wire (the horizontal and vertical parts in the first case, and the straight and curved parts in the second)?

[3] **Problem 18.** Consider two conducting spheres of radius r separated by a distance $a \gg r$. The spheres can be thought of as the two plates of a parallel plate capacitor.

(a) By assuming the surface charge density on each sphere is uniform, estimate the capacitance.

(b) In reality, the surface charges on each sphere will be distorted by the other sphere. The surface charges assumed in part (a) will induce changes in the surface charges, represented by an image charge for each sphere. But these image charges will induce further image charges, and so on, yielding an infinite series for the charge distribution and hence the capacitance. Find the first two terms in the series for the capacitance.

(c) If $a/r = 10$, roughly estimate the error in neglecting the other terms.

[3] **Problem 19.** USAPhO 2022, problem A2. A computational problem involving surface tension.

Example 4

Find the leading interaction force between a dipole of dipole moment p and a grounded conducting sphere of radius r, separated by a distance $R \gg r$. What if the sphere is electrically neutral instead?

Solution

Place the origin at the center of the sphere and orient the z-axis to pass through the dipole. We can regard the dipole $p = qd$ as a combination of two charges

$$q \text{ at } z = R, \quad -q \text{ at } z = R + d$$

where d is very small. In the grounded case, this induces two image charges in the sphere,

$$\frac{qr}{R} \text{ at } z = \frac{r^2}{R}, \quad -\frac{qr}{R+d} \text{ at } z = \frac{r^2}{R+d}$$

approximately separated by r^2/R^2. We can now use Coulomb's law four times, but that's a bit tedious. Instead, decompose the image charges into a dipole moment and a net charge,

$$p' = \frac{pr^3}{R^3}, \quad Q' = \frac{qr}{R} - \frac{qr}{R+d} \approx \frac{pr}{R^2}.$$

We can place both of these at the origin, because this slight displacement will only affect the answer by subleading terms in r/R. Then the corresponding fields, far along the z-axis, are

$$E_{p'}(z) = \frac{2kpr^3}{R^3 z^3}, \quad E_{Q'}(z) = \frac{kpr}{R^2 z^2}.$$

The first term is negligible compared to the second, due to the many powers of R and z in the denominator. Thus, keeping only the second term, the force on the original dipole is

$$F = p \frac{d}{dz} E(z) \bigg|_{z=R} = -\frac{2kp^2 r}{R^5}$$

which falls off very quickly with distance. This derivation illustrates a common subtlety: it might not always be obvious how far to approximate. We threw away terms subleading in r/R, because we only wanted the leading contribution. But if we had applied that principle to the image charges at the first step, we would have thrown out the tiny net charge Q', which actually provides the dominant contribution to the force, because of how tiny p' is.

Now, the situation for a neutral sphere is completely different. By the logic of problem 4, there's a third image at the center of the sphere to enforce neutrality,

$$-\frac{pr}{R^2} \text{ at } z = 0.$$

The image charges can now be decomposed into a combination of two dipole moments. We already saw the first one p' above, while the second is, to leading order

$$p'' \approx \frac{pr}{R^2} \frac{r^2}{R} = \frac{pr^3}{R^3}$$

with the same magnitude and direction as p'. Thus, this system of image charges has approximate dipole moment $2p'$. The corresponding force is

$$F = p \frac{d}{dz} \frac{4\,kpr^3}{R^3 z^3} \bigg|_{z=R} = -\frac{12\,kp^2 r^3}{R^7}$$

which falls off even more quickly with distance. In this derivation, we didn't have to worry too much about getting p'' exactly right, because there was no net charge ("monopole") term that could've overwhelmed the dipole field, so all other field contributions are safely suppressed by more powers of r/R. (Of course, if p'' had come out pointing the opposite direction to p', so that the two almost cancelled, we would've had to be more careful.)

The lesson of this example is *not* to just use exact expressions and Taylor expand at the end. Here, that brute force approach would have required Taylor expanding six Coulomb's law forces out to order $1/R^7$, which is extraordinarily tedious. Instead, to approximate properly, we have to think carefully in every case. Incidentally, when applied to a polar and neutral nonpolar molecule, the $1/R^7$ force above is called the Debye force; it is one of the "van der Waals forces" which are often vaguely described in chemistry classes.

Example 5

Estimate the interaction force between a point charge q and a thin conducting rod of length ℓ, which is a distance $L \gg \ell$ from the charge and oriented along the separation between them.

Solution

The interaction occurs because the point charge induces negative charges on the near end of the rod, and positive charges on the far end. These charges are then acted on by the electric field of the point charge, causing a force.

To get a very crude estimate, let's just suppose that charge Q appears on the far end and charge $-Q$ appears on the near end. The resulting field produced in the middle is

$$E \sim \frac{kQ}{\ell^2}.$$

On the other hand, this needs to cancel a field from the point charge of

$$E \sim \frac{kq}{L^2}$$

which tells us that $Q \sim (\ell/L)^2 q$. The force on the induced charges is

$$F \sim kqQ \left(\frac{1}{L^2 + \ell^2} - \frac{1}{L^2} \right) \sim -\frac{kqQ\ell^2}{L^4} \sim -\frac{kq^2\ell^4}{L^6}.$$

Again, the force is attractive, and falls off quickly with distance.

[3] **Problem 20** (Physics Cup 2017)**.** Estimate the interaction force between a point charge q and an infinitely thin circular neutral conducting disc of radius r if the charge is at the axis of the disc, and the distance between the disc and the charge is $L \gg r$.

Example 6

Find the charge distribution on a conducting disc of radius R and total charge Q.

Solution

In general, there are very few situations where the charge distribution on a conductor can be found explicitly. As you've seen, some of the simplest examples can be solved with image charges. Some more complex, two-dimensional examples can be solved with a mathematical technique called conformal mapping. And this special example can be solved with a neat trick.

Consider a uniformly charged spherical shell centered on the origin, and consider a point P inside the shell, on the xy plane. The electric field at point P is zero, by the shell theorem. Recall that in the usual proof of the shell theorem, one draws two cones opening out of P in opposite directions. The charges contained in each cone produce canceling electric fields.

Now imagine shrinking the spherical shell towards the xy plane, so it becomes elliptical. The crucial insight is that the shell theorem argument above still works, for points on the xy plane. When we squash the shell all the way down to the xy plane, it becomes a disc, with zero electric field on it. This is thus a valid charge distribution for a disc-shaped conductor, and by the uniqueness theorem, it's the only one.

By keeping track of how much charge gets squashed to radius $[r, r + dr]$, we find $\sigma(r) \propto R/\sqrt{R^2 - r^2}$, and fixing the proportionality constant gives

$$\sigma(r) = \frac{Q}{4\pi R\sqrt{R^2 - r^2}}.$$

4 Electrical Conduction

We now leave the world of electrostatics and consider magnetostatics, the study of steady currents.

Idea 6

In a conductor with conductivity σ, the current density is

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

Alternatively, $\mathbf{E} = \rho\mathbf{J}$ where ρ is the resistivity. The current and charge density satisfy

$$\nabla \cdot \mathbf{J} = -\frac{\partial\rho}{\partial t}.$$

The current passing through a surface S at a given time is

$$I = \int_S \mathbf{J} \cdot d\mathbf{S}.$$

Since $\mathbf{J} \propto \mathbf{E}$, we have Ohm's law $V = IR$, where V is the voltage drop across the resistor. The power dissipated in a resistor is $P = IV$. The resistance R adds in series, while $1/R$ adds in parallel.

[2] **Problem 21** (HRK)**.** A battery causes a current to run through a loop of wire.

(a) Suppose the wire makes a sharp corner. How do the charges know to turn around there?

(b) A copper wire with conductivity σ is joined to an iron wire with conductivity $\sigma' < \sigma$. For the current in both sections to be the same, the electric field in the iron wire must be higher. How does that happen?

In general, the surface charge distribution in a DC circuit can be quite complex; the aspects shown in these questions are just the beginning. For more about this, see this paper and this paper.

[2] **Problem 22** (PPP 22). Two students, living in neighboring rooms, decided to economize by connecting their ceiling lights in series. They agreed they would each install a 100 W bulb in their own rooms and that they would pay equal shares of the electricity bill. However, both tries to get better lighting at the other's expense. The first student installed a 200 W bulb, while the second student installed a 50 W bulb. Which student subsequently failed their final exams?

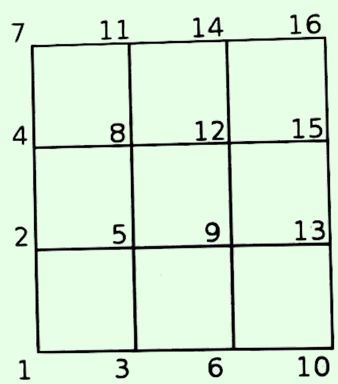
To warm up for DC circuits, we'll consider some resistor network problems.

Idea 7

If any two points in a resistor network are at the same potential, nothing will change if the two points are connected together and treated as one. More generally, the resistance of any resistor directly connecting the two points may be changed freely.

Example 7

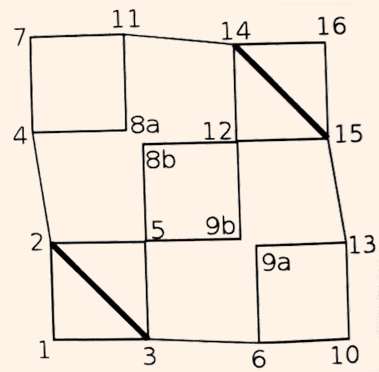
Consider the 3×3 grid below, where every edge is a resistor R.



Find the equivalent resistance between nodes 1 and 16.

Solution

By the above idea, we can short together nodes 2/3, and 14/15, by the diagonal symmetry of the network. Next, we can break nodes 8 and 9 into two pieces.



This is valid because the separated nodes $8a/8b$ and $9a/9b$ still have the same potential in the new network, by the diagonal symmetry. (This is using the above idea in reverse.) Now, the circuit has been reduced to combinations of series and parallel resistors. The resistance between 1 and 2/3 is $R/2$. The resistance between 2/3 and 14/15 is the combination of three networks in parallel, and finally the resistance of 14/15 and 16 is $R/2$. Thus,

$$R_{\text{eq}} = \left(\frac{1}{2} + \left(\frac{1}{3} + \frac{1}{3} + \frac{1}{2} \right)^{-1} + \frac{1}{2} \right) R = \frac{13}{7} R.$$

You won't see any resistor problems as complicated as this one for the rest of the training, because they're kind of contrived; the point of this example was just to show multiple uses of symmetry techniques.

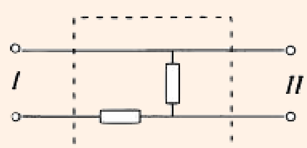
Example 8: PPP 23

A black box contains a resistor network and has two output terminals.

If a battery of voltage V is connected across the first terminal, the voltage across the second terminal is $V/2$. If a battery of voltage V is connected across the second terminal, the voltage across the first terminal is V. Find one possible configuration of the resistors inside the box.

Solution

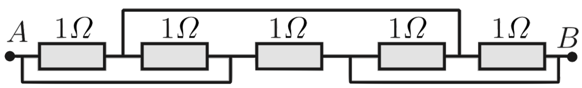
A simple configuration with two equal resistors works.



When a battery is connected across II, the horizontal resistor doesn't do anything. When a battery is connected across I, the two resistors comprise a voltage divider.

[2] **Problem 23.** USAPhO 2007, problem A1.

[2] **Problem 24** (IPhO 1996)**.** Consider the following resistor network.



Find the equivalent resistance between A and B.

[3] **Problem 25.** Consider a cube of side length L whose edges are resistors of resistance R.

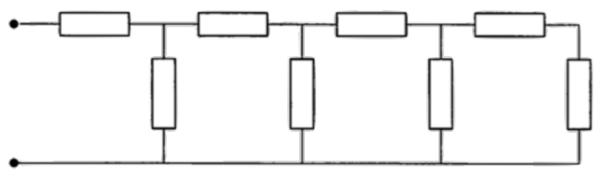
(a) Compute the resistance between two vertices a distance $\sqrt{3}L$ apart.

(b) Compute the resistance between two vertices a distance $\sqrt{2}L$ apart.

(c) Compute the resistance between two vertices a distance L apart.

(d) Generalize to vertices $\sqrt{n}L$ apart on an n-dimensional cube. (Give your answer in the form of a summation.)

[2] **Problem 26** (PPP 158)**.** Consider the circuit below, where every resistor is $1\,\Omega$.

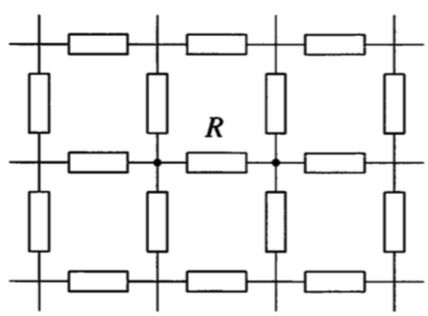


(a) Find the equivalence resistance between the input terminals.

(b) Do the same in the case where the chain is infinitely long.

[3] **Problem 27** (PPP 159-161)**.** Superposition can be a useful trick to analyze circuit networks.

(a) Consider an infinite two-dimensional grid of identical resistors R.



Find the equivalent resistance between two neighboring points by considering the superposition of a current I flowing into one point, and an equal current I flowing out the other.

(b) What would the equivalent resistance be if the resistor directly connecting the two neighboring points was removed?

(c) Now consider an icosahedron of identical resistors R. By superposing appropriate current distributions, find the equivalent resistance between two neighboring vertices.

Idea 8

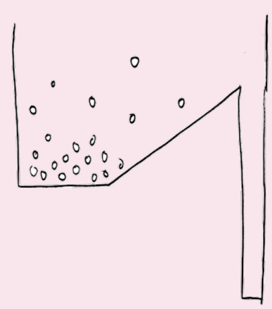
In a circuit of resistors and batteries, Kirchoff's loop rule states that the sum of the voltage drops around a loop is zero. Kirchoff's junction rule states that the net current flowing into a vertex is zero. (This is technically nonzero, because of the effect of problem 21, but negligible because wires have tiny capacitance.)

Remark

If the sum of the voltage drops around a loop is zero, then why would current ever want to flow? After all, if you had a circular tube of water, the water would never flow, because the net drop in height along the circle is zero. The reason current flows in circuits with batteries is that within the battery, charges are moved from lower to higher electric potential energy, just like how a pump could be used to move water upward to start a liquid circuit, by an "electromotive force".

But this immediately raises the question: what *is* this specific force? It can't be the electric force, because we just established that it's pointing the wrong way. It's not a magnetic effect. For some setups, it is literally a mechanical force like a pump: in the Van der Graaff generator, a motor drives the charges on a statically charged conveyor belt to higher potential. But that's not how batteries work.

In a battery, there is no specific force pushing charges from low to high electric potential. Instead, the charges just jiggle around randomly, and the result emerges from the effects of their many collisions. To understand this, consider a gravitational analogy.

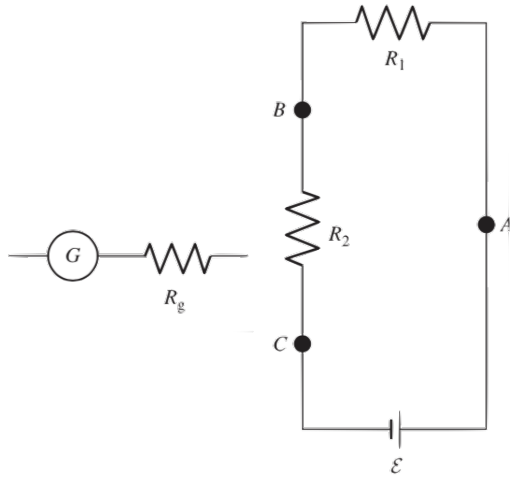


Consider an ideal gas at temperature T released in the trough shown above. The gas molecules will randomly collide, sometimes being propelled upward by chance. Sometimes, a gas molecule will climb the hill and fall into the deep hole, at which point it is unlikely to come out again. Thus, if the hole begins empty, it is energetically favorable for gas molecules to fill it. But there is no attractive force pulling molecules up along the slope! Gravity always points down; molecules go up the slope when they are randomly bounced that way.

This is essentially how the potential in an initially neutral battery is set up. The hole corresponds to the lower energy state an electron can reach inside the anode, but there is no long-range force pushing it there, just the average effect of random collisions.

[2] **Problem 28** (Purcell 4.10). The basic ingredient in older voltmeters and ammeters is the galvanometer, a device to measure very small currents. (It works via magnetic effects, but the exact

mechanism isn't important here.) Inherent in any galvanometer is some resistance R_g, so a physical galvanometer can be represented by the system shown below.



Consider a circuit such as the one shown, with all quantities unknown. We want to measure the current flowing across point A and the voltage difference between points B and C. Given a galvanometer with known R_g, and also a supply of known resistors (ranging from much smaller to much larger than R_g), how can you accomplish these two tasks? Explain how to construct your two devices (called an ammeter and voltmeter), and also how you should insert them in the given circuit. You will need to make sure that you (a) affect the given circuit as little as possible, and (b) don't destroy your galvanometer by passing more current through it than it can handle.

Remark

Occasionally, you might see Olympiad problems where a voltmeter is connected in series. The most common voltmeters are handheld digital multimeters, where the voltmeter setting presents a resistance of about $10\,\mathrm{M}\Omega$. Thus, for such problems, you should just treat the voltmeter like a high-resistance resistor.

Is this realistic? Well, it certainly happens every day, in almost every introductory physics lab in the world. But no professional would ever do this on purpose, because voltmeters aren't designed to be used this way. There is no guarantee that the resistance of the voltmeter is a constant. Instead, for most digital multimeters, there is a complex circuit inside that adjusts the internal resistance depending on the input and the configuration settings. You probably won't break the voltmeter when you put it in series, but you won't get reliable results either.

[2] **Problem 29.** USAPhO Quarterfinal 2009, problems 3 and 4.

[3] **Problem 30.** INPhO 2021, problem 1. A nice problem on practical circuit measurements. Note that the question statement is a bit vague. You are supposed to keep track of quantities of order R_A/R and R/R_V, but you are allowed to neglect quantities as small as R_A/R_V.

Electromagnetism II: Electricity

Chapters 3 and 4 of Purcell cover the material presented here, as does chapter 6 of Wang and Ricardo, volume 2. Image charges are covered in more detail in section 3.2 of Griffiths. For an array of interesting physical examples, see chapters II-6 through II-9 of the Feynman lectures. There is a total of **77** points.

1 The Method of Images

Idea 1

The method of images can be used in some highly symmetric situations to compute the electric field in the vicinity of a conductor. Specifically, consider any configuration of static charges and take any equipotential surface containing some of the charges. Then the resulting field configuration outside that surface is the field configuration we would have if that surface bounded a conductor. This is simply because it has constant potential on the conductor surface, so it must be the right answer by the uniqueness theorem.

[4] **Problem 1.** The simplest application of the method of images is the case of a charge q a distance a from an infinite grounded conducting plane. This problem explores some of its subtleties, assuming you've already read the basic treatment in section 3.4 of Purcell.

(a) Find the force on the charge.

(b) Find the work needed to move the charge out to infinity. (Answer: $q^2/16\pi\epsilon_0 a$.)

(c) Find the total potential energy of the charges on the conducting plane, i.e. the potential energy associated only with their interaction with each other. (Answer: $q^2/16\pi\epsilon_0 a$.)

(d) Now suppose there is another parallel grounded conducting plane on the other end of the charge, a distance b away. How many image charges are needed now? Draw some of them.

(e) A conducting plane forces the electric field to be perpendicular to it. Suppose we modified the plane so that the electric field was instead always *parallel* to it. Find the force on the charge.

Solution. (a) The field above the plate from the screening charges on the plate can be mimicked by placing an image charge $-q$ below the plane. This follows because in both cases, the plane of the plate is an equipotential of potential 0. Thus, the force on the charge is $q^2/(16\pi\epsilon_0 a^2)$ pointing towards the plane.

(b) We just integrate the expression from before to get $q^2/(16\pi\epsilon_0 a)$.

It's tempting to just write down $q^2/(8\pi\epsilon_0 a)$ because this is the energy of two charges $\pm q$ separated by $2a$, but we must remember that the image charge isn't a real charge. It instead describes the effects of all the screening charges on the plane. You might think it takes work to move the screening charges too, but they can move "for free", since the electric field is perpendicular to the plane throughout this entire process.

(c) Suppose we freeze the plane's screening charges in place, then move the point charge out to infinity. In this case the image charge is stationary, so the work needed is $q^2/(8\pi\epsilon_0 a)$.

Let's think about what this means. There are two components to the initial potential energy: the energy U_1 of the point charge interacting with the screening charges, and the energy U_2 of the screening charges interacting with each other.

In part (b), we showed that it takes work $q^2/(16\pi\epsilon_0 a)$ to move the point charge out, after which point all charges in the problem are widely separated. So

$$U_1 + U_2 = -\frac{q^2}{16\pi\epsilon_0 a}.$$

By the other argument we just made, if we freeze the screening charges and move the point charge away, we are left with just the screening charges. So

$$U_1 = -\frac{q^2}{8\pi\epsilon_0 a}.$$

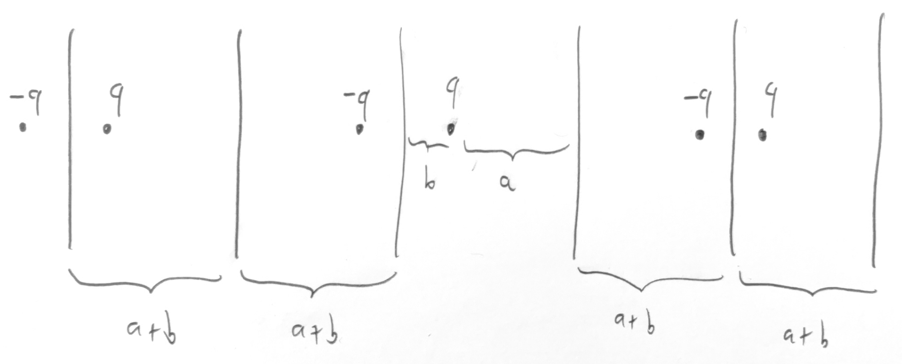
The quantity we're looking for is U_2, which is thus

$$U_2 = \frac{q^2}{16\pi\epsilon_0 a}.$$

Of course this energy is positive, because the screening charges repel each other.

(At first glance this argument may seem to contradict the statement made in part (b), which states that the screening charges cost zero energy to move, as we move the point charge. However, that statement is only true if the screening charges are always allowed to move, so that they preserve the boundary condition $E_\parallel = 0$. But above we instead considered an artificial situation where the screening charges were frozen, and this reasoning no longer holds.)

(d) We actually need infinitely many.



This is the same reason you see infinitely many images of yourself when between two mirrors.

(e) This boundary condition can be satisfied by placing a charge $+q$ at $z = -a$. The force on the charge is thus $kq^2/(4a^2)$ pointing *away* from the plane.

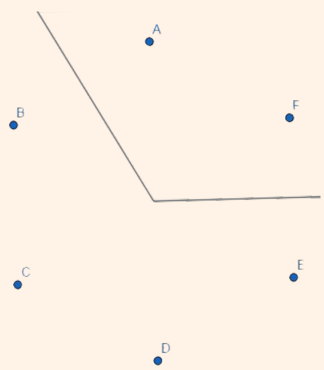
Example 1

Two grounded conducting half-planes intersect, so that in cylindrical coordinates, the equations describing the planes are $\theta = 0$ and $\theta = \theta_p = \pi/2$. A charge q is placed somewhere between the planes. Can the method of images be used to find the force on the charge? What if $\theta_p = 2\pi/3$, or for general θ_p?

Solution

We can solve the first case with three image charges. Suppose the charge q is at (x, y). Then we can reflect in the plane $\theta = 0$, adding a charge $-q$ at $(x, -y)$ to satisfy its boundary condition. Then we can reflect both the real charge and this image charge in the plane $\theta = \pi/2$ to satisfy that plane's boundary condition, adding a charge $-q$ at $(-x, y)$ and a charge q at (x, y).

But when the other plane is at $\theta = 2\pi/3$, there is no configuration of image charges that works. For concreteness, let's suppose the real charge is at point A, on the y-axis.



Reflecting in the $\theta = 0$ plane forces us to have an image charge $-q$ at D, reflecting in the $\theta = 2\pi/3$ plane yields an image charge q at E, and reflecting in the $\theta = 0$ plane again yields a $-q$ charge at F, which is *real* since it's in the same region as A. But this isn't allowed: the point of image charges is to provide an easy way of calculating the effects of screening charges on conducting surfaces on a given set of *real* charges (i.e. the charge at A), so it's not legal to introduce *new* real charges in the process. We would get the same conclusion if we reflected about the planes in a different order – we always need a charge at F. More generally, the method of images works for this problem if and only if $\theta_p = \pi/n$ for integer n.

[4] **Problem 2.** In this problem you'll develop the method of images for spheres.

(a) A point charge $-q$ is located at $x = a$ and a point charge Q is located at $x = A$. Show that the locus of points with $\phi = 0$ is a circle in the xy plane, and hence a spherical shell in space.

(b) Show that the center of the sphere is at the origin provided that

$$a = \frac{r^2}{A}, \qquad -q = -\frac{Qr}{A}$$

where r is the sphere's radius. These results will be used throughout the problem set.

(c) Now suppose a point charge Q is a distance b from the center of a spherical grounded conducting shell of radius r. Find the force on the charge, considering both the cases $b < r$ and $b > r$.

(d) The case $b < r$ is a bit confusing. On one hand, argue that the total charge is Q plus the image charge, and hence nonzero. On the other hand, argue that the total charge must be zero, by considering an appropriate Gaussian surface. One of these arguments is wrong – which one?

As an aside, the fundamental reason the method of images works for spheres is that electromagnetism has conformal symmetry, a symmetry under any local rescaling of space which preserves angles. (One example of a conformal transformation is inversion in Euclidean geometry.) The setup here is related to the conducting plane by such a transformation.

Solution. If you happen to know Euclidean geometry, you can very quickly solve parts (a) and (b) using properties of Apollonian circles. However, for everyone else, we'll present a straightforward solution using coordinates.

(a) The condition for the potential to vanish is

$$\frac{q}{\sqrt{(x-a)^2+y^2}} = \frac{Q}{\sqrt{(x-A)^2+y^2}}.$$

Squaring both sides and clearing denominators, we find

$$(q^2-Q^2)(x^2+y^2) + (q^2A^2 - Q^2a^2) + 2x(aQ^2 - Aq^2) = 0.$$

This has the form of a conic section, and since the coefficients of x^2 and y^2 are equal, it's a circle. (Strictly speaking, it could also be the empty set, since, for example, $x^2 + y^2 = -1$ has no solutions. But we know there have to be places where $\phi = 0$, because it's positive near the positive charge and negative near the negative charge, so it must cross zero by continuity.)

(b) The center of the sphere is at the origin if the coefficient of x vanishes,

$$aQ^2 = Aq^2.$$

Note that this forces a and A to have the same sign. Plugging this in and simplifying, we find

$$x^2 + y^2 = Aa$$

from which we conclude the radius is $r = \sqrt{Aa}$. If you know some Euclidean geometry, at this point you can also recognize that the two point charges are inversions of each other with respect to the sphere. By combining this with the first equation, conclude

$$q = \frac{Qr}{A}.$$

(c) We can treat both cases at once. The image charge is a distance $b' = r^2/b$ from the center of the shell, and its charge is $q' = -q\sqrt{b'/b} = -qr/b$. Therefore, the force on the charge is

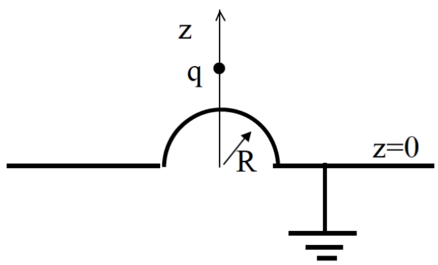
$$F = \frac{qq'}{4\pi\epsilon_0(b-b')^2} = \frac{q^2rb}{4\pi\epsilon_0(b^2-r^2)^2}.$$

It always attracts the point charge towards the nearest point on the surface of the sphere.

(d) The second argument is right. Since the conductor shields the details of the charges inside, the field outside the sphere must be spherically symmetric. But we also know the sphere is at zero potential, so the field outside must be exactly zero, so by using a spherical Gaussian surface, the total charge is zero.

On the other hand, the first argument seems to work for all the other examples we've seen: whenever we have an image charge Q', it seems to represent the effects of a total charge Q' on the conducting surface. But in this case, the method of images doesn't determine the charge distribution on the surface, because adding a constant charge density to the sphere's surface creates absolutely *no* field inside. So in reality, this constant density takes whatever value is necessary to make the sphere overall neutral, in accordance with the second argument.

[2] **Problem 3.** An infinite grounded conducting plane at $z = 0$ is deformed with a hemispherical bump of radius R centered at the origin, as shown. A charge q is placed at $z = a$ as shown.



Can the method of images be used to find the potential in the region with the charge? If so, specify the image charges; if not, explain why not.

Solution. It can be done with three image charges, all on the z-axis:

- A charge $-q$ at $z = -a$.
- A charge $-qR/a$ at $z = R^2/a$.
- A charge qR/a at $z = -R^2/a$.

These ensure that the voltage vanishes on both the whole plane $z = 0$ and on the sphere $r = R$.

[2] **Problem 4** (Purcell 3.50)**.** A point charge q is located a distance $b > r$ from the center of a *nongrounded* conducting spherical shell of radius r, which also has charge q. When b is close to r, the charge is attracted to the shell because it induces negative charge; when b is large the charge is clearly repelled. Find the value of b so that the point charge is in equilibrium. (Hint: you should have to solve a difficult polynomial equation. You can either use a computer or calculator, or use the fact that it contains a factor of $1 - x - x^2$.)

Solution. The image charge has value $-q' = -qr/b$, and is at position r^2/b. As of now, the spherical shell has total charge $-q'$ (by Gauss's law), so to compensate, we add a charge of value $q + q'$ at the center. Thus, to balance forces we have

$$\frac{q'}{(b - r^2/b)^2} = \frac{q + q'}{b^2}.$$

Defining $x = r/b$, this simplifies to

$$x = (1 + x)^3(1 - x)^2$$

which is a rather intimidating quintic equation. At this point you can pull out your calculator, but amazingly, it factors as

$$(1 - x - x^2)(1 + x - x^3) = 0.$$

The only root with $0 < x < 1$ is from the quadratic, $x = (\sqrt{5} - 1)/2$, giving

$$b = \frac{1 + \sqrt{5}}{2} r.$$

[2] **Problem 5.** Consider a neutral spherical conductor of radius R placed in a uniform external field $\mathbf{E}_0$. By using the method of images, argue that the field created by the sphere is the same as the field of an infinitesimal dipole at the center of the sphere, and find its dipole moment. (Hint: suppose the external field is sourced by two point charges that are very far away.)

Solution. Suppose $\mathbf{E}_0 = E_0\hat{\mathbf{z}}$. Let the conductor be centered at the origin. Now, consider a point charge Q at $z = -L$, and $-Q$ at $z = L$ for $L \gg R$. This clearly generates a uniform electric field near the origin (for $r \ll L$) with magnitude $2\frac{Q}{4\pi\epsilon_0 L^2}$. So we simply choose $Q = 2\pi\epsilon_0 L^2 E_0$. Now, the image charges are very close to the origin and form a dipole with moment $2Q'L' = 2(QR/L)(R^2/L) = \boxed{4\pi\epsilon_0 R^3 E_0}$.

By the way, you might have noticed that the uniform field near the sphere could also have been created by a single point charge $2Q$ at $z = -L$, with no need for a second point charge at $z = L$. In that case, the solution is a bit different but the conclusion is the same. Before, we had two image charges at $z = \pm L^2/R$. Now, there's no image charge at $z = L^2/R$, while the one at $z = -L^2/R$ has double the charge. However, there also needs to be an image charge at $z = 0$ with the opposite charge, since the sphere is overall neutral. These two image charges form an image dipole with double the charges and half the displacement, so the same dipole moment.

[5] **Problem 6** (Purcell 3.45). [A] Consider a point charge q located between two parallel infinite grounded conducting planes. The planes are a distance ℓ apart, and the point charge is a distance b from the left plane. The goal of this problem is to find the total charge induced on each plane.

(a) Argue that the total charge on each plane would not change if we replaced the point charge q with two point charges $q/2$, both a distance b from the left plane. By iterating this process, convert the point charge into a uniformly charged plane, and use this to get the answer.

(b) Alternatively, using image charges, show that the electric field on the inside surface of the left plane, perpendicular to the plane, at a point a distance r from the axis containing all the image charges, satisfies

$$4\pi\epsilon_0 E_\perp = \sum_{n=-\infty}^{\infty} \frac{2q(2n\ell + b)}{((2n\ell + b)^2 + r^2)^{3/2}}.$$

(c) Since $\sigma = -\epsilon_0 E_\perp$, we can integrate both sides to find the total charge on the left plane. However, the integral of each term by itself is simply q, so the series doesn't converge. To get the result, do the following steps in this specific order: group the terms $\pm n$ together, then integrate them only out to a distance $R \gg b$, then sum over the values of $|n|$, then take the limit $R \to \infty$. You should get a finite result that matches that of part (a). As you'll probably see in the process, if you do the steps in any other order, you'll get a nonsensical answer.

Those concerned with mathematical rigor might be bothered by the many choices made in part (c). You might ask, couldn't we have gotten a different result by changing how we did the computation? In fact, by the Riemann rearrangement theorem, we could have gotten almost *any* result. But the way we did it is the physically correct way. It roughly sums the terms "in to out", which respects the fact that real plates are finite. Closely related ideas are used to "cancel infinities" in quantum field theory, in a process known as renormalization. We'll see another example in **X1**.

Solution. (a) Consider the induced charge distribution for one point charge q. By the superposition principle, the boundary conditions in this case are also satisfied if we take half that charge distribution, and center it about each of the charges $q/2$. By uniqueness, this is the solution.

Extrapolating to infinitely many charges, we get a uniform plane of charge. For the two plates to be at the same voltage, the electric fields to the left and right of the plane must have ratio $b/(\ell - b)$. Then the charges have the ratio $b/(\ell - b)$ and sum to $-q$, so the charge on the near plate is $-q(\ell - b)/\ell$, while the charge on the far plate is $-qb/\ell$.

(b) We do this by summing over image charges. As we can see from the figure in the solution to problem 1, there are infinitely many image charges. The $n = 0$ term in the sum corresponds to the image charge and real charge closest to the left plane. The $n = 1$ term corresponds to the two image charges a distance $2\ell + b$ from the left plane, while the $n = -1$ term corresponds to the image charges a distance $2\ell - b$ from the left plane, and so on.

(c) Applying Gauss's law, $\sigma = \epsilon_0 E_\perp$, and grouping the terms $\pm n$ together, we have a total charge on the left plane of

$$\begin{aligned} Q &= \int_0^\infty (-\epsilon_0 E_\perp) \cdot 2\pi r\, dr \\ &= q \int_0^\infty \left[-\frac{br}{(b^2 + r^2)^{3/2}} + \sum_{n=1}^\infty \left(\frac{(2n\ell - b)r}{((2n\ell - b)^2 + r^2)^{3/2}} - \frac{(2n\ell + b)r}{((2n\ell + b)^2 + r^2)^{3/2}} \right) \right] dr. \end{aligned}$$

The first term integrates to 1, so we will deal with the sum. Consider one term for some given n, and say we integrate out to some finite but large R. Note that we can only assume $b \ll R$, not $n\ell \ll R$. We will drop all factors of b^2. The term integrates out to

$$\begin{aligned} & -\frac{2n\ell - b}{\sqrt{4n^2\ell^2 + R^2 - 4n\ell b}} + \frac{2n\ell + b}{\sqrt{4n^2\ell^2 + R^2 + 4n\ell b}} \\ \approx & -\frac{2n\ell - b}{\sqrt{R^2 + 4n^2\ell^2}} \left(1 + \frac{1}{2} \frac{4n\ell b}{R^2 + 4n^2\ell^2} \right) + \frac{2n\ell + b}{\sqrt{R^2 + 4n^2\ell^2}} \left(1 - \frac{1}{2} \frac{4n\ell b}{R^2 + 4n^2\ell^2} \right) \\ = & \frac{1}{\sqrt{R^2 + 4n^2\ell^2}} \left(b - 2n\ell + 2n\ell + b + \frac{-2n\ell b(2n\ell - b) - 2n\ell b(2n\ell + b)}{R^2 + 4n^2\ell^2} \right) \\ = & \frac{2R^2 b}{(R^2 + 4n^2\ell^2)^{3/2}} \\ = & \frac{2b}{R} \frac{1}{(1 + (2n\ell/R)^2)^{3/2}}. \end{aligned}$$

Since $R \to \infty$, this sum in this limit can be written as an integral, giving

$$\frac{2b}{R} \int_0^\infty \frac{1}{(1 + (2n\ell/R)^2)^{3/2}}\, dn = \frac{2b}{R} \frac{R}{2\ell} \int_0^\infty \frac{dx}{1 + x^2} = \frac{b}{\ell}.$$

Therefore, the total charge is simply $\boxed{-q(1-b/\ell)}$.

If you found the analysis in part (c) quite tricky, know that you're in good company. About 10 papers have been written about this exact system in the American Journal of Physics alone!

2 Capacitors

Idea 2

There are multiple definitions of capacitance. For a single, isolated conductor with charge Q, the self-capacitance is defined as

$$Q = C\phi$$

where ϕ is the potential difference between the conductor and infinity. But for a set of two isolated conductors with charges $\pm Q$, you can also define a "mutual" capacitance by

$$Q = C\phi$$

where ϕ is the potential difference between the two conductors. When someone talks about a "capacitance" without specification, such as in idea 4, they probably mean this latter one.

Idea 3

The definitions of C above are only useful when you have only one or two conductors in the problem, respectively. In a situation with more than two, it's very tricky to use the above definitions, because all the conductors will affect each other; even a neutral conductor will have an effect since there will be induced charges on its surface.

Instead, it's better to revert to more general principles. The underlying principle behind capacitance is linearity: the charges are linearly related to the fields. For multiple capacitors, the most general possible linear relation is

$$Q_i = \sum_j C_{ij}\phi_j$$

where conductor i has charge Q_i and potential ϕ_i, the potential is taken to be zero at infinity, and the C_{ij} are called general coefficients of capacitance, or in electrical engineering, the Maxwell capacitance matrix. Similarly, inverting this relation,

$$\phi_i = \sum_j p_{ij}Q_j$$

where the p_{ij} are called coefficients of potential.

In Olympiad physics, you'll almost never want to compute the C_{ij} or p_{ij} explicitly. Instead, the point here is that if you're given the charges and want the potentials, or vice versa, you can build up the answer you want using the principle of superposition, computing all the fields you need, e.g. using Gauss's law.

Remark

General capacitance coefficients are discussed further in section 3.6 of Purcell. One nontrivial fact is that $C_{ij} = C_{ji}$, which is proven by energy conservation in problem 3.64 of Purcell. Capacitance coefficients can be clunky to work with. For example, suppose you want to compute the familiar capacitance of a system of two conductors. By definition, we have

$$Q_1 = C_{11}\phi_1 + C_{12}\phi_2, \quad Q_2 = C_{21}\phi_1 + C_{22}\phi_2.$$

An ordinary two-plate capacitor corresponds to the special case of opposite charges on the plates, so we write $Q = Q_1 = -Q_2$. There is a potential difference V across the plates, so $\phi_1 = \phi_2 + V$, and plugging this in gives

$$Q = (C_{11} + C_{12})\phi_2 + C_{11}V, \quad -Q = (C_{22} + C_{21})\phi_2 + C_{21}V.$$

Eliminating ϕ_2 from the system of equations above, we find the familiar mutual capacitance

$$C = \frac{Q}{V} = \frac{C_{11}C_{22} - C_{12}^2}{C_{11} + C_{22} + 2C_{12}}$$

where we used $C_{12} = C_{21}$. This is quite an inconvenient formula, so as a result we won't consider general capacitance coefficients any further, except briefly for practice in problem 8.

[2] **Problem 7** (Purcell 3.21)**.** Consider a capacitor made of four parallel plates with large area A, evenly spaced with small separation s. The first and third are connected by a wire, as are the second and fourth. What is the capacitance of the system?

Solution. Say we charge the conductors to equal and opposite charges. Then by symmetry, the surface charges are

$$\sigma_1, -\sigma_2, \sigma_2, -\sigma_1$$

reading left to right (1 to 4). The field in between the first and the second plates is σ_1/ϵ_0, and the field between the second and the third plates is $(\sigma_1 - \sigma_2)/\epsilon_0$. Since 1 and 3 are connected, and 2 and 4 are, we must have that

$$\sigma_1 s = (\sigma_2 - \sigma_1)s \implies \sigma_2 = 2\sigma_1.$$

Thus, the potential difference is $\sigma_1 s/\epsilon_0$, so the capacitance is $C = Q/\phi = 3\sigma_1 A/(\sigma_1 s/\epsilon_0) = 3\epsilon_0 A/s$.

[3] **Problem 8.** Consider two concentric spherical metal shells, with radii $a < b$.

(a) Compute their capacitance using Gauss's law.

(b) Compute their capacitance by computing the four capacitance coefficients, verifying that $C_{12} = C_{21}$ along the way, and using the result for C above.

Solution. (a) Let the shells have charge $\pm Q$. The field between the shells is $(Q/4\pi\epsilon_0 r^2)\hat{\mathbf{r}}$, so

$$V = \frac{Q}{4\pi\epsilon_0}\left(\frac{1}{a} - \frac{1}{b}\right).$$

Thus the capacitance is

$$C = \frac{Q}{V} = 4\pi\epsilon_0 \frac{ab}{b-a}.$$

(b) Let the first capacitor be the inner shell. If only the outer shell is charged, with charge Q, then $\phi_1 = \phi_2 = Q/4\pi\epsilon_0 b$. The general capacitance equations in this case are

$$0 = C_{11}\phi_1 + C_{12}\phi_2, \quad Q = C_{21}\phi_1 + C_{22}\phi_2$$

from which we see that

$$C_{11} + C_{12} = 0, \quad C_{21} + C_{22} = 4\pi\epsilon_0 b.$$

Now suppose only the inner shell is charged, with charge Q. In this case we have $\phi_1 = Q/4\pi\epsilon_0 a$ while $\phi_2 = Q/4\pi\epsilon_0 b$, so

$$Q = C_{11}\phi_1 + C_{12}\phi_2, \quad 0 = C_{21}\phi_1 + C_{22}\phi_2$$

from which we see that

$$\frac{C_{11}}{a} + \frac{C_{12}}{b} = 4\pi\epsilon_0, \quad \frac{C_{21}}{a} + \frac{C_{22}}{b} = 0.$$

Solving these four equations for the capacitance coefficients gives

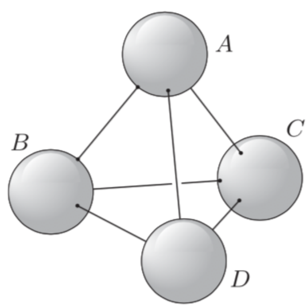
$$C_{11} = 4\pi\epsilon_0 \frac{ab}{b-a}, \quad C_{12} = C_{21} = -4\pi\epsilon_0 \frac{ab}{b-a}, \quad C_{22} = 4\pi\epsilon_0 \frac{b^2}{b-a}.$$

Plugging into the general formula, we have

$$C = \frac{4\pi\epsilon_0}{b-a} \frac{(ab)b^2 - (ab)^2}{ab + b^2 - 2ab} = \frac{4\pi\epsilon_0}{b-a} \frac{b^2 a(b-a)}{b(b-a)} = 4\pi\epsilon_0 \frac{ab}{b-a}.$$

This is certainly a longer route to get to the same conclusion!

[3] **Problem 9** (MPPP 152)**.** Four identical non-touching metal spheres are positioned at the vertices of a regular tetrahedron, as shown.



A charge $4q$ given to sphere A raises it to a potential V. Sphere A can also be raised to potential V if it and one of the other spheres are each charged with $3q$. What must be the size of equal charges given to A and two other spheres for the potential of A to again be raised to V? What if all four spheres are used?

Solution. Suppose sphere A has charge $4y + 3x + 3x$, and spheres B and C have charges of $3x$. We can write this a superposition of three cases, $(4y, 0, 0, 0) + (3x, 3x, 0, 0) + (3x, 0, 3x, 0)$. The potential at A is then just $(y/q)V + 2(x/q)V$ because of linearity (one can use the capacitance coefficients to formalize this). Therefore, we want to set $V = (y/q) + 2(x/q)V$, so $2x + y = q$. We also want $4y + 3x + 3x = 3x$, so $y = -3x/4$. Thus, $2x - 3x/4 = q$, so $x = 4q/5$ and $y = -3q/5$. Thus, the charges are $\boxed{12q/5}$.

We know consider the case of all four spheres being used. Suppose sphere A has charge $4y + 3x + 3x + 3x$, and all others have charge $3x$. The potential at A is then $(y/q)V + 3(x/q)V = V$, so $3x + y = q$. Also, $4y + 9x = 3x$, so $y = -3x/2$. Thus, $3x - 3x/2 = q$, so $x = 2q/3$ and $y = -q$. Thus, the charge is $3x = \boxed{2q}$.

[3] **Problem 10.** USAPhO 2008, problem A1.

Idea 4

A two-plate capacitor with voltage difference V and mutual capacitance C stores energy

$$U = \frac{1}{2}QV = \frac{1}{2}CV^2.$$

Many circuits have multiple two-plate capacitors. In general, these need to be handled with the capacitance coefficients introduced in idea 3. But in practice, capacitors used in circuits are designed to produce fields confined within themselves, so that different capacitors don't interact with each other. In that case, we can just use mutual capacitance throughout, and C adds in parallel, while $1/C$ adds in series. (But this not work if, e.g. you put one capacitor inside another, in which case you should think about the charges and fields directly.)

[2] **Problem 11** (Purcell 3.24)**.** Some estimates involving capacitance.

(a) Estimate the capacitance of the Earth.

(b) Make a rough estimate of the capacitance of the human body.

(c) By shuffling over a nylon rug on a dry winter day, you can easily charge yourself up to a couple of kilovolts, as shown by the length of the spark when your hand comes too close to a grounded conductor. How much energy would be dissipated in such a spark?

Solution. (a) The Earth is a sphere of radius of order 10^7 m, so

$$C = 4\pi\epsilon_0 r \sim 10^{-3}\,\text{F}.$$

Notably, we can make "bigger" capacitors in the lab! Nonetheless, a huge amount of charge can be delivered to the Earth, such as by lightning strikes. This is because the voltage of the Earth is also huge, which is possible because its huge size means the corresponding electric fields aren't that big.

(b) A human is approximately a sphere of radius 0.5 m. Then, $C = 4\pi\epsilon_0 r \sim 5 \times 10^{-11}\,\text{F}$.

(c) We have $U = \frac{1}{2}CV^2 \sim 10^{-4}\,\text{J}$.

[2] **Problem 12.** The total energy can also be found by integrating the electric field energy,

$$U = \frac{\epsilon_0}{2}\int E^2\,dV.$$

(a) Show that this agrees with $U = CV^2/2$ for a parallel plate capacitor.

(b) Show that this agrees with $U = CV^2/2$ for a capacitor made of concentric spheres.

The general proof is more advanced, but if you're interested, one slick method is given in problem 1.33 of Purcell.

Solution. (a) Let the plate area be A and the distance between them be d. Then

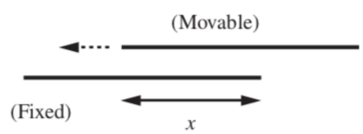
$$U = \frac{\epsilon_0}{2} E^2(Ad) = \frac{\sigma^2}{2\epsilon_0} Ad = \frac{C}{2} \frac{\sigma^2 d^2}{\epsilon_0^2} = \frac{CV^2}{2}.$$

(b) Let the radii be R_1 and R_2 and the charges be $\pm Q$. The field is $Q/(4\pi\epsilon_0 r^2)$, so

$$U = \frac{\epsilon_0}{2} \int \frac{Q^2}{16\pi^2\epsilon_0^2} \frac{dV}{r^4} = \frac{Q^2}{32\pi^2\epsilon_0} \int_{R_1}^{R_2} \frac{4\pi r^2\, dr}{r^4} = \frac{Q^2}{8\pi\epsilon_0} \left(\frac{1}{R_1} - \frac{1}{R_2} \right).$$

On the other hand, this should be equal to $U = QV/2$, which follows directly from the result of problem 8.

[3] **Problem 13** (Purcell 3.26)**.** A parallel-plate capacitor consists of a fixed plate and a movable plate that is allowed to slide in the direction parallel to the plates. Let x be the distance of overlap.



The separation between the plates is fixed. Let $C(x)$ be the capacitance.

(a) Assume the plates are electrically isolated, so that their charges $\pm Q$ are constant. By differentiating the energy, find the leftward force on the movable plate in terms of Q and $C(x)$.

(b) Now assume the plates are connected to a battery, so that their potential difference ϕ is held constant. Find the leftward force on the movable plate, in terms of ϕ and $C(x)$.

(c) If the movable plate is held in place, the two answers above should be equal because nothing is moving. Verify that this is the case, being careful with signs.

(d) In terms of electric fields, why is there a force on the movable plate? Does the effect invoked in the answer to this part change the conclusion of parts (a) through (c) at all?

Solution. (a) The energy as a function of x is

$$U(x) = \frac{Q^2}{2C}$$

where we understand that C is also a function of x. Thus, the force on the plate is

$$F = -\frac{dU}{dx} = \frac{Q^2}{2} \frac{d}{dx} \left(-\frac{1}{C} \right) = \frac{Q^2}{2C^2} \frac{dC}{dx}.$$

(b) Here, the energy is $U(x) = \frac{1}{2} C\phi^2$, so naively we have

$$F = -\frac{dU}{dx} = -\frac{\phi^2}{2} \frac{dC}{dx}.$$

This is negative, while the answer to part (a) is positive. The reason is that U should reflect the total energy of the system – and in this case, the system must include the battery that does work to maintain the potential difference ϕ.

Say x increases by dx. Let the change in capacitance be dC. Then, $dQ = \phi\, dC$. Thus, the work the battery does is

$$dW = \phi\, dQ = \phi^2\, dC.$$

Therefore, if F is the net force the plate feels, we have that

$$dW = F\, dx + dU \implies F = \frac{1}{2}\phi^2 \frac{dC}{dx}.$$

(c) Let F_Q be the first force, and F_ϕ the second. We have that

$$F_Q/F_\phi = \frac{Q^2}{\phi^2 C^2} = 1.$$

If we didn't account for the subtlety in part (b), we would have gotten -1 here.

(d) The force is due to the fringe fields, as the non-fringe field is perfectly vertical. Thus, we have managed to compute a fringe field effect using energy, even though we were able to completely ignore the fringe fields in the energy calculation! This is yet another example of how conservation laws can hand you information that's very hard to get otherwise.

Idea 5: Dielectrics

While we're on the subject of capacitors, it's useful to introduce dielectrics. A dielectric is an insulator which polarizes in the presence of an electric field, with positive charges displaced slightly along the field. The resulting electric dipoles distributed throughout the material in turn create a field that tends to weaken the original applied field within the material.

Each part of a dielectric polarizes based on the local electric field, but that electric field depends on the applied field, and the polarization of *every other* piece of the dielectric. Thus, solving for the electric field for a general dielectric geometry is very difficult, and usually not possible in closed form, just like how it's usually not possible to solve for the field of a charged conductor. In Olympiad physics, you will almost always consider highly symmetric situations, where a dielectric simply reduces the applied electric field everywhere inside by a factor of κ, called the dielectric constant. (We'll consider some trickier situations in **E8**.)

Consider a parallel plate capacitor with charge $\pm Q$ on each plate. If a dielectric is inserted with the charge kept the same, then the field inside is reduced by a factor of κ. Thus, the capacitance $C = Q/V$ increases by a factor of κ. Dielectrics may increase the amount of energy that can be stored in a capacitor, which is typically limited by the voltage V_0 where electrical breakdown occurs. So if V_0 stays the same, the maximal stored energy $U = CV_0^2/2$ goes up by a factor of κ.

Plugging in the definition of C, this result implies that the energy density in the capacitor is $\kappa\epsilon_0 E^2$. But we showed in **E1** that the energy density of the electric field is only $\epsilon_0 E^2$. The extra energy is stored in the dielectric material itself: it takes energy to separate positive and negative charges within the dielectric, as if we were stretching many microscopic springs. This potential energy is released when the capacitor is discharged.

3 Tricky Problems

Example 2: PPP 151

A closed body with conducting surface F has self-capacitance C. The surface is now dented so that the new surface F^* is entirely inside F. Prove that the capacitance has decreased.

Solution

The energy stored in the capacitor is $U = Q^2/2C$. Therefore, if we give the capacitor a fixed charge Q, proving that F^* has lower C is equivalent to showing that it takes positive work to dent the foil from F to F^*. It's cleaner to show the other direction, i.e. that starting from F^*, we can get to F while only lowering the energy.

Suppose without loss of generality that F is infinitesimally larger than F^*. (We can break any finite change into infinitesimal stages and repeat this argument.) We can go from F^* to F by just taking each charge on the surface and moving it outward until it hits F. This lowers the energy because the electric field is always directed outward, as we proved in **E1**.

At this point, the charges lie on F, but they don't have the right distribution, i.e. F is not an equipotential. Now we let the charges spontaneously redistribute themselves so that F is again an equipotential. This again lowers the energy, proving the desired result.

Example 3

Are there charge distributions that aren't spherically symmetric, but which produce an *exact* $\hat{\mathbf{r}}/r^2$ field outside of them?

Solution

If you know a bit about the multipole expansion, this might seem like a daunting question. To make the field exactly $\hat{\mathbf{r}}/r^2$, you need to make sure the charge distribution has no dipole moment, no quadrupole moment, no octupole moment, and so on to infinity, and it seems impossible to satisfy all of these constraints without spherical symmetry. But we have *already* seen an example of such a charge distribution earlier in the problem set!

Recall that when we treated the method of images for spheres, we found that in some situations, the complicated charge densities on conducting spheres were exactly the same as those produced by a fictitious image charge inside the sphere, and generally away from its center. If we place the origin at that image charge, then we have an example of a charge distribution that is perfectly $\hat{\mathbf{r}}/r^2$ far away, but which isn't spherically symmetric. (The general solution is given here.)

[2] **Problem 14.** Consider a set of n conducting, very large parallel plates, placed in zero external electric field. The plates are given charges Q_i. If the left ends of the plates are at locations x_i, and the plates have thickness d_i, what is the total charge on the left end of the leftmost plate, and the right end of the rightmost plate?

Solution. If you try to solve this one directly, the calculations can get pretty messy, but there's a simple solution using basic facts about capacitors. First, consider the two leftmost plates. Since they're conductors, the electric field must vanish inside them. Now split the charge in the problem into two parts: (1) the charge on the right side of the leftmost plate, and the left side of the second leftmost plate, and (2) all other charge.

The electric field due to (2) has some uniform value E_{ext} within the first and second plates. Therefore, in order for the electric field to vanish in both of those plates, the electric field due to (1) has to be the *same* in both plates. This is only possible if the two charges in (1) are *opposite*, which corresponds to those sides forming a parallel plate capacitor.

We can then repeat this reasoning, which shows that the charges on the right side of plate i must cancel with the charges on the left side of plate $i+1$, producing no field anywhere except in between plates i and $i+1$. All that is left is the left end of the leftmost plate and the right end of the rightmost plate, which must have total charge $\sum_i Q_i$. Finally, in order to ensure no field in any plate, these two charges must be equal, so the answer is that both have total charge $(\sum_i Q_i)/2$.

[2] **Problem 15** (Purcell 3.9)**.** A conducting spherical shell has charge Q and radius R_1. A larger concentric conducting spherical shell has charge $-Q$ and radius R_2.

(a) If the outer shell is grounded, explain why nothing happens to the charge on it.

(b) If instead the inner shell is grounded, e.g. by connecting it to ground by a very thin wire that passes through a very small hole in the outer shell, find its final charge.

(c) It's not so clear why charge would leave the inner shell in part (b), thinking in terms of forces. A small bit of positive charge will certainly want to hop on the wire and follow the electric field across the gap to the larger shell. But when it gets to the larger shell, it seems like it has no reason to keep going to infinity, because the field is zero outside. And, even worse, the field will point *inward* once some positive charge has moved away from the shells. So it seems like the field will drag back any positive charge that has left. Does charge actually leave the inner shell? If so, what's wrong with the above reasoning?

Solution. (a) The potential at the outer shell due to itself is $-Q/4\pi\epsilon_0 R_2$ and the potential due to the inner shell is $Q/4\pi\epsilon_0 R_2$, so it is zero overall. Thus, the outer shell is already effectively grounded.

(b) The potential at the inner shell due to itself is $Q'/4\pi\epsilon_0 R_1$ and the potential to the outer shell is $-Q/4\pi\epsilon_0 R_2$. Since the total must be zero, $Q' = R_1 Q/R_2$.

(c) The key mistake is that the positive charges are repelled also by the charges behind it in the wire. So yes, eventually the field due to the shells may even become inward, there is a whole line of plus charge behind a given charge that force it forward.

Another way of saying this is that a wire has negligible capacitance; like a thin metal pipe of water, it cannot store extra net charge but can only let charge move rigidly through the entire thing. It is energetically favorable for this to happen, so even if some charges don't want to move forward, their neighbors will push them forward.

[2] **Problem 16.** The usual expression for the capacitance of a parallel-plate capacitor is $A\epsilon_0/d$. However, in reality the field within the capacitor is not perfectly uniform, and there are fringe fields outside. Taking these effects into account, is the true capacitance higher or lower than $A\epsilon_0/d$?

Solution. For concreteness, suppose the plates are circular disks, with charge $\pm Q$, and consider the electric field along the axis of symmetry. In the naive derivation, we assume the charge density on the plates is uniform. Then we approximate the plates as infinite in order to use Gauss's law to conclude that the field inside is $\sigma/\epsilon_0 = Q/A\epsilon_0$. This is inaccurate for two reasons:

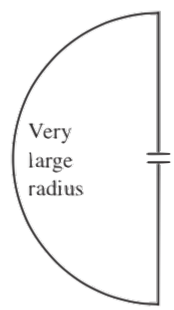
- The plates are not actually infinite, so the field on the symmetry axis should actually be smaller.
- The charge distribution is not actually uniform. Instead, since the like charges on each plate repel each other, some charge gets pushed outward. This further decreases the field on the symmetry axis.

Therefore, the field on the symmetry axis is less than σ/ϵ_0, so the voltage drop is lower, which implies the capacitance is higher.

[2] **Problem 17** (Purcell 4.16)**.** In a parallel plate capacitor, the quantity $\int \mathbf{E} \cdot d\mathbf{s}$ should be equal to V for any path that connects the two plates.

A charged capacitor can be discharged by attaching a wire to the external surfaces of the plates. No matter how one attaches the wire, $\int \mathbf{E} \cdot d\mathbf{s}$ along the wire should be equal to V. And as we've argued in **E2**, this is sufficient to cause charges to move along the wire, even if the electric field points in the "wrong" direction at some points along the wire, because the wire has negligible capacitance: charges within it move rigidly, each pushing the next one and pulling the previous one.

But it's puzzling how this works for a capacitor, because the electric field is supposed to be essentially zero just outside it. Consider two possible limiting cases for the wire's shape.



In each case, explain qualitatively how $\int \mathbf{E} \cdot d\mathbf{s}$ can be equal to V. In particular, how large are the contributions from the distinct segments of the wire (the horizontal and vertical parts in the first case, and the straight and curved parts in the second)?

Solution. In the first case, the horizontal parts of the wire contribute almost nothing. That's because the radial part of the electric field vanishes within the conductor plates themselves (since they must be equipotentials), and the horizontal path is right next to the plates. Therefore, the contribution is almost entirely from the vertical segment.

You might be wondering how this is possible, because at the edge of the plates, there seems to be less charge nearby, so the electric field should be smaller than near the middle of the plates. The resolution is that the surface charge density near the edge of plates is a lot higher than the surface charge density near the middle, because like charges repel.

It's interesting to compare this to the case of two parallel plates with *uniform* charge density. In that case, the vertical segment contributes roughly $V/2$. To see why, consider putting a second, identical parallel plate capacitor directly to the left of the first one. Now the vertical segment is in the middle of a big capacitor, and has voltage drop V. So each of the two halves of that big capacitor contributes $V/2$ to the vertical segment. However, the total voltage drop is still V, because for uniform charge density there are substantial horizontal fields, so that the horizontal segments contribute roughly $V/4$ each.

In the second case, the result is due to the far-field behavior. When you zoom out, the capacitor looks like a dipole, so the field at long distances is a dipole field. Now, the dipole field falls off as $1/r^3$, and the circumference of the curved part is proportional to r, so the contribution of this part goes as $r/r^3 \to 0$ as $r \to \infty$. So all the contribution is from the straight part.

To see how this can be the case, note that the vertical field just above the capacitor plates is negligible; the dipole field only kicks in once we're far enough so that the plates look small, i.e. subtending a small angle from our perspective. If the plates are squares of side length a, this occurs at a distance of order a. Then

$$\int \mathbf{E} \cdot d\mathbf{s} \sim 2 \int_a^\infty \frac{p}{2\pi\epsilon_0 r^3}\, dr = \frac{2}{\pi} \frac{p}{\epsilon_0 a^2}$$

where p is the electric dipole moment. If the plates are separated by a distance $d \ll a$, then $p = Qd = \sigma a^2 d$, giving

$$\frac{2}{\pi} \frac{p}{\epsilon_0 a^2} = \frac{2}{\pi} \frac{\sigma d}{\epsilon_0} = \frac{2}{\pi} V$$

which is on the order of the voltage across the capacitor plates. Of course, we didn't get precisely V because we made a lot of approximations in the calculation, but this illustrates the conceptual point: the full integral of $\mathbf{E} \cdot d\mathbf{s}$ can indeed be equal to V, and most of the contribution to this integral comes from the part of the vertical wire which is a distance of order a from the capacitor.

[3] **Problem 18.** Consider two conducting spheres of radius r separated by a distance $a \gg r$. The spheres can be thought of as the two plates of a parallel plate capacitor.

(a) By assuming the surface charge density on each sphere is uniform, estimate the capacitance.

(b) In reality, the surface charges on each sphere will be distorted by the other sphere. The surface charges assumed in part (a) will induce changes in the surface charges, represented by an image charge for each sphere. But these image charges will induce further image charges, and so on, yielding an infinite series for the charge distribution and hence the capacitance. Find the first two terms in the series for the capacitance.

(c) If $a/r = 10$, roughly estimate the error in neglecting the other terms.

Solution. (a) The potential of a single sphere, relative to infinity, is $V = Q/4\pi\epsilon_0 r$. Hence the two spheres, at this level of approximation, have potentials $\pm Q/4\pi\epsilon_0 r$, and the capacitance is

$$C = \frac{Q}{V_1 - V_2} = 2\pi\epsilon_0 r.$$

This is ignoring any interaction between the charges on different spheres.

(b) The uniform charge on one sphere looks like a point charge to the other sphere, by the shell theorem, and this induces an image charge. The same applies for the spheres in reverse. The image charges are smaller than the original ones by a factor of r/a, by the result of problem 2.

Now, these "first-order" image charges on each sphere in turn induce second-order image charges in the other sphere, which are smaller by another factor of r/a, and so on. Once we sum up an infinite series of image charges, we get the true charge configuration, from which we can compute the exact capacitance, which is now a power series in r/a.

In this part we're only interested in the first two terms. We found the first term in part (a), and the next term comes from accounting for the "first-order" image charges. Let the spheres have charges $\pm Q$. The negative sphere induces an image charge of Qr/a within the positive sphere, so that the negative sphere plus the image charge yield a total of zero potential on the positive sphere. So the total potential of the positive sphere is still $V = Q/4\pi\epsilon_0 r$, except that the total charge on the sphere is not Q, but $Q(1 + r/a)$. Applying the same reasoning in reverse to the other sphere,

$$C = \frac{Q(1 + r/a)}{V_1 - V_2} = 2\pi\epsilon_0 r(1 + r/a + O((r/a)^2)).$$

It makes sense the true capacitance is higher than the naive guess of part (a), because in part (a) we demanded a uniform charge density, and charges can lower their energy further if we let them spread out.

(c) As we've seen, the answer is a power series in r/a. So if we include the zeroth and first order terms, the second order and higher terms are suppressed by at least two powers, $(r/a)^2 = 1\%$, so we expect our approximation to be good within a few percent.

[3] **Problem 19.** USAPhO 2022, problem A2. A computational problem involving surface tension.

Example 4

Find the leading interaction force between a dipole of dipole moment p and a grounded conducting sphere of radius r, separated by a distance $R \gg r$. What if the sphere is electrically neutral instead?

Solution

Place the origin at the center of the sphere and orient the z-axis to pass through the dipole. We can regard the dipole $p = qd$ as a combination of two charges

$$q \text{ at } z = R, \quad -q \text{ at } z = R + d$$

where d is very small. In the grounded case, this induces two image charges in the sphere,

$$\frac{qr}{R} \text{ at } z = \frac{r^2}{R}, \quad -\frac{qr}{R+d} \text{ at } z = \frac{r^2}{R+d}$$

approximately separated by r^2/R^2. We can now use Coulomb's law four times, but that's a bit tedious. Instead, decompose the image charges into a dipole moment and a net charge,

$$p' = \frac{pr^3}{R^3}, \quad Q' = \frac{qr}{R} - \frac{qr}{R+d} \approx \frac{pr}{R^2}.$$

We can place both of these at the origin, because this slight displacement will only affect the answer by subleading terms in r/R. Then the corresponding fields, far along the z-axis, are

$$E_{p'}(z) = \frac{2kpr^3}{R^3 z^3}, \quad E_{Q'}(z) = \frac{kpr}{R^2 z^2}.$$

The first term is negligible compared to the second, due to the many powers of R and z in the denominator. Thus, keeping only the second term, the force on the original dipole is

$$F = p \frac{d}{dz} E(z) \bigg|_{z=R} = -\frac{2kp^2 r}{R^5}$$

which falls off very quickly with distance. This derivation illustrates a common subtlety: it might not always be obvious how far to approximate. We threw away terms subleading in r/R, because we only wanted the leading contribution. But if we had applied that principle to the image charges at the first step, we would have thrown out the tiny net charge Q', which actually provides the dominant contribution to the force, because of how tiny p' is.

Now, the situation for a neutral sphere is completely different. By the logic of problem 4, there's a third image at the center of the sphere to enforce neutrality,

$$-\frac{pr}{R^2} \text{ at } z = 0.$$

The image charges can now be decomposed into a combination of two dipole moments. We already saw the first one p' above, while the second is, to leading order

$$p'' \approx \frac{pr}{R^2} \frac{r^2}{R} = \frac{pr^3}{R^3}$$

with the same magnitude and direction as p'. Thus, this system of image charges has approximate dipole moment $2p'$. The corresponding force is

$$F = p \frac{d}{dz} \frac{4\,kpr^3}{R^3 z^3} \bigg|_{z=R} = -\frac{12\,kp^2 r^3}{R^7}$$

which falls off even more quickly with distance. In this derivation, we didn't have to worry too much about getting p'' exactly right, because there was no net charge ("monopole") term that could've overwhelmed the dipole field, so all other field contributions are safely suppressed by more powers of r/R. (Of course, if p'' had come out pointing the opposite direction to p', so that the two almost cancelled, we would've had to be more careful.)

The lesson of this example is *not* to just use exact expressions and Taylor expand at the end. Here, that brute force approach would have required Taylor expanding six Coulomb's law forces out to order $1/R^7$, which is extraordinarily tedious. Instead, to approximate properly, we have to think carefully in every case. Incidentally, when applied to a polar and neutral nonpolar molecule, the $1/R^7$ force above is called the Debye force; it is one of the "van der Waals forces" which are often vaguely described in chemistry classes.

Example 5

Estimate the interaction force between a point charge q and a thin conducting rod of length ℓ, which is a distance $L \gg \ell$ from the charge and oriented along the separation between them.

Solution

The interaction occurs because the point charge induces negative charges on the near end of the rod, and positive charges on the far end. These charges are then acted on by the electric field of the point charge, causing a force.

To get a very crude estimate, let's just suppose that charge Q appears on the far end and charge $-Q$ appears on the near end. The resulting field produced in the middle is

$$E \sim \frac{kQ}{\ell^2}.$$

On the other hand, this needs to cancel a field from the point charge of

$$E \sim \frac{kq}{L^2}$$

which tells us that $Q \sim (\ell/L)^2 q$. The force on the induced charges is

$$F \sim kqQ\left(\frac{1}{L^2+\ell^2} - \frac{1}{L^2}\right) \sim -\frac{kqQ\ell^2}{L^4} \sim -\frac{kq^2\ell^4}{L^6}.$$

Again, the force is attractive, and falls off quickly with distance.

[3] **Problem 20** (Physics Cup 2017)**.** Estimate the interaction force between a point charge q and an infinitely thin circular neutral conducting disc of radius r if the charge is at the axis of the disc, and the distance between the disc and the charge is $L \gg r$.

Solution. The interaction is because charges redistribute on the disc to keep it an equipotential. As an extremely rough approximation, suppose that charge Q appears at the center of the disc and charge $-Q$ appears on the rim. Then essentially by dimensional analysis, the electric field in the disc is

$$E \sim \frac{kQ}{r^2}.$$

On the other hand, the electric field due to the point charge along the disc is of order

$$E \sim \frac{kQ}{L^2}\frac{r}{L}$$

where the r/L factor is from projecting the field along the disc. Then

$$Q \sim -q\frac{r^3}{L^3}.$$

The force is, by Coulomb's law,

$$F \sim kqQ\left(\frac{1}{L^2} - \frac{L}{(L^2+r^2)^{3/2}}\right) \sim \frac{kq^2r^5}{L^7}.$$

A more accurate estimate would get the numeric prefactors.

Example 6

Find the charge distribution on a conducting disc of radius R and total charge Q.

Solution

In general, there are very few situations where the charge distribution on a conductor can be found explicitly. As you've seen, some of the simplest examples can be solved with image charges. Some more complex, two-dimensional examples can be solved with a mathematical technique called conformal mapping. And this special example can be solved with a neat trick.

Consider a uniformly charged spherical shell centered on the origin, and consider a point P inside the shell, on the xy plane. The electric field at point P is zero, by the shell theorem. Recall that in the usual proof of the shell theorem, one draws two cones opening out of P in opposite directions. The charges contained in each cone produce canceling electric fields.

Now imagine shrinking the spherical shell towards the xy plane, so it becomes elliptical. The crucial insight is that the shell theorem argument above still works, for points on the xy plane. When we squash the shell all the way down to the xy plane, it becomes a disc, with zero electric field on it. This is thus a valid charge distribution for a disc-shaped conductor, and by the uniqueness theorem, it's the only one.

By keeping track of how much charge gets squashed to radius $[r, r+dr]$, we find $\sigma(r) \propto R/\sqrt{R^2 - r^2}$, and fixing the proportionality constant gives

$$\sigma(r) = \frac{Q}{4\pi R\sqrt{R^2 - r^2}}.$$

4 Electrical Conduction

We now leave the world of electrostatics and consider magnetostatics, the study of steady currents.

Idea 6

In a conductor with conductivity σ, the current density is

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

Alternatively, $\mathbf{E} = \rho \mathbf{J}$ where ρ is the resistivity. The current and charge density satisfy

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}.$$

The current passing through a surface S at a given time is

$$I = \int_S \mathbf{J} \cdot d\mathbf{S}.$$

Since $\mathbf{J} \propto \mathbf{E}$, we have Ohm's law $V = IR$, where V is the voltage drop across the resistor. The power dissipated in a resistor is $P = IV$. The resistance R adds in series, while $1/R$ adds in parallel.

[2] **Problem 21** (HRK). A battery causes a current to run through a loop of wire.

(a) Suppose the wire makes a sharp corner. How do the charges know to turn around there?

(b) A copper wire with conductivity σ is joined to an iron wire with conductivity $\sigma' < \sigma$. For the current in both sections to be the same, the electric field in the iron wire must be higher. How does that happen?

In general, the surface charge distribution in a DC circuit can be quite complex; the aspects shown in these questions are just the beginning. For more about this, see this paper and this paper.

Solution. (a) The first charges to make it there don't; they just stop at the surface of the wire, due to the attraction from the protons. Once this charge builds up at the kink, it repels the next electrons so that they automatically turn around. This typically occurs extremely quickly, as the relevant timescale is the RC of the wire and C is tiny. The amount of charge required is very small, less than a few hundred electrons.

(b) It's the same story as part (a). The first charges to reach the iron will start moving slower, because the fields are the same. This then causes a buildup of charge at the interface between them, which increases the field in the iron and decreases the field in the copper. In the steady state, the current densities in both are equal. Again, this occurs very quickly and requires very little charge.

[2] **Problem 22** (PPP 22). Two students, living in neighboring rooms, decided to economize by connecting their ceiling lights in series. They agreed they would each install a 100 W bulb in their own rooms and that they would pay equal shares of the electricity bill. However, both tries to get better lighting at the other's expense. The first student installed a 200 W bulb, while the second student installed a 50 W bulb. Which student subsequently failed their final exams?

Solution. A standard bulb is designed to be hooked up in parallel with other bulbs, across some fixed voltage V. Since $P = V^2/R$, higher wattage bulbs have lower resistance.

Since the bulbs are in series, then have the same current through them. Since $P = I^2R$, that means the bulb with the *higher* wattage rating draws *less* power, so the first student fails. The second student's bulb burns four times brighter. (The fact that the brightness of a bulb depends on what else is attached is also a reason why hooking lights up in series is impractical.)

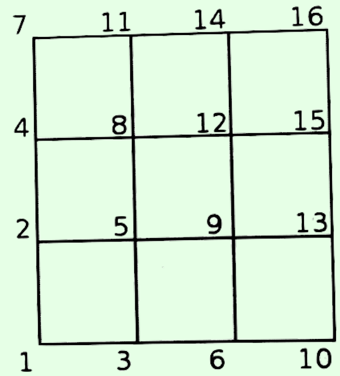
To warm up for DC circuits, we'll consider some resistor network problems.

Idea 7

If any two points in a resistor network are at the same potential, nothing will change if the two points are connected together and treated as one. More generally, the resistance of any resistor directly connecting the two points may be changed freely.

Example 7

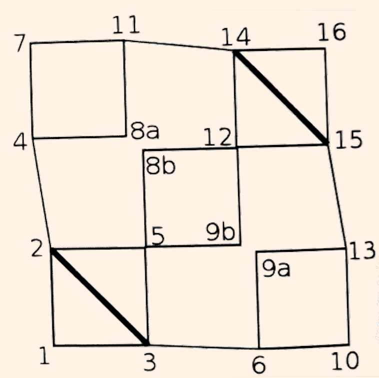
Consider the 3×3 grid below, where every edge is a resistor R.



Find the equivalent resistance between nodes 1 and 16.

Solution

By the above idea, we can short together nodes 2/3, and 14/15, by the diagonal symmetry of the network. Next, we can break nodes 8 and 9 into two pieces.



This is valid because the separated nodes $8a/8b$ and $9a/9b$ still have the same potential in the new network, by the diagonal symmetry. (This is using the above idea in reverse.) Now, the circuit has been reduced to combinations of series and parallel resistors. The resistance between 1 and 2/3 is $R/2$. The resistance between 2/3 and 14/15 is the combination of three networks in parallel, and finally the resistance of 14/15 and 16 is $R/2$. Thus,

$$R_{\text{eq}} = \left(\frac{1}{2} + \left(\frac{1}{3} + \frac{1}{3} + \frac{1}{2}\right)^{-1} + \frac{1}{2}\right) R = \frac{13}{7} R.$$

You won't see any resistor problems as complicated as this one for the rest of the training, because they're kind of contrived; the point of this example was just to show multiple uses of symmetry techniques.

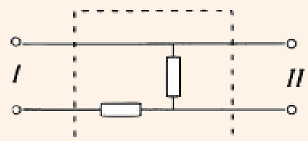
Example 8: PPP 23

A black box contains a resistor network and has two output terminals.

If a battery of voltage V is connected across the first terminal, the voltage across the second terminal is $V/2$. If a battery of voltage V is connected across the second terminal, the voltage across the first terminal is V. Find one possible configuration of the resistors inside the box.

Solution

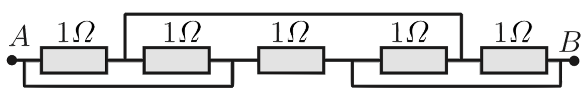
A simple configuration with two equal resistors works.



When a battery is connected across II, the horizontal resistor doesn't do anything. When a battery is connected across I, the two resistors comprise a voltage divider.

[2] **Problem 23.** USAPhO 2007, problem A1.

[2] **Problem 24** (IPhO 1996)**.** Consider the following resistor network.



Find the equivalent resistance between A and B.

Solution. See the official solutions here.

[3] **Problem 25.** Consider a cube of side length L whose edges are resistors of resistance R.

(a) Compute the resistance between two vertices a distance $\sqrt{3}L$ apart.

(b) Compute the resistance between two vertices a distance $\sqrt{2}L$ apart.

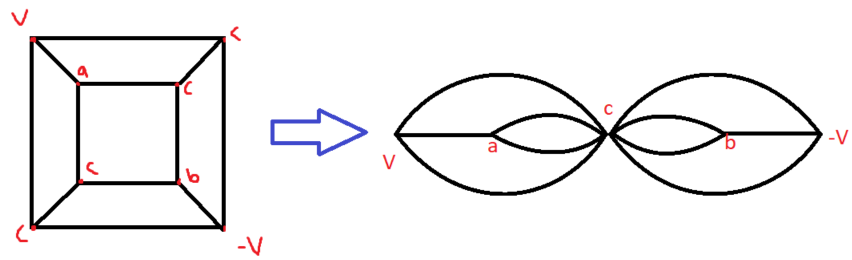
(c) Compute the resistance between two vertices a distance L apart.

(d) Generalize to vertices $\sqrt{n}L$ apart on an n-dimensional cube. (Give your answer in the form of a summation.)

Solution. (a) Let the two vertices be A and B. Let the vertices distance 1 from A be labeled a and distance two labeled b. Note that all the vertices labeled a have the same potential by symmetry, and same for b. Thus, we can treat all the vertices labeled the same thing as one vertex.

We have 3 connections from A to a, 6 from a to b, and 3 from b to B. Thus, our resistance is $R/3 + R/6 + R/3 = \boxed{5R/6}$.

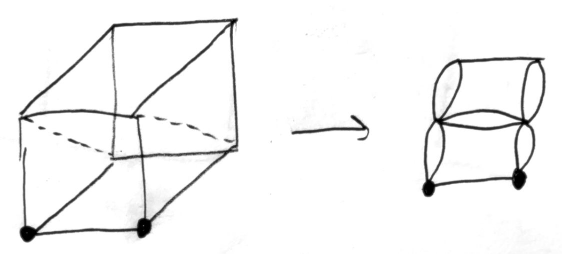
(b) Apply potential V and $-V$ to the vertices, and label the rest of the vertices as shown.



We claim that all the vertices labeled c have potential 0. The idea is then that negating the potentials V and $-V$ must negate the potentials at all vertices. But negating is equivalent to a simple reflection that preserves the locations of the vertices labeled c. Thus the only option is that their potential is 0. Therefore, all the cs can be treated as one vertex, and we have the drawn equivalent circuit. This is a combination of parallel and series, and we compute the answer to be

$$2 \cdot \frac{1}{1+1+\frac{1}{1+1/2}} R = \boxed{3R/4}.$$

(c) In a similar fashion, the points labeled the same below have the same potential, and on the right is the equivalent circuit.



This is again just a series/parallel problem, and we compute the answer to be

$$R \frac{1}{\frac{1}{\frac{1}{2}+\frac{1}{\frac{1}{\frac{1}{2}}+\frac{1}{\frac{1}{2}+1+\frac{1}{2}}}+\frac{1}{2}}+\frac{1}{1}} = \boxed{7R/12}.$$

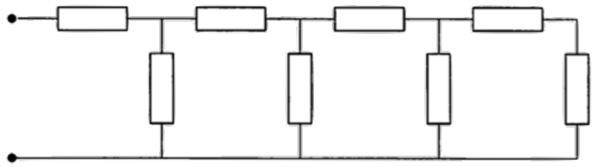
(d) The coordinates take the form $(x_1, x_2, \ldots, x_n)$ where x_i is zero or one. We consider the vertices $(0, 0, \ldots, 0)$ and $(1, 1, \ldots, 1)$. The first vertex is connected to all the vertices with one 1, of which there are n. By symmetry, these are all at the same voltage. Next, these vertices are connected to all the vertices with two 1's, of which there are $\binom{n}{2}$, and so on.

We hence have $n + 1$ effective vertices of different voltages. Consider the vertex representing points with k 1's. The number of connections to points with $k + 1$ 1's is $\binom{n}{k}(n - k)$. Then by adding series and parallel resistances,

$$R_{\text{eq}} = R \sum_{k=0}^{n} \left(\binom{n}{k} (n-k) \right)^{-1}.$$

For example, this recovers the result of part (a) for $n = 3$.

[2] **Problem 26** (PPP 158)**.** Consider the circuit below, where every resistor is $1\,\Omega$.



(a) Find the equivalence resistance between the input terminals.

(b) Do the same in the case where the chain is infinitely long.

Solution. (a) Suppose the rightmost resistor has unit current flowing down, and let V be the potential difference across A and B.

Number the resistors, starting from the back, and going to the left. Let I_k be the current in R_k. We claim that

$$I_k = F_k$$

where the F_k are the Fibonacci numbers $F_1 = F_2 = 1$ and $F_n = F_{n-1} + F_{n-2}$. This is true by induction. Note that

$$I_{2k} = I_{2k-1} + I_{2k-2}$$

by junction law. By the loop law,

$$I_{2k+1} = I_{2k} + I_{2k-1}.$$

Note that all currents are either to the right or down. Let the number of vertical resistors be n. Then, the potential difference is

$$V = I_{2n-1} + I_{2n} = I_{2n+1} = F_{2n+1}.$$

The equivalent resistance is

$$R = (F_{2n+1}/F_{2n})\,\Omega.$$

(b) By taking the limit $n \to \infty$ above, we get the golden ratio,

$$R = \frac{1+\sqrt{5}}{2}\,\Omega.$$

There's another, slicker way to do this. In the infinite case, if we let the answer be R, then we have

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

which is equivalent to the quadratic

$$R^2 - R - 1 = 0, \quad R = \frac{1 \pm \sqrt{5}}{2}.$$

Then taking the positive root gives the answer.

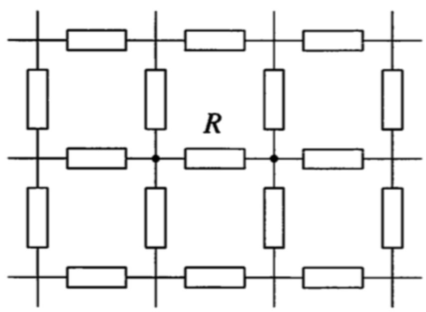
This presents a question: what's the deal with the negative root? Both the positive and negative roots are technically equally good. In fact, there exist circuit elements with negative resistance, as you'll see in **E6**, though they need an active source of power to maintain.

However, the positive root is still better in an important way. Suppose you start out with something on the right end, with (possibly negative) equivalent resistance R_0. If we attach two $1\,\Omega$ resistors on the left, as in the diagram, then we get some new resistance R_1, and attaching two more gives R_2, and so on. The two roots for R found above are the two fixed points for this iteration. However, the positive root is more physical because it's the *stable* fixed point. If you start out with any resistance other than $(1-\sqrt{5})/2$, the recursion will converge to $(1+\sqrt{5})/2$. (For example, in part (a), we showed this happens if you start with $1\,\Omega$.) Only if you start out with the exact negative resistance $(1-\sqrt{5})/2$ will you stay there.

Therefore, even though a mathematically infinite network doesn't have anything "on the right end" (since it has no end at all), it's still meaningful to say that "the" resistance is $(1+\sqrt{5})/2$. When we introduce infinite objects in physics, we usually do so just to get a mathematically tractable approximation for a real, finite object. And for almost any long, but finite chain, you'll get an answer near $(1+\sqrt{5})/2$, so that's the useful answer in the infinite case.

[3] **Problem 27** (PPP 159-161). Superposition can be a useful trick to analyze circuit networks.

(a) Consider an infinite two-dimensional grid of identical resistors R.



Find the equivalent resistance between two neighboring points by considering the superposition of a current I flowing into one point, and an equal current I flowing out the other.

(b) What would the equivalent resistance be if the resistor directly connecting the two neighboring points was removed?

(c) Now consider an icosahedron of identical resistors R. By superposing appropriate current distributions, find the equivalent resistance between two neighboring vertices.

Solution. (a) If we have a current I flowing into any point, then by symmetry, a current $I/4$ flows out along each of the resistors connected to that point. Thus, when we superpose a current I entering one point and a current I leaving another point, the current in the resistor between the points is $I/2$, so the voltage difference between the points is $\Delta V = IR/2$. Thus, the total resistance is $R_{\text{eq}} = \Delta V/I = \boxed{R/2}$.

(b) Suppose the answer is R'. From part (a), R' and R attached in parallel give $R/2$. Thus,

$$\frac{1}{R'} + \frac{1}{R} = \frac{2}{R},$$

so $R' = \boxed{R}$.

(c) Consider the current distribution where I flows into a vertex, and $I/11$ leaves from each other vertex. We see that the current in each edge coming out from the source vertex is $I/5$. So superposing a similar but with signs flipped current distribution for an adjacent vertex, we see that $(12/11)IR_{\text{eff}} = 2RI/5$, so $R_{\text{eff}} = \boxed{11R/30}$.

Idea 8

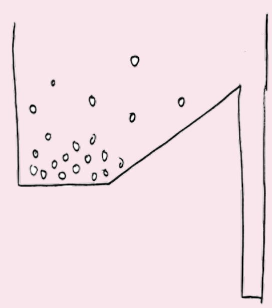
In a circuit of resistors and batteries, Kirchoff's loop rule states that the sum of the voltage drops around a loop is zero. Kirchoff's junction rule states that the net current flowing into a vertex is zero. (This is technically nonzero, because of the effect of problem 21, but negligible because wires have tiny capacitance.)

Remark

If the sum of the voltage drops around a loop is zero, then why would current ever want to flow? After all, if you had a circular tube of water, the water would never flow, because the net drop in height along the circle is zero. The reason current flows in circuits with batteries is that within the battery, charges are moved from lower to higher electric potential energy, just like how a pump could be used to move water upward to start a liquid circuit, by an "electromotive force".

But this immediately raises the question: what *is* this specific force? It can't be the electric force, because we just established that it's pointing the wrong way. It's not a magnetic effect. For some setups, it is literally a mechanical force like a pump: in the Van der Graaff generator, a motor drives the charges on a statically charged conveyor belt to higher potential. But that's not how batteries work.

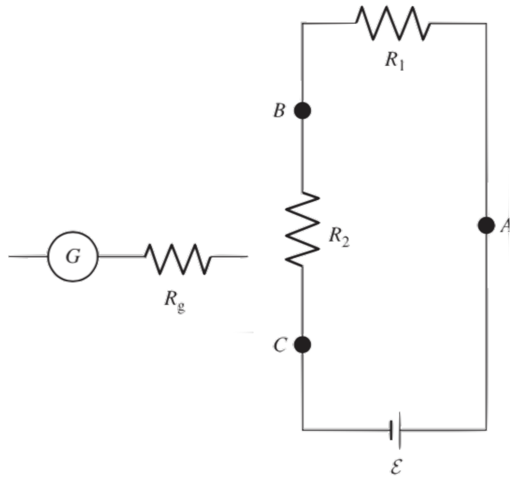
In a battery, there is no specific force pushing charges from low to high electric potential. Instead, the charges just jiggle around randomly, and the result emerges from the effects of their many collisions. To understand this, consider a gravitational analogy.



Consider an ideal gas at temperature T released in the trough shown above. The gas molecules will randomly collide, sometimes being propelled upward by chance. Sometimes, a gas molecule will climb the hill and fall into the deep hole, at which point it is unlikely to come out again. Thus, if the hole begins empty, it is energetically favorable for gas molecules to fill it. But there is no attractive force pulling molecules up along the slope! Gravity always points down; molecules go up the slope when they are randomly bounced that way.

This is essentially how the potential in an initially neutral battery is set up. The hole corresponds to the lower energy state an electron can reach inside the anode, but there is no long-range force pushing it there, just the average effect of random collisions.

[2] **Problem 28** (Purcell 4.10). The basic ingredient in older voltmeters and ammeters is the galvanometer, a device to measure very small currents. (It works via magnetic effects, but the exact mechanism isn't important here.) Inherent in any galvanometer is some resistance R_g, so a physical galvanometer can be represented by the system shown below.



Consider a circuit such as the one shown, with all quantities unknown. We want to measure the current flowing across point A and the voltage difference between points B and C. Given a galvanometer with known R_g, and also a supply of known resistors (ranging from much smaller to much larger than R_g), how can you accomplish these two tasks? Explain how to construct your two devices (called an ammeter and voltmeter), and also how you should insert them in the given circuit. You will need to make sure that you (a) affect the given circuit as little as possible, and (b) don't destroy your galvanometer by passing more current through it than it can handle.

Solution. We explain how to create an ammeter and a voltmeter. As usual, attach the ammeter in series, and the voltmeter in parallel.

To create an ammeter, connect a resistor $R \ll R_g$ in parallel with the galvanometer (including R_g obviously). Suppose a current I passes through the device. If $N = R_g/R$, then its not hard to see that the current through galvanometer is about I/N, and the system acts very much like a wire. Therefore, to get I, multiply the reading by N.

To create a voltmeter, attach a resistor $R \gg R_g$ in series to the galvanometer (again, attach this contraption in parallel to the circuit). If the voltage drop we want to measure is V, then the current will be about V/R, so we can use the current to find V.

Remark

Occasionally, you might see Olympiad problems where a voltmeter is connected in series. The most common voltmeters are handheld digital multimeters, where the voltmeter setting presents a resistance of about $10\,\mathrm{M}\Omega$. Thus, for such problems, you should just treat the voltmeter like a high-resistance resistor.

Is this realistic? Well, it certainly happens every day, in almost every introductory physics lab in the world. But no professional would ever do this on purpose, because voltmeters aren't designed to be used this way. There is no guarantee that the resistance of the voltmeter is a constant. Instead, for most digital multimeters, there is a complex circuit inside that adjusts the internal resistance depending on the input and the configuration settings. You probably won't break the voltmeter when you put it in series, but you won't get reliable results either.

[2] **Problem 29.** USAPhO Quarterfinal 2009, problems 3 and 4.

[3] **Problem 30.** INPhO 2021, problem 1. A nice problem on practical circuit measurements. Note that the question statement is a bit vague. You are supposed to keep track of quantities of order R_A/R and R/R_V, but you are allowed to neglect quantities as small as R_A/R_V.

Solution. See the official solutions here.

Electromagnetism III: Magnetostatics

Chapters 4 and 6 of Purcell cover DC circuits and magnetostatics, as does chapter 5 of Griffiths. For advanced circuits techniques, see chapter 9 of Wang and Ricardo, volume 2. Chapter 5 of Purcell famously derives magnetic forces from Coulomb's law and relativity. It's beautiful, but not required to understand chapter 6; we will cover relativistic electromagnetism in depth in **R3**. For further discussion, see chapters II-12 through II-15 of the Feynman lectures. There is a total of **87** points.

1 Static DC Circuits

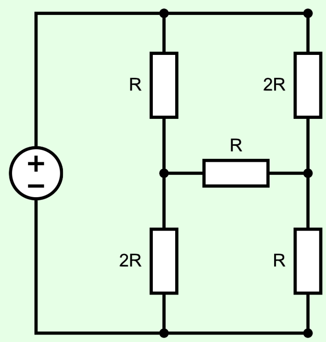
We continue with DC circuits, in more complex setups than in **E2**.

Idea 1

When analyzing circuits, it is sometimes useful to parametrize the currents in the circuits in terms of the current in each independent loop. This is typically more efficient, because it enforces Kirchoff's junction rule automatically, leading to fewer equations.

Example 1: Imbalanced Wheatstone Bridge

Find the current through the following circuit, if the battery has voltage V.



Solution

This circuit can't be simplified using series and parallel combinations, so instead we use Kirchoff's rules directly. From the diagram, we see the circuit has three loops. Let I_1 be the clockwise current on the left loop, I_2 be the clockwise current through the top-right loop, and I_3 be the clockwise current through the bottom-right loop. For instance, this means that the current flowing downward through the top-left resistor is $I_1 - I_2$.

The three Kirchoff's loop rule equations are

$$3I_1R - I_2R - 2I_3R = V,$$
$$4I_2R - I_1R - I_3R = 0,$$
$$4I_3R - 2I_1R - I_2R = 0.$$

Adding the last two equations shows that

$$I_1 = I_2 + I_3$$

and plugging this back in shows that $3I_2 = 2I_3$, so we have

$$I_2 = \frac{2}{5}I_1, \quad I_3 = \frac{3}{5}I_1.$$

Since the answer to the question is just I_1, we can now plug this back into the first equation,

$$\frac{V}{R} = 3I_1 - I_2 - 2I_3 = \left(3 - \frac{2}{5} - \frac{6}{5}\right) I_1 = \frac{7}{5}I_1.$$

This gives the answer, $5V/7R$.

Incidentally, the Wheatstone bridge is a famous circuit with the same topology. We note that the current through the middle resistor is zero when the ratios between the top and bottom resistances match on both sides of it. Hence if three of these outer resistances are known, we can adjust one of them until the current through the middle resistor vanishes, thereby measuring the fourth resistor.

Idea 2

Since Kirchoff's loop equations are linear, currents and voltages in a DC circuit with multiple batteries can be found by superposing the currents and voltages due to each battery alone.

Idea 3: Thevenin's Theorem and Norton's Theorem

Consider any system of batteries and resistors, with two external terminals A and B. Suppose that when a current I is sent into A and out of B, then a voltage difference $V = V_A - V_B$ appears. From an external standpoint, the function $V(I)$ is all we can measure.

Now, by the linearity of Kirchoff's rules, $V(I)$ is a linear function, so we can write

$$V(I) = V_{\text{eq}} + IR_{\text{eq}}.$$

In other words, $V(I)$ is exactly the same as if the entire system were a resistor R_{eq} in series with a battery with emf V_{eq} (with the positive end pointing towards A). This generalizes the idea of replacing a system of resistors with an equivalent resistance, and is known as Thevenin's theorem.

We can also flip this around. Note that $I(V)$ must also be a linear function, and we can write

$$I(V) = I_{\text{eq}} + \frac{V}{R_{\text{eq}}}.$$

This is precisely the $I(V)$ of an ideal current source I_{eq} (sending current towards B) in parallel with a resistor R_{eq}. (An ideal current source makes a fixed current flow through it, just like a battery creates a fixed voltage across it.) This is known as Norton's theorem.

Since these functions are inverses of each other, you can see that the R_{eq}'s in both equations above are the same (both are equal to the ordinary equivalent resistance), and $V_{\text{eq}} = -I_{\text{eq}}R_{\text{eq}}$.

Example 2

Consider some batteries connected in parallel, with emfs $\mathcal{E}_i$ and internal resistances R_i. What is the Thevenin equivalent of this circuit?

Solution

The equivalent resistance is simply

$$R_{\text{eq}} = \left(\sum_i \frac{1}{R_i} \right)^{-1}.$$

To infer V_{eq}, we just need one more $V(I)$ value. The most convenient is to set $V = 0$, shorting all of the batteries. Each battery alone would produce a current of $\mathcal{E}_i / R_i$, so

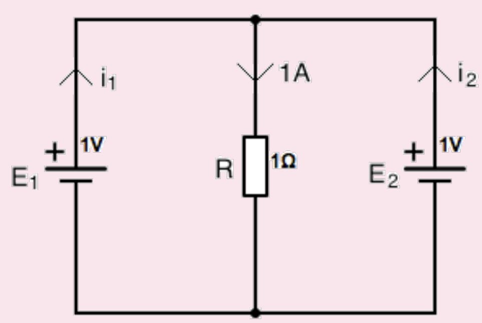
$$0 = V_{\text{eq}} + \left(\sum_i \frac{\mathcal{E}_i}{R_i} \right) R_{\text{eq}}.$$

Thus, we have

$$V_{\text{eq}} = \left(\sum_i \frac{\mathcal{E}_i}{R_i} \right) \left(\sum_j \frac{1}{R_j} \right)^{-1}.$$

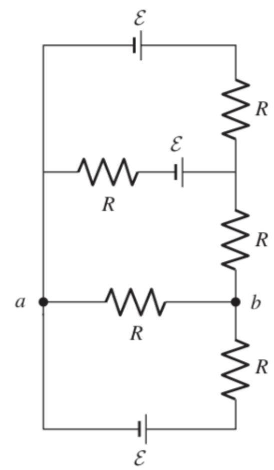
Remark

With ideal batteries, it's easy to set up circuits that don't make any sense.



For example, in the above circuit, Kirchoff's rules don't determine the currents; they only say that $i_1 + i_2 = 1\,\text{A}$. If the emfs of the batteries were different, the situation would be even worse: the equations would be contradictory, with no solution at all! In real life, this is avoided because all batteries have some internal resistance. Adding such a resistance to each battery, no matter how small, resolves the problem and gives a unique solution.

[2] **Problem 1** (Purcell 4.12). Consider the circuit below.



(a) Find the potential difference between points a and b.

(b) Find the equivalent Thevenin resistance and emf between points a and b.

[2] **Problem 2** (Wang)**.** A circuit containing batteries and resistors has two terminals. When an ideal ammeter is connected between them, the reading is I_1. When a resistor R is connected between them, the current through the resistor is I_2, in the same direction. What would be the reading V of an ideal voltmeter connected between them?

[3] **Problem 3.** USAPhO 2015, problem A2.

Now we give a few problems on current flow through continuous objects. Fundamentally, all one needs for these problems is the definition $\mathbf{J} = \sigma \mathbf{E}$, and superposition.

Example 3

Consider two long, concentric cylindrical shells of radii $a < b$ and length L. The volume between the two shells is filled with material with conductivity $\sigma(r) = k/r$. What is the resistance between the shells, and the charge density?

Solution

To find the resistance, we compute the current I when a voltage V is applied between the shells. By symmetry, in the steady state the current density must be

$$\mathbf{J}(\mathbf{r}) = \frac{I}{2\pi r L}\,\hat{\mathbf{r}}.$$

On the other hand, we also know that

$$V = \int \mathbf{E} \cdot d\mathbf{r} = \int_a^b \frac{I}{2\pi r L \sigma}\, dr = \frac{I(b-a)}{2\pi k L}$$

from which we conclude

$$R = \frac{b-a}{2\pi k L}.$$

Note that the radial electric field between the shells is constant, so

$$\mathbf{E}(\mathbf{r}) = \frac{V}{b-a}\,\hat{\mathbf{r}}.$$

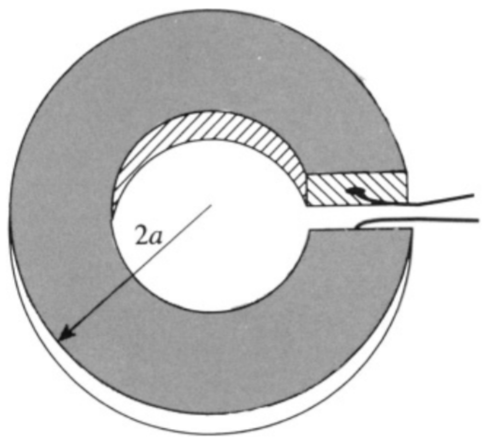
This means that in the steady state, there must be a nonzero charge density between the shells. (If there weren't, then we would have $E(r) \propto 1/r$, rather than a constant.)

To find the charge density explicitly, it's easiest to use Gauss's law in differential form in cylindrical coordinates. We use the form of the divergence derived in **E1**,

$$\nabla \cdot \mathbf{E} = \frac{1}{r}\frac{\partial(rE_r)}{\partial r} + (\text{other terms}) = \frac{1}{r}\frac{V}{b-a} = \frac{\rho}{\epsilon_0}$$

thus showing that the charge density is proportional to $1/r$. Of course, we could also get this result by applying Gauss's law in integral form, to concentric spheres.

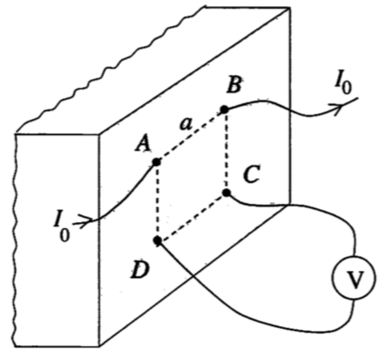
[2] **Problem 4** (Cahn). A washer is made of a material of resistivity ρ. It has a square cross section of length a on a side, and its outer radius is $2a$. A small slit is made on one side and wires are connected to the faces exposed.



Since the washer has an irregular shape, the current distribution inside it is complicated: it spreads out from the first wire, goes around the washer, and converges into the second wire. However, the situation is simpler if we glue a perfectly conducting square plate, of side length a, to each exposed face. Find the resistance in this case.

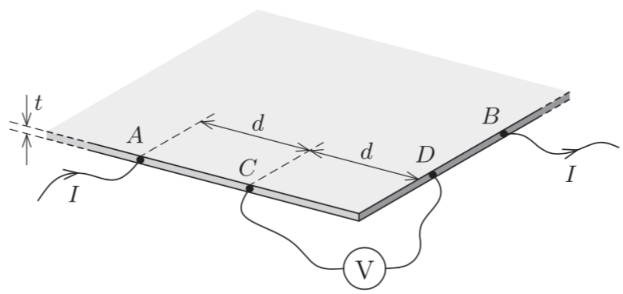
[3] **Problem 5** (BAUPC 1995). An electrical signal can be transferred between two metallic objects buried in the ground, where the current passes through the Earth itself. Assume that these objects are spheres of radius r, separated by a horizontal distance $L \gg r$, and suppose both objects are buried a depth much greater than L in the ground. If the Earth has uniform resistivity ρ, find the approximate resistance between the terminals. (Hint: consider the superposition principle.)

[3] **Problem 6** (PPP 162). A plane divides space into two halves. One half is filled with a homogeneous conducting medium, and physicists work in the other. They mark the outline of a square of side a on the plane and let a current I_0 in and out at two of its neighboring corners. Meanwhile, the measure the potential difference V between the two other corners.



Find the resistivity ρ of the medium.

[3] **Problem 7** (MPPP 174)**.** We aim to measure the resistivity of the material of a large, thin, homogeneous square metal plate, of which only one corner is accessible. To do this, we chose points A, B, C and D on the side edges of the plate that form the corner.



Points A and B are both $2d$ from the corner, whereas C and D are each a distance d from it. The length of the plate's sides is much greater than d, which, in turn, is much greater than the thickness t of the plate. If a current I enters the plate at point A, and leaves it at B, then the reading on a voltmeter connected between C and D is V. Find the resistivity ρ of the plate material.

Remark

Setups like those in the previous two problems are commonly used to measure resistivities, but why do they use a complicated "four terminal" setup? Wouldn't it have been easier to just attach two terminals, send a current I through them, and measure the voltage drop V? The problem with this is that it also picks up the resistance R of the contacts between the terminals and the material, along with the resistances of the wires. By having a pair of terminals measure voltage alone, drawing negligible current, we avoid this problem.

[4] **Problem 8.** [A] This problem is just for fun; the techniques used here are too advanced to appear on Olympiads. We will prove Rayleigh's monotonicity law, which states that increasing the resistance of any part of a resistor network increases the equivalent resistance between any two points. This may seem obvious, but it's actually tricky to prove. The following is the slickest way.

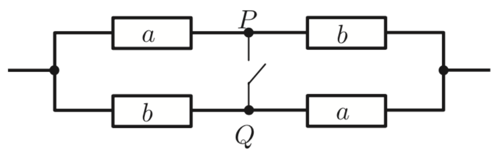
(a) Consider a graph of resistors, where a battery is attached across two of the vertices, fixing their voltages. Write an expression for the total power dissipated, assuming the voltages at each vertex are V_i and the resistances are R_{ij}.

(b) The voltages V_i at all the other vertices are determined by Kirchoff's rules. But suppose you didn't know that, or didn't want to set up those equations. Remarkably, it turns out that you can derive the exact same results by simply treating the voltages V_i as free to vary, and setting them to minimize the total power dissipated! Show this result. (This is an example of a variational principle, like the principle of least action in mechanics.)

(c) For any network of resistors, show that $P = V^2/R$ when V is the battery voltage applied across two vertices, R is the equivalent resistance between them, and P is the total power dissipated in the resistors. (This is intuitive, but it's worth showing in detail to assist with the next part.)

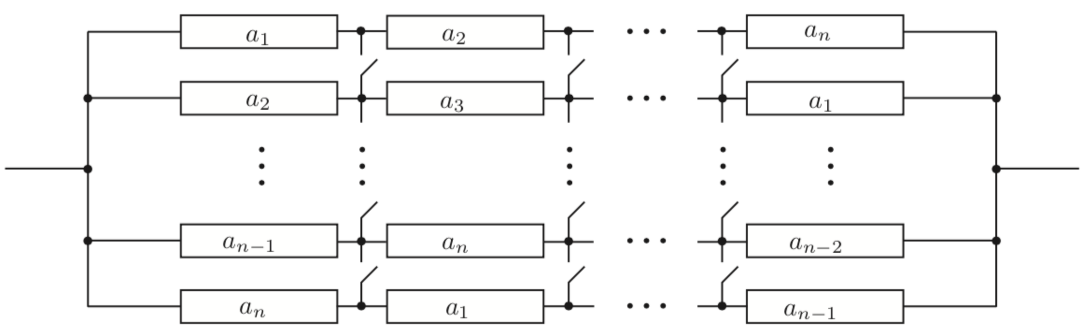
(d) By combining all of these results, prove Rayleigh's monotonicity law.

(e) We can use Rayleigh's monotonicity law to prove some mathematical results. Consider the resistor network shown below, where the variables label the resistances.



By considering the resistances before and after closing the switch PQ, show that the arithmetic mean of two numbers is at least the geometric mean.

(f) Consider the resistor network shown below.



By closing all the switches, show that the arithmetic mean of n numbers is at least the harmonic mean.

Remark

You might think that Rayleigh's monotonicity law is too obvious to require a proof; if you decrease a resistance, how could the net resistance possibly go up? In fact, this kind of non-monotonicity occurs very often! For example, Braess's paradox is that fact that adding more roads can slow down traffic, even when the total number of cars stays the same. A U.S. Physics Team coach has argued that allowing more team strategies can make a basketball team score less. For more on this subject, see the paper *Paradoxical behaviour of mechanical and electrical networks* or this video.

Remark

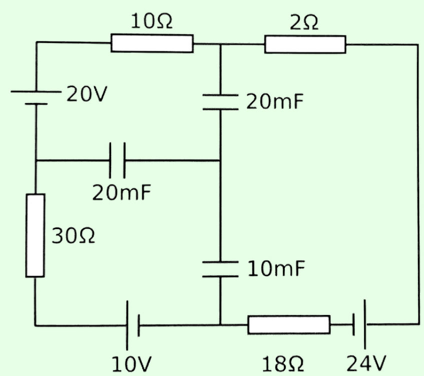
Circuit questions can get *absurdly* hard, but at some point they start being more about mathematical tricks than physics. As a result, I haven't included any such problems here; they tend not to appear on the USAPhO or IPhO, or in college physics, or in real life, or really anywhere besides a few competitions. On the other hand, you might find such questions fun! For some examples, see the Physics Cup problems 2013.6, 2017.2, 2018.1, and 2019.4.

2 RC Circuits

Next we'll briefly cover RC circuits, our first exposure to a situation genuinely changing in time.

Example 4: CPhO

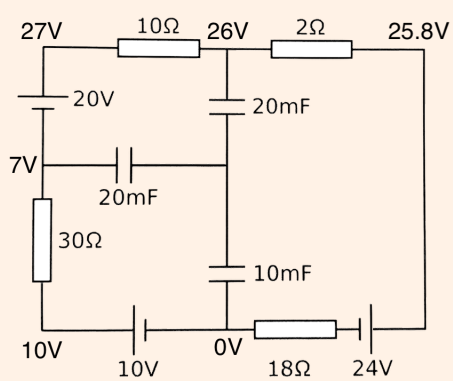
The capacitors in the circuit shown below were initially neutral. Then, the circuit is allowed to reach the steady state.



After a long time, what is the charge stored on the 10 mF capacitor?

Solution

After a long time, no current flows through the capacitors, so there is effectively a single loop in the circuit. It has a total resistance 60 Ω and a total emf 6 V, so the current is $I = 0.1\,\text{A}$. Using this, we can straightforwardly label the voltages everywhere on the outer loop.



To finish the problem, we need to know the voltage V_0 of the central node, so we need one more equation. That equation is charge conservation: the fact that the central part of the circuit, containing the inner plates of the three capacitors, begins and remains uncharged. Suppressing units, this means

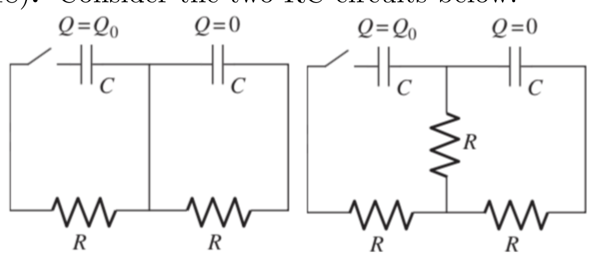
$$20(26 - V_0) + 20(7 - V_0) + 10(0 - V_0) = 0, \quad V_0 = \frac{66}{5}\,\text{V}$$

from which we read off the answer,

$$Q = CV = 0.132\,\text{C}.$$

[3] **Problem 9.** USAPhO 1997, problem A3.

[3] **Problem 10** (Purcell 4.18)**.** Consider the two RC circuits below.

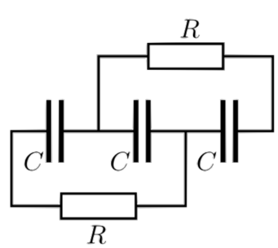


(a) The circuit shown below contains two identical capacitors and two identical resistors, with initial charges as shown above at left. If the switch is closed at $t = 0$, find the charges on the capacitors as functions of time.

(b) Now consider the same setup with an extra resistor, as shown above at right. Find the maximum charge that the right capacitor achieves. (Hint: the methods of **M4** can be useful.)

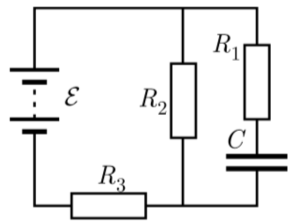
[3] **Problem 11.** USAPhO 2004, problem A1.

[3] **Problem 12** (Kalda)**.** Three identical capacitors are placed in series and charged with a battery of emf $\mathcal{E}$. Once they are fully charged, the battery is removed, and simultaneously two resistors are connected as shown.



Find the heat dissipated on each of the resistors after a long time.

[3] **Problem 13** (Kalda)**.** Find the time constant of the RC circuit shown below.



[3] **Problem 14** (MPPP 175/176)**.** A metal sphere of radius R has charge Q and hangs on an insulating cord. It slowly loses charge because air has a conductivity σ. In all cases, neglect any magnetic or radiation effects.

(a) Find the time for the charge to halve.

(b) You should have found that the time is independent of the radius R of the sphere, which follows directly from dimensional analysis. Can you show that, in fact, it is completely independent of the shape? (This doesn't just follow from dimensional analysis, because the shape might be described by dimensionless numbers, such as the eccentricity of an ellipsoid.)

(c) Air has a conductivity of $\sigma \sim 10^{-13}\,\Omega^{-1}\mathrm{m}^{-1}$, while water has a conductivity of $\sigma \sim 10^{-2}\,\Omega^{-1}\mathrm{m}^{-1}$. About how long does the charge on an object last, if it is in air or water?

This problem generalizes USAPhO 2010, problem A2, which you can compare.

[5] **Problem 15.** IPhO 1993, problem 1. A really neat question with real-world relevance.

[5] **Problem 16.** IPhO 2007, problem "orange". A combination of mechanics and RC circuits.

3 Computing Magnetic Fields

Idea 4

The Biot–Savart law is

$$\mathbf{B} = \frac{\mu_0 I}{4\pi} \oint \frac{d\mathbf{s} \times \mathbf{r}}{r^3}.$$

As a consequence, we have Ampere's law,

$$\oint \mathbf{B} \cdot d\mathbf{s} = \mu_0 I, \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{J}$$

as well as Gauss's law for magnetism,

$$\oint \mathbf{B} \cdot d\mathbf{S} = 0, \quad \nabla \cdot \mathbf{B} = 0.$$

Idea 5

The force on a stationary wire carrying current I in a magnetic field $\mathbf{B}$ is

$$\mathbf{F} = I \int d\mathbf{s} \times \mathbf{B}.$$

The energy of a magnetic field is

$$U = \frac{1}{2\mu_0} \int B^2 \, dV.$$

The magnetic dipole moment of a planar current loop of area A and current I is $m = IA$, with $\mathbf{m}$ directed perpendicular to the loop by the right-hand rule.

Idea 6: Magnetic Dipoles

Far from a magnetic dipole with magnetic moment m, its magnetic field is just the same as the electric field of an electric dipole,

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0 m}{4\pi r^3}(2\cos\theta\,\hat{\mathbf{r}} + \sin\theta\,\hat{\boldsymbol{\theta}}) = \frac{\mu_0}{4\pi r^3}(3(\mathbf{m}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}} - \mathbf{m}).$$

As with the electric dipole field, you don't need to memorize this, but you should remember that it's proportional to the dipole moment, falls off as $1/r^3$, and be able to sketch it.

You should have already seen basic examples of using the Biot–Savart law in Halliday and Resnick, such as the field of a circular ring of current on its axis. We'll start with some problems that are similarly straightforward, but more technically complex.

[3] **Problem 17.** This is a key question which will help you understand idea 7. A spherical shell with radius R and uniform surface charge density σ spins with angular frequency ω about a diameter.

(a) Find the magnetic field at the sphere's center.

(b) Find the magnetic dipole moment of the sphere.

(c) It can be shown that (1) the magnetic field inside the sphere is uniform, and (2) the magnetic field outside the sphere is exactly that of a magnetic dipole. (It requires doing some obnoxious integrals, as can be seen in section 5.4 of Griffiths.) Using this information, make a qualitatively accurate sketch of the field.

(d) There is a closely related question in electrostatics: suppose we had two spherical shells of the same radius R, and surface charge densities $\pm\sigma$, and the shells were displaced by a small distance $d \ll R$. Qualitatively, what would the electric field of this setup look like? How would it differ from your answer to part (c)? You can neglect the region between the shells.

[2] **Problem 18** (Purcell 6.12)**.** A ring with radius R carries a current I. Show that the magnetic field due to the ring, at a point in the plane of the ring, a distance r from the center, is given by

$$B = \frac{\mu_0 I}{2\pi} \int_0^\pi \frac{(R - r\cos\theta)R\,d\theta}{(r^2 + R^2 - 2rR\cos\theta)^{3/2}}.$$

In the $r \gg R$ limit, show that

$$B \approx \frac{\mu_0}{4\pi}\frac{m}{r^3}$$

where $m = IA$ is the magnetic dipole moment of the ring, as expected from idea 6.

[3] **Problem 19** (Purcell 6.14)**.** Consider a square loop with current I and side length a centered at the origin, with sides parallel to the x and y axes. Show that the magnetic field at $r\hat{\mathbf{x}}$ is $B \approx (\mu_0/4\pi)(m/r^3)$ for $r \gg a$, as expected from idea 6. Be careful with factors of 2!

[3] **Problem 20.** USAPhO 2012, problem A3.

Idea 7: Magnetic Monopoles

Far away from the center of the dipole, the magnetic field of a magnetic dipole has the same form as the electric field of an electric dipole. Therefore, we can often replace a magnetic dipole m with a fictitious pair of "magnetic charges" $\pm q_m$ separated by d, where $q_m d = m$. This is called a "Gilbert dipole", in contrast to a true "Amperian dipole".

This was the default way to think about magnets in the 1800s, but was largely removed from American textbooks in the 1950s because it's misleading in general: magnetic charges don't actually exist in magnets, and applying this analogy will give the wrong fields inside the dipole, as you saw in problem 17 and will see another way in problem 21. However, if we only care about the field outside the magnet, the analogy works, and it's often the fastest way to solve problems. We'll return to this idea in greater depth in **E8**.

[3] **Problem 21.** USAPhO 2015, problem B2. A key problem which illustrates idea 7.

We now give a few arguments for computing fields using symmetry.

Example 5: PPP 31

An electrically charged conducting sphere "pulses" radially, i.e. its radius changes periodically with a fixed amplitude. What is the net pattern of radiation from the sphere?

Solution

There is no radiation. By spherical symmetry, the magnetic field can only point radially. But then this would produce a magnetic flux through a Gaussian sphere centered around the pulsing sphere, which would violate Gauss's law for magnetism. So there is no magnetic field at all, and since radiation always needs both electric and magnetic fields (as you'll see in **E7**), there is no radiation at all. In fact, outside the sphere the electric field is always exactly equal to $Q/4\pi\epsilon_0 r^2$, in accordance with Coulomb's law.

Example 6

Find the magnetic field of an infinite cylindrical solenoid, of radius R and n turns per unit length, carrying current I.

Solution

Orient the solenoid along the vertical direction and use cylindrical coordinates. By symmetry, the field must be independent of z. Now consider the radial component of the magnetic field B_r. Turning the solenoid upside-down is equivalent to reversing the current. But the former does not flip B_r while the latter does, so we must have $B_r = 0$.

Now, by rotational symmetry, the tangential component B_ϕ must be uniform. But then Ampere's law on any circular loop gives $B_\phi(2\pi r) = 0$, so we must have $B_\phi = 0$ as well.

The only thing left to consider is B_z. By applying Ampere's law to small vertical rectangles, we see that B_z is constant unless that rectangle crosses the surface of the solenoid. Furthermore, B_z must be zero far from the solenoid, so it must be zero everywhere outside the solenoid. Now, for a rectangle of height h that does cross the surface, Ampere's law gives

$$\oint \mathbf{B} \cdot d\mathbf{s} = B_z^{\text{in}} h = \mu_0 I_{\text{enc}} = \mu_0 n I h$$

which tells us that $B_z^{\text{in}} = \mu_0 n I$.

Example 7

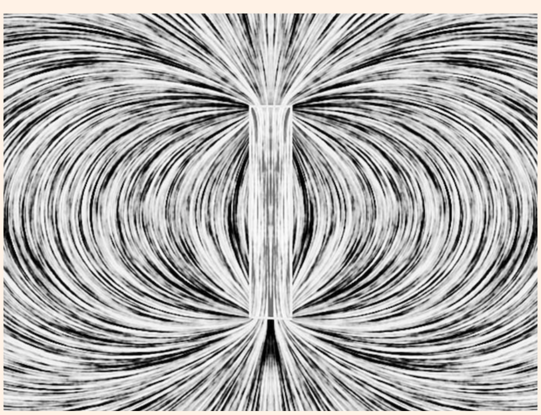
Now suppose the solenoid has finite length $L \gg R$. What do the fringe fields look like?

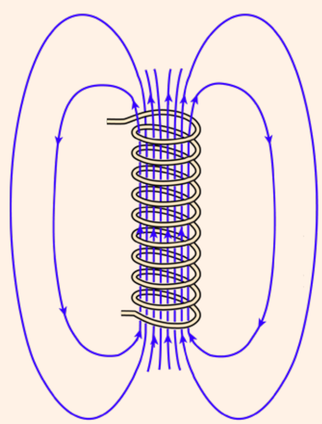
Solution

In principle we could solve for the exact fringe field by applying the Biot–Savart law to the solenoid wire, but that would be rather complicated. Instead, let's approximate the solenoid as a stack of $N = nL$ evenly spaced circular wire loops. Each one of these loops is a magnetic dipole $\mu = \pi R^2 I$, so the field of each loop well outside of it is just a dipole field.

Summing up all of these dipole fields is still complicated, so let's use idea 7. We can replace each wire loop with a pair of magnetic charges $\pm q_m$ separated by d, with the same magnetic dipole moment $\mu = q_m d$. If we choose $d = 1/n$, then the charges of adjacent dipoles cancel, leaving only charges $q_m = \pm n\mu = \pm \pi R^2 n I$ on the ends.

Thus, the fringe field of a solenoid, at distances much greater than R, looks like the electric field of two point charges! This is confirmed by a numeric calculation shown at left below.





This may come as a surprise to you if you've read basic, algebra-based introductory physics textbooks. Many of them contain hand-drawn diagrams like the one shown at right above, where all the magnetic flux comes neatly out the ends of the solenoids, in straight lines. In reality, the field sprays out almost spherically symmetrically from the end, with only half the flux actually going out through the end face, while the rest exits downward through the sides. (You will show this more directly with a slick argument in problem 23.)

We can also be more quantitative. Suppose the solenoid is vertical and centered at $z = 0$. Then the field at a radius r from the solenoid axis, at $z = 0$, is

$$\mathbf{B}(r) = \mu_0 n I \hat{\mathbf{z}} \times \begin{cases} 1 & r < R \\ -2R^2/L^2 & R \ll r \ll L \\ -R^2 L/4r^3 & L \ll r \end{cases}$$

where the first line is the usual solenoid field, the second line is from applying Coulomb's law to our dipole analogy (which is only valid when $R \ll r$), and the third is from the dipole field of the two charges (only valid when $L \ll r$). As expected, in the limit $L \gg R$, the fringe field outside the solenoid is negligible. Another way of phrasing the result is that most of the upward flux through the solenoid returns through a downward field which mainly extends out to $r \sim L$. You can see all of these features in the accurate drawing above.

We can draw two lessons from this example. First, misleading diagrams are a common problem in introductory textbooks. A general rule is that the more basic a textbook is, the more pictures it'll have, but the less useful they'll be. Second, the analogy between Ampere and Gilbert dipoles is quite useful, and shows up frequently in tricky Olympiad problems.

Remark: Real Solenoids

Real solenoids are even more complicated. First, we didn't account for the discreteness of the wires. We just treated them as forming a uniform current per length $K = nI$, which is how we wrote $I_{\text{enc}} = nIh$. This is valid when you don't care about looking too closely, i.e. if your distance to any wire is much larger than the wire spacing $1/n$.

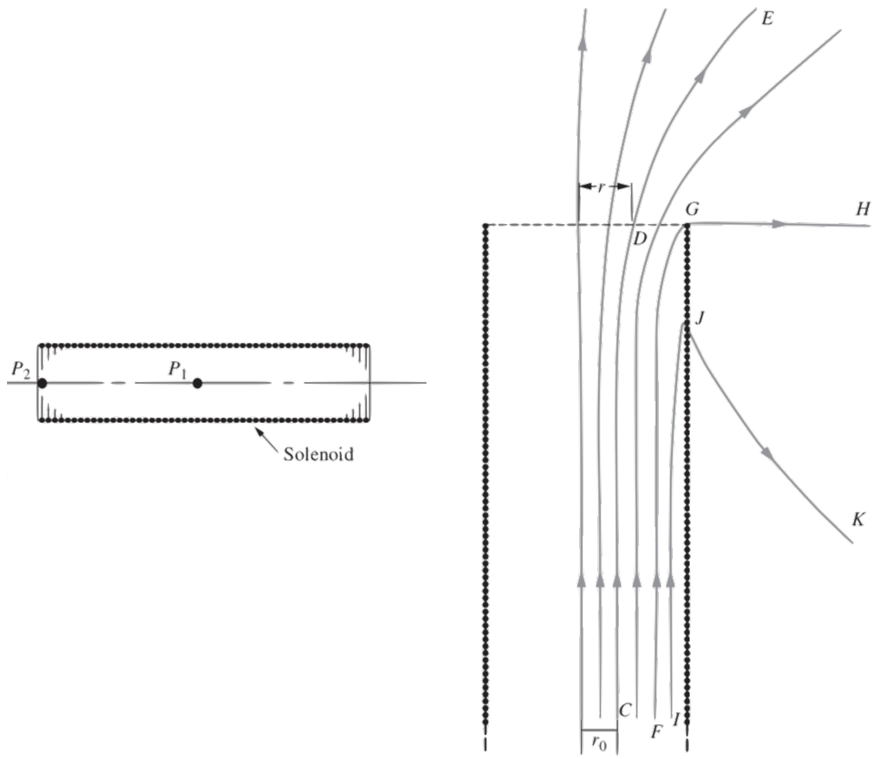
Second, the fact that solenoids are made by winding real wires means there is another contribution to the current, even in the limit $n \to \infty$. The wires are wound with a small slope, since a net current I still has to move along the solenoid. Another way of saying this is that the current per length along the solenoid surface is $\mathbf{K} = nI\hat{\boldsymbol{\theta}} + (I/2\pi R)\hat{\mathbf{z}}$. This causes a tangential magnetic field $B_\phi = \mu_0 I/2\pi r$ outside the solenoid. Thus, in practice many solenoids are "counterwound": half the wires are wound evenly spaced going up the axis, and the other half are wound evenly spaced going back down the axis, which closes the loop and cancels this unwanted field.

[2] **Problem 22.** A toroidal solenoid is created by wrapping N turns of wire around a torus with a rectangular cross section. The height of the torus is h, and the inner and outer radii are a and b.

(a) In the ideal case, the magnetic field vanishes everywhere outside the toroid, and is purely tangential inside the toroid. Find the magnetic field inside the toroid.

(b) There is another small contribution to the magnetic field due to the winding effect mentioned above. Roughly what does the resulting extra magnetic field look like? If you didn't want this additional field, how would you design the solenoid to get rid of it?

[3] **Problem 23** (Purcell 6.63). A number of simple facts about the fields of solenoids can be found by using superposition. The idea is that two solenoids of the same diameter, and length L, if joined end to end, make a solenoid of length $2L$. Two semi-infinite solenoids butted together make an infinite solenoid, and so on.



Prove the following facts.

(a) In the finite-length solenoid shown at left above, the magnetic field on the axis at the point P_2 at one end is approximately half the field at the point P_1 in the center. (Is it slightly more than half, or slightly less than half?)

(b) In the semi-infinite solenoid shown at right above, the field line FGH, which passes through the very end of the winding, is a straight line from G out to infinity.

(c) The flux through the end face of the semi-infinite solenoid is half the flux through the coil at a large distance back in the interior.

(d) Any field line that is a distance r_0 from the axis far back in the interior of the coil exits from the end of the coil at a radius $r_1 = \sqrt{2}r_0$, assuming $\sqrt{2}r_0$ is less than the solenoid radius.

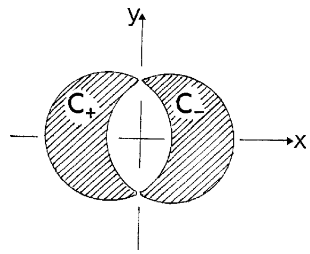
[3] **Problem 24** (MPPP 160). Two infinite parallel wires, a distance d apart, carry electric currents along the z-axis with equal magnitudes but opposite directions. We can find the shape of the magnetic field lines with a neat trick, which only works for "two-dimensional" setups like this one, where the fields lie in the xy plane and don't depend on z.

(a) Argue that if we rotated $\mathbf{B}$ by $90°$ in the xy plane at each point, it would produce a valid electrostatic field $\mathbf{E}$. (Hint: consider rotating the $\mathbf{B}$ field of each wire individually.)

(b) Argue that the field lines of $\mathbf{B}$ are the same as the equipotentials of this artificial $\mathbf{E}$, and use this to find the field lines.

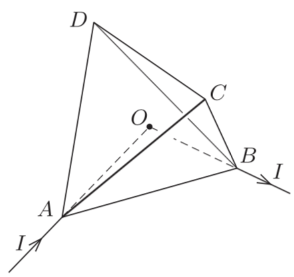
This trick is also useful for fluids in two dimensions, where it swaps vortices with sources and sinks.

[2] **Problem 25** (IPhO 1996)**.** Two straight, long conductors C_+ and C_-, insulated from each other, carry current I in the positive and the negative $\hat{\mathbf{z}}$ direction respectively. The cross sections of the conductors are circles of diameter D in the xy plane, with a distance $D/2$ between the centers.



The current in each conductor is uniformly distributed. Find the magnetic field in the space between the conductors.

[3] **Problem 26** (MPPP 157)**.** A regular tetrahedron is made of a wire with constant resistance per unit length. A long, straight wire sends current I into one vertex, and another long, straight wire removes it from another vertex, as shown.



Find the magnetic field at the center of the tetrahedron.

[5] **Problem 27.** APhO 2013, problem 1. A neat question on a cylindrical RC circuit that uses many of the techniques we've covered so far.

[5] **Problem 28.** EuPhO 2023, problem 3. Another neat question, in a setup where an eddy current can be computed exactly. For this problem, you won't need to know anything about magnetism besides the fact that the magnetic force per unit charge is $\mathbf{v} \times \mathbf{B}$.

Electromagnetism III: Magnetostatics

Chapters 4 and 6 of Purcell cover DC circuits and magnetostatics, as does chapter 5 of Griffiths. For advanced circuits techniques, see chapter 9 of Wang and Ricardo, volume 2. Chapter 5 of Purcell famously derives magnetic forces from Coulomb's law and relativity. It's beautiful, but not required to understand chapter 6; we will cover relativistic electromagnetism in depth in **R3**. For further discussion, see chapters II-12 through II-15 of the Feynman lectures. There is a total of **87** points.

1 Static DC Circuits

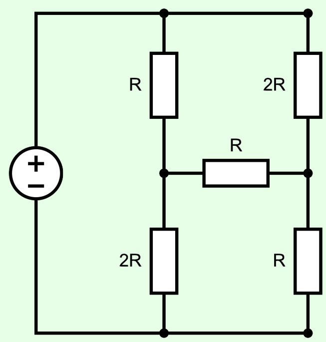
We continue with DC circuits, in more complex setups than in **E2**.

Idea 1

When analyzing circuits, it is sometimes useful to parametrize the currents in the circuits in terms of the current in each independent loop. This is typically more efficient, because it enforces Kirchoff's junction rule automatically, leading to fewer equations.

Example 1: Imbalanced Wheatstone Bridge

Find the current through the following circuit, if the battery has voltage V.



Solution

This circuit can't be simplified using series and parallel combinations, so instead we use Kirchoff's rules directly. From the diagram, we see the circuit has three loops. Let I_1 be the clockwise current on the left loop, I_2 be the clockwise current through the top-right loop, and I_3 be the clockwise current through the bottom-right loop. For instance, this means that the current flowing downward through the top-left resistor is $I_1 - I_2$.

The three Kirchoff's loop rule equations are

$$3I_1R - I_2R - 2I_3R = V,$$
$$4I_2R - I_1R - I_3R = 0,$$
$$4I_3R - 2I_1R - I_2R = 0.$$

Adding the last two equations shows that

$$I_1 = I_2 + I_3$$

and plugging this back in shows that $3I_2 = 2I_3$, so we have

$$I_2 = \frac{2}{5} I_1, \quad I_3 = \frac{3}{5} I_1.$$

Since the answer to the question is just I_1, we can now plug this back into the first equation,

$$\frac{V}{R} = 3I_1 - I_2 - 2I_3 = \left(3 - \frac{2}{5} - \frac{6}{5}\right) I_1 = \frac{7}{5} I_1.$$

This gives the answer, $5V/7R$.

Incidentally, the Wheatstone bridge is a famous circuit with the same topology. We note that the current through the middle resistor is zero when the ratios between the top and bottom resistances match on both sides of it. Hence if three of these outer resistances are known, we can adjust one of them until the current through the middle resistor vanishes, thereby measuring the fourth resistor.

Idea 2

Since Kirchoff's loop equations are linear, currents and voltages in a DC circuit with multiple batteries can be found by superposing the currents and voltages due to each battery alone.

Idea 3: Thevenin's Theorem and Norton's Theorem

Consider any system of batteries and resistors, with two external terminals A and B. Suppose that when a current I is sent into A and out of B, then a voltage difference $V = V_A - V_B$ appears. From an external standpoint, the function $V(I)$ is all we can measure.

Now, by the linearity of Kirchoff's rules, $V(I)$ is a linear function, so we can write

$$V(I) = V_{\text{eq}} + IR_{\text{eq}}.$$

In other words, $V(I)$ is exactly the same as if the entire system were a resistor R_{eq} in series with a battery with emf V_{eq} (with the positive end pointing towards A). This generalizes the idea of replacing a system of resistors with an equivalent resistance, and is known as Thevenin's theorem.

We can also flip this around. Note that $I(V)$ must also be a linear function, and we can write

$$I(V) = I_{\text{eq}} + \frac{V}{R_{\text{eq}}}.$$

This is precisely the $I(V)$ of an ideal current source I_{eq} (sending current towards B) in parallel with a resistor R_{eq}. (An ideal current source makes a fixed current flow through it, just like a battery creates a fixed voltage across it.) This is known as Norton's theorem.

Since these functions are inverses of each other, you can see that the R_{eq}'s in both equations above are the same (both are equal to the ordinary equivalent resistance), and $V_{\text{eq}} = -I_{\text{eq}}R_{\text{eq}}$.

Example 2

Consider some batteries connected in parallel, with emfs $\mathcal{E}_i$ and internal resistances R_i. What is the Thevenin equivalent of this circuit?

Solution

The equivalent resistance is simply

$$R_{\text{eq}} = \left(\sum_i \frac{1}{R_i}\right)^{-1}.$$

To infer V_{eq}, we just need one more $V(I)$ value. The most convenient is to set $V = 0$, shorting all of the batteries. Each battery alone would produce a current of $\mathcal{E}_i/R_i$, so

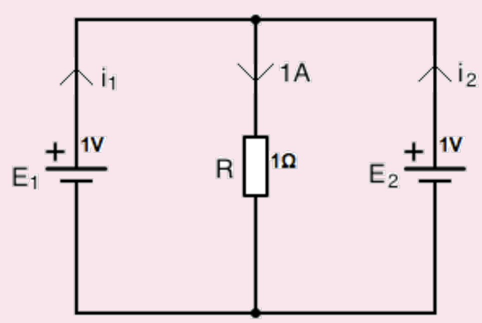
$$0 = V_{\text{eq}} + \left(\sum_i \frac{\mathcal{E}_i}{R_i}\right) R_{\text{eq}}.$$

Thus, we have

$$V_{\text{eq}} = \left(\sum_i \frac{\mathcal{E}_i}{R_i}\right)\left(\sum_j \frac{1}{R_j}\right)^{-1}.$$

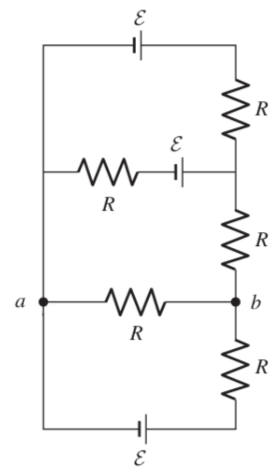
Remark

With ideal batteries, it's easy to set up circuits that don't make any sense.



For example, in the above circuit, Kirchoff's rules don't determine the currents; they only say that $i_1 + i_2 = 1\,\text{A}$. If the emfs of the batteries were different, the situation would be even worse: the equations would be contradictory, with no solution at all! In real life, this is avoided because all batteries have some internal resistance. Adding such a resistance to each battery, no matter how small, resolves the problem and gives a unique solution.

[2] **Problem 1** (Purcell 4.12). Consider the circuit below.



(a) Find the potential difference between points a and b.

(b) Find the equivalent Thevenin resistance and emf between points a and b.

Solution. (a) We'll use loop currents, with positive being clockwise. Let the loop currents be I_1, I_2, I_3, from top to bottom. We see that

$$\begin{aligned}\mathcal{E} - \mathcal{E} - 2RI_1 + RI_2 &= 0\\ \mathcal{E} - 3RI_2 + RI_3 + RI_1 &= 0\\ -\mathcal{E} - 2RI_3 + RI_2 &= 0,\end{aligned}$$

or

$$\begin{aligned}I_2 &= 2I_1\\ 3I_2 - I_3 - I_1 &= \mathcal{E}/R\\ I_2 - 2I_3 &= \mathcal{E}/R.\end{aligned}$$

This can be solved to give $I_1 = \mathcal{E}/8R$, $I_2 = \mathcal{E}/4R$, and $I_1 = -3\mathcal{E}/8R$. We see that $V_b - V_a = (I_2 - I_3)R = \boxed{5\mathcal{E}/8}$.

(b) We'll do this in two ways for variety. First, note that we already found the voltage between a and b in part (a), and this is precisely the Thevenin emf, $V_{\text{eff}} = 5\mathcal{E}/8$. The Thevenin resistance is simply the equivalent resistance between a and b. By a straightforward application of the series and parallel rules, this is $R_{\text{eff}} = \boxed{3R/8}$.

Second, suppose we short points a and b with a wire. Then by Thevenin's theorem, the current flowing through that wire should be $I = V_{\text{eff}}/R_{\text{eff}}$. We already know V_{eff} from part (a). To compute the current, we just use Kirchoff's loop rules again; these are now as follows.

$$\begin{aligned}\mathcal{E} - \mathcal{E} - 2RI_1 + RI_2 &= 0\\ \mathcal{E} - 2RI_2 + RI_1 &= 0\\ -\mathcal{E} - RI_3 &= 0\end{aligned}$$

Solving these equations gives $I_1 = \mathcal{E}/3R$, $I_2 = 2\mathcal{E}/3R$, and $I_3 = -\mathcal{E}/R$. The current through the wire is now $I_2 - I_3 = 5\mathcal{E}/3R$. Thus, $R_{\text{eff}} = (5\mathcal{E}/8)/(5\mathcal{E}/3R) = 3R/8$.

[2] **Problem 2** (Wang)**.** A circuit containing batteries and resistors has two terminals. When an ideal ammeter is connected between them, the reading is I_1. When a resistor R is connected between them, the current through the resistor is I_2, in the same direction. What would be the reading V of an ideal voltmeter connected between them?

Solution. We consider the Thevenin equivalent, i.e. the function $V(I)$. The first piece of information tells us that when $V = 0$, $I = I_1$. The second tells us that when $V = -I_2 R$, then $I = I_2$. Thus,

$$0 = V_{\text{eq}} + I_1 R_{\text{eq}}, \quad -I_2 R = V_{\text{eq}} + I_2 R_{\text{eq}}.$$

Solving this system of equations gives

$$V_{\text{eq}} = \frac{I_1 I_2 R}{I_2 - I_1}, \quad R_{\text{eq}} = \frac{I_2 R}{I_1 - I_2}.$$

When an ideal voltmeter is connected, we have $I = 0$, so

$$V = V_{\text{eq}} = \frac{I_1 I_2 R}{I_2 - I_1}.$$

Note that your answer may differ by a harmless sign, which ultimately depends on your sign conventions for I_1 and I_2 (i.e. which terminal is A and which terminal is B).

[3] **Problem 3.** USAPhO 2015, problem A2.

Now we give a few problems on current flow through continuous objects. Fundamentally, all one needs for these problems is the definition $\mathbf{J} = \sigma \mathbf{E}$, and superposition.

Example 3

Consider two long, concentric cylindrical shells of radii $a < b$ and length L. The volume between the two shells is filled with material with conductivity $\sigma(r) = k/r$. What is the resistance between the shells, and the charge density?

Solution

To find the resistance, we compute the current I when a voltage V is applied between the shells. By symmetry, in the steady state the current density must be

$$\mathbf{J}(\mathbf{r}) = \frac{I}{2\pi r L}\,\hat{\mathbf{r}}.$$

On the other hand, we also know that

$$V = \int \mathbf{E} \cdot d\mathbf{r} = \int_a^b \frac{I}{2\pi r L \sigma}\,dr = \frac{I(b-a)}{2\pi k L}$$

from which we conclude

$$R = \frac{b-a}{2\pi k L}.$$

Note that the radial electric field between the shells is constant, so

$$\mathbf{E}(\mathbf{r}) = \frac{V}{b-a}\,\hat{\mathbf{r}}.$$

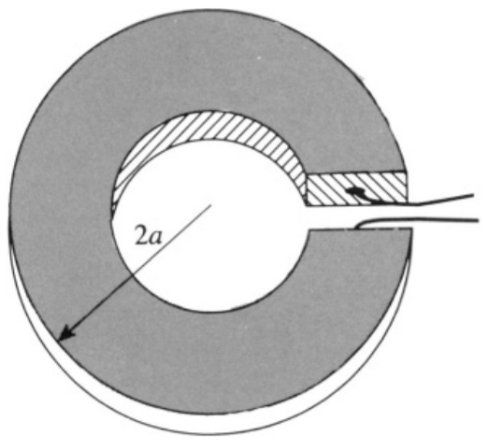
This means that in the steady state, there must be a nonzero charge density between the shells. (If there weren't, then we would have $E(r) \propto 1/r$, rather than a constant.)

To find the charge density explicitly, it's easiest to use Gauss's law in differential form in cylindrical coordinates. We use the form of the divergence derived in **E1**,

$$\nabla \cdot \mathbf{E} = \frac{1}{r} \frac{\partial(rE_r)}{\partial r} + (\text{other terms}) = \frac{1}{r} \frac{V}{b-a} = \frac{\rho}{\epsilon_0}$$

thus showing that the charge density is proportional to $1/r$. Of course, we could also get this result by applying Gauss's law in integral form, to concentric spheres.

[2] **Problem 4** (Cahn)**.** A washer is made of a material of resistivity ρ. It has a square cross section of length a on a side, and its outer radius is $2a$. A small slit is made on one side and wires are connected to the faces exposed.



Since the washer has an irregular shape, the current distribution inside it is complicated: it spreads out from the first wire, goes around the washer, and converges into the second wire. However, the situation is simpler if we glue a perfectly conducting square plate, of side length a, to each exposed face. Find the resistance in this case.

Solution. Since the plates are conducting, the potential doesn't depend on the radius r at the plates themselves, so by rotational symmetry, it doesn't depend on r anywhere in the washer. Therefore, there is no radial current; all the current flows tangentially, so we can think of the washer as a set of radial rings in parallel.

Split the washer into a bunch of radial rings with width dr. We see that r ranges from a to $2a$. Each little ring has resistance $\rho(2\pi r)/(a dr)$, and they are all effectively connected in parallel. Thus,

$$\frac{1}{R} = \int_a^{2a} \rho^{-1} \frac{a}{2\pi r} dr = \rho^{-1} \frac{a}{2\pi} \log 2,$$

giving the answer,

$$R = \frac{2\pi}{\log 2} \frac{\rho}{a}.$$

[3] **Problem 5** (BAUPC 1995)**.** An electrical signal can be transferred between two metallic objects buried in the ground, where the current passes through the Earth itself. Assume that these objects

are spheres of radius r, separated by a horizontal distance $L \gg r$, and suppose both objects are buried a depth much greater than L in the ground. If the Earth has uniform resistivity ρ, find the approximate resistance between the terminals. (Hint: consider the superposition principle.)

Solution. We can consider one object at a time, and then use superposition to find the combined effect of both. Suppose that current I comes out from one of the objects. Placing this object at the origin, we have

$$\mathbf{J} = \frac{I}{4\pi r^2}\hat{\mathbf{r}}, \quad \mathbf{E} = \frac{\rho I}{4\pi r^2}\hat{\mathbf{r}}.$$

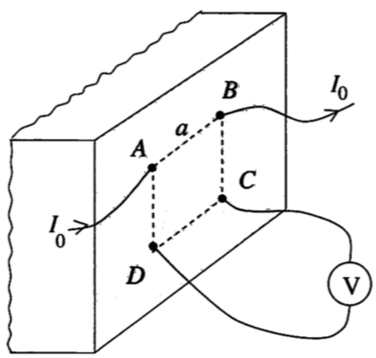
Therefore, the potential difference between this object and where the other object would be is

$$V = \frac{\rho I}{4\pi}\int_r^L \frac{1}{r^2} \approx \frac{\rho I}{4\pi r}$$

where we used $r \ll L$. Finally, the other object takes in current I, with its $\mathbf{J}$, $\mathbf{E}$, and V superposing with the first object. Thus, the total potential difference is $\rho I/2\pi r$, so

$$R = \frac{V}{I} = \frac{\rho}{2\pi r}.$$

[3] **Problem 6** (PPP 162)**.** A plane divides space into two halves. One half is filled with a homogeneous conducting medium, and physicists work in the other. They mark the outline of a square of side a on the plane and let a current I_0 in and out at two of its neighboring corners. Meanwhile, the measure the potential difference V between the two other corners.



Find the resistivity ρ of the medium.

Solution. Consider the case where there is current coming in just at A. Then, $\mathbf{J}$ always points radially outward from A and has (hemi)spherical symmetry, with magnitude

$$J \cdot 2\pi r^2 = I_0 \implies J = \frac{I_0}{2\pi r^2}.$$

Then we have

$$\mathbf{E} = \frac{I_0}{2\pi\sigma r^2}\hat{\mathbf{r}}$$

where σ is the conductivity of the material, which implies

$$V_D - V_C = \int_a^{\sqrt{2}a} \frac{I_0}{2\pi\sigma r^2}dr = \frac{I_0}{2\pi\sigma a}(1 - 1/\sqrt{2}).$$

Similarly, for the case where the current is coming out of B, we have

$$V_D - V_C = \frac{I_0}{2\pi\sigma a}(1 - 1/\sqrt{2}).$$

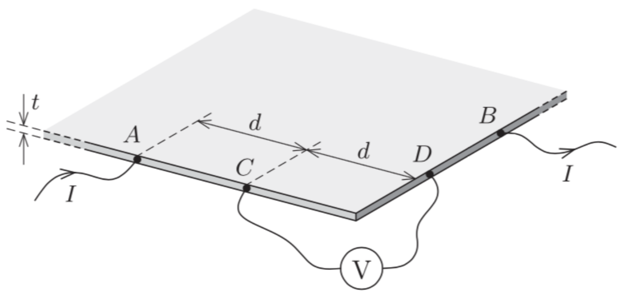
The actual voltage drop is the superposition of the two,

$$\Delta V = \frac{I_0 \rho}{2\pi a}(2 - \sqrt{2})$$

where $\rho = 1/\sigma$ is the resistivity. Then ρ can be calculated as

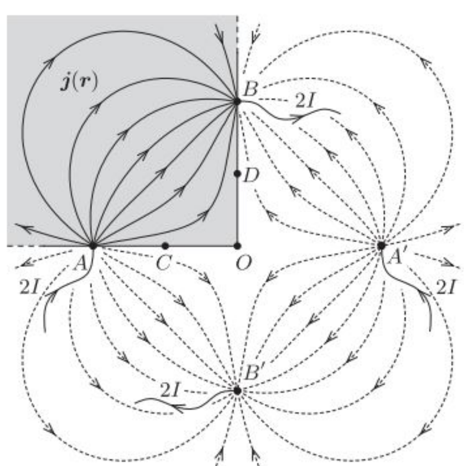
$$\rho = \frac{2\pi a \Delta V}{I_0(2 - \sqrt{2})} = \frac{\pi a(2 + \sqrt{2})\Delta V}{I_0}.$$

[3] **Problem 7** (MPPP 174)**.** We aim to measure the resistivity of the material of a large, thin, homogeneous square metal plate, of which only one corner is accessible. To do this, we chose points A, B, C and D on the side edges of the plate that form the corner.



Points A and B are both $2d$ from the corner, whereas C and D are each a distance d from it. The length of the plate's sides is much greater than d, which, in turn, is much greater than the thickness t of the plate. If a current I enters the plate at point A, and leaves it at B, then the reading on a voltmeter connected between C and D is V. Find the resistivity ρ of the plate material.

Solution. The reason this problem is a lot harder than the previous one is that there is a nontrivial boundary condition, namely that the current density at the edges of the plate is parallel to the plate. The key insight is that we can use the following "image current" configuration to automatically satisfy the original problem's boundary conditions, but on an infinite plate.



We have a new current source and sink respectively at the reflections of A and B in O. The current sources and sinks all have magnitude $2I$, rather than I, because only half of the currents at A and B actually enter and exit the physical plate, shaded in gray.

Now, if a current $2I$ enters the plate, the current at a distance r is $\frac{2I}{2\pi r t}$, so the electric field at a distance r is $\frac{\rho I}{\pi r t}$ (pointing radially outward), so the potential is $\frac{\rho I}{\pi t}\log(r_0/r)$ for some arbitrary r_0, which we'll take to be the same for all current sources and sinks. Then

$$V_C = \frac{\rho I}{\pi t}(\log(r_0/d) + \log(r_0/3d) - 2\log(r_0/\sqrt{5}d)) = \frac{\rho I}{\pi t}\log(5/3).$$

By symmetry, $V_C = -V_D$. Thus,

$$V = 2V_C = \frac{2\rho I}{\pi t}\log(5/3), \qquad \boxed{\rho = \frac{\pi t}{2\log(5/3)}\frac{V}{I}.}$$

Remark

Setups like those in the previous two problems are commonly used to measure resistivities, but why do they use a complicated "four terminal" setup? Wouldn't it have been easier to just attach two terminals, send a current I through them, and measure the voltage drop V? The problem with this is that it also picks up the resistance R of the contacts between the terminals and the material, along with the resistances of the wires. By having a pair of terminals measure voltage alone, drawing negligible current, we avoid this problem.

[4] **Problem 8.** [A] This problem is just for fun; the techniques used here are too advanced to appear on Olympiads. We will prove Rayleigh's monotonicity law, which states that increasing the resistance of any part of a resistor network increases the equivalent resistance between any two points. This may seem obvious, but it's actually tricky to prove. The following is the slickest way.

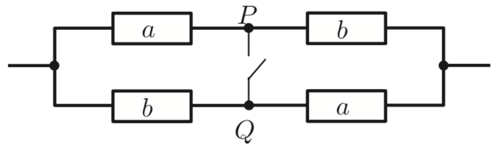
(a) Consider a graph of resistors, where a battery is attached across two of the vertices, fixing their voltages. Write an expression for the total power dissipated, assuming the voltages at each vertex are V_i and the resistances are R_{ij}.

(b) The voltages V_i at all the other vertices are determined by Kirchoff's rules. But suppose you didn't know that, or didn't want to set up those equations. Remarkably, it turns out that you can derive the exact same results by simply treating the voltages V_i as free to vary, and setting them to minimize the total power dissipated! Show this result. (This is an example of a variational principle, like the principle of least action in mechanics.)

(c) For any network of resistors, show that $P = V^2/R$ when V is the battery voltage applied across two vertices, R is the equivalent resistance between them, and P is the total power dissipated in the resistors. (This is intuitive, but it's worth showing in detail to assist with the next part.)

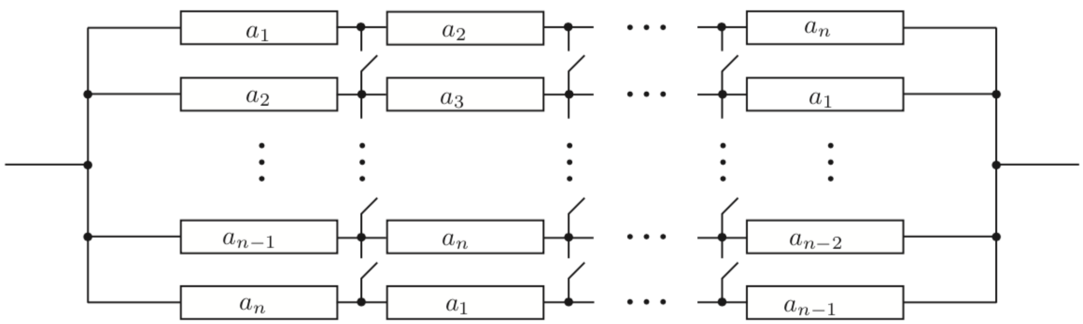
(d) By combining all of these results, prove Rayleigh's monotonicity law.

(e) We can use Rayleigh's monotonicity law to prove some mathematical results. Consider the resistor network shown below, where the variables label the resistances.



By considering the resistances before and after closing the switch PQ, show that the arithmetic mean of two numbers is at least the geometric mean.

(f) Consider the resistor network shown below.



By closing all the switches, show that the arithmetic mean of n numbers is at least the harmonic mean.

Solution. (a) The power is

$$P = \sum_{i<j} \frac{(V_i - V_j)^2}{R_{ij}}.$$

Here the sum over $i < j$ counts all pairs of vertices once. If there is no direct connection between i and j, the resistance R_{ij} is infinite.

(b) The power is minimized when its derivative is zero, and we are free to vary all voltages except for the two points where the battery is connected. Let V_i be one of these voltages. Then

$$\frac{\partial P}{\partial V_i} = \sum_{j \neq i} \frac{2(V_i - V_j)}{R_{ij}} = 0.$$

Now compare this to how we would solve the problem using Kirchoff's laws. The fact that the sum of the voltage drops along a loop is zero is already accounted for, because we already have specified the voltages at each vertex. The only new equations we would write down would be charge conservation at each vertex,

$$\sum_{j \neq i} I_{ij} = 0.$$

However, applying Ohm's law, we see this is precisely the equation that power minimization has given us!

(c) By the definition of the equivalent resistance, $V = IR$ where I is the total current going through the circuit. By the definition of power, the power put in by the battery is $P = IV$, since any current going through the circuit must go through the battery. By conservation of energy, the power dissipated in the circuit is equal to the power put in by the battery. So the power dissipated is $P = IV = V^2/R$.

(d) Put a battery of voltage V across the points we are considering. By part (c) Rayleigh's monotonicity law is equivalent to the statement that, if we increase any of the R_{ij}, the total power P dissipated in the resistor network goes down.

We can account for the effect of increasing one of the R_{ij} in two steps. First, suppose we do so while artificially keeping all the voltages V_i constant. Then by part (a), P decreases. Second, in reality the voltages quickly rearrange themselves to satisfy Kirchoff's laws, which we saw in part (b) is equivalent to minimizing the power. So this further rearrangement can only further decrease P. This shows the desired result.

(e) Before closing the switch, the resistance is

$$R_i = \frac{a+b}{2}.$$

After closing the switch, the resistance is

$$R_f = 2\left(\frac{1}{a} + \frac{1}{b}\right)^{-1} = \frac{2ab}{a+b}.$$

Closing the switch is equivalent to decreasing R_{PQ} from infinity to zero, so $R_f \leq R_i$ by Rayleigh's monotonicity law. This gives

$$\sqrt{ab} \leq \frac{a+b}{2}$$

which is the AM-GM inequality.

(f) Before closing the switches,

$$R_i = \frac{1}{n}\sum_i a_i$$

which is the arithmetic mean. After closing the switches,

$$R_f = n\left(\sum_i \frac{1}{a_i}\right)^{-1}$$

which is the harmonic mean. Thus, the arithmetic mean is at least the harmonic mean.

Remark

You might think that Rayleigh's monotonicity law is too obvious to require a proof; if you decrease a resistance, how could the net resistance possibly go up? In fact, this kind of non-monotonicity occurs very often! For example, Braess's paradox is that fact that adding more roads can slow down traffic, even when the total number of cars stays the same. A U.S. Physics Team coach has argued that allowing more team strategies can make a basketball team score less. For more on this subject, see the paper *Paradoxical behaviour of mechanical and electrical networks* or this video.

Remark

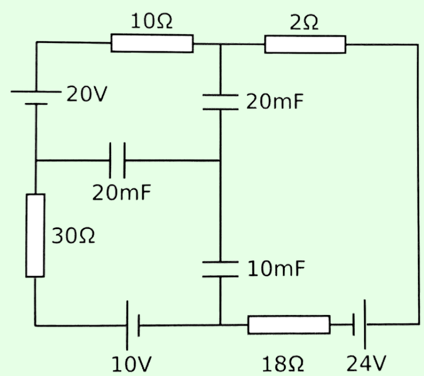
Circuit questions can get *absurdly* hard, but at some point they start being more about mathematical tricks than physics. As a result, I haven't included any such problems here; they tend not to appear on the USAPhO or IPhO, or in college physics, or in real life, or really anywhere besides a few competitions. On the other hand, you might find such questions fun! For some examples, see the Physics Cup problems 2013.6, 2017.2, 2018.1, and 2019.4.

2 RC Circuits

Next we'll briefly cover RC circuits, our first exposure to a situation genuinely changing in time.

Example 4: CPhO

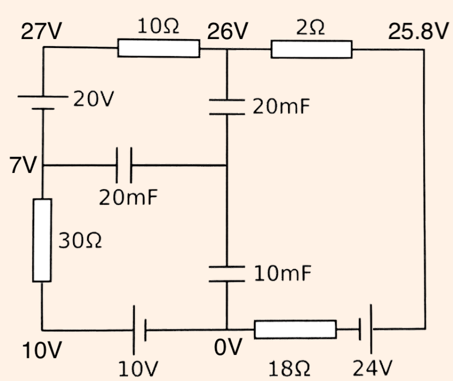
The capacitors in the circuit shown below were initially neutral. Then, the circuit is allowed to reach the steady state.



After a long time, what is the charge stored on the 10 mF capacitor?

Solution

After a long time, no current flows through the capacitors, so there is effectively a single loop in the circuit. It has a total resistance $60\,\Omega$ and a total emf $6\,\mathrm{V}$, so the current is $I = 0.1\,\mathrm{A}$. Using this, we can straightforwardly label the voltages everywhere on the outer loop.



To finish the problem, we need to know the voltage V_0 of the central node, so we need one more equation. That equation is charge conservation: the fact that the central part of the circuit, containing the inner plates of the three capacitors, begins and remains uncharged. Suppressing units, this means

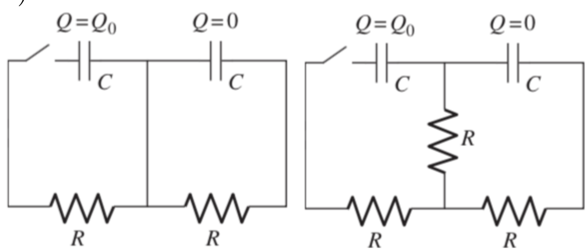
$$20(26 - V_0) + 20(7 - V_0) + 10(0 - V_0) = 0, \quad V_0 = \frac{66}{5}\,\text{V}$$

from which we read off the answer,

$$Q = CV = 0.132\,\text{C}.$$

[3] **Problem 9.** USAPhO 1997, problem A3.

[3] **Problem 10** (Purcell 4.18)**.** Consider the two RC circuits below.



(a) The circuit shown below contains two identical capacitors and two identical resistors, with initial charges as shown above at left. If the switch is closed at $t = 0$, find the charges on the capacitors as functions of time.

(b) Now consider the same setup with an extra resistor, as shown above at right. Find the maximum charge that the right capacitor achieves. (Hint: the methods of **M4** can be useful.)

Solution. (a) Let the two loop currents be I_1 and I_2, both counterclockwise. The loop equations are $Q_1/C = I_1 R$ and $Q_2/C = I_2 R$. We also have $I_k = -\dot{Q}_k$. Thus, $Q_k + RC\dot{Q}_k = 0$ for $k = 1, 2$. Based on the initial conditions, we see then that the solutions are $Q_1(t) = Q_0 e^{-t/RC}$ and $Q_2(t) = 0$. (The simple reason $Q_2(t)$ is zero is because the middle wire effectively shorts out the right half of the circuit.)

(b) Again, with the same setup of variables, we get that

$$Q_1/C - 2I_1R + I_2R = 0$$
$$Q_2/C - 2I_2R + I_1R = 0.$$

This is a system of two linear differential equations, which can be solved using the methods of **M4**. However, in this case we can just add and subtract the equations, giving

$$(Q_1 + Q_2)/C - (I_1 + I_2)R = 0, \quad (Q_1 - Q_2)/C - 3(I_1 - I_2)R = 0.$$

That is, the sum of the two acts like an RC circuit with time constant RC, while the difference acts like one with time constant $3RC$. (These are the "normal modes".) By superposing these solutions and fitting the initial conditions, we get

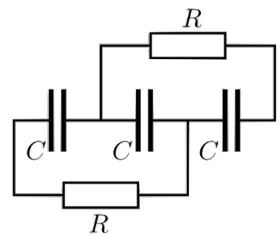
$$Q_1(t) = \frac{Q_0}{2}(e^{-t/RC} + e^{-t/3RC}), \quad Q_2(t) = \frac{Q_0}{2}(e^{-t/RC} - e^{-t/3RC}).$$

We want to maximize $|Q_2|$, so setting the derivative to zero gives $t = \frac{3}{2}RC\log(3)$, so

$$\boxed{|Q_2|_{\max} = \frac{Q_0}{3\sqrt{3}}}.$$

[3] **Problem 11.** USAPhO 2004, problem A1.

[3] **Problem 12** (Kalda)**.** Three identical capacitors are placed in series and charged with a battery of emf $\mathcal{E}$. Once they are fully charged, the battery is removed, and simultaneously two resistors are connected as shown.



Find the heat dissipated on each of the resistors after a long time.

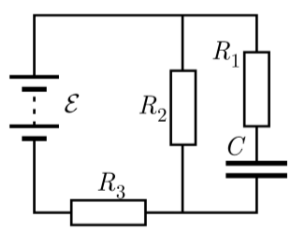
Solution. In the beginning, the charges on the plates are $\mathcal{E}C/3, -\mathcal{E}C/3, \mathcal{E}C/3, -\mathcal{E}C/3, \mathcal{E}C/3, -\mathcal{E}C/3$. After a long time, let the charges on the plates be $q_1, -q_1, q_2, -q_2, q_3, -q_3$. Note that all currents are 0 now, so we may effectively ignore the resistors and treat the wires as zero resistance. Therefore, the potential at points connected by wires is the same, so $q_1 = -q_2 = q_3$. Also, by charge conservation on the two disjoint pieces ($q_1, -q_2, q_3$ and $-q_1, q_2, -q_3$), we see

$$q_1 - q_2 + q_3 = \mathcal{E}C/3,$$

which implies $|q_1| = |q_2| = |q_3| = \mathcal{E}C/9$. The energy is $\sum \frac{1}{2}Q^2/C$, so the charges dropping by a factor of 3 means we lose 8/9 of the original total energy, so each resistor dissipates 4/9 of the original total energy. This is

$$\frac{4}{9} \cdot 3 \cdot \frac{(\mathcal{E}C/3)^2}{2C} = \boxed{\frac{2}{27}\mathcal{E}^2 C}.$$

[3] **Problem 13** (Kalda)**.** Find the time constant of the RC circuit shown below.



Solution. For the purposes of computing the time constant, it is equivalent to assume the capacitor is already charged, then take out the battery and see how it *discharges*. Thus all that matters is the resistance between the capacitor plates, which is

$$R = R_1 + \frac{R_2R_3}{R_2+R_3} = \frac{R_1R_2 + R_1R_3 + R_2R_3}{R_2+R_3},$$

so $\tau = \boxed{C\frac{R_1R_2 + R_1R_3 + R_2R_3}{R_2+R_3}}$.

[3] **Problem 14** (MPPP 175/176). A metal sphere of radius R has charge Q and hangs on an insulating cord. It slowly loses charge because air has a conductivity σ. In all cases, neglect any magnetic or radiation effects.

(a) Find the time for the charge to halve.

(b) You should have found that the time is independent of the radius R of the sphere, which follows directly from dimensional analysis. Can you show that, in fact, it is completely independent of the shape? (This doesn't just follow from dimensional analysis, because the shape might be described by dimensionless numbers, such as the eccentricity of an ellipsoid.)

(c) Air has a conductivity of $\sigma \sim 10^{-13}\,\Omega^{-1}\text{m}^{-1}$, while water has a conductivity of $\sigma \sim 10^{-2}\,\Omega^{-1}\text{m}^{-1}$. About how long does the charge on an object last, if it is in air or water?

This problem generalizes USAPhO 2010, problem A2, which you can compare.

Solution. (a) We can analyze this as an RC circuit. (The circuit is completed by the "sphere at infinity".) The capacitance is the self-capacitance of the sphere,

$$C = 4\pi\epsilon_0 R.$$

The resistance is the resistance between the sphere and infinity. The air can be thought of as a set of resistors in series, with each resistor being a spherical shell of air. Then

$$R_{\text{eq}} = \int dR = \frac{1}{\sigma}\int_R^\infty \frac{dr}{4\pi r^2} = \frac{1}{4\pi\sigma R}.$$

This gives a time constant of

$$\tau = RC = \frac{\epsilon_0}{\sigma}.$$

Therefore, the time is

$$t = \frac{\epsilon_0}{\sigma}\log 2.$$

(b) Of course, dimensional analysis doesn't work, because there might be dimensionless parameters describing a general shape (e.g. the eccentricity of an ellipsoid). Instead we use the following more general argument. We note that

$$I = \oint \mathbf{J}\cdot d\mathbf{S}, \quad \Phi_E = \oint \mathbf{E}\cdot d\mathbf{S}$$

over any surface completely enclosing the object. The right-hand sides are related by $\mathbf{J} = \sigma\mathbf{E}$, and Gauss's law gives $\Phi_E = Q/\epsilon_0$. Combining these gives

$$\dot{Q} = -\frac{\sigma}{\epsilon_0}Q$$

so the charge decreases exponentially with timescale ϵ_0/σ, completely independently of the shape. (Of course, the sphere is still special, because with the sphere we are guaranteed there are *no* magnetism or radiation effects (why?). For a general shape, we have to assume these effects are negligible.)

(c) The relevant timescale is ϵ_0/σ. Thus we find $t \sim 10\,\text{s}$ for air, and $t \sim 1\,\text{ns}$ for water.

[5] **Problem 15.** IPhO 1993, problem 1. A really neat question with real-world relevance.

[5] **Problem 16.** IPhO 2007, problem "orange". A combination of mechanics and RC circuits.

3 Computing Magnetic Fields

Idea 4

The Biot–Savart law is

$$\mathbf{B} = \frac{\mu_0 I}{4\pi} \oint \frac{d\mathbf{s} \times \mathbf{r}}{r^3}.$$

As a consequence, we have Ampere's law,

$$\oint \mathbf{B} \cdot d\mathbf{s} = \mu_0 I, \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{J}$$

as well as Gauss's law for magnetism,

$$\oint \mathbf{B} \cdot d\mathbf{S} = 0, \quad \nabla \cdot \mathbf{B} = 0.$$

Idea 5

The force on a stationary wire carrying current I in a magnetic field $\mathbf{B}$ is

$$\mathbf{F} = I \int d\mathbf{s} \times \mathbf{B}.$$

The energy of a magnetic field is

$$U = \frac{1}{2\mu_0} \int B^2 \, dV.$$

The magnetic dipole moment of a planar current loop of area A and current I is $m = IA$, with $\mathbf{m}$ directed perpendicular to the loop by the right-hand rule.

Idea 6: Magnetic Dipoles

Far from a magnetic dipole with magnetic moment m, its magnetic field is just the same as the electric field of an electric dipole,

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0 m}{4\pi r^3}(2\cos\theta\,\hat{\mathbf{r}} + \sin\theta\,\hat{\boldsymbol{\theta}}) = \frac{\mu_0}{4\pi r^3}(3(\mathbf{m} \cdot \hat{\mathbf{r}})\hat{\mathbf{r}} - \mathbf{m}).$$

As with the electric dipole field, you don't need to memorize this, but you should remember that it's proportional to the dipole moment, falls off as $1/r^3$, and be able to sketch it.

You should have already seen basic examples of using the Biot–Savart law in Halliday and Resnick, such as the field of a circular ring of current on its axis. We'll start with some problems that are similarly straightforward, but more technically complex.

[3] **Problem 17.** This is a key question which will help you understand idea 7. A spherical shell with radius R and uniform surface charge density σ spins with angular frequency ω about a diameter.

(a) Find the magnetic field at the sphere's center.

(b) Find the magnetic dipole moment of the sphere.

(c) It can be shown that (1) the magnetic field inside the sphere is uniform, and (2) the magnetic field outside the sphere is exactly that of a magnetic dipole. (It requires doing some obnoxious integrals, as can be seen in section 5.4 of Griffiths.) Using this information, make a qualitatively accurate sketch of the field.

(d) There is a closely related question in electrostatics: suppose we had two spherical shells of the same radius R, and surface charge densities $\pm\sigma$, and the shells were displaced by a small distance $d \ll R$. Qualitatively, what would the electric field of this setup look like? How would it differ from your answer to part (c)? You can neglect the region between the shells.

Solution. (a) First, we find the field due to a ring of counterclockwise current I with radius a in the $z = 0$ plane at a point directly above the center at some height z. Using the Biot–Savart law, we see that

$$\mathbf{B} = \frac{\mu_0 I}{4\pi} \frac{2\pi a \frac{a}{\sqrt{a^2+z^2}}}{a^2 + z^2}\, \hat{\mathbf{z}} = \frac{\mu_0 I}{2} \frac{a^2}{(a^2 + z^2)^{3/2}}\, \hat{\mathbf{z}}.$$

Let us work in spherical coordinates with the axis of rotation being the z axis. Then, at angle θ, we essentially have a ring of charge of radius $a = R\sin\theta$, $z = R\cos\theta$, and

$$dI = \sigma R(d\theta)\omega(R\sin\theta) = \sigma R^2 \omega \sin\theta\, d\theta.$$

Therefore,

$$d\mathbf{B} = \frac{\mu_0 \sigma R^2 \omega \sin\theta\, d\theta}{2} \frac{R^2 \sin^2\theta}{R^3}\, \hat{\mathbf{z}} = \frac{\hat{\mathbf{z}}}{2} \mu_0 \sigma\omega R \sin^3\theta\, d\theta.$$

Integrating from 0 to π to obtain the full field,

$$\mathbf{B} = \frac{\hat{\mathbf{z}}}{2}\, \mu_0 \sigma\omega \int_0^\pi \sin^3\theta\, d\theta = \frac{2}{3} \mu_0 \sigma\omega R\, \hat{\mathbf{z}}.$$

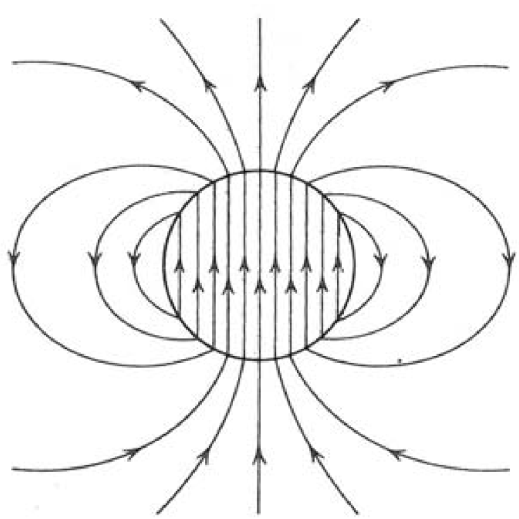
(b) The magnetic dipole moment of a slice is

$$d\mathbf{m} = \hat{\mathbf{z}}\pi(R\sin\theta)^2 \omega\sigma R^2 \sin\theta\, d\theta.$$

Integrating this gives

$$\mathbf{m} = \frac{4}{3}\pi\omega\sigma R^4\, \hat{\mathbf{z}}.$$

(c) The field is as shown below.



The key feature is that the field lines of the dipole outside and the uniform field inside match up perfectly, so that every field line forms a closed loop; this is Gauss's law for magnetism.

(d) In this case, it is straightforward to show by the shell theorem that the electric field is uniform inside (this is essentially a problem in **E1**), and exactly an electric dipole field outside. The crucial difference is that the uniform field inside points *along* the dipole moment, rather than against it. As a result, $\mathbf{E}$ switches directions upon entering the sphere. This reflects the fact that $\nabla \times \mathbf{E} = 0$, so that the line integral of $\mathbf{E}$ along any closed path has to vanish. If we follow a field line, this is only possible if the field switches direction. In addition, compared to the magnetic case, the field lines don't close up, because $\nabla \cdot \mathbf{E}$ isn't zero.

[2] **Problem 18** (Purcell 6.12). A ring with radius R carries a current I. Show that the magnetic field due to the ring, at a point in the plane of the ring, a distance r from the center, is given by

$$B = \frac{\mu_0 I}{2\pi} \int_0^\pi \frac{(R - r\cos\theta)R\, d\theta}{(r^2 + R^2 - 2rR\cos\theta)^{3/2}}.$$

In the $r \gg R$ limit, show that

$$B \approx \frac{\mu_0}{4\pi} \frac{m}{r^3}$$

where $m = IA$ is the magnetic dipole moment of the ring, as expected from idea 6.

Solution. Let the ring be centered at the origin, and let the field point be $a\hat{\mathbf{x}}$, and say we are at an angle θ. Then, $d\mathbf{l} = (-R\sin\theta)\,\hat{\mathbf{x}} + (R\cos\theta)\,\hat{\mathbf{y}}$, and

$$\mathbf{r} = a\,\hat{\mathbf{x}} - R(\cos\theta\,\hat{\mathbf{x}} + \sin\theta\,\hat{\mathbf{y}}) = (a - R\cos\theta)\,\hat{\mathbf{x}} + (-R\sin\theta)\,\hat{\mathbf{y}}.$$

Note that $r = |\mathbf{r}| = \sqrt{a^2 + R^2 - 2aR\cos\theta}$. Therefore, by the Biot–Savart law,

$$d\mathbf{B} = \frac{\mu_0 I}{4\pi} \frac{R(R - a\cos\theta)\hat{\mathbf{z}}\, d\theta}{(a^2 + R^2 - 2aR\cos\theta)^{3/2}},$$

so integrating from 0 to 2π and noting that θ and $-\theta$ contribute the same, we arrive at the desired result. Now, to take the $r \gg R$ limit cleanly and consistently, it's best to nondimensionalize everything. Defining $x = R/r \ll 1$, we can pull dimensionful factors out of the integral to get

$$B = \frac{\mu_0 I}{2\pi} \frac{rR}{r^3} \int_0^\pi \frac{x - \cos\theta}{(1 + x^2 - 2x\cos\theta)^{3/2}}\, d\theta.$$

Now, it's not immediately obvious to what order in x we should expand in. If we already know the answer is proportional to $1/r^3$, then we can see the answer must be first order in x. But if we didn't know that, we could expand to zeroth order, giving

$$B = \frac{\mu_0 I}{2\pi} \frac{R}{r^2} \int_0^\pi (-\cos\theta)\, d\theta = 0.$$

The fact that the answer vanishes means we need to go to higher order to find the true answer. At first order, applying the binomial theorem, the integrand is

$$(x - \cos\theta)(1 + x^2 - 2x\cos\theta)^{-3/2} \approx (x - \cos\theta)(1 + 3x\cos\theta) \approx -\cos\theta + x(1 - 3\cos^2\theta)$$

where we threw away higher order terms throughout. Then

$$B = \frac{\mu_0 I}{2\pi} \frac{R^2}{r^3} \int_0^\pi (1 - 3\cos^2\theta)\, d\theta.$$

Using the fact that $\cos^2\theta$ averages to 1/2 over a cycle, the integral is $-\pi/2$, giving

$$B = \frac{\mu_0 I}{4\pi}\frac{\pi R^2}{r^3} = \frac{\mu_0}{4\pi}\frac{m}{r^3}$$

as desired.

[3] **Problem 19** (Purcell 6.14)**.** Consider a square loop with current I and side length a centered at the origin, with sides parallel to the x and y axes. Show that the magnetic field at $r\hat{\mathbf{x}}$ is $B \approx (\mu_0/4\pi)(m/r^3)$ for $r \gg a$, as expected from idea 6. Be careful with factors of 2!

Solution. This calculation is a bit subtle. It is tempting to ignore the sides parallel to $\hat{\mathbf{x}}$, because the current is almost parallel to $\mathbf{r}$, so $d\mathbf{s} \times \mathbf{r}$ is small; more precisely, it is suppressed by a power of a/r. The sides parallel to $\hat{\mathbf{y}}$ do each give much larger contributions, but they have opposite sign and nearly cancel out, suppressing their sum by a power of a/r. So all four sides need to be considered.

First consider the segments parallel to $\hat{\mathbf{x}}$. We get a factor of $(a/2)/r$ from the $\sin\theta$ factor in the cross product. Similarly, a appears in the Biot–Savart integral in the denominator; however, its effect here would give higher-order terms in a/r, which we don't want to keep since they're much smaller than the final answer. So the segments each contribute equally,

$$B_1 = -\frac{\mu_0 I}{4\pi}\left(\frac{a\frac{a/2}{r}}{r^2} + \frac{a\frac{a/2}{r}}{r^2}\right) = -\frac{\mu_0 I}{4\pi}\frac{a^2}{r^3}.$$

Next, the segments parallel to $\hat{\mathbf{y}}$ contribute

$$B_2 = \frac{\mu_0 I}{4\pi}\left(\frac{a}{(r-a)^2} - \frac{a}{(r+a)^2}\right) = \frac{\mu_0 I}{4\pi}\frac{2a^2}{r^3}$$

where we work to the same accuracy as for B_1. Adding the two contributions gives the desired result. If you forget to count B_1, you'll get an answer that is two times too big.

[3] **Problem 20.** USAPhO 2012, problem A3.

Idea 7: Magnetic Monopoles

Far away from the center of the dipole, the magnetic field of a magnetic dipole has the same form as the electric field of an electric dipole. Therefore, we can often replace a magnetic dipole m with a fictitious pair of "magnetic charges" $\pm q_m$ separated by d, where $q_m d = m$. This is called a "Gilbert dipole", in contrast to a true "Amperian dipole".

This was the default way to think about magnets in the 1800s, but was largely removed from American textbooks in the 1950s because it's misleading in general: magnetic charges don't actually exist in magnets, and applying this analogy will give the wrong fields inside the dipole, as you saw in problem 17 and will see another way in problem 21. However, if we only care about the field outside the magnet, the analogy works, and it's often the fastest way to solve problems. We'll return to this idea in greater depth in **E8**.

[3] **Problem 21.** USAPhO 2015, problem B2. A key problem which illustrates idea 7.

We now give a few arguments for computing fields using symmetry.

Example 5: PPP 31

An electrically charged conducting sphere "pulses" radially, i.e. its radius changes periodically with a fixed amplitude. What is the net pattern of radiation from the sphere?

Solution

There is no radiation. By spherical symmetry, the magnetic field can only point radially. But then this would produce a magnetic flux through a Gaussian sphere centered around the pulsing sphere, which would violate Gauss's law for magnetism. So there is no magnetic field at all, and since radiation always needs both electric and magnetic fields (as you'll see in **E7**), there is no radiation at all. In fact, outside the sphere the electric field is always exactly equal to $Q/4\pi\epsilon_0 r^2$, in accordance with Coulomb's law.

Example 6

Find the magnetic field of an infinite cylindrical solenoid, of radius R and n turns per unit length, carrying current I.

Solution

Orient the solenoid along the vertical direction and use cylindrical coordinates. By symmetry, the field must be independent of z. Now consider the radial component of the magnetic field B_r. Turning the solenoid upside-down is equivalent to reversing the current. But the former does not flip B_r while the latter does, so we must have $B_r = 0$.

Now, by rotational symmetry, the tangential component B_ϕ must be uniform. But then Ampere's law on any circular loop gives $B_\phi(2\pi r) = 0$, so we must have $B_\phi = 0$ as well.

The only thing left to consider is B_z. By applying Ampere's law to small vertical rectangles, we see that B_z is constant unless that rectangle crosses the surface of the solenoid. Furthermore, B_z must be zero far from the solenoid, so it must be zero everywhere outside the solenoid. Now, for a rectangle of height h that does cross the surface, Ampere's law gives

$$\oint \mathbf{B} \cdot d\mathbf{s} = B_z^{\text{in}} h = \mu_0 I_{\text{enc}} = \mu_0 n I h$$

which tells us that $B_z^{\text{in}} = \mu_0 n I$.

Example 7

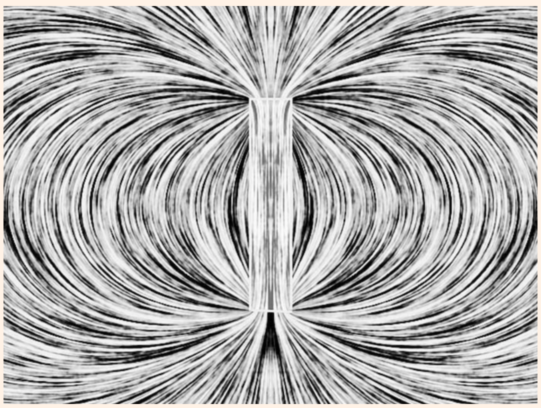
Now suppose the solenoid has finite length $L \gg R$. What do the fringe fields look like?

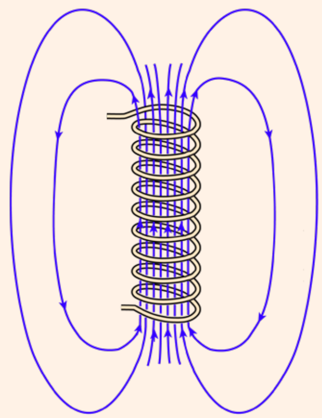
Solution

In principle we could solve for the exact fringe field by applying the Biot–Savart law to the solenoid wire, but that would be rather complicated. Instead, let's approximate the solenoid as a stack of $N = nL$ evenly spaced circular wire loops. Each one of these loops is a magnetic dipole $\mu = \pi R^2 I$, so the field of each loop well outside of it is just a dipole field.

Summing up all of these dipole fields is still complicated, so let's use idea 7. We can replace each wire loop with a pair of magnetic charges $\pm q_m$ separated by d, with the same magnetic dipole moment $\mu = q_m d$. If we choose $d = 1/n$, then the charges of adjacent dipoles cancel, leaving only charges $q_m = \pm n\mu = \pm \pi R^2 n I$ on the ends.

Thus, the fringe field of a solenoid, at distances much greater than R, looks like the electric field of two point charges! This is confirmed by a numeric calculation shown at left below.





This may come as a surprise to you if you've read basic, algebra-based introductory physics textbooks. Many of them contain hand-drawn diagrams like the one shown at right above, where all the magnetic flux comes neatly out the ends of the solenoids, in straight lines. In reality, the field sprays out almost spherically symmetrically from the end, with only half the flux actually going out through the end face, while the rest exits downward through the sides. (You will show this more directly with a slick argument in problem 23.)

We can also be more quantitative. Suppose the solenoid is vertical and centered at $z = 0$. Then the field at a radius r from the solenoid axis, at $z = 0$, is

$$\mathbf{B}(r) = \mu_0 n I \hat{\mathbf{z}} \times \begin{cases} 1 & r < R \\ -2R^2/L^2 & R \ll r \ll L \\ -R^2 L/4r^3 & L \ll r \end{cases}$$

where the first line is the usual solenoid field, the second line is from applying Coulomb's law to our dipole analogy (which is only valid when $R \ll r$), and the third is from the dipole field of the two charges (only valid when $L \ll r$). As expected, in the limit $L \gg R$, the fringe field outside the solenoid is negligible. Another way of phrasing the result is that most of the upward flux through the solenoid returns through a downward field which mainly extends out to $r \sim L$. You can see all of these features in the accurate drawing above.

We can draw two lessons from this example. First, misleading diagrams are a common problem in introductory textbooks. A general rule is that the more basic a textbook is, the more pictures it'll have, but the less useful they'll be. Second, the analogy between Ampere and Gilbert dipoles is quite useful, and shows up frequently in tricky Olympiad problems.

Remark: Real Solenoids

Real solenoids are even more complicated. First, we didn't account for the discreteness of the wires. We just treated them as forming a uniform current per length $K = nI$, which is how we wrote $I_{\text{enc}} = nIh$. This is valid when you don't care about looking too closely, i.e. if your distance to any wire is much larger than the wire spacing $1/n$.

Second, the fact that solenoids are made by winding real wires means there is another contribution to the current, even in the limit $n \to \infty$. The wires are wound with a small slope, since a net current I still has to move along the solenoid. Another way of saying this is that the current per length along the solenoid surface is $\mathbf{K} = nI\hat{\boldsymbol{\theta}} + (I/2\pi R)\hat{\mathbf{z}}$. This causes a tangential magnetic field $B_\phi = \mu_0 I/2\pi r$ outside the solenoid. Thus, in practice many solenoids are "counterwound": half the wires are wound evenly spaced going up the axis, and the other half are wound evenly spaced going back down the axis, which closes the loop and cancels this unwanted field.

[2] **Problem 22.** A toroidal solenoid is created by wrapping N turns of wire around a torus with a rectangular cross section. The height of the torus is h, and the inner and outer radii are a and b.

(a) In the ideal case, the magnetic field vanishes everywhere outside the toroid, and is purely tangential inside the toroid. Find the magnetic field inside the toroid.

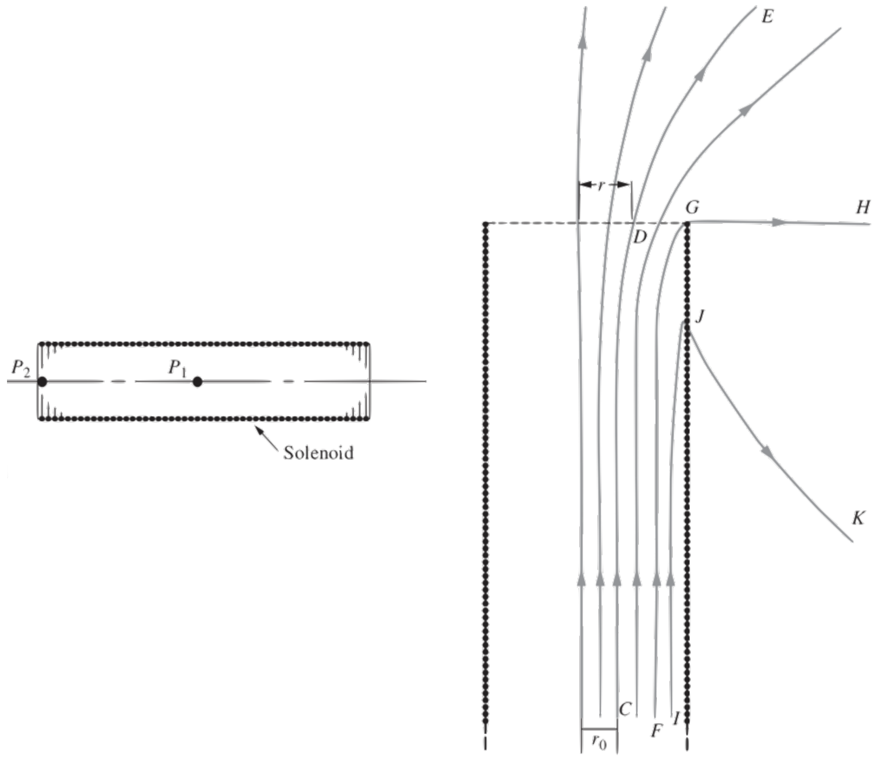
(b) There is another small contribution to the magnetic field due to the winding effect mentioned above. Roughly what does the resulting extra magnetic field look like? If you didn't want this additional field, how would you design the solenoid to get rid of it?

Solution. (a) Applying Ampere's law on a circular loop gives $B(r)(2\pi r) = \mu_0 NI$, so

$$B(r) = \frac{\mu_0 NI}{2\pi r}.$$

(b) Note that the twisting of the wire adds an effective small current in the tangential direction. This looks like a current loop, so, e.g. it produces a magnetic field pointing vertically through the toroid's hole. We can remove it by using a bunch of current loops instead of a single winding wire, or by using counterwinding: after winding the wire around the toroid once clockwise, wind it around again counterclockwise.

[3] **Problem 23** (Purcell 6.63)**.** A number of simple facts about the fields of solenoids can be found by using superposition. The idea is that two solenoids of the same diameter, and length L, if joined end to end, make a solenoid of length $2L$. Two semi-infinite solenoids butted together make an infinite solenoid, and so on.



Prove the following facts.

(a) In the finite-length solenoid shown at left above, the magnetic field on the axis at the point P_2 at one end is approximately half the field at the point P_1 in the center. (Is it slightly more than half, or slightly less than half?)

(b) In the semi-infinite solenoid shown at right above, the field line FGH, which passes through the very end of the winding, is a straight line from G out to infinity.

(c) The flux through the end face of the semi-infinite solenoid is half the flux through the coil at a large distance back in the interior.

(d) Any field line that is a distance r_0 from the axis far back in the interior of the coil exits from the end of the coil at a radius $r_1 = \sqrt{2}r_0$, assuming $\sqrt{2}r_0$ is less than the solenoid radius.

Solution. (a) Let $\mathbf{B}_1$ and $\mathbf{B}_2$ be the fields at these points, respectively. Note that $\mathbf{B}_1$ is close to the ideal value $\mu_0 nI$, but smaller because the solenoid is not infinite. Now glue two of these solenoids together end-to-end, and consider the field at the center of this new, bigger solenoid. By superposition, it is $2\mathbf{B}_2$, but also, it is close to $\mu_0 nI$, and it is slightly closer to $\mu_0 nI$ than $\mathbf{B}_1$ is, since the combined solenoid is longer. Therefore, $\mathbf{B}_2$ is slightly more than half of $\mathbf{B}_1$.

(b) Let G' be the reflection of G in the axis. Say the field line GH comes out at an angle θ. Then, at G', it also comes out with an angle θ. Now, making a copy and rotating $180°$ and flipping the current direction, the field at G becomes one pointing at θ above the horizontal (coming from G at the original), and one at angle $\pi - \theta$ to the horizontal (coming from G' in the copy). Therefore, the field there would be non-zero right outside the solenoid, unless $\theta = 0$, in which case the fields cancel.

(c) Do the same procedure as in (a), and the flux at the glue points gets doubled to what it was originally. However, now we have an infinite solenoid, so double the flux through the end is equal to the flux in the middle.

(d) Note that (c) holds even if we take a constant disk of radius a as our surface to take the flux over. Note that the flux through the disk at the edge with radius r is the same as at the middle with radius r_0 (same field lines). However, if we draw a disk of radius r at the middle, it will have twice the flux as it did at the top, or twice the flux as with r_0. However, here in the middle, the magnetic field is essentially constant, so the areas must be twice each other, so $\pi r^2 = 2\pi r_0^2$, or $r = \sqrt{2} r_0$.

[3] **Problem 24** (MPPP 160)**.** Two infinite parallel wires, a distance d apart, carry electric currents along the z-axis with equal magnitudes but opposite directions. We can find the shape of the magnetic field lines with a neat trick, which only works for "two-dimensional" setups like this one, where the fields lie in the xy plane and don't depend on z.

(a) Argue that if we rotated $\mathbf{B}$ by 90° in the xy plane at each point, it would produce a valid electrostatic field $\mathbf{E}$. (Hint: consider rotating the $\mathbf{B}$ field of each wire individually.)

(b) Argue that the field lines of $\mathbf{B}$ are the same as the equipotentials of this artificial $\mathbf{E}$, and use this to find the field lines.

This trick is also useful for fluids in two dimensions, where it swaps vortices with sources and sinks.

Solution. (a) First, we can get the intuition using a single wire. In this case,

$$\mathbf{B} = \frac{\mu_0 I}{2\pi r}\hat{\boldsymbol{\theta}}$$

in cylindrical coordinates. Upon a 90° rotation, $\hat{\boldsymbol{\theta}}$ turns into $\hat{\mathbf{r}}$, giving

$$\mathbf{E} = \frac{\mu_0 I}{2\pi r}\hat{\mathbf{r}}$$

which is a valid electrostatic field, as it's simply the electric field of an charged wire with linear charge density $\lambda = \mu_0 \epsilon_0 I$. So by superposition, rotating the $\mathbf{B}$ field of the two wires would also give a valid electrostatic field. (Of course, this isn't really physically meaningful, since electric and magnetic fields don't even have the same units. It's just a mathematical trick.)

We can also prove the correspondence more generally. The key criterion for a valid magnetostatic field is $\nabla \cdot \mathbf{B} = 0$, which for such two-dimensional setups is $\partial_x B_x + \partial_y B_y = 0$. Now, when we rotate by 90°, we define an electric field by $E_y = B_x$ and $E_x = -B_y$, which implies $\partial_x E_y - \partial_y E_x = 0$. But in such a two-dimensional setup, this is equivalent to $\nabla \times \mathbf{E} = 0$, which is the condition to have a valid electrostatic field.

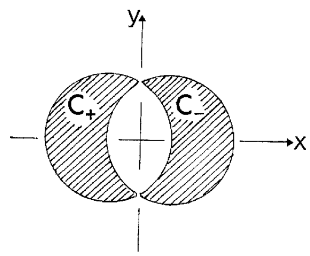
(b) The field lines of $\mathbf{B}$ are always parallel to $\mathbf{B}$. Now, this artificial $\mathbf{E}$ is always perpendicular to $\mathbf{B}$, and equipotentials are always perpendicular to $\mathbf{E}$, so the equipotentials follow the magnetic field lines.

On the other hand, we know precisely what the potential is in this problem. By integrating the $1/r$ field, the potential is proportional to $\log r$, so

$$V(r) \propto \log(r_+) - \log(r_-) = \log(r_+/r_-)$$

where r_+ and r_- are the distances to the two wires. So the equipotentials have constant r_+/r_-. We've already found, when investigating the method of images for spheres in **E2**, that this implies the equipotentials are *circles*, specifically circles of Appolonius. So the magnetic field lines are circles!

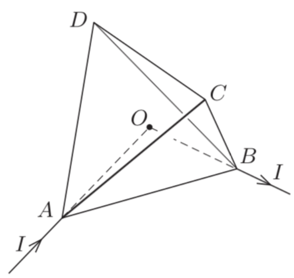
[2] **Problem 25** (IPhO 1996)**.** Two straight, long conductors C_+ and C_-, insulated from each other, carry current I in the positive and the negative $\hat{\mathbf{z}}$ direction respectively. The cross sections of the conductors are circles of diameter D in the xy plane, with a distance $D/2$ between the centers.



The current in each conductor is uniformly distributed. Find the magnetic field in the space between the conductors.

Solution. See page 12 of the official solutions here.

[3] **Problem 26** (MPPP 157)**.** A regular tetrahedron is made of a wire with constant resistance per unit length. A long, straight wire sends current I into one vertex, and another long, straight wire removes it from another vertex, as shown.



Find the magnetic field at the center of the tetrahedron.

Solution. By symmetry C and D are at the same potential, so $I_{DC} = 0$. Then the current from A to B just splits up into three branches, which have resistances $R_{ACB} = R_{ADB} = 2R_{AB}$. Therefore, the currents are

$$I_{AB} = \frac{1}{2}I, \quad I_{AC} = I_{AD} = I_{CB} = I_{DB} = \frac{1}{4}I.$$

The field at O due to the current along AD is directed along the vector $\overrightarrow{CB}$. Similarly, the magnetic field due to the current along AC is directed along $\overrightarrow{BD}$, and so on. By repeating this reasoning for all five contributions, we find that the magnetic field at O is proportional to

$$2\overrightarrow{DC} + \overrightarrow{BD} + \overrightarrow{CB} + \overrightarrow{AD} + \overrightarrow{CA} = 2\overrightarrow{DC} + \overrightarrow{CD} + \overrightarrow{CD} = 0$$

so there is no field at O.

[5] **Problem 27.** APhO 2013, problem 1. A neat question on a cylindrical RC circuit that uses many of the techniques we've covered so far.

[5] **Problem 28.** EuPhO 2023, problem 3. Another neat question, in a setup where an eddy current can be computed exactly. For this problem, you won't need to know anything about magnetism besides the fact that the magnetic force per unit charge is $\mathbf{v} \times \mathbf{B}$.

Solution. See the official solutions here.

Electromagnetism IV: Lorentz Force

The problems here mostly use material covered in previous problem sets, though chapter 5 of Purcell covers relativistic field transformations. For further interesting physical examples, see chapter II-29 of the Feynman lectures. There is a total of **81** points.

1 Electrostatic Forces

Idea 1: Lorentz Force

A charge q in an electromagnetic field experiences the force

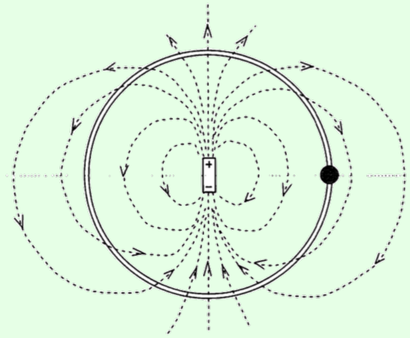
$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

In particular, a stationary wire carrying current I in a magnetic field experiences the force

$$\mathbf{F} = I \int d\mathbf{s} \times \mathbf{B}.$$

Example 1: PPP 183

A small charged bead can slide on a circular, frictionless insulating ring. A point-like electric dipole is fixed at the center of the circle with the dipole's axis lying in the plane of the circle. Initially the bead is in the plane of symmetry of the dipole, as shown.



Ignoring gravity, how does the bead move after it is released? How would the bead move if the ring weren't there?

Solution

Set up spherical coordinates so that the dipole is in the $\hat{\mathbf{z}}$ direction. Then

$$V(r, \theta) = \frac{kp\cos\theta}{r^2}.$$

Since the ring fixes r, the potential on the ring is just proportional to $\cos\theta$, which is in turn proportional to z. But a potential linear in z is equivalent to a uniform downward field, so the bead oscillates like the mass of a pendulum, with amplitude $\pi/2$.

The answer remains the same when the ring is removed! Conservation of energy states that

$$\frac{kqp\cos\theta}{r^2} + \frac{1}{2}mv^2 = 0.$$

Let N be the normal force. Then accounting for radial forces gives

$$N + q\frac{\partial V}{\partial r} = \frac{mv^2}{r}.$$

However, plugging in our conservation of energy result for v^2 shows that $N = 0$, so the ring doesn't actually do anything, and it may be removed without effect.

Example 2

A parallel plate capacitor with separation d and area A is attached to a battery of voltage V. One plate moves towards the other with uniform speed v. Verify that energy is conserved.

Solution

The capacitance is $C = A\epsilon_0/d$. The power supplied by the battery is

$$P_{\text{batt}} = IV = V\,\frac{dQ}{dt} = V^2\frac{dC}{dt}.$$

On the other hand, the rate of change of the energy stored in the capacitor is

$$P_{\text{cap}} = \frac{d}{dt}\left(\frac{1}{2}CV^2\right) = \frac{1}{2}V^2\frac{dC}{dt}.$$

At first glance, there seems to be a problem. But then we remember that there is an attractive force between the plates, so the plates do work on whatever is moving them together,

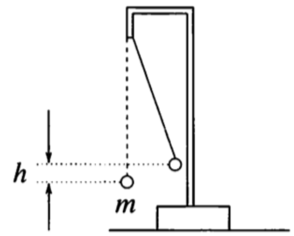
$$P_{\text{mech}} = Fv = \frac{QE}{2}\,v = \frac{QV}{2d}\,v = \frac{1}{2}CV^2\,\frac{v}{d} = \frac{1}{2}V^2\frac{dC}{dt}.$$

where E is the electric field inside the capacitor. Thus, $P_{\text{batt}} = P_{\text{cap}} + P_{\text{mech}}$ as required.

Technically there's energy in the magnetic field too, but it's smaller than the electric field energy by v^2/c^2, and thus negligible unless you're moving the plates so fast that relativity comes into play. Most problems in this problem set ignore such relativistic effects.

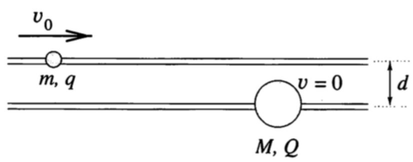
[2] **Problem 1** (PPP 193)**.** Two positrons are at opposite corners of a square of side a. The other two corners of the square are occupied by protons. All particles have charge q, and the proton mass M is much larger than the positron mass m. Find the approximate speeds of the particles much later.

[3] **Problem 2** (PPP 114)**.** A small positively charged ball of mass m is suspended by an insulating thread of negligible mass. Another positively charged small ball is moved very slowly from a large distance until it is in the original position of the first ball. As a result, the first ball rises by h.



How much work has been done?

[3] **Problem 3** (PPP 71)**.** Two small beads slide without friction, one on each of two long horizontal parallel fixed rods a distance d apart.



The masses of the beads are m and M and they carry charges q and Q. Initially, the larger mass M is at rest and the other one is far away approaching it at a speed v_0. For what values of v_0 does the smaller bead ever get to the right of the larger bead?

[2] **Problem 4** (PPP 192)**.** Classically, a conductor is made of nuclei of positive charge fixed in place, and electrons that are free to move.

(a) Consider a solid conductor in a gravitational field $\mathbf{g}$. Argue that the electric field inside the conductor is *not* zero; find out what it is.

(b) Now suppose a positron is placed at the center of a hollow spherical conductor in a gravitational field $\mathbf{g}$. Find its initial acceleration.

[3] **Problem 5.** USAPhO 2008, problem B2. You may ignore part (c), which was removed in the final version of the exam, though you can also do it for extra practice.

[3] **Problem 6.** USAPhO 2019, problem B1.

[5] **Problem 7.** IPhO 2004, problem 1. A nice question on the dynamics of a multi-part system.

2 The Lorentz Force

Idea 2

Some questions below will involve special relativity. The Lorentz force law as written in idea 1 is still valid as long as $\mathbf{F}$ is interpreted as $d\mathbf{p}/dt$, where the relativistic momentum is

$$\mathbf{p} = \gamma m \mathbf{v}, \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}.$$

The relativistic energy is also modified to

$$E = \gamma mc^2 = mc^2 + \frac{1}{2} mv^2 + \ldots.$$

We will return to this subject in more detail in **R2**, but for now this is all you need.

Example 3: Kalda 163

A beam of electrons, of mass m and charge q, is emitted with a speed v almost parallel to a uniform magnetic field $\mathbf{B}$. The initial velocities of the electrons have an angular spread of $\alpha \ll 1$, but after a distance L the electrons converge again. Neglecting the interaction between the electrons, what is L?

Solution

Consider an electron initially traveling at an angle α to the magnetic field. This electron has a speed $v_\parallel = v\cos\alpha \approx v$ parallel to the field, and a speed $v_\perp v \sin\alpha \approx v\alpha$ perpendicular to the field. The component $v_\parallel$ always stays the same, while $v_\perp$ rotates, so the electron spirals along the field lines.

The acceleration of the electron is

$$a_\perp = \frac{F}{m} = \frac{qv_\perp B}{m}.$$

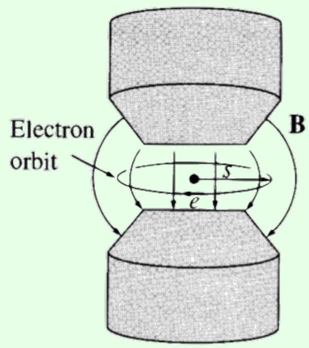
The perpendicular velocity component rotates through a circle in velocity space of circumference $2\pi v_\perp$. After one such circle, the total perpendicular displacement is zero, so the beam refocuses. Thus we have

$$L = \frac{2\pi v_\perp}{a} v_\parallel \approx \frac{2\pi m v}{qB}.$$

In other words, this setup acts like a magnetic "lens".

Example 4: Griffiths 7.50

In a "betatron", electrons move in circles in a magnetic field. When the magnetic field is slowly increased, the accompanying electric field will impart tangential acceleration.



Suppose the field always has the same spatial profile $B(r,t) = B_0(r)f(t)$. For what $B_0(r)$ is it possible for an electron to start at rest in zero magnetic field, and then move in a circle of constant radius as the field is increased?

Solution

The electrons experience a tangential force

$$\dot{p} = qE = q\frac{\dot{\Phi}_B}{2\pi r} = \frac{qr}{2}\dot{B}_{\text{av}}$$

where B_{av} is the average field over the orbit. Since the particles start from rest in zero field, we can integrate this to find

$$p = \frac{qr}{2}B_{\text{av}}.$$

On the other hand, the standard result for cyclotron motion is $p = qrB$, which means we must have $B = B_{\text{av}}/2$, i.e. the field at any radius is half the average magnetic field inside,

$$B(r) = \frac{1}{2}\frac{1}{\pi r^2}\int_0^r B(r')(2\pi r')\,dr'.$$

This rearranges slightly to give

$$r^2 B(r) = \int_0^r r' B(r')\,dr'.$$

Differentiating both sides with respect to r, we have

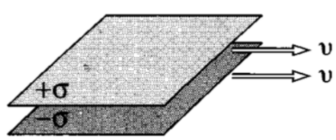
$$2rB(r) + r^2 B'(r) = rB(r)$$

which simplifies to

$$\frac{dB}{B} = -\frac{dr}{r}$$

which means the field profile should be $B_0(r) \propto 1/r$. (Of course, a real betatron might differ since it only needs to obey $B = B_{\text{av}}/2$ at the radii where electrons will be orbiting.)

[3] **Problem 8** (Griffiths 5.17). A large parallel plate capacitor with uniform surface charge σ on the upper plate and $-\sigma$ on the lower is moving with a constant speed v as shown.



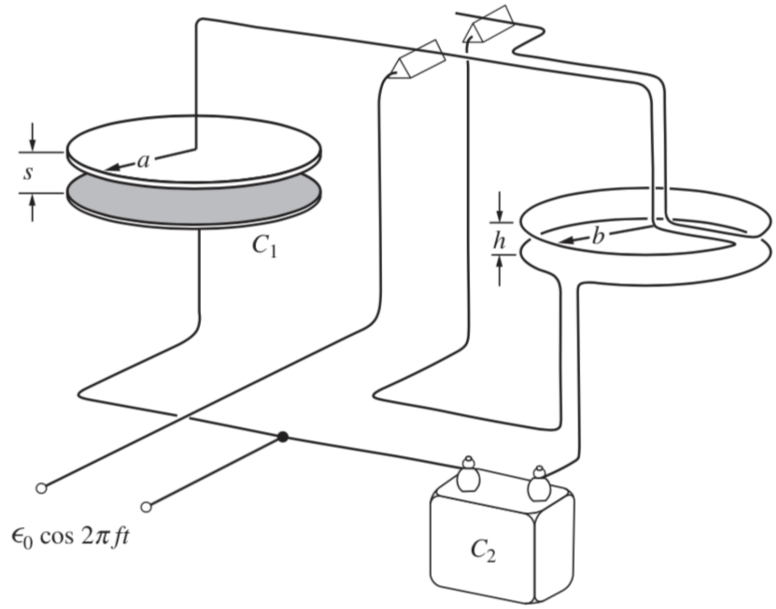
(a) Find the magnetic field between the plates and also above and below them.

(b) Find the magnetic force per unit area on the upper plate, including its direction.

(c) What happens to the net force between the plates in the limit $v \to c$? Explain your result using some basic ideas from special relativity.

[3] **Problem 9.** EFPhO 2012, problem 7. An elegant Lorentz force problem with wires. (If you enjoy this problem, consider looking at IdPhO 2020, problem 1B, which has a similar setup but requires three-dimensional reasoning. The official solutions are here.)

[4] **Problem 10** (Purcell 6.35/INPhO 2008.6). Consider the arrangement shown below.



The force between capacitor plates is balanced against the force between parallel wires carrying current in the same direction. A voltage alternating sinusoidally with angular frequency ω is applied to the parallel-plate capacitor C_1 and also to the capacitor C_2, and the current is equal to the current through the rings. Assume that $s \ll a$ and $h \ll b$.

Suppose the weights of both sides are adjusted to balance without any applied voltage, and C_2 is adjusted so that the time-averaged downward forces on both sides are equal. Show that

$$\frac{1}{\sqrt{\mu_0 \epsilon_0}} = \sqrt{2\pi}\, a\omega \sqrt{\frac{b}{h}\frac{C_2}{C_1}}.$$

The left-hand side is equal to c, as we'll show in **E7**, so this setup measures the speed of light.

[3] **Problem 11.** An electron beam is accelerated from rest by applying an electric field E for a time t, and subsequently guided by magnetic fields. These magnetic fields are produced with a series of coils, which carry currents I_i.

Now suppose the apparatus is repurposed to shoot proton beams. Suppose a proton beam is accelerated from rest by applying an electric field E for a time t (in the opposite direction). Let an electron have mass m and a proton have mass M.

(a) Find the currents I_i needed so that the proton follows the same trajectory the electron did, assuming V is small enough that both the electron and proton are nonrelativistic.

(b) How does the answer change if relativistic corrections are accounted for?

[5] **Problem 12.** IPhO 2000, problem 2. A solid question on the Lorentz force with real-world relevance. Requires a little relativity, namely the expressions for relativistic momentum/energy.

[4] **Problem 13.** IPhO 1996, problem 2. An elegant problem on particles in a magnetic field. (There's a deeper principle behind the solution to this problem; see **R3** for more discussion.)

3 Magnetic Moments

[3] **Problem 14.** Consider a current loop I in the xy plane in a constant magnetic field $\mathbf{B}$.

(a) Show that the net force on the loop is zero.

(b) Show that the torque is

$$\boldsymbol{\tau} = \mathbf{m} \times \mathbf{B}$$

where the magnetic moment is

$$\mathbf{m} = IA\hat{\mathbf{z}}$$

where A is the area of the loop. For simplicity, you can show this in the case where the current loop is a square of side length L, whose sides are aligned with the x and y axes. (The proof for a general loop shape requires some vector calculus, but you can attempt it for a challenge. You'll need the double cross product identity, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = 0$.)

Idea 3

The force on a small magnetic dipole $\mathbf{m}$ is

$$\mathbf{F} = \nabla(\mathbf{m} \cdot \mathbf{B}).$$

If there are no other currents at the dipole's location, so that $\nabla \times \mathbf{B} = 0$, this is equivalent to $\mathbf{F} = (\mathbf{m} \cdot \nabla)\mathbf{B}$, which is sometimes easier to evaluate.

As in problem 14, this can be shown relatively easily for a square loop, and requires some tricky vector calculus for a general current distribution. Both the force and torque on a magentic dipole can be found by differentiating the potential energy

$$U = -\mathbf{m} \cdot \mathbf{B}.$$

All of these results also hold for electric dipoles, if we replace $\mathbf{m}$ with $\mathbf{p}$ and $\mathbf{B}$ with $\mathbf{E}$.

Remark

The expression for the potential energy above is notoriously subtle. Here's the problem: we know the Lorentz force on a charge is $q\mathbf{v} \times \mathbf{B}$, which means magnetic fields never do work. So how can they be associated with a nonzero potential energy?

There are two levels of explanation. First, suppose the magnetic dipole is made of charges moving in a loop. When such a current loop is placed in a magnetic field, and moved or rotated, mechanical work can be done on the loop. But at the same time, there will be an induced emf in the loop, which speeds up or slows down the current. The work done by these two effects perfectly cancels, so that the energy of the loop stays constant. For this kind of dipole, the expression for U doesn't indicate the total energy, but only the "mechanical" potential energy, in the sense that differentiating it gives the right forces and torques. (Some further discussion of this point is in chapter II-15 of the Feynman lectures.)

On the other hand, the magnetic dipole moment of a common bar magnet doesn't come from charges moving in a loop! Instead, it comes from the intrinsic magnetic dipole moments of the unpaired electrons in the magnet. These kinds of dipole moments aren't composed of any moving subcomponents; they are an elementary and immutable property of the electron, like its mass or charge. In these cases, $U = -\mathbf{m} \cdot \mathbf{B}$ really is the total energy, and the magnetic field *can* do work. You won't hear much about these elementary dipole moments in introductory books, because they can only be properly understood by combining relativity and quantum mechanics, but they're responsible for most magnetic phenomena.

Example 5

If a magnet is held over a table, it can pick up a paper clip. If the paper clip is removed, it can pick up another paper clip just as well, and this process can seemingly continue forever without any effect on the magnet. Since the magnet does work on each paper clip, doesn't this mean a permanent magnet is an infinite energy source?

Solution

This is the kind of question that makes magnets feel so mysterious. They're basically the only everyday example of a long range force besides gravity (in fact, Kepler once thought the Sun acted on the planets like a giant magnet), and as such they've inspired countless attempts at perpetual motion machines. For centuries, many people have spent years of their lives trying to get elaborations of this example to work.

To see why this doesn't work for a bar magnet, just replace the word "magnet" with "charge". It's true that a positive charge can attract a negative charge to it. And if the negative charge is then removed, the positive charge can then attract another negative charge to it. But conservation of energy isn't violated, because the force from the positive charge is conservative: the work it does on the negative charge to draw it close is precisely the opposite of the work an external agent needs to do to pull it away. The force of a magnet on a paper clip is also conservative.

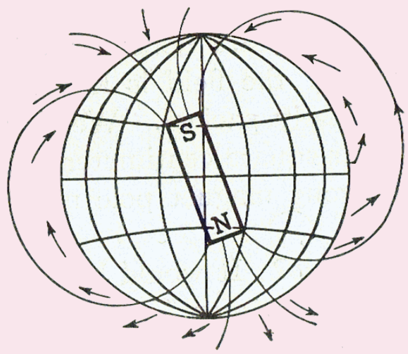
It's also interesting to consider a slightly different case. Unlike a bar magnet, an electromagnet (i.e. a magnet created by moving current in a loop) can be turned on and off with the flick of a switch. Therefore, we might suspect that the following is a perpetual motion machine:

1. Turn on the electromagnet, which costs energy E_0.
2. Use it to lift a paper clip, increasing its potential energy by mgh.
3. Turn off the electromagnet, which costs energy E_0, while holding the paper clip.
4. Move the paper clip away; we've managed to raise it higher for free.

To see the problem, note that the attractive force between the magnet and paper clip arises because the magnet induces a magnetic dipole moment in the paper clip, leading to a $(\mathbf{m} \cdot \nabla)\mathbf{B}$ force. As the paper clip moves toward the magnet, its own dipole moment causes a changing magnetic flux through the electromagnet, and thus an emf against the current. Therefore, it costs extra energy to keep the current in the electromagnet steady. Since the $q\mathbf{v} \times \mathbf{B}$ Lorentz force doesn't do work, that energy must be precisely mgh, so nothing comes for free.

Remark

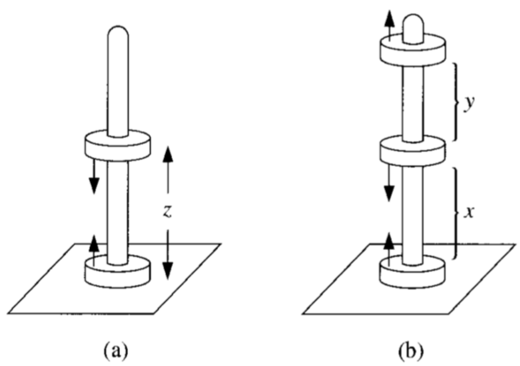
A compass needle is essentially a small magnetic dipole, whose dipole moment points towards the end painted red. We can also approximate the Earth's magnetic field as a dipole field.



Since the tangential component of this dipole field points north, the red end of the compass points towards the geographic north pole, which is the Earth's magnetic south pole.

By the way, a cheap compass calibrated to work in America or Europe won't work well in Australia. The reason is that the Earth's magnetic field also has a radial component, which acts to tip the compass needle up or down. The needle needs to be appropriately weighted to stay horizontal, so that it can freely rotate, but the side that needs to be weighted differs between the hemispheres.

[3] **Problem 15** (Griffiths 6.23)**.** A familiar toy consists of donut-shaped permanent magnets which slide frictionlessly on a vertical rod.

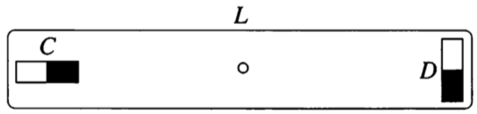


Treat the magnets as dipoles with mass m_d and dipole moment $\mathbf{m}$, with directions as shown above.

(a) If you put two back-to-back magnets on the rod, the upper one will "float". At what height z does it float?

(b) If you now add a third magnet parallel to the bottom one as shown, find the ratio x/y of the two heights, using only a scientific calculator. (Answer: 0.85.)

[3] **Problem 16** (PPP 89)**.** Two identical small bar magnets are placed on opposite ends of a rod of length L as shown.



(a) Show that the torques the magnets exert on each other are *not* equal and opposite.

(b) Suppose the rod is pivoted at its center, and the magnets are attached to the rod so that they can spin about their centers. If the magnets are released, the result of part (a) implies that they will begin spinning. Explain how this can be consistent with energy and angular momentum conservation, treating the latter quantitatively.

4 Point Charges

In this section we'll give a sampling of classic problems involving just point charges in fields; these will be a bit more mathematically advanced than the others in this problem set.

[3] **Problem 17.** A point charge q of mass m is released from rest a distance d from a grounded conducting plane. Find the time until the point charge hits the plane. (Hint: use Kepler's laws.)

[3] **Problem 18.** A point charge of mass m and charge q is released from rest at the origin in the fields $\mathbf{E} = E\hat{\mathbf{x}}$, $\mathbf{B} = B\hat{\mathbf{y}}$. Find its position as a function of time by solving the differential equations given by Newton's second law, $\mathbf{F} = m\mathbf{a}$.

[3] **Problem 19** (Wang)**.** Two identical particles of mass m and charge q are placed in the xy plane with a uniform magnetic field $B\hat{\mathbf{z}}$. The particles have paths $\mathbf{r}_1(t)$ and $\mathbf{r}_2(t)$. Neglect relativistic effects, but account for the interaction between the charges.

(a) Write down a differential equation describing the evolution of the separation $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$.

(b) Suppose that the initial conditions have been set up so that the particles orbit each other in a circle in the xy plane, with constant separation d. What is the smallest d for which this motion is possible?

[4] **Problem 20.** [A] Consider a point charge of mass m and charge q in the field of a magnetic monopole at the origin,

$$\mathbf{B} = \frac{g}{r^2}\hat{\mathbf{r}}.$$

In this problem we'll investigate the strange motion that results.

(a) Argue that the speed v is constant.

(b) Show that the angular momentum $\mathbf{L}$ of the charge is *not* conserved, but that

$$\mathbf{V} = \mathbf{L} - qg\hat{\mathbf{r}}$$

is. The second term is the angular momentum stored in the fields of the charge and monopole.

(c) Show that the charge moves on the surface of a cone. (Hint: in spherical coordinates where the z-axis is parallel to $\mathbf{V}$, consider $\mathbf{V} \cdot \hat{\boldsymbol{\phi}}$.) Sketch some typical trajectories.

One can do problem 18 slickly using field transformations, an advanced subject we will cover in **R3**.

Idea 4: Field Transformations

If the electromagnetic field is $(\mathbf{E}, \mathbf{B})$ in one reference frame, then in a reference frame moving with velocity $\mathbf{v}$ with respect to this frame, the components of the field parallel to $\mathbf{v}$ are

$$E'_{\parallel} = E_{\parallel}, \quad B'_{\parallel} = B_{\parallel}$$

while the components perpendicular are

$$\mathbf{E}'_{\perp} = \gamma(\mathbf{E}_{\perp} + \mathbf{v} \times \mathbf{B}), \quad \mathbf{B}'_{\perp} = \gamma\left(\mathbf{B}_{\perp} - \frac{\mathbf{v}}{c^2} \times \mathbf{E}\right).$$

Remark

The nonrelativistic limit of the field transformation is useful, but one has to be careful in deriving it. You might think, what's the need for care? Can't we just send $c \to \infty$, Taylor expand the above expressions, and call it a day? The problem with this reasoning is that there's no such thing as setting $c \to \infty$. You can't change a fundamental constant, and moreover this statement isn't even dimensionally correct, as noted in **P1**. What we really mean by the nonrelativistic limit is restricting our attention to some subset of possible situations, within which relativistic effects don't matter.

For example, if we have a bunch of point charges with typical speed v, then the nonrelativistic limit is considering only situations where v/c is small. In other words, we are taking $v/c \to 0$, not $c \to \infty$. Since the magnetic field of a point charge is v/c^2 times the electric field, the magnetic field ends up small. Now if we also consider boosts with small speeds v, then expanding the field transformations to lowest order in v/c gives

$$\mathbf{E}' = \mathbf{E}, \quad \mathbf{B}' = \mathbf{B} - \frac{\mathbf{v}}{c^2} \times \mathbf{E}.$$

This is the nonrelativistic limit for situations where $E/B \gg c$, also called the electric limit.

However, there's another possibility. Suppose that we have a bunch of neutral wires. In this case, it's the electric fields that are small, $E/B \ll c$. Using this in the transformations above, we arrive at the distinct result

$$\mathbf{B}' = \mathbf{B}, \quad \mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}$$

which apply for situations where $E/B \ll c$, also called the magnetic limit.

You might think we could improve the approximation by combining the two,

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}, \quad \mathbf{B}' = \mathbf{B} - \frac{\mathbf{v}}{c^2} \times \mathbf{E}$$

but this isn't self-consistent. For example, if you apply a Galilean boost with speed v, and then a boost with speed $-v$, you don't get back the same fields you started with! A sensible Galilean limit is only possible if $E/B \gg c$ or $E/B \ll c$, which are called the electric and magnetic limits, discussed further in this classic paper. It's only in relativity that E and B can be treated on an equal footing.

[3] **Problem 21.** Using the Galilean field transformations to solve problem 18.

(a) In the magnetic limit, show that the Lorentz force stays the same between frames, as it should. Then use the field transformations to find an appropriate reference frame where the problem becomes easy.

(b) In the electric limit, show that the Lorentz force stays the same up to terms that are order $(v/c)^2$ smaller, assuming $B/E \sim v/c^2$. (This is fine, since we're taking the limit $v/c \to 0$ anyway.) Then use the field transformations to find an appropriate reference frame where the problem becomes easy.

(c) The solutions you found in parts (a) and (b) should look very different, even though you should have found only one type of behavior in problem 18. In fact, there is a critical value of E/B separating the two kinds of behavior. What is this critical value, and why didn't you run into it when solving problem 18?

5 Continuous Systems

Example 6: The Drude Model

Model a conductor as a set of electrons, of charge q, mass m, and number density n, which are completely free. Assume that in every small time interval dt, each electron has a probability dt/τ of hitting a lattice ion, which randomizes the direction of its velocity. Under these assumptions, compute the resistivity of the material.

Solution

First, suppose the electrons have some average momentum $\langle \mathbf{p} \rangle$ each. Because the collisions randomize the velocity, the average momentum falls exponentially with timescale τ,

$$\frac{d\langle \mathbf{p} \rangle}{dt} = -\frac{\langle \mathbf{p} \rangle}{\tau}.$$

On the other hand, if there is an applied field, a force term appears on the right,

$$\frac{d\langle \mathbf{p} \rangle}{dt} = -\frac{\langle \mathbf{p} \rangle}{\tau} + q\mathbf{E}$$

since $\mathbf{F} = d\mathbf{p}/dt$ for each individual electron. In the steady state,

$$\langle \mathbf{p} \rangle = q\mathbf{E}\tau.$$

The current density is

$$\mathbf{J} = nq\langle \mathbf{v} \rangle = \frac{nq\langle \mathbf{p} \rangle}{m} = \frac{nq^2\tau}{m}\mathbf{E}.$$

Thus, the resistivity in the Drude model is

$$\rho = \frac{m}{nq^2\tau}.$$

We can also compute the typical drift velocity,

$$v = \frac{qE\tau}{m}.$$

For values of m that give reasonable ρ, the value of v is a literal snail's pace, which is why people say that the electrons themselves move very slowly through a circuit. (Of course, a current can get started in a circuit much faster, because when a battery is attached, each moving electron pushes on the next one along the wire, and this wave of motion travels much faster than the electrons themselves.)

Remark: The Drude–Sommerfeld Model

Above we tacitly assumed there was a given probability of collision per unit time, but that's not right: when a particle flies through a medium, there is instead a given probability of collision per unit *length* it travels. These are equivalent for electrons moving at constant speed, but intuitively, we would expect electrons to have to accelerate starting from rest after each collision, in which case the two differ. To estimate this quickly, note that if the typical collision distance is ℓ, the kinetic energy picked up between collisions is $mv^2/2 \sim qE\ell$, giving typical speed $v \propto \sqrt{E}$. The analogue of Ohm's law would then be $I \propto \sqrt{V}$, completely contrary to observation!

The resolution is that electrons in solids really do effectively move with almost constant speed, even after collisions. This is a quantum mechanical effect, as explained in **X1**. The Pauli exclusion principle implies the electrons in the conductor have to occupy different quantum states, and the high density of electrons requires most of them to always have extremely high speeds, on the order of 1% of the speed of light! The drift velocity is merely the tiny amount by which their velocities are shifted on average.

[2] **Problem 22.** Consider Drude theory again, but now suppose there is also a fixed magnetic field $B\hat{\mathbf{z}}$. In this case, $\mathbf{J}$ is not necessarily parallel to $\mathbf{E}$, but the relation between the two can be described by the "tensor of resistivity". That is, the components are related by

$$E_i = \sum_{j \in \{x,y,z\}} \rho_{ij} J_j.$$

Calculate the coefficients ρ_{ij}. Express your answers in terms of the quantities

$$\rho_0 = \frac{m}{nq^2\tau}, \quad \omega_0 = \frac{qB}{m}$$

as well as the parameter τ.

Example 7: Griffiths 5.40

Since parallel currents attract, the currents within a single wire should contract. To estimate this, consider a long wire of radius r. Suppose the atomic nuclei are fixed and have uniform density, while the electrons move along the wire with speed v. Furthermore, assume that the electrons contract, filling a cylinder of radius $r' < r$ with uniform negative charge density, and that the wire is overall neutral. Find r'.

Solution

The contraction of the electrons produces an overall inward electric field, and hence an outward electric force on each electron, which balances the radially inward magnetic force. Specifically, equilibrium occurs when $E = vB$.

Let the charge densities of the nuclei and electrons be ρ_+ and ρ_-. The magnetic field at radius r is found by Ampere's law, which gives

$$(2\pi r)B = \mu_0(\rho_- v)(\pi r^2), \quad B = \frac{\mu_0 \rho_- v r}{2}.$$

The electric field at radius r is found by Gauss's law, which gives

$$(2\pi r)E = \frac{1}{\epsilon_0}(\rho_+ + \rho_-)\pi r^2, \quad E = \frac{1}{2\epsilon_0}(\rho_+ + \rho_-)r.$$

Note that both E and B are proportional to r. Then $E = vB$ can be satisfied at all r simultaneously, which confirms that our assumption that ρ_+ and ρ_- were uniform is self-consistent.

Plugging these results into $E = vB$ yields

$$\rho_+ + \rho_- = \rho_-(\epsilon_0 \mu_0 v^2) = \rho_- \frac{v^2}{c^2}.$$

This can be written in terms of the Lorentz factor of special relativity,

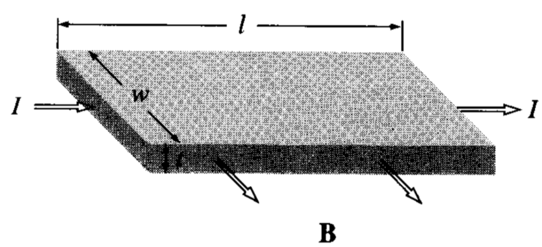
$$\rho_- = -\gamma^2 \rho_+, \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}.$$

Since the wire is overall neutral, $\rho_- r'^2 + \rho_+ r^2 = 0$, so

$$r' = \frac{r}{\gamma}.$$

For nonrelativistic motion, the contraction is extremely small. (However, in plasmas, where the positive charges are also free to move, this so-called pinch effect can be very significant.)

[2] **Problem 23** (Griffiths 5.41)**.** A current I flows to the right through a rectangular bar of conducting material, in the presence of a uniform magnetic field **B** pointing out of the page, as shown.



(a) If the moving charges are positive, in what direction are they deflected by the magnetic field? This deflection results in an accumulation of charge on the upper and lower surfaces of the bar, which in turn produces an electric force to counteract the magnetic one. Equilibrium occurs when the two exactly cancel. (This phenomenon is known as the Hall effect.)

(b) Find the resulting potential difference, called the Hall voltage, between the top and bottom of the bar, in terms of B, the speed v of the charges, and the dimensions of the bar.

(c) How would the answer change if the moving charges were negative?

When measurements were performed in the early 20th century, some metals were found to have *positive* moving charges! This "anomalous Hall effect" was solved by the quantum theory of solids, as you can learn in any solid state physics textbook. (It is related to the strange behavior you will see in problem 26.) Today, extensions of the Hall effect, such as the integer and fractional quantum Hall effects, remain active areas of research, and could be used to build quantum computers. We'll return to these effects in **X3**.

[3] **Problem 24.** USAPhO 1997, problem B1. A nice problem on the dynamics of a plasma.

[3] **Problem 25.** USAPhO 2019, problem A3. This is a tough but useful problem. The first half derives the so-called Child–Langmuir law, covered in problem 2.53 of Griffiths.

[3] **Problem 26.** USAPhO 2022, problem B3. About the weird behavior of electrons in solids.

Electromagnetism IV: Lorentz Force

The problems here mostly use material covered in previous problem sets, though chapter 5 of Purcell covers relativistic field transformations. For further interesting physical examples, see chapter II-29 of the Feynman lectures. There is a total of **81** points.

1 Electrostatic Forces

Idea 1: Lorentz Force

A charge q in an electromagnetic field experiences the force

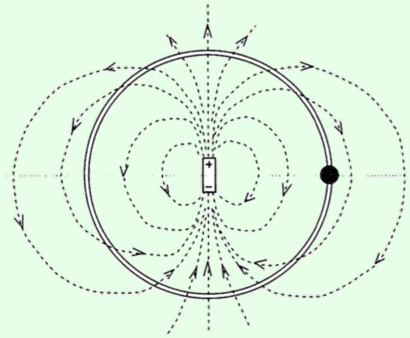
$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

In particular, a stationary wire carrying current I in a magnetic field experiences the force

$$\mathbf{F} = I \int d\mathbf{s} \times \mathbf{B}.$$

Example 1: PPP 183

A small charged bead can slide on a circular, frictionless insulating ring. A point-like electric dipole is fixed at the center of the circle with the dipole's axis lying in the plane of the circle. Initially the bead is in the plane of symmetry of the dipole, as shown.



Ignoring gravity, how does the bead move after it is released? How would the bead move if the ring weren't there?

Solution

Set up spherical coordinates so that the dipole is in the $\hat{\mathbf{z}}$ direction. Then

$$V(r, \theta) = \frac{kp\cos\theta}{r^2}.$$

Since the ring fixes r, the potential on the ring is just proportional to $\cos\theta$, which is in turn proportional to z. But a potential linear in z is equivalent to a uniform downward field, so the bead oscillates like the mass of a pendulum, with amplitude $\pi/2$.

The answer remains the same when the ring is removed! Conservation of energy states that

$$\frac{kqp\cos\theta}{r^2} + \frac{1}{2}mv^2 = 0.$$

Let N be the normal force. Then accounting for radial forces gives

$$N + q\frac{\partial V}{\partial r} = \frac{mv^2}{r}.$$

However, plugging in our conservation of energy result for v^2 shows that $N = 0$, so the ring doesn't actually do anything, and it may be removed without effect.

Example 2

A parallel plate capacitor with separation d and area A is attached to a battery of voltage V. One plate moves towards the other with uniform speed v. Verify that energy is conserved.

Solution

The capacitance is $C = A\epsilon_0/d$. The power supplied by the battery is

$$P_{\text{batt}} = IV = V\,\frac{dQ}{dt} = V^2\frac{dC}{dt}.$$

On the other hand, the rate of change of the energy stored in the capacitor is

$$P_{\text{cap}} = \frac{d}{dt}\left(\frac{1}{2}CV^2\right) = \frac{1}{2}V^2\frac{dC}{dt}.$$

At first glance, there seems to be a problem. But then we remember that there is an attractive force between the plates, so the plates do work on whatever is moving them together,

$$P_{\text{mech}} = Fv = \frac{QE}{2}\,v = \frac{QV}{2d}\,v = \frac{1}{2}CV^2\,\frac{v}{d} = \frac{1}{2}V^2\frac{dC}{dt}.$$

where E is the electric field inside the capacitor. Thus, $P_{\text{batt}} = P_{\text{cap}} + P_{\text{mech}}$ as required.

Technically there's energy in the magnetic field too, but it's smaller than the electric field energy by v^2/c^2, and thus negligible unless you're moving the plates so fast that relativity comes into play. Most problems in this problem set ignore such relativistic effects.

[2] **Problem 1** (PPP 193)**.** Two positrons are at opposite corners of a square of side a. The other two corners of the square are occupied by protons. All particles have charge q, and the proton mass M is much larger than the positron mass m. Find the approximate speeds of the particles much later.

Solution. The idea is that since the positrons are so light, they will be extremely far away before the protons hardly move. Let v_1 be their final speed. Then, energy conservation tells us that

$$\frac{kq^2}{a}\left(4 + \frac{2}{\sqrt{2}}\right) \approx \frac{kq^2}{\sqrt{2}a} + 2\left(\frac{1}{2}mv_1^2\right).$$

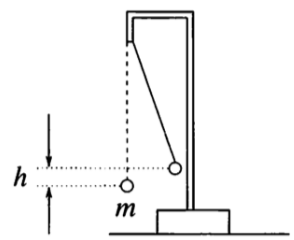
Solving for v_1 yields

$$v_1 = \sqrt{\frac{kq^2}{am}(4 + 1/\sqrt{2})}.$$

The speed of the protons can be calculated by assuming that the positrons didn't even exist, since by the time the protons move appreciably, the positrons are long gone away to a very far distance. Therefore, energy conservation again tells us that the final speed v_2 of the protons obeys

$$\frac{kq^2}{\sqrt{2}a} \approx 2\left(\frac{1}{2}Mv_2^2\right), \quad v_2 = \sqrt{\frac{kq^2}{\sqrt{2}\,aM}}.$$

[3] **Problem 2** (PPP 114)**.** A small positively charged ball of mass m is suspended by an insulating thread of negligible mass. Another positively charged small ball is moved very slowly from a large distance until it is in the original position of the first ball. As a result, the first ball rises by h.



How much work has been done?

Solution. Let r be the final separation of the balls, and let L be the length of the string. By basic trigonometry,

$$\frac{h}{r} = \sin\theta$$

where θ is half the angle of the string to the vertical. Letting the balls have charges q and Q and balancing forces, we have

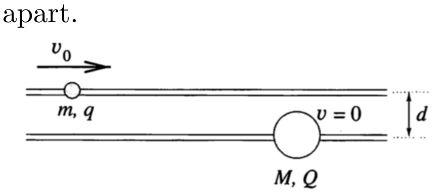
$$\frac{kqQ}{r^2}\cos\theta = mg\sin 2\theta = 2mg\sin\theta\cos\theta$$

from which we conclude

$$\frac{kqQ}{r} = 2mgr\sin\theta = 2mgh.$$

Furthermore, one of the balls has been raised by a height h during the process. Thus, the total work done is $\boxed{3mgh}$. It's neat how almost all the dimensionful quantities drop out in the final answer!

[3] **Problem 3** (PPP 71)**.** Two small beads slide without friction, one on each of two long horizontal parallel fixed rods a distance d apart.



The masses of the beads are m and M and they carry charges q and Q. Initially, the larger mass M is at rest and the other one is far away approaching it at a speed v_0. For what values of v_0 does the smaller bead ever get to the right of the larger bead?

Solution. When v_0 is just large enough for the small bead to get to the right of the big bead, when both beads end up side-to-side, the small bead's velocity should be just a bit greater than that of the big bead for it to get past. This means the minimum possible value v_m of v_0 should be just large enough to provide enough energy so that both beads can move together at some velocity v,

$$\frac{1}{2}mv_m^2 = \frac{kqQ}{d} + \frac{1}{2}(m+M)v^2.$$

Since the total horizontal momentum is conserved,

$$mv_m = (m+M)v.$$

Thus, we have

$$\frac{1}{2}\left(m - \frac{m^2}{m+M}\right)v_m^2 = \frac{kqQ}{d}, \quad v_m = \sqrt{\frac{2kqQ}{d}\frac{m+M}{mM}}.$$

[2] **Problem 4** (PPP 192)**.** Classically, a conductor is made of nuclei of positive charge fixed in place, and electrons that are free to move.

(a) Consider a solid conductor in a gravitational field $\mathbf{g}$. Argue that the electric field inside the conductor is *not* zero; find out what it is.

(b) Now suppose a positron is placed at the center of a hollow spherical conductor in a gravitational field $\mathbf{g}$. Find its initial acceleration.

Solution. (a) We usually argue that the electric field has to vanish to keep the electrons from accelerating. In this case, the electric field has to be nonzero, because otherwise the electrons will fall down. Specifically, there is a downward electric field of magnitude mg/e, where $e > 0$ is the magnitude of the electron charge and m is the electron mass. This comes out to about 6×10^{-11} V/m, which is quite small.

You might wonder how the forces on the positive ions are balanced, since they experience both downward gravitational and electrical forces. The answer is that they're locked into a lattice, and held up by internal stresses within the lattice. These ultimately come from whatever is keeping the conductor as a whole from falling, such as a normal force from the ground.

(b) The electric field found in part (a) also exists in a cavity, which one can argue by uniqueness or by the fact that the electric field is conservative. So the positron has an initial downward acceleration of $2g$. (We had to specify the positron was at the center, or else it would have an additional acceleration due to charge induction, which we could compute using image charges.)

[3] **Problem 5.** USAPhO 2008, problem B2. You may ignore part (c), which was removed in the final version of the exam, though you can also do it for extra practice.

[3] **Problem 6.** USAPhO 2019, problem B1.

[5] **Problem 7.** IPhO 2004, problem 1. A nice question on the dynamics of a multi-part system.

2 The Lorentz Force

Idea 2

Some questions below will involve special relativity. The Lorentz force law as written in idea 1 is still valid as long as $\mathbf{F}$ is interpreted as $d\mathbf{p}/dt$, where the relativistic momentum is

$$\mathbf{p} = \gamma m \mathbf{v}, \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}.$$

The relativistic energy is also modified to

$$E = \gamma mc^2 = mc^2 + \frac{1}{2}mv^2 + \ldots.$$

We will return to this subject in more detail in **R2**, but for now this is all you need.

Example 3: Kalda 163

A beam of electrons, of mass m and charge q, is emitted with a speed v almost parallel to a uniform magnetic field $\mathbf{B}$. The initial velocities of the electrons have an angular spread of $\alpha \ll 1$, but after a distance L the electrons converge again. Neglecting the interaction between the electrons, what is L?

Solution

Consider an electron initially traveling at an angle α to the magnetic field. This electron has a speed $v_\parallel = v\cos\alpha \approx v$ parallel to the field, and a speed $v_\perp v \sin\alpha \approx v\alpha$ perpendicular to the field. The component $v_\parallel$ always stays the same, while $v_\perp$ rotates, so the electron spirals along the field lines.

The acceleration of the electron is

$$a_\perp = \frac{F}{m} = \frac{qv_\perp B}{m}.$$

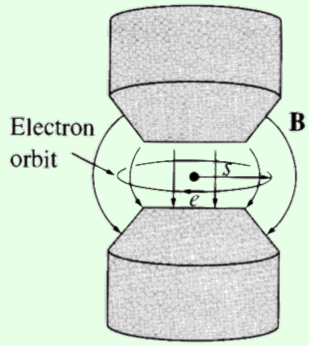
The perpendicular velocity component rotates through a circle in velocity space of circumference $2\pi v_\perp$. After one such circle, the total perpendicular displacement is zero, so the beam refocuses. Thus we have

$$L = \frac{2\pi v_\perp}{a} v_\parallel \approx \frac{2\pi mv}{qB}.$$

In other words, this setup acts like a magnetic "lens".

Example 4: Griffiths 7.50

In a "betatron", electrons move in circles in a magnetic field. When the magnetic field is slowly increased, the accompanying electric field will impart tangential acceleration.



Suppose the field always has the same spatial profile $B(r,t) = B_0(r) f(t)$. For what $B_0(r)$ is it possible for an electron to start at rest in zero magnetic field, and then move in a circle of constant radius as the field is increased?

Solution

The electrons experience a tangential force

$$\dot{p} = qE = q\frac{\dot{\Phi}_B}{2\pi r} = \frac{qr}{2}\dot{B}_{\text{av}}$$

where B_{av} is the average field over the orbit. Since the particles start from rest in zero field, we can integrate this to find

$$p = \frac{qr}{2}B_{\text{av}}.$$

On the other hand, the standard result for cyclotron motion is $p = qrB$, which means we must have $B = B_{\text{av}}/2$, i.e. the field at any radius is half the average magnetic field inside,

$$B(r) = \frac{1}{2}\frac{1}{\pi r^2}\int_0^r B(r')(2\pi r')\,dr'.$$

This rearranges slightly to give

$$r^2 B(r) = \int_0^r r' B(r')\,dr'.$$

Differentiating both sides with respect to r, we have

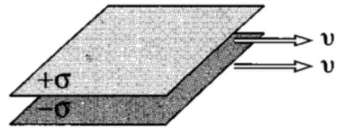
$$2rB(r) + r^2 B'(r) = rB(r)$$

which simplifies to

$$\frac{dB}{B} = -\frac{dr}{r}$$

which means the field profile should be $B_0(r) \propto 1/r$. (Of course, a real betatron might differ since it only needs to obey $B = B_{\text{av}}/2$ at the radii where electrons will be orbiting.)

[3] **Problem 8** (Griffiths 5.17). A large parallel plate capacitor with uniform surface charge σ on the upper plate and $-\sigma$ on the lower is moving with a constant speed v as shown.



(a) Find the magnetic field between the plates and also above and below them.

(b) Find the magnetic force per unit area on the upper plate, including its direction.

(c) What happens to the net force between the plates in the limit $v \to c$? Explain your result using some basic ideas from special relativity.

Solution. (a) Let $\hat{\mathbf{x}}$ be the direction of the velocity. Let $\hat{\mathbf{y}}$ point into the page, and let $\hat{\mathbf{z}}$ point up. The magnetic field due to the top plane is $-\frac{1}{2}\mu_0\sigma v\hat{\mathbf{y}}$ above the above plane and $\frac{1}{2}\mu_0\sigma v\hat{\mathbf{y}}$ below the above plane. Similarly for the lower plane, we have $\frac{1}{2}\mu_0\sigma v\hat{\mathbf{y}}$ above and $-\frac{1}{2}\mu_0\sigma v\hat{\mathbf{y}}$ below. Thus, the magnetic field is $\mu_0\sigma v\hat{\mathbf{y}}$ between the plates, and zero outside.

(b) The force per unit area (i.e. pressure) is $\sigma v\hat{\mathbf{x}} \times \frac{1}{2}\mu_0\sigma v\hat{\mathbf{y}} = \frac{1}{2}\mu_0\sigma^2 v^2\hat{\mathbf{z}}$. The factor of 1/2 is there since it only feels a force due to the contribution of the other plate; this is the essentially the same 1/2 as we found for the pressure on a conductor in **E1**.

(c) The forces balance when $v = c$,

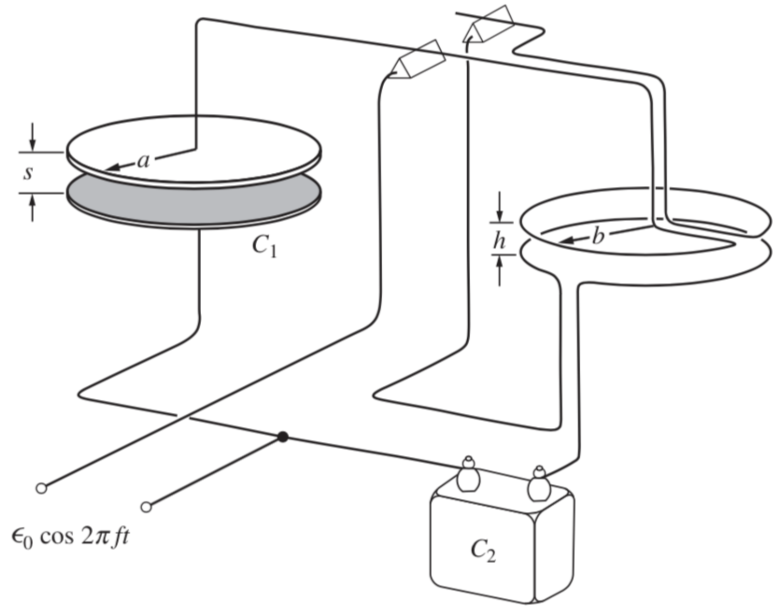
$$\frac{1}{2}\mu_0\sigma^2 c^2 = \frac{1}{2\epsilon_0}\sigma^2$$

because $c^2 = 1/\epsilon_0\mu_0$. Thus, as v increases to approach c, the attractive force between the plates gets smaller and smaller. If we invoke special relativity, this makes perfect sense. In the rest frame of the plates, there is only the attractive electrostatic force, so the plates move together. This implies that in the lab frame, the plates also have to move together, so the force must be attractive. But for high v, there's a lot of time dilation, so the plates move together more slowly. (This is partially due to the force decreasing, as derived here, and partially due to the higher "transverse mass" due to the plates' relativistic momentum, as we'll see in **R2**.)

[3] **Problem 9.** EFPhO 2012, problem 7. An elegant Lorentz force problem with wires. (If you enjoy this problem, consider looking at IdPhO 2020, problem 1B, which has a similar setup but requires three-dimensional reasoning. The official solutions are here.)

Solution. See the official solutions here.

[4] **Problem 10** (Purcell 6.35/INPhO 2008.6)**.** Consider the arrangement shown below.



The force between capacitor plates is balanced against the force between parallel wires carrying current in the same direction. A voltage alternating sinusoidally with angular frequency ω is applied to the parallel-plate capacitor C_1 and also to the capacitor C_2, and the current is equal to the current through the rings. Assume that $s \ll a$ and $h \ll b$.

Suppose the weights of both sides are adjusted to balance without any applied voltage, and C_2 is adjusted so that the time-averaged downward forces on both sides are equal. Show that

$$\frac{1}{\sqrt{\mu_0 \epsilon_0}} = \sqrt{2\pi}\, a\omega \sqrt{\frac{b}{h}\frac{C_2}{C_1}}.$$

The left-hand side is equal to c, as we'll show in **E7**, so this setup measures the speed of light.

Solution. It can be a little tricky to read the diagram. The key point is that the triangles are conductors. They represent the fulcrum of a see-saw, but they also allow the voltage to be applied across the capacitors on the left and right. The charge buildup on the capacitors on the left causes them to attract, while the current flowing through the circular wires on the right causes them to attract as well.

Let a current I flow on in the right-hand side. Since $h \ll b$, the magnetic field created by the bottom circular loop at a point on the top circular loop is approximately the same as that created by an infinite wire, $B = \mu_0 I/2\pi h$. Thus, the force between the wires is

$$F = (2\pi b I)\frac{\mu_0 I}{2\pi h} = \frac{\mu_0 b I^2}{h}.$$

This force oscillates over time. The charge on the capacitor C_2 is

$$Q_2(t) = C_2 \mathcal{E}_0 \cos(\omega t)$$

so that

$$\langle I^2(t)\rangle = C_2^2 \mathcal{E}_0^2 \omega^2 \langle \sin^2(\omega t)\rangle = \frac{C_2^2 \mathcal{E}_0^2 \omega^2}{2}.$$

Thus, the average force on the right is

$$\langle F\rangle = \frac{\mu_0 C_2^2 \mathcal{E}_0^2 \omega^2 b}{2h}.$$

On the left-hand side, the force between the plates is, by a result in **E1**,

$$F = \frac{1}{2}\epsilon_0 E^2 (\pi a^2)$$

where E is the electric field inside the plates, and we have

$$\langle E^2(t)\rangle = \frac{1}{s^2}\langle \mathcal{E}^2(t)\rangle = \frac{\mathcal{E}_0^2}{2s^2}.$$

Combining these results and eliminating s, since it doesn't appear in the final result,

$$\langle F\rangle = \frac{C_1^2 \mathcal{E}_0^2}{4\pi a^2 \epsilon_0}.$$

Equating the averaged forces gives

$$\frac{\mu_0 C_2^2 \omega^2 b}{h} = \frac{C_1^2}{2\pi a^2 \epsilon_0},$$

which is equivalent to the desired result.

[3] **Problem 11.** An electron beam is accelerated from rest by applying an electric field E for a time t, and subsequently guided by magnetic fields. These magnetic fields are produced with a series of coils, which carry currents I_i.

Now suppose the apparatus is repurposed to shoot proton beams. Suppose a proton beam is accelerated from rest by applying an electric field E for a time t (in the opposite direction). Let an electron have mass m and a proton have mass M.

(a) Find the currents I_i needed so that the proton follows the same trajectory the electron did, assuming V is small enough that both the electron and proton are nonrelativistic.

(b) How does the answer change if relativistic corrections are accounted for?

Solution. (a) The electron and proton have the same momentum p, and we have

$$\frac{dp}{dt} = qvB \sim qvI$$

since $B \propto I$. Now, the magnetic field can only rotate the particle's momentum. Suppose at some moment it is curving in a trajectory with radius of curvature r, and speed v. Then it has instantaneous angular velocity $\omega = v/r$ along the circle tangent to its trajectory, and

$$\frac{dp}{dt} = \omega p$$

in magnitude. Hence we have

$$qvI \sim \frac{v}{r}p$$

and since r is the same for both the electron and proton, we suppress it to give

$$I \sim \frac{p}{q}.$$

In other words, we have $I \propto 1/q$, so all that happens is that the currents should change sign to accommodate the proton.

(b) Every step in the solution to part (a) still works with relativity accounted for (the change of $p = mv$ to $p = \gamma mv$ doesn't matter, because we never used $p = mv$), so the answer is the same: we just flip the currents.

[5] **Problem 12.** IPhO 2000, problem 2. A solid question on the Lorentz force with real-world relevance. Requires a little relativity, namely the expressions for relativistic momentum/energy.

Solution. The official files are a mess; the solutions to this particular problem are here and here.

[4] **Problem 13.** IPhO 1996, problem 2. An elegant problem on particles in a magnetic field. (There's a deeper principle behind the solution to this problem; see **R3** for more discussion.)

3 Magnetic Moments

[3] **Problem 14.** Consider a current loop I in the xy plane in a constant magnetic field $\mathbf{B}$.

(a) Show that the net force on the loop is zero.

(b) Show that the torque is

$$\boldsymbol{\tau} = \mathbf{m} \times \mathbf{B}$$

where the magnetic moment is

$$\mathbf{m} = IA\hat{\mathbf{z}}$$

where A is the area of the loop. For simplicity, you can show this in the case where the current loop is a square of side length L, whose sides are aligned with the x and y axes. (The proof for a general loop shape requires some vector calculus, but you can attempt it for a challenge. You'll need the double cross product identity, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = 0$.)

Solution. (a) We see that

$$\mathbf{F} = I \oint d\mathbf{s} \times \mathbf{B} = I \left(\oint d\mathbf{s} \right) \times \mathbf{B} = 0,$$

as desired.

(b) The magnetic moment of the square is

$$\mathbf{m} = IL^2\hat{\mathbf{z}}.$$

The torque on a side of the square is

$$\boldsymbol{\tau} = \int \mathbf{r} \times d\mathbf{F} = I \int \mathbf{s} \times (d\mathbf{s} \times \mathbf{B}).$$

In particular, it's useful to pair the two sides parallel to the x axis. These have opposite currents and differ only by a translation $\Delta\mathbf{r} = L\hat{\mathbf{y}}$, so adding their contributions gives a torque

$$\boldsymbol{\tau} = -I \int_0^L (L\hat{\mathbf{y}}) \times (\hat{\mathbf{x}}\, dx \times \mathbf{B}) = -IL(\hat{\mathbf{y}} \times (\hat{\mathbf{x}} \times \mathbf{B})) \int_0^L dx = -IL^2(\hat{\mathbf{y}} \times (\hat{\mathbf{x}} \times \mathbf{B})).$$

Similarly, the torques due to the other two sides add up to

$$\boldsymbol{\tau} = IL^2(\hat{\mathbf{x}} \times (\hat{\mathbf{y}} \times \mathbf{B})).$$

Manually performing the cross products, we have

$$-\hat{\mathbf{y}} \times (\hat{\mathbf{x}} \times \mathbf{B}) = -B_y \hat{\mathbf{x}}, \quad \hat{\mathbf{x}} \times (\hat{\mathbf{y}} \times \mathbf{B}) = B_x \hat{\mathbf{y}}.$$

Adding these together gives exactly the desired result, $\boldsymbol{\tau} = \mathbf{m} \times \mathbf{B}$.

For completeness, we display a fully general, vector calculus solution, valid for any loop shape. We note that along the full, closed loop, the fundamental theorem of calculus implies

$$\oint d(\mathbf{s} \times (\mathbf{s} \times \mathbf{B})) = 0$$

simply because the closed loop integral of d(anything) is the net change in (anything) along the loop, which is zero. Expanding with the product rule gives

$$\oint d\mathbf{s} \times (\mathbf{s} \times \mathbf{B}) + \mathbf{s} \times (d\mathbf{s} \times \mathbf{B}) = 0.$$

Using these results and the double cross product identity, the torque is

$$\begin{aligned} \boldsymbol{\tau} &= I \oint \mathbf{s} \times (d\mathbf{s} \times \mathbf{B}) \\ &= -I \oint d\mathbf{s} \times (\mathbf{B} \times \mathbf{s}) - I \oint \mathbf{B} \times (\mathbf{s} \times d\mathbf{s}) \\ &= -\boldsymbol{\tau} - I\mathbf{B} \times \left(\oint \mathbf{s} \times d\mathbf{s} \right). \end{aligned}$$

Now, $\mathbf{s} \times d\mathbf{s} = 2\, d\mathbf{A}$, because as $\mathbf{s}$ moves a little along the loop it sweeps out a small triangle of area. Thus we have $2\boldsymbol{\tau} = 2I\mathbf{B} \times \mathbf{A}$, giving the result.

Idea 3

The force on a small magnetic dipole $\mathbf{m}$ is

$$\mathbf{F} = \nabla(\mathbf{m} \cdot \mathbf{B}).$$

If there are no other currents at the dipole's location, so that $\nabla \times \mathbf{B} = 0$, this is equivalent to $\mathbf{F} = (\mathbf{m} \cdot \nabla)\mathbf{B}$, which is sometimes easier to evaluate.

As in problem 14, this can be shown relatively easily for a square loop, and requires some tricky vector calculus for a general current distribution. Both the force and torque on a magentic dipole can be found by differentiating the potential energy

$$U = -\mathbf{m} \cdot \mathbf{B}.$$

All of these results also hold for electric dipoles, if we replace $\mathbf{m}$ with $\mathbf{p}$ and $\mathbf{B}$ with $\mathbf{E}$.

Remark

The expression for the potential energy above is notoriously subtle. Here's the problem: we know the Lorentz force on a charge is $q\mathbf{v} \times \mathbf{B}$, which means magnetic fields never do work. So how can they be associated with a nonzero potential energy?

There are two levels of explanation. First, suppose the magnetic dipole is made of charges moving in a loop. When such a current loop is placed in a magnetic field, and moved or rotated, mechanical work can be done on the loop. But at the same time, there will be an induced emf in the loop, which speeds up or slows down the current. The work done by these two effects perfectly cancels, so that the energy of the loop stays constant. For this kind of dipole, the expression for U doesn't indicate the total energy, but only the "mechanical" potential energy, in the sense that differentiating it gives the right forces and torques. (Some further discussion of this point is in chapter II-15 of the Feynman lectures.)

On the other hand, the magnetic dipole moment of a common bar magnet doesn't come from charges moving in a loop! Instead, it comes from the intrinsic magnetic dipole moments of the unpaired electrons in the magnet. These kinds of dipole moments aren't composed of any moving subcomponents; they are an elementary and immutable property of the electron, like its mass or charge. In these cases, $U = -\mathbf{m} \cdot \mathbf{B}$ really is the total energy, and the magnetic field *can* do work. You won't hear much about these elementary dipole moments in introductory books, because they can only be properly understood by combining relativity and quantum mechanics, but they're responsible for most magnetic phenomena.

Example 5

If a magnet is held over a table, it can pick up a paper clip. If the paper clip is removed, it can pick up another paper clip just as well, and this process can seemingly continue forever without any effect on the magnet. Since the magnet does work on each paper clip, doesn't this mean a permanent magnet is an infinite energy source?

Solution

This is the kind of question that makes magnets feel so mysterious. They're basically the only everyday example of a long range force besides gravity (in fact, Kepler once thought the Sun acted on the planets like a giant magnet), and as such they've inspired countless attempts at perpetual motion machines. For centuries, many people have spent years of their lives trying to get elaborations of this example to work.

To see why this doesn't work for a bar magnet, just replace the word "magnet" with "charge". It's true that a positive charge can attract a negative charge to it. And if the negative charge is then removed, the positive charge can then attract another negative charge to it. But conservation of energy isn't violated, because the force from the positive charge is conservative: the work it does on the negative charge to draw it close is precisely the opposite of the work an external agent needs to do to pull it away. The force of a magnet on a paper clip is also conservative.

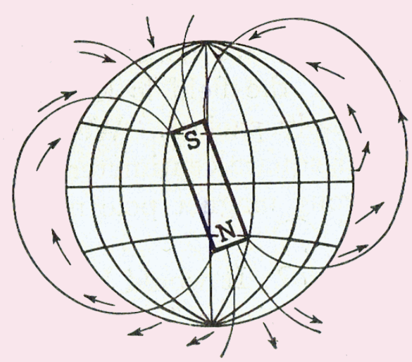
It's also interesting to consider a slightly different case. Unlike a bar magnet, an electromagnet (i.e. a magnet created by moving current in a loop) can be turned on and off with the flick of a switch. Therefore, we might suspect that the following is a perpetual motion machine:

1. Turn on the electromagnet, which costs energy E_0.
2. Use it to lift a paper clip, increasing its potential energy by mgh.
3. Turn off the electromagnet, which costs energy E_0, while holding the paper clip.
4. Move the paper clip away; we've managed to raise it higher for free.

To see the problem, note that the attractive force between the magnet and paper clip arises because the magnet induces a magnetic dipole moment in the paper clip, leading to a $(\mathbf{m}\cdot\nabla)\mathbf{B}$ force. As the paper clip moves toward the magnet, its own dipole moment causes a changing magnetic flux through the electromagnet, and thus an emf against the current. Therefore, it costs extra energy to keep the current in the electromagnet steady. Since the $q\mathbf{v}\times\mathbf{B}$ Lorentz force doesn't do work, that energy must be precisely mgh, so nothing comes for free.

Remark

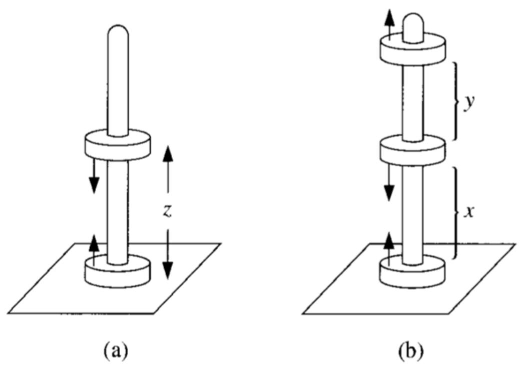
A compass needle is essentially a small magnetic dipole, whose dipole moment points towards the end painted red. We can also approximate the Earth's magnetic field as a dipole field.



Since the tangential component of this dipole field points north, the red end of the compass points towards the geographic north pole, which is the Earth's magnetic south pole.

By the way, a cheap compass calibrated to work in America or Europe won't work well in Australia. The reason is that the Earth's magnetic field also has a radial component, which acts to tip the compass needle up or down. The needle needs to be appropriately weighted to stay horizontal, so that it can freely rotate, but the side that needs to be weighted differs between the hemispheres.

[3] **Problem 15** (Griffiths 6.23). A familiar toy consists of donut-shaped permanent magnets which slide frictionlessly on a vertical rod.



Treat the magnets as dipoles with mass m_d and dipole moment $\mathbf{m}$, with directions as shown above.

(a) If you put two back-to-back magnets on the rod, the upper one will "float". At what height z does it float?

(b) If you now add a third magnet parallel to the bottom one as shown, find the ratio x/y of the two heights, using only a scientific calculator. (Answer: 0.85.)

Solution. (a) We know that the field from a magnetic dipole is

$$\mathbf{B} = \frac{\mu_0 m}{4\pi r^3}\left(2\cos\theta\,\hat{\mathbf{r}} + \sin\theta\,\hat{\boldsymbol{\theta}}\right).$$

Along the z-axis, this reduces to

$$B_z = \frac{\mu_0 m}{2\pi z^3}.$$

The force on the upper magnet must balance gravity, so

$$-\frac{\mu_0 m^2}{2\pi}\frac{d}{dz}\left(\frac{1}{z^3}\right) - m_d g = 0$$

which yields

$$z = \left(\frac{3\mu_0 m^2}{2\pi m_d g}\right)^{1/4}.$$

(b) The net force on the middle magnet comes from the field from the top and bottom magnets, along with gravity,

$$\frac{3\mu_0 m^2}{2\pi}\left(\frac{1}{x^4} - \frac{1}{y^4}\right) = m_d g.$$

Similarly, the top magnet, experiences forces from the bottom and middle magnets,

$$\frac{3\mu_0 m^2}{2\pi}\left(\frac{1}{y^4} - \frac{1}{(y+x)^4}\right) = m_d g.$$

Putting these two equations together yields

$$\frac{1}{x^4} - \frac{1}{y^4} = \frac{1}{y^4} - \frac{1}{(y+x)^4}.$$

Defining $\alpha = x/y$, we then need to solve

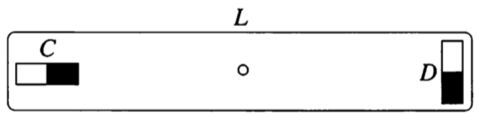
$$\alpha = \left(\frac{(1+\alpha)^4}{2(1+\alpha)^4 - 1}\right)^{1/4}.$$

We solve this using iteration, as introduced in **P1**. That is, we guess a reasonable value like $\alpha = 0.5$, then repeatedly plug in

$$\left(\frac{(1+\text{Ans})^4}{2(1+\text{Ans})^4 - 1}\right)^{1/4}$$

which yields $x/y = 0.85$.

[3] **Problem 16** (PPP 89)**.** Two identical small bar magnets are placed on opposite ends of a rod of length L as shown.



(a) Show that the torques the magnets exert on each other are *not* equal and opposite.

(b) Suppose the rod is pivoted at its center, and the magnets are attached to the rod so that they can spin about their centers. If the magnets are released, the result of part (a) implies that they will begin spinning. Explain how this can be consistent with energy and angular momentum conservation, treating the latter quantitatively.

Solution. (a) Referring to the dipole fields computed in **E1**, the field at D due to C is twice that at C due to D, so they can't possibly cancel. Worse, the directions of the torques are the same (both out of the page).

(b) Energy is conserved because there is an energy density $B^2/2\mu_0$ stored in the magnetic field of the two magnets. As the rotational kinetic energy of the system increases, the energy stored in the field decreases to compensate.

Angular momentum conservation holds for a different reason. While electromagnetic fields can store angular momentum too, they don't in this particular case. Instead, the answer is something more familiar. The magnets also exert forces on each other, so a force from the rod is necessary to keep the magnets in place. This implies the magnets exert a torque on the rod, which begins spinning in the opposite direction. Thus, angular momentum is conserved.

To show this quantitatively, set up coordinates with the origin at the center of the rod, and the z-axis pointing out of the page. The total torque on the two magnets is

$$\boldsymbol{\tau}_0 = \frac{3\mu_0}{4\pi}\frac{\mu^2}{L^3}\,\hat{\mathbf{z}}$$

where μ is the magnetic moment of each magnet. This is the rate of change of their spin angular momentum. Next, we consider forces. The force on magnet C due to D is

$$\mathbf{F}_{CD} = \mu\partial_x \mathbf{B} = \frac{\mu_0}{4\pi}\,\mu^2 \frac{d}{dx}\frac{1}{x^3}\bigg|_{x=L} = \frac{3\mu_0}{4\pi}\frac{\mu^2}{L^4}\,\hat{\mathbf{y}}.$$

This produces a torque on the rod, about its pivot point, of

$$\boldsymbol{\tau}_1 = -\frac{1}{2}\frac{3\mu_0}{4\pi}\frac{\mu^2}{L^3}\,\hat{\mathbf{z}}.$$

The force on magnet D due to magnet C is equal and opposite, and therefore provides an equal torque $\boldsymbol{\tau}_2$ on the rod. Therefore, the total rate of change of angular momentum is $\boldsymbol{\tau}_0 + \boldsymbol{\tau}_1 + \boldsymbol{\tau}_2 = 0$.

In introductory textbooks, you might have read that angular momentum is conserved as a consequence of the strong form of Newton's third law, which is that forces are equal and opposite, and always directed along the line separating two particles. As we've just seen, this isn't actually necessary: here we have an example of a force which isn't directed along the separation, but angular momentum is still conserved. In **E7** we'll see even more exotic examples, where even the weak form of Newton's third law (i.e. that forces are equal and opposite) breaks down, but momentum remains conserved anyway, as a consequence of electromagnetic fields carrying away the excess momentum. Generally speaking, the deeper you get into physics, the less important Newton's laws become, while conservation laws remain as important as ever.

4 Point Charges

In this section we'll give a sampling of classic problems involving just point charges in fields; these will be a bit more mathematically advanced than the others in this problem set.

[3] **Problem 17.** A point charge q of mass m is released from rest a distance d from a grounded conducting plane. Find the time until the point charge hits the plane. (Hint: use Kepler's laws.)

Solution. This is an incredibly classic problem, which has been appearing in various forms on competitions for decades. By using image charges, we see that the particle always experiences a force

$$F = \frac{kq^2}{4z^2}$$

directly towards the plane, where z is its separation from the plane. Let the particle impact the plane at point O.

This force has the form of an inverse-square law. In particular, we would get the exact same result if the force were always directed towards O (rather than always directed towards the plane),

$$\mathbf{F} = -\frac{kq^2}{4r^2}\hat{\mathbf{r}}.$$

But in this case, the problem is perfectly analogous to the central force of gravity, where O serves as the location of the Sun, and one of the foci of the charge's orbit. In particular, releasing the charge from near rest and waiting for it to hit the plane corresponds to performing the first half of an extremely eccentric elliptic orbit.

The trick is now to use Kepler's third law. If the charge had performed a circular orbit of radius d about O, then

$$\frac{kq^2}{4d^2} = \frac{mv^2}{d} = m\omega^2 d$$

which gives a period of

$$T = \frac{2\pi}{\omega} = 4\pi\sqrt{\frac{md^3}{kq^2}}.$$

We can use Kepler's third law to find the period of the eccentric elliptic orbit the charge actually follows. This orbit has semimajor axis $d/2$, so it has period

$$T' = \frac{T}{2\sqrt{2}}.$$

The actual path of the charge is only the first half of this orbit, so the answer is

$$\frac{T'}{2} = \frac{T}{4\sqrt{2}} = \frac{\pi}{\sqrt{2}}\sqrt{\frac{md^3}{kq^2}}.$$

Of course, the problem can also be solved by directly integrating the differential equation. If you do it that way, you'll get the same integral as in a similar example, given in **P1**.

[3] **Problem 18.** A point charge of mass m and charge q is released from rest at the origin in the fields $\mathbf{E} = E\hat{\mathbf{x}}$, $\mathbf{B} = B\hat{\mathbf{y}}$. Find its position as a function of time by solving the differential equations given by Newton's second law, $\mathbf{F} = m\mathbf{a}$.

Solution. We will assume non-relativistic motion throughout. Note that the motion is solely in the xz plane, since the electric and magnetic forces are in that plane. Newton's second law gives

$$\ddot{x} = \frac{q}{m}(E_0 - B_0\dot{z}),$$
$$\ddot{z} = \frac{q}{m}B_0\dot{x}.$$

Taking the time derivative of the first equation and plugging in into the second, we find

$$\dddot{x} = -\frac{q^2B_0^2}{m^2}\dot{x},$$

and along with the initial condition that $\dot{x}(0) = 0$, we see that

$$\dot{x} = v_0\sin(\omega t)$$

where v_0 is some yet to be determined velocity, and $\omega \equiv qB_0/m$. Integrating, and using the initial condition that $x(0) = 0$, we see that

$$x(t) = \frac{v_0}{\omega}(1 - \cos(\omega t)).$$

We also have that

$$\ddot{z} = \omega\dot{x} = \omega v_0\sin(\omega t).$$

Integrating twice and using the fact that $z(0) = \dot{z}(0) = 0$, we see that

$$z(t) = v_0 t - \frac{v_0}{\omega}\sin(\omega t).$$

All that is to be found now is v_0. Plugging our x and z into the first equation, we see that

$$v_0\omega\cos(\omega t) = \frac{q}{m}(E_0 - B_0v_0(1 - \cos(\omega t))) \implies v_0 = E_0/B_0.$$

Thus, our final solution is

$$x(t) = \frac{v_0}{\omega}(1 - \cos(\omega t)),$$
$$z(t) = v_0 t - \frac{v_0}{\omega}\sin(\omega t)$$

where $v_0 = E_0/B_0$ and $\omega = qB_0/m$.

Notice that while naively one might have thought the motion would be along $\mathbf{E}$, on average the particle actually moves along $\mathbf{E} \times \mathbf{B}$. This is actually quite general. For example, it remains true even if there's a bit of friction; the steady state velocity turns out to be along $\mathbf{E} \times \mathbf{B}$. Another example is how weather systems work. When I was a kid, I was always confused about how entire regions could have low or high pressure; would the wind just go along the pressure gradient to even it out? That doesn't happen because the Coriolis force deflects the wind sideways. In this case, the pressure gradient is acting like $\mathbf{E}$, and the Coriolis force behaves like a magnetic field $\mathbf{B} \parallel \boldsymbol{\omega} \parallel \hat{\mathbf{z}}$. The net effect is that in the steady state, wind tends to move *along* lines of constant pressure, not perpendicular to them. So a low pressure system stays low pressure but spins around.

[3] **Problem 19** (Wang)**.** Two identical particles of mass m and charge q are placed in the xy plane with a uniform magnetic field $B\hat{\mathbf{z}}$. The particles have paths $\mathbf{r}_1(t)$ and $\mathbf{r}_2(t)$. Neglect relativistic effects, but account for the interaction between the charges.

(a) Write down a differential equation describing the evolution of the separation $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$.

(b) Suppose that the initial conditions have been set up so that the particles orbit each other in a circle in the xy plane, with constant separation d. What is the smallest d for which this motion is possible?

Solution. (a) The equations of motion for the two particles are

$$m\ddot{\mathbf{r}}_1 = \frac{q^2}{4\pi\epsilon_0 r^3}\hat{\mathbf{r}} + q\dot{\mathbf{r}}_1 \times \mathbf{B}, \quad m\ddot{\mathbf{r}}_2 = -\frac{q^2}{4\pi\epsilon_0 r^3}\hat{\mathbf{r}} + q\dot{\mathbf{r}}_2 \times \mathbf{B}.$$

Subtracting the two, we have

$$m\ddot{\mathbf{r}} = \frac{q^2}{2\pi\epsilon_0 r^3}\mathbf{r} + q\dot{\mathbf{r}} \times \mathbf{B}.$$

(b) Note that since $\mathbf{B}$ is along the $\hat{\mathbf{z}}$ direction, and $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$ where $\boldsymbol{\omega}$ is also along the $\hat{\mathbf{z}}$ direction, all three vector terms in the above equation are parallel. So we have

$$\left(\frac{q^2}{2\pi\epsilon_0 r^3} + q\omega B + m\omega^2\right)\mathbf{r} = 0$$

and setting the term in parentheses to zero, we get

$$\omega = \frac{-(qB/m) \pm \sqrt{(qB/m)^2 - 2q^2/\pi m\epsilon_0 d^3}}{2}$$

where $\omega_c = qB/m$ is the usual cyclotron angular frequency. For this equation to have a solution, the discriminant must be nonnegative, so

$$\frac{q^2B^2}{m^2} \geq \frac{2q^2}{\pi m\epsilon_0 d^3}$$

which gives

$$d \geq \left(\frac{2m}{\pi\epsilon_0 B^2}\right)^{1/3}.$$

For smaller d, the charges will always fly apart, either due to electrostatic repulsion if they're slow, or the angular momentum barrier if they're fast.

[4] **Problem 20.** [A] Consider a point charge of mass m and charge q in the field of a magnetic monopole at the origin,

$$\mathbf{B} = \frac{g}{r^2}\hat{\mathbf{r}}.$$

In this problem we'll investigate the strange motion that results.

(a) Argue that the speed v is constant.

(b) Show that the angular momentum $\mathbf{L}$ of the charge is *not* conserved, but that

$$\mathbf{V} = \mathbf{L} - qg\hat{\mathbf{r}}$$

is. The second term is the angular momentum stored in the fields of the charge and monopole.

(c) Show that the charge moves on the surface of a cone. (Hint: in spherical coordinates where the z-axis is parallel to $\mathbf{V}$, consider $\mathbf{V} \cdot \hat{\boldsymbol{\phi}}$.) Sketch some typical trajectories.

Solution. (a) The force is $q\mathbf{v} \times \mathbf{B} \perp \mathbf{v}$, so no work is done on the particle, so its speed remains the same.

(b) Note that

$$\dot{\hat{\mathbf{r}}} = \frac{d}{dt}\frac{\mathbf{r}}{r} = \frac{\dot{\mathbf{r}}r - \dot{r}\mathbf{r}}{r^2} = \frac{\dot{\mathbf{r}}r - \mathbf{r}\frac{\mathbf{r}\cdot\dot{\mathbf{r}}}{r}}{r^2} = (\hat{\mathbf{r}}\cdot\hat{\mathbf{r}})\frac{\dot{\mathbf{r}}}{r} - \left(\hat{\mathbf{r}}\cdot\frac{\dot{\mathbf{r}}}{r}\right)\hat{\mathbf{r}} = \hat{\mathbf{r}} \times \left(\frac{\dot{\mathbf{r}}}{r} \times \hat{\mathbf{r}}\right) = \frac{1}{r^2}\mathbf{r} \times (\dot{\mathbf{r}} \times \hat{\mathbf{r}}).$$

We have $\dot{\mathbf{L}} = \boldsymbol{\tau} = \mathbf{r} \times (q\dot{\mathbf{r}} \times (g/r^2)\hat{\mathbf{r}}) = qg\dot{\hat{\mathbf{r}}}$, so $\mathbf{L} - qg\hat{\mathbf{r}}$ is conserved.

(c) Take coordinates so that $\mathbf{V}$ is directed along $\hat{\mathbf{z}}$ and the particle is instantaneously in the xz plane. Now take the y-component of the above equation, to give $L_y = 0$. In components, this tells us that $xp_z - zp_x = 0$, or in other words that $\dot{x}/\dot{z} = x/z$. By drawing similar triangles, this implies that the particle is momentarily moving so that x/z is conserved. By repeating this argument at all times, r/z is conserved, where r is the distance to the z-axis. This defines a cone.

In a typical trajectory, the charge spirals in towards the monopole along this cone, reaches some minimum distance from it, then turns around and spirals out. In fact, it turns out that if the cone is "cut and unfolded" and laid flat, the trajectory is a straight line! In other words, it is a geodesic on the cone.

One can do problem 18 slickly using field transformations, an advanced subject we will cover in **R3**.

Idea 4: Field Transformations

If the electromagnetic field is $(\mathbf{E}, \mathbf{B})$ in one reference frame, then in a reference frame moving with velocity $\mathbf{v}$ with respect to this frame, the components of the field parallel to $\mathbf{v}$ are

$$E'_\parallel = E_\parallel, \quad B'_\parallel = B_\parallel$$

while the components perpendicular are

$$\mathbf{E}'_\perp = \gamma(\mathbf{E}_\perp + \mathbf{v} \times \mathbf{B}), \quad \mathbf{B}'_\perp = \gamma\left(\mathbf{B}_\perp - \frac{\mathbf{v}}{c^2} \times \mathbf{E}\right).$$

Remark

The nonrelativistic limit of the field transformation is useful, but one has to be careful in deriving it. You might think, what's the need for care? Can't we just send $c \to \infty$, Taylor expand the above expressions, and call it a day? The problem with this reasoning is that there's no such thing as setting $c \to \infty$. You can't change a fundamental constant, and moreover this statement isn't even dimensionally correct, as noted in **P1**. What we really mean by the nonrelativistic limit is restricting our attention to some subset of possible situations, within which relativistic effects don't matter.

For example, if we have a bunch of point charges with typical speed v, then the nonrelativistic limit is considering only situations where v/c is small. In other words, we are taking $v/c \to 0$, not $c \to \infty$. Since the magnetic field of a point charge is v/c^2 times the electric field, the magnetic field ends up small. Now if we also consider boosts with small speeds v, then expanding the field transformations to lowest order in v/c gives

$$\mathbf{E}' = \mathbf{E}, \quad \mathbf{B}' = \mathbf{B} - \frac{\mathbf{v}}{c^2} \times \mathbf{E}.$$

This is the nonrelativistic limit for situations where $E/B \gg c$, also called the electric limit.

However, there's another possibility. Suppose that we have a bunch of neutral wires. In this case, it's the electric fields that are small, $E/B \ll c$. Using this in the transformations above, we arrive at the distinct result

$$\mathbf{B}' = \mathbf{B}, \quad \mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}$$

which apply for situations where $E/B \ll c$, also called the magnetic limit.

You might think we could improve the approximation by combining the two,

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}, \quad \mathbf{B}' = \mathbf{B} - \frac{\mathbf{v}}{c^2} \times \mathbf{E}$$

but this isn't self-consistent. For example, if you apply a Galilean boost with speed v, and then a boost with speed $-v$, you don't get back the same fields you started with! A sensible Galilean limit is only possible if $E/B \gg c$ or $E/B \ll c$, which are called the electric and magnetic limits, discussed further in this classic paper. It's only in relativity that E and B can be treated on an equal footing.

[3] **Problem 21.** Using the Galilean field transformations to solve problem 18.

(a) In the magnetic limit, show that the Lorentz force stays the same between frames, as it should. Then use the field transformations to find an appropriate reference frame where the problem becomes easy.

(b) In the electric limit, show that the Lorentz force stays the same up to terms that are order $(v/c)^2$ smaller, assuming $B/E \sim v/c^2$. (This is fine, since we're taking the limit $v/c \to 0$ anyway.) Then use the field transformations to find an appropriate reference frame where the problem becomes easy.

(c) The solutions you found in parts (a) and (b) should look very different, even though you should have found only one type of behavior in problem 18. In fact, there is a critical value of E/B separating the two kinds of behavior. What is this critical value, and why didn't you run into it when solving problem 18?

Solution. (a) Suppose a particle has velocity $\mathbf{u}$ in the original frame, so the force there is $\mathbf{F} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B})$. The force in the boosted frame is $\mathbf{F}' = q(\mathbf{E} + \mathbf{v} \times \mathbf{B} + (\mathbf{u} - \mathbf{v}) \times \mathbf{B}) = \mathbf{F}$.

We can find a frame where there's no electric field, by letting $\mathbf{E} + \mathbf{v} \times \mathbf{B} = 0$. In this frame, the particle just orbits in a circle. (Going back to the original frame just gives back the cycloid we found earlier.)

(b) We use the same setup as (a). The boosted force is

$$\mathbf{F}' = q(\mathbf{E} + (\mathbf{u} - \mathbf{v}) \times (\mathbf{B} - \mathbf{v} \times \mathbf{E}/c^2)) = \mathbf{F} + q\left(-\mathbf{v} \times \mathbf{B} + \frac{\mathbf{v} \times \mathbf{v} \times \mathbf{E}}{c^2} - \frac{\mathbf{u} \times \mathbf{v} \times \mathbf{E}}{c^2}\right).$$

The extra terms are all second order in v/c.

We can now find a frame where there's no magnetic field, by letting $\mathbf{B} - \mathbf{v} \times E/c^2 = 0$. In this frame, the particle just accelerates straight along $\mathbf{E}$. This indicates that in the original frame, the particle is always going along $\mathbf{E}$, while getting deflected a bit to the side by the magnetic field.

(c) As we found in problem 7, the characteristic velocity of the particle during the cycloid motion is $v_0 = E/B$. Hence in the electric limit $v_0 \gg c$, so our nonrelativistic solution must break down. Accounting for the full relativistic dynamics, it turns out it is harder to "turn around" a relativistic charge, so the particle is never turned around by the magnetic field.

To find the critical field value, we see the full relativistic field transformations allow us to remove $\mathbf{E}$ exactly when $E/B < c$, and to remove $\mathbf{B}$ exactly when $E/B > c$. Hence $E/B = c$ separates the two behaviors. Intuitively, as E/B increases up to c, the cycloid solution is stretched in the $\mathbf{E}$ direction more and more by relativistic effects, until when $E/B = c$ it is infinitely stretched.

5 Continuous Systems

Example 6: The Drude Model

Model a conductor as a set of electrons, of charge q, mass m, and number density n, which are completely free. Assume that in every small time interval dt, each electron has a probability dt/τ of hitting a lattice ion, which randomizes the direction of its velocity. Under these assumptions, compute the resistivity of the material.

Solution

First, suppose the electrons have some average momentum $\langle \mathbf{p} \rangle$ each. Because the collisions randomize the velocity, the average momentum falls exponentially with timescale τ,

$$\frac{d\langle \mathbf{p} \rangle}{dt} = -\frac{\langle \mathbf{p} \rangle}{\tau}.$$

On the other hand, if there is an applied field, a force term appears on the right,

$$\frac{d\langle \mathbf{p} \rangle}{dt} = -\frac{\langle \mathbf{p} \rangle}{\tau} + q\mathbf{E}$$

since $\mathbf{F} = d\mathbf{p}/dt$ for each individual electron. In the steady state,

$$\langle \mathbf{p} \rangle = q\mathbf{E}\tau.$$

The current density is

$$\mathbf{J} = nq\langle \mathbf{v} \rangle = \frac{nq\langle \mathbf{p} \rangle}{m} = \frac{nq^2\tau}{m}\mathbf{E}.$$

Thus, the resistivity in the Drude model is

$$\rho = \frac{m}{nq^2\tau}.$$

We can also compute the typical drift velocity,

$$v = \frac{qE\tau}{m}.$$

For values of m that give reasonable ρ, the value of v is a literal snail's pace, which is why people say that the electrons themselves move very slowly through a circuit. (Of course, a current can get started in a circuit much faster, because when a battery is attached, each moving electron pushes on the next one along the wire, and this wave of motion travels much faster than the electrons themselves.)

Remark: The Drude–Sommerfeld Model

Above we tacitly assumed there was a given probability of collision per unit time, but that's not right: when a particle flies through a medium, there is instead a given probability of collision per unit *length* it travels. These are equivalent for electrons moving at constant speed, but intuitively, we would expect electrons to have to accelerate starting from rest after each collision, in which case the two differ. To estimate this quickly, note that if the typical collision distance is ℓ, the kinetic energy picked up between collisions is $mv^2/2 \sim qE\ell$, giving typical speed $v \propto \sqrt{E}$. The analogue of Ohm's law would then be $I \propto \sqrt{V}$, completely contrary to observation!

The resolution is that electrons in solids really do effectively move with almost constant speed, even after collisions. This is a quantum mechanical effect, as explained in **X1**. The Pauli exclusion principle implies the electrons in the conductor have to occupy different quantum states, and the high density of electrons requires most of them to always have extremely high speeds, on the order of 1% of the speed of light! The drift velocity is merely the tiny amount by which their velocities are shifted on average.

[2] **Problem 22.** Consider Drude theory again, but now suppose there is also a fixed magnetic field $B\hat{\mathbf{z}}$. In this case, $\mathbf{J}$ is not necessarily parallel to $\mathbf{E}$, but the relation between the two can be described by the "tensor of resistivity". That is, the components are related by

$$E_i = \sum_{j \in \{x,y,z\}} \rho_{ij} J_j.$$

Calculate the coefficients ρ_{ij}. Express your answers in terms of the quantities

$$\rho_0 = \frac{m}{nq^2\tau}, \quad \omega_0 = \frac{qB}{m}$$

as well as the parameter τ.

Solution. The Lorentz force expression says

$$\frac{d\langle\mathbf{p}\rangle}{dt} = -\frac{\langle\mathbf{p}\rangle}{\tau} + q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

In the steady state, the left-hand side vanishes, so

$$\frac{\langle\mathbf{p}\rangle}{q\tau} = \mathbf{E} + \frac{1}{m}\langle\mathbf{p}\rangle \times \mathbf{B}.$$

Switching from $\langle\mathbf{p}\rangle$ to $\mathbf{J}$ and using the variables defined gives

$$\mathbf{E} = \rho_0\mathbf{J} - \rho_0\omega_0\tau\mathbf{J} \times \hat{\mathbf{z}}.$$

From this, we can directly read off the components of the resistivity,

$$\rho = \begin{pmatrix} \rho_0 & -\rho_0\omega_0\tau & \\ \rho_0\omega_0\tau & \rho_0 & \\ & & \rho_0 \end{pmatrix}.$$

When the electric field is in the $\hat{\mathbf{z}}$ direction, the magnetic field does nothing, which makes sense.

Example 7: Griffiths 5.40

Since parallel currents attract, the currents within a single wire should contract. To estimate this, consider a long wire of radius r. Suppose the atomic nuclei are fixed and have uniform density, while the electrons move along the wire with speed v. Furthermore, assume that the electrons contract, filling a cylinder of radius $r' < r$ with uniform negative charge density, and that the wire is overall neutral. Find r'.

Solution

The contraction of the electrons produces an overall inward electric field, and hence an outward electric force on each electron, which balances the radially inward magnetic force. Specifically, equilibrium occurs when $E = vB$.

Let the charge densities of the nuclei and electrons be ρ_+ and ρ_-. The magnetic field at radius r is found by Ampere's law, which gives

$$(2\pi r)B = \mu_0(\rho_- v)(\pi r^2), \quad B = \frac{\mu_0 \rho_- v r}{2}.$$

The electric field at radius r is found by Gauss's law, which gives

$$(2\pi r)E = \frac{1}{\epsilon_0}(\rho_+ + \rho_-)\pi r^2, \quad E = \frac{1}{2\epsilon_0}(\rho_+ + \rho_-)r.$$

Note that both E and B are proportional to r. Then $E = vB$ can be satisfied at all r simultaneously, which confirms that our assumption that ρ_+ and ρ_- were uniform is self-consistent.

Plugging these results into $E = vB$ yields

$$\rho_+ + \rho_- = \rho_-(\epsilon_0 \mu_0 v^2) = \rho_- \frac{v^2}{c^2}.$$

This can be written in terms of the Lorentz factor of special relativity,

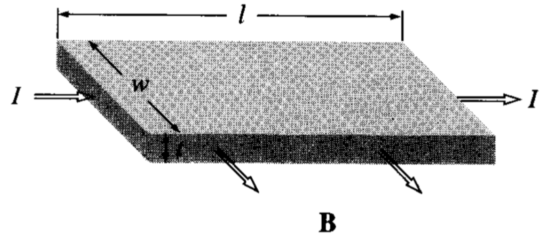
$$\rho_- = -\gamma^2 \rho_+, \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}.$$

Since the wire is overall neutral, $\rho_- r'^2 + \rho_+ r^2 = 0$, so

$$r' = \frac{r}{\gamma}.$$

For nonrelativistic motion, the contraction is extremely small. (However, in plasmas, where the positive charges are also free to move, this so-called pinch effect can be very significant.)

[2] **Problem 23** (Griffiths 5.41)**.** A current I flows to the right through a rectangular bar of conducting material, in the presence of a uniform magnetic field $\mathbf{B}$ pointing out of the page, as shown.



(a) If the moving charges are positive, in what direction are they deflected by the magnetic field? This deflection results in an accumulation of charge on the upper and lower surfaces of the bar, which in turn produces an electric force to counteract the magnetic one. Equilibrium occurs when the two exactly cancel. (This phenomenon is known as the Hall effect.)

(b) Find the resulting potential difference, called the Hall voltage, between the top and bottom of the bar, in terms of B, the speed v of the charges, and the dimensions of the bar.

(c) How would the answer change if the moving charges were negative?

When measurements were performed in the early 20th century, some metals were found to have *positive* moving charges! This "anomalous Hall effect" was solved by the quantum theory of solids, as you can learn in any solid state physics textbook. (It is related to the strange behavior you will see in problem 26.) Today, extensions of the Hall effect, such as the integer and fractional quantum Hall effects, remain active areas of research, and could be used to build quantum computers. We'll return to these effects in **X3**.

Solution. (a) By using the right-hand rule twice, we find they are deflected down.

(b) The electric field is $E = vB$, so $V = Eh = vBh$ where h is the thickness. Thus, in equilibrium, the bottom is at a higher potential.

(c) If the current stays the same, the charges move the other direction. Since both the charge and velocity flip, the Lorentz force stays the same, so the charges are still deflected down. Thus, the sign of the charge that accumulates on the bottom is flipped, so now the top is at a higher potential. Hence measuring the Hall voltage can be used to find the sign of the charge carriers in a material.

[3] **Problem 24.** USAPhO 1997, problem B1. A nice problem on the dynamics of a plasma.

[3] **Problem 25.** USAPhO 2019, problem A3. This is a tough but useful problem. The first half derives the so-called Child–Langmuir law, covered in problem 2.53 of Griffiths.

[3] **Problem 26.** USAPhO 2022, problem B3. About the weird behavior of electrons in solids.

Electromagnetism V: Induction

Chapter 7 of Purcell covers induction, as does chapter 7 of Griffiths, and chapter 8 of Wang and Ricardo, volume 2. For magnetism, see section 6.1 of Griffiths; for cool applications, see chapters II-16 and II-17 of the Feynman lectures. For a qualitative introduction to superconductivity, see appendix I of Purcell. There is a total of **87** points.

1 Motional EMF

Idea 1

If $\mathbf{F}$ is the force on a charge q, then the emf about a loop C is

$$\mathcal{E} = \frac{1}{q} \oint_C \mathbf{F} \cdot d\mathbf{s}.$$

For a moving closed loop in a time-independent magnetic field, the emf through the loop is

$$\mathcal{E} = -\frac{d\Phi}{dt}$$

where Φ is the magnetic flux through the loop. The direction of the emf produces a current that opposes the change in flux.

Example 1

A wire is bent into an arbitrary shape in the xy plane, so that its ends are at distances R_1 and R_2 from the z-axis. The wire is rotated about the z-axis with angular velocity ω, in a uniform magnetic field $B\,\hat{\mathbf{z}}$. Find the emf across the wire.

Solution

The emf is motional emf due to the magnetic force, so

$$\mathcal{E} = \int (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{r}.$$

The main point of this problem is to get you acquainted with some methods for manipulating vectors. First, we'll use components. Placing the origin along the axis of rotation, we have

$$\mathbf{v} = \mathbf{r} \times \boldsymbol{\omega} = (x\hat{\mathbf{x}} + y\hat{\mathbf{y}}) \times \omega\hat{\mathbf{z}} = \omega(y\hat{\mathbf{x}} - x\hat{\mathbf{y}})$$

for a point on the wire at $\mathbf{r}$. Evaluating the cross product with the magnetic field,

$$\mathbf{v} \times \mathbf{B} = \omega B(y\hat{\mathbf{x}} - x\hat{\mathbf{y}}) \times \hat{\mathbf{z}} = -\omega B(x\hat{\mathbf{x}} + y\hat{\mathbf{y}}) = -\omega B\mathbf{r}.$$

Therefore, we have

$$\mathcal{E} = -\omega B \int \mathbf{r} \cdot d\mathbf{r} = -\frac{\omega B}{2} \int_{R_1}^{R_2} d(r^2) = \frac{\omega B(R_1^2 - R_2^2)}{2}$$

which is completely independent of the wire's detailed shape.

Now let's solve the question again without components. Here it's useful to apply the double cross product, or "BAC-CAB" rule,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}).$$

If you want to show this for yourself, note that both sides are linear in $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, so it's enough to prove it for all combinations of unit vectors they could be; this just follows from casework. We can now simplify the emf integrand as

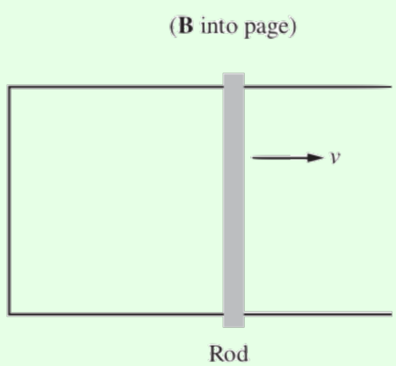
$$(\mathbf{r} \times \boldsymbol{\omega}) \times \mathbf{B} = \mathbf{B} \times (\boldsymbol{\omega} \times \mathbf{r}) = \boldsymbol{\omega}(\mathbf{B} \cdot \mathbf{r}) - \mathbf{r}(\mathbf{B} \cdot \boldsymbol{\omega}).$$

The first term is zero since $\mathbf{r}$ lies in the xy plane, while the second term is $-\omega B\mathbf{r}$. The rest of the solution follows as with the component method.

For problems that are essentially two-dimensional, there's not much difference in efficiency between the two methods, so you should use whatever you're more comfortable with. On the other hand, for problems with three-dimensional structure, components tend to get clunky.

Example 2: Purcell 7.2

A conducting rod is pulled to the right at speed v while maintaining a contact with two rails. A magnetic field points into the page.



An induced emf will cause a current to flow in the counterclockwise direction around the loop. Now, the magnetic force $q\mathbf{u} \times \mathbf{B}$ is perpendicular to the velocity $\mathbf{u}$ of the moving charges, so it can't do work on them. However, the magnetic force certainly looks like it's doing work. What's going on here? Is the magnetic force doing work or not? If not, then what is? There is definitely something doing work because the wire will heat up.

Solution

A perfectly analogous question is to imagine a block sliding down a ramp with friction, at a constant velocity. Heat is produced, so something is certainly doing work. We might suspect it's the normal force, because it has a horizontal component along the block's direction of horizontal travel. However, it also has a vertical component opposite the block's direction of vertical travel, so it of course performs no work. All it does is redirect the block's velocity; the ultimate source of energy is gravity.

Similarly, in this case, the current does not flow vertically (along the page), but also has a horizontal component because it is carried along with the rod. Just like the normal force in the ramp example, the magnetic force is perpendicular to the velocity, and does no work. It simply redirects the velocity created by whatever is pulling the rod to the right, which is the ultimate source of energy.

[2] **Problem 1** (Purcell). [A] Derive the result of idea 1 using the Lorentz force law as follows.

(a) Let the loop be C and let $\mathbf{v}$ be the velocity of each point on the loop. Argue that after a time dt, the change in flux is

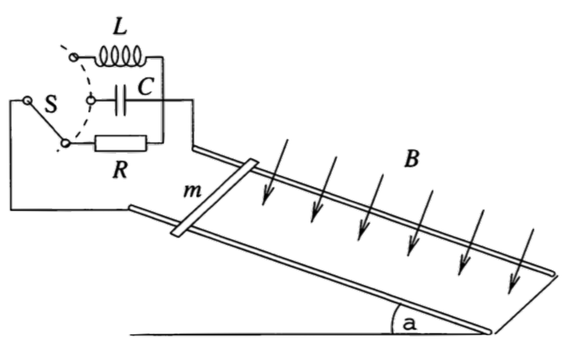
$$d\Phi = \oint_C \mathbf{B} \cdot ((\mathbf{v}\, dt) \times d\mathbf{s}).$$

(b) Using the identity $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = -\mathbf{c} \cdot (\mathbf{b} \times \mathbf{a})$, show that

$$\frac{d\Phi}{dt} = -\oint_C (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{s}$$

and use this to conclude the result.

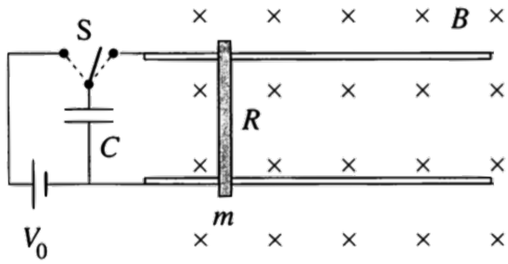
[3] **Problem 2** (PPP 167). A homogeneous magnetic field $\mathbf{B}$ is perpendicular to a track inclined at an angle α to the horizontal. A frictionless conducting rod of mass m and length ℓ straddles the two rails as shown.



How does the rod move, after being released from rest, if the circuit is closed by (a) a resistor of resistance R, (b) a capacitor of capacitance C, or (c) a coil of inductance L? In all cases, neglect the self-inductance of the closed loop formed, i.e. neglect the flux that its current puts through itself.

[3] **Problem 3.** USAPhO 2006, problem B1.

[3] **Problem 4** (PPP 168). One end of a conducting horizontal track is connected to a capacitor of capacitance C charged to voltage V_0. The inductance of the assembly is negligible. The system is placed in a uniform vertical magnetic field B, as shown.



A frictionless conducting rod of mass m, length ℓ, and resistance R is placed perpendicularly onto the track. The capacitor is charged so that the rod is repelled from the capacitor when the switch is turned. This arrangement is known as a railgun. Neglect self-inductance throughout this problem.

(a) What is the maximum velocity of the rod, and what is the maximum possible efficiency?

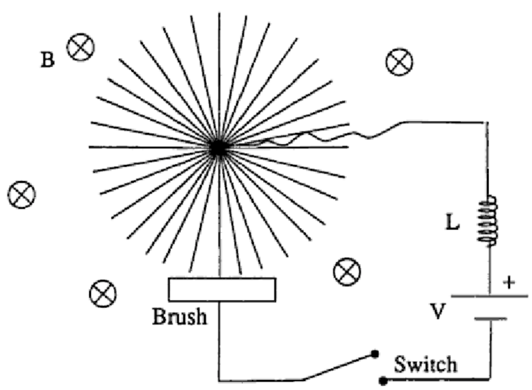
(b) At the end of this process, the rail is moving to the right. Therefore, by momentum conservation, something must have experienced a force towards the left. What is it? Answer this in both the case where the magnetic field is the same everywhere, and when it only overlaps the rails, as shown above.

[3] **Problem 5.** USAPhO 2012, problem B2.

Idea 2

Not all motional emfs can be found using $\mathcal{E} = -d\Phi/dt$. Sometimes, for more complex geometries where there is no clear "loop", it's easier to go back to the Lorentz force law.

[3] **Problem 6.** A wheel of radius R and moment of inertia J consisting of a large number of thin conducting spokes is free to rotate about an axle. A brush always makes electrical contact with one spoke at a time at the bottom of the wheel.



A battery with voltage V feeds current through an inductor L, into the axle, through the spoke, to the brush. There is a uniform magnetic field $\mathbf{B}$ pointing into the plane of the paper. At time $t = 0$ the switch is closed.

(a) Find the torque on the wheel and the motional emf along a spoke, as a function of the current I in the circuit and the angular velocity ω of the wheel.

(b) Solve for the full time evolution of $I(t)$ and $\omega(t)$. If there is a small amount of friction and resistance, then what will the final state of the system be?

This setup is an example of a homopolar motor.

[4] **Problem 7.** IPhO 1990, problem 2. A neat problem on an exotic propulsion mechanism called an electrodynamic tether, which also reviews **M6**.

2 Faraday's Law

Idea 3

Faraday's law states that even for a time-dependent magnetic field, we still have

$$\mathcal{E} = -\frac{d\Phi}{dt}.$$

In the case where the loop isn't moving but the magnetic field is changing, the emf is entirely provided by the electric field,

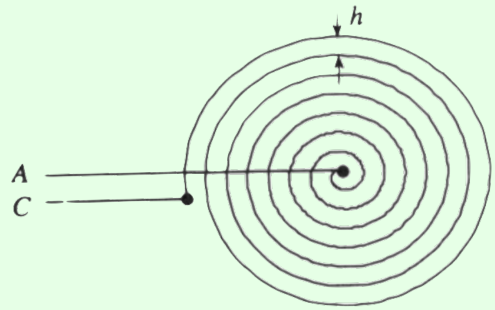
$$\mathcal{E} = \oint_C \mathbf{E} \cdot d\mathbf{s}.$$

Electric fields in the presence of changing magnetic fields can thus be nonconservative, i.e. they can have a nonzero closed line integral, a situation we haven't seen in any previous problem set. The differential form of Faraday's law is

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

Example 3

A flat metal spiral, with a constant distance h between coils, and $N \gg 1$ total turns is placed in a uniformly growing magnetic field $B(t) = \alpha t$ perpendicular to the plane of the spiral.



Find the emf induced between points A and C.

Solution

In theory, you can imagine connecting A and C and finding the flux through the resulting loop, but this is hard to visualize. A better way is to imagine turning the spiral into N concentric circles, connected in series. Then the emf is the sum of the emfs through each,

$$\mathcal{E} = \sum_{k=1}^{N} \pi (kh)^2 \alpha \approx \pi h^2 \alpha \int_0^N dk\, k^2 = \frac{\pi}{3} h^2 N^3 \alpha.$$

To see why this is valid, remember that the emfs are due to a nonconservative electric field, integrated along the length of the loop. Deforming it into a bunch of concentric circles doesn't significantly change $\mathbf{E} \cdot d\mathbf{s}$ along it, because N is large, so it doesn't change the answer much.

Remark: EMF vs. Voltage

We mentioned earlier in **E2** that we often care about electromotive forces, which just mean any forces that act on charges to push them around a circuit. The force due to a nonconservative electric field is another example.

When nonconservative electric fields are in play, the idea of "voltage" breaks down entirely, because you can't define it consistently. However, electrical engineers use a more pragmatic definition of voltage: to them, voltage is just whatever a voltmeter displays. In other words, what they call voltage is what we call electromotive force. This tends to lead to long and bitter semantic disputes, along with rather nonintuitive results, as you'll see below. For example, the "voltage" can be different for different voltmeters even if they are connected at the same points!

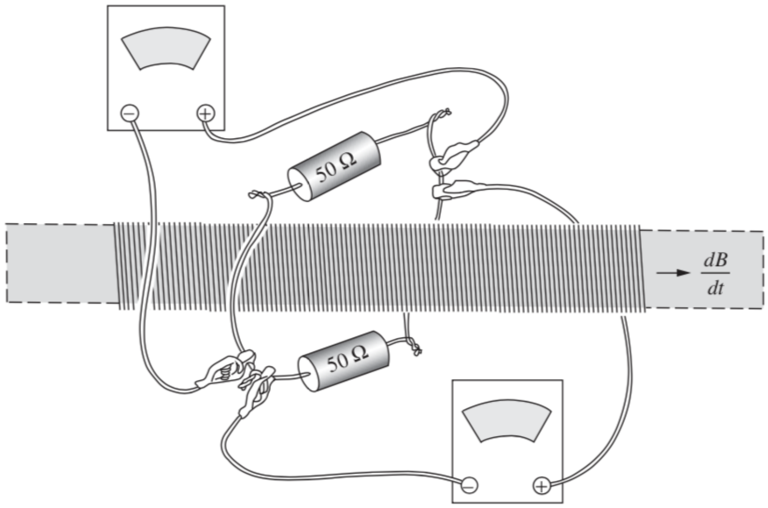
Despite this trouble, we'll go along with the standard electrical engineer nomenclature and refer to these emfs as voltages in later problem sets. For example, Kirchoff's loop rule should properly say that the sum of the voltage drops along a loop is not zero, but rather $-d\Phi/dt$. But it is conventional to move it to the other side and call it a "voltage drop" of $d\Phi/dt$.

Remark

When we apply Faraday's law, we often use Ampere's law (without the extra displacement current term) to calculate the magnetic field. This is not generally valid, but works if the currents are in the slowly changing "quasistatic" regime, which means radiation effects are negligible. All the problems below assume this, but we'll see more subtle examples in **E7**.

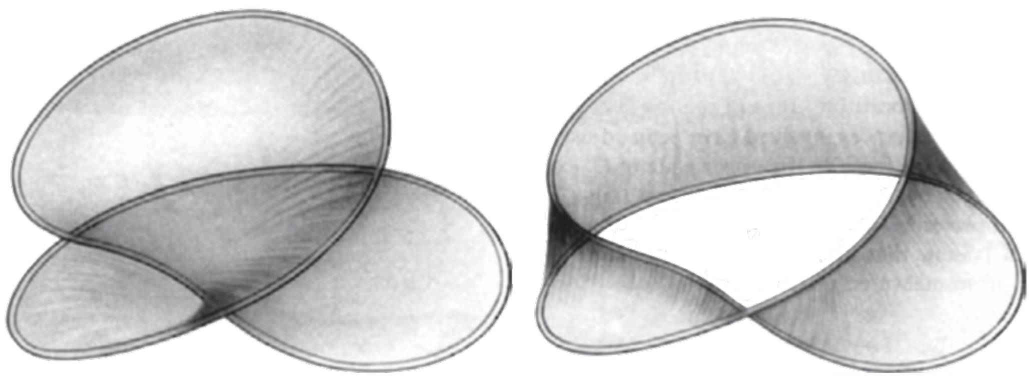
[2] **Problem 8** (Purcell 7.6)**.** An infinite cylindrical solenoid has radius R and n turns per unit length. The current grows linearly with time, according to $I(t) = Ct$. Assuming the electric field is cylindrically symmetric and purely tangential, find the electric field everywhere.

[2] **Problem 9** (Purcell 7.4)**.** Two voltmeters are attached around a solenoid with magnetic flux Φ.



Find the readings on the two voltmeters in terms of $d\Phi/dt$, paying attention to the signs.

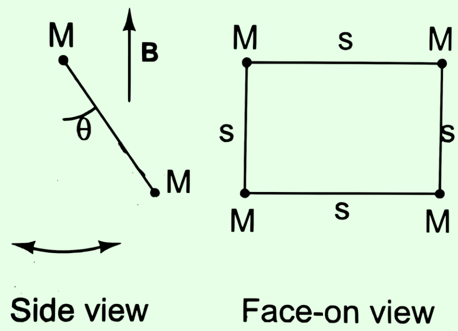
[2] **Problem 10** (Purcell 7.28). [A] Consider the loop of wire shown below.



Suppose we want to calculate the flux of $\mathbf{B}$ through this loop. Two surfaces bounded by the loop are shown above. Which, if either, is the correct surface to use? If each of the two turns in the loop are approximately circles of radius R, then what is the flux? Generalize to an N-turn coil.

Example 4

A square, rigid loop of wire has resistance R, sides of length s, and negligible mass. Point masses of mass M are attached at each corner. The top edge of the square loop is mounted so it is horizontal, and the loop may rotate as a frictionless pendulum about a fixed axis passing through this edge. Initially the pendulum is at rest at $\theta = 0$, and a uniform magnetic field $\mathbf{B}$ points horizontally through the loop. The magnetic field is then quickly rotated to the vertical direction, as shown.



Describe the subsequent evolution.

Solution

The rotation of the magnetic field provides a sharp impulse that causes the pendulum to start swinging. Letting ϕ be the angle of the field to the horizontal,

$$\mathcal{E} = -\frac{d(B_x s^2)}{dt} = -Bs^2 \frac{d(\cos\phi)}{dt}$$

and the torque about the axis of rotation is

$$\tau = (IsB_y)s = -\frac{s^4 B^2}{R} \sin\phi \frac{d(\cos\phi)}{dt}.$$

The total impulse delivered is

$$L = \int \tau \, dt = \frac{s^4 B^2}{R} \int_0^{\pi/2} \sin^2\phi \, d\phi = \frac{\pi}{4} \frac{s^4 B^2}{R}$$

which causes an initial angular velocity $\omega = L/(2Ms^2)$.

After the pendulum begins swinging, the presence of the magnetic field causes an effective drag force. To see this, note that now we have

$$\mathcal{E} = -Bs^2 \frac{d(\sin\theta)}{dt}$$

which implies

$$\tau = Is^2 B\cos\theta = -\frac{s^4 B^2}{R} \cos^2\theta \frac{d\theta}{dt}.$$

Therefore, the $\tau = I\alpha$ equation is

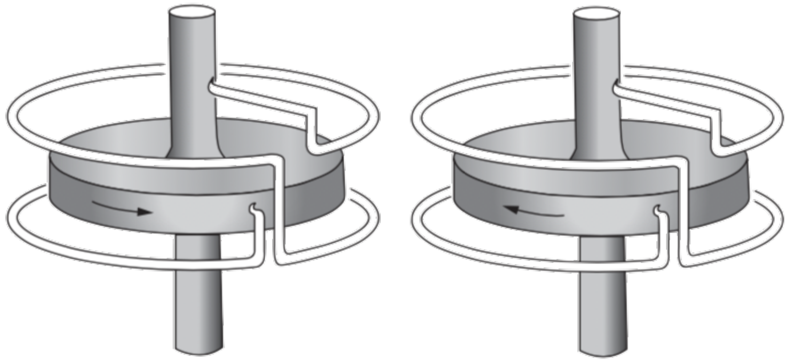
$$2Ms^2 \frac{d^2\theta}{dt^2} = -2Mgs\sin\theta - \frac{B^2 s^4}{R} \cos^2\theta \frac{d\theta}{dt}.$$

If we take the small angle approximation, then we recover ordinary damped harmonic oscillations, as covered in **M4**.

[3] **Problem 11.** USAPhO 2009, problem A1.

[3] **Problem 12.** USAPhO 1999, problem B2.

[3] **Problem 13** (Purcell). A dynamo is a generator that works as follows: a conductor is driven through a magnetic field, inducing an electromotive force in a circuit of which that conductor is part. The source of the magnetic field is the current that is caused to flow in that circuit by that electromotive force. An electrical engineer would call it a self-excited dynamo. One of the simplest dynamos conceivable is shown below.



It has only two essential parts. One part is a solid metal disk and axle which can be driven in rotation. The other is a two-turn "coil" which is stationary but is connected by sliding contacts, or "brushes", to the axle and to the rim of the revolving disk.

(a) One of the two devices pictured is, at least potentially, a dynamo. The other is not. Which is the dynamo?

A dynamo like the one above has a certain critical speed ω_0. If the disk revolves with an angular velocity less than ω_0, nothing happens. Only when that speed is attained is the induced $\mathcal{E}$ enough to make the current enough to make the magnetic field enough to induce an $\mathcal{E}$ of that magnitude. The critical speed can depend only on the size and shape of the conductors, the conductivity σ, and the constant μ_0. Let d be some characteristic dimension expression the size of the dynamo, such as the radius of the disk in our example.

(b) Show by a dimensional argument that ω_0 must be given by a relation of the form $\omega_0 = K/\mu_0\sigma d^2$ where K is some dimensionless numerical factor that depends only on the arrangement and relative size of the parts of the dynamo.

(c) Demonstrate this result again by using physical reasoning that relates the various quantities in the problem (R, $\mathcal{E}$, E, I, B, etc.). You can ignore all numerical factors in your calculations and absorb them into the constant K.

For a dynamo of modest size made wholly of copper, the critical speed would be practically unattainable. It is ferromagnetism that makes possible the ordinary DC generator by providing a magnetic field much stronger than the current in the coils, unaided, could produce. For an Earth-sized dynamo, however, the critical speed is much smaller. The Earth's magnetic field is produced by a nonferromagnetic dynamo involving motions in the fluid metallic core.

[3] **Problem 14.** USAPhO 2023, problem B1. A nice problem on a particular kind of motor, which reviews almost everything covered above in this problem set.

[3] **Problem 15** (MPPP 178)**.** In general, a magnet moving near a conductor is slowed down by induction effects. Suppose that inside a long vertical, thin-walled, brass tube a strong permanent magnet falls very slowly due to these effects, taking a time t to go from the top to the bottom.

(a) Let the magnet have mass m, and let the tube have resistivity ρ, thickness r, and length L. Suppose both the magnet and tube have radius approximately R, and let the magnet's length also be of order R. Let the typical magnetic fields produced at the magnet's surface have magnitude B_0. Find an estimate for t, up to dimensionless constants.

(b) If the experiment is repeated with a copper tube of the same length but a larger diameter, the magnet takes a time t' to fall through. How long does it take for the magnet to fall through the tubes if they are fitted inside each other? Neglect the mutual inductance of the tubes.

Remark

In this problem set, we presented motional emf first, and emf from a changing magnetic flux second. But historically, it went the other way around, as described here. Maxwell was aware of Faraday's experiments, which stated that $\mathcal{E} = -d\Phi/dt$ for stationary loops. He then demanded that this remain true for moving loops, and deduced that there must be a force per charge of $\mathbf{v} \times \mathbf{B}$. That is, Maxwell used Faraday's law to derive the Lorentz force! This is a reminder that the process of discovery is messy. When new physics is being found, the very same fact could be a law, a derived result, or simply true by definition, depending on where you start from. And it's not clear which it'll end up being until the dust settles.

3 Inductance

Idea 4: General Inductance

Consider a set of loops with fluxes Φ_i and currents I_i. By linearity, they are related by

$$\Phi_i = \sum_j L_{ij} I_j$$

where the L_{ij} are called the coefficients of inductance. It can be shown that $L_{ij} = L_{ji}$, and we call this quantity the mutual inductance of loops i and j. By Faraday's law, we have

$$\mathcal{E}_i = \sum_j L_{ij} \dot{I}_j.$$

In contrast with capacitance, we're usually concerned with the self-inductance $L_i = L_{ii}$ of single loops; these inductors provide an emf of $L\dot{I}$ each. However, mutual inductance effects can also impact how circuits behave, as we'll see in **E6**.

Remark

The inductance coefficients are similar to the capacitance coefficients in **E2**, but more useful. For capacitors, we are typically interested in configurations with one positive and one negative plate, and the capacitance of this object is related to all of the capacitance coefficients in a complicated way, as we saw in **E2**. But most inductors just use self-inductance, so the inductance we care about is simply one of the coefficients, L_{ii}. Moreover, the "mutual inductance" coefficients L_{ij} are also in the right form to be directly used, since they tell us how current changes in one part of the circuit impact emfs elsewhere.

A more general way to describe the difference is that $\mathcal{E}$ and $\dot{I}$ are directly measurable and controllable quantities, while the Q and V (i.e. the voltage relative to infinity) that the capacitance coefficients relate are less so.

Idea 5

The energy stored in a magnetic field is

$$U = \frac{1}{2\mu_0} \int B^2 \, dV$$

which implies the energy stored in an inductor is

$$U = \frac{1}{2} L I^2$$

where L is the self-inductance.

Example 5

Compute the self-inductance of a cylindrical solenoid of radius R, length $H \gg R$, and n turns per length.

Solution

One straightforward way to do this is to use the magnetic field energy. We have

$$U = \frac{1}{2\mu_0} (\mu_0 n I)^2 (\pi R^2 H)$$

and setting this equal to $LI^2/2$ gives

$$L = \pi \mu_0 n^2 R^2 H = \mu_0 N^2 \frac{\pi R^2}{H}$$

where N is the total number of turns.

We can also try to use the definition of inductance directly, $\Phi = LI$. But it's hard to imagine a surface bounded by the solenoid wires; as we saw in problem 10, even the case $N = 2$ is tricky! Instead it's better to use the form $\mathcal{E} = L\dot{I}$. We can then compute the emf across each turn of the solenoid individually, then add them together.

To compute the emf across one turn, we can replace it with a circular loop; this is valid because the emf ultimately comes from the local electric field, which shouldn't change too much if we deform the loop in this way. Then

$$|\mathcal{E}_{\text{loop}}| = \frac{d\Phi}{dt} = (\mu_0 n \dot{I})(\pi R^2).$$

The inductance is hence

$$L = \frac{N\mathcal{E}_{\text{loop}}}{\dot{I}} = (\mu_0 n N)(\pi R^2) = \mu_0 N^2 \frac{\pi R^2}{H}$$

as expected.

Example 6

Find the outward pressure at the walls of the solenoid in the previous example.

Solution

An outward pressure exists because of the Lorentz force of the the axial magnetic field of the solenoid acting on the circumferential currents at the walls. The force per length acting on a wire is IB, and the pressure is this quantity times the turns per length, so naively

$$P = (\mu_0 n I)(nI).$$

However, this is off by a factor of 2. To see why, consider a small Amperian rectangle that straddles the surface of the solenoid. The currents near this rectangle contribute axial magnetic fields of $\mu_0 nI/2$ inside and $-\mu_0 nI/2$ outside. Thus, the currents due to the entire rest of the solenoid contribute $\mu_0 nI/2$ both inside and outside. Since a wire can't exert a force on itself, only the latter field matters, so the true answer is

$$P = \frac{1}{2}\mu_0 n^2 I^2 = \frac{B^2}{2\mu_0}.$$

Remark: Electromagnetic Stress

The above example is like the one in **E1**, where we showed that the inward pressure on a conductor's surface due to electrostatic forces is $\epsilon_0 E^2/2$. In fact, there's a general principle behind both: electric and magnetic fields carry a tension per unit area (i.e. a negative pressure) of magnitude $\epsilon_0 E^2/2$ or $B^2/2\mu_0$ along their directions, and a repulsion per unit area (i.e. a positive pressure) $\epsilon_0 E^2/2$ or $B^2/2\mu_0$ perpendicular to their directions. Charges and currents, such as at the walls of a solenoid or the plates of a capacitor, cause discontinuities in $\mathbf{E}$ or $\mathbf{B}$ across them, leading to a net force on them.

This isn't mentioned in introductory electromagnetism books because the proper treatment of anisotropic pressure requires tensors. However, more advanced books will introduce the Maxwell stress tensor, from which the results above can be read off.

The great experimentalist Michael Faraday was a huge fan of these results. He viewed field lines as physical objects, which he called "lines of force", that carried tension along their lengths and repelled each other. He even presciently suggested that light consisted of waves propagating along lines of force, like waves on a string.

These days, we don't ascribe so much importance to field lines. The fundamental object is the field itself, and field lines are a secondary construction that often just add mathematical complication. For example, the field of a dipole is simple, but it's not so simple to solve for the corresponding field lines. Things get even more complicated in dynamic situations, where field lines can appear and disappear; Faraday viewed induction as a result of "cutting" magnetic field lines. And in **R3**, we'll show how fields transform between frames, which implies that the very existence of a field line can depend on the reference frame. Still, Faraday's intuition might be helpful occasionally, and it's still a useful tool in some subfields. For instance, in plasma physics, field lines can be used to visualize magnetic reconnection.

Remark

Recall the example in **E1** involving the force between two spherical balls of charge. There, we got the answer using a slightly tricky argument, where Newton's third law allowed us to use the shell theorem twice. But the idea of electromagnetic stress provides a straightforward alternative proof which also works for more general situations.

Suppose the two balls lie above and below the xy plane, and additional external forces hold them both at rest. Consider as a system everything at $z > 0$, which includes the second ball and a lot of empty space. The only external forces on this system are from the attractive $\epsilon E^2/2$ pressure at the xy plane, and the force F which holds the second ball in place. Since the momentum of the system is constant, these forces must cancel. Thus, to compute F we only need to know the electric field on the xy plane, for which we can clearly apply the shell theorem to both balls, replacing them with point charges.

The general idea is that electromagnetic forces on objects can be determined solely by the electromagnetic fields in the space around them. You can use this to find shorter solutions to some problems in **E1**.

[3] **Problem 16.** Consider a toroidal solenoid with a rectangular cross section of height h and width w, N turns, and inner radius R.

(a) Find the self-inductance by considering the magnetic flux.

(b) Now suppose the current increases at a constant rate dI/dt. Find the magnitude of the electric field at a height z above the center of the solenoid, assuming $h, w \ll R$. (Hint: write down the divergence and curl of $\mathbf{E}$ in terms of $\dot{\mathbf{B}}$ in general, and notice the similarities to the equations for $\mathbf{B}$ in terms of $\mathbf{J}$. This allows us to use the ideas of **E3** by analogy.)

(c) Verify that the two formulas for energy given in idea 5 are consistent in this setup.

Remark

In electromagnetism, we often have issues with divergences when we take idealized point sources. For example, the voltage near a point charge can become arbitrarily high. Similarly, the magnetic field diverges as you approach an idealized, infinitely-thin wire, which causes the self-inductance of wire loops to diverge. Of course, the resolution is that you don't actually get an infinite magnetic field as you approach a wire. A real wire has finite thickness, and its magnetic field instead goes to zero as you approach its center. (We didn't run into this problem for solenoids, because we modeled their wires as a uniform sheet of current, whose magnetic field isn't singular at all.) If a problem does involve a wire loop, it'll often circumvent this messy issue by just giving the self-inductance from the start.

[2] **Problem 17.** A wire of length ℓ is bent into a long "hairpin" shape, with two parallel straight edges of length $\ell/2$ separated by a distance $d \ll \ell$.

(a) Write down an integral expression for the self-inductance, neglecting the curved parts, and show that it diverges.

(b) Find a rough estimate for the self-inductance by taking the wire to have radius $r \ll d$ and ignoring any flux through the wire itself.

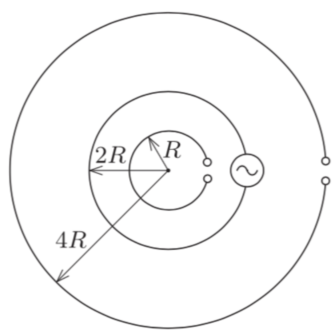
[3] **Problem 18.** Consider two concentric rings of radii r and $R \gg r$.

(a) Compute the mutual inductance by considering a current through the larger ring.

(b) Compute the mutual inductance by considering a current through the smaller ring, and verify your results agree. (Hint: this can be done without difficult integrals.)

In general, computing mutual inductance is a hard and practically important problem; there have been whole books written on the subject.

[2] **Problem 19** (MPPP 181)**.** Three nearly complete circular loops, with radii R, $2R$, and $4R$ are placed concentrically on a horizontal table, as shown.



A time-varying electric current is made to flow in the middle loop. Find the voltage induced in the largest loop at the moment when the voltage between the terminals of the smallest loop is V_0.

4 Magnetism

In this section we'll dip a little into atomic physics and the origin of magnetism. However, a proper understanding of this subject requires quantum mechanics, as we'll cover in **T3** and **X3**.

Idea 6

A spinning charged object carries a magnetic dipole moment $\boldsymbol{\mu}$ and angular momentum $\mathbf{J}$. If the object's mass and charge distributions are proportional, then $\boldsymbol{\mu}$ and $\mathbf{J}$ point in the same direction, and one can show that their ratio is always $\mu/J = q/2m$.

Example 7

Suppose the magnetic moment of an iron atom is due to a single unpaired electron, with angular momentum of order $\hbar$. The atoms are separated by a distance $d \sim 10^{-10}$ m. Estimate the maximum magnetic field an iron magnet can produce. How does this compare to the fields that can be produced in an electromagnet?

Solution

The answer doesn't scale significantly with the physical size of the iron magnet. To see this, think in terms of electric dipoles: if you have a giant cube of electric dipoles, it's equivalent to having a fixed surface charge density $\pm\sigma$ on two of the faces. The electric field produced by such a charge density near each face is of order σ/ϵ_0, independent of the size of the cube.

Therefore, the only things the magnetic field can depend on are μ_0, the magnetic dipole moment μ of a single atom, and d. By dimensional analysis,

$$B \sim \mu_0 \frac{\mu}{d^3}$$

which can also be thought of as $\mu_0 M$, where M is the magnetization density. Taking $\mu \sim e\hbar/m_e$ and plugging in the numbers gives $B \sim 10\,\text{T}$, which is the right order of magnitude.

Now consider the case of an electromagnet, where the field is produced by moving electrons with typical speed v, moving in a loop with typical size r. In a metal, there's on the order of one free electron per atom, so d is still the same. The difference is that the field made by each electron *does* scale with r, because each has magnetic moment

$$\mu = IA \sim \frac{ev}{r} r^2.$$

Compared to the previous result, this is larger by a factor of $mvr/\hbar$. The two are comparable, for $r \sim 1\,\text{m}$, if the electrons travel at the agonizingly slow velocity $v \sim 10^{-4}\,\text{m/s}$.

Therefore, you would get a magnetic field much larger than $10\,\text{T}$ if you could make the electrons go at a reasonable walking speed, but that's easier said than done. The largest steady magnetic fields made in the lab are only about $40\,\text{T}$. Such a field carries a pressure which would rip apart a solenoid made of coiled wire,

$$P = \frac{(40\,\text{T})^2}{2\mu_0} = 0.6\,\text{GPa}.$$

Instead, these fields are produced in Bitter electromagnets, which are solenoids made of thick metal plates, perforated with cooling channels to dissipate the enormous heat produced by resistance. It is possible to produce higher fields temporarily, but the results will be explosive.

[3] **Problem 20.** USAPhO 2021, problem A3. This covers a simple classical model of the electron.

[3] **Problem 21.** USAPhO 2007, problem B2. (Equation 10 of the official solution has a typo.)

[5] **Problem 22.** APhO 2013, problem 3. A solid question involving classical magnetic moments, which gives some intuition for the quantum behavior.

5 Superconductors

There are many tough Olympiad problems involving superconductors. Superconductors can be a bit intimidating at first, but they actually obey simple rules.

Idea 7

An ideal conductor has zero resistivity, which implies that the magnetic flux through any loop in the conductor is constant: attempting to change the flux instantly produces currents that cancel out the change. However, the flux can be nonzero.

A superconductor is an ideal conductor with the additional property that the magnetic field in the body of the superconductor is exactly zero, no matter what the initial conditions are; once an object becomes superconducting it forces all the existing flux out. This is known as the Meissner effect. It further implies that all the current in a superconductor is confined to its surface, and that the normal component of the magnetic field $B_\perp$ is zero on the surface. Many problems involving superconductors don't even use the Meissner effect, so they would also work for ideal conductors.

Example 8: PPP 153

A superconducting uniform spring has N turns of radius R, relaxed length x_0, and spring constant k. The two ends of the spring are connected by a wire, and a small, steady current I is made to flow through the spring. At equilibrium, what is the change in its length?

Solution

This question really is about ideal conductors, not just superconductors. The additional superconductivity property would tell us about the field *inside* the wires themselves (not the loops that the wires form), and thereby about some small screening currents on the surfaces of the wires. This is not important because the wires are thin compared to the spring as a whole.

In order to find the equilibrium length x_{eq}, we can use the principle of virtual work. We compute how the energy changes if we slightly perturb the system. At equilibrium, this change in energy should be zero.

We have $B = \mu_0 NI/x$, so the magnetic field energy is

$$U = \frac{B^2}{2\mu_0} V = \frac{AI^2}{x}, \quad A = \frac{\mu_0 \pi R^2 N^2}{2}.$$

Naively, this means the magnetic field energy decreases as x increases, so the spring would like to stretch. But this makes no sense, because we know that parallel currents attract, squeezing the spring. We have to recall that the spring is an ideal conductor, so when it is stretched or squeezed, the current changes to keep the flux the same. The flux is

$$\Phi_B = N(\pi R^2)B \propto \frac{I}{x}$$

so we have

$$I(x) = I\frac{x}{x_{\text{eq}}}, \quad U(x) = \frac{AI^2}{x_{\text{eq}}^2}x.$$

The other energy contribution is $k(x-x_0)^2/2$, so setting the derivative of energy to zero,

$$\frac{AI^2}{x_{\text{eq}}^2} = k(x_0 - x_{\text{eq}}).$$

Since the current is small, $x_0 \approx x_{\text{eq}}$, so we can replace x_{eq} with x_0 on the left-hand side, giving the answer,

$$x_{\text{eq}} = x_0 - \frac{AI^2}{x_0^2 k}.$$

As a sidenote, the original formulation of this question involved an external voltage source forcing the current I to be constant. However, in this case using energy conservation is more subtle because one has to account for the work done by the voltage source. Here we used a superconductor, which keeps the flux constant, so that the spring can be thought of as an isolated system. The final answers are the same, since in both cases we have the same magnetic forces, which determine the spring's compression.

Example 9

A long, thin cylinder of radius R is placed in a magnetic field B_0 parallel to its axis. The cylinder originally carries no current on its surface, and it is cooled until it reaches the superconducting state. Find the resulting distribution of current on its surface. Now suppose the external magnetic field is turned off; what is the new current distribution?

Solution

Solving this question requires using both properties. The Meissner effect tells us there is no magnetic field within the body of the cylinder itself (i.e. the region from $r = R$ to $r = R + dr$). The ideal conducting property tells us that the flux through a cross-section of the cylinder (i.e. the region from $r = 0$ to $r = R$) is constant, and hence equal to $\pi R^2 B_0$.

When the cylinder becomes superconducting, the Meissner effect kicks in, and the field within the body of the cylinder can be cancelled by a uniform surface current on the outer surface. By the same logic as we used to compute the field of a cylindrical solenoid, it is

$$K_{\text{out}} = -B_0/\mu_0.$$

To keep the flux constant, a compensating opposite current must appear on the inner surface,

$$K_{\text{in}} = B_0/\mu_0.$$

When we turn off the external magnetic field, the two properties imply

$$K_{\text{out}} = 0, \quad K_{\text{in}} = B_0/\mu_0$$

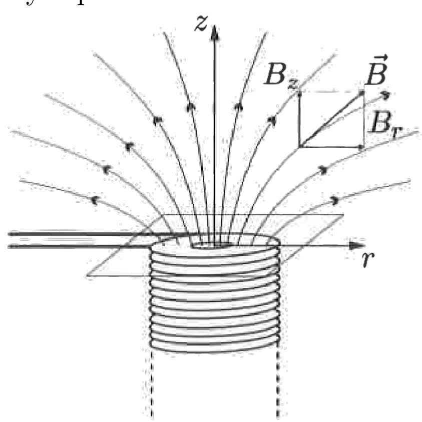
which you should check if you're not sure.

[3] **Problem 23** (MPPP 182). Two identical superconducting rings are initially very far from each other. The current in the first is I_0, but there is no current in the other. The rings are now slowly brought closer together. Find the current in the first ring when the current in the second is I_1.

[4] **Problem 24** (PPP 182, Russia 2006). A thin superconducting ring of radius r, mass m, and self-inductance L is supported by a piece of plastic just above the top of a long, cylindrical solenoid of radius $R \gg r$ and n turns per unit length. The ring and solenoid are coaxial. When the current in the solenoid is I_s, the magnetic field near the end of the solenoid is

$$B_z = B_0(1 - \alpha z), \quad B_r = B_0 \beta r$$

where we put the origin at the very top of the solenoid.



(a) Find an expression for B_0. (Keep your answers below in terms of B_0 to avoid clutter.)

(b) Find β in terms of α. What are their signs?

(c) Let I be the current through the ring. Suppose that initially $I_s = I = 0$. Find the value I_c of I_s when the ring lifts off the plastic.

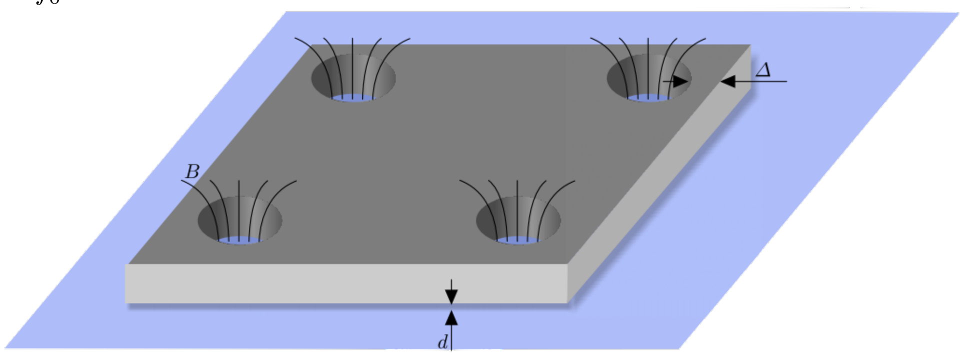
(d) Now the piece of plastic is removed and the ring is return to the same position. Initial conditions are set up so that $I_s = I_c$ and $I = 0$. The ring is released from rest. Find its subsequent motion, assuming for simplicity that the expressions for B_z and B_r above always hold. Express your final answers in terms of only α and g.

(e) In reality, the expressions for B_z and B_r break down if the ring moves too far. Consider part (d) again, but now suppose the *exact* expressions for B_z and B_r are used. Without solving any differential equations, will the resulting motion be qualitatively similar or not?

[4] **Problem 25.** IPhO 2012, problem 1C. A delightfully tricky problem that uses the properties of superconductors in a subtle way.

[4] **Problem 26.** EuPhO 2017, problem 3. Another tricky problem, using ideas we've seen before.

[4] **Problem 27** (Physics Cup 2013)**.** A rectangular superconducting plate of mass m has four identical circular holes, one near each corner, a distance Δ from the plate's edges. Each hole carries a magnetic flux Φ. The plate is put on a horizontal superconducting surface. The magnetic repulsion between the plate and the surface balances the weight of the plate when the width of the air gap beneath the plate is $d \ll \Delta$, and d is much smaller than the radii of the holes. The frequency of small vertical oscillations is f_0.



Next, a load of mass M is put on the plate, so that the load lays on the plate, and the plate levitates above the support. What is the new frequency of small oscillations?

[5] **Problem 28.** IPhO 1994, problem 2. This problem tests your intuition for induction, and is good preparation for **E6**.

Remark

In **E4**, we spent a lot of time applying $F = ma$ to charges. But in this problem set, we were somehow able to find how systems of charges behave using only Maxwell's equations, without ever explicitly referring to the forces on charges. Certainly this information has to be used implicitly somewhere, so what's going on?

To investigate this, let's do a careful derivation of Kirchoff's loop rule, for a series RLC circuit with a battery. By applying the work-kinetic energy theorem to a charge q as it goes around the circuit, from one capacitor plate to the other, we have

$$\int_C \mathbf{E} \cdot d\mathbf{s} + \int_C \mathbf{f} \cdot d\mathbf{s} = \frac{\Delta \mathrm{KE}}{q}$$

where $\mathbf{f}$ is any non-electric force per charge, and the line integrals follow the path C of the charge. By assumption, the battery and resistor contribute

$$\int_C \mathbf{f} \cdot d\mathbf{s} = \begin{cases} \mathcal{E} & \text{battery} \\ -IR & \text{resistor} \end{cases}$$

where the forces are due to chemical reactions (as covered in **E2**) or collisions with the ions (as covered in **E4**). Meanwhile, Faraday's law states

$$\oint \mathbf{E} \cdot d\mathbf{s} = \int_C \mathbf{E} \cdot d\mathbf{s} + \frac{Q}{C} = -\frac{d\Phi_B}{dt} = -\dot{I}L$$

where we need to add on Q/C to close the loop through the capacitor. Thus,

$$\mathcal{E} = \dot{I}L + IR + \frac{Q}{C} + \frac{\Delta \text{KE}}{q}.$$

Now, the key point is that in a conductor, the charges are extremely light and extremely numerous; it only takes a tiny amount of kinetic energy to get an enormous current. Therefore, the energy in any circuit is dominated by the energies stored in the inductor and capacitor, while the kinetic energy of the charges is negligible. We thus set the ΔKE term to zero to get the usual form of Kirchoff's loop rule.

Most books gloss over the derivation of Kirchoff's loop rule; for instance, Halliday, Resnick, and Krane merely prove it in the trivial case of an all-resistor circuit. Unfortunately, most purported "derivations" of it in other sources, or online, are simply wrong. For example, a common claim is that in the absence of inductors, Kirchoff's loop rule is nothing more than the statement that $\oint \mathbf{E} \cdot d\mathbf{s} = 0$. But this doesn't explain how the term Q/C can show up; since the electric field of a capacitor is conservative, its closed line integral always vanishes. The confusion only multiplies once inductors are in play.

As another note, if the work done on the charges is positive in some parts of the circuit, and negative in others, shouldn't the current wildly speed up and slow down as it goes through the wires? No, because as we saw in **E2**, charges strongly repel each other, so charge can't accumulate anywhere. More precisely it's because wires have negligible capacitance; in the fluid flow analogy, the fluid is incompressible.

To illustrate this point, consider a discharging RL circuit, where the inductor has no resistance. As the current in the inductor decreases, it induces an electric field along the inductor wires. The charges in the circuit then redistribute themselves as they flow; as a result, the electric field in the inductor wire is almost completely cancelled, while the induced emf $\dot{I}L$ appears across the resistor. It's just like how it's possible to pull on a massless rope attached to a massive block, even though the net force on a massless object always has to be zero – an internal tension force appears to transfer the force to the block.

Above, I say "almost" because the kinetic energy of the charges does play a small role. In other situations, it's possible for it to have a big effect. For example, if you really had a completely ideal wire loop, with no resistance and no capacitance, and twisted on itself so that it had no inductance, attached to an ideal battery, then the limiting factor which stops the current from becoming infinite is this inertia. The kinetic energy of charges is proportional to $v^2 \propto I^2$, so it acts like a very tiny inductance distributed throughout the wire (known as kinetic inductance), resisting changes in current. You'll see some examples in **ERev** where the motion of charges plays a direct role.

Electromagnetism V: Induction

Chapter 7 of Purcell covers induction, as does chapter 7 of Griffiths, and chapter 8 of Wang and Ricardo, volume 2. For magnetism, see section 6.1 of Griffiths; for cool applications, see chapters II-16 and II-17 of the Feynman lectures. For a qualitative introduction to superconductivity, see appendix I of Purcell. There is a total of **87** points.

1 Motional EMF

Idea 1

If $\mathbf{F}$ is the force on a charge q, then the emf about a loop C is

$$\mathcal{E} = \frac{1}{q} \oint_C \mathbf{F} \cdot d\mathbf{s}.$$

For a moving closed loop in a time-independent magnetic field, the emf through the loop is

$$\mathcal{E} = -\frac{d\Phi}{dt}$$

where Φ is the magnetic flux through the loop. The direction of the emf produces a current that opposes the change in flux.

Example 1

A wire is bent into an arbitrary shape in the xy plane, so that its ends are at distances R_1 and R_2 from the z-axis. The wire is rotated about the z-axis with angular velocity ω, in a uniform magnetic field $B\,\hat{\mathbf{z}}$. Find the emf across the wire.

Solution

The emf is motional emf due to the magnetic force, so

$$\mathcal{E} = \int (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{r}.$$

The main point of this problem is to get you acquainted with some methods for manipulating vectors. First, we'll use components. Placing the origin along the axis of rotation, we have

$$\mathbf{v} = \mathbf{r} \times \boldsymbol{\omega} = (x\hat{\mathbf{x}} + y\hat{\mathbf{y}}) \times \omega\hat{\mathbf{z}} = \omega(y\hat{\mathbf{x}} - x\hat{\mathbf{y}})$$

for a point on the wire at $\mathbf{r}$. Evaluating the cross product with the magnetic field,

$$\mathbf{v} \times \mathbf{B} = \omega B(y\hat{\mathbf{x}} - x\hat{\mathbf{y}}) \times \hat{\mathbf{z}} = -\omega B(x\hat{\mathbf{x}} + y\hat{\mathbf{y}}) = -\omega B\mathbf{r}.$$

Therefore, we have

$$\mathcal{E} = -\omega B \int \mathbf{r} \cdot d\mathbf{r} = -\frac{\omega B}{2} \int_{R_1}^{R_2} d(r^2) = \frac{\omega B(R_1^2 - R_2^2)}{2}$$

which is completely independent of the wire's detailed shape.

Now let's solve the question again without components. Here it's useful to apply the double cross product, or "BAC-CAB" rule,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}).$$

If you want to show this for yourself, note that both sides are linear in $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, so it's enough to prove it for all combinations of unit vectors they could be; this just follows from casework. We can now simplify the emf integrand as

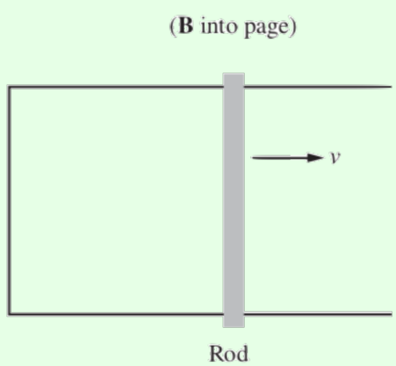
$$(\mathbf{r} \times \boldsymbol{\omega}) \times \mathbf{B} = \mathbf{B} \times (\boldsymbol{\omega} \times \mathbf{r}) = \boldsymbol{\omega}(\mathbf{B} \cdot \mathbf{r}) - \mathbf{r}(\mathbf{B} \cdot \boldsymbol{\omega}).$$

The first term is zero since $\mathbf{r}$ lies in the xy plane, while the second term is $-\omega B \mathbf{r}$. The rest of the solution follows as with the component method.

For problems that are essentially two-dimensional, there's not much difference in efficiency between the two methods, so you should use whatever you're more comfortable with. On the other hand, for problems with three-dimensional structure, components tend to get clunky.

Example 2: Purcell 7.2

A conducting rod is pulled to the right at speed v while maintaining a contact with two rails. A magnetic field points into the page.



An induced emf will cause a current to flow in the counterclockwise direction around the loop. Now, the magnetic force $q\mathbf{u} \times \mathbf{B}$ is perpendicular to the velocity $\mathbf{u}$ of the moving charges, so it can't do work on them. However, the magnetic force certainly looks like it's doing work. What's going on here? Is the magnetic force doing work or not? If not, then what is? There is definitely something doing work because the wire will heat up.

Solution

A perfectly analogous question is to imagine a block sliding down a ramp with friction, at a constant velocity. Heat is produced, so something is certainly doing work. We might suspect it's the normal force, because it has a horizontal component along the block's direction of horizontal travel. However, it also has a vertical component opposite the block's direction of vertical travel, so it of course performs no work. All it does is redirect the block's velocity; the ultimate source of energy is gravity.

Similarly, in this case, the current does not flow vertically (along the page), but also has a horizontal component because it is carried along with the rod. Just like the normal force in the ramp example, the magnetic force is perpendicular to the velocity, and does no work. It simply redirects the velocity created by whatever is pulling the rod to the right, which is the ultimate source of energy.

[2] **Problem 1** (Purcell)**.** [A] Derive the result of idea 1 using the Lorentz force law as follows.

(a) Let the loop be C and let $\mathbf{v}$ be the velocity of each point on the loop. Argue that after a time dt, the change in flux is

$$d\Phi = \oint_C \mathbf{B} \cdot ((\mathbf{v}\, dt) \times d\mathbf{s}).$$

(b) Using the identity $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = -\mathbf{c} \cdot (\mathbf{b} \times \mathbf{a})$, show that

$$\frac{d\Phi}{dt} = -\oint_C (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{s}$$

and use this to conclude the result.

Solution. (a) Consider a piece $d\mathbf{s}$ of the loop, and consider its motion over a time dt. The piece moves by $\mathbf{v}\, dt$, so we can construct a surface whose boundary is the new loop by considering the original surface, and appending these infinitesimal $d\mathbf{s}$ by $\mathbf{v}\, dt$ parallelograms to it. The amount of flux going through an infinitesimal parallelograms is $\mathbf{B} \cdot (\mathbf{v}\, dt \times d\mathbf{s})$. Integrating over the entire loop yields the desired result.

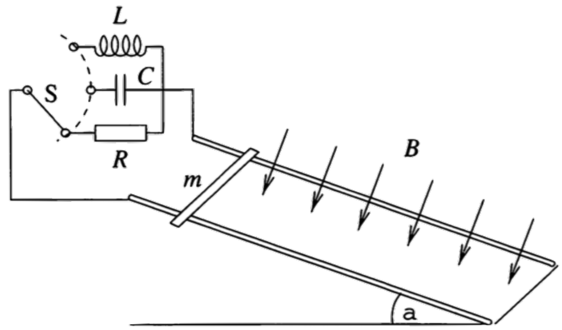
(b) Combining the previous parts, we have

$$d\Phi = -\oint_C (\mathbf{v}\, dt \times \mathbf{B}) \cdot d\mathbf{s}.$$

Dividing by dt yields the desired result.

You might also be wondering how to prove this identity. Note that $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is the volume of the parallelepiped (i.e. a three-dimensional parallelogram) whose edges are $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$. That's because the volume is the product of the area of the base and the height. Considering $\mathbf{b}$ and $\mathbf{c}$ to form the base gives a base area $|\mathbf{b} \times \mathbf{c}|$, and taking the dot product with $\mathbf{a}$ accounts for the height. The expression $\mathbf{c} \cdot (\mathbf{b} \times \mathbf{a})$ computes the same volume, up to a sign, with $\mathbf{a}$ and $\mathbf{b}$ forming the base. The correct sign can be found by considering a simple case, like the cube, and appropriately applying the right hand rule.

[3] **Problem 2** (PPP 167)**.** A homogeneous magnetic field $\mathbf{B}$ is perpendicular to a track inclined at an angle α to the horizontal. A frictionless conducting rod of mass m and length ℓ straddles the two rails as shown.



How does the rod move, after being released from rest, if the circuit is closed by (a) a resistor of resistance R, (b) a capacitor of capacitance C, or (c) a coil of inductance L? In all cases, neglect the self-inductance of the closed loop formed, i.e. neglect the flux that its current puts through itself.

Solution. Suppose the speed of the rod is v down the plane, and the current is I going from the side closer to the reader, to the side farther. By Newton's second law, we have

$$m\dot{v} = mg \sin\alpha - I\ell B.$$

The motional emf is $\mathcal{E} = \ell v B$, where positive $\mathcal{E}$ works to increase I. Thus, we have $\dot{\mathcal{E}} = \ell B\dot{v}$, so

$$\frac{m}{\ell B}\dot{\mathcal{E}} = mg \sin\alpha - I\ell B.$$

All that differs between the three parts is the expression for $\mathcal{E}$.

(a) Here we have $\mathcal{E} = IR$, so

$$\frac{m}{\ell B}R\dot{I} = mg \sin\alpha - I\ell B.$$

The solution to this is a decaying exponential that starts at 0 and asymptotes to $I_f = \frac{mg \sin\alpha}{\ell B}$. The velocity of the rod is $v = IR/\ell B$, so the terminal velocity is

$$v_f = \frac{Rmg \sin\alpha}{\ell^2 B^2}.$$

(b) Here we have $\mathcal{E} = Q/C$ where $\dot{Q} = I$, so $\dot{\mathcal{E}} = I/C$. Thus,

$$\frac{m}{\ell BC}I = mg \sin\alpha - I\ell B,$$

which implies the current is constant, and equal to

$$I = \frac{mg \sin\alpha}{\ell B + \frac{m}{\ell BC}}.$$

Note that

$$\ell B\dot{v} = \dot{\mathcal{E}} = \frac{I}{C}.$$

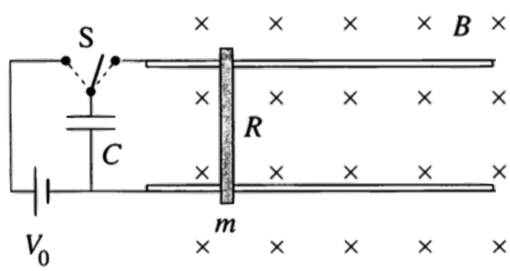
This implies that the motion is uniformly accelerated, with acceleration

$$a = \frac{mg \sin\alpha}{m + \ell^2 B^2 C}.$$

(c) Here $\mathcal{E} = L\dot{I}$, so
$$\frac{m}{\ell B} L\ddot{I} = mg\sin\alpha - I\ell B.$$
This is a simple harmonic motion equation with a shifted origin. Explicitly solving, using the usual techniques of **M1**, gives the general solution
$$I(t) = \frac{mg\sin\alpha}{\ell B} + I_0\cos(\omega t + \phi), \quad \omega^2 = \frac{\ell^2 B^2}{mL}.$$
The initial conditions are $I(0) = 0$ and $\dot{I}(0) = 0$ since $v(0) = 0$, so the particular solution is
$$I(t) = \frac{mg\sin\alpha}{\ell B}(1 - \cos(\omega t)).$$
Now, Faraday's law states that $\ell B\dot{x} = L\dot{I}$, and since $x(0) = 0$ and $I(0) = 0$, integrating gives
$$x(t) = \frac{L}{\ell B} I = \frac{mgL\sin\alpha}{\ell^2 B^2}(1 - \cos(\omega t)).$$

[3] **Problem 3.** USAPhO 2006, problem B1.

[3] **Problem 4** (PPP 168)**.** One end of a conducting horizontal track is connected to a capacitor of capacitance C charged to voltage V_0. The inductance of the assembly is negligible. The system is placed in a uniform vertical magnetic field B, as shown.



A frictionless conducting rod of mass m, length ℓ, and resistance R is placed perpendicularly onto the track. The capacitor is charged so that the rod is repelled from the capacitor when the switch is turned. This arrangement is known as a railgun. Neglect self-inductance throughout this problem.

(a) What is the maximum velocity of the rod, and what is the maximum possible efficiency?

(b) At the end of this process, the rail is moving to the right. Therefore, by momentum conservation, something must have experienced a force towards the left. What is it? Answer this in both the case where the magnetic field is the same everywhere, and when it only overlaps the rails, as shown above.

Solution. (a) Let I be the downward current in the rod, and let q be the charge on the capacitor. We see that $\dot{q} = -I$, and Kirchoff's loop rule gives
$$\frac{q}{C} - IR = \mathcal{E} = v\ell B.$$
Taking the derivative and plugging in $\dot{v} = I\ell B/m$ gives
$$\dot{I} = -\left(\frac{1}{RC} + \frac{\ell^2 B^2}{Rm}\right) I.$$

The initial condition is $I(0) = q/RC$, so

$$I(t) = \frac{q}{RC} \exp\left(-t\left(\frac{1}{RC} + \frac{\ell^2 B^2}{Rm}\right)\right).$$

Thus, integrating $\dot{v} = I\ell B/m$ and using $v(0) = 0$ gives

$$\begin{aligned} v(t) &= \frac{q\ell B}{RCm}\left(\frac{1}{RC} + \frac{\ell^2 B^2}{Rm}\right)^{-1}\left(1 - \exp\left(-t\left(\frac{1}{RC} + \frac{\ell^2 B^2}{Rm}\right)\right)\right) \\ &= \frac{V_0 \ell BC}{m + B^2\ell^2 C}\left(1 - \exp\left(-t\left(\frac{1}{RC} + \frac{\ell^2 B^2}{Rm}\right)\right)\right). \end{aligned}$$

Thus, the rod continually accelerates, asymptotically reaching a maximum speed of

$$v_{\max} = \frac{V_0 \ell BC}{m + B^2 \ell^2 C}.$$

The efficiency is the fraction of the initial energy converted to kinetic energy

$$\eta = \frac{mv_0^2/2}{CV_0^2/2} = \frac{m}{C}\frac{\ell^2 B^2 C^2}{(m + B^2\ell^2 C)^2} = \frac{1}{(p + 1/p)^2}$$

where $p = \frac{\sqrt{m}}{\sqrt{C}B\ell}$. Thus, by the AM-GM inequality, the maximum efficiency is 1/4.

(b) Momentum is conserved in both cases. When the magnetic field is uniform, it overlaps the left end of the circuit. The current in the rod implies a return current in the left end, and thus an opposite Lorentz force on it. If the circuit is held in place, the compensating leftward momentum goes to the Earth; if it isn't held in place, the whole circuit recoils to the left.

Now suppose the magnetic field is as shown in the figure, i.e. it doesn't overlap the left part of the circuit. (It does overlap the rails, but that doesn't produce a leftward Lorentz force and so is irrelevant.) To see how momentum is conserved, we need to remember that in electrostatics and magnetostatics, forces are ultimately between charges and currents. We get used to using the Lorentz force law with a given magnetic field, but that magnetic field has to be produced by some current. That current, in turn, can feel a force due to the magnetic field produced by the current in the railgun.

If the magnetic field were the same everywhere, then we could place the currents sourcing them very far away, and thus ignore this effect. (For example, the railgun could be between two distant, infinite uniform sheets of current.) But if the magnetic field is nonhomogeneous, as it is in this case, there must be current nearby. For example, the sudden decrease of the magnetic field shown in the figure above could be achieved by having an infinite sheet of current, which is cut perpendicularly by the rails, with surface current density pointing up the page.

Finally, the current through the rail creates a magnetic field at the current sheet that points into the page. And that implies a Lorentz force to the left, precisely balancing the rightward Lorentz force on the rail. Momentum is thus conserved; to see explicitly how Newton's third law holds up, see problem 5.50 of Griffiths.

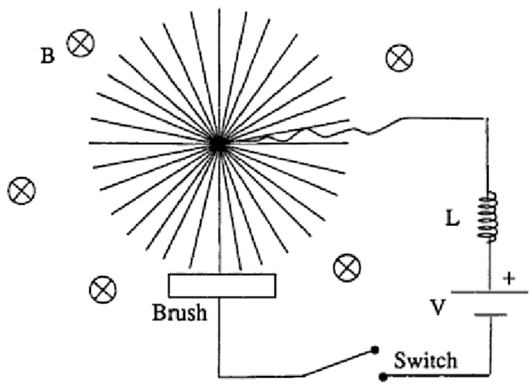
Incidentally, you might have heard the electromagnetic field can *also* carry momentum. Because of this, in general we shouldn't think of charges and currents interacting with each other, since their momentum won't be conserved; Newton's third law won't hold in general. Instead, charges and currents interact with the *field*, and the field then interacts with other charges and currents. However, we didn't need that subtlety for this problem, because there is no electromagnetic momentum at play. We'll see setups where it does matter in **E7**.

[3] **Problem 5.** USAPhO 2012, problem B2.

Idea 2

Not all motional emfs can be found using $\mathcal{E} = -d\Phi/dt$. Sometimes, for more complex geometries where there is no clear "loop", it's easier to go back to the Lorentz force law.

[3] **Problem 6.** A wheel of radius R and moment of inertia J consisting of a large number of thin conducting spokes is free to rotate about an axle. A brush always makes electrical contact with one spoke at a time at the bottom of the wheel.



A battery with voltage V feeds current through an inductor L, into the axle, through the spoke, to the brush. There is a uniform magnetic field $\mathbf{B}$ pointing into the plane of the paper. At time $t = 0$ the switch is closed.

(a) Find the torque on the wheel and the motional emf along a spoke, as a function of the current I in the circuit and the angular velocity ω of the wheel.

(b) Solve for the full time evolution of $I(t)$ and $\omega(t)$. If there is a small amount of friction and resistance, then what will the final state of the system be?

This setup is an example of a homopolar motor.

Solution. (a) By integrating the force along the wire, the torque is

$$\tau = \int_0^R IBr\,dr = \frac{IBR^2}{2}.$$

Similarly, the force per unit charge on a charge a distance r from the center of the disk is $vB = \omega rB$, so the motional emf is

$$\mathcal{E} = \int_0^R \omega Br\,dr = \frac{\omega BR^2}{2}.$$

(b) Newton's second law gives the time evolution of the wheel,

$$J\dot{\omega} = \frac{IBR^2}{2}.$$

Kirchoff's loop rule gives the time evolution of the circuit,

$$V = L\dot{I} + \frac{\omega BR^2}{2}.$$

When the switch is closed, the current and the wheel will start spinning up simultaneously. However, eventually the wheel will be rotating so fast that the back-emf starts to decrease the current through the inductor. After a while, this current goes *negative* and starts to slow down the wheel. Finally, once the wheel slows down enough, the current through the inductor can start increasing again. Then it turns out that both ω and I go to zero, and the process starts again. In other words, both ω and I oscillate in time.

Now let's see this quantitatively. Differentiating the first equation and plugging it into the second gives

$$\ddot{\omega} + \frac{1}{JL}\left(\frac{BR^2}{2}\right)^2 \omega = \frac{BR^2}{2JL}V$$

which is a simple harmonic motion equation with solution

$$\omega(t) = C\cos\Omega t + D\sin\Omega t + \frac{2V}{BR^2}, \quad \Omega = \frac{BR^2}{2\sqrt{JL}}.$$

We know that $\omega(0) = 0$, and that initially $I = 0$, which implies $\dot{\omega}(0) = 0$. Therefore,

$$\omega(t) = \frac{2V}{BR^2}(1 - \cos\Omega t).$$

Plugging this into Kirchoff's loop rule gives

$$I(t) = \frac{V}{\Omega L}\sin\Omega t.$$

Note that if there were a tiny bit of friction or resistance, then eventually these oscillations would damp out. We would then approach the steady state solution, which is where the back-emf balances the battery's emf and almost current flows at all, $I \approx 0$ and $\omega \approx 2V/BR^2$. (Or, if we used the motor to do work, then in the steady state the current would be nonzero and the angular velocity would be somewhat lower.)

[4] **Problem 7.** IPhO 1990, problem 2. A neat problem on an exotic propulsion mechanism called an electrodynamic tether, which also reviews **M6**.

2 Faraday's Law

Idea 3

Faraday's law states that even for a time-dependent magnetic field, we still have

$$\mathcal{E} = -\frac{d\Phi}{dt}.$$

In the case where the loop isn't moving but the magnetic field is changing, the emf is entirely provided by the electric field,

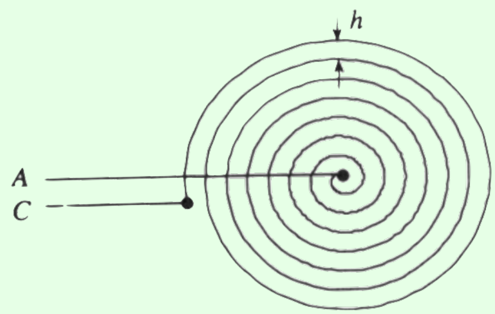
$$\mathcal{E} = \oint_C \mathbf{E} \cdot d\mathbf{s}.$$

Electric fields in the presence of changing magnetic fields can thus be nonconservative, i.e. they can have a nonzero closed line integral, a situation we haven't seen in any previous problem set. The differential form of Faraday's law is

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

Example 3

A flat metal spiral, with a constant distance h between coils, and $N \gg 1$ total turns is placed in a uniformly growing magnetic field $B(t) = \alpha t$ perpendicular to the plane of the spiral.



Find the emf induced between points A and C.

Solution

In theory, you can imagine connecting A and C and finding the flux through the resulting loop, but this is hard to visualize. A better way is to imagine turning the spiral into N concentric circles, connected in series. Then the emf is the sum of the emfs through each,

$$\mathcal{E} = \sum_{k=1}^{N} \pi (kh)^2 \alpha \approx \pi h^2 \alpha \int_0^N dk\, k^2 = \frac{\pi}{3} h^2 N^3 \alpha.$$

To see why this is valid, remember that the emfs are due to a nonconservative electric field, integrated along the length of the loop. Deforming it into a bunch of concentric circles doesn't significantly change $\mathbf{E} \cdot d\mathbf{s}$ along it, because N is large, so it doesn't change the answer much.

Remark: EMF vs. Voltage

We mentioned earlier in **E2** that we often care about electromotive forces, which just mean any forces that act on charges to push them around a circuit. The force due to a nonconservative electric field is another example.

When nonconservative electric fields are in play, the idea of "voltage" breaks down entirely, because you can't define it consistently. However, electrical engineers use a more pragmatic definition of voltage: to them, voltage is just whatever a voltmeter displays. In other words, what they call voltage is what we call electromotive force. This tends to lead to long and bitter semantic disputes, along with rather nonintuitive results, as you'll see below. For example, the "voltage" can be different for different voltmeters even if they are connected at the same points!

Despite this trouble, we'll go along with the standard electrical engineer nomenclature and refer to these emfs as voltages in later problem sets. For example, Kirchoff's loop rule should properly say that the sum of the voltage drops along a loop is not zero, but rather $-d\Phi/dt$. But it is conventional to move it to the other side and call it a "voltage drop" of $d\Phi/dt$.

Remark

When we apply Faraday's law, we often use Ampere's law (without the extra displacement current term) to calculate the magnetic field. This is not generally valid, but works if the currents are in the slowly changing "quasistatic" regime, which means radiation effects are negligible. All the problems below assume this, but we'll see more subtle examples in **E7**.

[2] **Problem 8** (Purcell 7.6)**.** An infinite cylindrical solenoid has radius R and n turns per unit length. The current grows linearly with time, according to $I(t) = Ct$. Assuming the electric field is cylindrically symmetric and purely tangential, find the electric field everywhere.

Solution. Note that $B = \mu_0 n I$ in the solenoid, so that the flux through a loop of radius r varies as

$$\frac{d\Phi_B}{dt} = \mu_0 n C \pi \times \begin{cases} r^2 & r < R \\ R^2 & r > R \end{cases}.$$

By assumption, the electric field is

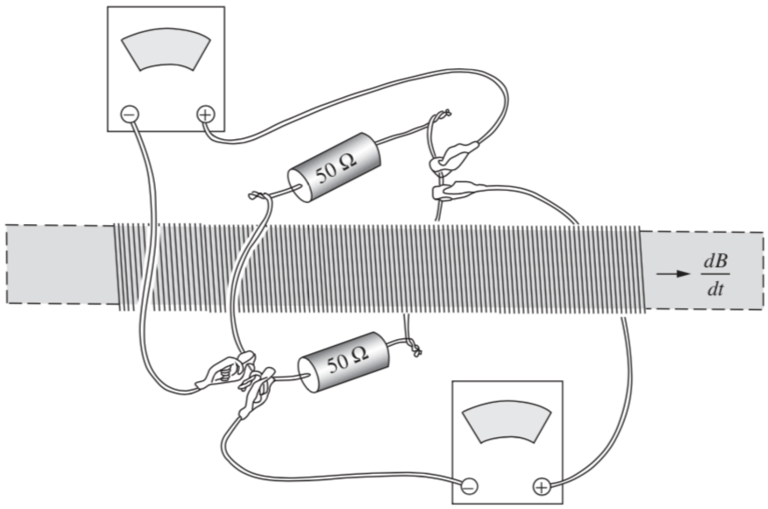
$$\mathbf{E} = E(r)\,\hat{\boldsymbol{\phi}}$$

and the emf is $2\pi s E$, so we conclude

$$E(r) = \frac{1}{2}\mu_0 n C \times \begin{cases} r & r < R \\ R^2/r & r > R \end{cases}.$$

Note that we had to assume $\mathbf{E} = E(r)\,\hat{\boldsymbol{\phi}}$. It's impossible to derive that from Maxwell's equations, because it's not true in general; as discussed in **E1**, you can get different results if you had different boundary conditions (such as the solenoid being inside a giant capacitor) or different initial conditions (such as somebody shining electromagnetic radiation on the solenoid using a flashlight). But this is the solution you get if none of that "extra" stuff is around.

[2] **Problem 9** (Purcell 7.4)**.** Two voltmeters are attached around a solenoid with magnetic flux Φ.



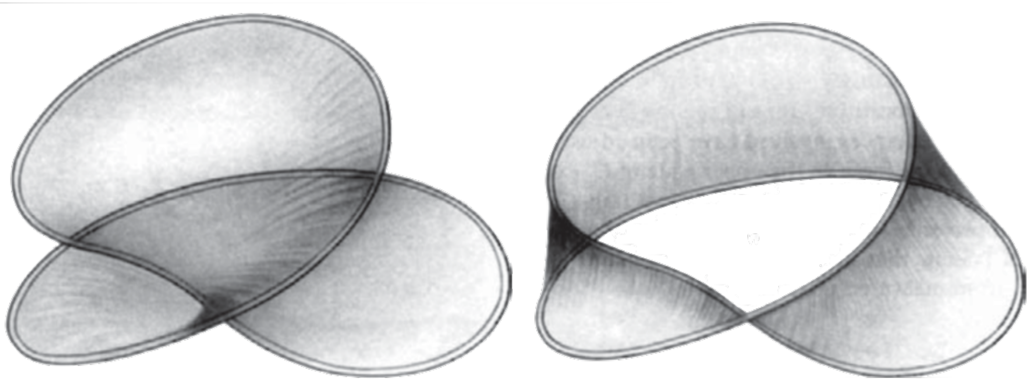
Find the readings on the two voltmeters in terms of $d\Phi/dt$, paying attention to the signs.

Solution. Let the resistance of each resistor be R. The current in the center loop with the two resistors is $I_0 = (d\Phi/dt)/(2R)$, and the emf across each resistor is $\mathcal{E}_0 = I_0 R = (1/2)d\Phi/dt$.

Each voltmeter is connected across one resistor. Now consider the loop formed by one voltmeter's wires, and the half of the center loop closest to it. There is no changing flux through either of these loops, so the integral of $\mathbf{E} \cdot d\mathbf{s}$ around them is zero. Thus, the emf across the resistor is balanced by the emf across the voltmeter, so each voltmeter reads $\pm\mathcal{E}_0$.

The subtlety is in the signs. Suppose that $d\Phi/dt$ is positive, as indicated in the diagram. Then the induced current in the top resistor is rightward, which means the right end of the resistor is at lower potential, which means the top voltmeter reads $-\mathcal{E}_0$. But the induced current in the bottom resistor is leftward, so by similar reasoning, the bottom voltmeter reads $\mathcal{E}_0$. So different voltmeters, with the same probes connected at the same points, can give different results!

[2] **Problem 10** (Purcell 7.28)**.** [A] Consider the loop of wire shown below.



Suppose we want to calculate the flux of $\mathbf{B}$ through this loop. Two surfaces bounded by the loop are shown above. Which, if either, is the correct surface to use? If each of the two turns in the loop are approximately circles of radius R, then what is the flux? Generalize to an N-turn coil.

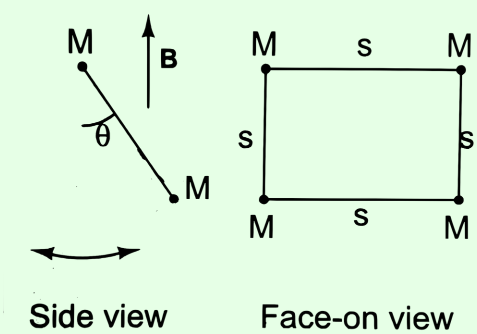
Solution. Remember that in the definition of the magnetic flux, one needs to define a normal vector $d\mathbf{S}$. This is arbitrary, since for any point on a surface there are two normal vectors which point in opposite directions. Applying Faraday's law requires making a consistent choice.

However, some surfaces are *nonorientable*, which means it is *impossible* to define the normal vector on the surface continuously. Concretely, what happens is that if we draw a normal vector at some point (arbitrarily picking up or down), and continuously extend this definition around the surface, we come back to the same point but with the normal vector pointing in the opposite direction. The surface on the right, which is a Mobius strip, has exactly this problem. For this reason, we can't define the flux through it at all! In order to apply Faraday's law (or Gauss's law, etc.) we always have to use orientable surfaces like the one on the left. Thankfully, for any closed loop, an orientable surface whose boundary is the loop always exists; it's called a Seifert surface.

The flux through the left surface is about $2\pi R^2 B$. In general, for N turns, we would get a flux of about $N\pi R^2 B$, though this gets hard to visualize in terms of surfaces.

Example 4

A square, rigid loop of wire has resistance R, sides of length s, and negligible mass. Point masses of mass M are attached at each corner. The top edge of the square loop is mounted so it is horizontal, and the loop may rotate as a frictionless pendulum about a fixed axis passing through this edge. Initially the pendulum is at rest at $\theta = 0$, and a uniform magnetic field $\mathbf{B}$ points horizontally through the loop. The magnetic field is then quickly rotated to the vertical direction, as shown.



Describe the subsequent evolution.

Solution

The rotation of the magnetic field provides a sharp impulse that causes the pendulum to start swinging. Letting ϕ be the angle of the field to the horizontal,

$$\mathcal{E} = -\frac{d(B_x s^2)}{dt} = -Bs^2 \frac{d(\cos\phi)}{dt}$$

and the torque about the axis of rotation is

$$\tau = (IsB_y)s = -\frac{s^4 B^2}{R} \sin\phi \frac{d(\cos\phi)}{dt}.$$

The total impulse delivered is

$$L = \int \tau \, dt = \frac{s^4 B^2}{R} \int_0^{\pi/2} \sin^2\phi \, d\phi = \frac{\pi}{4} \frac{s^4 B^2}{R}$$

which causes an initial angular velocity $\omega = L/(2Ms^2)$.

After the pendulum begins swinging, the presence of the magnetic field causes an effective drag force. To see this, note that now we have

$$\mathcal{E} = -Bs^2 \frac{d(\sin\theta)}{dt}$$

which implies

$$\tau = Is^2 B\cos\theta = -\frac{s^4 B^2}{R}\cos^2\theta\,\frac{d\theta}{dt}.$$

Therefore, the $\tau = I\alpha$ equation is

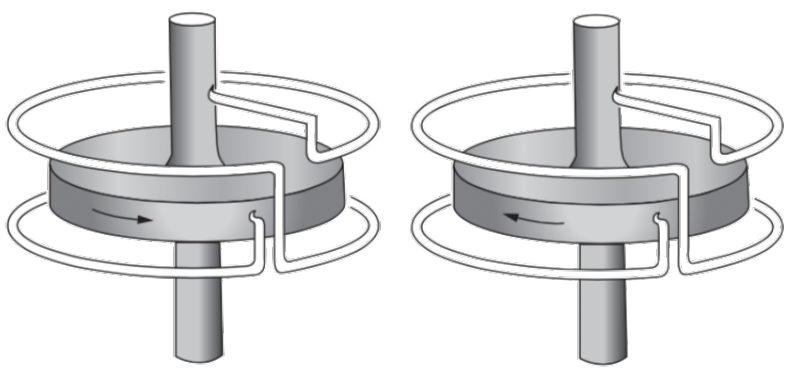
$$2Ms^2\frac{d^2\theta}{dt^2} = -2Mgs\,\sin\theta - \frac{B^2 s^4}{R}\cos^2\theta\,\frac{d\theta}{dt}.$$

If we take the small angle approximation, then we recover ordinary damped harmonic oscillations, as covered in **M4**.

[3] **Problem 11.** USAPhO 2009, problem A1.

[3] **Problem 12.** USAPhO 1999, problem B2.

[3] **Problem 13** (Purcell)**.** A dynamo is a generator that works as follows: a conductor is driven through a magnetic field, inducing an electromotive force in a circuit of which that conductor is part. The source of the magnetic field is the current that is caused to flow in that circuit by that electromotive force. An electrical engineer would call it a self-excited dynamo. One of the simplest dynamos conceivable is shown below.



It has only two essential parts. One part is a solid metal disk and axle which can be driven in rotation. The other is a two-turn "coil" which is stationary but is connected by sliding contacts, or "brushes", to the axle and to the rim of the revolving disk.

(a) One of the two devices pictured is, at least potentially, a dynamo. The other is not. Which is the dynamo?

A dynamo like the one above has a certain critical speed ω_0. If the disk revolves with an angular velocity less than ω_0, nothing happens. Only when that speed is attained is the induced $\mathcal{E}$ enough to make the current enough to make the magnetic field enough to induce an $\mathcal{E}$ of that magnitude. The critical speed can depend only on the size and shape of the conductors, the conductivity σ, and the constant μ_0. Let d be some characteristic dimension expression the size of the dynamo, such as the radius of the disk in our example.

(b) Show by a dimensional argument that ω_0 must be given by a relation of the form $\omega_0 = K/\mu_0\sigma d^2$ where K is some dimensionless numerical factor that depends only on the arrangement and relative size of the parts of the dynamo.

(c) Demonstrate this result again by using physical reasoning that relates the various quantities in the problem (R, $\mathcal{E}$, E, I, B, etc.). You can ignore all numerical factors in your calculations and absorb them into the constant K.

For a dynamo of modest size made wholly of copper, the critical speed would be practically unattainable. It is ferromagnetism that makes possible the ordinary DC generator by providing a magnetic field much stronger than the current in the coils, unaided, could produce. For an Earth-sized dynamo, however, the critical speed is much smaller. The Earth's magnetic field is produced by a nonferromagnetic dynamo involving motions in the fluid metallic core.

Solution. (a) We claim that the second device is a dynamo. Say a current is flowing that starts at the top contact point, flows down through the rod, then flows through the disk into the other contact point, then flows back. One can check that due to the wires, the magnetic field in the disk is pointing down. Now, as the charge passes into the disk, it has some tangential velocity due to the rotation, and in the first case, the $q\mathbf{v} \times \mathbf{B}$ force is pointing opposed to the flow of current, and in the second case, it is pointing in the same direction as the current. Therefore, a current is sustainable only in the second one.

(b) This is very routine; all you need to do is find the dimensions of σ, d, μ_0 and verify that the given combination is the only one that works.

(c) The key point is that there is resistance, which could hinder charge movement. The resistance goes like $R \sim 1/\sigma d$, so we have $V = IR$, or $Ed \sim I/\sigma d$, so $E \sim I/\sigma d^2$. We have $B \sim \mu_0 I/d$, and $E = vB \sim (\omega_0 d)\mu_0 I/d$. Therefore,

$$I/\sigma d^2 \sim \omega_0\mu_0 I,$$

or $\omega_0 \sim 1/\mu_0\sigma d^2$, as desired.

The coupling of motors and generators is a fascinating subject. A crude model of the power grid can be obtained by combining problems 6 and 13 (and also making all the currents AC). For example, when you start a washing machine in your house, AC power is used to start rotating the drum, which produces a back emf that ultimately slows down the rotation of a generator in a power plant. That generator's power is carefully adjusted to produce a reliable 60.0 Hz output frequency.

In reality, the power grid is comprised of many independent generators distributed across thousands of miles. Since they are all connected, they all rotate at roughly the same frequency; any slightly slower ones will be sped up by the rest. This massive entity is sometimes called "the world's largest machine".

[3] **Problem 14.** USAPhO 2023, problem B1. A nice problem on a particular kind of motor, which reviews almost everything covered above in this problem set.

[3] **Problem 15** (MPPP 178)**.** In general, a magnet moving near a conductor is slowed down by induction effects. Suppose that inside a long vertical, thin-walled, brass tube a strong permanent magnet falls very slowly due to these effects, taking a time t to go from the top to the bottom.

(a) Let the magnet have mass m, and let the tube have resistivity ρ, thickness r, and length L. Suppose both the magnet and tube have radius approximately R, and let the magnet's length also be of order R. Let the typical magnetic fields produced at the magnet's surface have magnitude B_0. Find an estimate for t, up to dimensionless constants.

(b) If the experiment is repeated with a copper tube of the same length but a larger diameter, the magnet takes a time t' to fall through. How long does it take for the magnet to fall through the tubes if they are fitted inside each other? Neglect the mutual inductance of the tubes.

Solution. (a) The magnetic flux through a horizontal slice of the tube near the magnet changes by order B_0R^2 when the magnet moves through a distance of about R, so

$$\mathcal{E} \sim \frac{B_0R^2}{R/v} \sim B_0vR.$$

This flux change mostly happens in a vertical section of tube of length about R. The resistance of this section of tube is

$$R \sim \rho\frac{R}{rR} \sim \frac{\rho}{r}.$$

Therefore, the power dissipated is

$$P \sim \frac{\mathcal{E}^2}{R} \sim \frac{B_0^2v^2R^2r}{\rho}.$$

In the steady state, this balances the rate of dissipation of gravitational potential energy, $P \sim mgv$. Combining these gives

$$v \sim \frac{mg\rho}{R^2B_0^2r}.$$

This gives the estimate

$$t \sim \frac{L}{v} \sim \frac{LR^2B_0^2r}{mg\rho}.$$

This is a pretty rough estimate; for a quantitative treatment, see this paper.

(b) Assuming the tubes contain independent eddy currents, we can think of them as just two resistors in parallel. In parallel, the inverse resistivity $1/\rho$ adds. But $t \propto 1/\rho$, which means t adds. Thus, the new time is simply $t+t'$. (This is a bit slick, but if you're concerned about its correctness you can also just run through the derivation in part (a) twice to get the result.)

Remark

In this problem set, we presented motional emf first, and emf from a changing magnetic flux second. But historically, it went the other way around, as described here. Maxwell was aware of Faraday's experiments, which stated that $\mathcal{E} = -d\Phi/dt$ for stationary loops. He then demanded that this remain true for moving loops, and deduced that there must be a force per charge of $\mathbf{v} \times \mathbf{B}$. That is, Maxwell used Faraday's law to derive the Lorentz force! This is a reminder that the process of discovery is messy. When new physics is being found, the very same fact could be a law, a derived result, or simply true by definition, depending on where you start from. And it's not clear which it'll end up being until the dust settles.

3 Inductance

Idea 4: General Inductance

Consider a set of loops with fluxes Φ_i and currents I_i. By linearity, they are related by

$$\Phi_i = \sum_j L_{ij} I_j$$

where the L_{ij} are called the coefficients of inductance. It can be shown that $L_{ij} = L_{ji}$, and we call this quantity the mutual inductance of loops i and j. By Faraday's law, we have

$$\mathcal{E}_i = \sum_j L_{ij} \dot{I}_j.$$

In contrast with capacitance, we're usually concerned with the self-inductance $L_i = L_{ii}$ of single loops; these inductors provide an emf of $L\dot{I}$ each. However, mutual inductance effects can also impact how circuits behave, as we'll see in **E6**.

Remark

The inductance coefficients are similar to the capacitance coefficients in **E2**, but more useful. For capacitors, we are typically interested in configurations with one positive and one negative plate, and the capacitance of this object is related to all of the capacitance coefficients in a complicated way, as we saw in **E2**. But most inductors just use self-inductance, so the inductance we care about is simply one of the coefficients, L_{ii}. Moreover, the "mutual inductance" coefficients L_{ij} are also in the right form to be directly used, since they tell us how current changes in one part of the circuit impact emfs elsewhere.

A more general way to describe the difference is that $\mathcal{E}$ and $\dot{I}$ are directly measurable and controllable quantities, while the Q and V (i.e. the voltage relative to infinity) that the capacitance coefficients relate are less so.

Idea 5

The energy stored in a magnetic field is

$$U = \frac{1}{2\mu_0} \int B^2 \, dV$$

which implies the energy stored in an inductor is

$$U = \frac{1}{2} L I^2$$

where L is the self-inductance.

Example 5

Compute the self-inductance of a cylindrical solenoid of radius R, length $H \gg R$, and n turns per length.

Solution

One straightforward way to do this is to use the magnetic field energy. We have

$$U = \frac{1}{2\mu_0}(\mu_0 n I)^2(\pi R^2 H)$$

and setting this equal to $LI^2/2$ gives

$$L = \pi \mu_0 n^2 R^2 H = \mu_0 N^2 \frac{\pi R^2}{H}$$

where N is the total number of turns.

We can also try to use the definition of inductance directly, $\Phi = LI$. But it's hard to imagine a surface bounded by the solenoid wires; as we saw in problem 10, even the case $N = 2$ is tricky! Instead it's better to use the form $\mathcal{E} = L\dot{I}$. We can then compute the emf across each turn of the solenoid individually, then add them together.

To compute the emf across one turn, we can replace it with a circular loop; this is valid because the emf ultimately comes from the local electric field, which shouldn't change too much if we deform the loop in this way. Then

$$|\mathcal{E}_{\text{loop}}| = \frac{d\Phi}{dt} = (\mu_0 n \dot{I})(\pi R^2).$$

The inductance is hence

$$L = \frac{N\mathcal{E}_{\text{loop}}}{\dot{I}} = (\mu_0 n N)(\pi R^2) = \mu_0 N^2 \frac{\pi R^2}{H}$$

as expected.

Example 6

Find the outward pressure at the walls of the solenoid in the previous example.

Solution

An outward pressure exists because of the Lorentz force of the the axial magnetic field of the solenoid acting on the circumferential currents at the walls. The force per length acting on a wire is IB, and the pressure is this quantity times the turns per length, so naively

$$P = (\mu_0 n I)(nI).$$

However, this is off by a factor of 2. To see why, consider a small Amperian rectangle that straddles the surface of the solenoid. The currents near this rectangle contribute axial magnetic fields of $\mu_0 nI/2$ inside and $-\mu_0 nI/2$ outside. Thus, the currents due to the entire rest of the solenoid contribute $\mu_0 nI/2$ both inside and outside. Since a wire can't exert a force on itself, only the latter field matters, so the true answer is

$$P = \frac{1}{2}\mu_0 n^2 I^2 = \frac{B^2}{2\mu_0}.$$

Remark: Electromagnetic Stress

The above example is like the one in **E1**, where we showed that the inward pressure on a conductor's surface due to electrostatic forces is $\epsilon_0 E^2/2$. In fact, there's a general principle behind both: electric and magnetic fields carry a tension per unit area (i.e. a negative pressure) of magnitude $\epsilon_0 E^2/2$ or $B^2/2\mu_0$ along their directions, and a repulsion per unit area (i.e. a positive pressure) $\epsilon_0 E^2/2$ or $B^2/2\mu_0$ perpendicular to their directions. Charges and currents, such as at the walls of a solenoid or the plates of a capacitor, cause discontinuities in $\mathbf{E}$ or $\mathbf{B}$ across them, leading to a net force on them.

This isn't mentioned in introductory electromagnetism books because the proper treatment of anisotropic pressure requires tensors. However, more advanced books will introduce the Maxwell stress tensor, from which the results above can be read off.

The great experimentalist Michael Faraday was a huge fan of these results. He viewed field lines as physical objects, which he called "lines of force", that carried tension along their lengths and repelled each other. He even presciently suggested that light consisted of waves propagating along lines of force, like waves on a string.

These days, we don't ascribe so much importance to field lines. The fundamental object is the field itself, and field lines are a secondary construction that often just add mathematical complication. For example, the field of a dipole is simple, but it's not so simple to solve for the corresponding field lines. Things get even more complicated in dynamic situations, where field lines can appear and disappear; Faraday viewed induction as a result of "cutting" magnetic field lines. And in **R3**, we'll show how fields transform between frames, which implies that the very existence of a field line can depend on the reference frame. Still, Faraday's intuition might be helpful occasionally, and it's still a useful tool in some subfields. For instance, in plasma physics, field lines can be used to visualize magnetic reconnection.

Remark

Recall the example in **E1** involving the force between two spherical balls of charge. There, we got the answer using a slightly tricky argument, where Newton's third law allowed us to use the shell theorem twice. But the idea of electromagnetic stress provides a straightforward alternative proof which also works for more general situations.

Suppose the two balls lie above and below the xy plane, and additional external forces hold them both at rest. Consider as a system everything at $z > 0$, which includes the second ball and a lot of empty space. The only external forces on this system are from the attractive $\epsilon E^2/2$ pressure at the xy plane, and the force F which holds the second ball in place. Since the momentum of the system is constant, these forces must cancel. Thus, to compute F we only need to know the electric field on the xy plane, for which we can clearly apply the shell theorem to both balls, replacing them with point charges.

The general idea is that electromagnetic forces on objects can be determined solely by the electromagnetic fields in the space around them. You can use this to find shorter solutions to some problems in **E1**.

[3] **Problem 16.** Consider a toroidal solenoid with a rectangular cross section of height h and width w, N turns, and inner radius R.

(a) Find the self-inductance by considering the magnetic flux.

(b) Now suppose the current increases at a constant rate dI/dt. Find the magnitude of the electric field at a height z above the center of the solenoid, assuming $h, w \ll R$. (Hint: write down the divergence and curl of $\mathbf{E}$ in terms of $\dot{\mathbf{B}}$ in general, and notice the similarities to the equations for $\mathbf{B}$ in terms of $\mathbf{J}$. This allows us to use the ideas of **E3** by analogy.)

(c) Verify that the two formulas for energy given in idea 5 are consistent in this setup.

Solution. (a) Symmetry and Ampere's law imply that the field inside is $B = \frac{\mu_0 N I}{2\pi r}$ pointing in the $\hat{\boldsymbol{\phi}}$ direction. Hence the flux through a single loop is

$$\frac{\mu_0 N I h}{2\pi} \log \frac{R+w}{R}.$$

The inductance is

$$L = \frac{N\Phi}{I} = \frac{\mu_0 N^2 h}{2\pi} \log \frac{R+w}{R}.$$

(b) We know $\dot{\mathbf{B}}$ and want to find $\mathbf{E}$, which in this problem is determined by the equations

$$\nabla \times \mathbf{E} = -\dot{\mathbf{B}}, \quad \nabla \cdot \mathbf{E} = 0.$$

But we already know how to solve problems of this form, because the equations that govern the magnetic field in magnetostatic situations are

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \quad \nabla \cdot \mathbf{B} = 0.$$

So if we formally define a current $\mathbf{J}'$ by $-\dot{\mathbf{B}} = \mu_0 \mathbf{J}'$, and use the Biot–Savart law to solve for the corresponding $\mathbf{B}'$, then that quantity will be precisely the $\mathbf{E}$ we're looking for.

Now, $-\dot{\mathbf{B}}$ is localized within the toroid, so since we're assuming the toroid is thin, we can approximate $\mathbf{J}'$ as a ring of current $I' = L\dot{I}/\mu_0$. The resulting magnetic field is

$$\mathbf{B}' = \frac{\mu_0 I'}{2} \frac{R^2}{(z^2+R^2)^{3/2}} \hat{\mathbf{z}} = \frac{L\dot{I}}{2} \frac{R^2}{(z^2+R^2)^{3/2}} \hat{\mathbf{z}}$$

and this is the electric field we're looking for. (The sign isn't determined since the problem didn't specify which way the current in the solenoid went, but this doesn't affect the magnitude.)

(c) Let's compute the total energy of the magnetic field. The magnetic field outside the solenoid is zero, and the magnetic field inside is $B = \mu_0 NI/2\pi r$. Now consider cylindrical shells of radius r, thickness dr, and volume $dV = 2\pi rh\, dr$. The field energy is

$$U = \int \frac{B^2}{2\mu_0} dV = \int_R^{R+w} \frac{(\mu_0 NI)^2}{2\mu_0(2\pi r)^2}(2\pi rh\, dr) = \frac{\mu_0 N^2 I^2 h}{4\pi} \int_R^{R+w} \frac{dr}{r} = \frac{\mu_0 N^2 I^2 h}{4\pi} \log \frac{R+w}{R}.$$

Referring to the result of part (a), the expression $U = LI^2/2$ yields the same result.

Remark

In electromagnetism, we often have issues with divergences when we take idealized point sources. For example, the voltage near a point charge can become arbitrarily high. Similarly, the magnetic field diverges as you approach an idealized, infinitely-thin wire, which causes the self-inductance of wire loops to diverge. Of course, the resolution is that you don't actually get an infinite magnetic field as you approach a wire. A real wire has finite thickness, and its magnetic field instead goes to zero as you approach its center. (We didn't run into this problem for solenoids, because we modeled their wires as a uniform sheet of current, whose magnetic field isn't singular at all.) If a problem does involve a wire loop, it'll often circumvent this messy issue by just giving the self-inductance from the start.

[2] **Problem 17.** A wire of length ℓ is bent into a long "hairpin" shape, with two parallel straight edges of length $\ell/2$ separated by a distance $d \ll \ell$.

(a) Write down an integral expression for the self-inductance, neglecting the curved parts, and show that it diverges.

(b) Find a rough estimate for the self-inductance by taking the wire to have radius $r \ll d$ and ignoring any flux through the wire itself.

Solution. (a) We see the flux to be

$$\Phi = \int_0^d \frac{\mu_0 I}{2\pi} \left(\frac{1}{x} + \frac{1}{d-x} \right) \frac{\ell}{2}\, dx$$

which is infinite, because the integral of $1/x$ is logarithmically divergent.

(b) We replace the flux integral with

$$\Phi = \int_r^{d-r} \frac{\mu_0 I}{2\pi} \left(\frac{1}{x} + \frac{1}{d-x} \right) \frac{\ell}{2}\, dx.$$

This gives an inductance of

$$L = \frac{\Phi}{I} = \frac{\mu_0 \ell}{2\pi} \log\left(\frac{d-r}{r} \right) \approx \frac{\mu_0 \ell}{2\pi} \log(d/r).$$

This still diverges in the $d \to 0$ limit, as it should, but the presence of the logarithm means that the inductance doesn't depend that strongly on d, for realistic values. That's why we can often get away with not mentioning the details of the wire; you'll get a similar answer as long as it's thin.

[3] **Problem 18.** Consider two concentric rings of radii r and $R \gg r$.

(a) Compute the mutual inductance by considering a current through the larger ring.

(b) Compute the mutual inductance by considering a current through the smaller ring, and verify your results agree. (Hint: this can be done without difficult integrals.)

In general, computing mutual inductance is a hard and practically important problem; there have been whole books written on the subject.

Solution. (a) The field at the center is $\mu_0 I/2R$, so the flux through the small ring is $\Phi = (\mu_0 I/2R)\pi r^2$, so $L_{12} = \frac{\mu_0 \pi r^2}{2R}$.

(b) Consider the entire infinite plane the smaller ring lies in. The key idea is that the total flux through this plane is zero: every magnetic flux line due to the ring that goes up through the plane comes down through it somewhere else. Now decompose this plane into the part in the big ring and the part outside. We have

$$\Phi_{\text{plane}} = \Phi_{\text{in}} + \Phi_{\text{out}} = 0$$

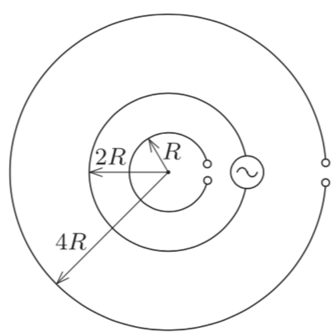
which means $\Phi_{\text{in}} = -\Phi_{\text{out}}$. This is useful because calculating Φ_{in} is very complicated. Calculating Φ_{out} is easy because the whole region is far from the small ring, so its field can be approximated as a dipole field.

The field at a distance s away in the plane of the rings is $\mathbf{B} = B_z \hat{\mathbf{z}}$ where $B_z = \frac{\mu_0 I \pi r^2}{4\pi s^3}$. The flux is then

$$\Phi_{\text{in}} = -\Phi_{\text{out}} = -\frac{\mu_0 I r^2}{4\pi} \int_R^\infty \frac{1}{s^3} 2\pi s \, ds = \frac{\mu_0 I \pi r^2}{2R}$$

Dividing by I gives the same result as (a), as expected.

[2] **Problem 19** (MPPP 181). Three nearly complete circular loops, with radii R, $2R$, and $4R$ are placed concentrically on a horizontal table, as shown.



A time-varying electric current is made to flow in the middle loop. Find the voltage induced in the largest loop at the moment when the voltage between the terminals of the smallest loop is V_0.

Solution. As you might have noticed from the previous problems, inductances generally scale with one power of length. In this case, the only length scale is the radii of the rings, so the mutual inductance of the $2R$ and $4R$ loops is twice that of the $2R$ and R loops by dimensional analysis. Since $V_0 = L_{R,2R} \dot{I}_{2R}$, we have $V_{4R} = L_{2R,4R} \dot{I}_{2R} = 2V_0$.

But there's a subtlety: in problem 17, we found that the self-inductance of a wire loop depended on the radius r of the wire, so why didn't we allow it in the dimensional analysis here? The point is that the self-inductance of a wire loop depends on the flux the loop puts through itself, so it depends on the very large magnetic fields very closer to the wire. But a mutual inductance only depends on the flux one loop puts through another, and here, all points of the second loop are far away from the first loop. So the wire radius doesn't matter, just like how it didn't in problem 18.

4 Magnetism

In this section we'll dip a little into atomic physics and the origin of magnetism. However, a proper understanding of this subject requires quantum mechanics, as we'll cover in **T3** and **X3**.

Idea 6

A spinning charged object carries a magnetic dipole moment $\boldsymbol{\mu}$ and angular momentum $\mathbf{J}$. If the object's mass and charge distributions are proportional, then $\boldsymbol{\mu}$ and $\mathbf{J}$ point in the same direction, and one can show that their ratio is always $\mu/J = q/2m$.

Example 7

Suppose the magnetic moment of an iron atom is due to a single unpaired electron, with angular momentum of order $\hbar$. The atoms are separated by a distance $d \sim 10^{-10}$ m. Estimate the maximum magnetic field an iron magnet can produce. How does this compare to the fields that can be produced in an electromagnet?

Solution

The answer doesn't scale significantly with the physical size of the iron magnet. To see this, think in terms of electric dipoles: if you have a giant cube of electric dipoles, it's equivalent to having a fixed surface charge density $\pm\sigma$ on two of the faces. The electric field produced by such a charge density near each face is of order σ/ϵ_0, independent of the size of the cube.

Therefore, the only things the magnetic field can depend on are μ_0, the magnetic dipole moment μ of a single atom, and d. By dimensional analysis,

$$B \sim \mu_0 \frac{\mu}{d^3}$$

which can also be thought of as $\mu_0 M$, where M is the magnetization density. Taking $\mu \sim e\hbar/m_e$ and plugging in the numbers gives $B \sim 10\,\mathrm{T}$, which is the right order of magnitude.

Now consider the case of an electromagnet, where the field is produced by moving electrons with typical speed v, moving in a loop with typical size r. In a metal, there's on the order of one free electron per atom, so d is still the same. The difference is that the field made by each electron *does* scale with r, because each has magnetic moment

$$\mu = IA \sim \frac{ev}{r} r^2.$$

Compared to the previous result, this is larger by a factor of $mvr/\hbar$. The two are comparable, for $r \sim 1$ m, if the electrons travel at the agonizingly slow velocity $v \sim 10^{-4}$ m/s.

Therefore, you would get a magnetic field much larger than $10\,\mathrm{T}$ if you could make the electrons go at a reasonable walking speed, but that's easier said than done. The largest steady magnetic fields made in the lab are only about $40\,\mathrm{T}$. Such a field carries a pressure which would rip apart a solenoid made of coiled wire,

$$P = \frac{(40\,\mathrm{T})^2}{2\mu_0} = 0.6\,\mathrm{GPa}.$$

Instead, these fields are produced in Bitter electromagnets, which are solenoids made of thick metal plates, perforated with cooling channels to dissipate the enormous heat produced by resistance. It is possible to produce higher fields temporarily, but the results will be explosive.

[3] **Problem 20.** USAPhO 2021, problem A3. This covers a simple classical model of the electron.

[3] **Problem 21.** USAPhO 2007, problem B2. (Equation 10 of the official solution has a typo.)

[5] **Problem 22.** APhO 2013, problem 3. A solid question involving classical magnetic moments, which gives some intuition for the quantum behavior.

5 Superconductors

There are many tough Olympiad problems involving superconductors. Superconductors can be a bit intimidating at first, but they actually obey simple rules.

Idea 7

An ideal conductor has zero resistivity, which implies that the magnetic flux through any loop in the conductor is constant: attempting to change the flux instantly produces currents that cancel out the change. However, the flux can be nonzero.

A superconductor is an ideal conductor with the additional property that the magnetic field in the body of the superconductor is exactly zero, no matter what the initial conditions are; once an object becomes superconducting it forces all the existing flux out. This is known as the Meissner effect. It further implies that all the current in a superconductor is confined to its surface, and that the normal component of the magnetic field $B_\perp$ is zero on the surface. Many problems involving superconductors don't even use the Meissner effect, so they would also work for ideal conductors.

Example 8: PPP 153

A superconducting uniform spring has N turns of radius R, relaxed length x_0, and spring constant k. The two ends of the spring are connected by a wire, and a small, steady current I is made to flow through the spring. At equilibrium, what is the change in its length?

Solution

This question really is about ideal conductors, not just superconductors. The additional superconductivity property would tell us about the field *inside* the wires themselves (not the loops that the wires form), and thereby about some small screening currents on the surfaces of the wires. This is not important because the wires are thin compared to the spring as a whole.

In order to find the equilibrium length x_{eq}, we can use the principle of virtual work. We compute how the energy changes if we slightly perturb the system. At equilibrium, this change in energy should be zero.

We have $B = \mu_0 NI/x$, so the magnetic field energy is

$$U = \frac{B^2}{2\mu_0} V = \frac{AI^2}{x}, \quad A = \frac{\mu_0 \pi R^2 N^2}{2}.$$

Naively, this means the magnetic field energy decreases as x increases, so the spring would like to stretch. But this makes no sense, because we know that parallel currents attract, squeezing the spring. We have to recall that the spring is an ideal conductor, so when it is stretched or squeezed, the current changes to keep the flux the same. The flux is

$$\Phi_B = N(\pi R^2) B \propto \frac{I}{x}$$

so we have

$$I(x) = I\frac{x}{x_{\mathrm{eq}}}, \quad U(x) = \frac{AI^2}{x_{\mathrm{eq}}^2} x.$$

The other energy contribution is $k(x - x_0)^2/2$, so setting the derivative of energy to zero,

$$\frac{AI^2}{x_{\mathrm{eq}}^2} = k(x_0 - x_{\mathrm{eq}}).$$

Since the current is small, $x_0 \approx x_{\mathrm{eq}}$, so we can replace x_{eq} with x_0 on the left-hand side, giving the answer,

$$x_{\mathrm{eq}} = x_0 - \frac{AI^2}{x_0^2 k}.$$

As a sidenote, the original formulation of this question involved an external voltage source forcing the current I to be constant. However, in this case using energy conservation is more subtle because one has to account for the work done by the voltage source. Here we used a superconductor, which keeps the flux constant, so that the spring can be thought of as an isolated system. The final answers are the same, since in both cases we have the same magnetic forces, which determine the spring's compression.

Example 9

A long, thin cylinder of radius R is placed in a magnetic field B_0 parallel to its axis. The cylinder originally carries no current on its surface, and it is cooled until it reaches the superconducting state. Find the resulting distribution of current on its surface. Now suppose the external magnetic field is turned off; what is the new current distribution?

Solution

Solving this question requires using both properties. The Meissner effect tells us there is no magnetic field within the body of the cylinder itself (i.e. the region from $r = R$ to $r = R + dr$). The ideal conducting property tells us that the flux through a cross-section of the cylinder (i.e. the region from $r = 0$ to $r = R$) is constant, and hence equal to $\pi R^2 B_0$.

When the cylinder becomes superconducting, the Meissner effect kicks in, and the field within the body of the cylinder can be cancelled by a uniform surface current on the outer surface. By the same logic as we used to compute the field of a cylindrical solenoid, it is

$$K_{\text{out}} = -B_0/\mu_0.$$

To keep the flux constant, a compensating opposite current must appear on the inner surface,

$$K_{\text{in}} = B_0/\mu_0.$$

When we turn off the external magnetic field, the two properties imply

$$K_{\text{out}} = 0, \quad K_{\text{in}} = B_0/\mu_0$$

which you should check if you're not sure.

[3] **Problem 23** (MPPP 182)**.** Two identical superconducting rings are initially very far from each other. The current in the first is I_0, but there is no current in the other. The rings are now slowly brought closer together. Find the current in the first ring when the current in the second is I_1.

Solution. Since the rings are ideal conductors, the flux through each ring is conserved. Now let the mutual inductance be M and the self-inductance be L, and let the final current through the first ring be I_f. The flux through the second ring is still zero, so

$$0 = I_f M + I_1 L.$$

Similarly, conservation of flux through the first ring gives

$$I_0 L = I_f L + I_1 M.$$

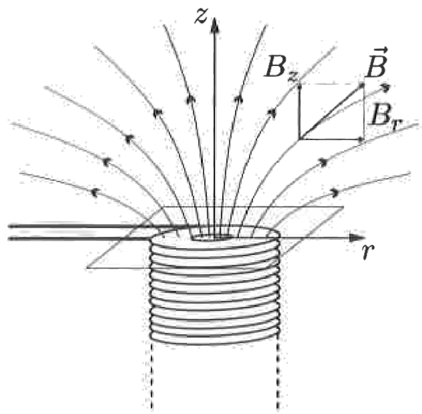
Solving for I_f gives a quadratic equation, with solution

$$I_f = \frac{I_0 + \sqrt{I_0^2 + 4I_1^2}}{2}.$$

[4] **Problem 24** (PPP 182, Russia 2006)**.** A thin superconducting ring of radius r, mass m, and self-inductance L is supported by a piece of plastic just above the top of a long, cylindrical solenoid of radius $R \gg r$ and n turns per unit length. The ring and solenoid are coaxial. When the current in the solenoid is I_s, the magnetic field near the end of the solenoid is

$$B_z = B_0(1 - \alpha z), \quad B_r = B_0 \beta r$$

where we put the origin at the very top of the solenoid.



(a) Find an expression for B_0. (Keep your answers below in terms of B_0 to avoid clutter.)

(b) Find β in terms of α. What are their signs?

(c) Let I be the current through the ring. Suppose that initially $I_s = I = 0$. Find the value I_c of I_s when the ring lifts off the plastic.

(d) Now the piece of plastic is removed and the ring is return to the same position. Initial conditions are set up so that $I_s = I_c$ and $I = 0$. The ring is released from rest. Find its subsequent motion, assuming for simplicity that the expressions for B_z and B_r above always hold. Express your final answers in terms of only α and g.

(e) In reality, the expressions for B_z and B_r break down if the ring moves too far. Consider part (d) again, but now suppose the *exact* expressions for B_z and B_r are used. Without solving any differential equations, will the resulting motion be qualitatively similar or not?

Solution. (a) The magnetic field in the center of the solenoid is $\mu_0 n I_s$. By symmetry and superposition, at the ends of the solenoid, the field at the axis is $\frac{1}{2}\mu_0 n I_s$ since the two ends can be put together to make a field of $\mu_0 n I_s$. (We previously saw this argument in **E3**.) Thus at the top of the solenoid, the field is $B_0 = \frac{1}{2}\mu_0 n I_s$.

(b) From Gauss's law, we know that $\oint \mathbf{B} \cdot d\mathbf{S} = 0$. Consider a cylinder of radius r and height z coaxial with the solenoid. The differences between the flux of the circular faces is $-\pi r^2(B_0 - B_0(1 - \alpha z)) = -\alpha z B_0 \pi r^2$. Since the flux that enters the cylinder must all leave through the side of the cylinder of area $2\pi r z$, we get

$$\alpha z B_0 \pi r^2 = 2\pi r h B_r = 2\pi r^2 z \beta B_0$$

from which we conclude

$$\beta = \alpha/2.$$

Since flux spreads out to the side, $\beta > 0$. Thus, we also have $\alpha > 0$, which makes sense since the axial field should weaken as we get further from the solenoid.

Equivalently, we could have used the differential form $\nabla \cdot \mathbf{B} = 0$, where $B_x = B_0 \beta x$ and $B_y = B_0 \beta y$. Using the form of the divergence in cylindrical coordinates gives $-\alpha B_0 + 2B_0\beta = 0$, which leads to the same result.

(c) The applied flux through the ring will be $\Phi_B = B_z \pi r^2$, and since the ring is an ideal conductor, the induced flux must cancel out the applied flux, since the initial flux was zero. By the definition of inductance, there is a current $I = \Phi_B / L$ in the ring.

The Lorentz force on the ring is $F = 2\pi r I B_r$, and the ring lifts off when it balances gravity,

$$2\pi r B_0 \beta r \frac{B_0(1-\alpha z)\pi r^2}{L} = mg.$$

Now setting $z = 0$ and using the result $\beta = \alpha/2$, we get

$$I_c = \frac{2}{\mu_0 n \pi r^2} \sqrt{\frac{mgL}{\alpha}}.$$

You could also get the same result by thinking in terms of energy, and using a "virtual work" argument, as in example 8. But that's not necessary; in that case, thinking about force was subtle (since there is an entire spring contracting at once) and thinking about energy was simple, while in this case, the energy is subtle (since there are two sources of magnetic field in the problem) and the force is simple.

(d) Now, the flux through the ring stays at $B_0 \pi r^2$, which means the induced current only balances the change in flux due to the change in B_z. By reasoning similar to the previous part, we have a force on the ring of

$$F = 2\pi r B_0 \beta r \frac{B_0(-\alpha z)\pi r^2}{L} - mg.$$

Combining the results of parts (a) and (c) to eliminate B_0 gives the simple result

$$F = mg(\alpha z - 1)$$

which is a simple harmonic motion equation with a shift, so that the equilibrium point is at $z = 1/\alpha$. Since the initial velocity is zero, we conclude

$$z(t) = \frac{1}{\alpha}\left(\cos(\sqrt{\alpha g}\, t) - 1\right).$$

(e) We have $\alpha \sim 1/R$ on dimensional grounds, so the above analysis shows that the ring drops by an amount of order R, i.e. by an amount comparable to the solenoid's width. That makes our assumptions about the forms of B_z and B_r dubious, because they only hold on scales small compared to R. The key question is: when we use the exact forms of B_z, do we still get some kind of oscillation (though not a simple harmonic one), or does the ring not oscillate at all?

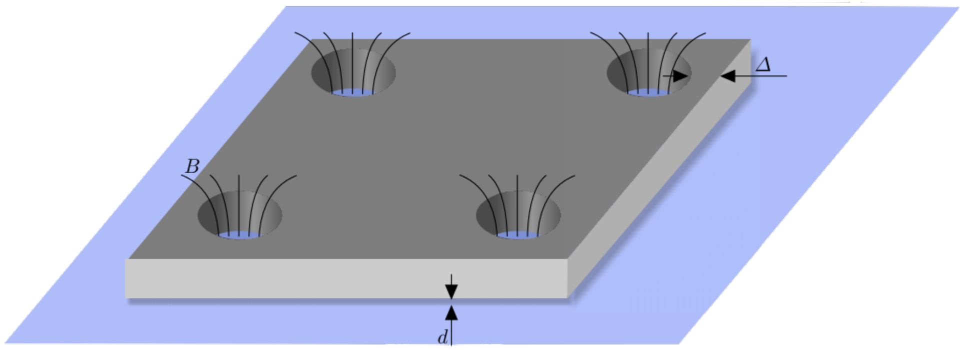
To answer this question, consider the highest upward force that can be exerted on the ring. Once the ring is well inside the solenoid, B_z reaches its maximum value of $2B_0 = \mu_0 n I_c$. At this point, if we still had $B_r = B_0 \beta r$, then the upward magnetic force would be mg, which is just enough to balance gravity. But as the ring goes further inside, the magnetic field spreads out less radially, so B_r is actually lower. Therefore, the upward magnetic force is always less than mg! The ring doesn't actually oscillate; it just falls faster and faster downward.

[4] **Problem 25.** IPhO 2012, problem 1C. A delightfully tricky problem that uses the properties of superconductors in a subtle way.

[4] **Problem 26.** EuPhO 2017, problem 3. Another tricky problem, using ideas we've seen before.

Solution. See the official solutions here. We foreshadowed the basic idea behind this problem, all the way back in **E2**.

[4] **Problem 27** (Physics Cup 2013)**.** A rectangular superconducting plate of mass m has four identical circular holes, one near each corner, a distance Δ from the plate's edges. Each hole carries a magnetic flux Φ. The plate is put on a horizontal superconducting surface. The magnetic repulsion between the plate and the surface balances the weight of the plate when the width of the air gap beneath the plate is $d \ll \Delta$, and d is much smaller than the radii of the holes. The frequency of small vertical oscillations is f_0.



Next, a load of mass M is put on the plate, so that the load lays on the plate, and the plate levitates above the support. What is the new frequency of small oscillations?

Solution. We think about the magnetic field energy present in between the plate and the superconducting surface. Since the field cannot penetrate the superconductor, it simply spreads out, so that it has magnitude $B(r) \sim \Phi/rd$ a distance r from the hole. The field energy is

$$U \sim \int B^2 \, dV \sim \int \frac{dV}{r^2 d^2} \sim \frac{1}{d} \int \frac{r\,dr}{r^2}.$$

The latter integral looks like it diverges, but we recall it is cut off by Λ on the upper end and the radii of the holes on the lower end. The value of the integral doesn't matter, because it's just some constant, and we're only interested in the dependence on d, which is

$$U \sim \frac{1}{d}.$$

This potential provides a restoring force against vertical displacements,

$$F \sim \frac{1}{d^2}.$$

If the total mass is m, that means that $d \sim 1/\sqrt{m}$ at equilibrium. About this equilibrium point, the effective spring constant is given by the derivative of F,

$$k \sim \frac{1}{d^3} \sim m^{3/2}.$$

Therefore, the oscillation frequency is

$$f \sim \sqrt{k/m} \sim m^{1/4}.$$

Therefore the new frequency is $f_0(1 + M/m)^{1/4}$. The fact that we didn't need to know any of the constants involved to get this scaling is a nice property of power-law potentials.

[5] **Problem 28.** IPhO 1994, problem 2. This problem tests your intuition for induction, and is good preparation for **E6**.

Remark

In **E4**, we spent a lot of time applying $F = ma$ to charges. But in this problem set, we were somehow able to find how systems of charges behave using only Maxwell's equations, without ever explicitly referring to the forces on charges. Certainly this information has to be used implicitly somewhere, so what's going on?

To investigate this, let's do a careful derivation of Kirchoff's loop rule, for a series RLC circuit with a battery. By applying the work-kinetic energy theorem to a charge q as it goes around the circuit, from one capacitor plate to the other, we have

$$\int_C \mathbf{E} \cdot d\mathbf{s} + \int_C \mathbf{f} \cdot d\mathbf{s} = \frac{\Delta \mathrm{KE}}{q}$$

where $\mathbf{f}$ is any non-electric force per charge, and the line integrals follow the path C of the charge. By assumption, the battery and resistor contribute

$$\int_C \mathbf{f} \cdot d\mathbf{s} = \begin{cases} \mathcal{E} & \text{battery} \\ -IR & \text{resistor} \end{cases}$$

where the forces are due to chemical reactions (as covered in **E2**) or collisions with the ions (as covered in **E4**). Meanwhile, Faraday's law states

$$\oint \mathbf{E} \cdot d\mathbf{s} = \int_C \mathbf{E} \cdot d\mathbf{s} + \frac{Q}{C} = -\frac{d\Phi_B}{dt} = -\dot{I}L$$

where we need to add on Q/C to close the loop through the capacitor. Thus,

$$\mathcal{E} = \dot{I}L + IR + \frac{Q}{C} + \frac{\Delta \mathrm{KE}}{q}.$$

Now, the key point is that in a conductor, the charges are extremely light and extremely numerous; it only takes a tiny amount of kinetic energy to get an enormous current. Therefore, the energy in any circuit is dominated by the energies stored in the inductor and capacitor, while the kinetic energy of the charges is negligible. We thus set the ΔKE term to zero to get the usual form of Kirchoff's loop rule.

Most books gloss over the derivation of Kirchoff's loop rule; for instance, Halliday, Resnick, and Krane merely prove it in the trivial case of an all-resistor circuit. Unfortunately, most purported "derivations" of it in other sources, or online, are simply wrong. For example, a common claim is that in the absence of inductors, Kirchoff's loop rule is nothing more than the statement that $\oint \mathbf{E} \cdot d\mathbf{s} = 0$. But this doesn't explain how the term Q/C can show up; since the electric field of a capacitor is conservative, its closed line integral always vanishes. The confusion only multiplies once inductors are in play.

As another note, if the work done on the charges is positive in some parts of the circuit, and negative in others, shouldn't the current wildly speed up and slow down as it goes through the wires? No, because as we saw in **E2**, charges strongly repel each other, so charge can't accumulate anywhere. More precisely it's because wires have negligible capacitance; in the fluid flow analogy, the fluid is incompressible.

To illustrate this point, consider a discharging RL circuit, where the inductor has no resistance. As the current in the inductor decreases, it induces an electric field along the inductor wires. The charges in the circuit then redistribute themselves as they flow; as a result, the electric field in the inductor wire is almost completely cancelled, while the induced emf $\dot{I}L$ appears across the resistor. It's just like how it's possible to pull on a massless rope attached to a massive block, even though the net force on a massless object always has to be zero – an internal tension force appears to transfer the force to the block.

Above, I say "almost" because the kinetic energy of the charges does play a small role. In other situations, it's possible for it to have a big effect. For example, if you really had a completely ideal wire loop, with no resistance and no capacitance, and twisted on itself so that it had no inductance, attached to an ideal battery, then the limiting factor which stops the current from becoming infinite is this inertia. The kinetic energy of charges is proportional to $v^2 \propto I^2$, so it acts like a very tiny inductance distributed throughout the wire (known as kinetic inductance), resisting changes in current. You'll see some examples in **ERev** where the motion of charges plays a direct role.

Electromagnetism VI: Circuits

AC circuits are covered in chapter 8 of Purcell, or chapter 10 of Wang and Ricardo, volume 2. Transmission lines, filters, and resonant cavities are covered physically in chapters II-22 and II-23 of the Feynman lectures, which will also build intuition for the next unit. Also see Jaan Kalda's circuits handout, an excellent resource which covers nonlinear circuit elements and much more. This problem set assumes knowledge about linear differential equations covered in **M1** and **M4**, but you can review the relevant material in chapter 4 of Morin. If you'd like to learn much more about circuits, from the electrical engineering perspective, a nice book is *Foundations of Analog and Digital Electronic Circuits* by Agarwal and Lang. There is a total of **84** points.

1 DC RLC Circuits

Idea 1

AC circuits correspond to driven damped oscillators by the analogies

$$Q \leftrightarrow x, \quad I \leftrightarrow v, \quad \dot{I} \leftrightarrow a, \quad L \leftrightarrow m, \quad R \leftrightarrow b, \quad C \leftrightarrow 1/k, \quad V_0 \leftrightarrow F_0.$$

More precisely, Kirchoff's loop equation in an AC circuit immediately becomes Newton's second law for a driven damped oscillator upon making these replacements.

Example 1

Consider a circuit with a battery of emf $\mathcal{E}$, a resistor R, and an inductor L in series, with zero initial current. Find the current $I(t)$ and verify that energy is conserved.

Solution

Kirchoff's loop equation is

$$\mathcal{E} = L\frac{dI}{dt} + IR.$$

To solve for the current, we can separate and integrate, giving

$$\frac{dt}{L} = \frac{dI}{\mathcal{E} - IR}$$

which yields

$$I(t) = \frac{\mathcal{E}}{R}(1 - e^{-(R/L)t}).$$

At long times, the inductor has no effect, since the current stops changing. To verify energy conservation we multiply Kirchoff's loop equation by I, since power is emf times current,

$$I\mathcal{E} = LI\frac{dI}{dt} + I^2R.$$

The left-hand side is the power output by the battery, and the two terms on the right-hand side represent the rate of increase in energy $LI^2/2$ stored in the inductor, and the power dissipated in the resistor, so all power is accounted for.

Example 2

An ideal battery of voltage $\mathcal{E}$ is suddenly connected to an ideal capacitor. After a short time, the capacitor has energy U. How much energy has been released by the battery?

Solution

This is a variant of the two capacitor paradox, which has essentially the same solution. First let's consider how the capacitor is charged up over time. Naively, since there's no resistance or inductance, the current in the circuit instantly becomes infinite, then instantly shuts off. This isn't realistic: to understand what's actually going on, we have to account for nonideal features of the circuit, such as resistance or self-inductance. For example, if the resistance dominates (overdamping), the capacitor charges up monotonically, as in an RC circuit. If the inductance dominates (underdamping), the capacitor voltage oscillates about $\mathcal{E}$, until eventually settling down due to the resistance.

At the end, the total energy on the capacitor is

$$\int V_C \, dQ = \int \frac{V_C}{C} \, dV_C = \frac{\mathcal{E}^2}{2C} = \frac{1}{2}\mathcal{E}Q$$

where Q is the total charge. But the work done by the battery is

$$\int \mathcal{E} \, dQ = \mathcal{E}Q$$

so the battery has released energy $2U$. Evidently, half of it is lost, no matter how close to "ideal" the circuit is! If the tiny resistance dominates, it is lost to heat in the circuit. If the tiny inductance dominates, then LC oscillations result, and energy can also be lost to electromagnetic radiation, as covered in **E7**.

In fact, this is just another example of the nonadiabatic processes you saw in **T1**. Instantly attaching a battery is the same kind of thing as instantly dropping a piston, and letting it bounce until it comes to a stop. Just like those nonadiabatic processes, attaching the battery in this way creates entropy; if the circuit and environment have temperature T, then

$$\Delta S = \frac{\Delta Q}{T} = \frac{\mathcal{E}Q}{2T}.$$

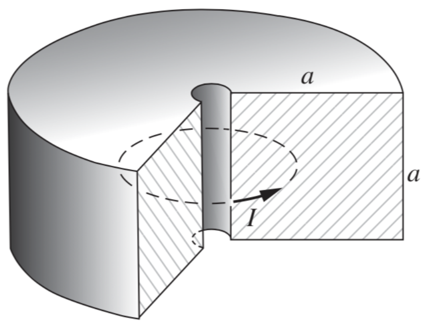
We can avoid wasting energy and producing entropy if we use an adjustable battery and gradually turn its voltage up, slowly enough so that the circuit is always near equilibrium. This is the electrical analogue of a smooth, adiabatic compression.

Remark

Suppose you wanted to account for the nonideal properties of a capacitor. In principle, the only way to get the answer exactly is to treat all the fields with Maxwell's equations and all the charges with Newton's laws. But we can often mimic nonideal effects by just adding resistors to the circuit. This has the benefit of staying within the "lumped element abstraction", where we can solve for everything with Kirchoff's laws, which are much simpler.

But where should we add the resistors? It depends on what nonideal effect we're trying to model. For example, if we want to account for the resistance of the wires, we should add a small resistance in series with the capacitor. But if instead the capacitor is slightly leaky, we should instead add a large resistance in parallel with the capacitor. If both effects are important, we should add both. And if radiation is the main way energy is lost, this can't really be modeled as a resistor at all, because the amount of radiated power depends on the rate of change of the current, not the current itself. In some cases, the lumped element abstraction simply can't be rescued.

[3] **Problem 1** (Purcell 7.46). We have found that in an LR circuit the current changes on the timescale L/R. In a large conducting body such as the metallic core of the Earth, the "circuit" is not easy to identify. Nevertheless, we can estimate the decay time. Suppose the current flows in a solid doughnut of square cross section, as shown, with conductivity σ.



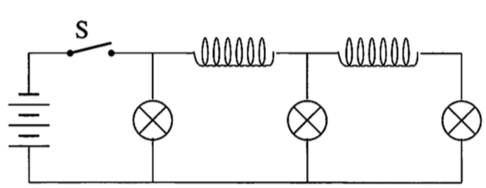
The current is spread out in some way over the cross section.

(a) Make a rough estimate of the resistance and inductance. For the latter, it may be easiest to estimate the magnetic field at the center of the donut first, then use that to estimate the total magnetic field energy.

(b) With these results, show that $\tau \sim \mu_0 a^2 \sigma$, which also follows from dimensional analysis.

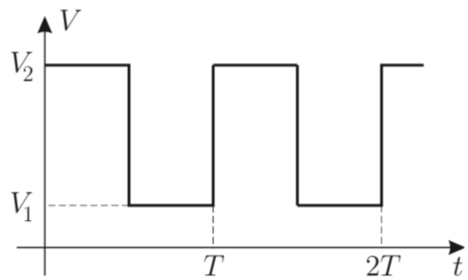
(c) Given that the radius of the Earth is $r \sim 3000\,\text{km}$ and $\sigma \sim 10^6\,(\Omega \cdot \text{m})^{-1}$, estimate τ.

[2] **Problem 2** (PPP 171). A circuit contains three identical lamps (modeled as resistors) and two identical inductors, as shown.



The switch S is closed for a long time, then suddenly opened. Immediately afterward, what are the relative brightnesses of the lamps?

[3] **Problem 3** (Kalda). A capacitor C and resistor R are connected in series. Rectangular voltage pulses are applied, as shown below.



After a long time, find the average power dissipated on the resistor if (a) $T \gg RC$ and (b) $T \ll RC$.

Remark

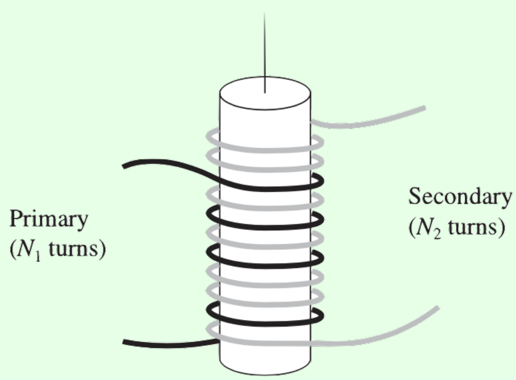
Out of all the analogies mentioned above, why is capacitance defined "backwards", so that $C \sim 1/k$? I actually have no idea, but one possibility is that large quantities should intuitively correspond to large objects. An object has to be physically large (and thereby expensive) to have a high C or a high L, and you can easily see this on your circuit board. (Of course, this doesn't explain everything; the largest R you can get is just a break in the circuit, which is neither large nor expensive.)

Another difference in the analogies is that for circuits we usually measure $I(t)$, while for mechanical oscillators we usually measure $x(t)$.

We now consider some problems involving mutual inductance.

Example 3: Griffiths 7.57

Two coils are wrapped around a cylindrical form so that the same flux passes through every turn of both coils, i.e. so that the mutual inductance is maximal. In practice this is achieved by inserting an iron core through the cylinder, which has the effect of forcing the magnetic flux to stay inside the cylinder.



The "primary" coil has N_1 turns and the secondary has N_2. If the current I in the primary is changing, show that the emf $\mathcal{E}_2$ in the secondary obeys

$$\frac{\mathcal{E}_2}{\mathcal{E}_1} = \frac{N_2}{N_1}$$

where $\mathcal{E}_1$ is the (back) emf of the primary.

Solution

Let Φ be the flux through a single loop of either coil due to the current in the primary. Then

$$\Phi_1 = N_1\Phi, \quad \Phi_2 = N_2\Phi.$$

By Faraday's law,

$$\mathcal{E}_1 = -N_1\frac{d\Phi}{dt}, \quad \mathcal{E}_2 = -N_2\frac{d\Phi}{dt}$$

which gives the desired result. This is a primitive transformer, a device for raising or lowering the emf of an alternating current source. By choosing the appropriate number of turns, any desired secondary emf can be obtained.

We can also solve this problem more formally using what we know about inductance, which will also tell us what happens when both currents are nonzero. The emfs obey

$$\mathcal{E}_1 = -L_1\frac{dI_1}{dt} - M\frac{dI_2}{dt}, \quad \mathcal{E}_2 = -L_2\frac{dI_2}{dt} - M\frac{dI_1}{dt}.$$

We showed in **E5** that $L_i = \mu_0 N_i^2 \pi R^2/H$ for a cylindrical solenoid. (Here, H stands for the length of the iron core, since this is the length over which the magnetic field exists.) As you'll show below, the maximum possible value of the mutual inductance, which is achieved by this ideal transformer, is $\sqrt{L_1L_2}$. Plugging in these results gives

$$\mathcal{E}_1 = -\left(\frac{\mu_0\pi R^2}{H}\right)\left(N_1^2\frac{dI_1}{dt} + N_1N_2\frac{dI_2}{dt}\right), \quad \mathcal{E}_2 = -\left(\frac{\mu_0\pi R^2}{H}\right)\left(N_2^2\frac{dI_2}{dt} + N_1N_2\frac{dI_1}{dt}\right).$$

This tells us the desired result holds for any values of the dI_i/dt.

This result is not surprising from the standpoint of Faraday's law. The flux change through any cross-section of the iron core is the same, so the induced emf around any circle around it is the same. Thus, the emf per turn is the same between the coils, $\mathcal{E}_1/N_1 = \mathcal{E}_2/N_2$, which again gives the desired result.

[3] **Problem 4.** Consider two inductors L_i, with mutual inductance M.

(a) Show that if the inductors have currents I_i, the total stored energy is

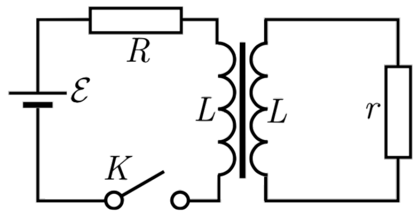
$$U = \frac{1}{2}L_1I_1^2 + \frac{1}{2}L_2I_2^2 + MI_1I_2.$$

Use this result to show that $|M| \leq \sqrt{L_1L_2}$.

(b) Suppose these two inductors are in series. Find their combined effective inductance.

(c) Suppose these two inductors are in parallel. Find their combined effective inductance.

[3] **Problem 5** (Kalda)**.** An electrical transformer is connected as shown.



Both windings of the transformer have the same number of loops and the self-inductance of both coils is equal to L. There is no leakage of the magnetic field lines from the core, so that the mutual inductance is also equal to L.

(a) Find the current in both loops immediately after the switch is closed.

(b) Find the currents as a function of time.

2 AC RLC Circuits and Impedance

Idea 2: Impedance

Current and voltage can be promoted to complex quantities,

$$V(t) = V_0 \cos(\omega t + \phi), \quad \tilde{V}(t) = \tilde{V}_0 e^{i\omega t}, \quad \tilde{V}_0 = V_0 e^{i\phi}$$

where the physical quantity is the real part. This is useful because we can relate $\tilde{V}$ and $\tilde{I}$ in all cases by $\tilde{V} = \tilde{I} Z$ where Z is the impedance, and

$$Z_R = R, \quad Z_C = \frac{1}{i\omega C}, \quad Z_L = i\omega L$$

for the three common circuit elements. Impedance is extremely useful for finding the steady state response of a circuit. If you're interested in the transients, you can find them by applying the techniques of **M4** to the Kirchoff's loop rule equation.

Idea 3: Power

Turning parameters complex and taking the real part works because we're dealing with linear equations. As a result, it doesn't work for energy or power, which are quadratic.

For instance, the power is not simply the real part of $\tilde{I}\tilde{V}$, but rather

$$P = IV = \mathrm{Re}(\tilde{I})\,\mathrm{Re}(\tilde{V}) = I_0 V_0 \cos(\omega t)\cos(\omega t + \phi)$$

where ϕ is the phase angle of Z. To compute the average power, note that

$$P = \frac{V_0^2}{|Z|}\cos(\omega t)(\cos(\omega t)\cos(\phi) - \sin(\omega t)\sin(\phi)).$$

The second term averages to zero, while $\cos^2(\omega t)$ averages to $1/2$ as usual, so

$$\overline{P} = \frac{1}{2}\frac{V_0^2}{|Z|}\cos(\phi) = \frac{1}{2} I_0 V_0 \cos(\phi).$$

We can decompose a general impedance as $Z = R + iX$, in which case $\cos\phi = R/|Z|$, and

$$\overline{P} = \frac{1}{2}\frac{I_0 V_0 R}{|Z|} = \frac{1}{2} I_0^2 R.$$

It's conventional to define $I_{\text{rms}}^2 = I_0^2/2$ to be the average value of I^2, giving

$$\overline{P} = I_{\text{rms}}^2 R = \frac{V_{\text{rms}}^2}{R}.$$

Example 4

Find the magnitude of the current through a series RLC circuit with AC voltage source $V_0 \cos\omega t$.

Solution

We promote the voltage and current to complex numbers,

$$V(t) = V_0 e^{i\omega t}.$$

Kirchoff's loop rule (subject to the caveats in **E5**) is

$$L\dot{I} + IR + \frac{Q}{C} = V_0 e^{i\omega t}.$$

This is quite similar to a damped driven harmonic oscillator, except that we want to get $I(t)$, rather than $Q(t)$. To get the steady state behavior, we guess

$$I(t) = I_0 e^{i\omega t}.$$

Then we have

$$\dot{I}(t) = (i\omega) I_0 e^{i\omega t}, \quad Q(t) = \frac{1}{i\omega} I_0 e^{i\omega t}.$$

Plugging this in, we find

$$\left(i\omega L + R + \frac{1}{i\omega C}\right) I_0 = V_0.$$

Solving for the magnitude of the current gives

$$|I_0| = \frac{|V_0|}{|i\omega L + R + 1/i\omega C|} = \frac{|V_0|}{\sqrt{R^2 + (\omega L - 1/\omega C)^2}}$$

which is maximized when $\omega = 1/\sqrt{LC}$, as we saw in **M4**. We could also have gotten straight to this last step by just using complex impedances.

Example 5

An imperfect voltage source consists of an ideal AC voltage source in series with an impedance Z_S. It is attached to a load of impedance Z_L. What value of Z_L maximizes the power transferred to the load?

Solution

Write the impedances as $Z_L = R_L + iX_L$. When the impedances are purely real, it's a familiar fact that the optimum is at $R_S = R_L$. We consider the case of general impedance here to illustrate how to work with power. First, the current has amplitude

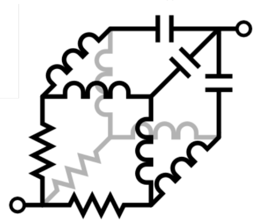
$$I_0 = \frac{|V|}{|Z_S + Z_L|}.$$

The average power dissipated in the load is

$$\overline{P} = \frac{1}{2} I_0^2 R_L \propto \frac{R_L}{|Z_S + Z_L|^2} = \frac{R_L}{(R_S + R_L)^2 + (X_S + X_L)^2}.$$

The denominator is minimized when $X_S + X_L = 0$, so the optimal real part is $R_L = R_S$ by the same logic as the purely real case. Thus, the highest power is achieved for $Z_L = Z_S^*$.

[1] **Problem 6.** Consider the cube of resistances R, capacitances C, and inductances L shown below.



Compute the impedance between the terminals.

[3] **Problem 7.** Consider an RLC circuit with a driving $V(t) = V_0 e^{i\omega t}$.

(a) Suppose the resistor, inductor, and capacitor are connected in *parallel*. Sketch the current $|I_0|$ through the driver as a function of ω, and compare it to the result for a standard series RLC circuit. Can you give a qualitative explanation for the difference?

(b) As we saw in **M4**, the quality factor for an oscillator quantifies how fast the energy in an *undriven* oscillator decays away. Specifically,

$$Q = \frac{\text{average energy stored in the oscillator}}{\text{average energy dissipated per radian}}.$$

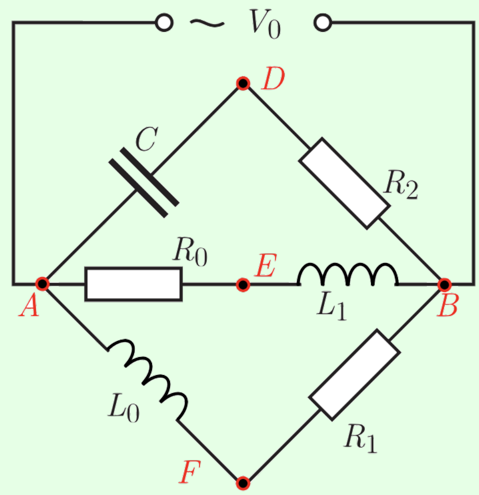
Find the quality factor for a series RLC circuit, and confirm your answer has correct dimensions.

(c) What is the condition for a series RLC circuit to be overdamped?

(d) Find the quality factor for a parallel RLC circuit. You should find that the quality factor increases as R is increased – why does this make sense?

Example 6: NBPhO 2018.7

Consider the following AC circuit.



The voltage difference between D and E has amplitude $V_{DE} = 7\,\text{V}$. Similarly, $V_{DF} = 15\,\text{V}$ and $V_{\text{EF}} = 20\,\text{V}$. What is the magnitude of V_0?

Solution

Treat the voltages as phasors, and let $V_A = 0$, so $V_B = V_0$. Then V_{AF} is perpendicular to V_{FB}, which means that V_F lies on the circle centered at $V_0/2$ with radius $V_0/2$. The exact same logic applies for V_E and V_D. Therefore, we know that V_{DE}, V_{DF} and V_{EF} form the three sides of a triangle, where the desired answer is the diameter of its circumcircle.

For a triangle with side lengths a, b, and c, and a circumcircle of radius R, Heron's formula states that the area is

$$A = \sqrt{p(p-a)(p-b)(p-c)}, \quad p = \frac{1}{2}(a+b+c).$$

The area is also given by

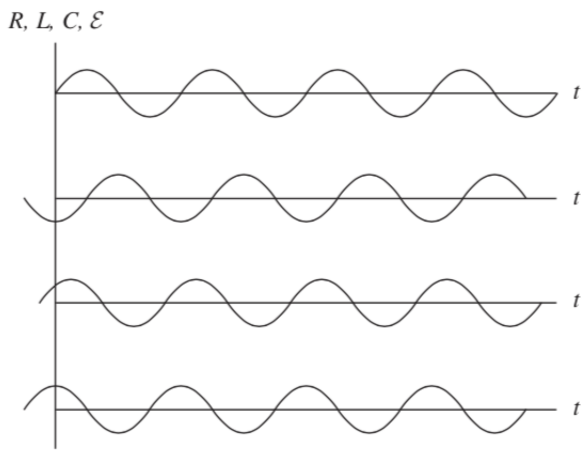
$$A = \frac{abc}{4R}.$$

Solving for the diameter, we have

$$2R = \frac{abc}{2\sqrt{p(p-a)(p-b)(p-c)}} = 25\,\text{V}.$$

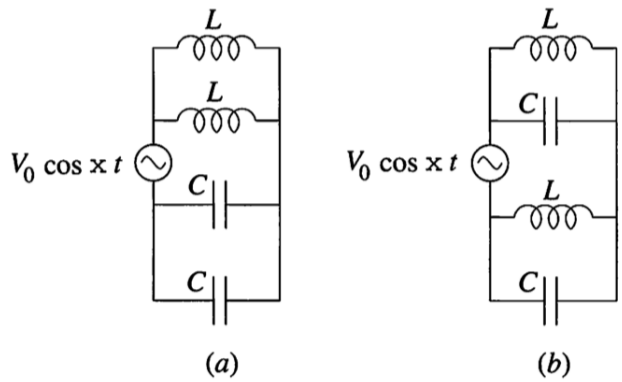
These geometric steps aren't that important; the key idea is thinking in terms of phasors.

[2] **Problem 8** (Purcell 8.26)**.** The four curves shown below are plots, in some order, of the applied voltage and the voltages across the resistor, inductor, and capacitor of a series RLC circuit.



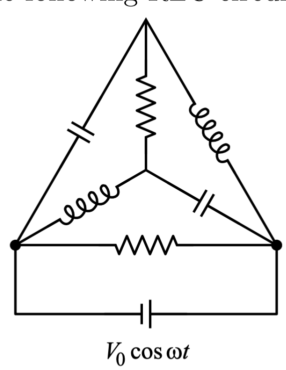
Which is which? Whose impedance is larger, the inductor's or the capacitor's?

[2] **Problem 9** (PPP 170)**.** Consider each of the following circuits.



In each case, find the amplitude of the current drawn from the source as a function of ω/ω_0, where $\omega_0 = 1/\sqrt{LC}$.

[3] **Problem 10** (BAUPC)**.** Consider the following RLC circuit.



The capacitors have capacitance C, the inductors have inductance L, and the resistors have resistance $R = \sqrt{L/C}$. Furthermore, the driving angular frequency is $\omega = 1/\sqrt{LC}$. Find the amplitude of the total current through the circuit.

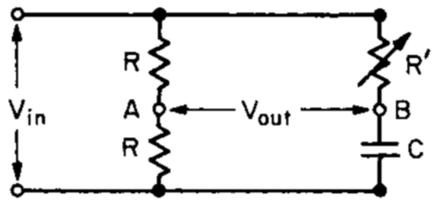
[3] **Problem 11.** USAPhO 2002, problem A1.

[3] **Problem 12.** USAPhO 2011, problem B1.

3 Electrical Engineering

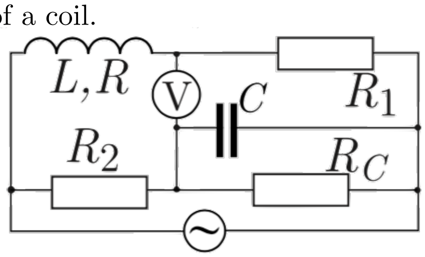
These next problems are about using RLC circuits for practical purposes. They don't require anything not already introduced in the previous section, but they represent a different way of thinking that it's crucial to get comfortable with.

[2] **Problem 13** (Feynman)**.** In electronic circuits it is often desired to provide a sinusoidal voltage of constant amplitude but variable phase. A circuit which accomplishes this is called a phase-shifting network. One example is shown below.



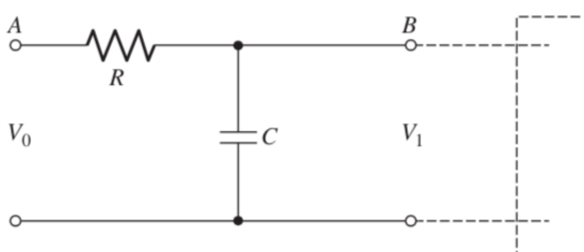
Show that the voltage measured between terminals A and B has half the amplitude of the input voltage, and a phase which may be adjusted by changing the resistance R'.

[3] **Problem 14** (Kalda)**.** The figure below shows a Maxwell's bridge, which is used for measuring the inductance L and resistance R of a coil.



To do this, the angular frequency ω is fixed and the known parameters R_1, R_2, R_C, and C are adjusted until the voltmeter reads zero. Once this is done, find R and L in terms of the other parameters.

[5] **Problem 15.** An alternating voltage $V_0 \cos \omega t$ is applied to the terminals at A. The terminals at B are connected to an audio amplifier of very high input impedance. (That is, current flow into the amplifier is negligible.)



This circuit is the most primitive of "low-pass" filters.

(a) Calculate the "gain" ratio $|\tilde{V}_1|/V_0$ in this filter. Show that for sufficiently high frequencies, the signal power is reduced by a factor of 4 for every doubling of the frequency.

(b) Design a low-pass filter without using a capacitor.

(c) Design a high-pass filter.

(d) Design a stronger low-pass filter, i.e. one which reduces the signal power by a greater factor for every doubling of the frequency.

(e) Design a band-pass filter, which suppresses both low and high frequencies, but has a constant gain for a wide range of medium frequencies. (It's okay if the constant gain is less than 1, as we can just pass the output through an amplifier.)

(f) Design a notch filter, which suppresses a very small range of frequencies, while letting all other frequencies through.

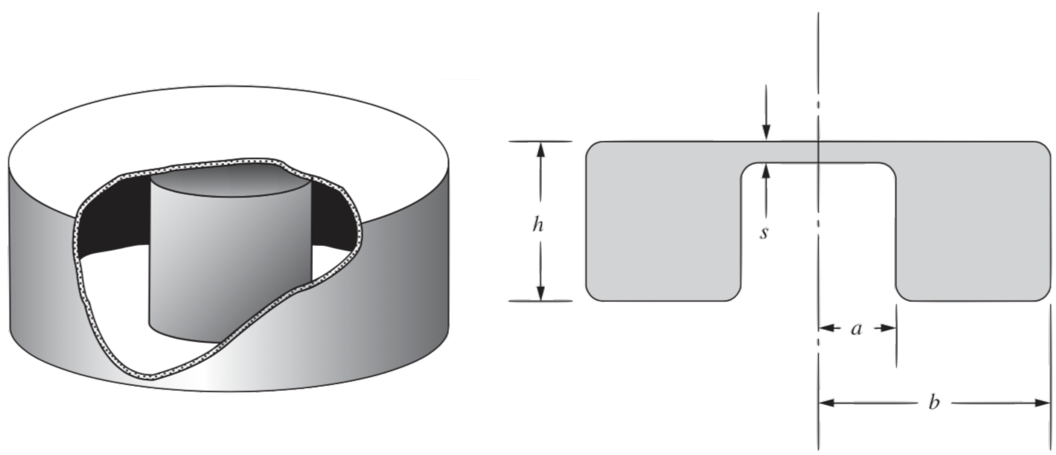
[4] **Problem 16.** IPhO 1984, problem 3. A nice, short problem on filters.

[2] **Problem 17.** Consider the same setup as problem 15, but with the resistor and capacitor switched.

(a) Assuming that $V_1 \ll V_0$, show that the output voltage is proportional to the derivative of the input voltage. Hence the circuit is a differentiator. (Can you relate this to the kind of filtering such a setup does?)

(b) Design a circuit whose output is proportional to the integral of V_0, again assuming $V_1 \ll V_0$.

[3] **Problem 18.** A resonant cavity of the form illustrated below is an essential part of many microwave oscillators. It can be regarded as a simple LC circuit.



(a) Assuming that $s \ll a, b, h$, estimate the lowest resonant angular frequency of the cavity by treating it as an LC circuit. It may be helpful to sketch the magnetic and electric fields.

(b) One of the most common types of cavity is a cylindrical cavity, i.e. a hollow cylinder. (It corresponds to taking $s = h$ in the above setup.) Assuming that $h \approx b$, find a reasonable estimate of the lowest resonant angular frequency ω.

Remark

In **E3**, we saw that for DC circuits, any system of resistors and ideal batteries with two ports is equivalent, from the perspective of anything connected across the ports, to either a single resistor and ideal battery in series (the Thevenin equivalent), or a single resistor and ideal current source in parallel (the Norton equivalent). From the ideas covered in this problem set, we also know that any system of resistors, inductors, and capacitors with two ports is equivalent, at a fixed angular frequency ω, to a single lumped element with impedance Z_{eq}. This in turn could be constructed out of a single resistor and inductor or capacitor in series.

This naturally leads to a more general question: it is possible to construct a simple "equivalent" circuit that has exactly the same $Z_{\text{eq}}(\omega)$, for *all* ω? The answer is yes. For example, consider the simple case of a circuit of only inductors and capacitors. Here's the rough idea: in this case, the equivalent impedance is always a pure imaginary, rational function of ω, meaning a ratio of two polynomials in ω. But rational functions can always be expanded in partial fractions. Assuming no multiple roots for simplicity, each term in the partial fraction decomposition can be mimicked with an LC circuit, and we get the sum by placing these circuits in series.

In electrical engineering, the general task of constructing a circuit with a prescribed $Z(\omega)$ is called network synthesis; the above example is called Foster's synthesis. These techniques can be used to construct filters more elaborate than the ones you explored in problem 15.

Remark

Power companies often transmit electricity with "three-phase power". This means that there are three "hot" electrical lines, carrying voltages

$$V_1(t) = V_0 \cos(\omega t), \quad V_2(t) = V_0 \cos(\omega t + 2\pi/3), \quad V_3(t) = V_0 \cos(\omega t + 4\pi/3).$$

Depending on your home, you might be able to connect to this three-phase power with special outlets, to use power tools. There are several advantages to three-phase power, but one is that it supplies a constant power, as $V_1^2 + V_2^2 + V_3^2$ is constant.

This shouldn't be confused with the three holes in an ordinary American wall outlet. In these outlets, one of the eyes is the "hot" one, with voltage $V_1(t)$, while the other eye and the mouth are both grounded. Appliances are powered by the voltage difference between the eyes. Appliances that use significant power and have metal exteriors have three-prong plugs. Here, the grounded "mouth" hole is connected directly to the exterior of the appliance, ensuring that it can't shock you, even if something goes wrong inside.

4 Normal Modes

Idea 4

A circuit with n independent loops has n normal modes. If we ignore resistances, the normal modes are pure sinusoids, though in all real circuits they exponentially damp over time. Just as in mechanics, the general solution for the behavior of a driven circuit is a superposition of normal mode currents and the response to the driving.

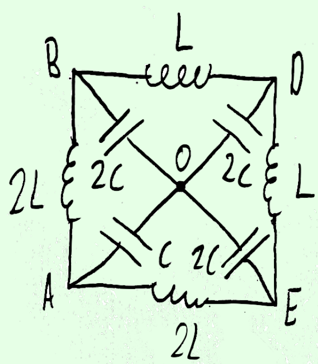
There are many ways to find the normal mode frequencies.

- One way is to pick any two points not directly connected by wires. We may imagine that across these points we have attached a current source $\tilde{I}$ which is doing nothing, $\tilde{I} = 0$. If a normal mode is present at angular frequency ω, then we can have $\tilde{V} \neq 0$, even though $\tilde{I} = 0$ because current is merely sloshing around inside the circuit. Thus, the equivalent impedance $Z(\omega)$ between these points is infinite.

- Another way is to pick two points directly connected by wires. We may imagine this wire is actually a voltage source $\tilde{V}$ which is doing nothing, $\tilde{V} = 0$. If a normal mode is present at angular frequency ω, then we can have a current $\tilde{I} \neq 0$ through the wire even though $\tilde{V} = 0$, so the equivalent impedance $Z(\omega)$ between these points is zero.
- Some LC circuits can be mapped to sets of masses and springs using the analogies in idea 1, which can help with guessing the normal modes.
- Finally, one may simply write down all of Kirchoff's loop equations, plug in $e^{i\omega t}$ time dependence, and look for a solution. This boils down to solving a system of n equations, or equivalently evaluating the determinant of an $n \times n$ matrix. This is rarely the best approach on an Olympiad.
- Not every problem benefits from using normal modes; for relatively simple circuits with special initial conditions, it may be better to solve Kirchoff's loop equations directly.

Example 7: Kalda 89

Find the normal mode frequencies of the circuit below.

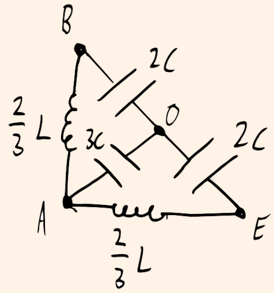


Solution

There are four independent Kirchoff's loop equations, so we expect four normal modes. One normal mode consists of current simply flowing uniformly along the outside, along the inductors. Since the capacitors aren't involved, this normal mode has $\omega_0 = 0$.

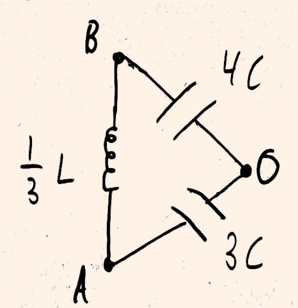
Now we apply the first technique listed above: we pick two points not directly connected with wires, and set the impedance to infinity. By symmetry, it's best to pick A and D. By symmetry, if any voltage is applied between A and D, the points B and E will be at the same voltage. Furthermore, this point will be at the same voltage as O, because the remaining circuit forms a balanced Wheatstone bridge, as introduced in **E3**. Identifying B, E, and O straightforwardly gives a simple LC circuit with $L_{\text{eff}} = (3/2)L$ and $C_{\text{eff}} = (2/3)C$, and resonant angular frequency $\omega_1 = 1/\sqrt{L_{\text{eff}} C_{\text{eff}}} = 1/\sqrt{LC}$.

This procedure only gave one of the three remaining normal modes, so we must have missed the other two because they have zero voltage difference between A and D. Therefore, to find the other two, we can join A and D, leading to the simpler equivalent circuit below.



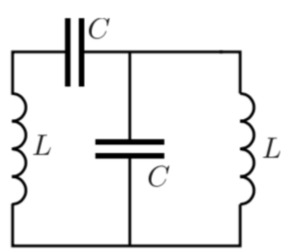
We now apply the same procedure between points B and E. This circuit is again a balanced Wheatstone bridge, so O and A are at the same voltage. We then have a simple LC circuit with $L_{\text{eff}} = (4/3)L$ and $C_{\text{eff}} = C$, giving $\omega_2 = \sqrt{3/4LC}$.

Again, we've missed a normal mode, so that remaining mode must have zero voltage difference between B and E. Joining them together leads to the final equivalent circuit below.



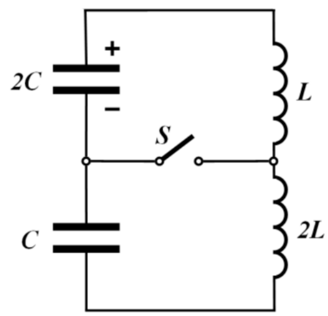
This is now a simple LC circuit with $L_{\text{eff}} = (1/3)L$ and $C_{\text{eff}} = (12/7)C$, giving the final resonant angular frequency $\omega_3 = \sqrt{7/4LC}$.

[2] **Problem 19** (Kalda)**.** Consider the LC circuit below.



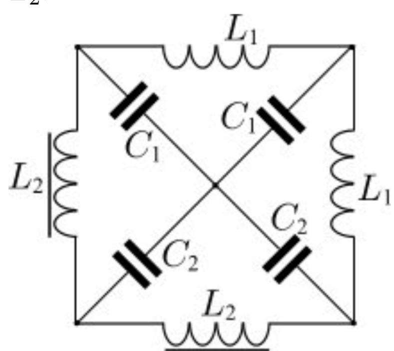
Show that the normal mode angular frequencies are $\omega = (\sqrt{5} \pm 1)/2\sqrt{LC}$.

[3] **Problem 20** (IPhO 2014)**.** Initially, the switch S is open in the circuit shown below.



The capacitor with capacitance $2C$ is given a charge q_0, and immediately begins to discharge. At the moment when the current through the inductors reaches its maximum value, the switch S is closed. Find the maximum current through the switch thereafter.

[5] **Problem 21** (Physics Cup 2012)**.** Find the angular frequencies of the normal modes of the circuit below, where $C_1 \ll C_2$ and $L_1 \ll L_2$.



You may give all of your answers to lowest order in C_1/C_2 and L_1/L_2.

5 Nonlinear Circuit Elements

In this section we'll introduce nonlinear circuit elements, focusing on diodes. More exotic circuit elements will be covered in **E7**.

Idea 5

Many nonlinear circuit elements can be described by a current-voltage characteristic $I(V)$. Such circuit elements have trivial time dependence, just like resistors, and working with them basically amounts to using Kirchoff's laws as usual, plugging in $I(V)$ where necessary.

Since the implementation details of such elements can be very complicated, and many draw power from external sources, it generally isn't productive to think of them "physically"; they are more like miniature computers than physical objects. One just has to take $I(V)$ as given and work directly with it. Some simple examples are:

- An ideal diode acts like a wire in one direction and a break in the other, so it has

$$I(V) = \begin{cases} \infty & V > 0, \\ 0 & V < 0. \end{cases}$$

- Sometimes one instead takes the $I(V)$ characteristic

$$I(V) = \begin{cases} \infty & V > V_0, \\ 0 & V < V_0 \end{cases}$$

which means that it "costs" voltage V_0 to go through the diode in the forward direction. More realistically, $I(V)$ smoothly increases when V passes V_0, but you don't often see this in Olympiad problems because it makes the math very messy.

- Zener diodes are bidirectional diodes. An ideal Zener diode has

$$I(V) = \begin{cases} \infty & V > V_0, \\ 0 & -V_0 < V < V_0, \\ -\infty & V < -V_0. \end{cases}$$

- Many familiar objects such as fuses (wires which break when I passes a threshold) and spark gaps (breaks that conduct when V passes a threshold) can be thought of as nonlinear circuit elements in the same way.

Analytically, these three cases are easily handled by casework. For instance, a diode acts just like a wire for positive V, and just like a break for negative V. In each case, the circuit is no more complicated than an ordinary one with linear circuit elements. Then you put the cases together to get the full behavior.

Example 8

A capacitor of capacitance C is charged so that its voltage is V_C. The capacitor is placed in series with a resistor R and a diode with $I(V)$ characteristic

$$I(V) = \begin{cases} \infty & V > V_0, \\ 0 & V < V_0. \end{cases}$$

The diode is oriented so that the initial voltage across it is positive. What happens next?

Solution

We use casework. If $V_c < V_0$, the voltage on the capacitor is not enough to get current to flow through the diode, so nothing happens. If $V_c > V_0$, current flows, and the diode acts like a battery of emf V_0 oriented in the opposite direction. This is just a discharging RC circuit, so the capacitor's voltage is

$$V(t) = (V_C - V_0)e^{-t/RC} + V_0.$$

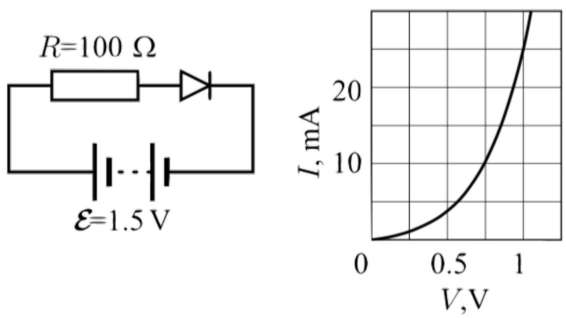
After a long time, the voltage on the capacitor falls to V_0 and current stops flowing.

Idea 6

It is difficult to solve a nonlinear circuit analytically if $I(V)$ is not very simple. In these cases:

- One can find the answer graphically as the intersection of $I(V)$ and another curve.
- One can solve for the answer iteratively on a calculator.
- If V stays within a narrow range, one can take a linear approximation to $I(V)$. This effectively replaces the element with a battery in series with a resistor, so the problem can be solved just like those in **E3**.

[2] **Problem 22** (Kalda)**.** Find the current in the circuit given below.



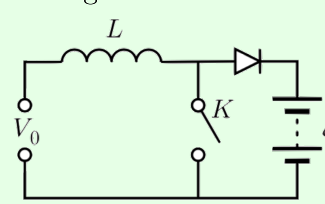
The nonlinear element is a diode with the $I(V)$ characteristic shown.

Idea 7

The power delivered to any circuit element is still $P = IV$. However, some nonlinear circuit elements can be active, providing net power to the circuit, like batteries.

Example 9: Kalda 64

The circuit below containing an ideal diode makes it possible to charge a rechargeable battery of voltage $\mathcal{E} = 12\,\mathrm{V}$ with a direct voltage source of a voltage $V_0 = 5\,\mathrm{V} < \mathcal{E}$.



To do this, the switch K is periodically opened and closed, with the opened and closed periods having equal length $\tau = 10\,\mathrm{ms}$. Find the average charging current assuming $L = 1\,\mathrm{H}$.

Solution

The intuition here is that, using an inductor and a switch, one can generate emfs larger than what we put in, because the current wants to keep flowing through the inductor when the switch is opened; this allows us to get enough emf to charge the battery. This idea is also used in the ignition coils of old-fashioned cars, where a voltage large enough to ionize air is produced, making a spark and starting the engine. There's also a fluid analogue, called the hydraulic ram, used to raise water. The point of the diode here is just to keep current from flowing the other way during the other half of the cycle.

When the switch is closed, no current can flow through the battery, and the current through the inductor builds up linearly, since there is an emf V_0 across the inductor. When the switch is opened, the emf across the inductor is $V_0 - \mathcal{E} = -7\,\text{V}$, causing its current to decrease while simultaneously charging the battery. After a time $(5/7)\tau$ with the switch open, the current through the inductor falls to zero, and the diode causes current to stop flowing.

Quantitatively, while the switch is closed, the current through the inductor builds up to $V_0\tau/L$. When the switch is open, current flows for a time $(5/7)\tau$, linearly falling to zero, so the total charge is

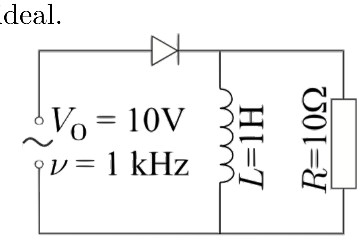
$$Q = \frac{1}{2}\frac{V_0\tau}{L}\frac{5}{7}\tau.$$

A cycle takes time 2τ, so

$$\overline{I} = \frac{Q}{2\tau} = \frac{5}{28}\frac{V_0\tau}{L} = 8.9\,\text{mA}.$$

By the way, your phone and laptop chargers probably have rectangular bricks containing a switched-mode power supply. This consists of one part that converts the AC wall power to DC, and a second part similar to the circuit above, but set up to output a *lower* DC voltage. You could also use a transformer to lower the AC voltage, but a switch-mode power supply is more space-efficient, and it easily copes with a range of input AC voltages and frequencies.

[3] **Problem 23** (Kalda). An alternating voltage $V = V_0\cos(2\pi\nu t)$ is applied to the leads of the circuit shown below. Treat the diode as ideal.



Assuming the current in the inductor begins at zero, what is the average current through the inductor at late times?

[3] **Problem 24.** EFPhO 2010, problem 9.

[3] **Problem 25.** EFPhO 2008, problem 6.

[3] **Problem 26.** EFPhO 2013, problem 8. This one has a nice mechanical analogy.

[3] **Problem 27.** IPhO 2001, problem 1c.

[3] **Problem 28.** USAPhO 2018, problem A2.

[4] **Problem 29.** EuPhO 2022, problem 2. A nice application of casework.

Electromagnetism VI: Circuits

AC circuits are covered in chapter 8 of Purcell, or chapter 10 of Wang and Ricardo, volume 2. Transmission lines, filters, and resonant cavities are covered physically in chapters II-22 and II-23 of the Feynman lectures, which will also build intuition for the next unit. Also see Jaan Kalda's circuits handout, an excellent resource which covers nonlinear circuit elements and much more. This problem set assumes knowledge about linear differential equations covered in **M1** and **M4**, but you can review the relevant material in chapter 4 of Morin. If you'd like to learn much more about circuits, from the electrical engineering perspective, a nice book is *Foundations of Analog and Digital Electronic Circuits* by Agarwal and Lang. There is a total of **84** points.

1 DC RLC Circuits

Idea 1

AC circuits correspond to driven damped oscillators by the analogies

$$Q \leftrightarrow x, \quad I \leftrightarrow v, \quad \dot{I} \leftrightarrow a, \quad L \leftrightarrow m, \quad R \leftrightarrow b, \quad C \leftrightarrow 1/k, \quad V_0 \leftrightarrow F_0.$$

More precisely, Kirchoff's loop equation in an AC circuit immediately becomes Newton's second law for a driven damped oscillator upon making these replacements.

Example 1

Consider a circuit with a battery of emf $\mathcal{E}$, a resistor R, and an inductor L in series, with zero initial current. Find the current $I(t)$ and verify that energy is conserved.

Solution

Kirchoff's loop equation is

$$\mathcal{E} = L\frac{dI}{dt} + IR.$$

To solve for the current, we can separate and integrate, giving

$$\frac{dt}{L} = \frac{dI}{\mathcal{E} - IR}$$

which yields

$$I(t) = \frac{\mathcal{E}}{R}(1 - e^{-(R/L)t}).$$

At long times, the inductor has no effect, since the current stops changing. To verify energy conservation we multiply Kirchoff's loop equation by I, since power is emf times current,

$$I\mathcal{E} = LI\frac{dI}{dt} + I^2R.$$

The left-hand side is the power output by the battery, and the two terms on the right-hand side represent the rate of increase in energy $LI^2/2$ stored in the inductor, and the power dissipated in the resistor, so all power is accounted for.

Example 2

An ideal battery of voltage $\mathcal{E}$ is suddenly connected to an ideal capacitor. After a short time, the capacitor has energy U. How much energy has been released by the battery?

Solution

This is a variant of the two capacitor paradox, which has essentially the same solution. First let's consider how the capacitor is charged up over time. Naively, since there's no resistance or inductance, the current in the circuit instantly becomes infinite, then instantly shuts off. This isn't realistic: to understand what's actually going on, we have to account for nonideal features of the circuit, such as resistance or self-inductance. For example, if the resistance dominates (overdamping), the capacitor charges up monotonically, as in an RC circuit. If the inductance dominates (underdamping), the capacitor voltage oscillates about $\mathcal{E}$, until eventually settling down due to the resistance.

At the end, the total energy on the capacitor is

$$\int V_C \, dQ = \int \frac{V_C}{C} \, dV_C = \frac{\mathcal{E}^2}{2C} = \frac{1}{2}\mathcal{E}Q$$

where Q is the total charge. But the work done by the battery is

$$\int \mathcal{E} \, dQ = \mathcal{E}Q$$

so the battery has released energy $2U$. Evidently, half of it is lost, no matter how close to "ideal" the circuit is! If the tiny resistance dominates, it is lost to heat in the circuit. If the tiny inductance dominates, then LC oscillations result, and energy can also be lost to electromagnetic radiation, as covered in **E7**.

In fact, this is just another example of the nonadiabatic processes you saw in **T1**. Instantly attaching a battery is the same kind of thing as instantly dropping a piston, and letting it bounce until it comes to a stop. Just like those nonadiabatic processes, attaching the battery in this way creates entropy; if the circuit and environment have temperature T, then

$$\Delta S = \frac{\Delta Q}{T} = \frac{\mathcal{E}Q}{2T}.$$

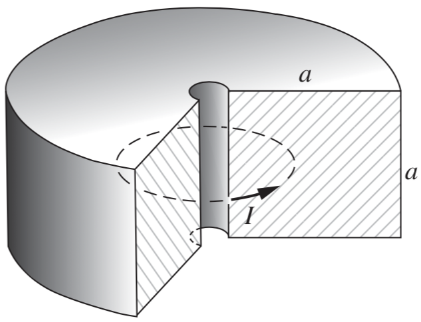
We can avoid wasting energy and producing entropy if we use an adjustable battery and gradually turn its voltage up, slowly enough so that the circuit is always near equilibrium. This is the electrical analogue of a smooth, adiabatic compression.

Remark

Suppose you wanted to account for the nonideal properties of a capacitor. In principle, the only way to get the answer exactly is to treat all the fields with Maxwell's equations and all the charges with Newton's laws. But we can often mimic nonideal effects by just adding resistors to the circuit. This has the benefit of staying within the "lumped element abstraction", where we can solve for everything with Kirchoff's laws, which are much simpler.

But where should we add the resistors? It depends on what nonideal effect we're trying to model. For example, if we want to account for the resistance of the wires, we should add a small resistance in series with the capacitor. But if instead the capacitor is slightly leaky, we should instead add a large resistance in parallel with the capacitor. If both effects are important, we should add both. And if radiation is the main way energy is lost, this can't really be modeled as a resistor at all, because the amount of radiated power depends on the rate of change of the current, not the current itself. In some cases, the lumped element abstraction simply can't be rescued.

[3] **Problem 1** (Purcell 7.46)**.** We have found that in an LR circuit the current changes on the timescale L/R. In a large conducting body such as the metallic core of the Earth, the "circuit" is not easy to identify. Nevertheless, we can estimate the decay time. Suppose the current flows in a solid doughnut of square cross section, as shown, with conductivity σ.



The current is spread out in some way over the cross section.

(a) Make a rough estimate of the resistance and inductance. For the latter, it may be easiest to estimate the magnetic field at the center of the donut first, then use that to estimate the total magnetic field energy.

(b) With these results, show that $\tau \sim \mu_0 a^2 \sigma$, which also follows from dimensional analysis.

(c) Given that the radius of the Earth is $r \sim 3000\,\text{km}$ and $\sigma \sim 10^6\,(\Omega \cdot \text{m})^{-1}$, estimate τ.

Solution. (a) We estimate

$$R \sim \frac{1}{\sigma} \frac{\text{length}}{\text{cross-sectional area}} \sim \frac{1}{\sigma} \frac{a}{a^2} \sim \frac{1}{\sigma a}.$$

The magnetic field at different points is on order $\mu_0 I/a$, so the magnetic energy is

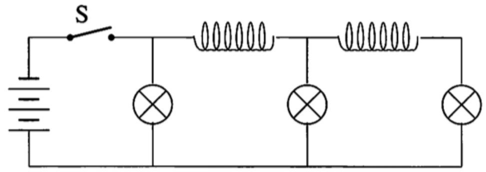
$$U_B \sim \frac{1}{2\mu_0} (\mu_0 I/a)^2 a^3$$

and equating this to $LI^2/2$ gives $L \sim \mu_0 a$.

(b) Combining these two gives $\tau \sim L/R \sim \mu_0 a^2 \sigma$, as desired.

(c) Plugging in the numbers gives $\tau \sim 3 \times 10^5$ years. This is much shorter than the Earth's life, so some energy source must actively drive the core.

[2] **Problem 2** (PPP 171)**.** A circuit contains three identical lamps (modeled as resistors) and two identical inductors, as shown.

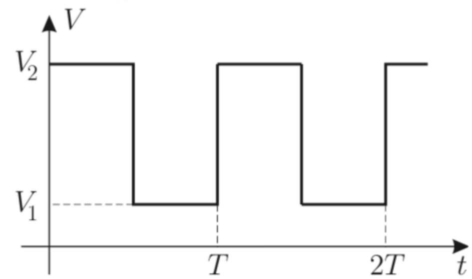


The switch S is closed for a long time, then suddenly opened. Immediately afterward, what are the relative brightnesses of the lamps?

Solution. After a long time with the switch closed, the inductors act like short circuits, so the currents in the bulbs are the same. Let this current be I. Then the currents in the two inductors are $2I$ and I.

When the switch is opened, the currents in the inductors remain the same, since they resist changes in current. This means the currents through the resistors are now $2I$, I, and I. Since $P = I^2 R$, the left bulb is 4 times brighter than the other two.

[3] **Problem 3** (Kalda)**.** A capacitor C and resistor R are connected in series. Rectangular voltage pulses are applied, as shown below.



After a long time, find the average power dissipated on the resistor if (a) $T \gg RC$ and (b) $T \ll RC$.

Solution. (a) In this case, there is sufficient time for the capacitor to reach equilibrium each time the voltage switches. The energy dissipated during a switch is

$$U = \int IV \, dt = \int V \, dQ.$$

During a switch, the voltage across the capacitor goes from $V_2 - V_1$ to zero linearly in the charge, while the total charge transferred is $\Delta Q = (V_2 - V_1)C$. Thus,

$$U = \overline{V}_R \, \Delta Q = \frac{1}{2}(V_2 - V_1)^2 C.$$

This happens every time $T/2$, so

$$\overline{P} = \frac{(V_2 - V_1)^2 C}{T}.$$

(b) In this case, the charge on the capacitor barely changes during each cycle. The average voltage across the capacitor is $(V_1 + V_2)/2$. Hence the magnitude of the voltage across the resistor is always approximately equal to $(V_2 - V_1)/2$, so

$$\overline{P} = \frac{V_R^2}{R} = \frac{(V_2 - V_1)^2}{4R}.$$

Remark

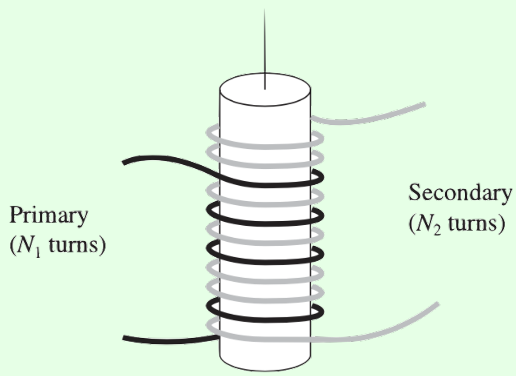
Out of all the analogies mentioned above, why is capacitance defined "backwards", so that $C \sim 1/k$? I actually have no idea, but one possibility is that large quantities should intuitively correspond to large objects. An object has to be physically large (and thereby expensive) to have a high C or a high L, and you can easily see this on your circuit board. (Of course, this doesn't explain everything; the largest R you can get is just a break in the circuit, which is neither large nor expensive.)

Another difference in the analogies is that for circuits we usually measure $I(t)$, while for mechanical oscillators we usually measure $x(t)$.

We now consider some problems involving mutual inductance.

Example 3: Griffiths 7.57

Two coils are wrapped around a cylindrical form so that the same flux passes through every turn of both coils, i.e. so that the mutual inductance is maximal. In practice this is achieved by inserting an iron core through the cylinder, which has the effect of forcing the magnetic flux to stay inside the cylinder.



The "primary" coil has N_1 turns and the secondary has N_2. If the current I in the primary is changing, show that the emf $\mathcal{E}_2$ in the secondary obeys

$$\frac{\mathcal{E}_2}{\mathcal{E}_1} = \frac{N_2}{N_1}$$

where $\mathcal{E}_1$ is the (back) emf of the primary.

Solution

Let Φ be the flux through a single loop of either coil due to the current in the primary. Then

$$\Phi_1 = N_1 \Phi, \quad \Phi_2 = N_2 \Phi.$$

By Faraday's law,

$$\mathcal{E}_1 = -N_1 \frac{d\Phi}{dt}, \quad \mathcal{E}_2 = -N_2 \frac{d\Phi}{dt}$$

which gives the desired result. This is a primitive transformer, a device for raising or lowering the emf of an alternating current source. By choosing the appropriate number of turns, any desired secondary emf can be obtained.

We can also solve this problem more formally using what we know about inductance, which will also tell us what happens when both currents are nonzero. The emfs obey

$$\mathcal{E}_1 = -L_1 \frac{dI_1}{dt} - M \frac{dI_2}{dt}, \quad \mathcal{E}_2 = -L_2 \frac{dI_2}{dt} - M \frac{dI_1}{dt}.$$

We showed in **E5** that $L_i = \mu_0 N_i^2 \pi R^2 / H$ for a cylindrical solenoid. (Here, H stands for the length of the iron core, since this is the length over which the magnetic field exists.) As you'll show below, the maximum possible value of the mutual inductance, which is achieved by this ideal transformer, is $\sqrt{L_1 L_2}$. Plugging in these results gives

$$\mathcal{E}_1 = -\left(\frac{\mu_0 \pi R^2}{H}\right)\left(N_1^2 \frac{dI_1}{dt} + N_1 N_2 \frac{dI_2}{dt}\right), \quad \mathcal{E}_2 = -\left(\frac{\mu_0 \pi R^2}{H}\right)\left(N_2^2 \frac{dI_2}{dt} + N_1 N_2 \frac{dI_1}{dt}\right).$$

This tells us the desired result holds for any values of the dI_i/dt.

This result is not surprising from the standpoint of Faraday's law. The flux change through any cross-section of the iron core is the same, so the induced emf around any circle around it is the same. Thus, the emf per turn is the same between the coils, $\mathcal{E}_1/N_1 = \mathcal{E}_2/N_2$, which again gives the desired result.

[3] **Problem 4.** Consider two inductors L_i, with mutual inductance M.

(a) Show that if the inductors have currents I_i, the total stored energy is

$$U = \frac{1}{2} L_1 I_1^2 + \frac{1}{2} L_2 I_2^2 + M I_1 I_2.$$

Use this result to show that $|M| \leq \sqrt{L_1 L_2}$.

(b) Suppose these two inductors are in series. Find their combined effective inductance.

(c) Suppose these two inductors are in parallel. Find their combined effective inductance.

Solution. (a) The differential work required to change the currents is

$$dU = (L_1 \dot{J}_1 + M \dot{J}_2)(J_1 dt) + (L_2 \dot{J}_2 + M \dot{J}_1)(J_2 dt) = L_1 J_1 dJ_1 + L_2 J_2 dJ_2 + M d(J_1 J_2)$$

where J_1, J_2 are the values of the currents at some intermediate time. Therefore, the total work required is

$$U = \frac{1}{2} L_1 I_1^2 + \frac{1}{2} L_2 I_2^2 + M I_1 I_2.$$

Let $x = I_2/I_1$, so

$$U \propto L_2 x^2 + 2Mx + L_1.$$

Physically, U must be positive for all x. Minimizing by setting the derivative to zero, we see the minimum of U is positive only if $|M| \leq \sqrt{L_1 L_2}$. The bound is saturated for inductors that practically overlap, or more generally for any configuration where all flux that goes through one inductor also goes through the other, such as the two coils of an ideal transformer.

(b) When the inductors are in series, they have the same current. The total emf is

$$\mathcal{E} = -\left(L_1\frac{dI}{dt} + M\frac{dI}{dt} + L_2\frac{dI}{dt} + M\frac{dI}{dt}\right) = -(L_1 + L_2 + 2M)\frac{dI}{dt}$$

which means

$$L_{\text{eff}} = L_1 + L_2 + 2M.$$

Incidentally, since L_{eff} must be positive, this implies the bound $M > -(L_1 + L_2)/2$, though this is weaker than the bound in part (a).

(c) When the inductors are in parallel, they have the same emf, so

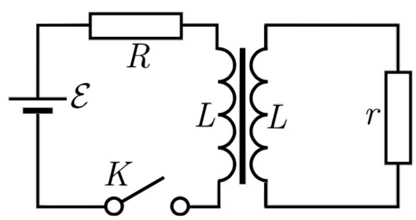
$$\mathcal{E} = -L_1\frac{dI_1}{dt} - M\frac{dI_2}{dt} = -L_2\frac{dI_2}{dt} - M\frac{dI_1}{dt}.$$

Solving the system gives

$$L_{\text{eff}} = \frac{L_1L_2 - M^2}{L_1 + L_2 - 2M}.$$

This also implies the bound $|M| < \sqrt{L_1L_2}$, as shown in a different way in part (a).

[3] **Problem 5** (Kalda)**.** An electrical transformer is connected as shown.



Both windings of the transformer have the same number of loops and the self-inductance of both coils is equal to L. There is no leakage of the magnetic field lines from the core, so that the mutual inductance is also equal to L.

(a) Find the current in both loops immediately after the switch is closed.

(b) Find the currents as a function of time.

Solution. (a) In a transformer, the flux is $\Phi = L(I_1 + I_2)$ and $d\Phi/dt = \mathcal{E}_L = L(dI_1/dt + dI_2/dt)$. Since $d\Phi/dt$ is finite, then initially $I_1 + I_2 = 0$. The voltage loop rule gives

$$\mathcal{E} = I_1R + L(dI_1/dt + dI_2/dt) = I_1R - I_2r$$

Using the initial condition of $I_1 = -I_2$ gives

$$I_1 = \frac{\mathcal{E}}{R + r}.$$

(b) Again,

$$\mathcal{E} = I_1R + L(dI_1/dt + dI_2/dt) \qquad L(dI_1/dt + dI_2/dt) + I_2r = 0$$

Let $I \equiv I_1 + I_2$. Combining the equations give

$$\frac{\mathcal{E}}{R} = I + \frac{L(R + r)}{Rr}\frac{dI}{dt}$$

$$I = \frac{\mathcal{E}}{R}(1 - e^{-t/\tau}) \qquad \tau \equiv L\left(\frac{1}{r} + \frac{1}{R}\right)$$

$$\frac{dI}{dt} = \frac{\mathcal{E}r}{L(R+r)} e^{-t/\tau}$$

Then with $I_1 = \mathcal{E}/R - \frac{L}{R} dI/dt$ and $I_2 = -\frac{L}{r} dI/dt$, we get

$$I_1 = \frac{\mathcal{E}}{R}\left(1 - \frac{r}{r+R} e^{-t/\tau}\right) \qquad I_2 = -\frac{\mathcal{E}}{(R+r)} e^{-t/\tau}.$$

2 AC RLC Circuits and Impedance

Idea 2: Impedance

Current and voltage can be promoted to complex quantities,

$$V(t) = V_0 \cos(\omega t + \phi), \quad \tilde{V}(t) = \tilde{V}_0 e^{i\omega t}, \quad \tilde{V}_0 = V_0 e^{i\phi}$$

where the physical quantity is the real part. This is useful because we can relate $\tilde{V}$ and $\tilde{I}$ in all cases by $\tilde{V} = \tilde{I} Z$ where Z is the impedance, and

$$Z_R = R, \quad Z_C = \frac{1}{i\omega C}, \quad Z_L = i\omega L$$

for the three common circuit elements. Impedance is extremely useful for finding the steady state response of a circuit. If you're interested in the transients, you can find them by applying the techniques of **M4** to the Kirchoff's loop rule equation.

Idea 3: Power

Turning parameters complex and taking the real part works because we're dealing with linear equations. As a result, it doesn't work for energy or power, which are quadratic.

For instance, the power is not simply the real part of $\tilde{I}\tilde{V}$, but rather

$$P = IV = \mathrm{Re}(\tilde{I})\,\mathrm{Re}(\tilde{V}) = I_0 V_0 \cos(\omega t)\cos(\omega t + \phi)$$

where ϕ is the phase angle of Z. To compute the average power, note that

$$P = \frac{V_0^2}{|Z|}\cos(\omega t)(\cos(\omega t)\cos(\phi) - \sin(\omega t)\sin(\phi)).$$

The second term averages to zero, while $\cos^2(\omega t)$ averages to $1/2$ as usual, so

$$\overline{P} = \frac{1}{2}\frac{V_0^2}{|Z|}\cos(\phi) = \frac{1}{2} I_0 V_0 \cos(\phi).$$

We can decompose a general impedance as $Z = R + iX$, in which case $\cos\phi = R/|Z|$, and

$$\overline{P} = \frac{1}{2}\frac{I_0 V_0 R}{|Z|} = \frac{1}{2} I_0^2 R.$$

It's conventional to define $I_{\text{rms}}^2 = I_0^2/2$ to be the average value of I^2, giving

$$\overline{P} = I_{\text{rms}}^2 R = \frac{V_{\text{rms}}^2}{R}.$$

Example 4

Find the magnitude of the current through a series RLC circuit with AC voltage source $V_0 \cos \omega t$.

Solution

We promote the voltage and current to complex numbers,

$$V(t) = V_0 e^{i\omega t}.$$

Kirchoff's loop rule (subject to the caveats in **E5**) is

$$L\dot{I} + IR + \frac{Q}{C} = V_0 e^{i\omega t}.$$

This is quite similar to a damped driven harmonic oscillator, except that we want to get $I(t)$, rather than $Q(t)$. To get the steady state behavior, we guess

$$I(t) = I_0 e^{i\omega t}.$$

Then we have

$$\dot{I}(t) = (i\omega) I_0 e^{i\omega t}, \quad Q(t) = \frac{1}{i\omega} I_0 e^{i\omega t}.$$

Plugging this in, we find

$$\left(i\omega L + R + \frac{1}{i\omega C} \right) I_0 = V_0.$$

Solving for the magnitude of the current gives

$$|I_0| = \frac{|V_0|}{|i\omega L + R + 1/i\omega C|} = \frac{|V_0|}{\sqrt{R^2 + (\omega L - 1/\omega C)^2}}$$

which is maximized when $\omega = 1/\sqrt{LC}$, as we saw in **M4**. We could also have gotten straight to this last step by just using complex impedances.

Example 5

An imperfect voltage source consists of an ideal AC voltage source in series with an impedance Z_S. It is attached to a load of impedance Z_L. What value of Z_L maximizes the power transferred to the load?

Solution

Write the impedances as $Z_L = R_L + iX_L$. When the impedances are purely real, it's a familiar fact that the optimum is at $R_S = R_L$. We consider the case of general impedance here to illustrate how to work with power. First, the current has amplitude

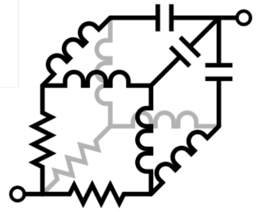
$$I_0 = \frac{|V|}{|Z_S + Z_L|}.$$

The average power dissipated in the load is

$$\overline{P} = \frac{1}{2} I_0^2 R_L \propto \frac{R_L}{|Z_S + Z_L|^2} = \frac{R_L}{(R_S + R_L)^2 + (X_S + X_L)^2}.$$

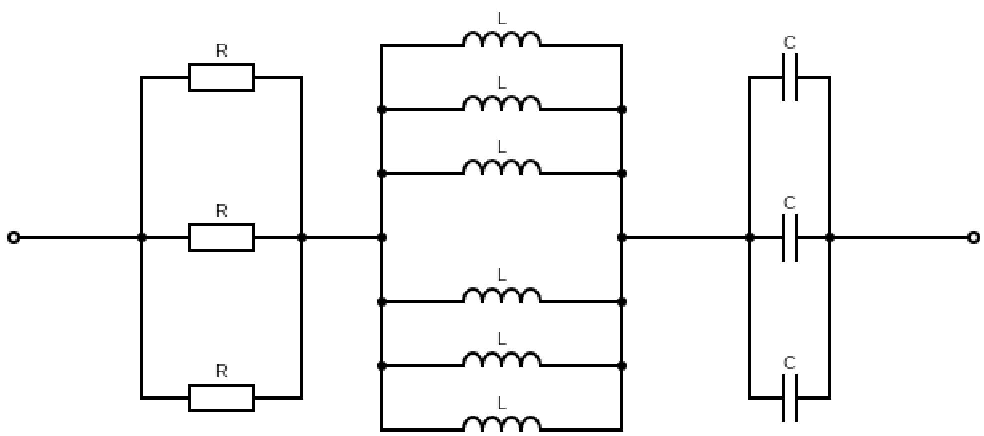
The denominator is minimized when $X_S + X_L = 0$, so the optimal real part is $R_L = R_S$ by the same logic as the purely real case. Thus, the highest power is achieved for $Z_L = Z_S^*$.

[1] **Problem 6.** Consider the cube of resistances R, capacitances C, and inductances L shown below.



Compute the impedance between the terminals.

Solution. By symmetry, the three points close to one end have the same voltage, and the three points close to the other end have the same as well. Thus, we simplify the diagram to the following:



We have impedance

$$Z = \frac{1}{3} R + \frac{1}{6} i\omega L + \frac{1}{3} \frac{1}{i\omega C}.$$

[3] **Problem 7.** Consider an RLC circuit with a driving $V(t) = V_0 e^{i\omega t}$.

(a) Suppose the resistor, inductor, and capacitor are connected in *parallel*. Sketch the current $|I_0|$ through the driver as a function of ω, and compare it to the result for a standard series RLC circuit. Can you give a qualitative explanation for the difference?

(b) As we saw in **M4**, the quality factor for an oscillator quantifies how fast the energy in an *undriven* oscillator decays away. Specifically,

$$Q = \frac{\text{average energy stored in the oscillator}}{\text{average energy dissipated per radian}}.$$

Find the quality factor for a series RLC circuit, and confirm your answer has correct dimensions.

(c) What is the condition for a series RLC circuit to be overdamped?

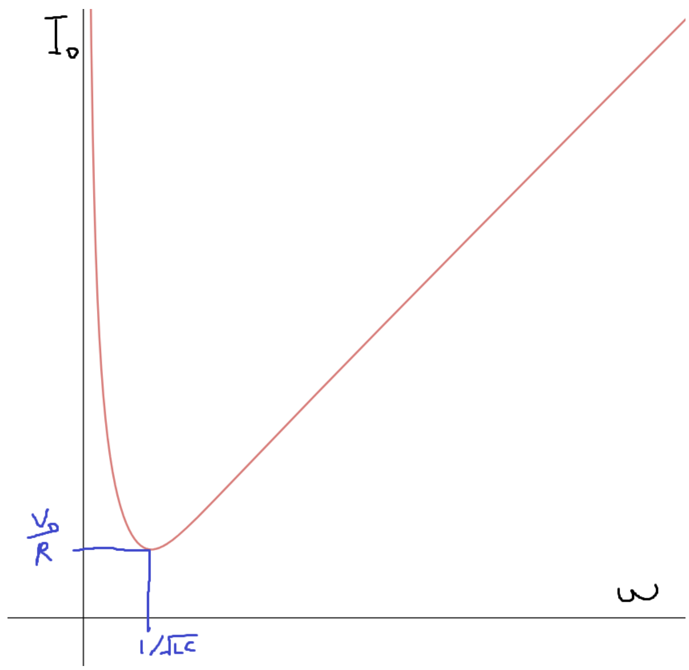
(d) Find the quality factor for a parallel RLC circuit. You should find that the quality factor increases as R is increased – why does this make sense?

Solution. (a) We know that $|I_0| = |V_0|/|Z|$, and adding the impedances in parallel gets

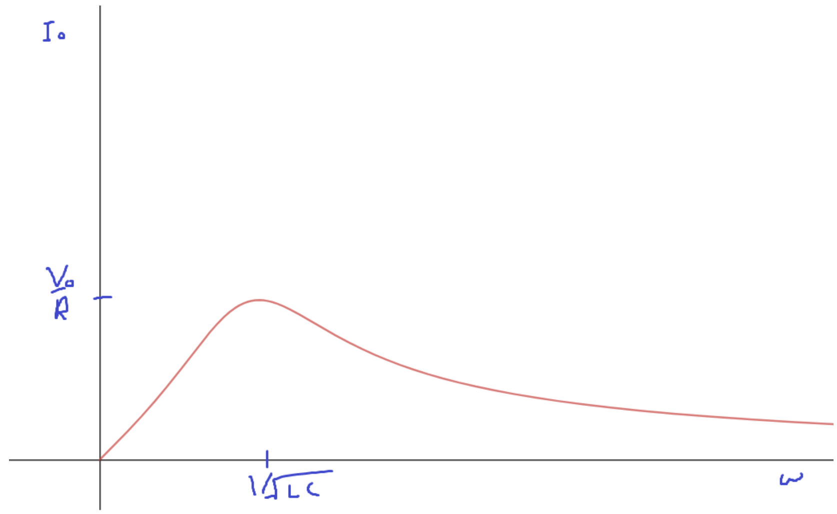
$$\frac{1}{Z} = \frac{1}{R} + \frac{1}{i\omega L} + i\omega C$$

$$|I_0| = \frac{|V_0|}{\omega LR}\sqrt{(\omega L)^2 + R^2(1-\omega^2 LC)^2}$$

This gives a graph for a parallel RLC circuit:



A series RLC circuit behaves the other way:



The intuition for a series RLC circuit is just as in **M4**, i.e. we get the most current when the system is driven at resonance, and the result is finite at resonance because the resistor absorbs the energy put in. But a parallel RLC circuit is backwards, because each component lets current through independently. The resistor's current contribution is independent of ω. At high frequencies, a capacitor in parallel can let a lot of current through, because it behaves like a wire. At low frequencies, an inductor can do the same. But at the resonant frequency, the contributions of the capacitor and inductor *cancel out*. Or in more formal language, the applied voltage excites the normal mode corresponding to the LC oscillation, which moves zero net current through the power source.

(b) Energy will be stored in the capacitor and inductor, of $Q^2/2C$ and $\frac{1}{2}LI^2$ respectively. As in any harmonic oscillator, these are equal on average, so the total energy is

$$E = 2 \times \frac{1}{2}L\langle I^2\rangle = L\langle I^2\rangle = \frac{1}{2}LI_0^2.$$

The power dissipated in the resistor is I^2R, so the average power is

$$\langle P\rangle = \langle I^2R\rangle = \frac{1}{2}I_0^2R.$$

Since a radian occurs in the time $1/\omega$,

$$Q = \frac{E}{\langle P\rangle/\omega} = \frac{\omega L}{R} = \frac{1}{R}\sqrt{\frac{L}{C}}$$

where we used $\omega \approx 1/\sqrt{LC}$. To check the dimensions, note that ωL and R are both impedances, so Q is dimensionless.

(c) Guessing $I(t) = I_0e^{i\omega t}$ in the Kirchoff loop rule (without power source) gets

$$i\omega L + R + \frac{1}{i\omega C} = 0$$

$$\omega^2LC - i\omega RC - 1 = 0$$

$$\omega = \frac{iRC \pm \sqrt{4LC - (RC)^2}}{2LC}$$

The oscillator is overdamped when there's no real component of ω which indicates no sinusoidal oscillation, so the condition is

$$\frac{R^2C}{4L} > 1.$$

In terms of the answer to part (b), this is $Q < 1/2$, and the requirement of a low quality factor makes sense. (However, we usually wouldn't describe such lossy systems in terms of a quality factor at all.)

(d) In the analysis for part (b), we used the fact that the current through each part of the series RLC circuit is equal. For the parallel RLC circuit, we use the fact that the voltages are equal. The average energy is

$$E = 2 \times \frac{1}{2}C\langle V^2\rangle = C\langle V^2\rangle = \frac{CV_0^2}{2}.$$

The power dissipated in the resistor is V^2/R, so the average power is

$$\langle P\rangle = \langle V^2/R\rangle = \frac{V_0^2}{2R}.$$

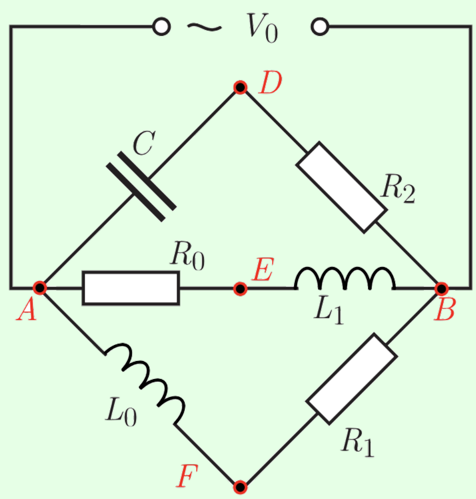
Since a radian occurs in the time $1/\omega$,

$$Q = \frac{E}{\langle P\rangle/\omega} = \omega RC = R\sqrt{\frac{C}{L}}.$$

This is the inverse of the quality factor of a series RLC circuit. It makes sense that the quality factor increases as R increases, because for very high R almost no current flows through the resistor. It's just the AC analogue of the usual result in DC circuits: since $P = I^2R = V^2/R$, high-resistance resistors dissipate more when in series, and less when in parallel.

Example 6: NBPhO 2018.7

Consider the following AC circuit.



The voltage difference between D and E has amplitude $V_{DE} = 7\,\text{V}$. Similarly, $V_{DF} = 15\,\text{V}$ and $V_{\text{EF}} = 20\,\text{V}$. What is the magnitude of V_0?

Solution

Treat the voltages as phasors, and let $V_A = 0$, so $V_B = V_0$. Then V_{AF} is perpendicular to V_{FB}, which means that V_F lies on the circle centered at $V_0/2$ with radius $V_0/2$. The exact same logic applies for V_E and V_D. Therefore, we know that V_{DE}, V_{DF} and V_{EF} form the three sides of a triangle, where the desired answer is the diameter of its circumcircle.

For a triangle with side lengths a, b, and c, and a circumcircle of radius R, Heron's formula states that the area is

$$A = \sqrt{p(p-a)(p-b)(p-c)}, \quad p = \frac{1}{2}(a+b+c).$$

The area is also given by

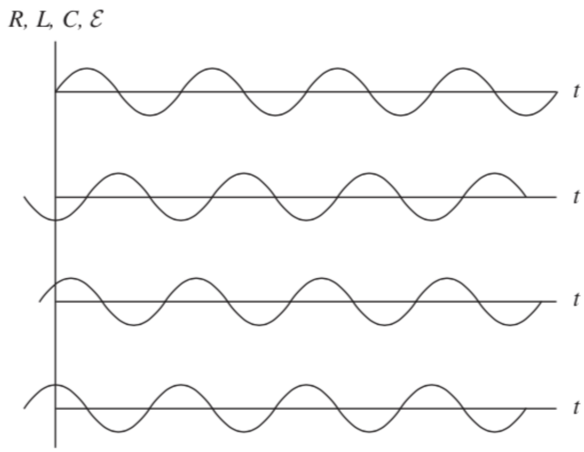
$$A = \frac{abc}{4R}.$$

Solving for the diameter, we have

$$2R = \frac{abc}{2\sqrt{p(p-a)(p-b)(p-c)}} = 25\,\mathrm{V}.$$

These geometric steps aren't that important; the key idea is thinking in terms of phasors.

[2] **Problem 8** (Purcell 8.26)**.** The four curves shown below are plots, in some order, of the applied voltage and the voltages across the resistor, inductor, and capacitor of a series RLC circuit.



Which is which? Whose impedance is larger, the inductor's or the capacitor's?

Solution. We have $V_R = RI$, $V_L = (i\omega L)I$, and $V_C = \frac{1}{i\omega C}I$. Thus, the graph of V_L is shifted by $\pi/2$ to the left of that of I, and V_C is shifted by $\pi/2$ to the right. The first, second, and fourth graphs are shifted relative to each other by multiples of $\pi/2$, so we conclude

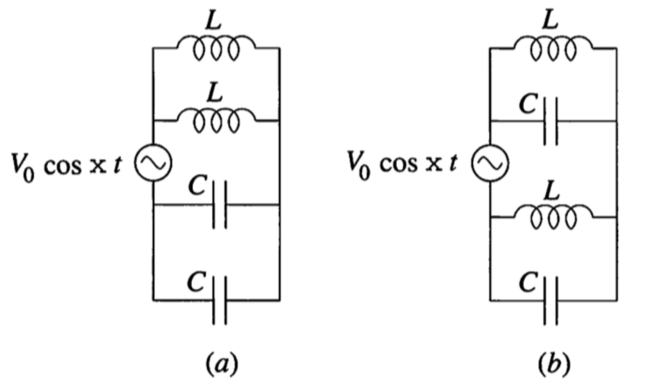
$$\text{the first graph is } V_R, \quad \text{the second graph is } V_C, \quad \text{the fourth graph is } V_L$$

from which we find by process of elimination that

$$\text{the third graph is } \mathcal{E}.$$

Note that when the inductor and capacitor have equal impedance, $\mathcal{E}$ and $I \propto V_R$ are in phase. When the inductor has a higher impedance, $\mathcal{E}$ looks more like V_L, which "leads" the current. When the capacitor has a higher impedance, $\mathcal{E}$ looks more like V_C, which "lags" the current. By comparing the first and third graphs, we see the phase of $\mathcal{E}$ slightly leads the current, so the inductor has higher impedance.

[2] **Problem 9** (PPP 170)**.** Consider each of the following circuits.



In each case, find the amplitude of the current drawn from the source as a function of ω/ω_0, where $\omega_0 = 1/\sqrt{LC}$.

Solution. We deal with each case separately.

(a) We see that the top two inductors are in parallel, and the bottom two capacitors are also in parallel, with the two systems being in series. Thus, the net impedance is

$$Z = \frac{1}{2}\left(i\omega L + \frac{1}{i\omega C}\right),$$

which has magnitude $\frac{1}{2\omega C}|(\omega/\omega_0)^2 - 1|$. Thus,

$$|I| = \frac{V_0}{|Z|} = \frac{2\omega_0 C V_0(\omega/\omega_0)}{|(\omega/\omega_0)^2 - 1|}.$$

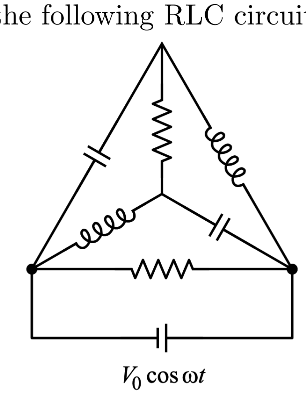
(b) The impedance here is

$$Z = 2\frac{1}{i\omega C + \frac{1}{i\omega L}},$$

so $|Z| = \frac{2\omega L}{|\omega^2/\omega_0^2 - 1|}$, so

$$|I| = \frac{V_0}{|Z|} = \frac{\omega_0 C V_0 |(\omega/\omega_0)^2 - 1|}{2(\omega/\omega_0)}.$$

[3] **Problem 10** (BAUPC)**.** Consider the following RLC circuit.



The capacitors have capacitance C, the inductors have inductance L, and the resistors have resistance $R = \sqrt{L/C}$. Furthermore, the driving angular frequency is $\omega = 1/\sqrt{LC}$. Find the amplitude of the total current through the circuit.

Solution. See the official solutions here (problem 4).

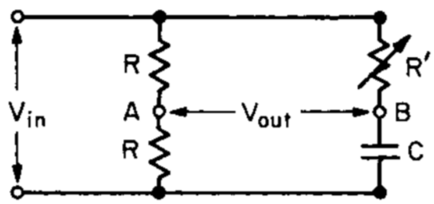
[3] **Problem 11.** USAPhO 2002, problem A1.

[3] **Problem 12.** USAPhO 2011, problem B1.

3 Electrical Engineering

These next problems are about using RLC circuits for practical purposes. They don't require anything not already introduced in the previous section, but they represent a different way of thinking that it's crucial to get comfortable with.

[2] **Problem 13** (Feynman)**.** In electronic circuits it is often desired to provide a sinusoidal voltage of constant amplitude but variable phase. A circuit which accomplishes this is called a phase-shifting network. One example is shown below.



Show that the voltage measured between terminals A and B has half the amplitude of the input voltage, and a phase which may be adjusted by changing the resistance R'.

Solution. Note that the current in the left branch is $I_1 = V_i/2R$, so $V_A = V_i - I_1 R = V_i/2$. We see that

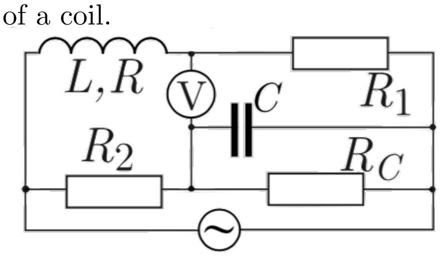
$$V_B = V_i - R' \frac{V_i}{R' + \frac{1}{i\omega C}} = V_i \left(1 - \frac{1}{1 + \frac{1}{i\omega CR'}} \right).$$

Thus,

$$V_B - V_A = V_i \left(\frac{1}{2} - \frac{1}{1 + \frac{1}{i\omega CR'}} \right) = -\frac{V_i}{2} \frac{1 - \frac{1}{i\omega CR'}}{1 + \frac{1}{i\omega CR'}}.$$

We clearly have $|V_B - V_A| = V_i/2$, and the second factor is a pure phase with phase $\phi = -2\tan^{-1}(1/\omega CR')$, which clearly can be adjusted to give any phase.

[3] **Problem 14** (Kalda)**.** The figure below shows a Maxwell's bridge, which is used for measuring the inductance L and resistance R of a coil.



To do this, the angular frequency ω is fixed and the known parameters R_1, R_2, R_C, and C are adjusted until the voltmeter reads zero. Once this is done, find R and L in terms of the other parameters.

Solution. Suppose the unknown coil has overall inductance Z. It suffices to find Z. We see that the voltage at the top end of the voltmeter is

$$V_1 = \mathcal{E}\frac{R_1}{Z + R_1}.$$

Let $Z_C = \frac{1}{\frac{1}{R_C} + i\omega C}$ denote the impedance of the resistive capacitor compound. The voltage at the other end of the voltmeter is then

$$V_2 = \mathcal{E}\frac{Z_C}{Z_C + R_2}.$$

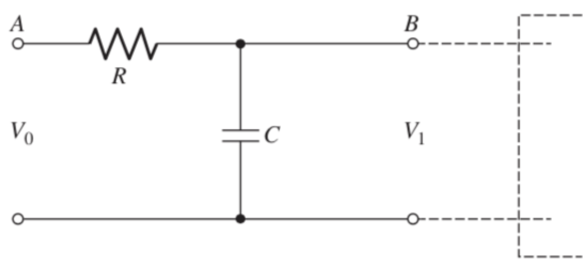
We have that $V_1 = V_2$, so $Z/R_1 + 1 = 1 + R_2/Z_C$, so

$$Z = R_1 R_2 (1/R_C + i\omega C).$$

On the other hand, we also have $Z = R + i\omega L$, so we can read off the answers,

$$R = R_1 R_2 / R_C, \quad L = R_1 R_2 C.$$

[5] **Problem 15.** An alternating voltage $V_0 \cos \omega t$ is applied to the terminals at A. The terminals at B are connected to an audio amplifier of very high input impedance. (That is, current flow into the amplifier is negligible.)



This circuit is the most primitive of "low-pass" filters.

(a) Calculate the "gain" ratio $|\tilde{V}_1|/V_0$ in this filter. Show that for sufficiently high frequencies, the signal power is reduced by a factor of 4 for every doubling of the frequency.

(b) Design a low-pass filter without using a capacitor.

(c) Design a high-pass filter.

(d) Design a stronger low-pass filter, i.e. one which reduces the signal power by a greater factor for every doubling of the frequency.

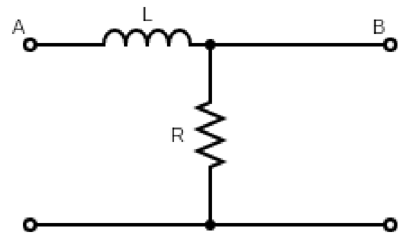
(e) Design a band-pass filter, which suppresses both low and high frequencies, but has a constant gain for a wide range of medium frequencies. (It's okay if the constant gain is less than 1, as we can just pass the output through an amplifier.)

(f) Design a notch filter, which suppresses a very small range of frequencies, while letting all other frequencies through.

Solution. (a) Letting the bottom be $V = 0$, we have $|V_1| = I|X_C|$ and $I = V_0/|X_C + R|$. This gives

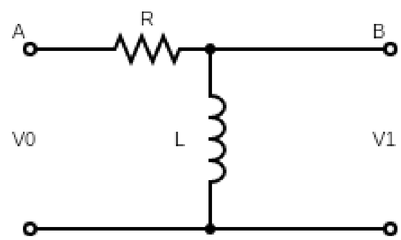
$$g = \frac{|V_1|}{V_0} = \frac{|X_C|}{|X_C + R|} = \frac{1}{\sqrt{\omega^2 R^2 C^2 + 1}}.$$

For high frequencies, $V_1 \approx V_0/\omega RC$, so the signal power $V_1^2/Z_1 \propto 1/\omega^2$, so doubling the frequency will reduce the power by a factor of 4.

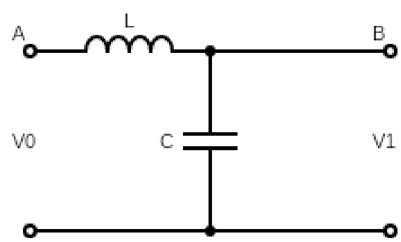


(b)

(c) Here's one possible answer. You could also take the low-pass filter in part (a) and switch the locations of the capacitor and resistor.



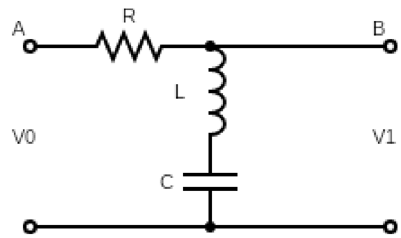
(d) This is the simplest answer.



(e) Attach a low pass filter to the output of a high pass filter to get a band-pass filter. There will then be a region in the middle with constant gain. For this to work straightforwardly, it is essential that the addition of the low pass filter doesn't affect the voltage output of the high pass filter, so that we can just multiply the gains. This occurs if the low pass filter draws negligible current from the output. (In terms of the example filters above, we need the resistance inside the low pass filter to be much higher than the resistance inside the high pass filter.)

Another option would be to attach the output to the resistor in a series RLC circuit, but then the "band" region would be too narrow.

(f) To suppress a certain frequency, make a circuit that looks like this:



Then make a chain of these circuits with varying values of R, L, C to suppress the desired frequencies (subject to the caveat in part (e)).

[4] **Problem 16.** IPhO 1984, problem 3. A nice, short problem on filters.

[2] **Problem 17.** Consider the same setup as problem 15, but with the resistor and capacitor switched.

(a) Assuming that $V_1 \ll V_0$, show that the output voltage is proportional to the derivative of the input voltage. Hence the circuit is a differentiator. (Can you relate this to the kind of filtering such a setup does?)

(b) Design a circuit whose output is proportional to the integral of V_0, again assuming $V_1 \ll V_0$.

Solution. (a) The output voltage will be

$$V_1 = V_0 \frac{\omega RC}{\sqrt{\omega^2 R^2 C^2 + 1}}.$$

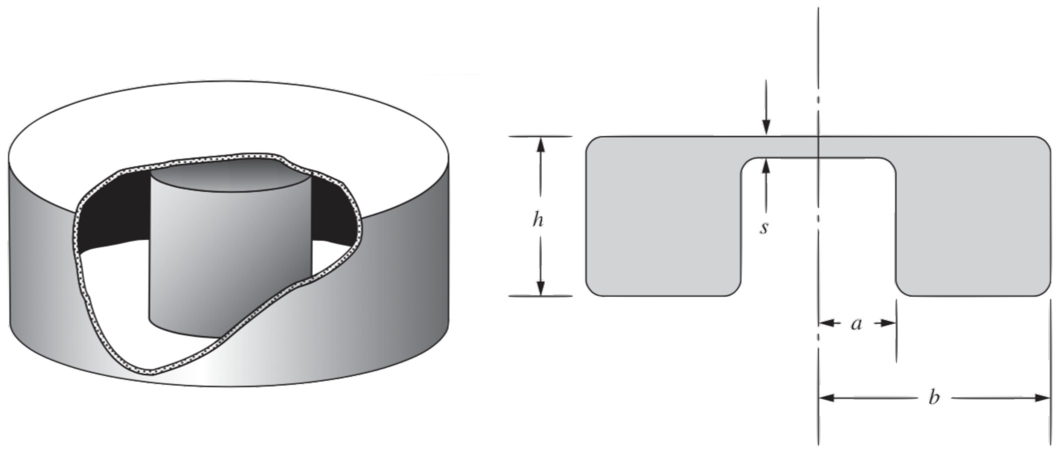
$V_1 \ll V_0$ means that $\omega RC \ll 1$, so $V_1 \approx V_0 \omega RC$. Since $dV_0/dt \propto \omega V_0$, we see that both V_1 and dV_0/dt are proportional to ωV_0. Since higher frequencies are emphasized, it's also a high pass filter.

(b) Now we want $V_1 \propto V_0/\omega$. Since the $X_L \propto \omega$ and $X_C \propto 1/\omega$, and we're looking for the opposite effect, it would make sense to try replacing the capacitor with an inductor.

$$V_1 = V_0 \frac{R}{\sqrt{R^2 + (\omega L)^2}}.$$

For $V_1 \ll V_0$, which indicates $R \ll \omega L$, we get $V_1 = V_0 R/\omega L$, which gets $V_1 \propto V_0/\omega$ as desired. This is also a low pass filter.

[3] **Problem 18.** A resonant cavity of the form illustrated below is an essential part of many microwave oscillators. It can be regarded as a simple LC circuit.



(a) Assuming that $s \ll a, b, h$, estimate the lowest resonant angular frequency of the cavity by treating it as an LC circuit. It may be helpful to sketch the magnetic and electric fields.

(b) One of the most common types of cavity is a cylindrical cavity, i.e. a hollow cylinder. (It corresponds to taking $s = h$ in the above setup.) Assuming that $h \approx b$, find a reasonable estimate of the lowest resonant angular frequency ω.

Solution. (a) The top of the small internal cylinder forms a small parallel plate capacitor with the top, with capacitance

$$C = \epsilon_0 \frac{\pi a^2}{s}.$$

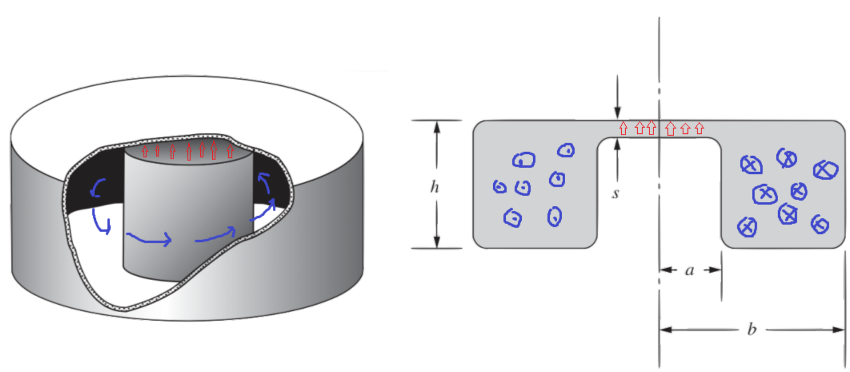
Meanwhile, the entire rest of the cavity looks like a toroidal solenoid with one turn, which we already know has an inductance of

$$L = \frac{\mu_0 h}{2\pi} \log \frac{b}{a}.$$

Therefore we have

$$\omega = \frac{1}{\sqrt{LC}} = \frac{1}{\sqrt{\frac{\mu_0 h}{2\pi} \log(b/a) \epsilon_0 (\pi a^2)/s}} = \frac{c}{a} \sqrt{\frac{2s}{h \log(b/a)}}$$

where c is the speed of light. The fields are sketched below.



(b) If we just plug in $s = h$ above, we get

$$\omega = \frac{c}{a}\sqrt{\frac{2}{\log(b/a)}}.$$

However, this result is nonsense, because it depends on a, which has no physical meaning when $s = h$. The problem is that our heuristic picture in (a) of how the current and charge is distributed only makes sense for $s \ll h$.

A complete and rather complicated analysis would show that the lowest resonant angular frequency is

$$\omega = c \times \min\left(1.841\left(\frac{1}{b^2} + \frac{2.912}{h^2}\right)^{1/2}, \frac{2.405}{b}\right).$$

In this case, we can get close by dimensional analysis, which tells us that $\omega \sim c/b$, since b is the only length scale in the problem. (Recall that we assumed $h \approx b$.)

Remark

In **E3**, we saw that for DC circuits, any system of resistors and ideal batteries with two ports is equivalent, from the perspective of anything connected across the ports, to either a single resistor and ideal battery in series (the Thevenin equivalent), or a single resistor and ideal current source in parallel (the Norton equivalent). From the ideas covered in this problem set, we also know that any system of resistors, inductors, and capacitors with two ports is equivalent, at a fixed angular frequency ω, to a single lumped element with impedance Z_{eq}. This in turn could be constructed out of a single resistor and inductor or capacitor in series.

This naturally leads to a more general question: it is possible to construct a simple "equivalent" circuit that has exactly the same $Z_{\text{eq}}(\omega)$, for *all* ω? The answer is yes. For example, consider the simple case of a circuit of only inductors and capacitors. Here's the rough idea: in this case, the equivalent impedance is always a pure imaginary, rational function of ω, meaning a ratio of two polynomials in ω. But rational functions can always be expanded in partial fractions. Assuming no multiple roots for simplicity, each term in the partial fraction decomposition can be mimicked with an LC circuit, and we get the sum by placing these circuits in series.

In electrical engineering, the general task of constructing a circuit with a prescribed $Z(\omega)$ is called network synthesis; the above example is called Foster's synthesis. These techniques can be used to construct filters more elaborate than the ones you explored in problem 15.

Remark

Power companies often transmit electricity with "three-phase power". This means that there are three "hot" electrical lines, carrying voltages

$$V_1(t) = V_0\cos(\omega t), \quad V_2(t) = V_0\cos(\omega t + 2\pi/3), \quad V_3(t) = V_0\cos(\omega t + 4\pi/3).$$

Depending on your home, you might be able to connect to this three-phase power with special outlets, to use power tools. There are several advantages to three-phase power, but one is that it supplies a constant power, as $V_1^2 + V_2^2 + V_3^2$ is constant.

This shouldn't be confused with the three holes in an ordinary American wall outlet. In these outlets, one of the eyes is the "hot" one, with voltage $V_1(t)$, while the other eye and the mouth are both grounded. Appliances are powered by the voltage difference between the eyes. Appliances that use significant power and have metal exteriors have three-prong plugs. Here, the grounded "mouth" hole is connected directly to the exterior of the appliance, ensuring that it can't shock you, even if something goes wrong inside.

4 Normal Modes

Idea 4

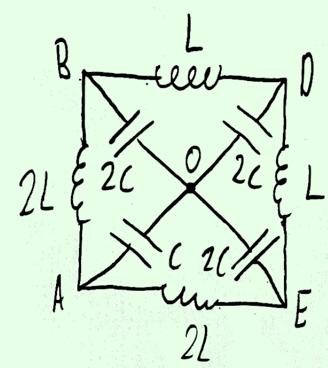
A circuit with n independent loops has n normal modes. If we ignore resistances, the normal modes are pure sinusoids, though in all real circuits they exponentially damp over time. Just as in mechanics, the general solution for the behavior of a driven circuit is a superposition of normal mode currents and the response to the driving.

There are many ways to find the normal mode frequencies.

- One way is to pick any two points not directly connected by wires. We may imagine that across these points we have attached a current source $\tilde{I}$ which is doing nothing, $\tilde{I} = 0$. If a normal mode is present at angular frequency ω, then we can have $\tilde{V} \neq 0$, even though $\tilde{I} = 0$ because current is merely sloshing around inside the circuit. Thus, the equivalent impedance $Z(\omega)$ between these points is infinite.
- Another way is to pick two points directly connected by wires. We may imagine this wire is actually a voltage source $\tilde{V}$ which is doing nothing, $\tilde{V} = 0$. If a normal mode is present at angular frequency ω, then we can have a current $\tilde{I} \neq 0$ through the wire even though $\tilde{V} = 0$, so the equivalent impedance $Z(\omega)$ between these points is zero.
- Some LC circuits can be mapped to sets of masses and springs using the analogies in idea 1, which can help with guessing the normal modes.
- Finally, one may simply write down all of Kirchoff's loop equations, plug in $e^{i\omega t}$ time dependence, and look for a solution. This boils down to solving a system of n equations, or equivalently evaluating the determinant of an $n \times n$ matrix. This is rarely the best approach on an Olympiad.
- Not every problem benefits from using normal modes; for relatively simple circuits with special initial conditions, it may be better to solve Kirchoff's loop equations directly.

Example 7: Kalda 89

Find the normal mode frequencies of the circuit below.

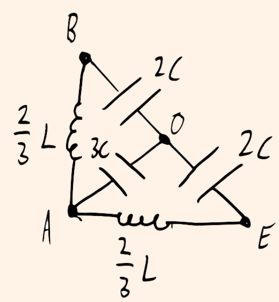


Solution

There are four independent Kirchoff's loop equations, so we expect four normal modes. One normal mode consists of current simply flowing uniformly along the outside, along the inductors. Since the capacitors aren't involved, this normal mode has $\omega_0 = 0$.

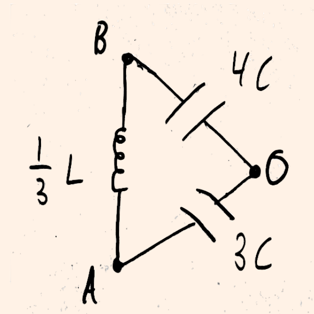
Now we apply the first technique listed above: we pick two points not directly connected with wires, and set the impedance to infinity. By symmetry, it's best to pick A and D. By symmetry, if any voltage is applied between A and D, the points B and E will be at the same voltage. Furthermore, this point will be at the same voltage as O, because the remaining circuit forms a balanced Wheatstone bridge, as introduced in **E3**. Identifying B, E, and O straightforwardly gives a simple LC circuit with $L_{\text{eff}} = (3/2)L$ and $C_{\text{eff}} = (2/3)C$, and resonant angular frequency $\omega_1 = 1/\sqrt{L_{\text{eff}}C_{\text{eff}}} = 1/\sqrt{LC}$.

This procedure only gave one of the three remaining normal modes, so we must have missed the other two because they have zero voltage difference between A and D. Therefore, to find the other two, we can join A and D, leading to the simpler equivalent circuit below.



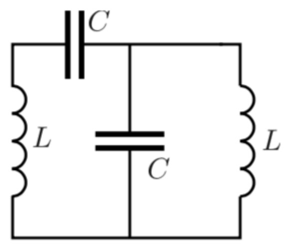
We now apply the same procedure between points B and E. This circuit is again a balanced Wheatstone bridge, so O and A are at the same voltage. We then have a simple LC circuit with $L_{\text{eff}} = (4/3)L$ and $C_{\text{eff}} = C$, giving $\omega_2 = \sqrt{3/4LC}$.

Again, we've missed a normal mode, so that remaining mode must have zero voltage difference between B and E. Joining them together leads to the final equivalent circuit below.



This is now a simple LC circuit with $L_{\text{eff}} = (1/3)L$ and $C_{\text{eff}} = (12/7)C$, giving the final resonant angular frequency $\omega_3 = \sqrt{7/4LC}$.

[2] **Problem 19** (Kalda). Consider the LC circuit below.



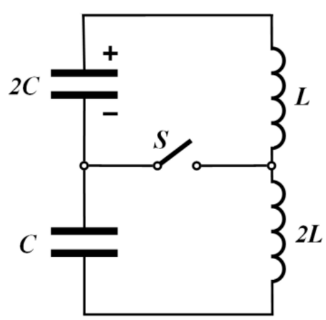
Show that the normal mode angular frequencies are $\omega = (\sqrt{5} \pm 1)/2\sqrt{LC}$.

Solution. We set the inductance between the two ends of the bottom left wire to be 0, so

$$i\omega L + \frac{1}{i\omega C} + \frac{1}{i\omega C + \frac{1}{i\omega L}} = 0.$$

Let $a = i\omega L$ and $b = \frac{1}{i\omega C}$. We have $a + b + 1/(1/a + 1/b) = 0$, so $(a/b)^2 + 3(a/b) + 1 = 0$, so $\omega^2 LC = -a/b = \frac{3\pm\sqrt{5}}{2}$. But note $(\sqrt{5}\pm 1)^2 = 2(3\pm\sqrt{5}) = 4\omega^2 LC$, which shows $\omega = (\sqrt{5}\pm 1)/2\sqrt{LC}$.

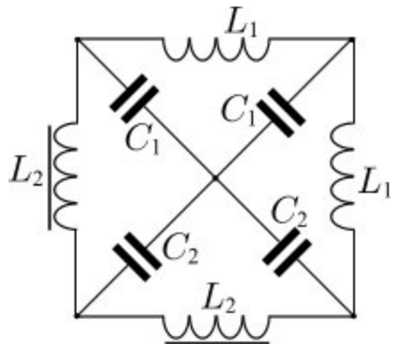
[3] **Problem 20** (IPhO 2014). Initially, the switch S is open in the circuit shown below.



The capacitor with capacitance $2C$ is given a charge q_0, and immediately begins to discharge. At the moment when the current through the inductors reaches its maximum value, the switch S is closed. Find the maximum current through the switch thereafter.

Solution. See the official solutions here.

[5] **Problem 21** (Physics Cup 2012). Find the angular frequencies of the normal modes of the circuit below, where $C_1 \ll C_2$ and $L_1 \ll L_2$.



You may give all of your answers to lowest order in C_1/C_2 and L_1/L_2.

Solution. See here for many solutions.

5 Nonlinear Circuit Elements

In this section we'll introduce nonlinear circuit elements, focusing on diodes. More exotic circuit elements will be covered in **E7**.

Idea 5

Many nonlinear circuit elements can be described by a current-voltage characteristic $I(V)$. Such circuit elements have trivial time dependence, just like resistors, and working with them basically amounts to using Kirchoff's laws as usual, plugging in $I(V)$ where necessary.

Since the implementation details of such elements can be very complicated, and many draw power from external sources, it generally isn't productive to think of them "physically"; they are more like miniature computers than physical objects. One just has to take $I(V)$ as given and work directly with it. Some simple examples are:

- An ideal diode acts like a wire in one direction and a break in the other, so it has

$$I(V) = \begin{cases} \infty & V > 0, \\ 0 & V < 0. \end{cases}$$

- Sometimes one instead takes the $I(V)$ characteristic

$$I(V) = \begin{cases} \infty & V > V_0, \\ 0 & V < V_0 \end{cases}$$

  which means that it "costs" voltage V_0 to go through the diode in the forward direction. More realistically, $I(V)$ smoothly increases when V passes V_0, but you don't often see this in Olympiad problems because it makes the math very messy.

- Zener diodes are bidirectional diodes. An ideal Zener diode has

$$I(V) = \begin{cases} \infty & V > V_0, \\ 0 & -V_0 < V < V_0, \\ -\infty & V < -V_0. \end{cases}$$

- Many familiar objects such as fuses (wires which break when I passes a threshold) and spark gaps (breaks that conduct when V passes a threshold) can be thought of as nonlinear circuit elements in the same way.

Analytically, these three cases are easily handled by casework. For instance, a diode acts just like a wire for positive V, and just like a break for negative V. In each case, the circuit is no more complicated than an ordinary one with linear circuit elements. Then you put the cases together to get the full behavior.

Example 8

A capacitor of capacitance C is charged so that its voltage is V_C. The capacitor is placed in series with a resistor R and a diode with $I(V)$ characteristic

$$I(V) = \begin{cases} \infty & V > V_0, \\ 0 & V < V_0. \end{cases}$$

The diode is oriented so that the initial voltage across it is positive. What happens next?

Solution

We use casework. If $V_c < V_0$, the voltage on the capacitor is not enough to get current to flow through the diode, so nothing happens. If $V_c > V_0$, current flows, and the diode acts like a battery of emf V_0 oriented in the opposite direction. This is just a discharging RC circuit, so the capacitor's voltage is

$$V(t) = (V_C - V_0)e^{-t/RC} + V_0.$$

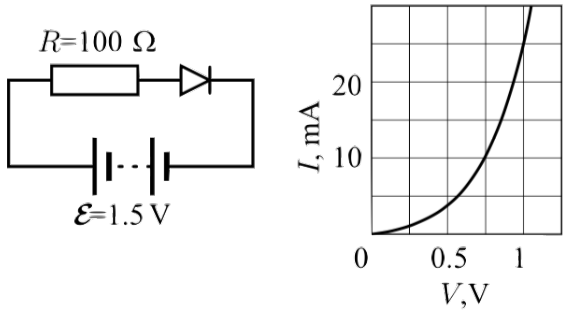
After a long time, the voltage on the capacitor falls to V_0 and current stops flowing.

Idea 6

It is difficult to solve a nonlinear circuit analytically if $I(V)$ is not very simple. In these cases:

- One can find the answer graphically as the intersection of $I(V)$ and another curve.
- One can solve for the answer iteratively on a calculator.
- If V stays within a narrow range, one can take a linear approximation to $I(V)$. This effectively replaces the element with a battery in series with a resistor, so the problem can be solved just like those in **E3**.

[2] **Problem 22** (Kalda). Find the current in the circuit given below.



The nonlinear element is a diode with the $I(V)$ characteristic shown.

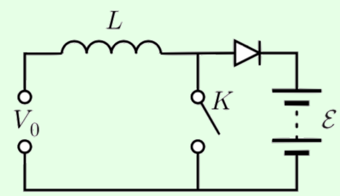
Solution. We see that $1.5 = 100I + V(I)$ where all numbers are in the proper SI units. We want to intersect $1.5 - 100I$ with $V(I)$. Doing this we see $I \approx \boxed{8\,\text{mA}}$.

Idea 7

The power delivered to any circuit element is still $P = IV$. However, some nonlinear circuit elements can be active, providing net power to the circuit, like batteries.

Example 9: Kalda 64

The circuit below containing an ideal diode makes it possible to charge a rechargeable battery of voltage $\mathcal{E} = 12\,\text{V}$ with a direct voltage source of a voltage $V_0 = 5\,\text{V} < \mathcal{E}$.



To do this, the switch K is periodically opened and closed, with the opened and closed periods having equal length $\tau = 10\,\text{ms}$. Find the average charging current assuming $L = 1\,\text{H}$.

Solution

The intuition here is that, using an inductor and a switch, one can generate emfs larger than what we put in, because the current wants to keep flowing through the inductor when the switch is opened; this allows us to get enough emf to charge the battery. This idea is also used in the ignition coils of old-fashioned cars, where a voltage large enough to ionize air is produced, making a spark and starting the engine. There's also a fluid analogue, called the hydraulic ram, used to raise water. The point of the diode here is just to keep current from flowing the other way during the other half of the cycle.

When the switch is closed, no current can flow through the battery, and the current through the inductor builds up linearly, since there is an emf V_0 across the inductor. When the switch is opened, the emf across the inductor is $V_0 - \mathcal{E} = -7\,\text{V}$, causing its current to decrease while simultaneously charging the battery. After a time $(5/7)\tau$ with the switch open, the current through the inductor falls to zero, and the diode causes current to stop flowing.

Quantitatively, while the switch is closed, the current through the inductor builds up to $V_0\tau/L$. When the switch is open, current flows for a time $(5/7)\tau$, linearly falling to zero, so the total charge is

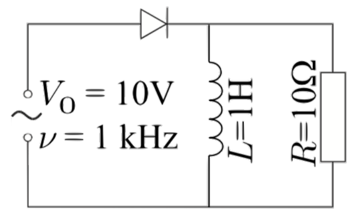
$$Q = \frac{1}{2}\frac{V_0\tau}{L}\frac{5}{7}\tau.$$

A cycle takes time 2τ, so

$$\overline{I} = \frac{Q}{2\tau} = \frac{5}{28}\frac{V_0\tau}{L} = 8.9\,\text{mA}.$$

By the way, your phone and laptop chargers probably have rectangular bricks containing a switched-mode power supply. This consists of one part that converts the AC wall power to DC, and a second part similar to the circuit above, but set up to output a *lower* DC voltage. You could also use a transformer to lower the AC voltage, but a switch-mode power supply is more space-efficient, and it easily copes with a range of input AC voltages and frequencies.

[3] **Problem 23** (Kalda)**.** An alternating voltage $V = V_0\cos(2\pi\nu t)$ is applied to the leads of the circuit shown below. Treat the diode as ideal.



Assuming the current in the inductor begins at zero, what is the average current through the inductor at late times?

Solution. Since $\omega L \gg R$, the inductor's current changes very slowly, so we can neglect its change over any one cycle. During some cycle, let's write the steady state current in the inductor as

$$\overline{I}_L = \alpha(V_0/R)$$

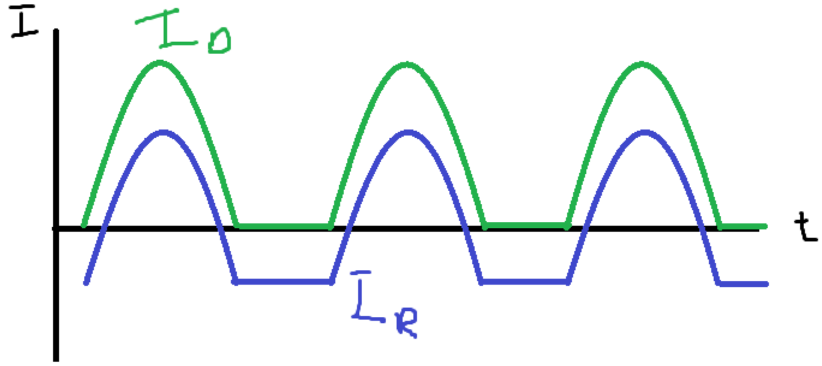
where $\alpha = 0$ in the beginning. Let the current through the resistor be I_R. The current through the diode is $I_D = \overline{I}_L + I_R$. When the diode is open, $I_D > 0$, the voltage across the inductor is

$$V_L = V_0\cos(2\pi\nu t).$$

During this time, the current through the diode is a shifted sinusoid,

$$I_D = \overline{I}_L + \frac{V_0}{R}\cos(2\pi\nu t).$$

The diode closes once I_D falls to zero. Thus, for $\alpha = 0$ the diode is closed half the time, while for $\alpha = 1$ the diode is never closed. The situation for $\alpha \approx 0.5$ is shown below.



When the diode is closed, the voltage across the inductor is

$$V_L = I_R R = -\overline{I}_L R.$$

The net change in $\overline{I}_L$ is one cycle is

$$\Delta \overline{I}_L = \frac{1}{L} \int_{\text{cycle}} V_L \, dt.$$

In the beginning, when $\alpha = 0$, this integral is positive because $V_L(t)$ looks like a sinusoid but with only the positive parts. As α increases, the integral begins to pick up part of the negative half of the sinusoid, but the overall integral is still positive, so α continues to increase. The final steady state is when $\alpha = 1$ and the current is open all the time. At this point, $\overline{I}_L = V_0/R = 1\,\text{A}$.

[3] **Problem 24.** EFPhO 2010, problem 9.

Solution. See the official solutions here.

[3] **Problem 25.** EFPhO 2008, problem 6.

Solution. See the official solutions here.

[3] **Problem 26.** EFPhO 2013, problem 8. This one has a nice mechanical analogy.

Solution. See the official solutions here.

[3] **Problem 27.** IPhO 2001, problem 1c.

[3] **Problem 28.** USAPhO 2018, problem A2.

[4] **Problem 29.** EuPhO 2022, problem 2. A nice application of casework.

Solution. See the official solutions here.

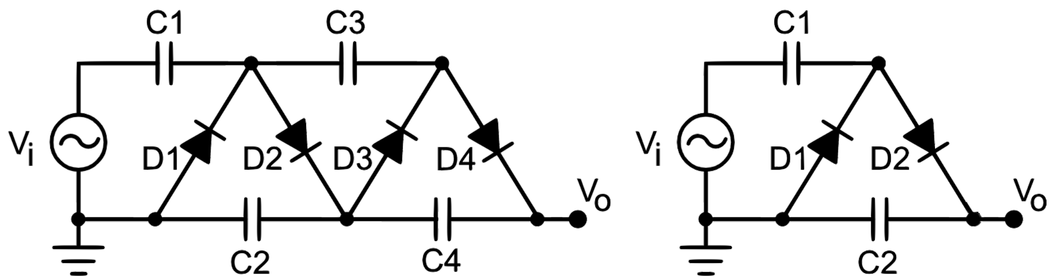
Electromagnetism VII: Electrodynamics

Chapter 9 of Purcell covers electromagnetic waves, and appendix H covers radiation by charges. For a pedagogical introduction with solved examples, see recitation 8 and recitation 9 of the MIT OCW 8.03 lectures. For more technical coverage, not necessarily relevant to the Olympiad, see chapters 7, 8, 10, 11 of Griffiths. For some lighter reading, see chapters I-28, I-32, II-18, II-20, II-21, and II-24 of the Feynman lectures. There is a total of **85** points.

1 More Nonlinear Circuit Elements

In this section we consider some more subtle applications of nonlinear circuit elements. First, we consider complex problems that use relatively familiar circuit elements.

[3] **Problem 1.** The setup below at left is called the Cockcroft–Walton voltage multiplier. It was used in 1932 to power the first particle accelerator.

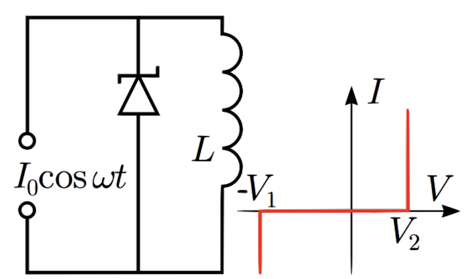


The four capacitors begin uncharged, and all have capacitance C. All four diodes are ideal. The voltage $V_i(t)$ alternates between V and $-V$. The output voltage is V_0.

(a) To warm up, consider the simpler setup shown at right above. Suppose the applied voltage V_i begins at $-V$. Describe how V_0 changes each time the applied voltage switches sign.

(b) Now consider the full setup, shown at left above. After a long time, what is V_0?

[3] **Problem 2** (NBPhO 2017). A Zener diode is connected to a source of alternating current as shown.



The inductance L of the inductor is such that $L\omega I_0 \gg V_1, V_2$ where V_1 and V_2 are the breakdown voltages, and $V_1 > V_2$. The $I(V)$ characteristic of the Zener diode is shown above. Assume that a long time has passed since the current source was first turned on.

(a) Find the average current through the inductor.

(b) Find the peak-to-peak amplitude of the current changes ΔI in the inductor.

[4] **Problem 3.** NBPhO 2016, problem 4. A rather complicated problem involving several exotic circuit elements.

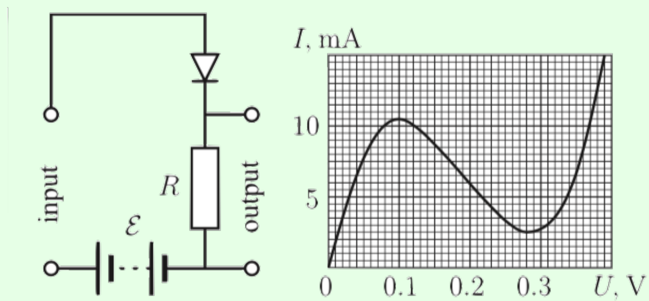
Next, we consider some qualitatively new behavior that can emerge from less familiar circuit elements, such as amplification, hysteresis, and instability.

Idea 1

Tunnel diodes are a variant of diodes, whose $I(V)$ rises, falls, and rises again. That is, they have a region with negative differential resistance, $dI/dV < 0$. This allows them to amplify signals, as we'll see below, and also can make them unstable.

Example 1: EFPhO 2003

The circuit below, containing a tunnel diode, acts as a simple amplifier.



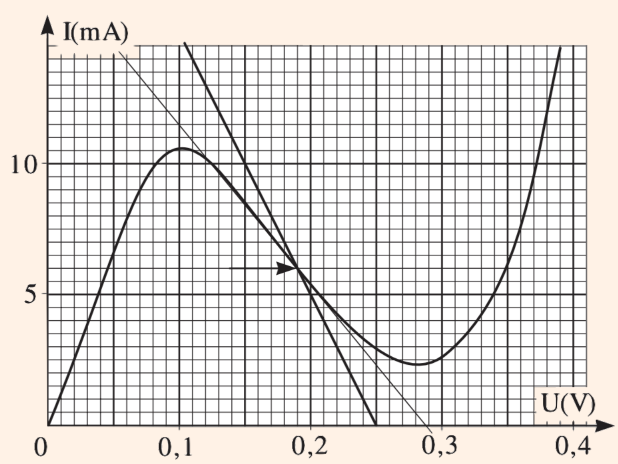
Here, $R = 10\,\Omega$ and $\mathcal{E} = 0.25\,\text{V}$. If a small signal voltage $V_{\text{in}}(t)$ is applied across the input, then an amplified and shifted version of the signal appears across the output. Find the amplification factor.

Solution

When a constant emf $\mathcal{E}$ is applied, Kirchoff's laws give

$$\mathcal{E} = I_0 R + V(I_0)$$

where $V(I)$ is the voltage characteristic of the diode.



By plotting $\mathcal{E} - IR$ on the graph above, we find I_0 at the intersection. (Notice the x-axis label: it is common to write a decimal point as a comma in Eastern Europe.) Now consider the effect of applying the signal voltage, which changes the current by ΔI,

$$\mathcal{E} + V_{\text{in}} = (I_0 + \Delta I)R + V(I_0 + \Delta I).$$

Since the signal voltage is small, we can Taylor expand the voltage characteristic, giving

$$V_{\text{in}} = \Delta I(R + V'(I_0)).$$

This in turn tells us that

$$V_{\text{out}} = (I_0 + \Delta I)R = V_{\text{out}}^0 + \frac{R}{R + V'(I_0)} V_{\text{in}}.$$

In other words, the change in V_{out} is just V_{in}, times the amplification factor

$$\frac{R}{R + V'(I_0)} = \frac{10}{10 - 16} = -\frac{5}{3}$$

where we read $V'(I_0)$ off the graph by drawing a tangent. The intuition here is that the circuit is like a voltage divider, but the tunnel diode acts like a negative resistance. If we had $V'(I_0)$ close to $-R$, for example, the amplification factor would have been huge. Since $V'(I_0)$ is more negative than $-R$, the signal ends up flipped.

[4] **Problem 4.** NBPhO 2020, problem 2. A comprehensive problem on the measurement and dynamics of tunnel diodes, which will give you a deeper understanding of negative resistance.

Idea 2

Op amps have four terminals, and output a voltage across the last two equal to the voltage across the first two, times a very large gain. Like tunnel diodes, op amps can be unstable: increasing the input increases the output, but this in turn could increase the input again. Thus, in practice, the output and input are always wired together in a way that produces negative feedback, with changes in the output acting to decrease the input. In this case, one can think of an op amp as a tool that tries to set the input voltages equal to each other.

The internals are somewhat complicated, consisting of a lot of resistors and transistors. In general, to analyze setups with multiple complex circuit elements like these, it's better to treat them as black boxes than to try to intuit what's going on at the level of individual subelements, or electric and magnetic fields. (Of course, engineers do need to understand circuit elements at these level to design them in the first place!)

[3] **Problem 5.** USAPhO 2016, problem A2. This problem is a nice introduction to op amps.

Idea 3

In some nonlinear circuit elements, the function $I(V)$ is multivalued. This indicates hysteresis: given V, the actual value of I depends on the history of the system. The same goes for when

$V(I)$ is multivalued.

[5] **Problem 6.** IPhO 2016, problem 2. This problem illustrates the previous idea with a thyristor. Print out the official answer sheet and record your answers on it.

2 Displacement Current

Idea 4

In general, Ampere's law is

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

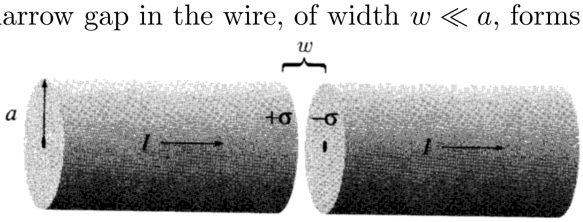
This is sometimes written in terms of a "displacement current" density $\mathbf{J}_d$, where

$$\nabla \times \mathbf{B} = \mu_0(\mathbf{J} + \mathbf{J}_d), \quad \mathbf{J}_d = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

In integral form, for a surface S bounded by a closed curve C,

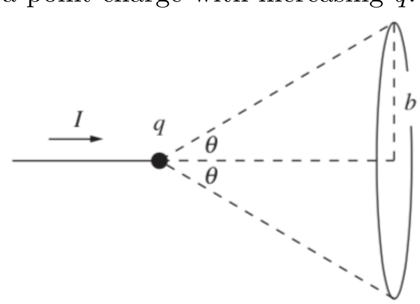
$$\oint_C \mathbf{B} \cdot d\mathbf{s} = \mu_0 \int_S (\mathbf{J} + \mathbf{J}_d) \cdot d\mathbf{S} = \mu_0 I + \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}.$$

[2] **Problem 7** (Griffiths 7.34)**.** A fat wire of radius a carries a constant current I uniformly distributed over its cross section. A narrow gap in the wire, of width $w \ll a$, forms a parallel plate capacitor.



Find the magnetic field in the gap, at a distance $s < a$ from the axis.

[3] **Problem 8** (Purcell 9.3)**.** A half-infinite wire carries constant current I from negative infinity to the origin, where it builds up at a point charge with increasing q. Consider the circle shown below.



Calculate the integral $\int \mathbf{B} \cdot d\mathbf{s}$ about this circle in three ways.

(a) Use the integrated form of Ampere's law, integrating over a surface which does not intersect the wire.

(b) Do the same, with a surface that does intersect the wire.

(c) Apply the Biot–Savart law to the current and displacement current.

In the previous problem, you should have found that the effect of the displacement current, in the Biot–Savart law, simply cancelled out everywhere. In fact, this cancellation is very general.

Idea 5

In any situation where $\mathbf{J}$ is constant, whether or not ρ is constant, Maxwell's equations are satisfied by applying Coulomb's law to ρ and the Biot–Savart law to $\mathbf{J}$. You can include displacement currents in the Biot–Savart integral too, but their contributions perfectly cancel.

To see why, note that

$$\nabla \times \mathbf{J}_d = \epsilon_0 \nabla \times \frac{\partial \mathbf{E}}{\partial t} = \epsilon_0 \frac{\partial}{\partial t}(\nabla \times \mathbf{E}) = -\epsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2}$$

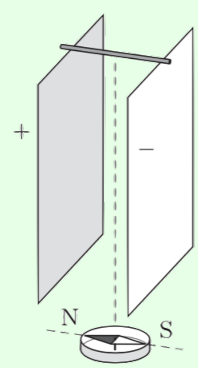
where we used Faraday's law. When the currents are constant, the magnetic fields are also constant, so the right-hand side vanishes. Then $\nabla \times \mathbf{J}_d = 0$. However, this means that $\mathbf{J}_d$ can always be written as a superposition of radial, spherically symmetric currents, and as we saw in the previous problem, such currents produce no magnetic fields.

This explains why we were able to get away with using Coulomb's law and the Biot–Savart law on previous problem sets, even in situations which were not exactly electrostatic or magnetostatic – all of these situations were "quasistatic". In general, the displacement current only matters in non-quasistatic situations involving rapid changes in $\mathbf{J}$, and hence rapid changes in $\mathbf{E}$ and $\mathbf{B}$. These are exactly the cases where significant electromagnetic radiation is produced, which is why radiation is covered in the last half of this problem set.

For more about this subtle point, see section 9.2 of Purcell, this paper and this paper.

Example 2: MPPP 190

A parallel plate capacitor is charged and positioned above a compass as shown.



The capacitor is discharged slowly when the tops of the plates are joined using a small conducting rod. Which way is the compass needle deflected during the discharge process?

Solution

Since the discharge is slow, the situation is quasistatic. (It would only be non-quasistatic in the case where the discharge time was comparable to the time it would take for light to cross the capacitor, a situation which is almost never achieved for RC circuits.) Then we know the magnetic field due to the displacement current cancels out everywhere, so only the current I in the rod matters. This current moves left to right, so by a straightforward application of the right-hand rule, we find that the compass is deflected east.

Things get more subtle if one insists on considering the displacement current anyway. A naive, incorrect argument would be to say that there is a total displacement current I going right to left inside the capacitor, and since this displacement current is closer than the current in the rod, it produces a larger magnetic field, deflecting the compass west. The problem with this reasoning is that it has ignored the displacement current due to the changes in the substantial fringe fields of the capacitor. When these are accounted for, the magnetic fields due to the displacement current cancel, as argued generally above.

[3] **Problem 9** (Griffiths 7.36). An alternating current $I = I_0 \cos(\omega t)$ flows down a long straight wire along the $\hat{\mathbf{z}}$ axis and returns along a coaxial conducting tube of radius a.

(a) Assuming the electric field goes to zero at infinity, show that

$$\mathbf{E} = \frac{\mu_0 I_0 \omega}{2\pi} \sin(\omega t) \log \frac{a}{r} \, \hat{\mathbf{z}}.$$

(b) Find the amplitude I_d of the total displacement current.

(c) Compute the ratio I_d/I_0. Show that I_d is only significant when ω is fast enough that the speed-of-light travel time from the wire to the tube is comparable to the period.

This shows another sense in which displacement current effects are only significant when radiation comes into play. If I_d/I were near one, we would have to then consider the magnetic fields induced by the changing electric fields associated with the displacement current, and then the displacement currents due to the changes in those magnetic fields, and so on. But this description of electric fields inducing magnetic fields and vice versa is just a description of an electromagnetic wave.

[3] **Problem 10.** [A] Consider an infinite thin solenoid which initially carries no current, and a loop of wire around this solenoid of enormous radius, say one light year. At some moment, a current is suddenly made to flow through the solenoid. (This cannot be done by simply attaching a battery somewhere, because it will take a long time for the current to turn on throughout the solenoid. So instead, consider a situation where many batteries arranged around the solenoid are all attached in at once, which can be achieved by machines which have synchronized their clocks beforehand.)

A magnetic field is hence quickly produced in the solenoid, so by Faraday's law,

$$\mathcal{E} = -\frac{d\Phi_B}{dt}$$

there should quickly be an emf in the loop of wire. But this seems to violate locality, because the motion of charges in the solenoid is quickly affecting the motion of charges in the wire loop, which is very far away. What's going on? Could there be something wrong with Faraday's law?

[3] **Problem 11** (Griffiths 7.64). [A] Setting $\mu_0 = \epsilon_0 = 1$, Maxwell's equations read

$$\nabla \cdot \mathbf{E} = \rho_e, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{B} = \mathbf{J}_e + \frac{\partial \mathbf{E}}{\partial t}$$

where ρ_e and $\mathbf{J}_e$ are the electric charge density and electric current density.

(a) Show that Maxwell's equations ensure the conservation of electric charge,

$$\dot{\rho}_e = -\nabla \cdot \mathbf{J}_e.$$

This is the continuity equation, and we saw versions of it for other conserved quantities in **T2**.

(b) Generalize Maxwell's equations to include a magnetic charge density ρ_m and a magnetic current density $\mathbf{J}_m$. Fix the signs by demanding that magnetic charge is conserved.

(c) Check that the resulting equations are invariant under the duality transformation

$$\begin{pmatrix} \mathbf{E}' \\ \mathbf{B}' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{B} \end{pmatrix}, \quad \begin{pmatrix} \rho'_e \\ \rho'_m \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \rho_e \\ \rho_m \end{pmatrix}$$

which rotates electricity into magnetism with angle θ.

(d) Write down the Lorentz force law for a particle with electric and magnetic charge, using the fact that it should be invariant under the duality transformation above.

Remark

The peculiar name of $\mathbf{J}_d$ is because Maxwell thought of it as a literal displacement of a jelly-like ether. In that era, all electromagnetic quantities, such as fields, charges, currents, polarizations, and magnetizations, were thought to reflect properties of a mechanical ether, such as local strains, displacements, and rotations. However, making this picture precise was known to be difficult even before the advent of relativity, which rendered ether obsolete. The best way to understand why physicists abandoned ether models is to have a look at their daunting complexity. For a nice overview, with diagrams, see chapter 4 of *The Maxwellians.*

3 Field Energy and Momentum

Idea 6

The Poynting vector

$$\mathbf{S} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0}$$

gives the flux density of the energy of an electromagnetic field. That is, the flux of $\mathbf{S}$ into a closed surface is the rate of change of total field energy within that surface.

[3] **Problem 12.** USAPhO 2010, problem B2.

[3] **Problem 13.** USAPhO 2013, problem B2.

Remark

It's unlikely that you'll see any examples besides the ones in the above two problems, because in almost all other setups, the Poynting vector depends sensitively on the fringe fields, which are very hard to calculate. (For some work in this direction, see *Energy transfer in electrical circuits: A qualitative account*.) In any case, the examples above illustrate the important point that the energy of a circuit does not flow along the wire, carried by the charges; instead it flows into circuit elements from the sides. This was an important early clue of the importance of the electromagnetic field.

Remark

As proven in Poynting's theorem, the Poynting vector indeed tells us about the net flow of energy. However, this would remain true if we added a constant vector to it, or more generally any divergence-free vector field, since these wouldn't change the net flow. So which option is the "correct" one? According to everything we've learned so far, there's no absolute way to choose, and we just use the Poynting vector because it's the simplest option. However, in general relativity, the flow of energy directly influences the curvature of spacetime, so there is an unambiguous correct answer, which is indeed the Poynting vector.

Example 3

Consider two charges q, at positions $r\hat{\mathbf{x}}$ and $r\hat{\mathbf{y}}$ respectively, both moving with speed v towards the origin. Show that the magnetic forces between them are *not* equal and opposite. That is, electromagnetic forces do not obey Newton's third law.

Solution

In order to find the $\mathbf{B}$ field produced by each charge at the location of the other, we use the Biot–Savart law and the right-hand rule. Then we use the Lorentz force and the right-hand rule again to find the magnetic forces on each charge.

For example, the $\mathbf{B}$ field produced by the first charge at the location of the second is along $-\hat{\mathbf{z}}$. Then the magnetic force on the second charge is parallel to $\hat{\mathbf{x}}$. The magnetic force on the first charge is parallel to $\hat{\mathbf{y}}$. And the force are definitely nonzero, so they can't be equal and opposite.

To explain this, we recall that the point of Newton's third law is just momentum conservation. This still holds, as long as one remembers that the field carries momentum of its own. (If we want to save some version of Newton's third law, we could say that the real action-reaction pairs are the forces between the charges and the field, not the charges with each other. But the real lesson is that Newton's third law is not fundamental, momentum conservation is.)

Idea 7

The momentum density of the electromagnetic field is

$$\mathbf{p} = \frac{\mathbf{S}}{c^2}.$$

In other words, momentum density and energy flux density are just proportional. As you will see in **R2**, this is true in general in relativity. The angular momentum density is $\mathbf{r} \times \mathbf{p}$. For an explicit derivation that these definitions ensure the total momentum and angular momentum are conserved, see section 8.2 of Griffiths. (You might think the definitions come out of nowhere; the straightforward way to find them is to apply Noether's theorem, as you will learn in a more advanced class.)

Remark

We have already seen an example of electromagnetic field momentum at work. Back in **E4**, you found that in the presence of a magnetic monopole, the mechanical angular momentum $\mathbf{L}$ of a point charge was not conserved, but $\mathbf{L} - qg\hat{\mathbf{r}}$ was. In fact, this second term turns out to be exactly the angular momentum of the field, so this conservation law is simply the conservation of total angular momentum. (If you'd like to verify this explicitly, it's easiest to use spherical coordinates with the monopole at the origin and the charge along the z-axis, but be warned, it's fairly messy.)

[3] **Problem 14** (Griffiths)**.** A long coaxial cable of length ℓ consists of an inner conductor of radius a and an outer conductor of radius b. The inner conductor carries a uniform charge per unit length λ, and a steady current I to the right; the outer conductor has the opposite charge and current.

(a) Find the electromagnetic momentum stored in the fields.

(b) In part (a) you should have found that the fields contain a nonzero momentum directed along the cable. However, this is puzzling because it appears that no net mass is transported along the cable. How is this paradox resolved? (Hint: it doesn't make sense to consider the cable in isolation, as nothing would be keeping the current going. Consider attaching a battery across the left end and a resistor across the right end.)

[3] **Problem 15.** In the early 20$^{\text{th}}$ century, physicists sought to explain the $E = mc^2$ rest energy in terms of electromagnetic field energy. As a concrete example, model a charged particle as a uniform spherical shell of radius a and charge q.

(a) Find the radius a so that the total field energy equals the rest energy associated with the electron mass m. Up to an $O(1)$ factor, this quantity is called the classical electron radius.

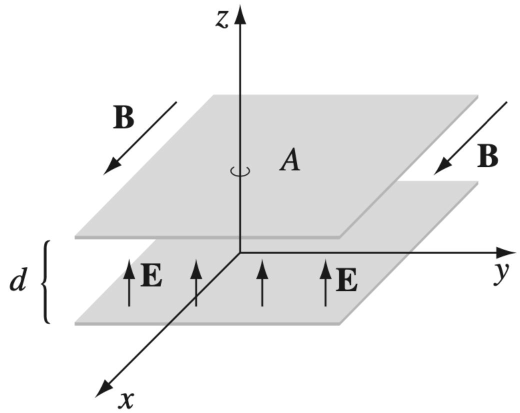
(b) If the shell moves with a small speed v, we expect to have $p = mv$, where p is the total field momentum. Show that instead, we have $p = (4/3)mv$. You may use the result

$$\mathbf{B} = \frac{\mathbf{v}}{c^2} \times \mathbf{E}$$

which we will prove in **R3**. Many complicated ideas were put forth to explain this infamous "4/3 problem", as recounted in chapter II-28 of the Feynman lectures.

For more about the “radius” of an electron, see this blog post. For a modern discussion of the resolution of the 4/3 problem, see this paper.

[3] **Problem 16** (Griffiths 8.6). A charged parallel plate capacitor is placed in a uniform magnetic field as shown.



(a) Find the electromagnetic momentum in the space between the plates.

(b) Now a resistive wire is connected between the plates, along the z-axis, so that the capacitor slowly discharges. The current through the wire will experience a magnetic force; show the total impulse equals the stored momentum.

(c) Alternatively, suppose we slowly reduced the magnetic field. Show that the total impulse delivered to the plates equals the stored momentum.

This calculation is standard and given in many textbooks, but it is actually completely wrong: we have ignored the fringe field, and when it is included the total electromagnetic momentum is half of what was naively calculated in part (a). The answer in part (b) is correct, but the other half of the impulse corresponds to a change in non-electromagnetic “hidden momentum”. The most basic example of hidden momentum is covered in example 12.12 of Griffiths. For a detailed analysis of the hidden momentum in this setup, see this paper.

[3] **Problem 17.** USAPhO 2004, problem B2. (This is a classic setup which also appears on USAPhO 2020, problem A1, and INPhO 2020, problem 2. But note that the official solution to USAPhO 2020, problem A1 has typos.)

4 Electromagnetic Waves

Idea 8

Maxwell’s equations have propagating wave solutions of the form

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \mathbf{B} = \mathbf{B}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$$

where $\mathbf{E}$ and $\mathbf{B}$ are in phase, perpendicular in direction, and have magnitudes $E_0 = cB_0$. The propagation direction $\mathbf{k}$ is along $\mathbf{E}\times\mathbf{B}$, and the wave speed is

$$c = \frac{\omega}{k} = \frac{1}{\sqrt{\mu_0\epsilon_0}}.$$

Example 4

Verify explicitly that in the absence of charges and currents, the electromagnetic field above satisfies Maxwell's equations.

Solution

First let's consider Gauss's law, $\nabla \cdot \mathbf{E} = 0$. Splitting everything explicitly into components,

$$\begin{aligned}\nabla \cdot \mathbf{E} &= e^{-i\omega t}\left(\frac{\partial}{\partial x}(E_{0,x}e^{i\mathbf{k}\cdot\mathbf{r}}) + \frac{\partial}{\partial y}(E_{0,y}e^{i\mathbf{k}\cdot\mathbf{r}}) + \frac{\partial}{\partial z}(E_{0,z}e^{i\mathbf{k}\cdot\mathbf{r}})\right) \\ &= e^{-i\omega t}\left(E_{0,x}\frac{\partial}{\partial x}e^{i\mathbf{k}\cdot\mathbf{r}} + E_{0,y}\frac{\partial}{\partial y}e^{i\mathbf{k}\cdot\mathbf{r}} + E_{0,z}\frac{\partial}{\partial z}e^{i\mathbf{k}\cdot\mathbf{r}}\right) \\ &= e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\left(iE_{0,x}k_x + iE_{0,y}k_y + iE_{0,z}k_z\right) \\ &= i\mathbf{k}\cdot\mathbf{E} = 0\end{aligned}$$

since $\mathbf{k}$ is perpendicular to $\mathbf{E}_0$. This is another example of a lesson we saw in **M4**. Namely, when everything is a complex exponential, differentiation is very easy. For an complex exponential in time, $e^{i\omega t}$, differentiation with respect to time is just multiplication by $i\omega$. Similarly, for a field which is a complex exponential in space, $e^{i\mathbf{k}\cdot\mathbf{r}}$, the divergence $(\nabla\cdot)$ becomes $(i\mathbf{k}\cdot)$.

By similar reasoning, Gauss's law for magnetism is satisfied. Next, we check Ampere's law,

$$\nabla \times \mathbf{B} = \mu_0\epsilon_0\frac{\partial \mathbf{E}}{\partial t}.$$

By the same logic as above, the curl becomes $(i\mathbf{k}\times)$, while the time derivative becomes multiplication by $-i\omega$, giving

$$i\mathbf{k}\times\mathbf{B} = (-i\omega)\mu_0\epsilon_0\mathbf{E}.$$

Because $\mathbf{k}$, $\mathbf{E}$, and $\mathbf{B}$ are all mutually perpendicular, the directions of both sides match. Then all that remains is to check the magnitudes,

$$kB_0 = \omega\mu_0\epsilon_0E_0.$$

By plugging in results from above, this reduces to

$$c^2 = \frac{1}{\mu_0\epsilon_0}$$

which matches what we said above. (Or, if we didn't know what c case, this logic would have been a way to derive it, as Maxwell did.) The verification of Faraday's law is similar. Note that the displacement current term was essential; it wouldn't have been possible to get electromagnetic wave solutions without it.

[3] **Problem 18.** Consider the energy and momentum of the electromagnetic wave in idea 8.

(a) Show that the spatial average of the energy density is $\epsilon_0E_0^2/2$. (Be careful with factors of 2.)

(b) Compute the spatial average of the momentum density $\langle \mathbf{p} \rangle$ using idea 7.

(c) Confirm that $E = pc$ for an electromagnetic wave.

[3] **Problem 19.** The intensity of sunlight at noon is approximately $1\,\mathrm{kW/m^2}$.

(a) Compute the rms magnetic field strength.

(b) Compute the radiation pressure acting on a mirror lying on the ground.

(c) In terms of the Lorentz force, how is this pressure exerted on the particles in the mirror?

[3] **Problem 20** (Purcell 9.7)**.** Consider the sum of two oppositely-traveling electromagnetic waves, with electric fields

$$\mathbf{E}_1 = E_0 \cos(kz - \omega t)\hat{\mathbf{x}}, \quad \mathbf{E}_2 = E_0 \cos(kz + \omega t)\hat{\mathbf{x}}.$$

(a) Write down the magnetic field.

(b) Draw plots of the energy density $U(z, t)$ for $\omega t \in \{0, \pi/4, \pi/2, 3\pi/4, \pi\}$.

(c) On top of these plots, draw the direction and magnitude of the Poynting vector. Convince yourself that the Poynting vector accurately describes how the energy sloshes back and forth.

Idea 9: Larmor Formula

An accelerating charge produces electromagnetic radiation, with power

$$P = \frac{q^2 a^2}{6\pi\epsilon_0 c^3}.$$

We'll derive it properly in **R3**, but a lot of it can be motivated with the techniques of **P1**.

The power could only depend on q, ϵ_0, μ_0, and properties of the particle's motion. The only combinations of the first three parameters that get rid of the electromagnetic units are q^2/ϵ_0 and $1/\sqrt{\epsilon_0\mu_0} = c$. Since energy is proportional to the electric and magnetic fields squared, and these fields are proportional to q, the answer must be proportional to q^2/ϵ_0.

Radiation can't result from uniform velocity, by Lorentz invariance; another way to see this is that with only v and c, there is no way to write down an expression for power with the right units! The next simplest option is radiation from acceleration, from which the most general result is $P = (q^2 a^2/\epsilon_0 c^3) f(v/c)$. The fact that acceleration is squared is also natural, because acceleration is a vector, so this is the simplest way to get a rotationally invariant result. The proper derivation shows that $f(0) = 1/6\pi$. When v/c is substantial, there are relativistic corrections, which we will consider in **R3**.

[2] **Problem 21** (Purcell H.2)**.** A common classical model of an electron in an atom is to imagine it is a mass on a spring, where the spring force is due to the atomic nucleus. Suppose that such an electron, with charge e, is vibrating in simple harmonic motion with angular frequency ω and amplitude A.

(a) Find the average rate of energy loss by radiation.

(b) If no energy is supplied to make up the loss, how long will it take the oscillator's energy to fall to $1/e$ of its initial value?

Numerically, this is an extremely small time, so classical models of the atom are not realistic. We will see in **X1** that in quantum mechanics this problem is solved because in the ground state the electron does not move around the atom, but rather occupies a standing wave.

[3] **Problem 22** (Purcell H.3)**.** A plane electromagnetic wave with angular frequency ω and electric field amplitude E_0 is incident on an atom. As in problem 21, we model the electron as a simple harmonic oscillator, with mass m and natural angular frequency ω_0.

(a) First suppose that $\omega \gg \omega_0$. Argue that in this case, the "spring" force on the electron can be neglected. Find the average power radiated by the electron, and show that it is equal to the power incident on a disc of area

$$\sigma = \frac{8\pi}{3}\left(\frac{e^2}{4\pi\epsilon_0 mc^2}\right)^2.$$

This is the Thomson scattering cross section. To an electromagnetic wave, each electron looks like it has this area.

(b) Now suppose $\omega \ll \omega_0$, yielding Rayleigh scattering, which describes the scattering of visible light by air. In this case, show that $\sigma \propto \omega^4$. This sharp frequency dependence explains why the sky is blue. (Why isn't it violet?)

(c) Explain the meaning of the common phrase "red sky at night, sailor's delight; red sky in morning, sailor's warning". (Hint: in the cultures where this saying is used, weather patterns usually move from west to east.)

For some further discussion of Rayleigh scattering, see section 9.4 of The Art of Insight. For more about colors in the atmosphere, see this nice video.

[3] **Problem 23.** USAPhO 2016, problem B2.

Remark

We noted in **M7** that clouds are visible because the radiation scattered by a small droplet of n water molecules grows as n^2. To understand why, note that each of the molecules performs independent Rayleigh scattering, as computed above. For separated molecules, the energy scattered just adds. However, for nearby molecules the electromagnetic waves scattered interfere constructively, so the amplitude grows as n and hence the energy scattered as n^2.

This quadratic enhancement breaks down once the droplets exceed the wavelength λ of the light. This means the maximum possible enhancement is larger for larger wavelengths, acting against the ω^4 dependence of Rayleigh scattering. This is why clouds are white, not blue.

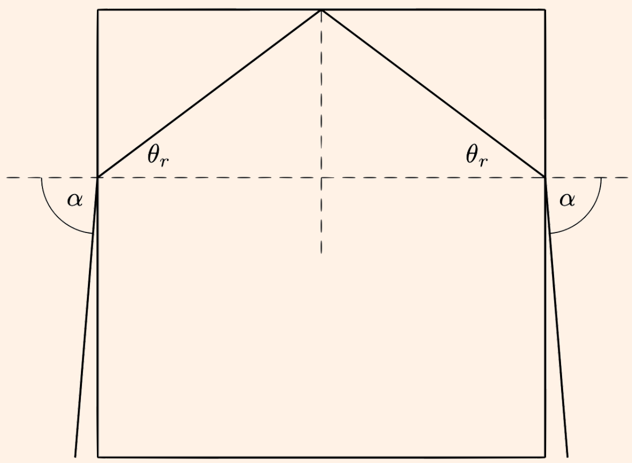
Radiation pressure can also have mechanical effects.

Example 5: NBPhO 2018.6

A laser pointer of power P is directed at a glass cube, with refractive index $n > \sqrt{2}$. The surface of the cube has an anti-reflective coating, so there is no partial reflection when light enters or exits it; the laser pointer only refracts. What is the maximum force the laser pointer can exert on the cube?

Solution

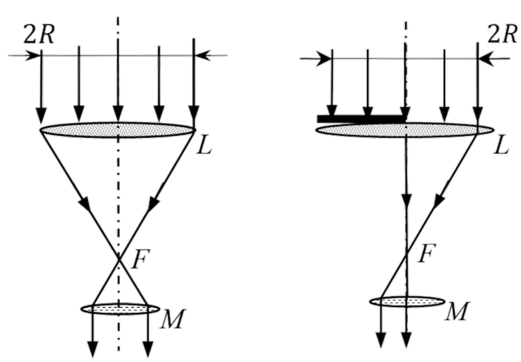
The force is due to a change in momentum of the light. The greatest possible force is attained if the direction of the light is reversed, which can occur as shown, in the limit $\alpha \to 90°$.



Assuming $n > \sqrt{2}$, we then have $\theta_r < 45°$, and then the laser internally reflects when it hits the top surface of the cube. It exits in the opposite direction it came in.

If the laser pointer has power P, then the momentum of the laser beam per time is P/c. The momentum is reversed, so the force is $2P/c$.

[3] **Problem 24** (IZhO 2022)**.** In 2018, the Nobel Prize in physics was awarded to Arthur Ashkin for the creation of the "laser tweezer", a device that allows one to hold and move transparent microscopic objects with the help of light. In one such device, a parallel beam of light from a laser passes through a converging lens L and hits a microparticle M, which can also be considered a converging lens. Point F is the common focus of L and M.



The light intensity in the beam is $I = 1.00\,\mu\text{W/cm}^2$, the beam radius is $R = 1.00\,\text{cm}$, and the focal length of the lens L is $F = 10.0\,\text{cm}$. Ignore the absorption and reflection of light.

(a) Calculate the force acting on the microparticle, in the setup shown at left above.

(b) Next, the left half of the lens L is covered by a diaphragm, as shown at right above. Calculate the force acting on the microparticle in the transverse direction of the beam.

[3] **Problem 25** (Feynman). In one proposed means of space propulsion, a spaceship of mass $10^3\,\text{kg}$ carries a thin sheet of area $100\,\text{m}^2$. The sheet is made of highly reflective plastic film, and can be used as a solar radiation pressure "sail". The spaceship travels in a circular orbit of radius r, which is initially equal to the Earth's orbit radius, where the intensity of sunlight is $1400\,\text{W/m}^2$. Assume the spaceship is moving nonrelativistically and the gravitational effect of the Earth is negligible.

(a) Find the angle at which the sail should be pointed to maximize dr/dt.

(b) Assuming the sail is pointed this way, find the numeric value of dr/dt.

(c) If this continues for a very long time, then r will grow as $r \propto t^n$. Find the value of n.

Finally, we'll consider electromagnetic wave propagation in transmission lines.

[4] **Problem 26** (Griffiths 7.62, Crawford 4.8). A certain transmission line is constructed from two thin metal ribbons, of width w, a very small distance $h \ll w$ apart. The current travels down one strip and back along the other. In each case it spreads out uniformly over the surface of the ribbon.

(a) Find the capacitance per unit length $\mathcal{C}$, and the inductance per unit length $\mathcal{L}$.

(b) Argue that the speed of propagation of electromagnetic waves through this transmission line is of order $1/\sqrt{\mathcal{LC}}$, and evaluate this quantity.

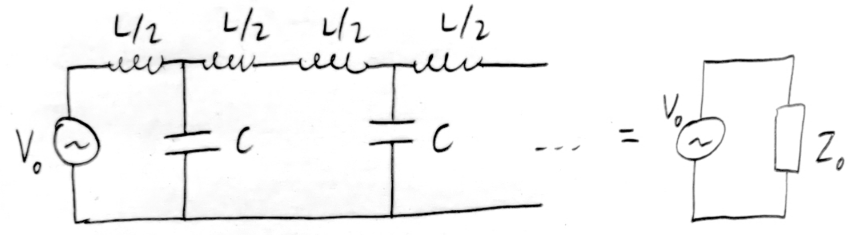
(c) Repeat the first two parts for a coaxial transmission line, consisting of two cylinders of radii $a < b$ with the same axis of symmetry.

(d) Repeat the first two parts for a parallel-wire transmission line, consisting of two wires of radius r whose axes are a distance $D \gg r$ apart.

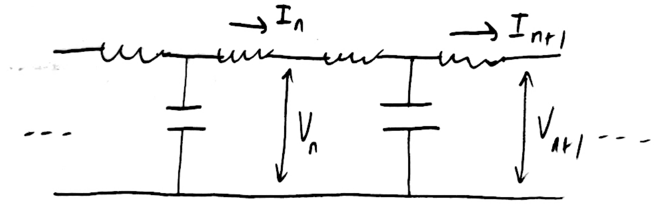
You should find that in all cases, $1/\sqrt{\mathcal{LC}}$ is the same, yielding the same speed for electromagnetic waves. This isn't a coincidence, and applies for transmission lines with conductors of any shape, though the general proof requires some elaborate vector calculus, as you can see here.

[4] **Problem 27.** In this problem, we treat electromagnetic wave propagation through a transmission line using a "lumped element" approach, where the line is replaced with discrete capacitors and inductors, as shown. (This is an example of a network synthesis, mentioned in **E6**.)

(a) Calculate the characteristic impedance $Z_0(\omega)$ of the entire network, as shown below.



(b) The diagram below shows two adjacent sections of the ladder.



Find the ratio of the complex voltage amplitudes V_{n+1}/V_n.

(c) The AC driving attempts to create electromagnetic waves which travel through the network, to the right. It turns out that above a certain critical angular frequency ω_c, waves will not travel through the ladder network. Find ω_c. (Hint: this can be done using either the result of part (a) or part (b).)

(d) For angular frequencies $\omega \ll \omega_c$, waves travel through the ladder with a constant speed. Find this speed, assuming each segment of the ladder has physical length ℓ. (Hint: the speed of a wave obeys $v = d\omega/dk$.)

(e) You should have found in one of the earlier parts that the impedance of this infinite network can be a real number, even though it's made of parts which all have imaginary impedance. That sounds strange, but what's even stranger is that we *should* be able to handle this infinite circuit by taking the limit of progressively larger finite circuits, just as we did for a similar network of resistors in **E2**. But for *any* finite LC network, the impedance will be imaginary, so the limit must be imaginary too! On one hand, we should trust the finite result because all real circuits are finite. On the other hand, the real impedance we get for the infinite result certainly can be measured in real life. So what's going on?

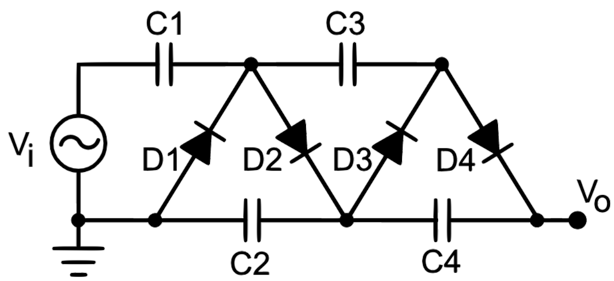
Electromagnetism VII: Electrodynamics

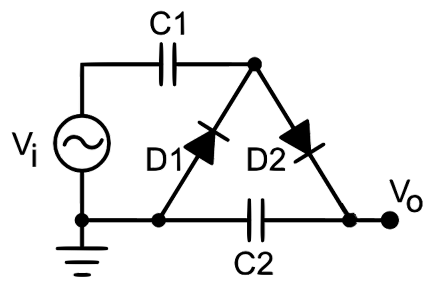
Chapter 9 of Purcell covers electromagnetic waves, and appendix H covers radiation by charges. For a pedagogical introduction with solved examples, see recitation 8 and recitation 9 of the MIT OCW 8.03 lectures. For more technical coverage, not necessarily relevant to the Olympiad, see chapters 7, 8, 10, 11 of Griffiths. For some lighter reading, see chapters I-28, I-32, II-18, II-20, II-21, and II-24 of the Feynman lectures. There is a total of **85** points.

1 More Nonlinear Circuit Elements

In this section we consider some more subtle applications of nonlinear circuit elements. First, we consider complex problems that use relatively familiar circuit elements.

[3] **Problem 1.** The setup below at left is called the Cockcroft–Walton voltage multiplier. It was used in 1932 to power the first particle accelerator.





The four capacitors begin uncharged, and all have capacitance C. All four diodes are ideal. The voltage $V_i(t)$ alternates between V and $-V$. The output voltage is V_0.

(a) To warm up, consider the simpler setup shown at right above. Suppose the applied voltage V_i begins at $-V$. Describe how V_0 changes each time the applied voltage switches sign.

(b) Now consider the full setup, shown at left above. After a long time, what is V_0?

Solution. Since everything in this problem is ideal, the diodes let an arbitrary amount of current through whenever the potential difference across them is positive. In other words, they make sure the potential difference across them is zero or negative.

(a) When the applied voltage becomes $-V$, the potential at the right plate of C1 is lowered. This causes current to flow through D1, until the potential of that plate is nonnegative. When the applied voltage switches from $-V$ to V, the potential on the right plate of C1 is raised. This causes current to flow through D2, until the potential on the right plate of C2 is at least as high as that on the right plate of C1. (We don't have to worry much about the left plates; these are connected to ground, so they just automatically pick up opposite charges to the right plates.) Thus, the potentials on these right plates evolve as follows.

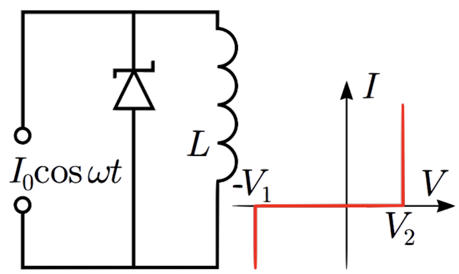
1. $V_i = -V$. This switches the potentials to $(-V, 0)$, which means current flows through D1 until they are $(0, 0)$.
2. $V_i = V$. This switches the potentials to $(2V, 0)$. Current flows through D2 until they are (V, V).

3. $V_i = -V$. This switches the potentials to $(-V, V)$, so current flows through D1 until they are $(0, V)$.
4. $V_i = V$. This switches the potentials to $(2V, V)$, so current flows through D2 until they are $(3V/2, 3V/2)$.

The process continues similarly for more steps. The output voltage begins at zero, then becomes V for two steps, then $3V/2$, then $7V/4$, then $15V/8$, and so on, asymptotically approaching $2V$. So this setup is a voltage doubler.

(b) The actual sequence of events is complicated, but the full setup is essentially two copies of the reduced setup. After a long time, the output voltage is $4V$. For a neat animation of how this works, using the fluid flow analogy for circuits, see this video.

[3] **Problem 2** (NBPhO 2017). A Zener diode is connected to a source of alternating current as shown.



The inductance L of the inductor is such that $L\omega I_0 \gg V_1, V_2$ where V_1 and V_2 are the breakdown voltages, and $V_1 > V_2$. The $I(V)$ characteristic of the Zener diode is shown above. Assume that a long time has passed since the current source was first turned on.

(a) Find the average current through the inductor.

(b) Find the peak-to-peak amplitude of the current changes ΔI in the inductor.

Solution. See the official solutions here.

[4] **Problem 3.** NBPhO 2016, problem 4. A rather complicated problem involving several exotic circuit elements.

Solution. See the official solutions here.

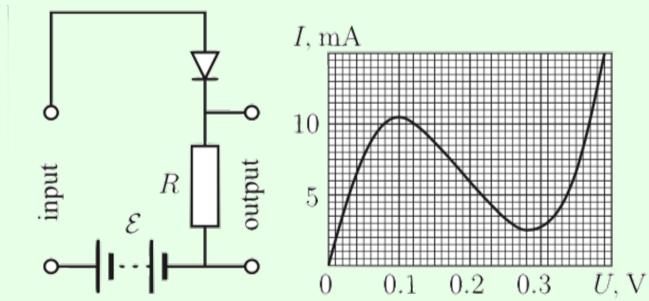
Next, we consider some qualitatively new behavior that can emerge from less familiar circuit elements, such as amplification, hysteresis, and instability.

Idea 1

Tunnel diodes are a variant of diodes, whose $I(V)$ rises, falls, and rises again. That is, they have a region with negative differential resistance, $dI/dV < 0$. This allows them to amplify signals, as we'll see below, and also can make them unstable.

Example 1: EFPhO 2003

The circuit below, containing a tunnel diode, acts as a simple amplifier.



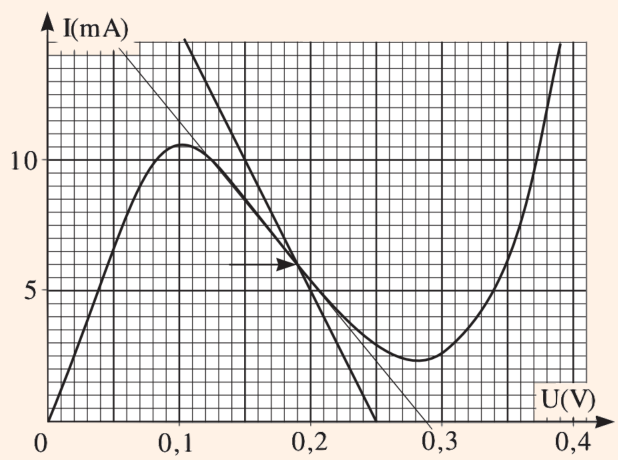
Here, $R = 10\,\Omega$ and $\mathcal{E} = 0.25\,\text{V}$. If a small signal voltage $V_{\text{in}}(t)$ is applied across the input, then an amplified and shifted version of the signal appears across the output. Find the amplification factor.

Solution

When a constant emf $\mathcal{E}$ is applied, Kirchoff's laws give

$$\mathcal{E} = I_0 R + V(I_0)$$

where $V(I)$ is the voltage characteristic of the diode.



By plotting $\mathcal{E} - IR$ on the graph above, we find I_0 at the intersection. (Notice the x-axis label: it is common to write a decimal point as a comma in Eastern Europe.) Now consider the effect of applying the signal voltage, which changes the current by ΔI,

$$\mathcal{E} + V_{\text{in}} = (I_0 + \Delta I)R + V(I_0 + \Delta I).$$

Since the signal voltage is small, we can Taylor expand the voltage characteristic, giving

$$V_{\text{in}} = \Delta I(R + V'(I_0)).$$

This in turn tells us that

$$V_{\text{out}} = (I_0 + \Delta I)R = V_{\text{out}}^0 + \frac{R}{R + V'(I_0)} V_{\text{in}}.$$

In other words, the change in V_{out} is just V_{in}, times the amplification factor

$$\frac{R}{R+V'(I_0)} = \frac{10}{10-16} = -\frac{5}{3}$$

where we read $V'(I_0)$ off the graph by drawing a tangent. The intuition here is that the circuit is like a voltage divider, but the tunnel diode acts like a negative resistance. If we had $V'(I_0)$ close to $-R$, for example, the amplification factor would have been huge. Since $V'(I_0)$ is more negative than $-R$, the signal ends up flipped.

[4] **Problem 4.** NBPhO 2020, problem 2. A comprehensive problem on the measurement and dynamics of tunnel diodes, which will give you a deeper understanding of negative resistance.

Solution. See the official solutions here. For more about stabilizing circuits with negative differential resistance, see problem 82 of Kalda's circuits handout.

Idea 2

Op amps have four terminals, and output a voltage across the last two equal to the voltage across the first two, times a very large gain. Like tunnel diodes, op amps can be unstable: increasing the input increases the output, but this in turn could increase the input again. Thus, in practice, the output and input are always wired together in a way that produces negative feedback, with changes in the output acting to decrease the input. In this case, one can think of an op amp as a tool that tries to set the input voltages equal to each other.

The internals are somewhat complicated, consisting of a lot of resistors and transistors. In general, to analyze setups with multiple complex circuit elements like these, it's better to treat them as black boxes than to try to intuit what's going on at the level of individual subelements, or electric and magnetic fields. (Of course, engineers do need to understand circuit elements at these level to design them in the first place!)

[3] **Problem 5.** USAPhO 2016, problem A2. This problem is a nice introduction to op amps.

Idea 3

In some nonlinear circuit elements, the function $I(V)$ is multivalued. This indicates hysteresis: given V, the actual value of I depends on the history of the system. The same goes for when $V(I)$ is multivalued.

[5] **Problem 6.** IPhO 2016, problem 2. This problem illustrates the previous idea with a thyristor. Print out the official answer sheet and record your answers on it.

2 Displacement Current

Idea 4

In general, Ampere's law is

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

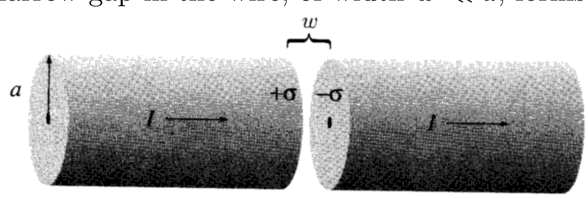
This is sometimes written in terms of a "displacement current" density $\mathbf{J}_d$, where

$$\nabla \times \mathbf{B} = \mu_0(\mathbf{J} + \mathbf{J}_d), \quad \mathbf{J}_d = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

In integral form, for a surface S bounded by a closed curve C,

$$\oint_C \mathbf{B} \cdot d\mathbf{s} = \mu_0 \int_S (\mathbf{J} + \mathbf{J}_d) \cdot d\mathbf{S} = \mu_0 I + \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}.$$

[2] **Problem 7** (Griffiths 7.34)**.** A fat wire of radius a carries a constant current I uniformly distributed over its cross section. A narrow gap in the wire, of width $w \ll a$, forms a parallel plate capacitor.

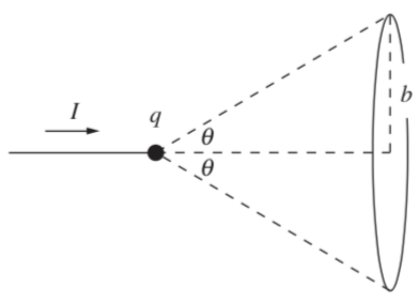


Find the magnetic field in the gap, at a distance $s < a$ from the axis.

Solution. The field in that region is $E = \sigma/\epsilon_0$, so the displacement current density is $d\sigma/dt = I/(\pi R^2)$, so we might as well assume that there is no gap. We see then that by Ampere's law that

$$B \cdot 2\pi s = \mu_0 (s^2/a^2) I, \quad B(s) = \frac{\mu_0 I}{2\pi} \frac{s}{a^2}.$$

[3] **Problem 8** (Purcell 9.3)**.** A half-infinite wire carries constant current I from negative infinity to the origin, where it builds up at a point charge with increasing q. Consider the circle shown below.



Calculate the integral $\int \mathbf{B} \cdot d\mathbf{s}$ about this circle in three ways.

(a) Use the integrated form of Ampere's law, integrating over a surface which does not intersect the wire.

(b) Do the same, with a surface that does intersect the wire.

(c) Apply the Biot–Savart law to the current and displacement current.

Solution. Define coordinates such that the z axis is anti-parallel to the wire, and use spherical coordinates (r, θ, ϕ) with respect to this choice of the z axis (we won't need θ, so the choice of x and y axes is unimportant).

(a) Consider the surface $r = R$ and $0 \leq \theta \leq \theta_0$ where we redefine the θ in the problem to θ_0. Note that $\mathbf{E} = \frac{q}{4\pi\epsilon_0 R^2}\hat{\mathbf{r}}$, so the displacement current at this surface is $\mathbf{J}_d = \frac{I}{4\pi R^2}\hat{\mathbf{r}}$. We have

$$\int \mathbf{B} \cdot d\mathbf{s} = \int \mu_0 \mathbf{J}_d \cdot d\mathbf{S}$$

where the second integral is over the surface we just defined. Note that the surface area of this surface is $2\pi R^2(1 - \cos\theta_0)$, so we have

$$\int \mathbf{B} \cdot d\mathbf{s} = \mu_0 I \, \frac{1 - \cos\theta_0}{2}.$$

(b) Instead use the surface $\theta_0 \leq \theta \leq \pi$, and account for the current that pierces the surface. In the end, we get

$$\mu_0 I - \mu_0 I \, \frac{1 + \cos\theta_0}{2} = \mu_0 I \, \frac{1 - \cos\theta_0}{2},$$

which is what we got before.

(c) Since the displacement current is spherically symmetric, it doesn't produce any magnetic field at all, as we saw in an example in **E3**. So we consider just the wire.

Note that all the contributions from the infinitesimal pieces point in the $\hat{\boldsymbol{\phi}}$ direction, so we see that the field is

$$B = \frac{\mu_0 I}{4\pi} \int_{-\infty}^{-b\cot\theta_0} \frac{b\,dz}{(b^2 + z^2)^{3/2}} = \frac{\mu_0 I}{4\pi b} \int_{-\infty}^{-\cot\theta_0} \frac{dk}{(1 + k^2)^{3/2}} = \frac{\mu_0 I}{4\pi b}\,(1 - \cos\theta_0).$$

Thus, the integral is

$$\int \mathbf{B} \cdot d\mathbf{s} = \frac{\mu_0 I}{2}\,(1 - \cos\theta_0)$$

just as found in the previous parts.

In the previous problem, you should have found that the effect of the displacement current, in the Biot–Savart law, simply cancelled out everywhere. In fact, this cancellation is very general.

Idea 5

In any situation where $\mathbf{J}$ is constant, whether or not ρ is constant, Maxwell's equations are satisfied by applying Coulomb's law to ρ and the Biot–Savart law to $\mathbf{J}$. You can include displacement currents in the Biot–Savart integral too, but their contributions perfectly cancel.

To see why, note that

$$\nabla \times \mathbf{J}_d = \epsilon_0 \nabla \times \frac{\partial \mathbf{E}}{\partial t} = \epsilon_0 \frac{\partial}{\partial t}(\nabla \times \mathbf{E}) = -\epsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2}$$

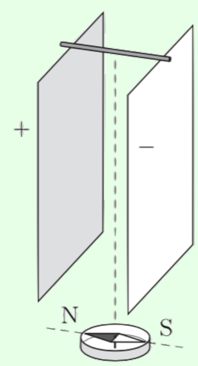
where we used Faraday's law. When the currents are constant, the magnetic fields are also constant, so the right-hand side vanishes. Then $\nabla \times \mathbf{J}_d = 0$. However, this means that $\mathbf{J}_d$ can always be written as a superposition of radial, spherically symmetric currents, and as we saw in the previous problem, such currents produce no magnetic fields.

This explains why we were able to get away with using Coulomb's law and the Biot–Savart law on previous problem sets, even in situations which were not exactly electrostatic or magnetostatic – all of these situations were "quasistatic". In general, the displacement current only matters in non-quasistatic situations involving rapid changes in $\mathbf{J}$, and hence rapid changes in $\mathbf{E}$ and $\mathbf{B}$. These are exactly the cases where significant electromagnetic radiation is produced, which is why radiation is covered in the last half of this problem set.

For more about this subtle point, see section 9.2 of Purcell, this paper and this paper.

Example 2: MPPP 190

A parallel plate capacitor is charged and positioned above a compass as shown.



The capacitor is discharged slowly when the tops of the plates are joined using a small conducting rod. Which way is the compass needle deflected during the discharge process?

Solution

Since the discharge is slow, the situation is quasistatic. (It would only be non-quasistatic in the case where the discharge time was comparable to the time it would take for light to cross the capacitor, a situation which is almost never achieved for RC circuits.) Then we know the magnetic field due to the displacement current cancels out everywhere, so only the current I in the rod matters. This current moves left to right, so by a straightforward application of the right-hand rule, we find that the compass is deflected east.

Things get more subtle if one insists on considering the displacement current anyway. A naive, incorrect argument would be to say that there is a total displacement current I going right to left inside the capacitor, and since this displacement current is closer than the current in the rod, it produces a larger magnetic field, deflecting the compass west. The problem with this reasoning is that it has ignored the displacement current due to the changes in the substantial fringe fields of the capacitor. When these are accounted for, the magnetic fields due to the displacement current cancel, as argued generally above.

[3] **Problem 9** (Griffiths 7.36). An alternating current $I = I_0 \cos(\omega t)$ flows down a long straight wire

along the $\hat{\mathbf{z}}$ axis and returns along a coaxial conducting tube of radius a.

(a) Assuming the electric field goes to zero at infinity, show that

$$\mathbf{E} = \frac{\mu_0 I_0 \omega}{2\pi} \sin(\omega t) \log \frac{a}{r} \, \hat{\mathbf{z}}.$$

(b) Find the amplitude I_d of the total displacement current.

(c) Compute the ratio I_d/I_0. Show that I_d is only significant when ω is fast enough that the speed-of-light travel time from the wire to the tube is comparable to the period.

This shows another sense in which displacement current effects are only significant when radiation comes into play. If I_d/I were near one, we would have to then consider the magnetic fields induced by the changing electric fields associated with the displacement current, and then the displacement currents due to the changes in those magnetic fields, and so on. But this description of electric fields inducing magnetic fields and vice versa is just a description of an electromagnetic wave.

Solution. (a) Use an Amperian loop that is a rectangle extending from $s = r$ to some value of s bigger than a, and with height z. We see that

$$E_z = -\frac{d\Phi_B}{dt} = -\frac{d}{dt} \int_s^a \frac{\mu_0 I}{2\pi s'} \, z \, ds' = \frac{\mu_0 I_0 \omega}{2\pi} \sin(\omega t) \log \frac{a}{r},$$

as desired.

(b) Note that $\mathbf{J}_d = \epsilon_0 \frac{\mu_0 I_0 \omega^2}{2\pi} \cos(\omega t) \log(a/s) \, \hat{\mathbf{z}}$. Thus, the amplitude is

$$I_d = \frac{\omega^2 I}{2\pi c^2} \int_0^a \log(a/s) 2\pi s \, ds = \frac{\omega^2 I a^2}{4c^2}.$$

(c) The ratio goes as $\omega^2 a^2/c^2$, so I_d is significant once $1/\omega \sim a/c$. Up to constants, these are the period and the speed of light travel time, as desired.

By the way, you might be wondering why this problem uses a cylindrical geometry, while the geometries of a spherical or infinite parallel-plane capacitor are more symmetric. Things go wrong in those cases because you need some path for the current to get from one plate to the other. For the spherical case, you can only maintain the symmetry if the current flows radially in a spherically symmetric manner. But in such cases, the magnetic field is always just zero by symmetry and Gauss's law, as we argued in **E3**, which ruins the point of the problem. In the parallel plate case, you could imagine the current goes from one plate to another "at infinity", but, as we saw in a similar problem in **E1**, the precise way you define what's happening at infinity will affect the answer! The general lesson is that cylindrical geometries are often nice for demonstrating theoretical points, because they are infinite in one direction (so you can have the current return there) but not others (so you can still unambiguously define the potential to be "zero far away"). In fact, as you get further in physics, nice setups become increasingly rare, and it will often be the case that only *one* setup works for demonstrating a point without technical complications.

[3] **Problem 10.** [A] Consider an infinite thin solenoid which initially carries no current, and a loop of wire around this solenoid of enormous radius, say one light year. At some moment, a current is suddenly made to flow through the solenoid. (This cannot be done by simply attaching a battery

somewhere, because it will take a long time for the current to turn on throughout the solenoid. So instead, consider a situation where many batteries arranged around the solenoid are all attached in at once, which can be achieved by machines which have synchronized their clocks beforehand.)

A magnetic field is hence quickly produced in the solenoid, so by Faraday's law,

$$\mathcal{E} = -\frac{d\Phi_B}{dt}$$

there should quickly be an emf in the loop of wire. But this seems to violate locality, because the motion of charges in the solenoid is quickly affecting the motion of charges in the wire loop, which is very far away. What's going on? Could there be something wrong with Faraday's law?

Solution. Faraday's law is fine. It is completely compatible with relativity – after all, the fact that Maxwell's equations obey the postulates of relativity is how we discovered it in the first place!

The real resolution is that *both* $\mathcal{E}$ and Φ_B remain zero for about a year. At the moment the current is switched on, the changing current produces a pulse of outward-moving, radially symmetric electromagnetic radiation. This radiation has a downward-pointing magnetic field and tangential-pointing electric field. The downward-pointing magnetic field exactly cancels the flux from the upward-pointing solenoid field, so that Φ_B is exactly zero until the pulse passes by the wire. (In terms of field lines, each magnetic field line going up through the solenoid returns downward. The downwardly returning field lines move outward at speed c.) As the pulse passes by the wire, Φ_B starts to change. Accordingly, an emf appears at the very same moment, due to the tangential electric field in the pulse. So Faraday's law is satisfied the whole time, but in quite a subtle way!

[3] **Problem 11** (Griffiths 7.64). [A] Setting $\mu_0 = \epsilon_0 = 1$, Maxwell's equations read

$$\nabla \cdot \mathbf{E} = \rho_e, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{B} = \mathbf{J}_e + \frac{\partial \mathbf{E}}{\partial t}$$

where ρ_e and $\mathbf{J}_e$ are the electric charge density and electric current density.

(a) Show that Maxwell's equations ensure the conservation of electric charge,

$$\dot{\rho}_e = -\nabla \cdot \mathbf{J}_e.$$

This is the continuity equation, and we saw versions of it for other conserved quantities in **T2**.

(b) Generalize Maxwell's equations to include a magnetic charge density ρ_m and a magnetic current density $\mathbf{J}_m$. Fix the signs by demanding that magnetic charge is conserved.

(c) Check that the resulting equations are invariant under the duality transformation

$$\begin{pmatrix} \mathbf{E}' \\ \mathbf{B}' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{B} \end{pmatrix}, \quad \begin{pmatrix} \rho'_e \\ \rho'_m \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \rho_e \\ \rho_m \end{pmatrix}$$

which rotates electricity into magnetism with angle θ.

(d) Write down the Lorentz force law for a particle with electric and magnetic charge, using the fact that it should be invariant under the duality transformation above.

Solution. (a) Taking the time derivative of Gauss's law gives

$$\dot{\rho}_e = \nabla \cdot \dot{\mathbf{E}}.$$

By taking the divergence of Faraday's law, we have

$$\nabla \cdot (\nabla \times \mathbf{B}) = \nabla \cdot \mathbf{J}_e + \nabla \cdot \dot{\mathbf{E}}.$$

The left-hand side vanishes, because the divergence of a curl of *any* vector field is always zero. We thus have

$$0 = \nabla \cdot \mathbf{J}_e + \dot{\rho}_e$$

as desired.

(b) Almost by definition, we will want to have $\nabla \cdot \mathbf{B} = \rho_m$. By symmetry we want something like

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \pm \mathbf{J}_m.$$

By enforcing that the divergence of the right side is 0, we learn that it is

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} - \mathbf{J}_m.$$

(c) The equations can be written succinctly as

$$\nabla \cdot \begin{pmatrix} \mathbf{E} \\ \mathbf{B} \end{pmatrix} = \begin{pmatrix} \rho_e \\ \rho_m \end{pmatrix}.$$

Applying the rotation matrix to both sides shows that the Gauss's laws are satisfied in the primed setup. Similarly, the other two can be written as

$$\nabla \times \begin{pmatrix} \mathbf{E} \\ \mathbf{B} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \partial_t \begin{pmatrix} \mathbf{E} \\ \mathbf{B} \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \partial_t \begin{pmatrix} \mathbf{J}_e \\ \mathbf{J}_m \end{pmatrix}.$$

Again, the result is clear by applying the rotation transformation.

(d) We see that $\mathbf{F} = q_e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) + q_m(\mathbf{B} - \mathbf{v} \times \mathbf{E})$. We see that

$$\begin{aligned} \mathbf{F}' = &(q_e \cos\theta + q_m \sin\theta)\left[(\mathbf{E}\cos\theta + \mathbf{B}\sin\theta) + \mathbf{v} \times (\mathbf{B}\cos\theta - \mathbf{E}\sin\theta)\right] \\ &+ (q_m \cos\theta - q_e \sin\theta)\left[(\mathbf{B}\cos\theta - \mathbf{E}\sin\theta) - \mathbf{v} \times (\mathbf{E}\cos\theta + \mathbf{B}\sin\theta)\right] \\ = &\, q_e \mathbf{E}(\cos^2\theta + \sin^2\theta) + \mathbf{v} \times \mathbf{B}(\cos^2\theta + \sin^2\theta) \\ &+ q_m \mathbf{B}(\sin^2\theta + \cos^2\theta) + \mathbf{v} \times (-\mathbf{E})(\sin^2\theta + \cos^2\theta) = \mathbf{F} \end{aligned}$$

as desired.

Remark

The peculiar name of $\mathbf{J}_d$ is because Maxwell thought of it as a literal displacement of a jelly-like ether. In that era, all electromagnetic quantities, such as fields, charges, currents, polarizations, and magnetizations, were thought to reflect properties of a mechanical ether, such as local strains, displacements, and rotations. However, making this picture precise was known to be difficult even before the advent of relativity, which rendered ether obsolete. The best way to understand why physicists abandoned ether models is to have a look at their daunting complexity. For a nice overview, with diagrams, see chapter 4 of *The Maxwellians*.

3 Field Energy and Momentum

Idea 6

The Poynting vector

$$\mathbf{S} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0}$$

gives the flux density of the energy of an electromagnetic field. That is, the flux of $\mathbf{S}$ into a closed surface is the rate of change of total field energy within that surface.

[3] **Problem 12.** USAPhO 2010, problem B2.

[3] **Problem 13.** USAPhO 2013, problem B2.

Remark

It's unlikely that you'll see any examples besides the ones in the above two problems, because in almost all other setups, the Poynting vector depends sensitively on the fringe fields, which are very hard to calculate. (For some work in this direction, see *Energy transfer in electrical circuits: A qualitative account*.) In any case, the examples above illustrate the important point that the energy of a circuit does not flow along the wire, carried by the charges; instead it flows into circuit elements from the sides. This was an important early clue of the importance of the electromagnetic field.

Remark

As proven in Poynting's theorem, the Poynting vector indeed tells us about the net flow of energy. However, this would remain true if we added a constant vector to it, or more generally any divergence-free vector field, since these wouldn't change the net flow. So which option is the "correct" one? According to everything we've learned so far, there's no absolute way to choose, and we just use the Poynting vector because it's the simplest option. However, in general relativity, the flow of energy directly influences the curvature of spacetime, so there is an unambiguous correct answer, which is indeed the Poynting vector.

Example 3

Consider two charges q, at positions $r\hat{\mathbf{x}}$ and $r\hat{\mathbf{y}}$ respectively, both moving with speed v towards the origin. Show that the magnetic forces between them are *not* equal and opposite. That is, electromagnetic forces do not obey Newton's third law.

Solution

In order to find the $\mathbf{B}$ field produced by each charge at the location of the other, we use the Biot–Savart law and the right-hand rule. Then we use the Lorentz force and the right-hand rule again to find the magnetic forces on each charge.

For example, the $\mathbf{B}$ field produced by the first charge at the location of the second is along $-\hat{\mathbf{z}}$. Then the magnetic force on the second charge is parallel to $\hat{\mathbf{x}}$. The magnetic force on the first charge is parallel to $\hat{\mathbf{y}}$. And the force are definitely nonzero, so they can't be equal and opposite.

To explain this, we recall that the point of Newton's third law is just momentum conservation. This still holds, as long as one remembers that the field carries momentum of its own. (If we want to save some version of Newton's third law, we could say that the real action-reaction pairs are the forces between the charges and the field, not the charges with each other. But the real lesson is that Newton's third law is not fundamental, momentum conservation is.)

Idea 7

The momentum density of the electromagnetic field is

$$\mathbf{p} = \frac{\mathbf{S}}{c^2}.$$

In other words, momentum density and energy flux density are just proportional. As you will see in **R2**, this is true in general in relativity. The angular momentum density is $\mathbf{r} \times \mathbf{p}$. For an explicit derivation that these definitions ensure the total momentum and angular momentum are conserved, see section 8.2 of Griffiths. (You might think the definitions come out of nowhere; the straightforward way to find them is to apply Noether's theorem, as you will learn in a more advanced class.)

Remark

We have already seen an example of electromagnetic field momentum at work. Back in **E4**, you found that in the presence of a magnetic monopole, the mechanical angular momentum $\mathbf{L}$ of a point charge was not conserved, but $\mathbf{L} - qg\hat{\mathbf{r}}$ was. In fact, this second term turns out to be exactly the angular momentum of the field, so this conservation law is simply the conservation of total angular momentum. (If you'd like to verify this explicitly, it's easiest to use spherical coordinates with the monopole at the origin and the charge along the z-axis, but be warned, it's fairly messy.)

[3] **Problem 14** (Griffiths)**.** A long coaxial cable of length ℓ consists of an inner conductor of radius a and an outer conductor of radius b. The inner conductor carries a uniform charge per unit length λ, and a steady current I to the right; the outer conductor has the opposite charge and current.

(a) Find the electromagnetic momentum stored in the fields.

(b) In part (a) you should have found that the fields contain a nonzero momentum directed along the cable. However, this is puzzling because it appears that no net mass is transported along the cable. How is this paradox resolved? (Hint: it doesn't make sense to consider the cable in isolation, as nothing would be keeping the current going. Consider attaching a battery across the left end and a resistor across the right end.)

Solution. (a) Set up the obvious cylindrical coordinates, with $\hat{\mathbf{z}}$ directed to the right. In between the tubes the fields are

$$\mathbf{E} = \frac{1}{2\pi\epsilon_0}\frac{\lambda}{s}\,\hat{\mathbf{s}}, \quad \mathbf{B} = \frac{\mu_0}{2\pi}\frac{I}{s}\,\hat{\boldsymbol{\phi}},$$

and they are zero everywhere else. Therefore, the momentum is $\int \epsilon_0(\mathbf{E}\times\mathbf{B})\,dV$ or

$$\mathbf{p} = \hat{\mathbf{z}}\,\frac{\mu_0 I\lambda}{4\pi^2}\int_a^b \frac{1}{s^2}\ell\,2\pi s\,ds = \frac{\mu_0 I\lambda\ell}{2\pi}\log(b/a)\,\hat{\mathbf{z}}.$$

(b) As this process goes on, the battery loses energy and the resistor gains energy (i.e. heats up). In other words, energy is being transported along the cable, just as we would expect from idea 7. But because $E = mc^2$, that means the resistor is gaining mass while the battery is losing mass, so the center of mass of the system is going to the right, as it must!

One might further ask where the initial momentum comes from. If the entire apparatus is on a table, the initial impulse comes from friction between the circuit elements and the table. If the apparatus were on a frictionless cart instead, the cart would have recoiled to the left when the battery was attached, so that the net momentum would remain zero.

[3] **Problem 15.** In the early 20$^{\text{th}}$ century, physicists sought to explain the $E = mc^2$ rest energy in terms of electromagnetic field energy. As a concrete example, model a charged particle as a uniform spherical shell of radius a and charge q.

(a) Find the radius a so that the total field energy equals the rest energy associated with the electron mass m. Up to an $O(1)$ factor, this quantity is called the classical electron radius.

(b) If the shell moves with a small speed v, we expect to have $p = mv$, where p is the total field momentum. Show that instead, we have $p = (4/3)mv$. You may use the result

$$\mathbf{B} = \frac{\mathbf{v}}{c^2}\times\mathbf{E}$$

which we will prove in **R3**. Many complicated ideas were put forth to explain this infamous "4/3 problem", as recounted in chapter II-28 of the Feynman lectures.

For more about the "radius" of an electron, see this blog post. For a modern discussion of the resolution of the 4/3 problem, see this paper.

Solution. (a) The electrostatic energy of a uniform spherical shell of radius a is

$$U = \frac{1}{2}qV = \frac{q^2}{8\pi\epsilon_0 a}$$

where the factor of 1/2 avoids double counting the energy. Setting $U = mc^2$ gives

$$a = \frac{q^2}{8\pi\epsilon_0 mc^2}.$$

(b) Note that

$$\mathbf{S} = \frac{1}{\mu_0} \mathbf{E} \times (\mathbf{v} \times \mathbf{E})/c^2.$$

Set up spherical coordinates where $\mathbf{v} \parallel \hat{\mathbf{z}}$. We then have that

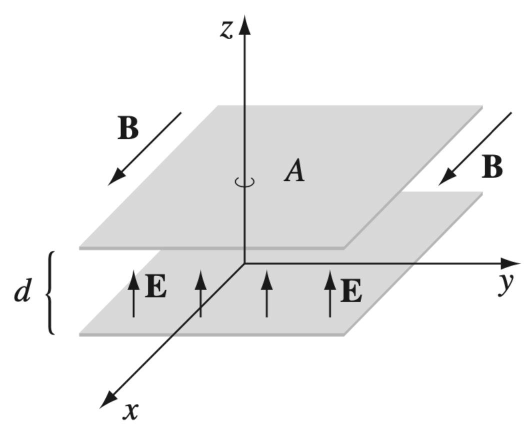
$$\begin{aligned}
\int \mathbf{E} \times (\mathbf{v} \times \mathbf{E}) dV &= \int \big(\mathbf{v}(E^2) - \mathbf{E}(\mathbf{v} \cdot \mathbf{E})\big)\, dV \\
&= (kq)^2 \hat{\mathbf{z}} \int_a^\infty \int_0^\pi \int_0^{2\pi} \left[v/r^4 - v \cos^2\theta / r^4 \right] r^2 \sin\theta \, d\phi \, d\theta \, dr \\
&= 2\pi (kq)^2 v \int_a^\infty \frac{1}{r^2} \int_0^\pi (\sin\theta - \cos^2\theta \sin\theta) \, d\theta \, dr \\
&= 2\pi (kq)^2 v \int_a^\infty \frac{4}{3} \frac{dr}{r^2} \\
&= \frac{8}{3} \frac{(kq)^2 \pi v}{a}
\end{aligned}$$

where $k = 1/4\pi\epsilon_0$. The momentum is then

$$p = \frac{1}{c^4 \mu_0} \frac{8}{3} \left(\frac{q}{4\pi\epsilon_0} \right)^2 \frac{\pi v}{a} = \frac{4}{3} \frac{1}{c^2} \frac{q^2}{8\pi\epsilon_0 a} v = \frac{4}{3} m v$$

as stated.

[3] **Problem 16** (Griffiths 8.6). A charged parallel plate capacitor is placed in a uniform magnetic field as shown.



(a) Find the electromagnetic momentum in the space between the plates.

(b) Now a resistive wire is connected between the plates, along the z-axis, so that the capacitor slowly discharges. The current through the wire will experience a magnetic force; show the total impulse equals the stored momentum.

(c) Alternatively, suppose we slowly reduced the magnetic field. Show that the total impulse delivered to the plates equals the stored momentum.

This calculation is standard and given in many textbooks, but it is actually completely wrong: we have ignored the fringe field, and when it is included the total electromagnetic momentum is half of what was naively calculated in part (a). The answer in part (b) is correct, but the other half of the impulse corresponds to a change in non-electromagnetic "hidden momentum". The most basic example of hidden momentum is covered in example 12.12 of Griffiths. For a detailed analysis of the hidden momentum in this setup, see this paper.

Solution. (a) Using the standard formula,

$$\mathbf{p} = \epsilon_0(\mathbf{E} \times \mathbf{B})Ad = \epsilon_0 EBAd\,\hat{\mathbf{y}}.$$

(b) The impulse $\mathbf{j}$ is

$$\mathbf{j} = \int_0^\infty I(\boldsymbol{\ell} \times \mathbf{B})\,dt = -(d\hat{\mathbf{y}})\int_0^\infty B\,dQ.$$

Performing the integral, the total impulse is

$$\mathbf{j} = BQd\,\hat{\mathbf{y}} = \epsilon_0 EBAd\,\hat{\mathbf{y}}$$

in agreement with part (a).

(c) By Faraday's law, a nonconservative electric field is generated in the setup, which pushes the plates with a net force. Note that when the situation is symmetric, the electric field is $\mathbf{E}' = (1/2)\dot{B}d\,\hat{\mathbf{y}}$ at the bottom plate, and $-\mathbf{E}'$ at the top plate. So the total impulse is

$$\mathbf{j} = \int (Q\mathbf{E}') + ((-Q)(-\mathbf{E}'))\,dt = \int \dot{B}Qd\hat{\mathbf{y}}\,dt = BQd\hat{\mathbf{y}} = \epsilon_0 EBAd\hat{\mathbf{y}}$$

in agreement with parts (a) and (b). The answer is the same if the setup were asymmetric, i.e. if the fields had been $\mathbf{E}' + \mathbf{E}_0$ and $-\mathbf{E}' + \mathbf{E}_0$ at the top and bottom plates, because $\mathbf{E}_0$ would not contribute to the net force.

[3] **Problem 17.** USAPhO 2004, problem B2. (This is a classic setup which also appears on USAPhO 2020, problem A1, and INPhO 2020, problem 2. But note that the official solution to USAPhO 2020, problem A1 has typos.)

4 Electromagnetic Waves

Idea 8

Maxwell's equations have propagating wave solutions of the form

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \mathbf{B} = \mathbf{B}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$$

where $\mathbf{E}$ and $\mathbf{B}$ are in phase, perpendicular in direction, and have magnitudes $E_0 = cB_0$. The propagation direction $\mathbf{k}$ is along $\mathbf{E} \times \mathbf{B}$, and the wave speed is

$$c = \frac{\omega}{k} = \frac{1}{\sqrt{\mu_0\epsilon_0}}.$$

Example 4

Verify explicitly that in the absence of charges and currents, the electromagnetic field above satisfies Maxwell's equations.

Solution

First let's consider Gauss's law, $\nabla \cdot \mathbf{E} = 0$. Splitting everything explicitly into components,

$$\begin{aligned}\nabla \cdot \mathbf{E} &= e^{-i\omega t}\left(\frac{\partial}{\partial x}(E_{0,x}e^{i\mathbf{k}\cdot\mathbf{r}}) + \frac{\partial}{\partial y}(E_{0,y}e^{i\mathbf{k}\cdot\mathbf{r}}) + \frac{\partial}{\partial z}(E_{0,z}e^{i\mathbf{k}\cdot\mathbf{r}})\right) \\ &= e^{-i\omega t}\left(E_{0,x}\frac{\partial}{\partial x}e^{i\mathbf{k}\cdot\mathbf{r}} + E_{0,y}\frac{\partial}{\partial y}e^{i\mathbf{k}\cdot\mathbf{r}} + E_{0,z}\frac{\partial}{\partial z}e^{i\mathbf{k}\cdot\mathbf{r}}\right) \\ &= e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\left(iE_{0,x}k_x + iE_{0,y}k_y + iE_{0,z}k_z\right) \\ &= i\mathbf{k}\cdot\mathbf{E} = 0\end{aligned}$$

since $\mathbf{k}$ is perpendicular to $\mathbf{E}_0$. This is another example of a lesson we saw in **M4**. Namely, when everything is a complex exponential, differentiation is very easy. For an complex exponential in time, $e^{i\omega t}$, differentiation with respect to time is just multiplication by $i\omega$. Similarly, for a field which is a complex exponential in space, $e^{i\mathbf{k}\cdot\mathbf{r}}$, the divergence $(\nabla\cdot)$ becomes $(i\mathbf{k}\cdot)$.

By similar reasoning, Gauss's law for magnetism is satisfied. Next, we check Ampere's law,

$$\nabla \times \mathbf{B} = \mu_0\epsilon_0\frac{\partial \mathbf{E}}{\partial t}.$$

By the same logic as above, the curl becomes $(i\mathbf{k}\times)$, while the time derivative becomes multiplication by $-i\omega$, giving

$$i\mathbf{k}\times\mathbf{B} = (-i\omega)\mu_0\epsilon_0\mathbf{E}.$$

Because $\mathbf{k}$, $\mathbf{E}$, and $\mathbf{B}$ are all mutually perpendicular, the directions of both sides match. Then all that remains is to check the magnitudes,

$$kB_0 = \omega\mu_0\epsilon_0E_0.$$

By plugging in results from above, this reduces to

$$c^2 = \frac{1}{\mu_0\epsilon_0}$$

which matches what we said above. (Or, if we didn't know what c case, this logic would have been a way to derive it, as Maxwell did.) The verification of Faraday's law is similar. Note that the displacement current term was essential; it wouldn't have been possible to get electromagnetic wave solutions without it.

[3] **Problem 18.** Consider the energy and momentum of the electromagnetic wave in idea 8.

(a) Show that the spatial average of the energy density is $\epsilon_0E_0^2/2$. (Be careful with factors of 2.)

(b) Compute the spatial average of the momentum density $\langle\mathbf{p}\rangle$ using idea 7.

(c) Confirm that $E = pc$ for an electromagnetic wave.

Solution. The key pitfall here is that we have to take into account the fact that for all nonlinear quantities such as these, we need to look only at the real part of the above formulas.

(a) The energy is $u = \frac{1}{2\epsilon_0}(E^2 + c^2B^2)$. The average value of E^2 is $E_0^2/2$ by the usual $\cos^2(kx)$ averaging trick, and the average value of B^2 is $B_0^2/2$. There are two terms, so the average of their sum is $\epsilon_0 E_0^2/2$.

(b) It's $\langle \epsilon_0 \mathbf{E} \times \mathbf{B} \rangle$. Again the product has the spatial form $\cos^2(kx)$, so the averaging over space gives a factor of $1/2$, giving $\epsilon_0 E_0^2/2c$.

(c) This is equivalent to showing that the energy density is equal to the momentum density times c, which is indeed true from our results above.

[3] **Problem 19.** The intensity of sunlight at noon is approximately $1\,\mathrm{kW/m^2}$.

(a) Compute the rms magnetic field strength.

(b) Compute the radiation pressure acting on a mirror lying on the ground.

(c) In terms of the Lorentz force, how is this pressure exerted on the particles in the mirror?

Solution. (a) Note that the intensity is the Poynting vector. We have $\langle S \rangle = \frac{c}{\mu_0} B_{\mathrm{rms}}^2$, so

$$B_{\mathrm{rms}} \approx 2 \times 10^{-6}\,\mathrm{T}.$$

(b) We get a factor of 2 because the radiation bounces off the mirror, giving

$$P = \frac{2S}{c} = 7 \times 10^{-6}\,\mathrm{Pa}.$$

(c) The basic idea is that the particles are accelerated in the direction of $\mathbf{E}$, and thus feel a force in the direction $(\mathbf{v} \times \mathbf{B}) \parallel (\mathbf{E} \times \mathbf{B}) \parallel \mathbf{S}$.

However, making this more quantitative is more subtle. For an ideal free particle, $\mathbf{a}$ is in phase with $\mathbf{E}$, which means $\mathbf{v}$ is 90° out of phase with $\mathbf{E}$. Since $\mathbf{B}$ is in phase with $\mathbf{E}$, the magnetic force $\mathbf{v} \times \mathbf{B}$ has time dependence of the form $\cos(\omega t)\sin(\omega t)$, which averages to zero.

On the other hand, suppose the particle is attached to a damped harmonic oscillator, and ω is at the resonant angular frequency. Then from **M4** results, $\mathbf{v}$ is instead in phase with $\mathbf{E}$, which means the magnetic force has time dependence $\cos^2(\omega t)$, which doesn't average to zero.

The point is, you need some kind of other force at play to produce a phase shift between $\mathbf{a}$ and $\mathbf{E}$, or else the force averages to zero. And this makes perfect sense from an energy conservation standpoint: a nonzero average force means momentum is taken out of the radiation, which means part of it is absorbed. This is only possible if the absorbed energy goes somewhere else, e.g. in the damped harmonic oscillator case it is dissipated by the damping force.

This raises yet another question: how it is possible for an isolated charge to scatter radiation, in Thomson or Compton scattering? The reason is that there *is* another force at play, namely the radiation reaction force acting on the particle. You can read more about this subtle force in section 11.2 of Griffiths.

[3] **Problem 20** (Purcell 9.7)**.** Consider the sum of two oppositely-traveling electromagnetic waves, with electric fields

$$\mathbf{E}_1 = E_0 \cos(kz - \omega t)\hat{\mathbf{x}}, \quad \mathbf{E}_2 = E_0 \cos(kz + \omega t)\hat{\mathbf{x}}.$$

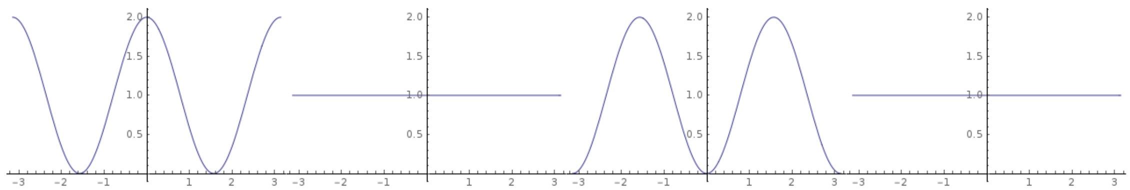
(a) Write down the magnetic field.

(b) Draw plots of the energy density $U(z,t)$ for $\omega t \in \{0, \pi/4, \pi/2, 3\pi/4, \pi\}$.

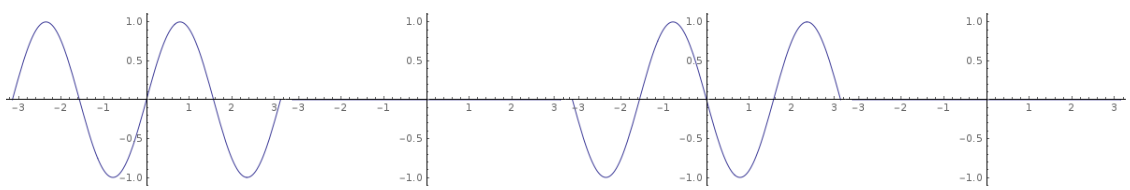
(c) On top of these plots, draw the direction and magnitude of the Poynting vector. Convince yourself that the Poynting vector accurately describes how the energy sloshes back and forth.

Solution. (a) We note that $\mathbf{B} = (E_0/c)(\cos(kz-\omega t) - \cos(kz+\omega t))\,\hat{\mathbf{y}} = (2E_0/c)\sin(kz)\sin(\omega t)\,\hat{\mathbf{y}}$.

(b) For the purposes of the following plots and calculations, we set $k = \omega = E_0 = B_0 = c = \mu_0 = \epsilon_0 = 1$. Note that $\mathbf{E} = \cos(z-t) + \cos(z+t) = 2\cos(z)\cos(t)$. Thus, $U(z,t) = 2(\sin^2(z)\sin^2(t) + \cos^2(z)\cos^2(t))$. The energy plots are:



(c) The Poynting vector is $\hat{\mathbf{z}}\sin(2z)\cos(2t)$. The plots are:



Idea 9: Larmor Formula

An accelerating charge produces electromagnetic radiation, with power

$$P = \frac{q^2 a^2}{6\pi\epsilon_0 c^3}.$$

We'll derive it properly in **R3**, but a lot of it can be motivated with the techniques of **P1**.

The power could only depend on q, ϵ_0, μ_0, and properties of the particle's motion. The only combinations of the first three parameters that get rid of the electromagnetic units are q^2/ϵ_0 and $1/\sqrt{\epsilon_0\mu_0} = c$. Since energy is proportional to the electric and magnetic fields squared, and these fields are proportional to q, the answer must be proportional to q^2/ϵ_0.

Radiation can't result from uniform velocity, by Lorentz invariance; another way to see this is that with only v and c, there is no way to write down an expression for power with the right units! The next simplest option is radiation from acceleration, from which the most general result is $P = (q^2a^2/\epsilon_0 c^3) f(v/c)$. The fact that acceleration is squared is also natural, because acceleration is a vector, so this is the simplest way to get a rotationally invariant result. The proper derivation shows that $f(0) = 1/6\pi$. When v/c is substantial, there are relativistic corrections, which we will consider in **R3**.

[2] **Problem 21** (Purcell H.2). A common classical model of an electron in an atom is to imagine it is a mass on a spring, where the spring force is due to the atomic nucleus. Suppose that such an electron, with charge e, is vibrating in simple harmonic motion with angular frequency ω and amplitude A.

(a) Find the average rate of energy loss by radiation.

(b) If no energy is supplied to make up the loss, how long will it take the oscillator's energy to fall to $1/e$ of its initial value?

Numerically, this is an extremely small time, so classical models of the atom are not realistic. We will see in **X1** that in quantum mechanics this problem is solved because in the ground state the electron does not move around the atom, but rather occupies a standing wave.

Solution. (a) The average value of a^2 is $A^2\omega^4/2$, giving

$$\langle P\rangle = \frac{e^2 A^2 \omega^4}{12\pi\epsilon_0 c^3}.$$

(b) If m is the mass of the electron, then the energy of the system is $E = \frac{1}{2}m\omega^2 A^2$. We see that

$$\dot{E} = -\frac{e^2\omega^4}{12\pi\epsilon_0 c^3}\frac{2}{m\omega^2}E = -\frac{e^2\omega^2}{6m\pi\epsilon_0 c^3}E.$$

This is an exponential decay, with characteristic time

$$t = \frac{6m\pi\epsilon_0 c^3}{e^2\omega^2}.$$

[3] **Problem 22** (Purcell H.3). A plane electromagnetic wave with angular frequency ω and electric field amplitude E_0 is incident on an atom. As in problem 21, we model the electron as a simple harmonic oscillator, with mass m and natural angular frequency ω_0.

(a) First suppose that $\omega \gg \omega_0$. Argue that in this case, the "spring" force on the electron can be neglected. Find the average power radiated by the electron, and show that it is equal to the power incident on a disc of area

$$\sigma = \frac{8\pi}{3}\left(\frac{e^2}{4\pi\epsilon_0 mc^2}\right)^2.$$

This is the Thomson scattering cross section. To an electromagnetic wave, each electron looks like it has this area.

(b) Now suppose $\omega \ll \omega_0$, yielding Rayleigh scattering, which describes the scattering of visible light by air. In this case, show that $\sigma \propto \omega^4$. This sharp frequency dependence explains why the sky is blue. (Why isn't it violet?)

(c) Explain the meaning of the common phrase "red sky at night, sailor's delight; red sky in morning, sailor's warning". (Hint: in the cultures where this saying is used, weather patterns usually move from west to east.)

For some further discussion of Rayleigh scattering, see section 9.4 of The Art of Insight. For more about colors in the atmosphere, see this nice video.

Solution. (a) As seen in **M4**, the amplitude of a driven harmonic oscillator is

$$A = \frac{F_0/m}{\sqrt{(\omega_0^2 - \omega^2)^2 + (b\omega/m)^2}}.$$

Here, $b = 0$ and since $\omega \gg \omega_0$, we have $A = eE_0/m\omega^2$. Putting this into our answer in part a of the previous problem gets

$$\langle P \rangle = \frac{e^2\omega^4}{12\pi\epsilon_0 c^3}\frac{e^2E_0^2}{m^2\omega^4} = \frac{e^4E_0^2}{12\pi\epsilon_0 m^2c^3}.$$

Since for electromagnetic radiation, $\langle S \rangle = \frac{1}{2}\epsilon_0 c E_0^2$ so $\langle P \rangle = \sigma\langle S \rangle$, we can put this in the above expression to get

$$\sigma = \frac{e^4E_0^2}{12\pi\epsilon_0 m^2c^3}\frac{2}{\epsilon_0 c E_0^2} = \frac{8\pi}{3}\left(\frac{e^2}{4\pi\epsilon_0 mc^2}\right)^2.$$

(b) Now, our expression for A will be $eE_0/m\omega_0^2$. Referring to part (a), we have $\sigma \propto A^2$, so

$$\sigma = \frac{8\pi}{3}\left(\frac{e^2}{4\pi\epsilon_0 mc^2}\right)^2\frac{\omega^4}{\omega_0^4}.$$

In other words, higher frequencies are scattered much more. The atmosphere scatters most of the blue light from the Sun, and some of it hits your eyes, making the sky look blue.

The reason the sky isn't purple is actually a bit complex. At the simplest level, it's because the Sun doesn't produce all that much purple light. However, the physiology of color vision also plays a role; see here for details.

(c) See this nice explanation for details.

[3] **Problem 23.** USAPhO 2016, problem B2.

Remark

We noted in **M7** that clouds are visible because the radiation scattered by a small droplet of n water molecules grows as n^2. To understand why, note that each of the molecules performs independent Rayleigh scattering, as computed above. For separated molecules, the energy scattered just adds. However, for nearby molecules the electromagnetic waves scattered interfere constructively, so the amplitude grows as n and hence the energy scattered as n^2.

This quadratic enhancement breaks down once the droplets exceed the wavelength λ of the light. This means the maximum possible enhancement is larger for larger wavelengths, acting against the ω^4 dependence of Rayleigh scattering. This is why clouds are white, not blue.

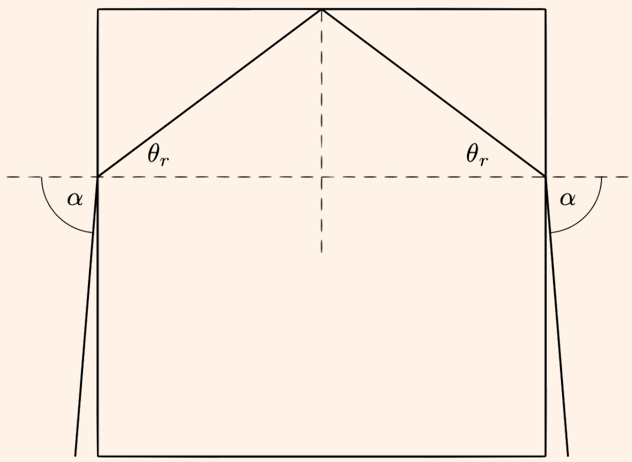
Radiation pressure can also have mechanical effects.

Example 5: NBPhO 2018.6

A laser pointer of power P is directed at a glass cube, with refractive index $n > \sqrt{2}$. The surface of the cube has an anti-reflective coating, so there is no partial reflection when light enters or exits it; the laser pointer only refracts. What is the maximum force the laser pointer can exert on the cube?

Solution

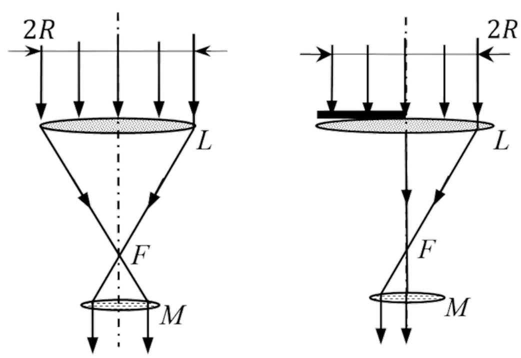
The force is due to a change in momentum of the light. The greatest possible force is attained if the direction of the light is reversed, which can occur as shown, in the limit $\alpha \to 90°$.



Assuming $n > \sqrt{2}$, we then have $\theta_r < 45°$, and then the laser internally reflects when it hits the top surface of the cube. It exits in the opposite direction it came in.

If the laser pointer has power P, then the momentum of the laser beam per time is P/c. The momentum is reversed, so the force is $2P/c$.

[3] **Problem 24** (IZhO 2022)**.** In 2018, the Nobel Prize in physics was awarded to Arthur Ashkin for the creation of the "laser tweezer", a device that allows one to hold and move transparent microscopic objects with the help of light. In one such device, a parallel beam of light from a laser passes through a converging lens L and hits a microparticle M, which can also be considered a converging lens. Point F is the common focus of L and M.



The light intensity in the beam is $I = 1.00\,\mu\mathrm{W/cm}^2$, the beam radius is $R = 1.00\,\mathrm{cm}$, and the focal length of the lens L is $F = 10.0\,\mathrm{cm}$. Ignore the absorption and reflection of light.

(a) Calculate the force acting on the microparticle, in the setup shown at left above.

(b) Next, the left half of the lens L is covered by a diaphragm, as shown at right above. Calculate the force acting on the microparticle in the transverse direction of the beam.

Solution. See the official solutions here.

[3] **Problem 25** (Feynman)**.** In one proposed means of space propulsion, a spaceship of mass 10^3 kg carries a thin sheet of area $100\,\mathrm{m}^2$. The sheet is made of highly reflective plastic film, and can be used as a solar radiation pressure "sail". The spaceship travels in a circular orbit of radius r, which is initially equal to the Earth's orbit radius, where the intensity of sunlight is $1400\,\mathrm{W/m}^2$. Assume the spaceship is moving nonrelativistically and the gravitational effect of the Earth is negligible.

(a) Find the angle at which the sail should be pointed to maximize dr/dt.

(b) Assuming the sail is pointed this way, find the numeric value of dr/dt.

(c) If this continues for a very long time, then r will grow as $r \propto t^n$. Find the value of n.

Solution. (a) We have that $E = -GMm/2r$, so increasing radius as fast as possible is the same as imparting the most energy to the spaceship; in other words, we want to maximize $\mathbf{F}\cdot\mathbf{v}$. The velocity is almost purely tangential, so only the tangential component of the force contributes to the power.

Let the normal vector of the sail be at an angle θ to the radial direction. To find the net force on the sail, we can think of the reflection process in two steps: absorbing the light and reemitting it. Absorbing the light yields an outward radial force F. Reemitting it yields a force F of the same magnitude, directed at an angle 2θ to the radial direction.

Only the second force can contribute to the power, so

$$P = \mathbf{F}\cdot\mathbf{v} = Fv\,\sin 2\theta.$$

Finally, since the radial area covered by the sail is $A\cos\theta$, we have $F = I(A\cos\theta)/c$, so

$$P = \frac{2IAv}{c}\sin\theta\cos^2\theta.$$

Setting the derivative to zero, the maximum is at $\theta = \sin^{-1}(1/\sqrt{3}) = 35.3^\circ$, giving power

$$P_{\max} = \beta\,\frac{IAv}{c}, \quad \beta = \frac{4}{3\sqrt{3}} \approx 0.77.$$

Strictly speaking, the answer is very slightly different, by corrections of order v/c, since the spaceship is moving, but we'll neglect this here.

(b) The power is

$$P = \frac{dE}{dt} = \frac{GMm}{2r^2}\frac{dr}{dt}.$$

We also know from force balance that

$$\frac{GMm}{r^2} = \frac{mv^2}{r}$$

since the radial force from the sail is negligible, and combining these results gives

$$\frac{dr}{dt} = \frac{\beta IA}{c}\frac{2r}{mv} = \frac{2\beta IA}{m\omega c}$$

where $\omega = 2\pi/(1\,\text{year})$ is the angular velocity of the Earth. Plugging in the numbers gives $dr/dt = 4\,\mathrm{m/s}$.

(c) We have

$$\frac{dr}{dt} \propto \frac{Ir}{v} \propto \frac{(1/r^2)\, r}{1/\sqrt{r}} = \frac{1}{\sqrt{r}}.$$

Separating and integrating gives $t \propto r^{3/2}$ in the long run (i.e. when the contribution from the initial condition is negligible), so $n = 2/3$.

Finally, we'll consider electromagnetic wave propagation in transmission lines.

[4] **Problem 26** (Griffiths 7.62, Crawford 4.8)**.** A certain transmission line is constructed from two thin metal ribbons, of width w, a very small distance $h \ll w$ apart. The current travels down one strip and back along the other. In each case it spreads out uniformly over the surface of the ribbon.

(a) Find the capacitance per unit length $\mathcal{C}$, and the inductance per unit length $\mathcal{L}$.

(b) Argue that the speed of propagation of electromagnetic waves through this transmission line is of order $1/\sqrt{\mathcal{LC}}$, and evaluate this quantity.

(c) Repeat the first two parts for a coaxial transmission line, consisting of two cylinders of radii $a < b$ with the same axis of symmetry.

(d) Repeat the first two parts for a parallel-wire transmission line, consisting of two wires of radius r whose axes are a distance $D \gg r$ apart.

You should find that in all cases, $1/\sqrt{\mathcal{LC}}$ is the same, yielding the same speed for electromagnetic waves. This isn't a coincidence, and applies for transmission lines with conductors of any shape, though the general proof requires some elaborate vector calculus, as you can see here.

Solution. Suppose the length is ℓ.

(a) We have $C = \epsilon_0 w\ell/h$, so $\mathcal{C} = \epsilon_0 w/h$. Similarly, $L = \frac{\mu_0 h}{w}\ell$, so $\mathcal{L} = \mu_0 h/w$.

(b) Note that the time scale of this LC circuit is $\sqrt{LC}$. Thus, the transmission velocity is on the order of $\ell/\sqrt{LC} = 1/\sqrt{\mathcal{LC}} = c$. Alternatively, this is easily implied by dimensional analysis.

(c) The capacitance of two coaxial cylinders can be found by giving the inner cylinder a charge Q and using $C = Q/V$, where

$$V = \int E\, ds = \int_a^b \frac{(Q/\ell)}{2\pi\epsilon_0 r} dr = \frac{Q/\ell}{2\pi\epsilon_0} \log(b/a)$$

and therefore

$$\mathcal{C} = \frac{2\pi\epsilon_0}{\log(b/a)}.$$

To find the inductance for a transmission line setup, the current will flow parallel to the axis, so by Ampere's law the field inside the region is $B = \mu_0 I/2\pi r$, and using $L = \Phi/I$ where Φ will be the flux going around the inner cylinder will give

$$\Phi = \int_a^b \ell \frac{\mu_0 I}{2\pi r}\, dr = \frac{\mu_0 I \ell}{2\pi} \log\left(\frac{b}{a}\right)$$

and therefore

$$\mathcal{L} = \frac{\mu_0}{2\pi} \log\left(\frac{b}{a}\right).$$

Thus we get $1/\sqrt{\mathcal{LC}} = c$.

(d) The capacitance can be found similarly (factors of 2 appear because the negative charge/opposing current will contribute to the E and B fields too),

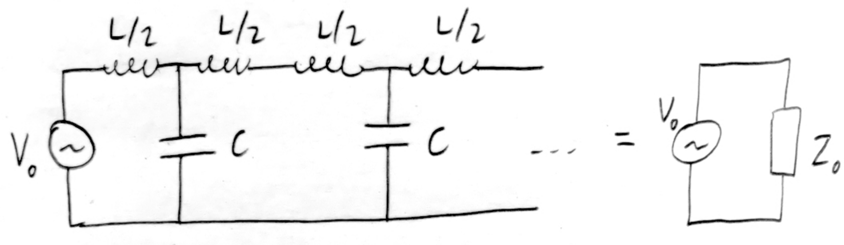
$$V = \int_r^D \frac{2Q/\ell}{2\pi\epsilon_0 r}\, dr, \quad \mathcal{C} = \frac{\pi\epsilon_0}{\log(D/r)}.$$

$$\Phi_B = \int_r^D \frac{2\mu_0 I}{2\pi r}\, dr, \quad \mathcal{L} = \frac{\mu_0}{\pi}\log\left(\frac{D}{r}\right).$$

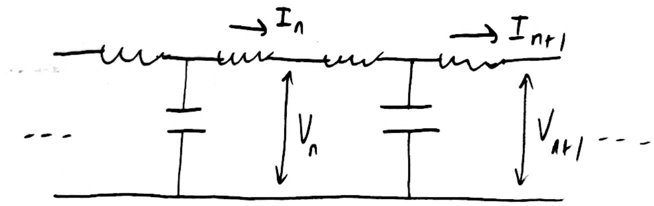
This gives $1/\sqrt{\mathcal{LC}} = c$.

[4] **Problem 27.** In this problem, we treat electromagnetic wave propagation through a transmission line using a "lumped element" approach, where the line is replaced with discrete capacitors and inductors, as shown. (This is an example of a network synthesis, mentioned in **E6**.)

(a) Calculate the characteristic impedance $Z_0(\omega)$ of the entire network, as shown below.



(b) The diagram below shows two adjacent sections of the ladder.



Find the ratio of the complex voltage amplitudes V_{n+1}/V_n.

(c) The AC driving attempts to create electromagnetic waves which travel through the network, to the right. It turns out that above a certain critical angular frequency ω_c, waves will not travel through the ladder network. Find ω_c. (Hint: this can be done using either the result of part (a) or part (b).)

(d) For angular frequencies $\omega \ll \omega_c$, waves travel through the ladder with a constant speed. Find this speed, assuming each segment of the ladder has physical length ℓ. (Hint: the speed of a wave obeys $v = d\omega/dk$.)

(e) You should have found in one of the earlier parts that the impedance of this infinite network can be a real number, even though it's made of parts which all have imaginary impedance. That sounds strange, but what's even stranger is that we *should* be able to handle this infinite circuit by taking the limit of progressively larger finite circuits, just as we did for a similar network of resistors in **E2**. But for *any* finite LC network, the impedance will be imaginary, so the limit must be imaginary too! On one hand, we should trust the finite result because all real circuits are finite. On the other hand, the real impedance we get for the infinite result certainly can be measured in real life. So what's going on?

Solution. (a) The impedance of the infinite ladder doesn't change if we add another unit onto the left. Let the inductor have impedance $Z_1/2$ and let the capacitor have impedance Z_2. Then

$$\frac{Z_1}{2} + \frac{1}{\frac{1}{Z_2} + \frac{1}{Z_1/2+Z_0}} = Z_0$$

which can be solved to give

$$Z_0 = \sqrt{(Z_1/2)^2 + Z_1 Z_2}.$$

Since we have $Z_1 = i\omega L$ and $Z_2 = 1/i\omega C$, we have

$$Z_0 = \sqrt{\frac{L}{C} - \frac{\omega^2 L^2}{4}}.$$

(b) Each segment sees an impedance Z_0 to its right, so

$$V_n = I_n Z_0, \quad V_{n+1} = I_{n+1} Z_0.$$

On the other hand, we also have

$$V_{n+1} - V_n = \frac{I_n Z_1}{2} + \frac{I_{n+1} Z_1}{2}$$

and solving these equations yields

$$\frac{V_{n+1}}{V_n} = \frac{Z_0 + Z_1/2}{Z_0 - Z_1/2} = \frac{\sqrt{L/C - \omega^2 L^2/4} - i\omega L/2}{\sqrt{L/C - \omega^2 L^2/4} + i\omega L/2} = \frac{\sqrt{4/\omega^2 LC - 1} - i}{\sqrt{4/\omega^2 LC - 1} + i}.$$

(c) First we'll find the critical angular frequency using part (b). When the square root is a real number, the numerator and denominator have equal magnitudes, so $|V_{n+1}| = |V_n|$, indicating wave propagation. When the square root is imaginary, the wave instead exponentially decays. The cutoff is when

$$4/\omega^2 LC - 1 = 0$$

which gives

$$\omega_c = \frac{2}{\sqrt{LC}}.$$

To derive the same conclusion using the result of part (a), note that the impedance Z_0 becomes real when $\omega < 2/\sqrt{LC}$. How could one get a real impedance, which signals energy loss, if there are no resistors anywhere in the circuit? It can only happen if the driver can create electromagnetic waves, which then propagate through the network; since the network is infinite, this energy never returns to the driver. Because waves can appear for $\omega < 2/\sqrt{LC}$, we again conclude that $\omega_c = 2/\sqrt{LC}$.

(d) In this limit, we have

$$\frac{V_{n+1}}{V_n} \approx \frac{2/\omega\sqrt{LC} - i}{2/\omega\sqrt{LC} + i}$$

and so across each unit, there is a phase shift of

$$\delta = \omega\sqrt{LC}.$$

Since wavenumber is phase shift per distance, $k = \delta/\ell = \omega\sqrt{LC}/\ell$, which means

$$v = \frac{d\omega}{dk} = \frac{\ell}{\sqrt{LC}}.$$

That is, waves in a transmission line travel with a constant speed, as we already found in problem 26. If we further plug in the L and C found in that problem, we would recover the speed of light.

(e) For an ideal, finite LC network, the finite result is perfectly correct: the impedance is pure imaginary. The network can't absorb net energy, because in the steady state energy propagates through the network, bounces off the other end, and comes back to the voltage source. However, when we're using transmission lines in practice, we put a load on the other end (i.e. a resistance) that absorbs the incoming wave. This introduces a real impedance to the finite circuit, and the limiting procedure works just fine, recovering a real impedance in the infinite limit.

But you care about the mathematics, you might object in the following way: in the infinite network analysis, we never needed to use the fact that a real impedance was at the end, because there *was* no end. You get the infinite network either by taking the limit of finite LC circuits, or by taking the limit of finite LC circuits terminated by a resistor, so how do we mathematically choose which limit gives the right answer?

The resolution is that the former limit does not even exist: as the size of the LC network is increased, the impedance keeps bounces around, never settling down to a limit. Physically, this is because the total length of the network is changing, which changes the phase shift of the wave once it gets back to the voltage source. It's analogous to trying to compute $\lim_{a\to\infty} \int_0^a e^{ix}\, dx$.

In order for the limit to be defined, we need to introduce resistances somewhere. For example, we could add a small resistance Δr to every inductor, which is also perfectly realistic. The limit is now well-defined, since the waves gradually decay away, and in the infinite limit we get some impedance $Z(\Delta r, \omega)$. Finally, taking the limit $\Delta r \to 0$ recovers the infinite result we derived earlier. It's precisely the same result as taking the infinite limit of LC networks terminated by resistors. We could also get the same result by giving the capacitors the small resistance.

The general lesson here is that in math, often limits are undefined, but in physics we can come up with "regulators" that make the limits defined, and which correspond to real-world effects. We then compute the real-world result by taking the limit, and then removing the regulator. The miracle is that very often, the final result does not depend on the regulator at all! This is very surprising to the mathematician, but very natural to the physicist: it simply says that the behavior of real objects, which always come with all of these regulating imperfections, can't possibly depend on the fine details of how we choose to model them.

Electromagnetism VIII: Synthesis

Electromagnetism in matter is covered in chapters 10 and 11 of Purcell. For more on dielectrics, see chapters II-10 and II-11 of the Feynman lectures. Electromagnetism in matter is covered in greater detail in chapters 4, 6, and 9 of Griffiths, and chapters I-31 and II-32 through II-37 of the Feynman lectures. For an enlightening overview of the history of magnetism, see chapter 1 of *Magnetism and Magnetic Materials* by Coey. There is a total of **97** points.

1 Polarization

In **E2**, we introduced the basics of dielectrics. To review: when a dielectric is placed in an electric field, dipoles inside align with the field, reducing the field value. In very symmetrical situations, the field is simply reduced by a factor of the dielectric constant $\kappa = \epsilon/\epsilon_0$. The total energy density within a dielectric is $\epsilon E^2/2$. This section is about problems which require more than these few facts. To answer them, we need to think about the charge bound to the dielectric itself.

Idea 1: Bound Charge

The polarization $\mathbf{P}$ of a material is its electric dipole moment per unit volume. It corresponds to a "bound" charge density

$$\rho_b = -\nabla \cdot \mathbf{P}$$

within the dielectric, as well a bound surface charge density

$$\sigma_b = \mathbf{P} \cdot \hat{\mathbf{n}}$$

on its surface.

Example 1

Find the electric field of a sphere with uniform polarization $\mathbf{P}$ and radius R.

Solution

There is no bound charge density inside the sphere, but a bound surface charge density

$$\sigma_b = \mathbf{P} \cdot \hat{\mathbf{r}} = P \cos\theta$$

on its surface. We could apply Coulomb's law to this charge density, but an easier method is to recall that polarization just means an internal displacement of charge. This surface charge density precisely corresponds to having two uniformly charged balls of total charge $\pm Q$ displaced by a tiny amount $\mathbf{d}$ so that $Q\mathbf{d} = (4\pi R^3/3)\mathbf{P}$.

By the shell theorem, the resulting field inside is uniform, and points against $\mathbf{P}$,

$$\mathbf{E} = -\frac{\mathbf{P}}{3\epsilon_0},$$

and the field outside is exactly a dipole field, with dipole moment $\mathbf{p} = (4\pi R^3/3)\mathbf{P}$.

[2] **Problem 1.** An infinite cylindrical rod of radius R has a uniform polarization $\mathbf{P}$.

(a) If $\mathbf{P}$ is perpendicular to the rod's axis, describe $\mathbf{E}$ outside, and find the value of $\mathbf{E}$ inside.

(b) If $\mathbf{P}$ is parallel to the rod's axis, find the electric field everywhere.

Now let's think about how polarization arises in the first place.

Idea 2: Electric Susceptibility

A small number of insulators are "ferroelectrics", whose crystal structure lets them maintain a preferred polarization $\mathbf{P}$ in the absence of external electric fields. (You can suppose that the preceding example and problem were implicitly about ferroelectrics.) But the vast majority of insulators are dielectrics, whose polarization is related to the total electric field by

$$\mathbf{P} = \epsilon_0 \chi_e \mathbf{E}, \quad \epsilon = \epsilon_0(1 + \chi_e) = \epsilon_0 \kappa$$

where χ_e is the electric susceptibility and κ is the dielectric constant. The susceptibility is nonnegative, except in some very exotic materials. The key difficulty is that above, $\mathbf{E}$ is the *total* electric field, including that due to the bound charge, which in turn depends on $\mathbf{P}$.

Example 2

A point charge q is inside a dielectric sphere of radius R with dielectric constant κ. Find the electric field and charge density everywhere.

Solution

This is one of the simple symmetric cases where the electric field in the dielectric is simply reduced by a factor of κ,

$$\mathbf{E} = \frac{q\,\hat{\mathbf{r}}}{4\pi\epsilon_0 r^2} \times \begin{cases} 1/\kappa & r < R \\ 1 & r > R \end{cases}.$$

Inside the sphere, this corresponds to an electric polarization

$$\mathbf{P} = \frac{q}{4\pi r^2}\frac{\chi_e}{\kappa}\,\hat{\mathbf{r}}.$$

To check that this solution is actually right, we need to ensure the original point charge q, plus the bound charge, indeed generates the claimed electric field.

The divergence of $\mathbf{P}$ is zero everywhere besides the origin, where negative bound charge piles up to cancel some of the charge q. The charge at the origin is thus

$$q - q_b = q\left(1 - \frac{\chi_e}{\kappa}\right) = q\left(1 - \frac{\kappa - 1}{\kappa}\right) = \frac{q}{\kappa}$$

which is consistent with Gauss's law for $\mathbf{E}$ there. At the surface of the sphere, there is a positive bound surface charge density

$$\sigma_b = \frac{q}{4\pi R^2}\frac{\chi_e}{\kappa}$$

which cancels the negative bound charge at the origin. Thus, by the shell theorem, the electric field outside the sphere is indeed that of the point charge q alone.

Example 3

A dielectric sphere of radius R and dielectric constant κ is placed in a uniform field $\mathbf{E}_0$, and as a result develops a uniform polarization $\mathbf{P}$. Find $\mathbf{P}$ and the field inside the sphere.

Solution

In example 1, we found the electric field due to the polarized sphere itself, which we'll call $\mathbf{E}_p$. Here, we must remember that the polarization is produced in response to the *total* electric field inside the sphere,

$$\mathbf{P} = \chi_e \epsilon_0 \mathbf{E}, \quad \mathbf{E} = \mathbf{E}_0 + \mathbf{E}_p.$$

Using our previous result for $\mathbf{E}_p$ and solving the system, we find

$$\mathbf{E} = \frac{3}{\kappa+2}\,\mathbf{E}_0, \quad \mathbf{P} = 3\,\frac{\kappa-1}{\kappa+2}\,\epsilon_0 \mathbf{E}_0.$$

The polarizability α of each atom is defined as the dipole moment per applied field,

$$\mathbf{p} = \alpha \mathbf{E}_0$$

so we have shown above that

$$\alpha = \frac{3\epsilon_0}{n}\frac{\kappa-1}{\kappa+2}$$

where n is the number density of atoms. This is the Clausius–Mossotti formula; it relates the macroscopically measurable parameter κ to the microscopic parameter α.

[2] **Problem 2** (Purcell 10.10)**.** Assume that the uniform field $\mathbf{E}_0$ that causes the electric field in example 2 is produced by large capacitor plates very far away. The field lines tangent to the sphere hit each of the distant capacitor plates in a circle of radius r. Find r in terms of R and κ.

Idea 3

The "free" charge density ρ_f is the part of the charge density that isn't bound, so that

$$\rho = \rho_b + \rho_f.$$

If we take the divergence of $\mathbf{P} = \epsilon_0 \chi_e \mathbf{E}$, we get $\rho_b = -\chi_e \rho$ inside a uniform dielectric, so

$$\rho = \rho_f / \kappa.$$

That is, a uniform dielectric "screens" charges embedded within it, reducing it by a factor of κ. That's exactly what we saw in example 2, and it also tells us that a conductor can be viewed as a dielectric with $\kappa \to \infty$, because conductors completely expel electric fields.

The difficulty in dealing with dielectrics is when κ changes in space, such as at the boundary of a dielectric, where a bound surface charge density σ_b can appear. If all the free charges in a problem are outside of uniform dielectrics, bound charges only appear on their surfaces.

[3] **Problem 3.** A version of the method of images, introduced in **E2**, works for dielectrics. Let's suppose there is vacuum at $z > 0$, a dielectric κ at $z < 0$, and a point charge q a distance d above

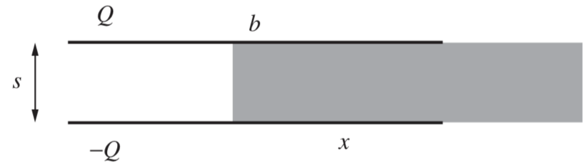
the plane $z = 0$. We need to find the surface bound charge density σ_b that appears on the plane.

(a) Let E_0^z be the z-component of the electric field due to the point charge alone. At a given point just below the plane $z = 0$, find E^z in terms of E_0^z and σ_b.

(b) Use this result to solve for σ_b in terms of E_0^z and κ.

(c) Your answer will be exactly the same as what one gets for a conductor at $z < 0$, multiplied by a κ-dependent constant. Using this information, characterize the image charge and find the force on the real charge.

[3] **Problem 4** (Purcell 10.2). A rectangular capacitor with side lengths a and b has separation $s \ll a, b$. It is partially filled with a dielectric with dielectric constant κ. The overlap distance is x.



The capacitor is isolated and has constant charge Q.

(a) What is the energy stored in the system?

(b) Using the result of part (a), what is the force on the dielectric? Which direction does it point?

(c) Is your answer to part (b) affected by the presence of fringe fields near the interface?

[3] **Problem 5** (Griffiths 4.28). Two long coaxial cylindrical metal tubes of inner radius a and outer radius b stand vertically in a tank of dielectric oil, with susceptibility χ_e and mass density ρ. The inner one is maintained at potential V, and the outer one is grounded. To what height h does the oil rise in the space between the tubes?

2 Magnetization

Idea 4

As discussed in **E5**, materials contain two kinds of magnetic dipole moments: the “orbital” part, due to moving electrons, and the “spin” part, due to the electrons’ intrinsic magnetic moments. For most materials, in the absence of external magnetic fields, these dipole moments point in random directions, and thus sum to zero on average.

When such a material is placed in a magnetic field, two things happen at once:

- The spins partially align with the field, producing a net dipole moment along $\mathbf{B}$.
- The orbits are affected by the changing field in accordance with Lenz’s law, and thus produce a net dipole moment *against* $\mathbf{B}$.

These effects are often comparable in size. If the first is more important, the material is paramagnetic, and if the second is more important, it is diamagnetic.

This can be a bit tricky to remember, because it seems opposite to the definition of a dielectric, where the internal electric dipoles try to align *with* the external field. The reason it makes sense is that inside an electric dipole, the electric field points against the dipole moment, while inside a magnetic dipole, the magnetic field points with the dipole moment, as discussed in **E3**. So, both dielectrics and diamagnets try to reduce the applied field within them.

Idea 5: Bound Currents

The magnetization $\mathbf{M}$ of a material is its magnetic dipole moment per unit volume. It corresponds to a bound current density

$$\mathbf{J}_b = \nabla \times \mathbf{M}$$

as well as a surface bound current density

$$\mathbf{K}_b = \mathbf{M} \times \hat{\mathbf{n}}$$

on its surface.

Example 4

Find the magnetic field of a sphere with uniform magnetization $\mathbf{M}$ and radius R.

Solution

In this case $\mathbf{J}_b$ is zero in the sphere, while at the sphere's surface,

$$\mathbf{K}_b = \mathbf{M} \times \hat{\mathbf{r}} = M \sin\theta\, \hat{\boldsymbol{\phi}}$$

where we worked in spherical coordinates and aligned $\mathbf{M}$ with the z-axis. However, this is precisely the current density of a rotating, uniformly charged sphere, as we discussed in **E3**. Scaling the constants appropriately, we find that inside,

$$\mathbf{B} = \frac{2}{3}\mu_0 \mathbf{M}$$

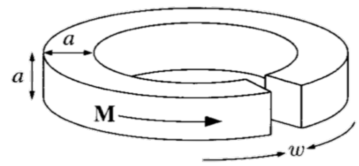
which should be compared with example 3. Outside, the field is exactly a magnetic dipole field, with $\mathbf{m} = (4\pi R^3/3)\mathbf{M}$.

[2] **Problem 6.** An infinite cylindrical rod of radius R has a fixed, uniform magnetization $\mathbf{M}$.

(a) If $\mathbf{M}$ is parallel to the rod's axis, find the magnetic field everywhere.

(b) If $\mathbf{M}$ is perpendicular to the rod's axis, describe $\mathbf{B}$ outside, and find the value of $\mathbf{B}$ inside.

[2] **Problem 7** (Griffiths 6.10)**.** A rod of length L and square cross section of side a is given a uniform longitudinal magnetization $\mathbf{M}$ and then bent into a circle with a narrow gap of width w.



Find the magnetic field at the center of the gap, assuming $w \ll a \ll L$.

Idea 6: Magnetic Susceptibility

Permanent magnets, such as the ones on your refrigerator, or the strong neodymium magnets popular in science toys, are made of "hard" ferromagnets. These are materials whose crystal structure lets them maintain a fixed magnetization $\mathbf{M}$ in the absence of external magnetic fields. (The preceding example and problems were implicitly about hard ferromagnets.)

In most other materials, the magnetization is related to the magnetic field in the material by

$$\mathbf{M} = \frac{1}{\mu_0} \frac{\chi_m}{1 + \chi_m} \mathbf{B}, \quad \mu = \mu_0(1 + \chi_m) = \mu_0 \mu_r$$

where χ_m is the magnetic susceptibility, μ is called the permeability, and μ_r is called the relative permeability. We'll explain later why χ_m isn't defined the same way as χ_e.

- Diamagnets have $-1 < \chi_m < 0$ and paramagnets have $\chi_m > 0$.
- Most materials are weakly diamagnetic ($|\chi_m| \ll 1$), but some are weakly paramagnetic.
- As discussed in **E5**, a superconductor totally expels magnetic fields, and thus can be viewed as a "perfect diamagnet" with $\chi_m = -1$ and hence $\mu = 0$.
- A "soft" ferromagnet (such as iron) is strongly paramagnetic, with $\chi_m \gg 1$.
- It is impossible to have $\chi_m < -1$, as then the energy density $B^2/2\mu$ would be negative. The material would spontaneously develop arbitrarily large B, and blow itself up.

Because magnetization can arise from freely moving electrons, bound electrons orbiting, or the spin of electrons, these ideas can be applied to both conductors and insulators. As always, we must be careful to remember that $\mathbf{B}$ is the total magnetic field, due to both whatever is outside the material, and the magnetization of the material itself.

Remark: Estimating Susceptibility

Why is it that many common solids have $|\chi_m| \ll 1$, but χ_e of order 1? Atoms contain a few valence electrons of charge q orbiting with radius of order a_0, the Bohr radius. To very roughly estimate electric and magnetic susceptibility, it's easiest to consider the extreme case where the field is so strong that the atom is about to fall apart.

The electrons are bound by an electric field $E_{\text{max}} \sim q/\epsilon_0 a_0^2$, so the atom will fall apart if the external field is much larger than this. And when the electron orbits are completely deformed, they will provide an electric dipole moment $p_{\max} \sim qa_0$. So the polarizability is of order

$$\alpha \sim \frac{p_{\max}}{E_{\max}} \sim \epsilon_0 a_0^3.$$

Then the electric susceptibility is

$$\chi_e = \frac{P}{\epsilon_0 E} \sim \frac{n\alpha}{\epsilon_0} \sim 1$$

because the number density of atoms is $n \sim a_0^{-3}$. So, in a completely typical insulator, χ_e is of order 1, which is indeed what we observe.

On the other hand, magnetic susceptibility is penalized by the fact that magnetic forces are suppressed by a factor of the electron speed v. The field strength at which the magnetic force is as strong as that of the binding electric field is $B_{\max} \sim E_{\max}/v$. At this point, the electrons provide the maximum possible magnetic moment by all orbiting in the same direction, so

$$\mu_{\max} \sim IA \sim (qv/a_0)(a_0^2) \sim qva_0.$$

To leading order in χ_m, the magnetic susceptibility is

$$\chi_m \approx \frac{\mu_0 M}{B} \sim \frac{\mu_0 \mu_{\max} n}{B_{\max}} \sim \epsilon_0 \mu_0 v^2 = \frac{v^2}{c^2} \sim \alpha^2 \sim 10^{-4}$$

where α is the fine structure constant, introduced in **P1**. And indeed, this estimate result matches experimental results. Magnetic susceptibility is typically small because relativistic effects for valence electrons are weak.

Example 5

An infinite solenoid with n turns per length and current I is filled with material with magnetic susceptibility χ_m. Find the magnetic field inside.

Solution

The magnetic field inside has contributions from the solenoid wire and the magnetization,

$$B = \mu_0(nI + M) = \mu_0 nI + \frac{\chi_m}{1+\chi_m} B.$$

Solving for B yields

$$B = (1+\chi_m)nI = \mu nI$$

which can be a significant enhancement if the material is a soft ferromagnet.

[2] **Problem 8.** A sphere of magnetic susceptibility χ_m is placed in a uniform field $\mathbf{B}_0$, and as a result develops a uniform magnetization $\mathbf{M}$. Find $\mathbf{M}$ and the field $\mathbf{B}$ inside the sphere. Then check the limiting cases of a superconductor and a soft ferromagnet. Do they make sense?

Idea 7: The H Field

Historically, magnetism was formulated in terms of the field

$$\mathbf{H} = \frac{1}{\mu_0}\mathbf{B} - \mathbf{M}.$$

Using this field gives a close analogy between electrostatics and magnetostatics. Note that

$$\nabla \times \mathbf{H} = \mathbf{J} - \mathbf{J}_b$$

so that $\mathbf{H}$ has no curl if there are no currents around besides the bound current. (This extra current is sometimes called "free" current $\mathbf{J}_f$.) In addition,

$$\nabla \cdot \mathbf{H} = -\nabla \cdot \mathbf{M}$$

which is analogous to how a polarization yields a charge density, $\nabla \cdot \mathbf{E} = -\nabla \cdot \mathbf{P}/\epsilon_0$. Therefore, any magnetostatic problem without free current can be mapped to an electrostatic one via

$$\mathbf{E} \leftrightarrow \mathbf{H}, \quad \mathbf{P} \leftrightarrow \mu_0 \mathbf{M}, \quad \epsilon_0 \leftrightarrow \mu_0, \quad \rho_b \leftrightarrow \rho_m$$

where $\rho_m = -\nabla \cdot \mathbf{M}$ is the "magnetic charge density".

This is the mathematical formalization of the idea of Gilbert dipoles, introduced in **E3**, which replace a true magnetic dipole with a pair of fictitious magnetic charges. At the time, we remarked that this gives you the correct magnetic field outside of a magnet, but not inside. The underlying reason is this analogy is actually computing $\mu_0\mathbf{H}$, not $\mathbf{B}$. To get the correct $\mathbf{B}$ within a magnetized material, we have to compute $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$. Heuristically, $\mu_0\mathbf{M}$ is the magnetic field due to dipole moments *right* at that location, while $\mu_0\mathbf{H}$ is the contribution from all other currents and magnetic dipole moments.

Magnetic quantities were originally defined to be simple in terms of $\mathbf{H}$, with

$$\mathbf{M} = \chi_m \mathbf{H}, \quad \mathbf{B} = \mu \mathbf{H}.$$

In fact, some even call $\mathbf{H}$ "the" magnetic field, but we'll just call it "$\mathbf{H}$". As always, $\mathbf{H}$ here is the *entire* $\mathbf{H}$ field, due to both the material itself and the field the material is in. Note that a soft ferromagnet can be thought of as a material within which $\mathbf{H}$ is approximately zero.

Remark: The History of H

If you learned physics in the United States, you might have found the preceding idea unfamiliar, because it has been systematically removed from the introductory curriculum.

The reason comes down to history. First, it's worth noting that the choice between covering $\mathbf{B}$ and $\mathbf{H}$ isn't obvious. Today we would say the $\mathbf{B}$ field is more fundamental, because it is what determines the Lorentz force on a charge, and the force and torque on a dipole moment, and thereby the reading on a magnetometer. But for most of history, one could just as well argue that it is $\mu_0\mathbf{H}$ that determines these things. After all, we can't embed a measuring device within a magnet. Doing so would require hollowing out a hole, which would remove $\mathbf{M}$ there, which would render $\mathbf{B}$ just equal to $\mu_0 \mathbf{H}$.

Therefore, choosing between $\mathbf{B}$ and $\mu_0 \mathbf{H}$ requires measuring some interaction where particles pass *through* a magnet. In the 1930s, this became possible with neutron scattering, since neutrons are electrically neutral but carry a magnetic dipole moment $\mathbf{m}$. As discussed here, the Nobel laureates Bloch and Schwinger disagreed on whether the interaction energy was $-\mathbf{m} \cdot \mathbf{B}$ or $-\mathbf{m} \cdot (\mu_0 \mathbf{H})$, which corresponds to treating $\mathbf{m}$ as Amperian or Gilbertian, respectively. The controversy remained open for decades, but was eventually settled by data.

At the same time, there was a growing consensus in the United States that magnetic poles were a "useless concept" which should be banished from teaching entirely. So, when the modern American physics curriculum was set in the 1960s, that's exactly what happened. For example, Halliday, Resnick, and Krane spends only a few pages covering magnetization, and half of them are spent admonishing the reader that magnetic poles don't exist. If you were educated in America, you probably only heard about them for a day in middle school.

Unfortunately, removing magnetic poles from the curriculum has some real costs. Experimentalists still think in terms of poles and $\mathbf{H}$, because it's harder to visualize how complicated currents source $\mathbf{B}$. Moreover, it's harder to do anything with $\mathbf{B}$ without vector calculus background. That's why some other countries' introductory physics courses put poles first. (But some teachers don't clearly explain how $\mathbf{B}$ and $\mu_0 \mathbf{H}$ differ, leading to confusion later.)

In this problem set, I'll tell you only what you need to know about $\mathbf{H}$ to solve theoretical problems. If you try to dive deeper into how experimentalists use it, you'll run into a lot more historical cruft. For instance, they tend to prefer the "Gaussian" system, where annoying factors of 4π are inserted into Maxwell's equations to make the Coulomb and Biot–Savart laws slightly simpler. Also, they use "cgs" (centimeter-gram-second) units, so all units need to be rescaled by some number of powers of 10. Worst of all, they have totally different units for $\mathbf{B}$ (Gauss), $\mathbf{H}$ (Oersted), and $\mathbf{M}$ (emu/cm^3), which are tricky to relate. I wouldn't recommend learning any of this unless you have to for your job.

Example 6

Using the $\mathbf{H}$ field, recompute the magnetic fields inside a uniformly magnetized sphere, and a rod magnetized parallel to and perpendicular to its axis.

Solution

In the first section, we found that inside these objects,

$$\mathbf{E} = \begin{cases} -\mathbf{P}/3\epsilon_0 & \text{sphere,} \\ -\mathbf{P}/2\epsilon_0 & \text{cylinder, perpendicular} \\ 0 & \text{cylinder, parallel} \end{cases}$$

Using the analogy above, we immediately conclude

$$\mathbf{H} = \begin{cases} -\mathbf{M}/3 & \text{sphere} \\ -\mathbf{M}/2 & \text{cylinder, perpendicular} \\ 0 & \text{cylinder, parallel} \end{cases}$$

The minus signs make sense because the "magnetic charge" accumulates on the side that **M** points to, and produces an **H** field in the opposite direction. Converting back to **B** yields

$$\mathbf{B} = \begin{cases} 2\mu_0\mathbf{M}/3 & \text{sphere} \\ \mu_0\mathbf{M}/2 & \text{cylinder, perpendicular} \\ \mu_0\mathbf{M} & \text{cylinder, parallel} \end{cases}$$

which precisely matches what we found in example 4 and problem 6, with much less effort.

Remark: Demagnetizing Fields

At a given point in an isolated magnet, $\mu_0\mathbf{H}$ is the part of **B** due to the rest of the magnet. However, the above example shows that **H** always points *against* the direction of **M**, so a permanent magnet is always trying to demagnetize itself! Similarly, electrically polarized materials carry an internal "depolarization" field. This is why, in the absence of external fields, the vast majority of materials have zero polarization and magnetization.

Before the advent of very effective "hard" ferromagnets, like neodymium magnets, magnets had to be shaped to avoid this effect, e.g. by making them into long bars or horseshoes. Even so, the demagnetization effect would make the field produced by the magnet a little less than you would expect. For an average-shaped bar magnet made in the 1950s, the magnetic poles are effectively not at the ends, but rather 10% to 20% closer together.

This bit of historical trivia is irrelevant today, but strangely, it has stuck around in the Indian physics curriculum. Every Indian physics textbook demands its students memorize the ratio of the "magnetic length" and "geometric length" of a bar magnet, as if it were a fundamental constant of nature rather than an obsolete rule of thumb. But different Indian books don't even agree on what the ratio is, with HC Verma giving 84%, various JEE prep sources stating 4/5, 5/6, or 7/8, and none whatsoever explaining where the number comes from.

Idea 8: Magnetic Energy, Force, and Torque

The appropriate magnetic energy density depends on the material.

- If the magnetization is permanent, as in a hard ferromagnet, we should use the same potential energy introduced in **E4**. Specifically, the potential energy density is $-\mathbf{M} \cdot \mathbf{B}$.
- For all other materials, where the magnetization is induced by the presence of other fields, the total energy density is $B^2/2\mu$.

If you use one of these formulas where the other applies, you'll typically be off by a factor of 2.

Regardless of how the magnetization arises, the resulting force and torque in an external field $\mathbf{B}$ are given by the formulas introduced in **E4**. Specifically, the torque density is $\mathbf{M} \times \mathbf{B}$, and the force density is $\nabla(\mathbf{M} \cdot \mathbf{B})$, where the ∇ only acts on $\mathbf{B}$. Alternatively, if you're using the magnetic pole trick, a magnetic charge q_m feels a force $q_m\mathbf{B}$.

[3] **Problem 9.** A version of the method of images works for magnetic materials. Let's suppose there is vacuum at $z > 0$, and a material of relative permeability μ_r at $z < 0$. When using the method of images, we only care about the field at $z > 0$, where $\mathbf{B}$ and $\mathbf{H}$ are proportional. So we can directly use the analogy between $\mathbf{H}$ and $\mathbf{E}$.

(a) Suppose a magnetic charge q_m is a distance d above the plane. By recycling your answer to problem 3, find the magnetic charge q'_m of the image. What does it become if the material is a superconductor, or a soft ferromagnet?

(b) Of course, magnetic charges don't actually exist, so let's instead suppose a permanent magnetic dipole moment $\mathbf{m}$ was a distance d above the plane, with $\mathbf{m}$ pointing towards the plane. Characterize the image dipole, and find the force on the real dipole.

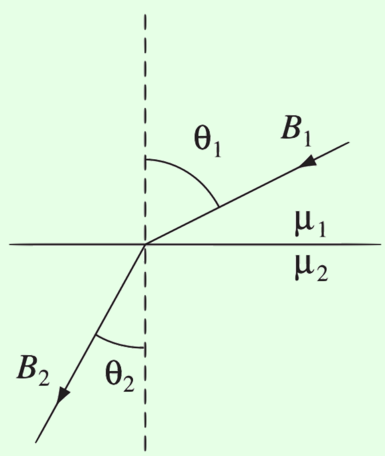
(c) To be even more concrete, consider a very long permanent magnet of cross-sectional area A and uniform magnetization M along its length. When one end of the magnet is placed flat against an iron plate, what is the force between them?

[2] **Problem 10.** AuPhO 2019, problem 13. A neat explanation of how a fridge magnet works; for this problem it will be useful to consult the answer sheet.

[5] **Problem 11.** Physics Cup 2024, problem 3. This relatively straightforward problem reviews almost everything we've covered so far.

Example 7: Griffiths 6.27

How does a magnetic field line bend when it passes from one medium to another?



Solution

We say the field lines "bend" because of Gauss's law for magnetism: they can't start or end, so each one has to keep on going. Let's suppose the figure above is drawn in the xz plane. Applying Gauss's law for magnetism in a small pillbox spanning the interface gives

$$B_1^z = B_2^z.$$

On the other hand, since $\nabla \times \mathbf{H}$ is zero (assuming no additional, "free" current is around), considering an Amperian loop spanning the interface gives

$$H_1^x = H_2^x.$$

Combining these results gives

$$\frac{\tan\theta_2}{\tan\theta_1} = \frac{\mu_2}{\mu_1}.$$

In other words, when a field line enters a medium with higher μ, it bends away from the normal, and when it enters a medium with lower μ, it bends towards the normal.

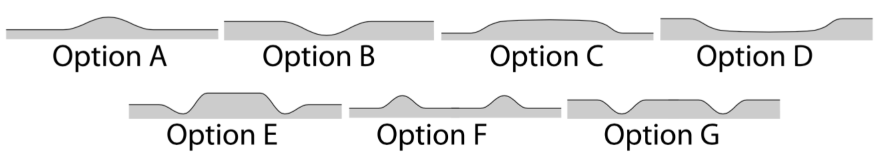
This statement has two limiting cases which will be important later.

- A magnetic field line can't enter a superconductor ($\mu_2 = 0$) at all, so field lines approaching a superconductor bend away, to become tangent to them ($\theta_1 \to 90^\circ$).
- A magnetic field line entering a soft ferromagnet ($\mu_2 \to \infty$) bends towards it to enter along the normal direction ($\theta_1 \to 0^\circ$), similar to how electric field lines approach conductors. It's also possible for θ_1 to be nonzero if $\theta_2 \to 90^\circ$, but we won't see any examples of this.

You can see both of these behaviors in the limiting cases of problem 9. In general, we conclude that magnetic field lines are "attracted" to regions of higher μ, which makes sense because it helps minimize the energy. Soft ferromagnets tend to keep magnetic field lines within themselves, which is why they're used in transformers.

[2] **Problem 12** (IPhO 2012 Experiment)**.** Water is a diamagnetic substance. A powerful cylindrical magnet with field B is placed below the water surface.

(a) Which of the following shows the resulting shape of the water surface?



The magnet is roughly 2/3 as wide as each of these sketches.

(b) Let ρ be the density of the water. If the maximum change in height of the water surface has magnitude h, find an approximate expression for the magnetic susceptibility χ_m of water.

For a very closely related, but more extreme problem, see EuPhO 2018, problem 2.

[3] **Problem 13.** EFPhO 2004, problem 6. A cute exercise with permanent magnets.

[5] **Problem 14.** IPhO 2022, problem 1. A series of exercises on spherical magnets, which uses almost everything covered in this section.

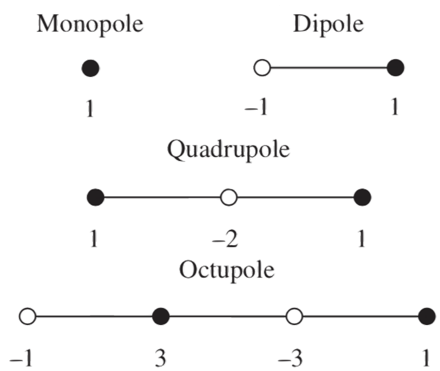
[4] **Problem 15.** Physics Cup 2012, problem 2. If you only know what's taught in American introductory courses, this problem is basically impossible. If you only know what's stated explicitly in Griffiths, it's very hard. But if you've internalized the intuition of the above examples, and the relevant section of **E5** on superconductors, it should be relatively approachable.

[5] **Problem 16.** Physics Cup 2018, problem 3. A substantially tougher problem which requires solving some differential equations. I recommend starting from the fifth hint.

3 Multipoles

In this section, we explore some of the physics of dipoles and higher multipoles.

[3] **Problem 17** (Purcell 10.27)**.** Two monopoles of opposite sign form a dipole, two dipoles of opposite sign for a quadrupole, and so on. Hence we can construct arbitrarily high multipoles using the rows of Pascal's triangle.



The field of a dipole falls as $1/r^3$, a quadrupole as $1/r^4$, and an octupole as $1/r^5$.

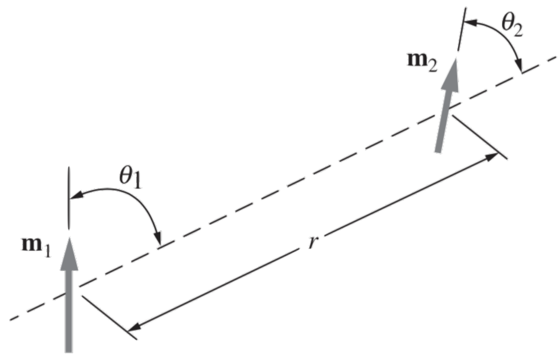
(a) To warm up, verify explicitly that the quadrupole field along the axis of the quadrupole starts at $1/r^4$, i.e. that all lower terms cancel.

(b) Prove that this cancellation occurs for general multipoles along their axis.

(c) [A] The magnitude and orientation of a dipole is specified by a vector, with three components. How many numbers are necessary to specify the magnitude and orientation of a quadrupole? (The linear quadrupoles here are just a special case of a general quadrupole.) Try to generalize to arbitrary multipoles.

To learn how to decompose an arbitrary charge distribution into multipoles, see section 3.4 of Griffiths.

[3] **Problem 18** (Purcell 11.23)**.** Consider two magnetic dipoles with coplanar dipole moments.



Show that the associated potential energy is

$$U = \frac{\mu_0 m_1 m_2}{4\pi r^3} \left(\sin\theta_1 \sin\theta_2 - 2\cos\theta_1 \cos\theta_2 \right).$$

For what orientations is this potential energy maximized or minimized?

[2] **Problem 19** (Purcell 11.36)**.** Three magnetic compasses are placed at the corners of a horizontal equilateral triangle. As in any ordinary compass, each compass needle is a magnetic dipole constrained to rotate in a horizontal plane. The Earth's magnetic field has been shielded. What orientation will the compass needles eventually assume? Does your result also hold for regular N-gons?

[3] **Problem 20.** Some questions about forces between dipoles and other multipoles.

(a) Above, you've shown that the force between permanent magnetic dipoles falls off as $1/r^4$. How about two permanent electric dipoles?

(b) How about a permanent dipole and a permanent quadrupole?

(c) How about two permanent quadrupoles?

(d) Now consider an ion and a neutral atom. The electric field of the ion polarizes the atom; the field of that induced dipole then reacts on the ion. Show that the resulting force is attractive and falls as $1/r^5$.

4 Electromagnetic Waves in Matter

In this section, you will work out some of the theory of electromagnetic waves in matter.

Idea 9

In the absence of any free charge or current, Maxwell's equations in matter are identical to Maxwell's equations in vacuum, except that ϵ_0 and μ_0 are related by ϵ and μ, so the waves propagate with speed $1/\sqrt{\epsilon\mu} = c/n$, with $E = (c/n)B$.

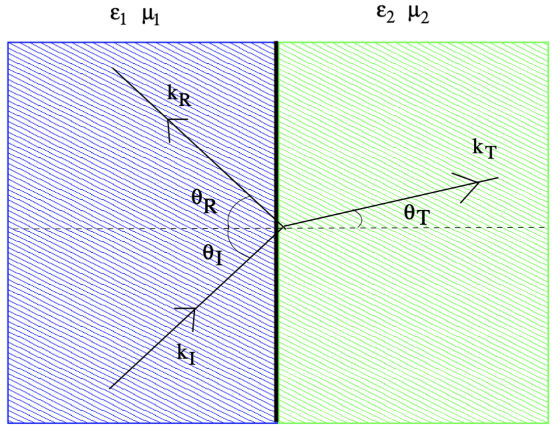
[5] **Problem 21.** Suppose the regions $x < 0$ and $x > 0$ are filled with material with permittivities ϵ_1 and ϵ_2, both with permeability μ_0. (This is typical; if you don't count permanent magnets, most objects have permeability about μ_0.) We send in an incident wave from the left with electric field

$$\mathbf{E}_i e^{i(\mathbf{k}_i \cdot \mathbf{x} - \omega_i t)}.$$

The wave will be both transmitted and reflected at the interface, so the total electric field is

$$\mathbf{E} = \begin{cases} \mathbf{E}_i e^{i(\mathbf{k}_i \cdot \mathbf{x} - \omega_i t)} + \mathbf{E}_r e^{i(\mathbf{k}_r \cdot \mathbf{x} - \omega_r t)} & x < 0, \\ \mathbf{E}_t e^{i(\mathbf{k}_t \cdot \mathbf{x} - \omega_t t)} & x > 0. \end{cases}$$

The angles with the normal are θ_i, θ_r, and θ_t as shown.



(a) Argue that by continuity of the field at the boundary,

$$\omega_i = \omega_r = \omega_t.$$

(b) Suppose the y-axis is oriented so that $\mathbf{k}_i \cdot \hat{\mathbf{y}} = 0$. Argue that

$$\mathbf{k}_r \cdot \hat{\mathbf{y}} = \mathbf{k}_t \cdot \hat{\mathbf{y}} = 0, \quad \mathbf{k}_i \cdot \hat{\mathbf{z}} = \mathbf{k}_r \cdot \hat{\mathbf{z}} = \mathbf{k}_t \cdot \hat{\mathbf{z}}.$$

From these conditions, derive the laws of reflection and refraction,

$$\theta_i = \theta_r, \quad n_1 \sin\theta_i = n_2 \sin\theta_t.$$

Note that neither this part nor the previous part require Maxwell's equations; they hold for *all* kinds of waves as long as we define $n_i \propto 1/v_i$.

(c) Argue that at the boundary, $\mathbf{E}_\parallel$ and $B_\perp$ must be continuous in general. In this case, because both sides have the same permittivity μ_0, there is no bound current, so $\mathbf{B}_\parallel$ is also continuous.

(d) Now suppose the electric fields are polarized along the $\mathbf{y}$ axis, so $\mathbf{E}_i$, $\mathbf{E}_r$, and $\mathbf{E}_t$ are all parallel to the y-axis. Then continuity of $\mathbf{E}_\parallel$ gives

$$E_i + E_r = E_t.$$

Using continuity of $\mathbf{B}_\parallel$, show that

$$\frac{E_r}{E_i} = \frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t}, \quad \frac{E_t}{E_i} = \frac{2n_1 \cos\theta_i}{n_1 \cos\theta_i + n_2 \cos\theta_t}.$$

These are the Fresnel equations for normal polarized light. (Hint: this is a bit messy, so you can warm up with the case $\theta_i = 0$.)

(e) If $n_1 > n_2$, then total internal reflection occurs when

$$\sin\theta_i > \frac{n_2}{n_1}$$

and the wave is totally reflected. Nonetheless, E_t is nonzero in this regime. To make sense of this, show that $\mathbf{k}_t \cdot \mathbf{x}$ is imaginary in this regime, indicating that the "transmitted" wave does not propagate in the region $x > 0$, but rather exponentially decays.

[5] **Problem 22.** In most common materials, $\mu \approx \mu_0$ while ϵ depends on frequency. We'll investigate the origin of this frequency dependence below.

(a) Model an electron in an atom as a mass m with charge q attached to a spring, with natural angular frequency ω_0 and a damping force $-m\gamma\mathbf{v}$, in an electric field $\mathbf{E}_0 e^{-i\omega t}$. Write down the equation of motion for the electron.

(b) The atomic polarizability is $\mathbf{p} = \alpha\mathbf{E}$. Show that

$$\alpha = \frac{q^2/m}{-\omega^2 + \omega_0^2 - i\gamma\omega}.$$

(c) For a gas with small number density n, the Clausius–Mossotti formula reduces to

$$\epsilon = \epsilon_0 + n\alpha.$$

Therefore, the permittivity is generally a complex number. The wavevector and angular frequency are related by $k^2 = \mu\epsilon\omega^2$. Explain why the fact that ϵ is complex indicates that waves can be absorbed.

From this point on, you may approximate γ as small.

(d) What value of ω maximizes the absorption rate of the electromagnetic waves? Roughly how many wavelengths does such a wave propagate before being mostly absorbed?

(e) What value of ω maximizes the speed of the electromagnetic waves, and what is that speed?

(f) Transparent objects such as glass can be modeled as having a very high resonant frequency, much higher than that of visible light. Does blue light or red light refract more when passing from air to glass?

The intuitive reason that these electrons can affect the propagation speed of light is because they emit secondary electromagnetic waves that are out of phase with the original wave; this "pushes" the phase of the composite wave forward or backward, affecting the phase velocity. A nice explanation of this can be found in chapter I.31 of the Feynman lectures.

[5] **Problem 23.** IPhO 2002, problem 1. A neat application of electromagnetic waves in matter.

[5] **Problem 24.** APhO 2007, problem 2. A problem on an exotic negative index of refraction.

Remark

Above, we considered the response of a medium composed of atoms, obeying $p = \alpha E$. However, this relation is just an approximation, like Hooke's law. For larger electric fields, higher order terms are necessary,

$$p = \alpha E + \alpha' E^2 + \ldots$$

which lead to strange effects, studied in the field of nonlinear optics. For example, suppose we send in light of angular frequency ω. Then

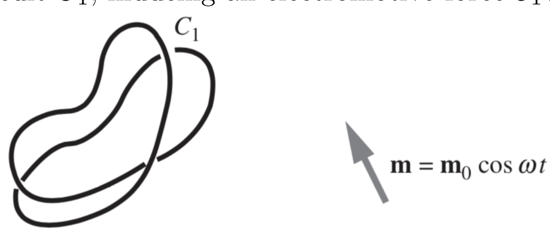
$$E^2 \propto \cos^2(\omega t) = \frac{1 + \cos(2\omega t)}{2}.$$

That means that a nonlinear medium can respond to light at angular frequency ω by oscillating, and hence emitting light, at angular frequency 2ω. This phenomenon is called frequency doubling, or second-harmonic generation, and converts red light to blue. Similarly, for a cubic nonlinearity, you can use trigonometric identities to show that frequency tripling can occur.

5 Electromagnetic Systems

In this section we'll consider problems that use everything we've covered, with a focus on technological applications and systems with multiple moving parts.

[2] **Problem 25** (Purcell 11.19)**.** A magnetic dipole $\mathbf{m}$ oscillates so that $\mathbf{m}(t) = \mathbf{m}_0 \cos \omega t$. Some of its flux links the nearby circuit C_1, inducing an electromotive force $\mathcal{E}_1 \sin \omega t$.



If a current I_1 flowed in C_1, then the resulting field at the location of the dipole would be $\mathbf{B}_1$. Show that $\mathcal{E}_1 = (\omega/I_1)\mathbf{B}_1 \cdot \mathbf{m}_0$. (Hint: recall the results involving mutual inductance in **E5**.)

[3] **Problem 26.** EFPhO 2007, problem 3. A problem on focusing particles with electric fields.

[4] **Problem 27.** IPhO 2004, problem 3. A practical problem which also reviews damped/driven oscillations.

[4] **Problem 28.** EFPhO 2014, problem 1. A challenging problem about a complex nonlinear circuit.

[5] **Problem 29.** Physics Cup 2020, problem 1. (Unfortunately, the word "dielectric" has two common distinct meanings, and you'll have to tell from context which is meant. When this problem states that the rod is "dielectric", it means that the rod always has zero charge and current density everywhere, i.e. it is an insulator with zero electric susceptibility. Alternatively, you can suppose the rod might have some electric susceptibility, but it's too thin to have an effect on the dynamics of the metal balls.)

Electromagnetism VIII: Synthesis

Electromagnetism in matter is covered in chapters 10 and 11 of Purcell. For more on dielectrics, see chapters II-10 and II-11 of the Feynman lectures. Electromagnetism in matter is covered in greater detail in chapters 4, 6, and 9 of Griffiths, and chapters I-31 and II-32 through II-37 of the Feynman lectures. For an enlightening overview of the history of magnetism, see chapter 1 of *Magnetism and Magnetic Materials* by Coey. There is a total of **97** points.

1 Polarization

In **E2**, we introduced the basics of dielectrics. To review: when a dielectric is placed in an electric field, dipoles inside align with the field, reducing the field value. In very symmetrical situations, the field is simply reduced by a factor of the dielectric constant $\kappa = \epsilon/\epsilon_0$. The total energy density within a dielectric is $\epsilon E^2/2$. This section is about problems which require more than these few facts. To answer them, we need to think about the charge bound to the dielectric itself.

Idea 1: Bound Charge

The polarization $\mathbf{P}$ of a material is its electric dipole moment per unit volume. It corresponds to a "bound" charge density

$$\rho_b = -\nabla \cdot \mathbf{P}$$

within the dielectric, as well a bound surface charge density

$$\sigma_b = \mathbf{P} \cdot \hat{\mathbf{n}}$$

on its surface.

Example 1

Find the electric field of a sphere with uniform polarization $\mathbf{P}$ and radius R.

Solution

There is no bound charge density inside the sphere, but a bound surface charge density

$$\sigma_b = \mathbf{P} \cdot \hat{\mathbf{r}} = P \cos\theta$$

on its surface. We could apply Coulomb's law to this charge density, but an easier method is to recall that polarization just means an internal displacement of charge. This surface charge density precisely corresponds to having two uniformly charged balls of total charge $\pm Q$ displaced by a tiny amount $\mathbf{d}$ so that $Q\mathbf{d} = (4\pi R^3/3)\mathbf{P}$.

By the shell theorem, the resulting field inside is uniform, and points against $\mathbf{P}$,

$$\mathbf{E} = -\frac{\mathbf{P}}{3\epsilon_0},$$

and the field outside is exactly a dipole field, with dipole moment $\mathbf{p} = (4\pi R^3/3)\mathbf{P}$.

[2] **Problem 1.** An infinite cylindrical rod of radius R has a uniform polarization $\mathbf{P}$.

(a) If $\mathbf{P}$ is perpendicular to the rod's axis, describe $\mathbf{E}$ outside, and find the value of $\mathbf{E}$ inside.

(b) If $\mathbf{P}$ is parallel to the rod's axis, find the electric field everywhere.

Solution. (a) The resulting bound charge is equivalent to having two uniform cylinders of total linear charge density λ separated by $\mathbf{d}$ where $\lambda\mathbf{d} = (\pi R^2)\mathbf{P}$.

Outside the rod, this is equivalent to two lines of charge density $\pm\lambda$ separated by d. As for inside, note that for a single cylinder with charge density ρ, Gauss's law tells us that the field is $\mathbf{E} = \rho\mathbf{r}/(2\epsilon_0)$. Thus, superposing the positive and negative cylinders, the field inside the rod is

$$\mathbf{E} = -\frac{\rho\mathbf{s}}{2\epsilon_0} = -\frac{\mathbf{P}}{2\epsilon_0}$$

which is uniform. (There's a little paradox here: why doesn't the limit $R \to \infty$ of this answer coincide with the $R \to \infty$ limit of the result of example 1? As was discussed in a related context in **E1**, the issue is that when a charge configuration is infinite, the answer is ambiguous and depends on boundary conditions. In this problem, we are implicitly adopting boundary conditions that yield a field with cylindrical symmetry.)

(b) In this case there's no bound charge anywhere. (You can imagine charge displaced a bit from infinity in one direction to infinity in the other direction.) So there is no electric field at all.

Now let's think about how polarization arises in the first place.

Idea 2: Electric Susceptibility

A small number of insulators are "ferroelectrics", whose crystal structure lets them maintain a preferred polarization $\mathbf{P}$ in the absence of external electric fields. (You can suppose that the preceding example and problem were implicitly about ferroelectrics.) But the vast majority of insulators are dielectrics, whose polarization is related to the total electric field by

$$\mathbf{P} = \epsilon_0\chi_e\mathbf{E}, \quad \epsilon = \epsilon_0(1+\chi_e) = \epsilon_0\kappa$$

where χ_e is the electric susceptibility and κ is the dielectric constant. The susceptibility is nonnegative, except in some very exotic materials. The key difficulty is that above, $\mathbf{E}$ is the *total* electric field, including that due to the bound charge, which in turn depends on $\mathbf{P}$.

Example 2

A point charge q is inside a dielectric sphere of radius R with dielectric constant κ. Find the electric field and charge density everywhere.

Solution

This is one of the simple symmetric cases where the electric field in the dielectric is simply reduced by a factor of κ,

$$\mathbf{E} = \frac{q\,\hat{\mathbf{r}}}{4\pi\epsilon_0 r^2} \times \begin{cases} 1/\kappa & r < R \\ 1 & r > R \end{cases}.$$

Inside the sphere, this corresponds to an electric polarization

$$\mathbf{P} = \frac{q}{4\pi r^2} \frac{\chi_e}{\kappa} \,\hat{\mathbf{r}}.$$

To check that this solution is actually right, we need to ensure the original point charge q, plus the bound charge, indeed generates the claimed electric field.

The divergence of $\mathbf{P}$ is zero everywhere besides the origin, where negative bound charge piles up to cancel some of the charge q. The charge at the origin is thus

$$q - q_b = q\left(1 - \frac{\chi_e}{\kappa}\right) = q\left(1 - \frac{\kappa - 1}{\kappa}\right) = \frac{q}{\kappa}$$

which is consistent with Gauss's law for $\mathbf{E}$ there. At the surface of the sphere, there is a positive bound surface charge density

$$\sigma_b = \frac{q}{4\pi R^2} \frac{\chi_e}{\kappa}$$

which cancels the negative bound charge at the origin. Thus, by the shell theorem, the electric field outside the sphere is indeed that of the point charge q alone.

Example 3

A dielectric sphere of radius R and dielectric constant κ is placed in a uniform field $\mathbf{E}_0$, and as a result develops a uniform polarization $\mathbf{P}$. Find $\mathbf{P}$ and the field inside the sphere.

Solution

In example 1, we found the electric field due to the polarized sphere itself, which we'll call $\mathbf{E}_p$. Here, we must remember that the polarization is produced in response to the *total* electric field inside the sphere,

$$\mathbf{P} = \chi_e \epsilon_0 \mathbf{E}, \quad \mathbf{E} = \mathbf{E}_0 + \mathbf{E}_p.$$

Using our previous result for $\mathbf{E}_p$ and solving the system, we find

$$\mathbf{E} = \frac{3}{\kappa + 2}\,\mathbf{E}_0, \quad \mathbf{P} = 3\,\frac{\kappa - 1}{\kappa + 2}\,\epsilon_0 \mathbf{E}_0.$$

The polarizability α of each atom is defined as the dipole moment per applied field,

$$\mathbf{p} = \alpha \mathbf{E}_0$$

so we have shown above that

$$\alpha = \frac{3\epsilon_0}{n} \frac{\kappa - 1}{\kappa + 2}$$

where n is the number density of atoms. This is the Clausius–Mossotti formula; it relates the macroscopically measurable parameter κ to the microscopic parameter α.

[2] **Problem 2** (Purcell 10.10)**.** Assume that the uniform field $\mathbf{E}_0$ that causes the electric field in example 2 is produced by large capacitor plates very far away. The field lines tangent to the sphere

hit each of the distant capacitor plates in a circle of radius r. Find r in terms of R and κ.

Solution. The field lines are tangent at the widest part of the sphere. Consider a Gaussian surface which is bounded by a distant capacitor plate, a horizontal slice through the middle of the sphere, and all of these field lines. Using the results of example 3, the charge contained inside is

$$Q = \pi R^2 \, \frac{3(\kappa - 1)}{\kappa + 2} \, \epsilon_0 E_0 - \pi r^2 \epsilon E_0$$

where the first term is from cutting the polarized sphere. The flux through this surface is

$$\Phi = \frac{3}{\kappa + 2} \, \pi R^2 E_0.$$

Applying Gauss's law, we have

$$r = \sqrt{\frac{3\kappa}{\kappa + 2}} \, R.$$

Idea 3

The "free" charge density ρ_f is the part of the charge density that isn't bound, so that

$$\rho = \rho_b + \rho_f.$$

If we take the divergence of $\mathbf{P} = \epsilon_0 \chi_e \mathbf{E}$, we get $\rho_b = -\chi_e \rho$ inside a uniform dielectric, so

$$\rho = \rho_f / \kappa.$$

That is, a uniform dielectric "screens" charges embedded within it, reducing it by a factor of κ. That's exactly what we saw in example 2, and it also tells us that a conductor can be viewed as a dielectric with $\kappa \to \infty$, because conductors completely expel electric fields.

The difficulty in dealing with dielectrics is when κ changes in space, such as at the boundary of a dielectric, where a bound surface charge density σ_b can appear. If all the free charges in a problem are outside of uniform dielectrics, bound charges only appear on their surfaces.

[3] **Problem 3.** A version of the method of images, introduced in **E2**, works for dielectrics. Let's suppose there is vacuum at $z > 0$, a dielectric κ at $z < 0$, and a point charge q a distance d above the plane $z = 0$. We need to find the surface bound charge density σ_b that appears on the plane.

(a) Let E_0^z be the z-component of the electric field due to the point charge alone. At a given point just below the plane $z = 0$, find E^z in terms of E_0^z and σ_b.

(b) Use this result to solve for σ_b in terms of E_0^z and κ.

(c) Your answer will be exactly the same as what one gets for a conductor at $z < 0$, multiplied by a κ-dependent constant. Using this information, characterize the image charge and find the force on the real charge.

Solution. (a) By an elementary application of Gauss's law, the result is

$$E^z = E_0^z - \frac{\sigma_b}{2\epsilon_0}.$$

(b) By the definition of χ_e, we know that just under the plane,

$$\sigma_b = P^z = \epsilon_0 \chi_e E^z.$$

Combining this with the result of part (a) and solving for σ_b gives

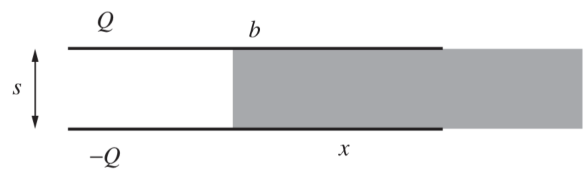
$$\sigma_b = \frac{\chi_e}{\chi_e + 2}(2\epsilon_0 E_0^z) = \frac{\kappa - 1}{\kappa + 1}(2\epsilon_0 E_0^z).$$

(c) A conductor corresponds to the limit $\kappa \to \infty$, where we simply have $\sigma_b = 2\epsilon_0 E_0^z$, and the electric field inside the conductor vanishes. Evidently, for a general dielectric this result is multiplied by $(\kappa-1)/(\kappa+1)$. We therefore conclude that the image charge is $-q(\kappa-1)/(\kappa+1)$, a distance d below the plane. The force on the real charge is given by Coulomb's law,

$$F = \frac{q^2}{16\pi\epsilon_0 d^2}\frac{\kappa - 1}{\kappa + 1}$$

and is directed towards the dielectric.

[3] **Problem 4** (Purcell 10.2)**.** A rectangular capacitor with side lengths a and b has separation $s \ll a, b$. It is partially filled with a dielectric with dielectric constant κ. The overlap distance is x.



The capacitor is isolated and has constant charge Q.

(a) What is the energy stored in the system?

(b) Using the result of part (a), what is the force on the dielectric? Which direction does it point?

(c) Is your answer to part (b) affected by the presence of fringe fields near the interface?

Solution. (a) The system consists of two capacitors in parallel, with capacitances $C_1 = \epsilon_0(b - x)a/s$ and $C_2 = \kappa\epsilon_0 xa/s$. Thus,

$$C = \epsilon_0(a/s)((\kappa - 1)x + b)$$

which gives

$$U = \frac{Q^2}{2C} = \frac{Q^2 s}{2\epsilon_0 a(b + (\kappa - 1)x)}.$$

(b) Note that

$$F = -\frac{dU}{dx} = \frac{Q^2 s(\kappa - 1)}{2\epsilon_0 a(b + (\kappa - 1)x)^2}.$$

The sign is positive, so it points in direction of increasing x, so the slab is pulled in.

(c) Fringe fields don't change the result of part (b). The presence of fringe fields does change the energy found in part (a), but this has essentially no effect on the *derivative* of the energy, because shifting the dielectric just shifts the fringe field over essentially unchanged.

Of course, from a force perspective, *all* of the force is due to the fringe fields, because those are the only fields with a horizontal component; this paper gives such a calculation. The fact that you can get the same answer, by using an energy-based derivation that doesn't depend on the fringe fields, or by a force-based derivation that relies entirely on the fringe fields, is just another example of conservation of energy giving us nontrivial information.

[3] **Problem 5** (Griffiths 4.28)**.** Two long coaxial cylindrical metal tubes of inner radius a and outer radius b stand vertically in a tank of dielectric oil, with susceptibility χ_e and mass density ρ. The inner one is maintained at potential V, and the outer one is grounded. To what height h does the oil rise in the space between the tubes?

Solution. The field in the region with no oil is $E = \frac{\lambda}{2\pi\epsilon_0 r}$, and with the oil is $E' = \frac{\lambda'}{2\pi\epsilon r}$ where λ' is the free charge density. Thus,

$$V = \frac{\lambda}{2\pi\epsilon_0}\log(b/a),$$

and equating with the oil part, we get that $\lambda' = \kappa\lambda$, as expected. Now, the total charge on this effective capacitor is

$$Q = \lambda' h + \lambda(\ell - h) = \lambda(\chi_e h + \ell),$$

so

$$C = \frac{Q}{V} = 2\pi\epsilon_0 \frac{\chi_e h + \ell}{\log(b/a)}.$$

We know the net force is $\frac{1}{2}V^2(dC/dh)$ (note that there is not a minus sign here because of the work done by the battery, as explained in a problem in **E2**). The gravitational force is $\rho\pi g h(b^2 - a^2)$, so equating and solving for h gives

$$h = \frac{\epsilon_0 \chi_e V^2}{\rho(b^2 - a^2) g \log(b/a)}.$$

2 Magnetization

Idea 4

As discussed in **E5**, materials contain two kinds of magnetic dipole moments: the "orbital" part, due to moving electrons, and the "spin" part, due to the electrons' intrinsic magnetic moments. For most materials, in the absence of external magnetic fields, these dipole moments point in random directions, and thus sum to zero on average.

When such a material is placed in a magnetic field, two things happen at once:

- The spins partially align with the field, producing a net dipole moment along $\mathbf{B}$.
- The orbits are affected by the changing field in accordance with Lenz's law, and thus produce a net dipole moment *against* $\mathbf{B}$.

These effects are often comparable in size. If the first is more important, the material is paramagnetic, and if the second is more important, it is diamagnetic.

This can be a bit tricky to remember, because it seems opposite to the definition of a dielectric, where the internal electric dipoles try to align *with* the external field. The reason it makes sense is that inside an electric dipole, the electric field points against the dipole moment, while inside a magnetic dipole, the magnetic field points with the dipole moment, as discussed in **E3**. So, both dielectrics and diamagnets try to reduce the applied field within them.

Idea 5: Bound Currents

The magnetization $\mathbf{M}$ of a material is its magnetic dipole moment per unit volume. It corresponds to a bound current density

$$\mathbf{J}_b = \nabla \times \mathbf{M}$$

as well as a surface bound current density

$$\mathbf{K}_b = \mathbf{M} \times \hat{\mathbf{n}}$$

on its surface.

Example 4

Find the magnetic field of a sphere with uniform magnetization $\mathbf{M}$ and radius R.

Solution

In this case $\mathbf{J}_b$ is zero in the sphere, while at the sphere's surface,

$$\mathbf{K}_b = \mathbf{M} \times \hat{\mathbf{r}} = M \sin\theta \, \hat{\boldsymbol{\phi}}$$

where we worked in spherical coordinates and aligned $\mathbf{M}$ with the z-axis. However, this is precisely the current density of a rotating, uniformly charged sphere, as we discussed in **E3**. Scaling the constants appropriately, we find that inside,

$$\mathbf{B} = \frac{2}{3} \mu_0 \mathbf{M}$$

which should be compared with example 3. Outside, the field is exactly a magnetic dipole field, with $\mathbf{m} = (4\pi R^3/3)\mathbf{M}$.

[2] **Problem 6.** An infinite cylindrical rod of radius R has a fixed, uniform magnetization $\mathbf{M}$.

(a) If $\mathbf{M}$ is parallel to the rod's axis, find the magnetic field everywhere.

(b) If $\mathbf{M}$ is perpendicular to the rod's axis, describe $\mathbf{B}$ outside, and find the value of $\mathbf{B}$ inside.

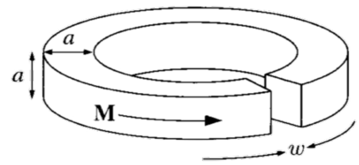
Solution. Let $\hat{\mathbf{z}}$ point along the axis of the rods.

(a) The surface bound current density is $\mathbf{K}_b = M \hat{\boldsymbol{\theta}}$, which is simply that of an infinite solenoid. So the magnetic field is $\mu_0 \mathbf{M}$ inside the rod, and zero outside.

(b) Let's say that $\mathbf{M}$ is parallel to $\hat{\mathbf{x}}$. Then in cylindrical coordinates, we have $\mathbf{K}_b = M \sin\theta \, \hat{\mathbf{z}}$. This current density is equivalent to superposing two cylinders carrying uniform current density $\pm J \, \hat{\mathbf{z}}$, separated by d along the $\hat{\mathbf{y}}$ direction, where $Jd = M$.

Outside the rod, the cylinders can be replaced with wires carrying current $I = \pi R^2 J$, and the corresponding magnetic field was found in a problem in **E3**. Inside the rod, superposing the magnetic fields of the cylinders yields $\mathbf{B} = \mu_0 \mathbf{M}/2$.

[2] **Problem 7** (Griffiths 6.10)**.** A rod of length L and square cross section of side a is given a uniform longitudinal magnetization $\mathbf{M}$ and then bent into a circle with a narrow gap of width w.



Find the magnetic field at the center of the gap, assuming $w \ll a \ll L$.

Solution. First, suppose there was no gap. Before the iron rod was bent, its uniform magnetization $\mathbf{M}$ corresponded to a bound current density $\mathbf{K}_b = \mathbf{M}$ everywhere along its surface, directed circumferentially. Since $a \ll L$, this remains approximately true after bending the rod. (A small volume bound current density $\mathbf{J}_b$ appears, but we neglect this.)

Therefore, the current density is the same as that of a toroidal solenoid with current I and n turns per length, where $M = In$. The field inside is therefore $\mu_0 M$, directed in the $\hat{\boldsymbol{\theta}}$ direction and zero everywhere outside the rod.

Now let's account for the gap. Adding the gap is equivalent to superposing an opposite magnetization at the gap. Since $w \ll a$, we can treat it as an $a \times a$ square current loop, with current $I = mw$. By the Biot–Savart law, the field due to such a loop at the center is

$$B_{\text{loop}} = \frac{2\sqrt{2}\mu_0 M w}{\pi a}.$$

Combining the two gives a total field of

$$B = \mu_0 M \left(1 - \frac{2\sqrt{2} w}{\pi a}\right).$$

Idea 6: Magnetic Susceptibility

Permanent magnets, such as the ones on your refrigerator, or the strong neodymium magnets popular in science toys, are made of "hard" ferromagnets. These are materials whose crystal structure lets them maintain a fixed magnetization $\mathbf{M}$ in the absence of external magnetic fields. (The preceding example and problems were implicitly about hard ferromagnets.)

In most other materials, the magnetization is related to the magnetic field in the material by

$$\mathbf{M} = \frac{1}{\mu_0} \frac{\chi_m}{1 + \chi_m} \mathbf{B}, \quad \mu = \mu_0(1 + \chi_m) = \mu_0 \mu_r$$

where χ_m is the magnetic susceptibility, μ is called the permeability, and μ_r is called the relative permeability. We'll explain later why χ_m isn't defined the same way as χ_e.

- Diamagnets have $-1 < \chi_m < 0$ and paramagnets have $\chi_m > 0$.
- Most materials are weakly diamagnetic ($|\chi_m| \ll 1$), but some are weakly paramagnetic.

- As discussed in **E5**, a superconductor totally expels magnetic fields, and thus can be viewed as a "perfect diamagnet" with $\chi_m = -1$ and hence $\mu = 0$.
- A "soft" ferromagnet (such as iron) is strongly paramagnetic, with $\chi_m \gg 1$.
- It is impossible to have $\chi_m < -1$, as then the energy density $B^2/2\mu$ would be negative. The material would spontaneously develop arbitrarily large B, and blow itself up.

Because magnetization can arise from freely moving electrons, bound electrons orbiting, or the spin of electrons, these ideas can be applied to both conductors and insulators. As always, we must be careful to remember that $\mathbf{B}$ is the total magnetic field, due to both whatever is outside the material, and the magnetization of the material itself.

Remark: Estimating Susceptibility

Why is it that many common solids have $|\chi_m| \ll 1$, but χ_e of order 1? Atoms contain a few valence electrons of charge q orbiting with radius of order a_0, the Bohr radius. To very roughly estimate electric and magnetic susceptibility, it's easiest to consider the extreme case where the field is so strong that the atom is about to fall apart.

The electrons are bound by an electric field $E_{\text{max}} \sim q/\epsilon_0 a_0^2$, so the atom will fall apart if the external field is much larger than this. And when the electron orbits are completely deformed, they will provide an electric dipole moment $p_{\text{max}} \sim q a_0$. So the polarizability is of order

$$\alpha \sim \frac{p_{\text{max}}}{E_{\text{max}}} \sim \epsilon_0 a_0^3.$$

Then the electric susceptibility is

$$\chi_e = \frac{P}{\epsilon_0 E} \sim \frac{n\alpha}{\epsilon_0} \sim 1$$

because the number density of atoms is $n \sim a_0^{-3}$. So, in a completely typical insulator, χ_e is of order 1, which is indeed what we observe.

On the other hand, magnetic susceptibility is penalized by the fact that magnetic forces are suppressed by a factor of the electron speed v. The field strength at which the magnetic force is as strong as that of the binding electric field is $B_{\text{max}} \sim E_{\text{max}}/v$. At this point, the electrons provide the maximum possible magnetic moment by all orbiting in the same direction, so

$$\mu_{\text{max}} \sim IA \sim (qv/a_0)(a_0^2) \sim q v a_0.$$

To leading order in χ_m, the magnetic susceptibility is

$$\chi_m \approx \frac{\mu_0 M}{B} \sim \frac{\mu_0 \mu_{\text{max}} n}{B_{\text{max}}} \sim \epsilon_0 \mu_0 v^2 = \frac{v^2}{c^2} \sim \alpha^2 \sim 10^{-4}$$

where α is the fine structure constant, introduced in **P1**. And indeed, this estimate result matches experimental results. Magnetic susceptibility is typically small because relativistic effects for valence electrons are weak.

Example 5

An infinite solenoid with n turns per length and current I is filled with material with magnetic susceptibility χ_m. Find the magnetic field inside.

Solution

The magnetic field inside has contributions from the solenoid wire and the magnetization,

$$B = \mu_0(nI + M) = \mu_0 nI + \frac{\chi_m}{1+\chi_m}B.$$

Solving for B yields

$$B = (1+\chi_m)nI = \mu nI$$

which can be a significant enhancement if the material is a soft ferromagnet.

[2] **Problem 8.** A sphere of magnetic susceptibility χ_m is placed in a uniform field $\mathbf{B}_0$, and as a result develops a uniform magnetization $\mathbf{M}$. Find $\mathbf{M}$ and the field $\mathbf{B}$ inside the sphere. Then check the limiting cases of a superconductor and a soft ferromagnet. Do they make sense?

Solution. In this case, by the result of example 4, we have

$$\mathbf{B} = \mathbf{B}_0 + \frac{2}{3}\mu_0\mathbf{M} = \mathbf{B}_0 + \frac{2}{3}\frac{\chi_m}{1+\chi_m}\mathbf{B}$$

inside the sphere. Solving for $\mathbf{B}$ and then for $\mathbf{M}$ gives

$$\mathbf{B} = \frac{1+\chi_m}{1+\chi_m/3}\mathbf{B}_0, \quad \mathbf{M} = \frac{\chi_m}{1+\chi_m/3}\frac{\mathbf{B}_0}{\mu_0}.$$

For a superconductor, $\chi_m = -1$, we have

$$\mathbf{B} = 0, \quad \mathbf{M} = -\frac{3\mathbf{B}_0}{2\mu_0}$$

which makes sense. For a soft ferromagnet, $\chi_m \to \infty$, we have

$$\mathbf{B} = 3\mathbf{B}_0, \quad \mathbf{M} = \frac{3\mathbf{B}_0}{\mu_0}.$$

This is a bit puzzling, because for an infinite solenoid the magnetic field inside was very large, but for a sphere it can apparently only be enhanced by a factor of 3.

As will be explained in example 6, the reason is that a sphere has a significant demagnetizing field. The magnetization tries to align with $\mathbf{B}_0$, but by the time it reaches $\mathbf{M} = 3\mathbf{B}_0/\mu_0$, the $\mathbf{H}$ field inside the sphere is already completely cancelled, so no further alignment can occur.

Idea 7: The H Field

Historically, magnetism was formulated in terms of the field

$$\mathbf{H} = \frac{1}{\mu_0}\mathbf{B} - \mathbf{M}.$$

Using this field gives a close analogy between electrostatics and magnetostatics. Note that

$$\nabla \times \mathbf{H} = \mathbf{J} - \mathbf{J}_b$$

so that $\mathbf{H}$ has no curl if there are no currents around besides the bound current. (This extra current is sometimes called "free" current $\mathbf{J}_f$.) In addition,

$$\nabla \cdot \mathbf{H} = -\nabla \cdot \mathbf{M}$$

which is analogous to how a polarization yields a charge density, $\nabla \cdot \mathbf{E} = -\nabla \cdot \mathbf{P}/\epsilon_0$. Therefore, any magnetostatic problem without free current can be mapped to an electrostatic one via

$$\mathbf{E} \leftrightarrow \mathbf{H}, \quad \mathbf{P} \leftrightarrow \mu_0 \mathbf{M}, \quad \epsilon_0 \leftrightarrow \mu_0, \quad \rho_b \leftrightarrow \rho_m$$

where $\rho_m = -\nabla \cdot \mathbf{M}$ is the "magnetic charge density".

This is the mathematical formalization of the idea of Gilbert dipoles, introduced in **E3**, which replace a true magnetic dipole with a pair of fictitious magnetic charges. At the time, we remarked that this gives you the correct magnetic field outside of a magnet, but not inside. The underlying reason is this analogy is actually computing $\mu_0 \mathbf{H}$, not $\mathbf{B}$. To get the correct $\mathbf{B}$ within a magnetized material, we have to compute $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$. Heuristically, $\mu_0 \mathbf{M}$ is the magnetic field due to dipole moments *right* at that location, while $\mu_0 \mathbf{H}$ is the contribution from all other currents and magnetic dipole moments.

Magnetic quantities were originally defined to be simple in terms of $\mathbf{H}$, with

$$\mathbf{M} = \chi_m \mathbf{H}, \quad \mathbf{B} = \mu \mathbf{H}.$$

In fact, some even call $\mathbf{H}$ "the" magnetic field, but we'll just call it "$\mathbf{H}$". As always, $\mathbf{H}$ here is the *entire* $\mathbf{H}$ field, due to both the material itself and the field the material is in. Note that a soft ferromagnet can be thought of as a material within which $\mathbf{H}$ is approximately zero.

Remark: The History of H

If you learned physics in the United States, you might have found the preceding idea unfamiliar, because it has been systematically removed from the introductory curriculum.

The reason comes down to history. First, it's worth noting that the choice between covering $\mathbf{B}$ and $\mathbf{H}$ isn't obvious. Today we would say the $\mathbf{B}$ field is more fundamental, because it is what determines the Lorentz force on a charge, and the force and torque on a dipole moment, and thereby the reading on a magnetometer. But for most of history, one could just as well argue that it is $\mu_0 \mathbf{H}$ that determines these things. After all, we can't embed a measuring device within a magnet. Doing so would require hollowing out a hole, which would remove $\mathbf{M}$ there, which would render $\mathbf{B}$ just equal to $\mu_0 \mathbf{H}$.

Therefore, choosing between $\mathbf{B}$ and $\mu_0 \mathbf{H}$ requires measuring some interaction where particles pass *through* a magnet. In the 1930s, this became possible with neutron scattering, since neutrons are electrically neutral but carry a magnetic dipole moment $\mathbf{m}$. As discussed here, the Nobel laureates Bloch and Schwinger disagreed on whether the interaction energy was $-\mathbf{m} \cdot \mathbf{B}$ or $-\mathbf{m} \cdot (\mu_0 \mathbf{H})$, which corresponds to treating $\mathbf{m}$ as Amperian or Gilbertian, respectively. The controversy remained open for decades, but was eventually settled by data.

At the same time, there was a growing consensus in the United States that magnetic poles were a "useless concept" which should be banished from teaching entirely. So, when the modern American physics curriculum was set in the 1960s, that's exactly what happened. For example, Halliday, Resnick, and Krane spends only a few pages covering magnetization, and half of them are spent admonishing the reader that magnetic poles don't exist. If you were educated in America, you probably only heard about them for a day in middle school.

Unfortunately, removing magnetic poles from the curriculum has some real costs. Experimentalists still think in terms of poles and $\mathbf{H}$, because it's harder to visualize how complicated currents source $\mathbf{B}$. Moreover, it's harder to do anything with $\mathbf{B}$ without vector calculus background. That's why some other countries' introductory physics courses put poles first. (But some teachers don't clearly explain how $\mathbf{B}$ and $\mu_0 \mathbf{H}$ differ, leading to confusion later.)

In this problem set, I'll tell you only what you need to know about $\mathbf{H}$ to solve theoretical problems. If you try to dive deeper into how experimentalists use it, you'll run into a lot more historical cruft. For instance, they tend to prefer the "Gaussian" system, where annoying factors of 4π are inserted into Maxwell's equations to make the Coulomb and Biot–Savart laws slightly simpler. Also, they use "cgs" (centimeter-gram-second) units, so all units need to be rescaled by some number of powers of 10. Worst of all, they have totally different units for $\mathbf{B}$ (Gauss), $\mathbf{H}$ (Oersted), and $\mathbf{M}$ (emu/cm^3), which are tricky to relate. I wouldn't recommend learning any of this unless you have to for your job.

Example 6

Using the $\mathbf{H}$ field, recompute the magnetic fields inside a uniformly magnetized sphere, and a rod magnetized parallel to and perpendicular to its axis.

Solution

In the first section, we found that inside these objects,

$$\mathbf{E} = \begin{cases} -\mathbf{P}/3\epsilon_0 & \text{sphere,} \\ -\mathbf{P}/2\epsilon_0 & \text{cylinder, perpendicular} \\ 0 & \text{cylinder, parallel} \end{cases}$$

Using the analogy above, we immediately conclude

$$\mathbf{H} = \begin{cases} -\mathbf{M}/3 & \text{sphere} \\ -\mathbf{M}/2 & \text{cylinder, perpendicular} \\ 0 & \text{cylinder, parallel} \end{cases}$$

The minus signs make sense because the "magnetic charge" accumulates on the side that $\mathbf{M}$ points to, and produces an $\mathbf{H}$ field in the opposite direction. Converting back to $\mathbf{B}$ yields

$$\mathbf{B} = \begin{cases} 2\mu_0\mathbf{M}/3 & \text{sphere} \\ \mu_0\mathbf{M}/2 & \text{cylinder, perpendicular} \\ \mu_0\mathbf{M} & \text{cylinder, parallel} \end{cases}$$

which precisely matches what we found in example 4 and problem 6, with much less effort.

Remark: Demagnetizing Fields

At a given point in an isolated magnet, $\mu_0\mathbf{H}$ is the part of $\mathbf{B}$ due to the rest of the magnet. However, the above example shows that $\mathbf{H}$ always points *against* the direction of $\mathbf{M}$, so a permanent magnet is always trying to demagnetize itself! Similarly, electrically polarized materials carry an internal "depolarization" field. This is why, in the absence of external fields, the vast majority of materials have zero polarization and magnetization.

Before the advent of very effective "hard" ferromagnets, like neodymium magnets, magnets had to be shaped to avoid this effect, e.g. by making them into long bars or horseshoes. Even so, the demagnetization effect would make the field produced by the magnet a little less than you would expect. For an average-shaped bar magnet made in the 1950s, the magnetic poles are effectively not at the ends, but rather 10% to 20% closer together.

This bit of historical trivia is irrelevant today, but strangely, it has stuck around in the Indian physics curriculum. Every Indian physics textbook demands its students memorize the ratio of the "magnetic length" and "geometric length" of a bar magnet, as if it were a fundamental constant of nature rather than an obsolete rule of thumb. But different Indian books don't even agree on what the ratio is, with HC Verma giving 84%, various JEE prep sources stating 4/5, 5/6, or 7/8, and none whatsoever explaining where the number comes from.

Idea 8: Magnetic Energy, Force, and Torque

The appropriate magnetic energy density depends on the material.

- If the magnetization is permanent, as in a hard ferromagnet, we should use the same potential energy introduced in **E4**. Specifically, the potential energy density is $-\mathbf{M}\cdot\mathbf{B}$.
- For all other materials, where the magnetization is induced by the presence of other fields, the total energy density is $B^2/2\mu$.

If you use one of these formulas where the other applies, you'll typically be off by a factor of 2.

Regardless of how the magnetization arises, the resulting force and torque in an external field $\mathbf{B}$ are given by the formulas introduced in **E4**. Specifically, the torque density is $\mathbf{M}\times\mathbf{B}$, and the force density is $\nabla(\mathbf{M}\cdot\mathbf{B})$, where the ∇ only acts on $\mathbf{B}$. Alternatively, if you're using the magnetic pole trick, a magnetic charge q_m feels a force $q_m\mathbf{B}$.

[3] **Problem 9.** A version of the method of images works for magnetic materials. Let's suppose there is vacuum at $z > 0$, and a material of relative permeability μ_r at $z < 0$. When using the method of images, we only care about the field at $z > 0$, where $\mathbf{B}$ and $\mathbf{H}$ are proportional. So we can directly use the analogy between $\mathbf{H}$ and $\mathbf{E}$.

(a) Suppose a magnetic charge q_m is a distance d above the plane. By recycling your answer to problem 3, find the magnetic charge q'_m of the image. What does it become if the material is a superconductor, or a soft ferromagnet?

(b) Of course, magnetic charges don't actually exist, so let's instead suppose a permanent magnetic dipole moment $\mathbf{m}$ was a distance d above the plane, with $\mathbf{m}$ pointing towards the plane. Characterize the image dipole, and find the force on the real dipole.

(c) To be even more concrete, consider a very long permanent magnet of cross-sectional area A and uniform magnetization M along its length. When one end of the magnet is placed flat against an iron plate, what is the force between them?

Solution. (a) When we apply the analogy described in idea 7, this problem becomes exactly the same as that problem, with χ_m corresponding to χ_e, and thus μ_r corresponding to κ. We conclude that the image magnetic charge is

$$q'_m = -q_m \frac{\mu_r - 1}{\mu_r + 1},$$

a distance d below the plane. Note that unlike the electric case, q'_m can have the same sign as q_m (for $\mu_r < 1$), or the opposite sign (for $\mu_r > 1$).

For a superconductor, $\mu_r = 0$, we have $q'_m = q_m$, which you might have already seen in a problem in **E5**. For a soft ferromagnet, $\mu_r \to \infty$, we have $q'_m = -q_m$. Both of these results are compatible with what we'd expect from example 7.

(b) In this case, there's an image dipole of magnitude

$$m' = m \frac{|\mu_r - 1|}{\mu_r + 1}.$$

For $\mu_r < 1$, it points in the opposite direction as the real dipole, while for $\mu_r > 1$, it points in the same direction.

Using the formula for the force on a dipole from **E4**, in the presence of the dipole field of the image, we find a force towards the plane of magnitude

$$F = \frac{3\mu_0}{2\pi} \frac{mm'}{(2d)^4} = \frac{3\mu_0}{32\pi} \frac{m^2}{d^4} \frac{\mu_r - 1}{\mu_r + 1}.$$

For $\mu_r < 1$, the force is instead repulsive.

(c) Here it's easiest to use the idea of magnetic charge. We can ignore the distant end of the magnet, because it's very far away. The end of the magnet touching the fridge has a magnetic charge density $\sigma_m = M$. For iron, which has $\mu_r \to \infty$, the resulting image has magnetic charge density $\sigma'_m = -M$, so it produces a magnetic field of magnitude $B' = \mu_0 \sigma'_m/2$ on each side of it. Thus, the interaction force is

$$F = B' \sigma_m A = \frac{\mu_0 M^2 A}{2}.$$

[2] **Problem 10.** AuPhO 2019, problem 13. A neat explanation of how a fridge magnet works; for this problem it will be useful to consult the answer sheet.

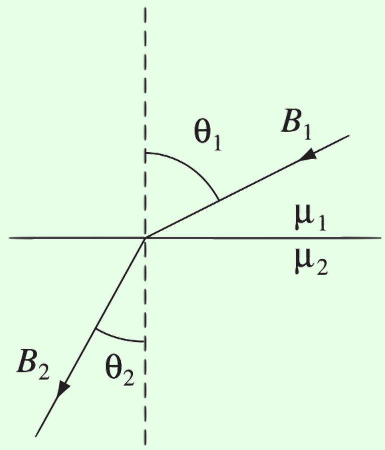
Solution. See the official solutions here.

[5] **Problem 11.** Physics Cup 2024, problem 3. This relatively straightforward problem reviews almost everything we've covered so far.

Solution. See the official solutions here.

Example 7: Griffiths 6.27

How does a magnetic field line bend when it passes from one medium to another?



Solution

We say the field lines "bend" because of Gauss's law for magnetism: they can't start or end, so each one has to keep on going. Let's suppose the figure above is drawn in the xz plane. Applying Gauss's law for magnetism in a small pillbox spanning the interface gives

$$B_1^z = B_2^z.$$

On the other hand, since $\nabla \times \mathbf{H}$ is zero (assuming no additional, "free" current is around), considering an Amperian loop spanning the interface gives

$$H_1^x = H_2^x.$$

Combining these results gives

$$\frac{\tan\theta_2}{\tan\theta_1} = \frac{\mu_2}{\mu_1}.$$

In other words, when a field line enters a medium with higher μ, it bends away from the normal, and when it enters a medium with lower μ, it bends towards the normal.

This statement has two limiting cases which will be important later.

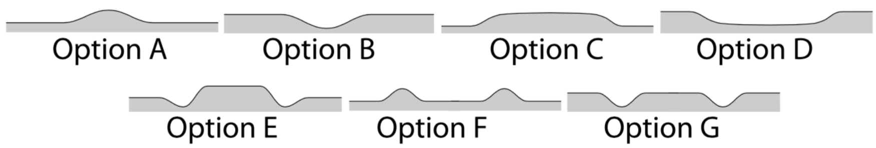
- A magnetic field line can't enter a superconductor ($\mu_2 = 0$) at all, so field lines approaching a superconductor bend away, to become tangent to them ($\theta_1 \to 90°$).

- A magnetic field line entering a soft ferromagnet ($\mu_2 \to \infty$) bends towards it to enter along the normal direction ($\theta_1 \to 0^\circ$), similar to how electric field lines approach conductors. It's also possible for θ_1 to be nonzero if $\theta_2 \to 90^\circ$, but we won't see any examples of this.

You can see both of these behaviors in the limiting cases of problem 9. In general, we conclude that magnetic field lines are "attracted" to regions of higher μ, which makes sense because it helps minimize the energy. Soft ferromagnets tend to keep magnetic field lines within themselves, which is why they're used in transformers.

[2] **Problem 12** (IPhO 2012 Experiment)**.** Water is a diamagnetic substance. A powerful cylindrical magnet with field B is placed below the water surface.

(a) Which of the following shows the resulting shape of the water surface?



The magnet is roughly 2/3 as wide as each of these sketches.

(b) Let ρ be the density of the water. If the maximum change in height of the water surface has magnitude h, find an approximate expression for the magnetic susceptibility χ_m of water.

For a very closely related, but more extreme problem, see EuPhO 2018, problem 2.

Solution. (a) For a diamagnetic substance, $\mu < \mu_0$, so the magnetic field energy is higher when water is present. The water surface is an equipotential, so a higher magnetic field energy at some points must be compensated by a lower gravitational potential energy. Thus, the answer is option D.

(b) Equating the change in gravitational potential energy with the change in field energy, both per volume, gives

$$\rho g h = \frac{B^2}{2\mu} - \frac{B^2}{2\mu_0} = \frac{B^2}{2\mu\mu_0}(\mu_0 - \mu) \approx \frac{B^2}{2\mu_0^2}(\mu_0 - \mu)$$

where the last step follows because $\mu \approx \mu_0$. We thus have

$$\chi_m = \frac{\mu - \mu_0}{\mu_0} = -\frac{2\mu_0 \rho g h}{B^2}.$$

With a strong magnet, and a measurement of h accurate to about 0.1 mm, one can indeed detect this effect. Note that if you treated the dipole moment as permanent, and used a potential energy density $-\mathbf{M} \cdot \mathbf{B}$, your answer here would be off by a factor of 2.

[3] **Problem 13.** EFPhO 2004, problem 6. A cute exercise with permanent magnets.

Solution. See the official solutions here.

[5] **Problem 14.** IPhO 2022, problem 1. A series of exercises on spherical magnets, which uses almost everything covered in this section.

[4] **Problem 15.** Physics Cup 2012, problem 2. If you only know what's taught in American introductory courses, this problem is basically impossible. If you only know what's stated explicitly in Griffiths, it's very hard. But if you've internalized the intuition of the above examples, and the relevant section of **E5** on superconductors, it should be relatively approachable.

Solution. See the official solutions here.

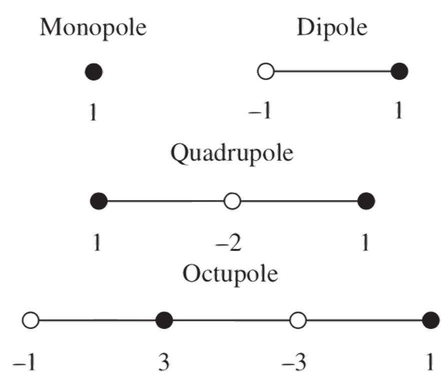
[5] **Problem 16.** Physics Cup 2018, problem 3. A substantially tougher problem which requires solving some differential equations. I recommend starting from the fifth hint.

Solution. See the official solutions here.

3 Multipoles

In this section, we explore some of the physics of dipoles and higher multipoles.

[3] **Problem 17** (Purcell 10.27). Two monopoles of opposite sign form a dipole, two dipoles of opposite sign for a quadrupole, and so on. Hence we can construct arbitrarily high multipoles using the rows of Pascal's triangle.



The field of a dipole falls as $1/r^3$, a quadrupole as $1/r^4$, and an octupole as $1/r^5$.

(a) To warm up, verify explicitly that the quadrupole field along the axis of the quadrupole starts at $1/r^4$, i.e. that all lower terms cancel.

(b) Prove that this cancellation occurs for general multipoles along their axis.

(c) **[A]** The magnitude and orientation of a dipole is specified by a vector, with three components. How many numbers are necessary to specify the magnitude and orientation of a quadrupole? (The linear quadrupoles here are just a special case of a general quadrupole.) Try to generalize to arbitrary multipoles.

To learn how to decompose an arbitrary charge distribution into multipoles, see section 3.4 of Griffiths.

Solution. Since we're lazy we'll set the Coulomb constant $k = 1$, and the unit of charge also to 1, as well as the unit of distance spacing.

(a) See the solution to (b).

(b) A simple way to do this is to reason inductively. For example, an octupole field is nothing more than two quadrupoles whose leading terms cancel, so the leading field of an octupole has to be at least one power lower in r.

However, we will give an explicit proof. A 2^N-pole can be constructed from $N+1$ charges, with charge j placed at $x=-j$ with charge $(-1)^j\binom{N}{j}$. Then the field at point x is

$$E(x)=\sum_{j=0}^{N}(-1)^j\binom{N}{j}\frac{1}{(x+j)^2}=x^{-2}\sum_{j=0}^{N}\binom{N}{j}\sum_{k=0}^{\infty}\binom{-2}{k}(j/x)^k.$$

We see that this can be split into sums of the form $f(k)=\sum_{j=0}^{N}(-1)^j\binom{N}{j}j^k$, and the coefficient of x^{-2-k} is some nonzero multiple times $f(k)$. So it suffices to show that $f(k)=0$ for all $k<N$, and $f(N)\neq 0$. This is an exercise in algebraic sums. The key idea is to define

$$g(k)=\sum_{j=0}^{N}(-1)^j\binom{N}{j}\binom{j}{k}=\sum_{j=k}^{N}(-1)^j\binom{N}{j}\binom{j}{k}.$$

We see that j^k can be written as a linear combination of $\binom{j}{0},\ldots,\binom{j}{k}$, so it suffices to show that $g(k)=0$ for all $k<N$, and that $g(N)\neq 0$. We see that

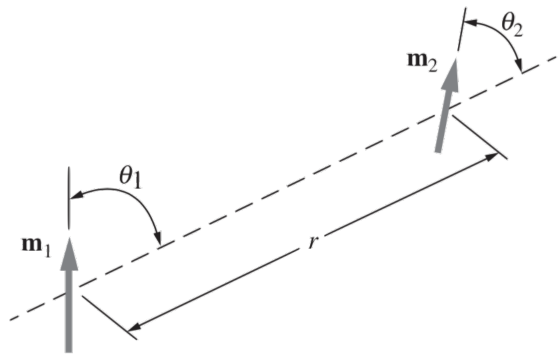
$$\begin{aligned}g(k)&=\sum_{j=k}^{N}(-1)^j\binom{N}{j}\binom{j}{k}\\&=\sum_{j=k}^{N}(-1)^j\binom{N}{k}\binom{N-k}{j-k}\\&=\binom{N}{k}\sum_{j=k}^{N}(-1)^j\binom{N-k}{j-k}\\&=\binom{N}{k}(-1)^k\cdot\mathbf{1}_{k=N}\end{aligned}$$

where we used the fact that $\sum_{\ell=0}^{M}(-1)^\ell\binom{M}{\ell}=\mathbf{1}_{M=0}$ (here $\mathbf{1}_S$ is 1 if and only if S is true, and is 0 otherwise), which follows from the binomial theorem. This completes the proof.

(c) Let's think of a general quadrupole as a superposition of two dipoles in opposite directions. Then there are three things that determine a quadrupole: the strength of the quadrupole moment (i.e. the prefactor of the $1/r^4$ field), the orientation of the first dipole, and the direction the second dipole is displaced from it. This is $1+2+2=5$ total parameters.

Similarly, to specify an octupole, we do the same above, then specify the direction the second quadrupole is displaced, giving $5+2=7$ parameters. In general, a 2^N-pole has $2N+1$ parameters.

[3] **Problem 18** (Purcell 11.23)**.** Consider two magnetic dipoles with coplanar dipole moments.



Show that the associated potential energy is

$$U = \frac{\mu_0 m_1 m_2}{4\pi r^3} \left(\sin\theta_1 \sin\theta_2 - 2\cos\theta_1 \cos\theta_2\right).$$

For what orientations is this potential energy maximized or minimized?

Solution. The magnetic field from a dipole pointing in the z direction is:

$$\mathbf{B} = \frac{\mu_0 m}{4\pi r^3}(2\cos\theta \hat{\mathbf{r}} + \sin\theta \hat{\boldsymbol{\theta}}).$$

Let $\hat{\mathbf{n}}$ be the unit vector perpendicular to $\hat{\mathbf{r}}$ ($\theta = -\pi/2$). Then the field of $\mathbf{m}_1$ is

$$\mathbf{B}_{12} = \frac{\mu_0 m_1}{4\pi r^3}(2\cos\theta_1 \hat{\mathbf{r}} + \sin\theta_1 \hat{\mathbf{n}})$$

The potential of a dipole in a field is $U = -\mathbf{m} \cdot \mathbf{B}$. Note that $\mathbf{m}_2 = m_2 \cos\theta_2 \hat{\mathbf{r}} - m_2 \sin\theta_2 \hat{\mathbf{n}}$.

$$U = -\mathbf{m}_2 \cdot \mathbf{B}_{12} = \frac{\mu_0 m_1 m_2}{4\pi r^3}(\sin\theta_1 \sin\theta_2 - 2\cos\theta_1 \cos\theta_2)$$

as desired. To extremize this expression, we set the partial derivatives with respect to θ_1 and θ_2 equal to zero. The results are

$$\cos\theta_1 \sin\theta_2 = -2\sin\theta_1 \cos\theta_2, \quad \sin\theta_1 \cos\theta_2 = -2\cos\theta_1 \sin\theta_2$$

which implies that

$$\cos\theta_1 \sin\theta_2 = \sin\theta_1 \cos\theta_2 = 0.$$

This can only hold if

$$\cos\theta_1 = \cos\theta_2 = 0 \text{ or } \sin\theta_1 = \sin\theta_2 = 0.$$

The first option leads to local maxima or minima, where both the angles are $\pm\pi/2$, with maximum/minimum energy occurring when the dipoles are anti-aligned/aligned. The second option leads to the *global* maximum and minimum,

$$\text{maximum: } (0,0) \text{ or } (\pi,\pi), \quad \text{minimum: } (0,\pi) \text{ or } (\pi,0)$$

where the dipoles are anti-aligned/aligned, along the direction of the separation between them.

[2] **Problem 19** (Purcell 11.36)**.** Three magnetic compasses are placed at the corners of a horizontal equilateral triangle. As in any ordinary compass, each compass needle is a magnetic dipole constrained to rotate in a horizontal plane. The Earth's magnetic field has been shielded. What orientation will the compass needles eventually assume? Does your result also hold for regular N-gons?

Solution. We claim they point in the direction of the tangents to the circumcircle of the triangle. In this case, the field at any one corner due to the compasses at the other corners points in the tangential direction, so the compasses are all aligned with the local fields.

We can show this claim by symmetry. Consider the field at a given corner of the triangle. Flipping about the axis that passes through this corner and the midpoint of the opposite side negates the dipole moments at the other two corners, so it must negate the field. But physically, the rotation operation negates the tangential component of the field. So there must only be a tangential component, i.e. the field at this corner is purely tangential. This argument holds unchanged for regular N-gons.

[3] **Problem 20.** Some questions about forces between dipoles and other multipoles.

(a) Above, you've shown that the force between permanent magnetic dipoles falls off as $1/r^4$. How about two permanent electric dipoles?

(b) How about a permanent dipole and a permanent quadrupole?

(c) How about two permanent quadrupoles?

(d) Now consider an ion and a neutral atom. The electric field of the ion polarizes the atom; the field of that induced dipole then reacts on the ion. Show that the resulting force is attractive and falls as $1/r^5$.

Solution. (a) The basic form of the fields and forces is identical, so the answer is the same.

(b) The field of a quadrupole goes like $1/r^4$, so energy of the dipole goes like $U \sim mB \sim 1/r^4$. Thus, the interaction energy in this case goes like $1/r^4$, for a force of $1/r^5$.

(c) A single quadrupole is two dipoles with moments $\mathbf{m}$ and $-\mathbf{m}$ separated by $d\mathbf{r}$ where the magnitude of the quadrupole moment is $\sim |\mathbf{m}||d\mathbf{r}|$. (Technically, the quadrupole moment is a tensor, and the sizes of its individual components depends on the relative orientation of $\mathbf{m}$ and $d\mathbf{r}$, but we won't worry about that detail here, since we're only looking for the scaling of the force with r.) The energy of the quadrupole is

$$\mathbf{m} \cdot \mathbf{B}(\mathbf{r} + d\mathbf{r}) - \mathbf{m} \cdot \mathbf{B} = \mathbf{m} \cdot ((\text{some sort of derivative of } \mathbf{B}) \cdot d\mathbf{r}).$$

The field of one quadrupole is $1/r^4$, so its derivative is $1/r^5$. Thus the energy of interaction goes like $1/r^5$, for a force of $1/r^6$.

(d) The field of the ion falls as $1/r^2$, so the dipole moment induced is $p \sim 1/r^2$. Furthermore, the dipole moment points along the field and hence the displacement between the ion and atom, indicating the force is attractive. The electric field from the dipole (and hence the force) goes as $p/r^3 \sim 1/r^5$. (You might wonder if the induced dipole then gives the ion itself a dipole moment. It does, but the resulting force is much weaker than the one we found here, between the induced dipole and the ion's overall charge.)

4 Electromagnetic Waves in Matter

In this section, you will work out some of the theory of electromagnetic waves in matter.

Idea 9

In the absence of any free charge or current, Maxwell's equations in matter are identical to Maxwell's equations in vacuum, except that ϵ_0 and μ_0 are related by ϵ and μ, so the waves propagate with speed $1/\sqrt{\epsilon\mu} = c/n$, with $E = (c/n)B$.

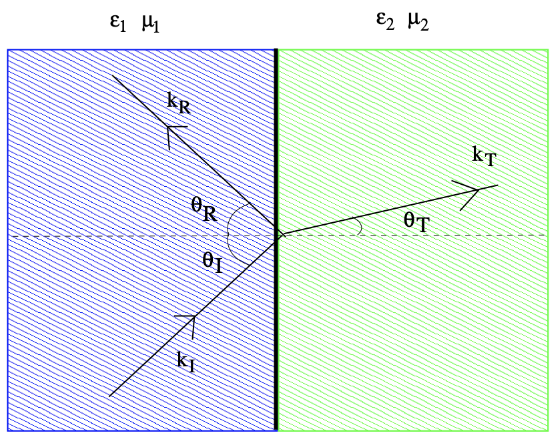
[5] **Problem 21.** Suppose the regions $x < 0$ and $x > 0$ are filled with material with permittivities ϵ_1 and ϵ_2, both with permeability μ_0. (This is typical; if you don't count permanent magnets, most objects have permeability about μ_0.) We send in an incident wave from the left with electric field

$$\mathbf{E}_i e^{i(\mathbf{k}_i \cdot \mathbf{x} - \omega_i t)}.$$

The wave will be both transmitted and reflected at the interface, so the total electric field is

$$\mathbf{E} = \begin{cases} \mathbf{E}_i e^{i(\mathbf{k}_i \cdot \mathbf{x} - \omega_i t)} + \mathbf{E}_r e^{i(\mathbf{k}_r \cdot \mathbf{x} - \omega_r t)} & x < 0, \\ \mathbf{E}_t e^{i(\mathbf{k}_t \cdot \mathbf{x} - \omega_t t)} & x > 0. \end{cases}$$

The angles with the normal are θ_i, θ_r, and θ_t as shown.



(a) Argue that by continuity of the field at the boundary,

$$\omega_i = \omega_r = \omega_t.$$

(b) Suppose the y-axis is oriented so that $\mathbf{k}_i \cdot \hat{\mathbf{y}} = 0$. Argue that

$$\mathbf{k}_r \cdot \hat{\mathbf{y}} = \mathbf{k}_t \cdot \hat{\mathbf{y}} = 0, \quad \mathbf{k}_i \cdot \hat{\mathbf{z}} = \mathbf{k}_r \cdot \hat{\mathbf{z}} = \mathbf{k}_t \cdot \hat{\mathbf{z}}.$$

From these conditions, derive the laws of reflection and refraction,

$$\theta_i = \theta_r, \quad n_1 \sin\theta_i = n_2 \sin\theta_t.$$

Note that neither this part nor the previous part require Maxwell's equations; they hold for *all* kinds of waves as long as we define $n_i \propto 1/v_i$.

(c) Argue that at the boundary, $\mathbf{E}_\parallel$ and $B_\perp$ must be continuous in general. In this case, because both sides have the same permittivity μ_0, there is no bound current, so $\mathbf{B}_\parallel$ is also continuous.

(d) Now suppose the electric fields are polarized along the $\mathbf{y}$ axis, so $\mathbf{E}_i$, $\mathbf{E}_r$, and $\mathbf{E}_t$ are all parallel to the y-axis. Then continuity of $\mathbf{E}_\parallel$ gives

$$E_i + E_r = E_t.$$

Using continuity of $\mathbf{B}_\parallel$, show that

$$\frac{E_r}{E_i} = \frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t}, \quad \frac{E_t}{E_i} = \frac{2n_1 \cos\theta_i}{n_1 \cos\theta_i + n_2 \cos\theta_t}.$$

These are the Fresnel equations for normal polarized light. (Hint: this is a bit messy, so you can warm up with the case $\theta_i = 0$.)

(e) If $n_1 > n_2$, then total internal reflection occurs when

$$\sin\theta_i > \frac{n_2}{n_1}$$

and the wave is totally reflected. Nonetheless, E_t is nonzero in this regime. To make sense of this, show that $\mathbf{k}_t \cdot \mathbf{x}$ is imaginary in this regime, indicating that the "transmitted" wave does not propagate in the region $x > 0$, but rather exponentially decays.

Solution. (a) Continuity of the field at the interface gives

$$E_i' e^{\omega_i t} + E_r' e^{\omega_r t} = E_t' e^{\omega_t t}.$$

This equation can only be satisfied if $\omega_i = \omega_r = \omega_t$, so that all three exponentials have the same time dependence.

Of course, the deeper reason behind this is just what we said in **M4** and will see again in **W1**. The differential equation the field obeys is linear, and has no explicit time dependence. Thus, it has solutions with uniform frequency everywhere.

(b) Again, look at the boundary and match parallel components of $\mathbf{E}$. By fixing z but varying y, we get an equation of the form

$$E_i' e^{\mathbf{k}_i \cdot y\hat{\mathbf{y}}} + E_r' e^{\mathbf{k}_r \cdot y\hat{\mathbf{y}}} = E_t' e^{\mathbf{k}_t \cdot y\hat{\mathbf{y}}}.$$

As before, for this to work, all the frequency factors of $\mathbf{k}_{\text{something}} \cdot \hat{\mathbf{y}}$ must be identical. We can do the same argument for $\hat{\mathbf{z}}$.

The first set gives that all the $\mathbf{k}$'s are in the same plane. The second gives that $k_i \sin\theta_i = k_r \sin\theta_r = k_t \sin\theta_t$. We have $\omega/k = c/n$, so $k = n\omega/c$. Since all the ω's are the same, this equation then reads

$$n_1 \sin\theta_i = n_1 \sin\theta_r = n_2 \sin\theta_t,$$

which is exactly what we want.

(c) For $B_\perp$, consider a thin Gaussian pillbox that straddles the boundary. By Gauss's law for magnetism, the magnetic flux through it must be zero. In the limit of a very thin pillbox, this ensures the continuity of $B_\perp$.

For $\mathbf{E}_\parallel$, consider a thin Amperian loop that straddles the boundary, and consider $\oint \mathbf{E} \cdot d\mathbf{s}$. As the width of the loop goes to zero, the magnetic flux through it goes to zero, so this integral must be zero. Taking loops of various orientations, this ensures the continuity of $\mathbf{E}_\parallel$.

Note that $E_\perp$ and $\mathbf{B}_\parallel$ need not be continuous, because we can have surface charges and currents at the boundary. Since both sides have the same μ_0, there are no surface currents, so $\mathbf{B}_\parallel$ is continuous.

(d) The continuity of $\mathbf{B}_\perp$ gives

$$B_i \cos\theta_i - B_r \cos\theta_r = B_t \cos\theta_t.$$

Since $B = En/c$, this means

$$E_i n_1 \cos\theta_i - E_r n_1 \cos\theta_r = E_t n_2 \cos\theta_t.$$

Now with the continuity of $\mathbf{E}_\parallel$ ($E_i + E_r = E_t$), and $\theta_i = \theta_r$, we have

$$E_i n_1 \cos\theta_i - E_r n_1 \cos\theta_i = E_i n_2 \cos\theta_t + E_r n_2 \cos\theta_t$$

which yields
$$\frac{E_r}{E_i} = \frac{n_1\cos\theta_i - n_2\cos\theta_t}{n_1\cos\theta_i + n_2\cos\theta_t}, \quad \frac{E_t}{E_i} = \frac{2n_1\cos\theta_i}{n_1\cos\theta_i + n_2\cos\theta_t}$$
as desired.

(e) As seen in part (b), we have

$$(k_i)_y = (k_t)_y, \quad (k_i)_z = (k_t)_z$$

The magnitudes of the wavenumbers are known:

$$k_t = \frac{\omega}{c} n_2, \quad k_i = \frac{\omega}{c} n_1, \quad k_t = k_i \frac{n_2}{n_1}$$

Then using the expression for the magnitude of $\mathbf{k}_t$:

$$(k_t)_x^2 = k_t^2 - (k_t)_y^2 - (k_t)_z^2 = \left(k_i \frac{n_2}{n_1}\right)^2 - (k_i)_y^2 - (k_i)_z^2$$

Since $\sin\theta_i$ represents $k_\parallel/k$, then for $\sin\theta_i = k_\parallel/k > n_2/n_1$:

$$(k_t)_x^2 = \left(k_i \frac{n_2}{n_1}\right)^2 - (k_i)_\parallel^2 < (k_i)_\parallel^2 - (k_i)_\parallel^2 = 0$$

from which we conclude that
$$(k_t)_x^2 < 0.$$

Thus the electric field will exponentially decay as it goes into the material.

[5] **Problem 22.** In most common materials, $\mu \approx \mu_0$ while ϵ depends on frequency. We'll investigate the origin of this frequency dependence below.

(a) Model an electron in an atom as a mass m with charge q attached to a spring, with natural angular frequency ω_0 and a damping force $-m\gamma\mathbf{v}$, in an electric field $\mathbf{E}_0 e^{-i\omega t}$. Write down the equation of motion for the electron.

(b) The atomic polarizability is $\mathbf{p} = \alpha \mathbf{E}$. Show that

$$\alpha = \frac{q^2/m}{-\omega^2 + \omega_0^2 - i\gamma\omega}.$$

(c) For a gas with small number density n, the Clausius–Mossotti formula reduces to

$$\epsilon = \epsilon_0 + n\alpha.$$

Therefore, the permittivity is generally a complex number. The wavevector and angular frequency are related by $k^2 = \mu\epsilon\omega^2$. Explain why the fact that ϵ is complex indicates that waves can be absorbed.

From this point on, you may approximate γ as small.

(d) What value of ω maximizes the absorption rate of the electromagnetic waves? Roughly how many wavelengths does such a wave propagate before being mostly absorbed?

(e) What value of ω maximizes the speed of the electromagnetic waves, and what is that speed?

(f) Transparent objects such as glass can be modeled as having a very high resonant frequency, much higher than that of visible light. Does blue light or red light refract more when passing from air to glass?

The intuitive reason that these electrons can affect the propagation speed of light is because they emit secondary electromagnetic waves that are out of phase with the original wave; this "pushes" the phase of the composite wave forward or backward, affecting the phase velocity. A nice explanation of this can be found in chapter I.31 of the Feynman lectures.

Solution. (a) We have

$$m\ddot{\mathbf{r}} = -m\omega_0^2\mathbf{r} - m\gamma\mathbf{v} + q\mathbf{E}_0 e^{-i\omega t}.$$

(b) Suppose $\mathbf{r} = \mathbf{r}_0 e^{i\omega t}$ where $\mathbf{r}_0$ is potentially complex. Then, we see that $\mathbf{E}_0 \parallel \mathbf{r}_0$ and

$$-m\omega^2\mathbf{r} = -m\omega_0^2\mathbf{r} - m\gamma i\omega\mathbf{r} + q(E_0/r_0)\mathbf{r}.$$

Thus,

$$E_0/r_0 = \frac{m(\omega_0^2 - \omega^2 + i\gamma\omega)}{q}.$$

We have that $p = -r_0 q/E_0$ which yields the result.

(c) If ϵ is complex, then with $\mu \approx \mu_0$ and ω^2 being real, then $k^2 = \mu\epsilon\omega^2$ will also be complex. Thus with a complex wavevector $\mathbf{k}$, the field of $\mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)}$ will exponentially decay since there will be a negative term from the complex wavevector $\mathbf{k}$.

(d) The absorption occurs exponentially from the complex component of kx. With $k = \omega\sqrt{\mu\epsilon} \approx \omega\sqrt{\mu_0\epsilon_0}(1 + \frac{n\alpha}{2\epsilon_0})$, the absorption rate is maximized when the complex component of k is maximized.

$$\beta \equiv \mathrm{Im}(k) = \mathrm{Im}\left(\frac{\omega n}{2c\epsilon_0}\alpha\right) = \frac{\omega n}{2c\epsilon_0}\frac{q^2/m}{(\omega_0^2 - \omega^2)^2 + (\omega\gamma)^2}(\gamma\omega)$$

$$= \frac{q^2\gamma n}{2mc\epsilon_0} \frac{\omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}$$

The maximum value of this occurs when

$$\frac{d\beta}{d\omega^2} \propto \left((\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2 - \omega^2(2(\omega^2 - \omega_0^2) + \gamma^2)\right) = 0$$

which simplifies to yield

$$\omega_0^4 - \omega^4 = 0.$$

Thus a angular frequency of $\omega = \omega_0$ will maximize the absorption rate of the electromagnetic wave.

The electric field will have a factor of $e^{-\beta x}$, and at $\omega = \omega_0$, $\beta = \frac{q^2 n}{2\gamma m c\epsilon_0}$. The value of the real wavevector $\mathrm{Re}\, k$ will be close to (note that $\mathrm{Re}(\alpha) = 0$ at $\omega = \omega_0$):

$$\mathrm{Re}(k) = \frac{\omega_0}{c}\left(1 + \mathrm{Re}\left(\frac{n\alpha}{2\epsilon_0}\right)\right) = \frac{\omega_0}{c}$$

Then for the wave to fall off by a factor of e, the wave will need to travel a distance of $\frac{1}{\beta}$, which is $\frac{1}{\beta\lambda} = \frac{k}{2\pi\beta}$ wavelengths. Thus,

$$\frac{k}{2\pi\beta} = \frac{\omega_0\gamma m\epsilon_0}{\pi q^2 n}$$

is the number of wavelengths it will travel before the amplitude gets reduced by a factor of e.

(e) The phase velocity is maximized when $\frac{\omega}{\mathrm{Re}\, k}$, or $\mathrm{Re}\, \frac{1}{\sqrt{\mu\epsilon}}$ is maximized.

$$v_p = \mathrm{Re}\, \frac{1}{\sqrt{\mu\epsilon}} \approx c\left(1 - \mathrm{Re}\, \frac{1}{2}\frac{n\alpha}{\epsilon_0}\right) = c + \frac{cq^2 n}{2m\epsilon_0} \frac{\omega^2 - \omega_0^2}{(\omega^2 - \omega_0^2)^2 + (\gamma\omega)^2}$$

Differentiating with respect to ω^2 and finding where it's zero yields

$$(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2 - (\omega^2 - \omega_0^2)(2(\omega^2 - \omega_0^2) + \gamma^2) = 0$$

$$(\omega^2 - \omega_0^2)^2 = \omega_0^2\gamma^2$$

$$\omega^2 = \omega_0^2 \pm \omega_0\gamma$$

Looking at the original, the smaller solution yields the minimum velocity, and the larger solution yields the maximum velocity (which happens to be greater than c). The maximum phase velocity is

$$v_{\max} = c + \frac{cq^2 n}{2m\epsilon_0} \frac{\omega_0\gamma}{(\omega_0\gamma)^2 + \gamma^2(\omega_0^2 + \omega_0\gamma)}$$

(f) From the previous part, we have

$$v_p = c - \frac{cq^2 n}{2m\epsilon_0} \frac{\omega_0^2 - \omega^2}{(\omega^2 - \omega_0^2)^2 + (\gamma\omega)^2}$$

and now we know that $\omega_0 \gg \omega$, so

$$\frac{v_p}{c} \approx 1 - \frac{q^2 n}{2m\epsilon_0} \frac{\omega_0^2 - \omega^2}{\omega_0^4 - 2\omega_0^2\omega^2 + (\gamma\omega)^2} \approx 1 - \frac{q^2 n}{2m\epsilon_0\omega_0^2}(1 + \omega^2/\omega_0^2).$$

Thus, increasing the frequency would decrease v_p and increase the index of refraction, so blue light would refract more.

[5] **Problem 23.** IPhO 2002, problem 1. A neat application of electromagnetic waves in matter.

[5] **Problem 24.** APhO 2007, problem 2. A problem on an exotic negative index of refraction.

Remark

Above, we considered the response of a medium composed of atoms, obeying $p = \alpha E$. However, this relation is just an approximation, like Hooke's law. For larger electric fields, higher order terms are necessary,

$$p = \alpha E + \alpha' E^2 + \ldots$$

which lead to strange effects, studied in the field of nonlinear optics. For example, suppose we send in light of angular frequency ω. Then

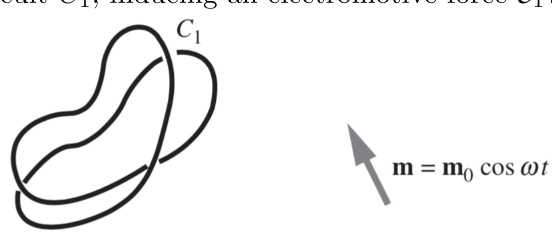
$$E^2 \propto \cos^2(\omega t) = \frac{1 + \cos(2\omega t)}{2}.$$

That means that a nonlinear medium can respond to light at angular frequency ω by oscillating, and hence emitting light, at angular frequency 2ω. This phenomenon is called frequency doubling, or second-harmonic generation, and converts red light to blue. Similarly, for a cubic nonlinearity, you can use trigonometric identities to show that frequency tripling can occur.

5 Electromagnetic Systems

In this section we'll consider problems that use everything we've covered, with a focus on technological applications and systems with multiple moving parts.

[2] **Problem 25** (Purcell 11.19). A magnetic dipole $\mathbf{m}$ oscillates so that $\mathbf{m}(t) = \mathbf{m}_0 \cos \omega t$. Some of its flux links the nearby circuit C_1, inducing an electromotive force $\mathcal{E}_1 \sin \omega t$.



If a current I_1 flowed in C_1, then the resulting field at the location of the dipole would be $\mathbf{B}_1$. Show that $\mathcal{E}_1 = (\omega/I_1)\mathbf{B}_1 \cdot \mathbf{m}_0$. (Hint: recall the results involving mutual inductance in **E5**.)

Solution. Let there be a flux Φ_1 in circuit 1 and $\Phi_2 = \mathbf{B}_1 \cdot \mathbf{m}_0 / I_2$ in circuit 2. Then because $L_{12} = L_{21}$, as stated in **E5**, we have

$$\Phi_1 / I_2 = \Phi_2 / I_1 \implies \Phi_1 = \mathbf{B}_1 \cdot \mathbf{m}_0 / I_2.$$

Then,

$$\mathcal{E}_1(t) = -d\Phi_1/dt = (\omega/I_1)\mathbf{B}_1 \cdot \mathbf{m}_0 \sin \omega t,$$

so $\mathcal{E}_1 = (\omega/I_1)\mathbf{B}_1 \cdot \mathbf{m}_0$.

[3] **Problem 26.** EFPhO 2007, problem 3. A problem on focusing particles with electric fields.

Solution. See the official solutions here.

[4] **Problem 27.** IPhO 2004, problem 3. A practical problem which also reviews damped/driven oscillations.

[4] **Problem 28.** EFPhO 2014, problem 1. A challenging problem about a complex nonlinear circuit.

Solution. See the official solutions here.

[5] **Problem 29.** Physics Cup 2020, problem 1. (Unfortunately, the word "dielectric" has two common distinct meanings, and you'll have to tell from context which is meant. When this problem states that the rod is "dielectric", it means that the rod always has zero charge and current density everywhere, i.e. it is an insulator with zero electric susceptibility. Alternatively, you can suppose the rod might have some electric susceptibility, but it's too thin to have an effect on the dynamics of the metal balls.)

Solution. See the official solutions here.