Electricity and Magnetism

Professor John Boccio

Your best friend in this class is mathematics.

Basics of multivariable and vector calculus will be taught using Cartesian coordinate systems.

My only assumption: You know single variable calculus.

We will do both integral and differential formulations of E&M

Goal: Derive and understand at Maxwell's equations

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho \qquad \vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \qquad \vec{\nabla} \times \vec{B} = \frac{4\pi}{c}\vec{J} + \frac{1}{c}\frac{\partial \vec{E}}{\partial t}$$

We want to be able to say what they really mean!

Review some basic concepts:

Space of Physics

Scalar = a number (temperature, minutes after noon, ...)
Vector = set of 3 numbers (characterizes a point in the 3-dimensional world)

The vector representing the location of a point in 3-dimensional space relative to a given origin is called the position vector. It is represented by the 3-tuple

$$\vec{r} = (x, y, z) = (x_1, x_2, x_3)$$

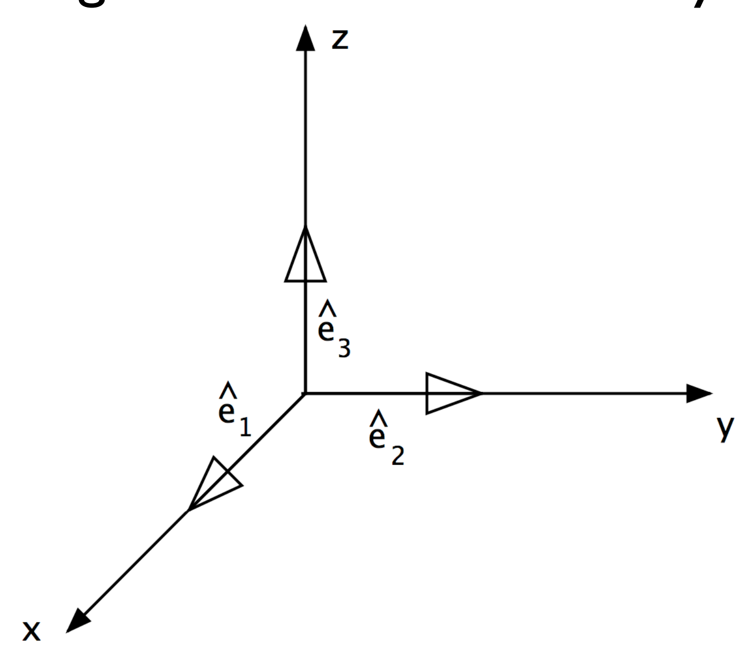
This is a shorthand notation for

$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k} = x_1\hat{e}_1 + x_2\hat{e}_2 + x_3\hat{e}_3 = x\hat{x} + y\hat{y} + z\hat{z}$$

which are all equivalent representations of the position vector. The three(for 3 dimensions) vectors

$$\hat{e}_1, \hat{e}_2, \hat{e}_3 \quad \text{or} \quad \hat{x}, \hat{y}, \hat{z} \quad \text{or} \quad \hat{i}, \hat{j}, \hat{k}$$

are unit vectors (length = 1) defining a right-handed coordinate system as show below:



The numbers (x,y,z) are called the components of the position vector with respect to the chosen unit vectors. The unit vectors are chosen to be orthogonal (perpendicular) for convenience as we shall see.

The length of the position vector is given (Pythagorean theorem) by

$$r = |\vec{r}| = \sqrt{x^2 + y^2 + z^2}$$

We define a **vector product operation**(symbol = $\cdot$) called the scalar or dot or inner product by the following relations

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij} = \text{Kronecker delta}$$

$$= \begin{cases} 1 & \text{i=j} \\ 0 & \text{otherwise} \end{cases}$$

In terms of the scalar product we have

$$r^2 = \vec{r} \cdot \vec{r} = \left(\sum_{i=1}^{3} x_i \hat{e}_i \right) \cdot \left(\sum_{j=1}^{3} x_j \hat{e}_j \right) = \sum_{i,j=1}^{3} x_i x_j \hat{e}_i \cdot \hat{e}_j = \sum_{i,j=1}^{3} x_i x_j \delta_{ij} = \sum_{i=1}^{3} x_i^2$$

We now define the Einstein summation convention,which assumes a summation anytime that a repeated index appears in the same term. Using this convention the above derivation looks like

$$r^2 = \vec{r} \cdot \vec{r} = \left(x_i \hat{e}_i \right) \cdot \left(x_j \hat{e}_j \right) = x_i x_j \hat{e}_i \cdot \hat{e}_j = x_i x_j \delta_{ij} = x_i x_i$$

We define a unit vector in the direction of the position vector by

$$\hat{e}_r = \hat{e}(\vec{r}) = \hat{r} = \frac{\vec{r}}{r} = \frac{x_i \hat{e}_i}{r} \rightarrow \vec{r} = r \hat{e}_r$$

This rule generalizes to an arbitrary vector as follows:

$$\vec{A} = A_i \hat{e}_i = A\hat{e}(\vec{A}) \rightarrow \hat{e}(\vec{A}) = \frac{\vec{A}}{A}$$

Vector Algebra

$$\vec{C} = \vec{A} + \vec{B} = \left(A_i + B_i\right)\hat{e}_i$$

$$\vec{C} = \lambda\vec{A} = \lambda A_i \hat{e}_i$$

$$\vec{0} = (0,0,0) = \text{zero or null vector}$$

$$-\vec{A} = (-A_i)\hat{e}_i = -A\hat{e}(\vec{A})$$

$$(-\vec{A}) + \vec{A} = \vec{0}$$

$$\hat{e}(-\vec{A}) = -\hat{e}(\vec{A})$$

Geometry of Space

For general vectors $\vec{A}$ and $\vec{B}$, the scalar product gives

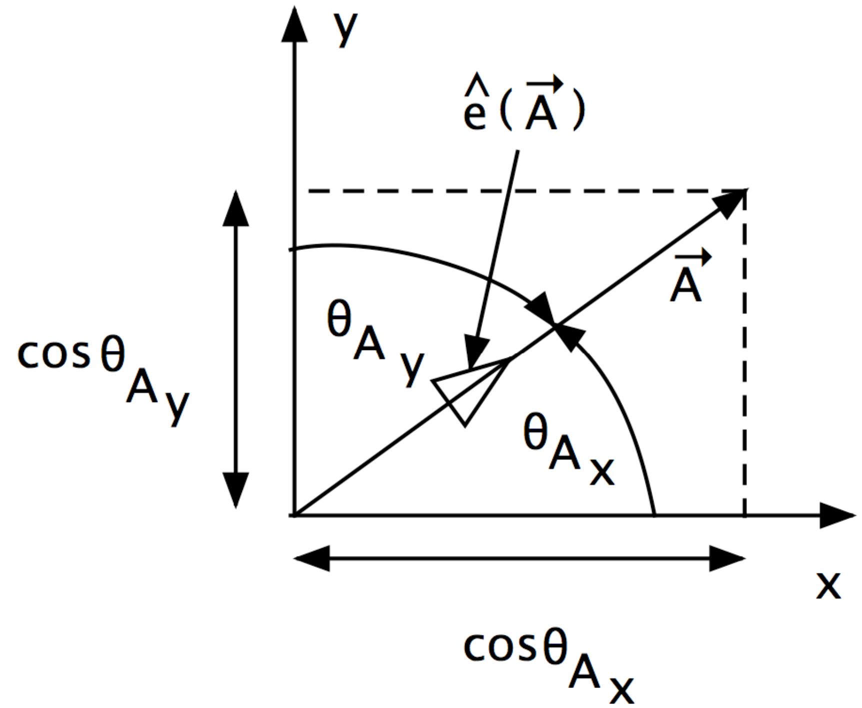
$$\vec{A} \cdot B = A_i B_j \hat{e}_i \cdot \hat{e}_j = A_i B_j \delta_{ij} = A_i B_i$$

and the length squared of a vector is $\vec{A} \cdot \vec{A} = A_i A_i$.

In 2-dimensions (completely general result because any 2 vectors form a 2-dimensional plane) we have

$$\hat{e}(\vec{A}) = \cos\theta_{A_x}\, \hat{e}_x + \sin\theta_{A_x}\, \hat{e}_y = \cos\theta_{A_x}\, \hat{e}_x + \cos\theta_{A_y}\, \hat{e}_y$$

as shown below.



These components are called the **direction cosines**. In general

$$\hat{e}(\vec{A}) \cdot \hat{e}_x = \cos\theta_{A_x} \quad \text{and so on}$$

Thus, we have

$$\hat{e}(\vec{A}) = \cos\theta_{A_x}\hat{e}_x + \cos\theta_{A_y}\hat{e}_y + \cos\theta_{A_z}\hat{e}_z$$

$$= (\hat{e}(\vec{A}) \cdot \hat{e}_i)\hat{e}_i$$

For a general vector

$$\vec{A} = A\hat{e}(\vec{A}) \rightarrow \vec{A} \cdot \hat{e}_i = i^{th} \text{ component} = A\hat{e}(\vec{A}) \cdot \hat{e}_i = A\cos\theta_{A_i}$$

This leads to the general rule

$$\vec{A} \cdot \vec{B} = AB\hat{e}(\vec{A}) \cdot \hat{e}(\vec{B}) = AB\cos\theta_{AB}$$

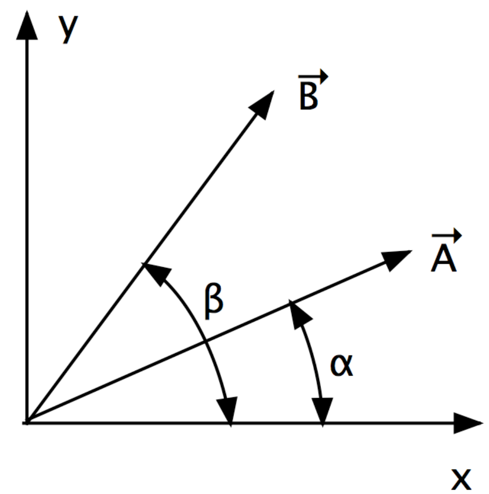
where

$$\theta_{AB} = \text{angle between } \vec{A} \text{ and } \vec{B}$$

So if two vectors are perpendicular or orthogonal, then their scalar product is equal to zero, $\theta_{AB} = 90°$.

Another way of writing this relation gives

$$\vec{A}\cdot\vec{B} = AB\cos(\beta-\alpha) = A_xB_x + A_yB_y = AB(\cos\alpha\cos\beta - \sin\alpha\sin\beta)$$



This implies that

$$\cos(\beta-\alpha) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$

which is a standard trigonometric identity.

Vector Product

We define the vector(cross) product of two vectors in terms of the standard cartesian unit vectors as follows:

$$\hat{e}_i \times \hat{e}_j = \varepsilon_{ijk}\hat{e}_k$$

resultant vector is perpendicular to plane of other two vectors

where ε_{ijk} is defined by

Levi-Civita symbol

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if ijk = even permutation of 123} \\ -1 & \text{if ijk = odd permutation of 123} \\ 0 & \text{otherwise} \end{cases}$$

This says that

$$\varepsilon_{123} = \varepsilon_{312} = \varepsilon_{231} = 1 \quad , \quad \varepsilon_{213} = \varepsilon_{132} = \varepsilon_{321} = -1$$

$$\varepsilon_{113} = \varepsilon_{112} = \varepsilon_{121} = \varepsilon_{131} = \varepsilon_{311} = \varepsilon_{211} = \varepsilon_{111} = \varepsilon_{221} = \varepsilon_{223} = \varepsilon_{212} = 0$$

$$\varepsilon_{232} = \varepsilon_{322} = \varepsilon_{122} = \varepsilon_{222} = \varepsilon_{331} = \varepsilon_{332} = \varepsilon_{313} = \varepsilon_{323} = \varepsilon_{133} = \varepsilon_{233} = \varepsilon_{333} = 0$$

Note: even permutation —> +1 and odd permutations —> -1

Let us work out some properties of vectors derivable from this vector product definition involving ε_{ijk}.
For two arbitrary vectors we have

$$\vec{A}\times\vec{B}=\varepsilon_{ijk}A_iB_j\hat{e}_k=\varepsilon_{ijk}A_jB_k\hat{e}_i$$

see lower right corner of slide

In 2-dimensions(for simplicity) using the diagram above we have

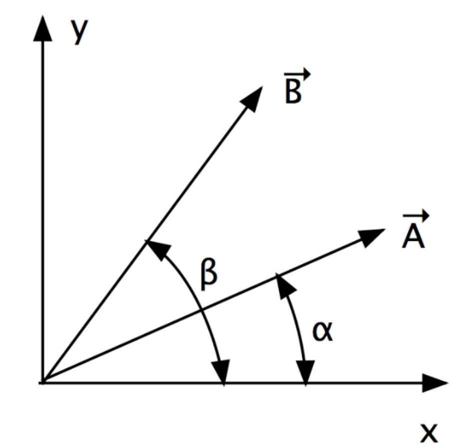
$$\vec{A}\times\vec{B}=(A_1\hat{e}_1+A_2\hat{e}_2)\times(B_1\hat{e}_1+B_2\hat{e}_2)=A_1B_1\hat{e}_1\times\hat{e}_1+A_1B_2\hat{e}_1\times\hat{e}_2+A_2B_1\hat{e}_2\times\hat{e}_1+A_2B_2\hat{e}_2\times\hat{e}_2$$
$$=(A_1B_2-A_2B_1)\hat{e}_1\times\hat{e}_2=(A_1B_2-A_2B_1)\hat{e}_3$$
$$=AB(\cos\alpha\sin\beta-\sin\beta\cos\alpha)\hat{e}_3=AB\sin(\beta-\alpha)\hat{e}_3$$

see figures below

So if two vectors are parallel, then their vector product is equal to zero. In general we have

$$\vec{A}\times\vec{B}=\varepsilon_{ijk}A_iB_j\hat{e}_k=\varepsilon_{123}A_1B_2\hat{e}_3+\varepsilon_{312}A_3B_1\hat{e}_2+\varepsilon_{231}A_2B_3\hat{e}_1+\varepsilon_{213}A_2B_1\hat{e}_3+\varepsilon_{132}A_1B_3\hat{e}_2+\varepsilon_{321}A_3B_2\hat{e}_1$$
$$=\left(A_2B_3-A_3B_2\right)\hat{e}_1+\left(A_3B_1-A_1B_3\right)\hat{e}_2+\left(A_1B_2-A_2B_1\right)\hat{e}_3$$

which is the standard form of the cross-product in 3-dimensions.

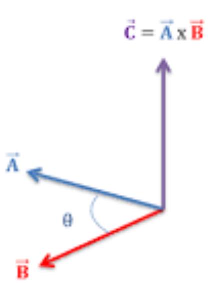


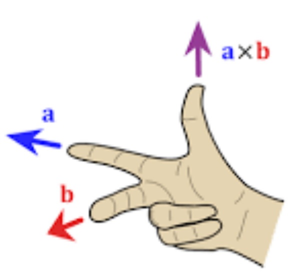
We note that the relation $\vec{A}\times\vec{B}=\varepsilon_{ijk}A_jB_k\hat{e}_i$ implies that $(\vec{A}\times\vec{B})_i=\varepsilon_{ijk}A_jB_k$.

Why should we introduce ε_{ijk} ?

(1) It is the easiest way to do complicated vector algebra in 3-dimensions
(2) For higher dimensions ε_{ijk} is the only way.
(3) It makes for a natural transition to other areas of mathematics (tensors) used in later courses.

Other Properties





$$\varepsilon_{mnk}\varepsilon_{ijk}=\delta_{mi}\delta_{nj}-\delta_{ni}\delta_{mj}$$
$$\varepsilon_{mjk}\varepsilon_{ijk}=2\delta_{mn}$$
$$\varepsilon_{ijk}\varepsilon_{ijk}=6$$

dummy indices inside summations

$$\varepsilon_{ijk}A_iB_j\hat{e}_k=\varepsilon_{jki}A_jB_k\hat{e}_i$$
$$=\varepsilon_{ijk}A_jB_k\hat{e}_i$$

since $\varepsilon_{ijk}=\varepsilon_{jki}$

Example of use in complex vector identity:

$$\vec{A}\times(\vec{B}\times\vec{C})=\vec{A}\times\left(\varepsilon_{ijk}B_jC_k\hat{e}_i\right)=\varepsilon_{ijk}B_jC_k\vec{A}\times\hat{e}_i=\varepsilon_{ijk}B_jC_k\varepsilon_{mnp}A_n(\hat{e}_i)_p\hat{e}_m=\varepsilon_{ijk}B_jC_k\varepsilon_{mnp}A_n\delta_{ip}\hat{e}_m$$

$$=\varepsilon_{ijk}B_jC_k\varepsilon_{mni}A_n\hat{e}_m=\varepsilon_{jki}\varepsilon_{mni}B_jC_kA_n\hat{e}_m=\left(\delta_{mj}\delta_{nk}-\delta_{mk}\delta_{nj}\right)B_jC_kA_n\hat{e}_m$$

$$=B_jC_kA_k\hat{e}_j-B_jC_kA_j\hat{e}_k=\left(B_j\hat{e}_j\right)\left(A_kC_k\right)-\left(C_k\hat{e}_k\right)\left(A_jB_j\right)$$

$$=\vec{B}(\vec{A}\cdot\vec{C})-\vec{C}(\vec{A}\cdot\vec{B})$$

go through step by step —>
after much practice this all becomes
second nature!

We also note that we can always write any vector in the following way:

$$\hat{e}=\text{arbitrary unit vector}$$

$$\vec{A}=\text{arbitrary vector}$$

$$\hat{e}\times\left(\vec{A}\times\hat{e}\right)=\vec{A}(\hat{e}\cdot\hat{e})-\hat{e}(\vec{A}\cdot\hat{e})=\vec{A}-\hat{e}(\vec{A}\cdot\hat{e})$$

$$\longrightarrow\quad\vec{A}=\hat{e}(\vec{A}\cdot\hat{e})+\hat{e}\times\left(\vec{A}\times\hat{e}\right)$$

Complex Numbers

The crucial quantity here is the number i. It has the property

$$i^2=-1,i^3=-i,i^4=+1,i^5=i,.........$$

We represent a complex number z as a 2-component object in a cartesian basis as

$$z=(x,y)=x+iy=\text{Real}(z)+i\text{Imag}(z)$$

Alternatively, we can represent it with a plane-polar basis as

$$x=r\cos\theta\ ,\ y=r\sin\theta\quad,\quad 0\le r\le\infty\ ,-\pi\le\theta\le\pi$$

$$r=\left(x^2+y^2\right)^{1/2}\quad,\quad\theta=\tan^{-1}\frac{y}{x}$$

$$z=r\left(\cos\theta+i\sin\theta\right)$$

Now using $\sin\theta = \theta - \frac{\theta^3}{3!} + \ldots. \quad , \quad \cos\theta = 1 - \frac{\theta^2}{2!} + \ldots.$ see next slide (added)

we get

$$z = r(\cos\theta + i\sin\theta) = r\left(1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \ldots..\right) = re^{i\theta}$$

This generalizes to let us write

$$\left(1 + im\theta + \frac{(im\theta)^2}{2!} + \frac{(im\theta)^3}{3!} + \ldots..\right) = e^{im\theta} = (\cos m\theta + i\sin m\theta)$$

We can also write

$$e^{im\theta} = (\cos m\theta + i\sin m\theta) = \left(e^{i\theta}\right)^m = (\cos m\theta + i\sin m\theta)^m$$

which can be used to get expressions for $\cos m\theta$, $\sin m\theta$, etc, as

$$\cos 2\theta + i\sin 2\theta = (\cos\theta + i\sin\theta)^2 = \cos^2\theta - \sin^2\theta + 2i\sin\theta\cos\theta$$

$$\cos 2\theta = \cos^2\theta - \sin^2\theta + 2i\sin\theta\cos\theta$$

$$\sin 2\theta = 2\sin\theta\cos\theta$$

We define the complex conjugate of z by

$$z^* = x - iy = re^{-i\theta} = r(\cos\theta - i\sin\theta)$$

$$zz^* = z^*z = r^2 = |z|^2$$

Note a few special results:

$$e^{2\pi ni} = 1 \rightarrow e^{i\theta} = e^{i(\theta+2\pi n)} \quad , \quad n = \text{integer}$$

real definition! $e^{\pi i} = -1 \quad , \quad e^{i\pi/2} = i$

Using the standard definitions of some mathematical functions we get:

$$z^{1/2} = \sqrt{z} = \left(re^{i\theta}\right)^{1/2} = \sqrt{r}e^{i\theta/2} = \left(re^{i(\theta+2n\pi)}\right)^{1/2} = \sqrt{r}e^{i(\theta/2+n\pi)} \quad \text{n=integer}$$

$$\log z = \log\left(re^{i\theta}\right) = \log r + i\theta = \log\left(re^{i(\theta+2n\pi)}\right) = \log r + i(\theta + 2n\pi) \quad \text{n=integer}$$

Definition: Power series:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots = \sum_{k=0}^{\infty} a_k x^k \qquad a_k = \left. \frac{d^k f(x)}{dx^k} \right|_{x=0}$$

Examples(the real definitions):

$$e^{\alpha x} = \sum_{k=0}^{\infty} \frac{1}{k!} \alpha^k x^k$$

$$\sin \alpha x = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} (-1)^k \alpha^{2k+1} x^{2k+1}$$

$$\cos \alpha x = \sum_{k=0}^{\infty} \frac{1}{(2k)!} (-1)^k \alpha^{2k} x^{2k}$$

More functions:
$$z_1 z_2 = r_1 r_2 e^{i(\theta_1+\theta_2)} \quad , \quad \frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}$$
$$z^{1/n} = r^{1/n} e^{i\theta/n} = r^{1/n}(\cos(\theta/n) + i\sin(\theta/n))$$
$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} \quad , \quad \cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$$

This last expression also works for complex angles
$$\cos(iy) = \frac{e^{-y} + e^{y}}{2} = \cosh y \to \cos(i) = \frac{e^{-1} + e^{1}}{2} = 1.54$$
$$\sin(iy) = \frac{e^{-y} - e^{y}}{2i} = i\sinh y$$

Complex Roots and Powers:
$$a^b = e^{\ell n a^b} = e^{b \ell n a}$$
$$\to i^{-2i} = e^{-2i\ell n i} = e^{-2i \ell n e^{i(\pi/2 \pm 2n\pi)}} e^{-2i(\ell n(1) + i(\pi/2 \pm 2n\pi))} = e^{\pi \pm 4n\pi}$$
$$= e^{\pi \pm 4n\pi}, e^{5\pi}, e^{-3\pi}, e^{9\pi}, \ldots = 23.14, \ldots\ldots\ldots$$

WOW!

Derivative

Given a function f(x) of one variable x, the derivative is defined as
$$\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$
We know that the derivative at a point is the slope, i.e., the rate of change of the function with respect to x. In other words, we can predict the change in the value of the function f(x) in the immediate vicinity of a point x_0, provided we have the derivative of f at that point:

$$df = \left(\frac{df}{dx}\right)_{x_0} dx = \left[\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}\right] dx$$

What if the function f has two variables, x and y, i.e., f=f(x,y)? Given a point (x_0,y_0), we would like to know the change when we simultaneously change x by dx and y by dy. To compute this change, we first keep the variable y fixed at the point y_0 and let x change by dx; the change in f, df, is given by

$$df = \left[\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}\right] dx \equiv \frac{\partial f}{\partial x} dx$$

Similarly, if we keep x fixed at x_0 and we change y by dy, the change in f is given by

$$df = \left[\lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}\right] dy \equiv \frac{\partial f}{\partial y} dy$$

These are called "partial" derivatives.

Example 1: $f(x,y) = xy \rightarrow \frac{\partial f}{\partial x} = y \quad and \quad \frac{\partial f}{\partial y} = x$

Example 2: $f(x,y) = 3x^2 y + 2y^3 \rightarrow \frac{\partial f}{\partial x} = 6xy \quad and \quad \frac{\partial f}{\partial y} = 3x^2 + 6y^2$

We now ask what happens when both x and y change by dx and dy, respectively:

$$\Delta f(x,y) = f(x + \Delta x, y + \Delta y) - f(x,y) = f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y) + f(x + \Delta x, y) - f(x,y)$$

Taking the limit as $\Delta x \rightarrow 0$ and $\Delta y \rightarrow 0$ we get

$$df(x,y) = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy$$

This result easily generalizes to any number of cartesian coordinates, i.e.,

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz$$

Vector Functions

Let us look at a vector function $\vec{A}(p)$ which depends on only one variable p. What is the change in this function as p changes by dp? This vector function is nothing more than three independent functions

$$A_x(p), A_y(p), A_z(p)$$

Thus,

$$\begin{aligned} d\vec{A} &= \vec{A}(p+dp) - \vec{A}(p) \\ &= \left(A_x(p+dp) - A_x(p)\right)\hat{x} + \left(A_y(p+dp) - A_y(p)\right)\hat{y} + \left(A_z(p+dp) - A_z(p)\right)\hat{z} \\ &= dA_x\hat{x} + dA_y\hat{y} + dA_z\hat{z} \end{aligned}$$

Now since $A_x(p), A_y(p), A_z(p)$ are scalar functions of only one variable, we have that

$$dA_i = \frac{\partial A_i}{\partial p} dp \quad , \quad i = 1,2,3$$

and we finally get

$$d\vec{A} = \left(\frac{\partial A_x}{\partial p}\hat{x} + \frac{\partial A_y}{\partial p}\hat{y} + \frac{\partial A_z}{\partial p}\hat{z} \right) dp$$

For example, suppose the position vector of a particle is given by

$$\vec{r} = 3t^2\hat{x} + 6t\hat{y} + (3 - 4t)\hat{z}$$

then the change in the position vector due to an infinitesimal change in time dt is then given by

$$d\vec{r} = \left(6t\hat{x} + 6\hat{y} - 4\hat{z}\right) dt$$

What happens if the vector function is a function of several variables? We illustrate what happens with a 3-dimensional vector function of two variables, say p and q.

$$\vec{A}(p,q) = A_x(p,q)\hat{x} + A_y(p,q)\hat{y} + A_z(p,q)\hat{z}$$

$$d\vec{A} = dA_x(p,q)\hat{x} + dA_y(p,q)\hat{y} + dA_z(p,q)\hat{z}$$

where

$$dA_i(p,q) = \frac{\partial A_i}{\partial p}dp + \frac{\partial A_i}{\partial q}dq$$

We now know how to compute the variation of any function, scalar or vector, in any number of Cartesian dimensions, with respect to the parameters the function depends on.

A very interesting - and physically meaningful - case is when one has a vector function that depends on position. An example is the gravitational field. A mass M creates a field around it. Any mass brought close to M will thus experience a force. We define the field so that it is dependent only on parameters associated with the source. In other words,

$$\vec{E}_G = -\frac{GM}{r^2}\hat{r}$$

a **field** is a physical quantity, represented by a number or another tensor, that has a value for each point in space and time

A field can be classified as a scalar field, a vector field, a spinor field or a tensor field according to whether the represented physical quantity is a scalar, a vector, a spinor, or a tensor, respectively.

Given the field, the force on any mass m is given by

we will see how later

$$\vec{F}_G = m\vec{E}_G = -\frac{GMm}{r^2}\hat{r}$$

In Cartesian coordinates, the field is written as

$$\vec{E}_G = -\frac{GM}{\left(x^2+y^2+z^2\right)^{3/2}}\left(x\hat{x}+y\hat{y}+z\hat{z}\right)$$

This is a field which depends on three variables x,y,z. The change in the field in the neighborhood of (x,y,z), i.e., when

$$x \rightarrow x+dx, y \rightarrow y+dy, z \rightarrow z+dz$$

is then

$$d\vec{E}_G = \left(dE_x\hat{x}+dE_y\hat{y}+dE_z\hat{z}\right) = \left(\frac{\partial E_x}{\partial x}\hat{x}+\frac{\partial E_y}{\partial y}\hat{y}+\frac{\partial E_z}{\partial z}\hat{z}\right)$$

where

$$dE_x = -GM\frac{\left(y^2+z^2-2x^2\right)dx - 3xydy - 3xzdz}{\left(x^2+y^2+z^2\right)^{5/2}}$$

$$dE_y = -GM\frac{-3xydx + \left(x^2+z^2-2y^2\right)dy - 3yzdz}{\left(x^2+y^2+z^2\right)^{5/2}}$$

$$dE_z = -GM\frac{-3xzdx - 3yzdy + \left(x^2+y^2-2z^2\right)dz}{\left(x^2+y^2+z^2\right)^{5/2}}$$

gets complicated quickly

The Gradient Operator

An operator is a "thing" that acts on an object. In this sense, the derivative operator is equal to

$$\frac{d}{dx}$$

Given any function f(x), this operator acts on the function to give a new function, the derivative of f(x). Thus,

$$\frac{d}{dx}f(x) \equiv \frac{df}{dx}$$

In the same way, we can think of the differential of a function of three variables as the dot product of two vectors: the vector

$$\vec{C} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$$

and the vector

$$d\vec{r} = (dx, dy, dz)$$

Then

$$df = \vec{C} \cdot d\vec{r}$$

In other words, for any function f(x,y,z) we can define the vector

$$\vec{C} = \hat{x}\frac{\partial f}{\partial x} + \hat{y}\frac{\partial f}{\partial y} + \hat{z}\frac{\partial f}{\partial z} = \left(\hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z}\right) f$$

We define the new operator to be equal to the quantity that acts on f:

$$\nabla = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z}$$

It is called the **gradient** operator. When it acts on a function, it gives a vector whose direction is the direction of maximum change of the function. Its magnitude is equal to the rate of change of the function in this direction of maximum change, with respect to changes in x, y, and z.

Example: a fly is in a room whose temperature is given by

$$T(x,y,z) = \frac{z}{x^2 + y^2 + \varepsilon}$$

i.e., the air is warmer as we go up towards the ceiling and warmer as we approach the point (0,0), which is the middle of the room. The parameter ε just prevents the temperature from diverging in the middle of the room. In which direction should the fly move in order to maximize the change of temperature it will experience? This direction is given by the gradient of the temperature:

$$\nabla T = \left(\hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z}\right)\frac{z}{x^2 + y^2 + \varepsilon}$$

$$= -\frac{2xz}{\left(x^2 + y^2 + \varepsilon\right)^2}\hat{x} - \frac{2yz}{\left(x^2 + y^2 + \varepsilon\right)^2}\hat{y} + \frac{1}{x^2 + y^2 + \varepsilon}\hat{z}$$

As expected, this direction is towards the center of the room in the x-y plane and upwards in the z-direction.

Some Properties of the Gradient:

It looks like a vector

It works like a vector

But it's not a real vector because it's meaningless by itself: it's an operator.

How it works:

It can act on both scalar and vector functions:

Acting on a scalar function:

∇f = a vector = gradient(f) = grad(f)

Acting on a vector function with dot product: divergence

$\nabla \cdot \vec{f}$ = a scalar = divergence($\vec{f}$) = div($\vec{f}$)

Acting on a vector function with cross product:

$\nabla \times \vec{f}$ = a vector = curl($\vec{f}$)

Divergence

Given a vector function

$$\vec{f}(x,y,z) = f_x(x,y,z)\hat{x} + f_y(x,y,z)\hat{y} + f_z(x,y,z)\hat{z}$$

we define its divergence as:

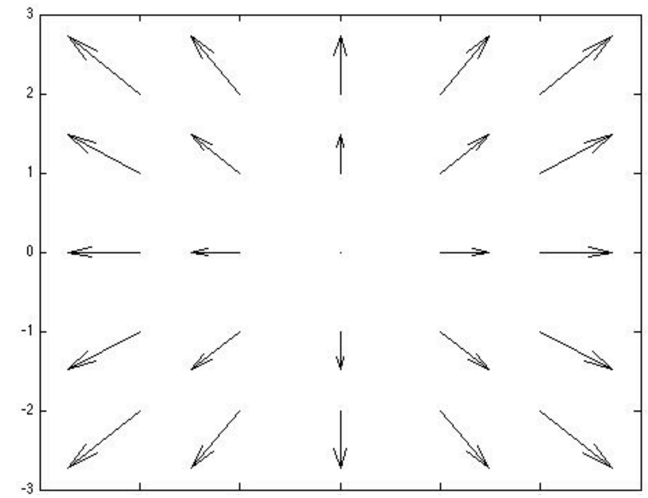
$$div(\vec{f}) = \nabla \cdot \vec{f} = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z}$$

Observations: The divergence is a scalar(dot product).

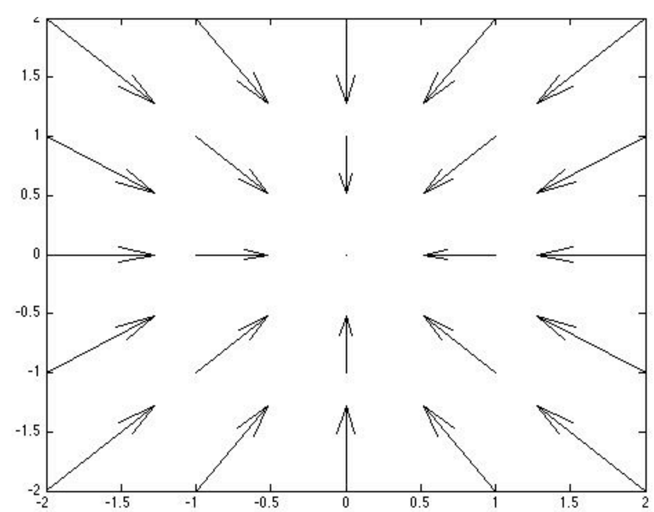
Geometrical interpretation: it measures how much the function "spreads around a point".

We can calculate the divergence and illustrate the corresponding gradient for some functions:

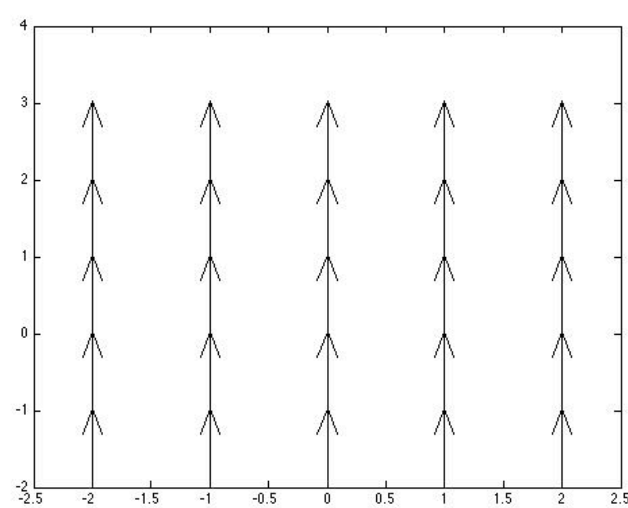
$$\vec{f}(x,y,z) = x\hat{x} + y\hat{y} + z\hat{z} \rightarrow div(\vec{f}) = 3 = a \text{ "faucet"}$$



$$\vec{f}(x,y,z) = -\hat{x} - y\hat{y} - z\hat{z} \rightarrow div(\vec{f}) = -3 = \text{ a "sink"}$$



$$\vec{f}(x,y,z) = \hat{y} \rightarrow div(\vec{f}) = 0$$



The plots (only in x-y plane) were illustrated using the **Octave** code below:

```
x=-2:1:2;
y=-2:1:2;
[X,Y]=meshgrid(x,y);
Z1=0.5*(X.^2+Y.^2);
[px,py]=gradient(Z1,1,1);
quiver(X,Y,px,py,'k');
Z1=-0.5*(X.^2+Y.^2);
[px,py]=gradient(Z1,1,1);
figure
quiver(X,Y,px,py,'k');
Z1=Y;
[px,py]=gradient(Z1,1,1);
figure
quiver(X,Y,px,py,'k');
```

one18.m

Is anyone interested in learning OCTAVE?

As we will see this corresponds to

div E <0 for –charge: sink
div E>0 for + charge: faucet

Curl

Given a vector function

$$\vec{f}(x,y,z) = f_x(x,y,z)\hat{x} + f_y(x,y,z)\hat{y} + f_z(x,y,z)\hat{z}$$

Using

$$\vec{A}\times\vec{B} = \varepsilon_{ijk}A_iB_j\hat{e}_k = \left(A_2B_3 - A_3B_2\right)\hat{e}_1 + \left(A_3B_1 - A_1B_3\right)\hat{e}_2 + \left(A_1B_2 - A_2B_1\right)\hat{e}_3$$

we define its curl as:

$$curl(\vec{f}) = \nabla\times\vec{f} = \varepsilon_{ijk}\nabla_i f_j\hat{e}_k = \left(\nabla_2 f_3 - \nabla_3 f_2\right)\hat{e}_1 + \left(\nabla_3 f_1 - \nabla_1 f_3\right)\hat{e}_2 + \left(\nabla_1 f_2 - \nabla_2 f_1\right)\hat{e}_3$$

$$= \left(\frac{\partial f_z}{\partial y} - \frac{\partial f_y}{\partial z}\right)\hat{x} + \left(\frac{\partial f_x}{\partial z} - \frac{\partial f_z}{\partial x}\right)\hat{y} + \left(\frac{\partial f_y}{\partial x} - \frac{\partial f_x}{\partial y}\right)\hat{z}$$

We note that this expression can also be written using a determinant as:

$$curl(\vec{f}) = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ f_x & f_y & f_z \end{vmatrix}$$

Observations: The curl is a vector(cross product).

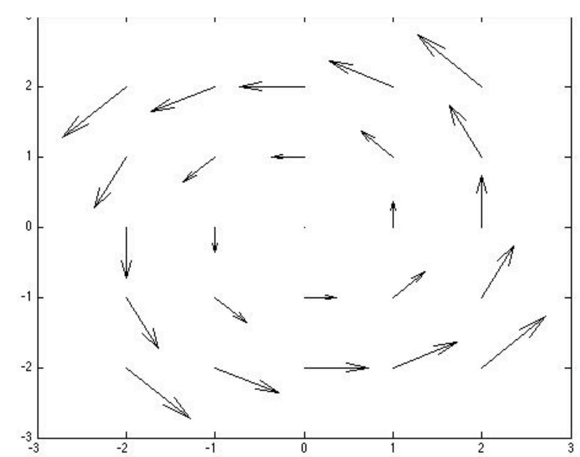
Geometrical interpretation: it measures how much the function "curls around a point".

Calculate the curl for the following function:

$$\vec{f}(x,y,z) = -y\hat{x} + x\hat{y}$$

$$curl(\vec{f}) = 2\hat{z}$$

This function has a non-zero curl. It is a vortex as shown below:

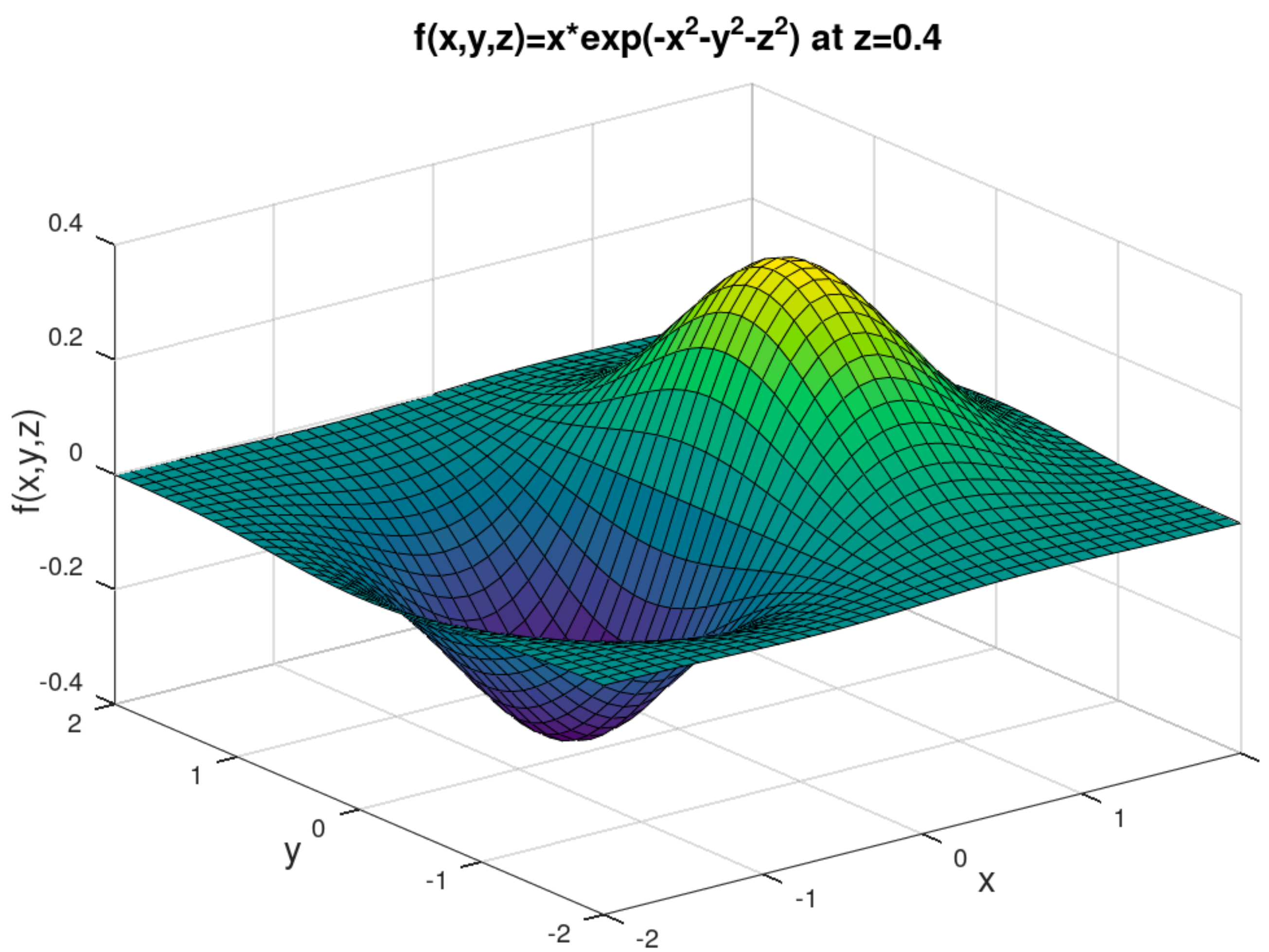


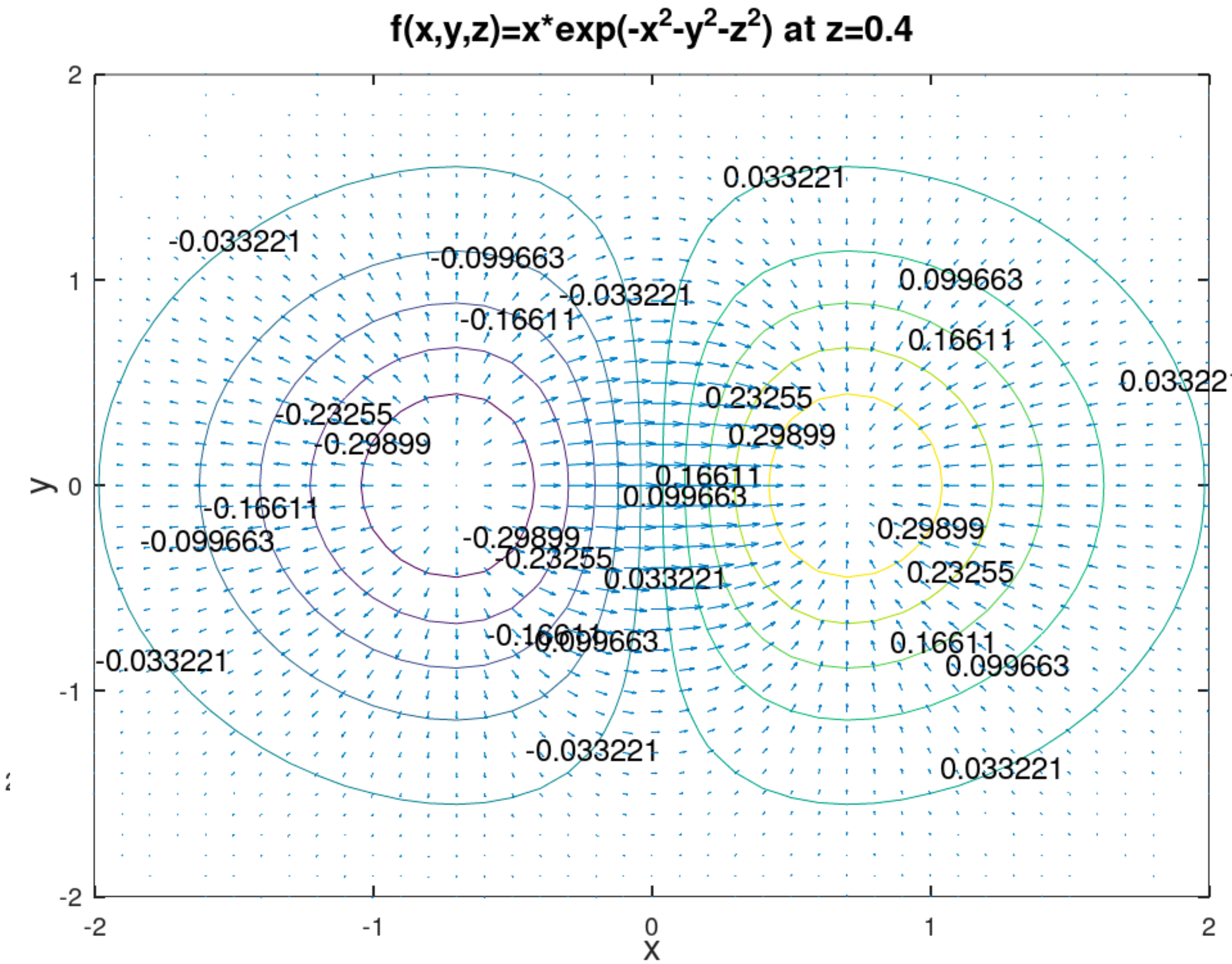
The figure was produced by the Octave code below:

```
x=-2:1:2;
y=-2:1:2;
[X,Y]=meshgrid(x,y);
px=-Y;
py=X;
quiver(X,Y,px,py,'k');
```

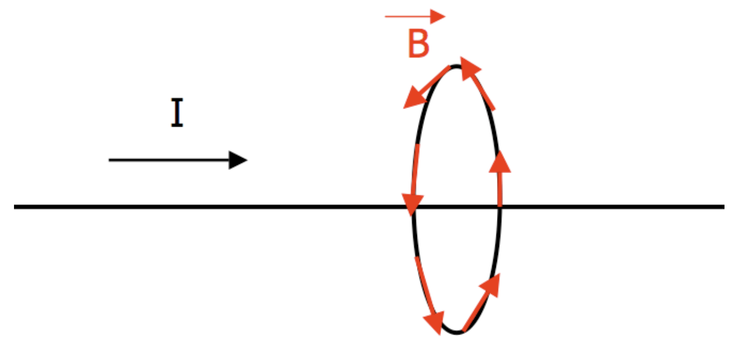
one19.m

gradient_ex.m





Example: As we will see later, the magnetic field around a current carrying wire looks like the figure below:



Clearly,

$$curl(\vec{B}) = \nabla \times \vec{B} \neq 0$$

Another vector identity:

$$\nabla \times \left(\nabla \times \vec{B}\right) = \nabla \times \left(\varepsilon_{ijk} \nabla_j B_k \hat{e}_i\right) = \left(\nabla \times \hat{e}_i\right)\left(\varepsilon_{ijk} \nabla_j B_k\right) = \left(\varepsilon_{lmn} \nabla_m \left(\hat{e}_i\right)_n \hat{e}_l\right)\left(\varepsilon_{ijk} \nabla_j B_k\right)$$

$$= \left(\varepsilon_{lmn} \nabla_m \delta_{in} \hat{e}_l\right)\left(\varepsilon_{ijk} \nabla_j B_k\right) = \left(\varepsilon_{lmi} \varepsilon_{ijk} \nabla_m \nabla_j B_k \hat{e}_l\right) = \varepsilon_{ilm} \varepsilon_{ijk} \nabla_m \nabla_j B_k \hat{e}_l$$

$$= \left(\delta_{lj} \delta_{mk} - \delta_{lk} \delta_{mj}\right) \nabla_m \nabla_j B_k \hat{e}_l = \nabla_k \nabla_j B_k \hat{e}_j - \nabla_j \nabla_j B_k \hat{e}_k = \nabla\left(\nabla \cdot \vec{B}\right) - \nabla^2 \vec{B}$$

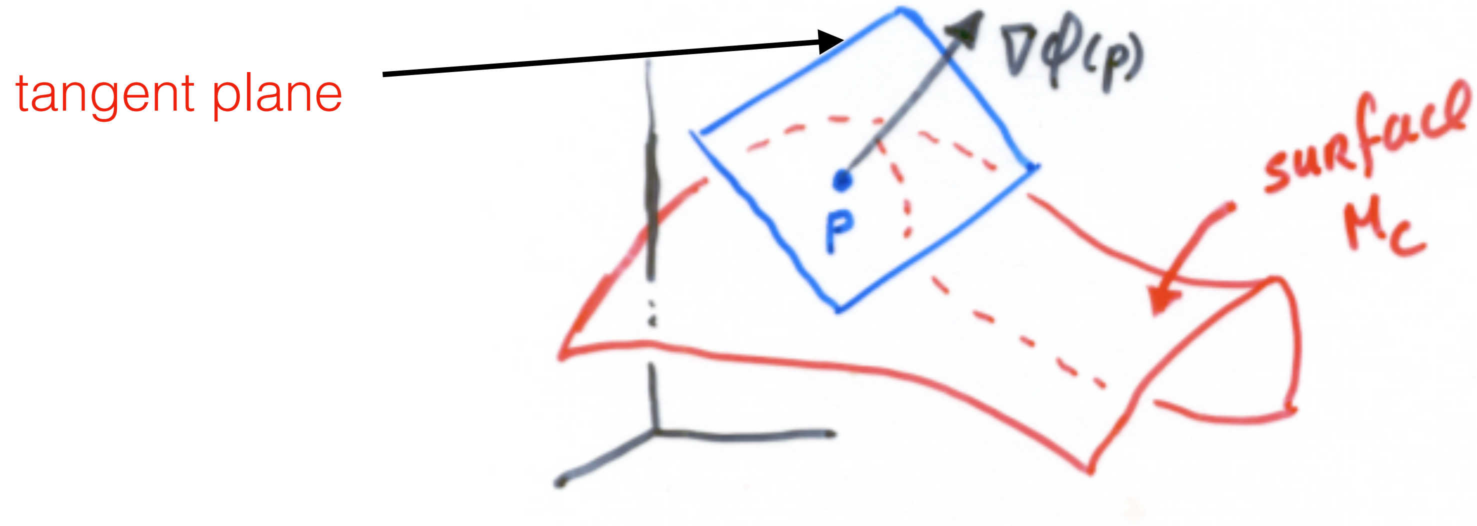
More Details

Let us look at the physical interpretation of Grad, Div and Curl.

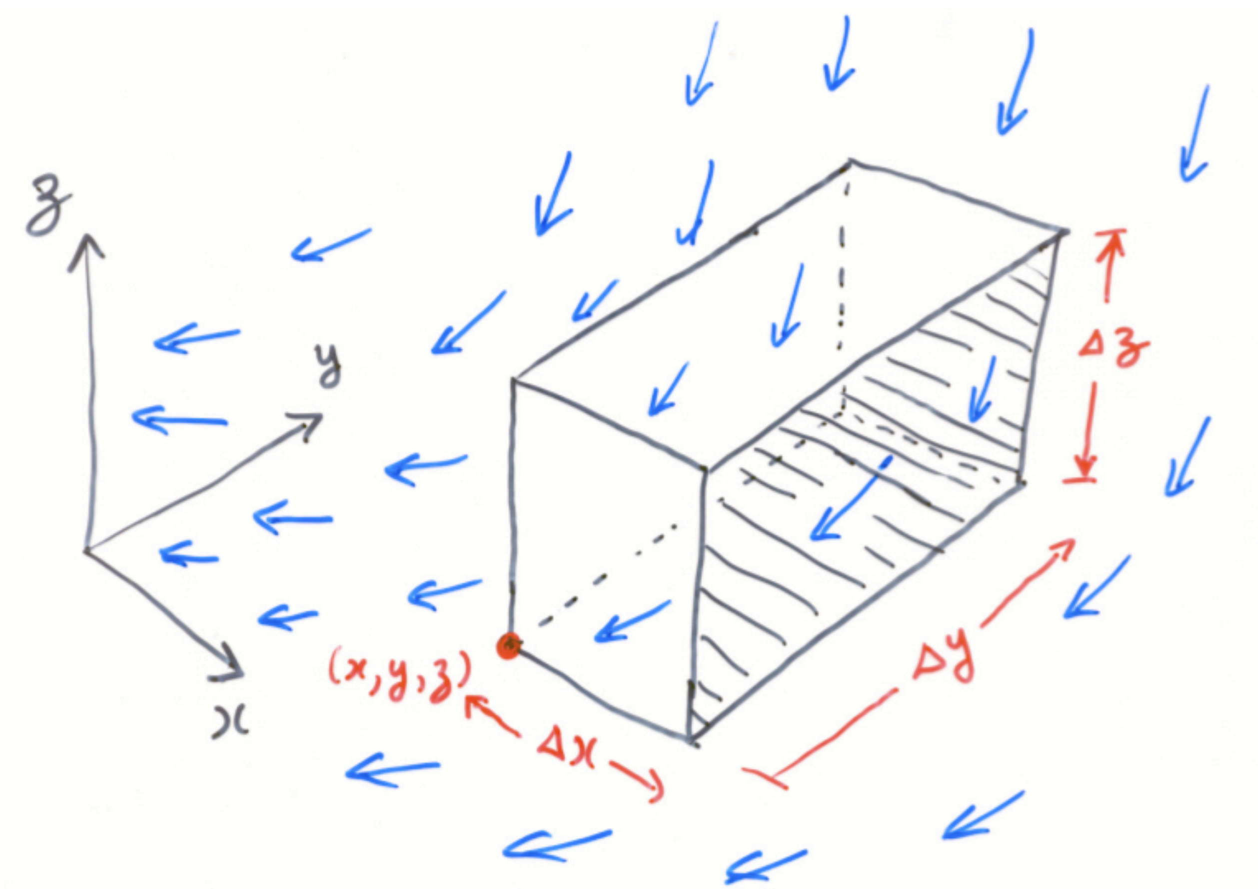
Grad: Consider the set of all points p such that

$$f(p) = c = \text{constant}$$

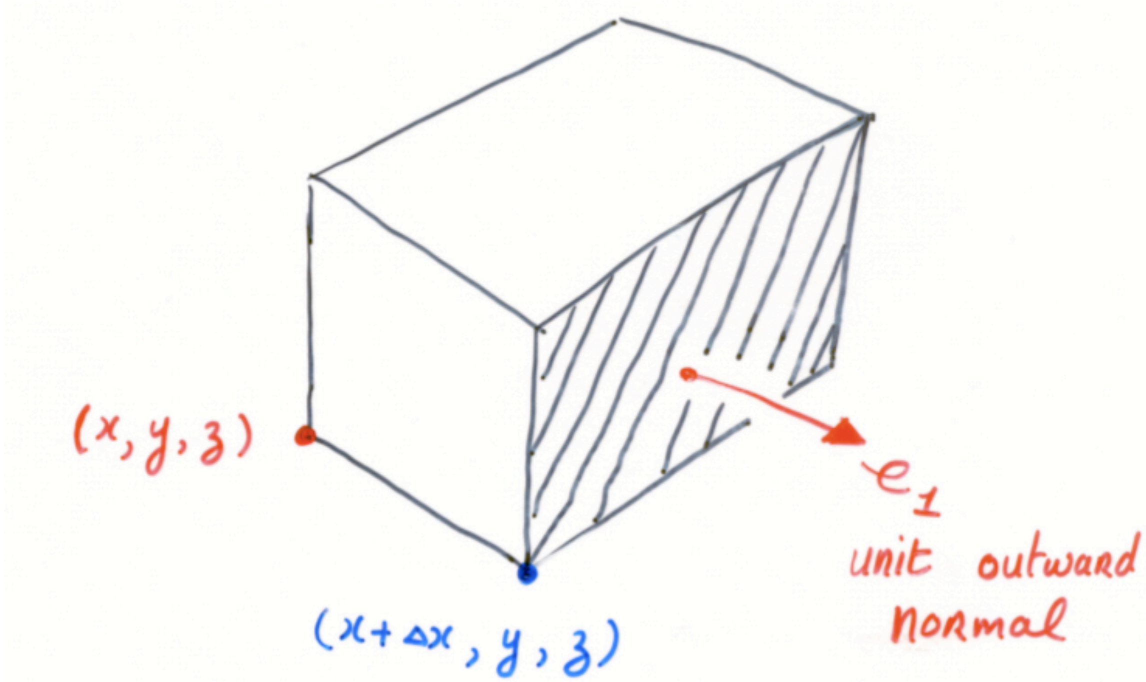
This set forms a surface M_c such that $\nabla f(p)$ is perpendicular to M_c at all p as shown below.



Div: Consider a vector field $\vec{f}(p)$ which is the velocity vector of a fluid moving in three dimensions.



Let us calculate the rate at which fluid is flowing out of the box per unit volume. Look at the shaded side



We approximate the velocity(vector) on this (small) shaded side by $\vec{f}(x+\Delta x, y, z)$. Therefore, the component of the velocity out over the shaded surface is $\hat{e}_1 \cdot \vec{f}(x+\Delta x, y, z) = f_1(x+\Delta x, y, z)$.

Hence, the rate at which fluid flows out over this shaded surface is

$f_1(x+\Delta x, y, z)\Delta y \Delta z$ = (speed in $\hat{e}_1$ direction)x(area)

Similarly, the rate at which fluid flows out over the back (in the $-\hat{e}_1$ direction) is

$$-f_1(x, y, z)\Delta y \Delta z$$

Therefore, the rate at which fluid flows out in the $\hat{e}_1$ direction per unit volume is

$$\frac{f_1(x+\Delta x, y, z)\Delta y \Delta z - f_1(x, y, z)\Delta y \Delta z}{\Delta x \Delta y \Delta z} = \frac{f_1(x+\Delta x, y, z) - f_1(x, y, z)}{\Delta x}$$

In the limit $\Delta x \to 0$, this becomes

$$\frac{\partial f_1(x, y, z)}{\partial x}$$

Thus, the rate at which fluid leaves (x,y,z) per unit volume in any direction is

$$\frac{\partial f_1(x, y, z)}{\partial x} + \frac{\partial f_2(x, y, z)}{\partial y} + \frac{\partial f_3(x, y, z)}{\partial z} = div(\vec{f}) = \nabla \cdot \vec{f}$$

Conclusion:

$$\nabla \cdot \vec{f}(x,y,z) = \begin{cases} \text{the rate at which "fluid"} \\ \text{is being created at (x,y,z)} \\ \text{per unit volume} \end{cases}$$

Curl:

Motivation: Imagine a rigid body rotating about the origin. Fix some particle p on this rigid body and let

$$\vec{r}(p,t) = \begin{cases} \text{the position vector} \\ \text{of the particle p} \\ \text{at time t} \end{cases}$$

Therefore,

$$\frac{d\vec{r}(p,t)}{dt} = \vec{v}(p,t) \begin{cases} \text{the velocity vector of} \\ \text{the particle p which is} \\ \text{at position } \vec{r}(p,t) \text{ at time t} \end{cases}$$

i.e., $\vec{v}(\vec{r}(p,t),t) = \dot{\vec{r}}(p,t)$

Fact: For each time t there exists a unique vector $\vec{\omega}(t)$, which depends on time only, such that

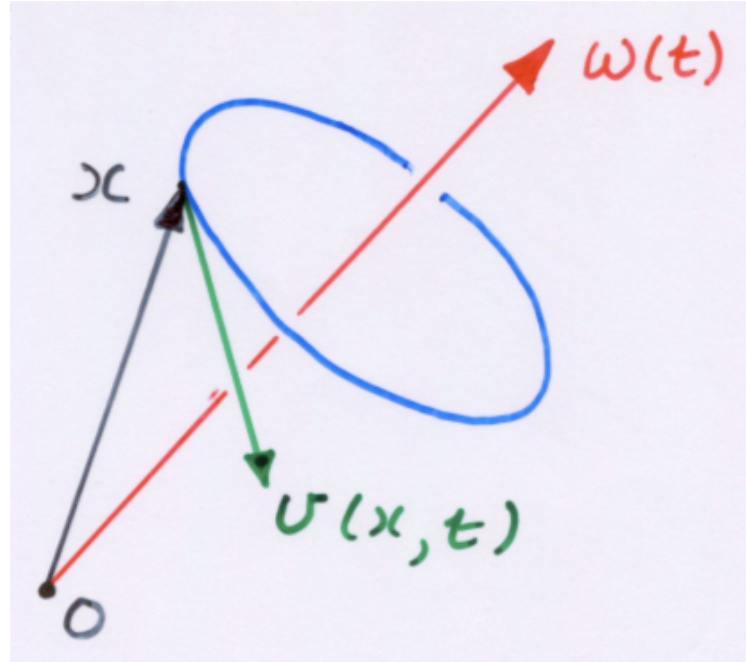
$$\vec{v}(\vec{r}(p,t),t) = \vec{\omega}(t) \times \vec{r}(p,t)$$

which we write

$$\vec{v}(\vec{r},t) = \vec{\omega}(t) \times \vec{r}$$

The vector $\vec{\omega}(t)$ is called the angular velocity vector of the rigid rotation at time t.

The direction of $\vec{\omega}(t)$ is called the instantaneous axis of rotation at time t. See figure below



Now keep t fixed. We get a vector field called the velocity field of the rotation at time t. It is given by

$$\vec{v}(\vec{r},t) = \vec{\omega}(t)\times\vec{r}$$

We then find

$$\frac{\partial\vec{v}(\vec{r},t)}{\partial x_j} = \frac{\partial\vec{\omega}(t)}{\partial x_j}\times\vec{r} + \vec{\omega}(t)\times\frac{\partial\vec{r}}{\partial x_j} = \frac{\partial\vec{\omega}(t)}{\partial x_j}\times\vec{r} + \vec{\omega}(t)\times\frac{\partial\sum_k x_k\hat{e}_k}{\partial x_j}$$

$$= \frac{\partial\vec{\omega}(t)}{\partial x_j}\times\vec{r} + \vec{\omega}(t)\times\sum_k\frac{\partial x_k}{\partial x_j}\hat{e}_k = \frac{\partial\vec{\omega}(t)}{\partial x_j}\times\vec{r} + \vec{\omega}(t)\times\sum_k\delta_{kj}\hat{e}_k$$

$$= \frac{\partial\vec{\omega}(t)}{\partial x_j}\times\vec{r} + \vec{\omega}(t)\times\hat{e}_j = \vec{\omega}(t)\times\hat{e}_j$$

since

$$\frac{\partial\vec{\omega}(t)}{\partial x_j} = 0$$

We now can show the following result:

$$\nabla\times\vec{v} = \varepsilon_{ijk}\frac{\partial v_k}{\partial x_j}\hat{e}_i = \varepsilon_{ijk}\left(\vec{\omega}\times\hat{e}_j\right)_k\hat{e}_i = \varepsilon_{ijk}\left(\varepsilon_{lmn}\omega_m\left(\hat{e}_j\right)_n\hat{e}_l\right)_k\hat{e}_i$$

$$= \varepsilon_{ijk}\varepsilon_{lmn}\omega_m\delta_{jn}\delta_{lk}\hat{e}_i = \varepsilon_{ijk}\varepsilon_{kmj}\omega_m\hat{e}_i = 2\delta_{im}\omega_m\hat{e}_i = 2\omega_i\hat{e}_i = 2\vec{\omega}$$

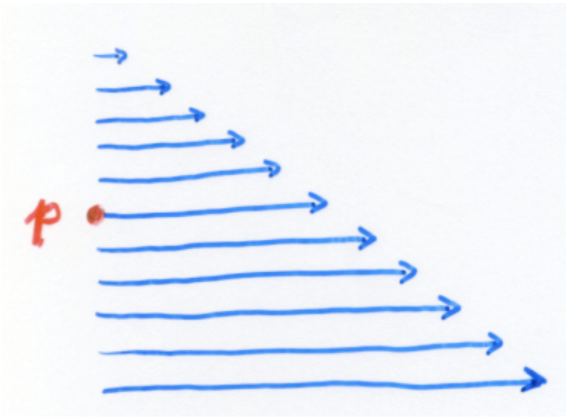
That is an amazing demonstration of the power of Einstein summation convention and the ε_{ijk} symbol.

This says that:

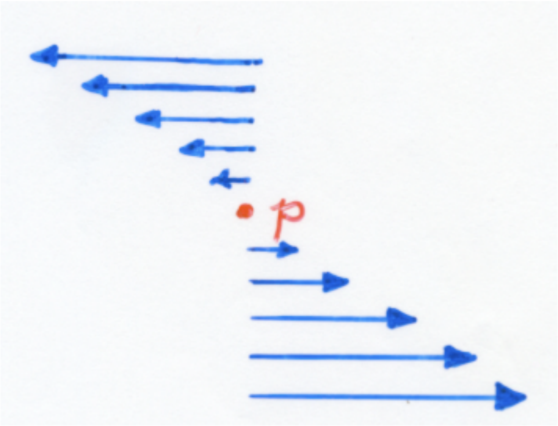
The curl of the velocity field of a rigid rotation about the origin is twice the angular velocity vector of that rotation.

In the case of an arbitrary vector field we can think of $\vec{f} = \vec{f}(p)$ as giving the velocity vector of a particle p of the fluid which is at $\vec{r}(p)$ at some fixed time.

Now imagine that an observer on p is moving along with the fluid. We get two pictures shown below.



The velocity field of the fluid from the point of view of an external observer



The velocity field of the fluid from the point of view of an observer moving along with p

Note that p = origin from the point of view of an observer moving along with p.

Now, suppose this observer moving along with p is a small ball. Then the motion of the fluid will cause this ball to rotate rigidly about p and

$$\left.\begin{array}{c}\text{the velocity field}\\ \text{of the fluid } \vec{v}\\ \text{near p}\end{array}\right\}\left\{\begin{array}{c}\text{the velocity field}\\ \text{of the rigid rotation of}\\ \text{the small ball about p}\end{array}\right.$$

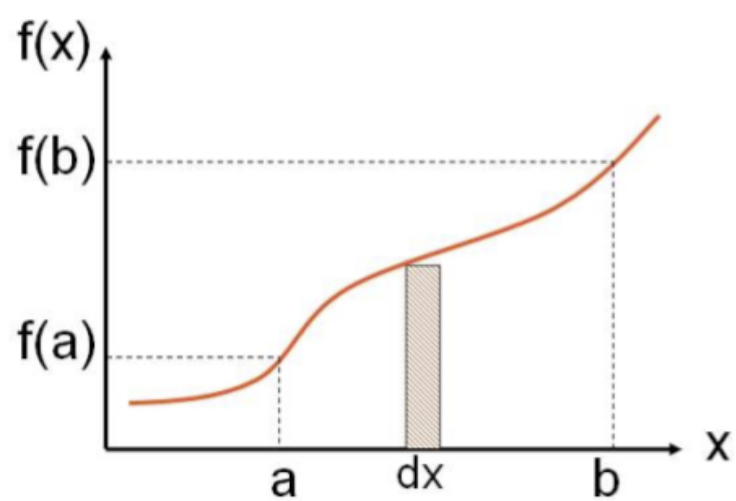
so that we have

$$\nabla\times\vec{v}(p)=\left\{\begin{array}{c}\text{twice the angular velocity}\\ \text{vector of the small ball which}\\ \text{is carried along with the fluid at p}\end{array}\right.$$

Conclusion:

$\nabla\times\vec{v}(p)$ measures the rotational (or curling) aspects of a fluid at p where $\vec{v}$ = the velocity field of this fluid

Fundamental theorems for gradient, divergence, and curl



From calculus you recall the fundamental theorem of calculus says

$$\int_a^b \frac{df(x)}{dx}dx = f(b)-f(a)$$

in other words, it's a connection between the rate of change of the function over the interval [a, b] and the values of the function at the endpoints (boundaries) of that interval.

There are equivalent "fundamental theorems" for line integrals, area integrals, and volume integrals.

In vector calculus we deal with different types of changes of scalar and vector fields, e.g.

$$\nabla\phi, \nabla\cdot\vec{v} \ \ and \ \ \nabla\times\vec{v}$$

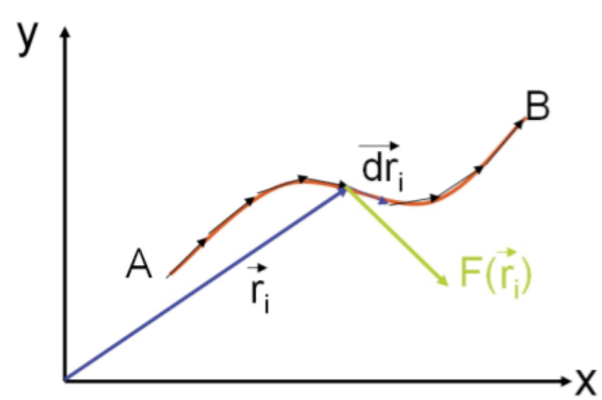
and each has its own theorem.

last year Mechanics class

You have discussed line integrals before(Physics 7; also reviewed in Mechanics Review Notes), mostly in the context of the work done along a path, but let's remind ourselves of the definition:

$$\int_A^B \vec{F}(\vec{r})\cdot d\vec{r} = \lim_{n\to\infty}\sum_{i=1}^{n}\int_A^B \vec{F}(\vec{r}_i)\cdot d\vec{r}_i$$

In other words we add up the area of all the little "rectangles" $\vec{F}(\vec{r}_i)\cdot d\vec{r}_i$ consisting of the vector $\vec{F}$ at the point $\vec{r}_i$ dotted into the path element $d\vec{r}_i$ as shown below



difficult to understand from formal definition; example clears things up!

Remember, the point is that, although we are doing an integral in a 2D space, we are constrained to move along a path, so there is only one real independent variable.

Example: Consider a triangle in the (x,y) plane with vertices at (-1,0), (1,0), and (0,1). Evaluate the closed line integral

$$I = \oint (-y\hat{x} + x\hat{y}) \cdot d\vec{r}$$

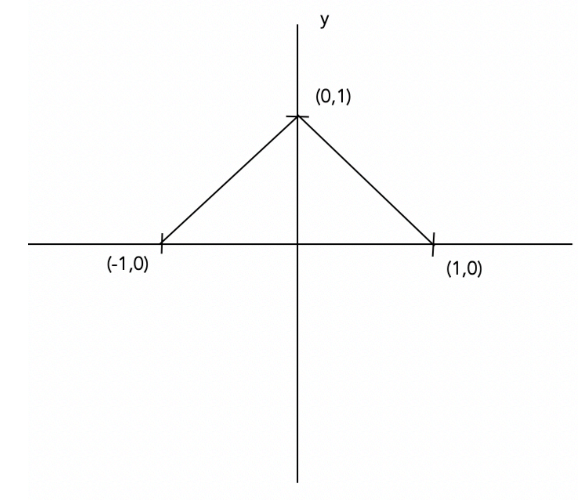
around the boundary of the triangle in the anticlockwise direction. Since the vector $\vec{r}$ has components (x, y), we have $d\vec{r} = dx\hat{x} + dy\hat{y}$ and

$$(-y\hat{x} + x\hat{y}) \cdot d\vec{r} = (-y\hat{x} + x\hat{y}) \cdot (dx\hat{x} + dy\hat{y}) = -ydx + xdy$$

On the path (-1,0) ➞ (1,0) we have y = 0, so the integral is

dy = 0 also

$$\int_{(-1,0)}^{(1,0)} (-ydx + xdy) = 0$$



On the path (1,0) ➞ (0,1) we have y = -x + 1, so integral is

$$\int_{(1,0)}^{(0,1)} (-ydx + xdy) = \int_{1}^{0} (-(1-x)dx) + \int_{0}^{1} ((1-y)dy) = \frac{1}{2} + \frac{1}{2} = 1$$

On the path (0,1) ➞ (-1,0) we have y = x + 1, so integral is

$$\int_{(0,1)}^{(-1,0} (-ydx + xdy) = \int_{0}^{-1} (-(1+x)dx) + \int_{1}^{0} ((y-1)dy) = \frac{1}{2} + \frac{1}{2} = 1$$

So total line integral is 0 + 1 + 1 = 2.

Let's go back and look at the leg from (1,0) ➞ (0,1) again. We could have "parameterized" this leg as x = 1 - t , y = t , 0 < t < 1. Then $\vec{r} = (1-t)\hat{x} + t\hat{y}$ and $d\vec{r} = (-\hat{x} + \hat{y})dt$. We can therefore write this integral as

$$\int_{(1,0)}^{(0,1)} \vec{F}(\vec{r}) \cdot d\vec{r} = \int_{0}^{1} (-t\hat{x} + (1-t)\hat{y}) \cdot (-\hat{x} + \hat{y})dt = \int_{0}^{1} (t + 1 - t)dt = 1$$

which is the same answer as above.

Think about that last one for a while!

Recall from calculus that the integral of an exact differential is independent of the path. This result can be expressed in our current particular context, line integrals of vector fields.

Suppose you have a vector field $\vec{F}$ which can be expressed as the gradient of a scalar field, $\vec{F} = \nabla\phi$. Then we have the fundamental theorem for gradients:

$$\int_a^b \nabla\phi \cdot d\vec{r} = \int_a^b d\phi = \phi(b) - \phi(a)$$

In other words, the integral

$$\int_a^b \vec{F} \cdot d\vec{r}$$

doesn't depend on the path between a and b. If $\vec{F} = \nabla\phi$ for some ϕ, $\vec{F}$ is **conservative**, and its integral is independent of path.

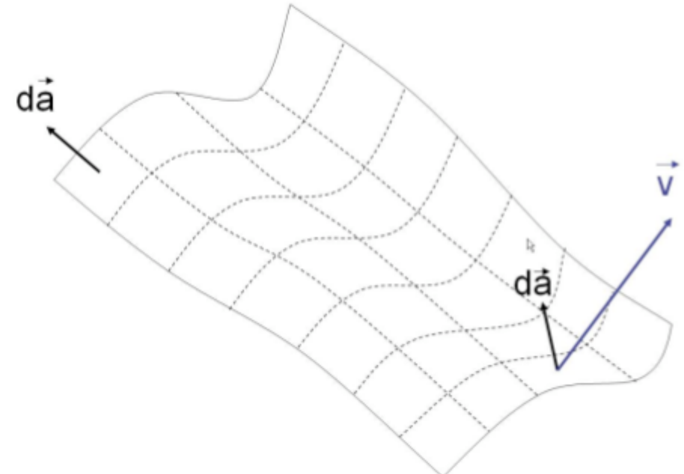
Now

$$\nabla \times \nabla\phi = \varepsilon_{ijk} \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_k}\phi = \varepsilon_{ijk} \frac{\partial^2\phi}{\partial x_j \partial x_k} = \frac{1}{2}\left(\varepsilon_{ijk} \frac{\partial^2\phi}{\partial x_j \partial x_k} + \varepsilon_{ikj} \frac{\partial^2\phi}{\partial x_k \partial x_j} \right)$$

$$= \frac{1}{2}\left(\varepsilon_{ijk} \frac{\partial^2\phi}{\partial x_j \partial x_k} - \varepsilon_{ijk} \frac{\partial^2\phi}{\partial x_k \partial x_j} \right) = \frac{\varepsilon_{ijk}}{2}\left(\frac{\partial^2\phi}{\partial x_j \partial x_k} - \frac{\partial^2\phi}{\partial x_k \partial x_j} \right) = 0$$

for any ϕ. Therefore, if $\nabla \times \vec{F} = 0$, then it is possible that $\vec{F} = \nabla\phi$. Thus, $\nabla \times \vec{F} = 0$ is a sufficient condition for a force to be conservative(derivable from a potential).

Surface integrals



We define the surface integral of the vector field $\vec{v}$ as

$$\int_A \vec{v} \cdot d\vec{A} = \lim_{n \to \infty} \sum_{i=1}^{n} \vec{v}(\vec{r}_i) \cdot d\vec{a}_i$$

where the infinitesimal directed area element $d\vec{a}_i$ is oriented perpendicular to the surface locally. Sometimes we call the integrand the "flux" of $\vec{v}$ through A,

$$d\Phi = \vec{v} \cdot d\vec{a}$$

This will be the flux of the electric field through a surface in this class, but it can be defined for any vector field.

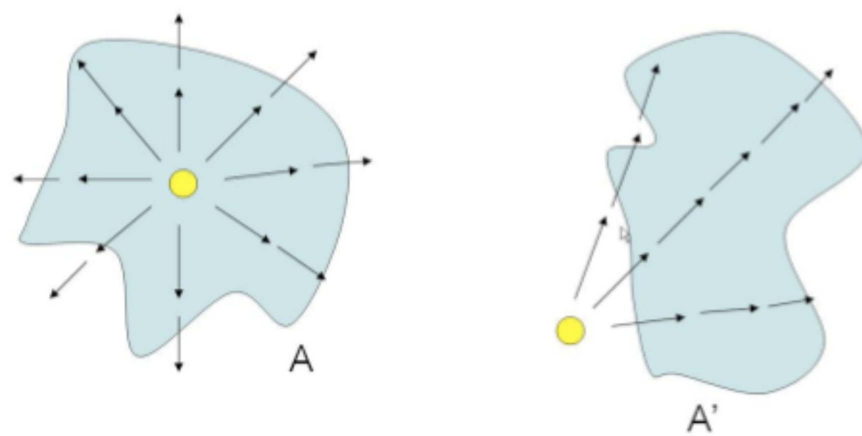
Another example: suppose $\vec{v}$ is a current density, i.e., represents the number of particles crossing a perpendicular plane per unit area per unit time. Then the flux $d\Phi = \vec{v} \cdot d\vec{a}$ is the number of particles passing though da per time. In general, the surface A may not be a plane but may have some curvature, but the total flux will just be

$$\Phi = \int_A \vec{v} \cdot d\vec{A}$$

Suppose A is a closed surface now, and there is a source of particles shooting out in all directions, penetrating the surface. Clearly the number of particles passing through the surface per time per area is nonzero. On the other hand, if the source is outside A', as shown, the flux through the surface looks like it might cancel on the left and right surfaces and be zero. In fact, we must somehow be able to express the conservation law that the particles which enter A' must eventually leave it:

what goes in, must come out

The answer is a very general relation between the surface integral of a vector field and the volume integral of its divergence $\nabla \cdot \vec{v}$.




```
Left:  particle source inside closed surface A. Flux is nonzero.
Right: particle source outside closed surface. Flux through A′ is zero.
```

Fundamental theorem for divergences: Gauss theorem.

Mathematically the divergence of $\vec{v}$ is just

$$\frac{\partial v_i}{\partial x_i} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}$$

Consider the volumes inside A and A' , A = boundary(V) and A' = boundary(V'). Remembering that we're thinking of $\vec{v}$ as a current density for now, let's ask ourselves how much comes out, passing through the surface. We can divide up the whole volume into infinitesimal cubes dxdydz and consider each cube independently. First imagine that $\vec{v}$ is constant. Then the surface integral over the cube will give zero, because the flux in one face will be exactly cancelled by the flux out the other face (signs of the dot product $\vec{v} \cdot d\vec{a}$ will be opposite). If $\vec{v}$ is not a constant, the vectors on opposite faces will not be quite equal, e.g. at x we have v_x but at x+dx we have $v_x + dv_x$. So the flow rate out of the cube in this direction will be equal to

$$\frac{\partial v_x}{\partial x} dx$$

Integrating over the whole cube will give the whole flow rate out:

$$\int_{cube} \nabla \cdot \vec{v}\, dxdydz = \int_{boundary(cube)} \vec{v} \cdot d\vec{a}$$

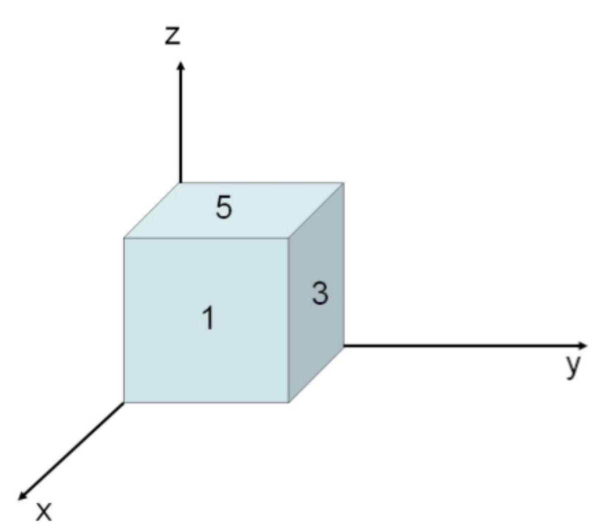
But if we consider any arbitrary volume τ to be a sum of such cubes, the surface integrals of the interior infinitesimal cubes will cancel, because their normals $\hat{a}$ are oppositely directed. Therefore the total volume integral over τ is similarly related to the total surface integral of boundary(τ):

$$\int_{\tau} \nabla \cdot \vec{v}\, d\tau = \int_{boundary(\tau)} \vec{v} \cdot d\vec{a}$$

This is called the divergence theorem or Gauss's theorem (not Gauss's law!).

This result is related to Gauss's law in electromagnetism where the source is an electric charge, and $\vec{v}$ the electric field. Note, however, Gauss's law is a statement about physics, and the relationship between the electric flux and its source, the electric charge. As we will see this depends sensitively on the $1/r^2$ nature of the Coulomb force.

Let's do an example where the total volume τ is an actual cube (side length = 1) as shown below.



We let

$$\vec{v} = y^2\hat{i} + (2xy + z^2)\hat{j} + 2yz\hat{k}$$

The divergence is

$$\nabla \cdot \vec{v} = \frac{\partial(y^2)}{\partial x} + \frac{\partial(2xy + z^2)}{\partial y} + \frac{\partial(2yz)}{\partial z} = 0 + 2x + 2y = 2(x + y)$$

The volume integral is

$$\int_\tau \nabla \cdot \vec{v}\, d\tau = 2\int_0^1\int_0^1\int_0^1 dxdydz(x + y) = 2$$

and the integrals through the faces are

(1) $d\vec{a} = dydz\,\hat{i}\,, x = 1, \vec{v} \cdot d\vec{a} = y^2 dydz \Rightarrow \int_0^1\int_0^1 y^2 dydz = \frac{1}{3}$

(1) $d\vec{a} = -dydz\,\hat{i}\,, x = 0, \vec{v} \cdot d\vec{a} = -y^2 dydz \Rightarrow -\int_0^1\int_0^1 y^2 dydz = -\frac{1}{3}$

(1) $d\vec{a} = dxdz\,\hat{j}, y = 1, \vec{v} \cdot d\vec{a} = (2xy + z^2)dxdz \Rightarrow \int_0^1\int_0^1 (2xy + z^2)dxdz = \frac{4}{3}$

(1) $d\vec{a} = -dxdz\,\hat{j}, y = 0, \vec{v} \cdot d\vec{a} = -z^2 dxdz \Rightarrow -\int_0^1\int_0^1 z^2 dxdz = -\frac{1}{3}$

(1) $d\vec{a} = dxdy\,\hat{k}, z = 1, \vec{v} \cdot d\vec{a} = 2yzdxdy \Rightarrow \int_0^1\int_0^1 2ydxdy = 1$

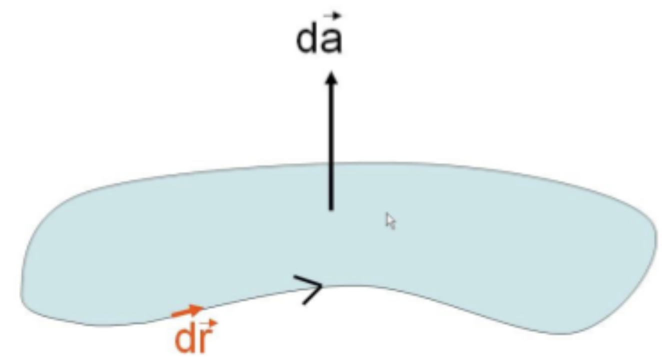
(1) $d\vec{a} = -dxdy\,\hat{k}, z = 0, \vec{v} \cdot d\vec{a} = -2yzdxdy = 0 \Rightarrow -\int_0^1\int_0^1 2yzdxdy = 0$

and the sum over the 6 faces is

$$\frac{1}{3} - \frac{1}{3} + \frac{4}{3} - \frac{1}{3} + 1 + 0 = 2$$

which is the same as the volume integral!

Fundamental theorem for curls: Stokes theorem



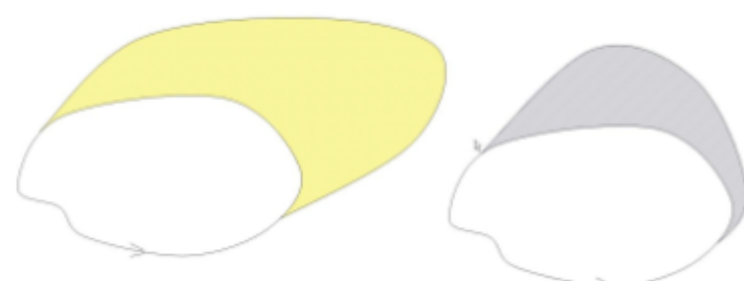
Directed area is perpendicular to loop according to right hand rule.

I will simply state it without proof here. In this case we consider a directed closed path in space, the boundary of a 2D area A.

$$\int_A (\nabla \times \vec{v}) \cdot d\vec{a} = \oint_{boundary(A)} \vec{v} \cdot d\vec{r}$$

Remarks (Stokes and divergence theorems):

Define the normal to the area according to the right hand rule using the sense of the loop, as in figure above.



Many surfaces can have the same boundary, see figure above.

If $\vec{v} = \nabla \times \vec{B}$ then using the divergence theorem

$$\int_{cube} \nabla \cdot \vec{v}\ dxdydz = \int_{boundary(cube)} \vec{v} \cdot d\vec{a}$$

and the fact that

$$\nabla\cdot\left(\nabla\times\vec{B}\right)=\nabla\cdot\left(\varepsilon_{ijk}\nabla_j B_k\hat{e}_i\right)=\left(\nabla\cdot\hat{e}_i\right)\left(\varepsilon_{ijk}\nabla_j B_k\right)=\varepsilon_{ijk}\nabla_i\nabla_j B_k=\varepsilon_{ijk}\frac{\partial^2 B_k}{\partial x_i\partial x_j}$$

$$=\frac{1}{2}\left(\varepsilon_{ijk}\frac{\partial^2 B_k}{\partial x_i\partial x_j}+\varepsilon_{jik}\frac{\partial^2 B_k}{\partial x_j\partial x_i}\right)=\frac{1}{2}\left(\varepsilon_{ijk}\frac{\partial^2 B_k}{\partial x_i\partial x_j}-\varepsilon_{ijk}\frac{\partial^2 B_k}{\partial x_j\partial x_i}\right)$$

$$=\frac{\varepsilon_{ijk}}{2}\left(\frac{\partial^2 B_k}{\partial x_i\partial x_j}-\frac{\partial^2 B_k}{\partial x_j\partial x_i}\right)=0$$

We then have

$$\int_{boundary(cube)}\left(\nabla\times\vec{B}\right)\cdot d\vec{a}=0$$

If $\vec{v}=\nabla\phi$, then $\nabla\times\vec{v}=\nabla\times\nabla\phi=0$. This says that $\vec{v}$ is conservative.

Divergenceless Fields

If $\nabla\cdot\vec{v}=0$ everywhere, the following conditions are equivalent:

$$\int_{closed\ surface}\vec{v}\cdot d\vec{a}=0$$

$\vec{v}$ is the curl of some vector, $\vec{v}=\nabla\times\vec{w}$. Note it is obviously not unique, we can add $\nabla\phi$ for any scalar field ϕ , and still maintain $\vec{v}=\nabla\times\vec{w}$. This will appear as "gauge symmetry" in electromagnetism.

Also note that

$$\int_{open\ surface}\vec{v}\cdot d\vec{a}$$

is independent of the surface for a given boundary line.

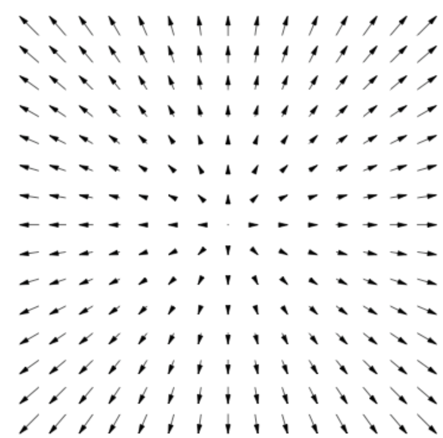
All this stuff will eventually make more sense when it comes up later in our studded of E&M

Curlless (conservative) fields

If $\nabla \times \vec{v} = 0$ everywhere, then $\oint \vec{v} \cdot d\vec{r} = 0$ by Stokes theorem.

If $\vec{v}$ is the gradient of some scalar $\vec{v} = -\nabla\phi$, then $\int_a^b \vec{v} \cdot d\vec{r}$ is independent of the path between a and b.

Return to earlier examples:



Example 1. "Diverging" radial field

$$\vec{v} = \vec{r} = (x, y) \qquad \nabla \cdot \vec{v} = 3 > 0$$

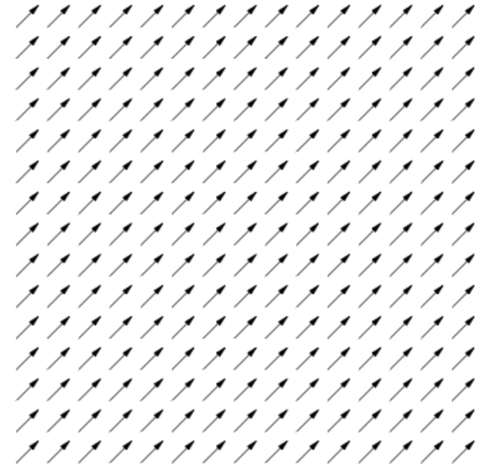
but

$$\nabla \times \vec{v} = 0$$

Vector field "spreads out" or "diverges". Flux

$$\int \vec{v} \cdot d\vec{a}$$

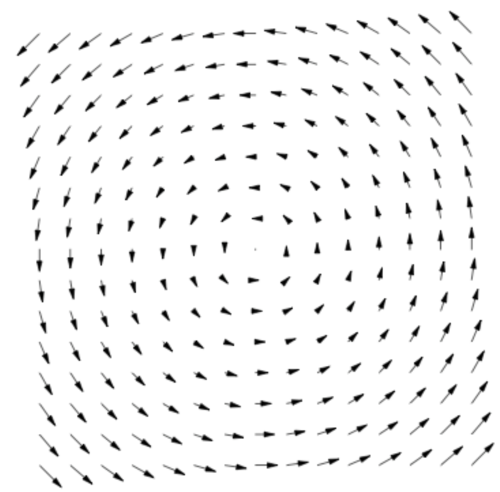
through closed surface surrounding $\vec{r} = 0$ is obviously nonzero.



Example 2. Constant vector field.

$$\vec{v} = (1,1) \qquad \nabla \cdot \vec{v} = 0$$

No spreading. Net flux through any closed surface is zero.



Example 3.

$$\vec{v} = (-y, x, 0) \qquad \nabla \cdot \vec{v} = 0 \qquad \nabla \times \vec{v} = (0,0,2) \neq 0$$

Line integral

$$\oint \vec{v} \cdot d\vec{r} = \oint r\hat{\theta} \cdot r\hat{\theta} d\theta = 2\pi r^2$$

Compare with integral of curl dotted into surface, take cylinder for simplicity:

$$\int (\nabla \times \vec{v}) \cdot d\vec{a} = 2\pi r^2$$

Stokes works!

Divergence of Coulomb/gravitational force of point charge/mass
The Dirac Delta Function

Consider the vector field $\vec{v}$ with the form

$$\vec{v} = \frac{\hat{r}}{r^2} = \frac{\vec{r}}{r^3}$$

i.e. it points radially outward and falls off as $1/r^2$.

This is the electric field of a point charge, or the gravitational field of a point mass. Here's a little paradox: CLAIM $\nabla \cdot \vec{v} = 0$:

$$\nabla \cdot \frac{(x,y,z)}{\left(x^2+y^2+z^2\right)^{3/2}} = \left(\frac{1}{\left(x^2+y^2+z^2\right)^{3/2}} - \frac{3x^2}{\left(x^2+y^2+z^2\right)^{5/2}} \right) + (x \leftrightarrow y) + (x \leftrightarrow z)$$

$$= \frac{-2x^2+y^2+z^2}{\left(x^2+y^2+z^2\right)^{5/2}} + (x \leftrightarrow y) + (x \leftrightarrow z) = 0$$

OK, so this field is divergenceless. It looks like something is flowing out, but on the other hand the magnitude of $\vec{v}$ is getting smaller as you go out, so maybe it's OK. BUT what about the divergence theorem? If I calculate

$$\int_{sphere} \vec{v} \cdot d\vec{a} = \int \frac{\hat{r}}{r^2} \cdot \hat{r} r^2 \sin\theta d\theta d\phi = 4\pi$$

It is not zero! So, is the divergence theorem wrong, or did we miss something?

The answer is that we were a little careless in calculating the divergence and declaring it to be zero. It is zero, almost everywhere. However, if you look at our equations exactly at the point x = y = z = 0, you will see that the derivatives diverge, i.e. are infinite (you may need to do L'Hospital's rule to convince yourself, but it's right).

So we have encountered a very peculiar situation, a function of space which is infinite at one point, but zero everywhere else. This type of function was explored in the physics context by Paul Dirac, so we refer to it as the Dirac delta-function, and give it the symbol δ, defined as

$$\delta(x) = 0 \quad , \quad x \neq 0 \quad and \quad \int_{-\infty}^{\infty} \delta(x)dx = 1$$

and

$$\delta^{(3)}(\vec{r}) = \delta(x)\delta(y)\delta(z)$$

Mathematically, it is not a function; it's called a distribution.

So since

$$\int_{\tau} \nabla \cdot \frac{\hat{r}}{r^2} dV = 4\pi$$

we have determined that

$$\nabla \cdot \frac{\hat{r}}{r^2} = 4\pi\delta^{(3)}(\vec{r})$$

Properties of δ-function

$$\delta_n(x) = 0 \quad , \quad x \neq 0$$

$$\int_{-\infty}^{\infty} \delta(x-a)f(x)dx = f(a)$$

$$\delta(ax) = \frac{1}{a}\delta(x)$$

$$\delta(g(x)) = \sum_{a_n} \frac{\delta(x-a_n)}{|g'(a_n)|}$$

where a_n are the zeros of $g(x) = 0$.

Example 1:

$$\int_0^5 \delta(x-3)x^3 dx = 27 \quad but \quad \int_0^2 \delta(x-3)x^3 dx = 0$$

Example 2:

$$\int_{-\infty}^{\infty} \delta(5x)(x+1)^3 dx = \frac{1}{5}\int_{-\infty}^{\infty} \delta(x)(x+1)^3 dx = \frac{1}{5}$$

Example 3:

$$\int_0^{\infty} \delta(x^2-1)x^3 dx = \int_0^{\infty}\left[\frac{\delta(x-1)}{|g'(1)|} + \frac{\delta(x+1)}{|g'(-1)|}\right]x^3 dx$$

$$= \frac{1}{2}\left[\int_0^{\infty} \delta(x-1)x^3 dx + \int_0^{\infty} \delta(x+1)x^3 dx\right] = \frac{1}{2}(1+0) = \frac{1}{2}$$

Taylor Series

A MaClaurin expansion for a function f(x) in 1-dimension relative to the origin x = 0 follows directly if the function f(x) has a power series expansion. Let

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

We then have

$$f(0) = a_0 \ , f'(0) = a_1 \ , \frac{1}{2!} f''(0) = a_2 , ...$$

or in general

$$\frac{1}{n!}f^{(n)}(0) = a_n$$

and thus

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

This is also called the Taylor expansion for f(x) about the origin x = 0. A Taylor series, in general, means a power series in powers of (x-a) where a = some constant. The derivation of the coefficients is identical to the last derivation except that we use x = a instead of x = 0.

Let

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n$$

We then have

$$f(a) = a_0 \ , f'(a) = a_1 \ , \frac{1}{2!}f''(a) = a_2, ...$$

or in general

$$\frac{1}{n!}f^{(n)}(a) = a_n$$

and thus

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n$$

Let us now determine f(x+h) using a neat method. Consider the following construction for an infinitesimal displacement

$$f(x+\Delta x) = f(x) + \Delta x \frac{df}{dx} = \left(1 + \Delta x \frac{d}{dx}\right) f(x)$$

where we have used the definition of the derivative(exact only in the limit of course). Now make another displacement

$$f(x+2\Delta x) = \left(1 + \Delta x \frac{d}{dx}\right) f(x+\Delta x) = \left(1 + \Delta x \frac{d}{dx}\right)^2 f(x)$$

Since a finite displacement h can always be constructed from an infinite number of consecutive infinitesimal displacements, we have

$$f(x+h)=\lim_{n\to\infty}\left(1+\frac{h}{n}\frac{d}{dx}\right)^n f(x)=\exp\left(h\frac{d}{dx}\right)f(x)$$

where we have used the interesting result that

$$\lim_{n\to\infty}\left(1+\frac{h}{n}Q\right)^n=\exp(hQ)$$

The differential operator

$$\exp\left(h\frac{d}{dx}\right)$$

is called the displacement operator. Since the exponential function has the power series representation

$$\exp(ax)=\sum_{k=0}^{\infty}\frac{(ax)^k}{k!}$$

we can write

$$f(x+h)=\exp\left(h\frac{d}{dx}\right)f(x)=\sum_{k=0}^{\infty}\frac{h^k}{k!}f^{(k)}(x)$$

Example:

$$\sin\omega(t+\tau)=\exp\left(\tau\frac{d}{dt}\right)\sin(\omega t)$$

$$=\left[1+\tau\frac{d}{dt}+\frac{1}{2!}\tau^2\frac{d^2}{dt^2}+\frac{1}{3!}\tau^3\frac{d^3}{dt^3}+....\right]\sin(\omega t)$$

$$=\left[\sin(\omega t)+\omega\tau\cos(\omega t)-\frac{1}{2!}(\omega\tau)^2\sin(\omega t)-\frac{1}{3!}(\omega\tau)^3\cos(\omega t)+....\right]$$

$$=\sin(\omega t)\left[1-\frac{1}{2!}(\omega\tau)^2+....\right]+\cos(\omega t)\left[\omega\tau-\frac{1}{3!}(\omega\tau)^3+....\right]$$

$$=\sin(\omega t)\cos(\omega\tau)+\cos(\omega t)\sin(\omega\tau)$$

which we know to be correct.

These results generalize easily to 3-dimensions as

$$f(\vec{r}+\vec{a})=\sum_{n=0}^{\infty}\frac{1}{n!}(\vec{a}\cdot\nabla)f(\vec{r})$$

and also can be used for displacements in time

$$f(t+\tau)=\exp\left(\tau\frac{d}{dt}\right)f(t)=\sum_{k=0}^{\infty}\frac{\tau^k}{k!}f^{(k)}(t)$$

Binomial Series

Consider the following function

$$f(x)=(1+x)^n$$

If we expand this out we get

$$f(x)=(1+x)^n=1+nx+\frac{n(n-1)}{2!}x^2+....=\sum_{m=0}^{\infty}\frac{n!}{m!(n-m)!}x^m=\sum_{m=0}^{\infty}\binom{n}{m}x^m$$

this is called the Binomial series and the ( ) is the Binomial coefficient. For n = integer, the series terminates at m = n.

In a similar way we can write(for n = integer),

$$(p+q)^n=\sum_{m=0}^{n}\frac{n!}{m!(n-m)!}p^m q^{n-m}=\sum_{m=0}^{n}\binom{n}{m}p^m q^{n-m}$$

This has an interesting interpretation in the following example. Suppose we have n coins that we will flip onto a table and we count the number heads and tails. Suppose the probability of getting a heads = p and the probability of getting a tails is q. Hence p+q = 1. This means we have

$$(p+q)^n=(1)^n=1=\sum_{m=0}^{n}\frac{n!}{m!(n-m)!}p^m q^{n-m}$$

Now the probability of throwing m heads and n-m tails is given by

$$p^m q^{n-m}$$

and the number of indistinguishable ways we can do this is

$$\frac{n!}{m!(n-m)!}$$

Therefore the total probability of throwing m heads and n-m tails is given by

$$P(n,m) = \frac{n!}{m!(n-m)!} p^m q^{n-m} = C(n,m) p^m q^{n-m}$$

Since the probability of something happening is sum of all of these probabilities, the sum must = 1 as we have already seen above.

Does this really work? Consider n = 3, so we can actually do it.

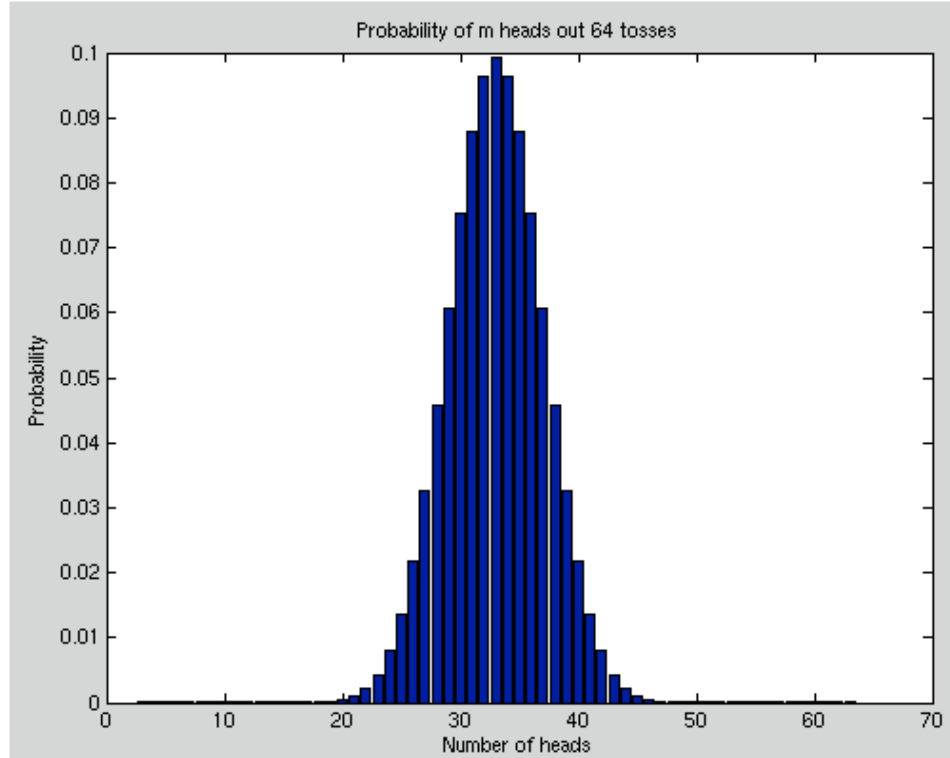
For n = 3, the number of possible outcomes = 2x2x2 = 8

ppp = 1 = C(3,3) - p(3,3) = 1/8
ppq,pqp,qpp = 3 = C(3,1) - p(3,2) = 3/8
qqp,qpq,pqq = 3 = C(3,2) - p(3,1) = 3/8
qqq = 1 = C(3,0) - p(3,0) = 1/8

So it does work!

Octave Program

```
% m-file coins1.m
% 64 coins with p = q = 1/2
p=0.5; N=64; pn=[];
for m=-N:2:N
 pn=[pn,(prod(1:N)*(0.5)^N)/ ...
       (prod(1:((N+m)/2))*prod(1:((N-m)/2)))];
end
bar(pn)
title('Probability of m heads out 64 tosses');
xlabel('Number of heads');
ylabel('Probability')
```



We note that the quantity n! that we have been using is usually defined only for positive integers. As we shall later in the course, there is a more general definition of this function, which reduces to the same values when n = integer. It is given by

$$s! = \int_0^\infty x^s e^{-x} dx = \Gamma(s+1)$$

for all s not equal to a negative integer. It is called the gamma function.

Taylor Expansion in more than 1 variable

We can generalize the Taylor expansion to more than 1 dimensions as follows:

$$\begin{aligned} f(x,y) &= f(a,b) + (x-a) f_x(a,b) + (y-b) f_y(a,b) \\ &+ \frac{1}{2!}(x-a)^2 f_{xx}(a,b) + \frac{2}{2!}(x-a)(y-b) f_{xy}(a,b) + \frac{1}{2!}(y-b)^2 f_{yy}(a,b) \\ &+ \frac{1}{3!}(x-a)^3 f_{xxx}(a,b) + \frac{3}{3!}(x-a)^2(y-b) f_{xxy}(a,b) + \frac{3}{3!}(x-a)(y-b)^2 f_{xyy}(a,b) + \frac{1}{3!}(y-b)^3 f_{yyy}(a,b) \\ &+ \ldots\ldots \end{aligned}$$

and so on.

The electromagnetic force: Ancient history…

500 B.C. – Ancient Greece
Amber (ελεχτρον="electron") attracts light objects
Iron rich rocks from μαγνεσια(Magnesia) attract iron

1730 - C. F. duFay: Two flavors of charges
Positive and negative

1766-1786 – Priestley/Cavendish/Coulomb
EM interactions follow an inverse square law:
Actual precision better than 2×10^9 !

$$F_{em} \propto \frac{q_1 q_2}{r^2}$$

1800 – Volta
Invention of the electric battery

Until now Electricity and Magnetism are disconnected!

1820 – Oersted and Ampere
Established first connection between electricity and magnetism

1831 – Faraday
Discovery of magnetic induction

1873 – Maxwell: Maxwell's equations
The birth of modern Electro-Magnetism

1887 – Hertz
Established connection between EM and radiation

1905 – Einstein
Special relativity makes connection between Electricity and Magnetism as natural as it can be!

Electricity and Magnetism are the same thing!

Electrostatics 101

The EM force acts on charges

Like charges repel

opposite charges attract

Electric charge is quantized (Millikan)

Multiples of the e = elementary charge

Oil drop experiment

e = $4.803\times10^{-10}\, esu(CGS)$, $1.602\times10^{-19}\, C(SI)$

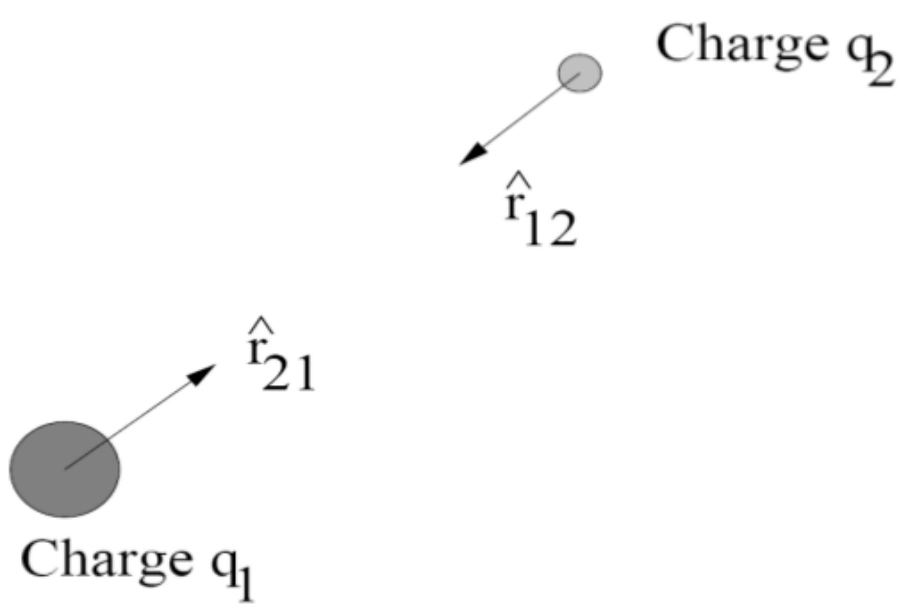
Q = -e ; Q = +e

Electric charge is conserved

In any isolated system, the total charge cannot change.

If the total charge of a system changes, then it means the system is not isolated and charges came in or escaped.

Coulomb's law



$$\vec{F}_2 = k\frac{q_1 q_2}{|r_{21}|^2}\hat{r}_{21}$$

What is force on charge q_1?

Where:

$\vec{F}_2$ is the force that the charge q_2 feels due to q_1

$\hat{r}_{21}$ is the unit vector going from q_1 to q_2

$$\vec{F}_2 = k\frac{q_1 q_2}{|r_{21}|^2}\hat{r}_{21}$$

Consequences:

Newton's third law: $\vec{F}_2 = -\vec{F}_1$

Like signs repel, opposite signs attract

Units: cgs vs SI

Units in cgs and SI (Sisteme Internationale)

	cgs	SI
Length	cm	m
Mass	g	Kg
Time	s	s
Charge	electrostatic units (e.s.u.)	Coulomb (C)
Current	e.s.u./s	Ampere (A)

In cgs the esu is defined so that k=1 in Coulomb's law →

$$1\ dyne = \frac{(1\ esu)^2}{(1\ cm)^2} \rightarrow 1\ esu = cm\sqrt{dyne}$$

esu is not fundamental!

In SI, the Ampere is a fundamental unit

$$k = \frac{1}{4\pi\varepsilon_0} = 8.99\times 10^9\ Newton - meter^2 / C^2$$

$$\varepsilon_0 = 8.8\times 10^{-12} C^2/(N - m^2) = \text{"}permittivity\ of\ free\ space$$

Practical info: cgs-SI conversion table

	SI Units		CGS units
Energy	1 Joule	=	10^7 erg
Force	1 Newton	=	10^5 dyne
Charge	1 Coulomb	=	"3" $\times 10^9$ esu
Current	1 Ampere	=	"3" $\times 10^9$ esu/sec
Potential	"3" $\times 10^2$ Volts	=	1 statvolt
Electric field	"3" $\times 10^4$ Volts/m	=	1 statvolt/cm
Magnetic field	1 Tesla	=	10^4 gauss
Capacitance	1 Farad	=	"9" $\times 10^{11}$ cm
Resistance	"9" $\times 10^{11}$ Ohm	=	1 sec/cm
Inductance	"9" $\times 10^{11}$ Henry	=	1 sec^2/cm

"3"=2.9979... =c

FAQ: why do we use cgs?

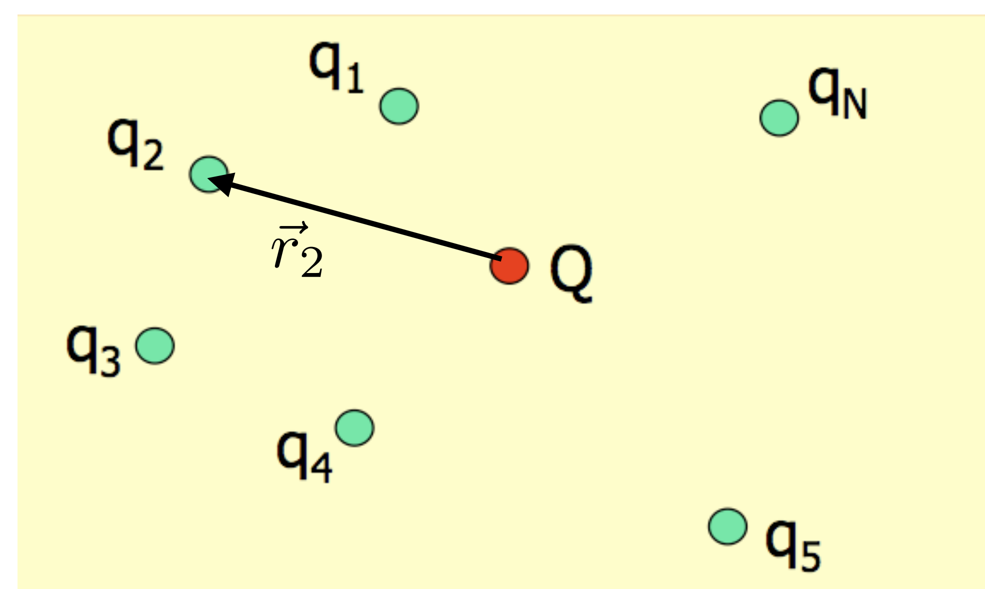
Theoretical physicists prefer "cgs"

"Constants =1"

The superposition principle: discrete charges

The force on the charge Q due to all the other charges is equal to the vector sum of the forces created by the individual charges:

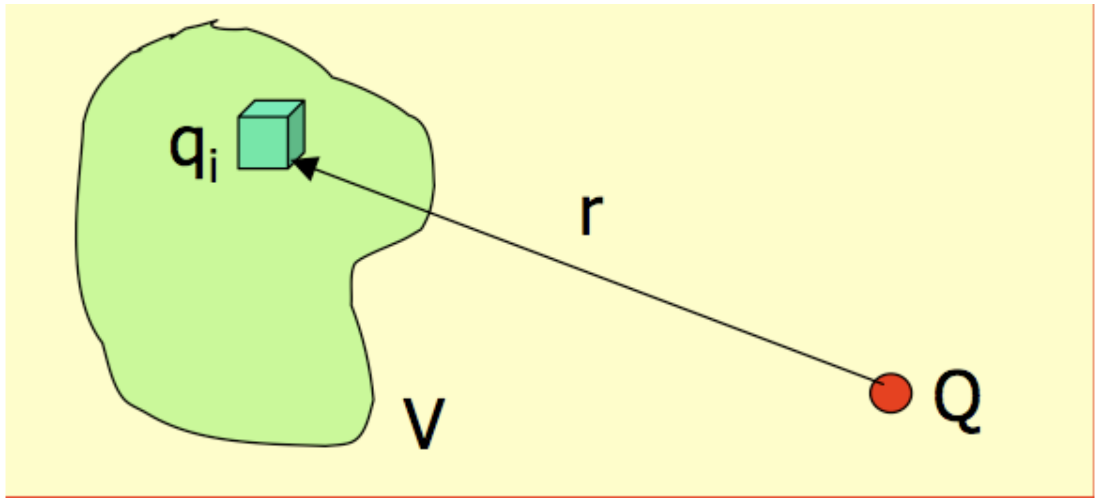
$$\vec{F}_Q = \frac{q_1 Q}{|r_1|^2}\hat{r}_1 + \frac{q_2 Q}{|r_2|^2}\hat{r}_2 + \ldots\ldots + \frac{q_N Q}{|r_N|^2}\hat{r}_N = \sum_{i=1}^{N} \frac{q_i Q}{|r_i|^2}\hat{r}_i$$



The superposition principle: continuous distribution of charges

What happens when the distribution of charges is continuous?

Take the limit for q ➞ dq and Σ ➞ integral:



$$\vec{F}_Q = \sum_{i=1}^{N} \frac{q_i Q}{|r_i|^2}\hat{r}_i \rightarrow \int_V \frac{dqQ}{|r|^2}\hat{r} = \int_V \frac{\rho dVQ}{|r|^2}\hat{r}$$

where ρ= charge per unit volume: the "volume charge density"

Charges are distributed inside a volume V: $\vec{F}_Q = \int_V \frac{\rho dVQ}{|r|^2}\hat{r}$

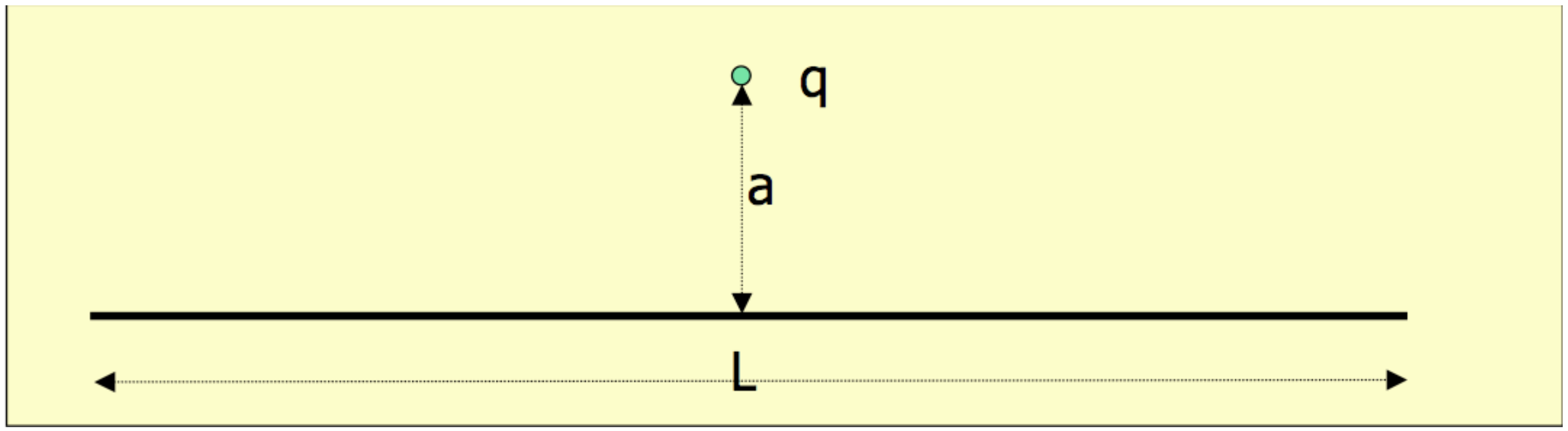
Charges are distributed on a surface A: $\vec{F}_Q = \int_A \frac{\sigma daQ}{|r|^2}\hat{r}$

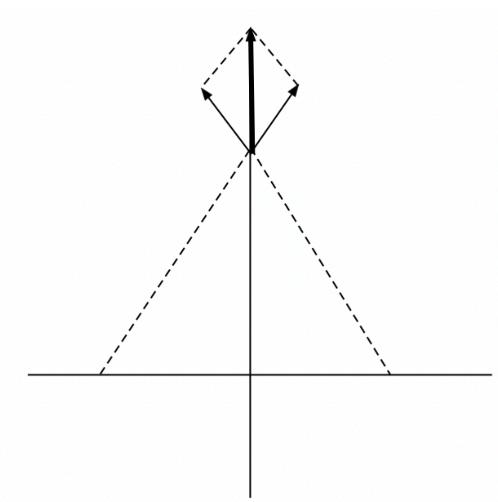
Charges are distributed on a line L: $\vec{F}_Q = \int_L \frac{\lambda dlQ}{|r|^2}\hat{r}$

Where:
ρ= charge per unit volume: "volume charge density"
σ= charge per unit area: "surface charge density"
λ= charge per unit length: "line charge density"

Application: charged rod

A rod of length L has a charge Q uniformly spread over it.
A test charge q is positioned at a distance a from the rod's midpoint.
What is the force F that the rod exerts on the charge q?

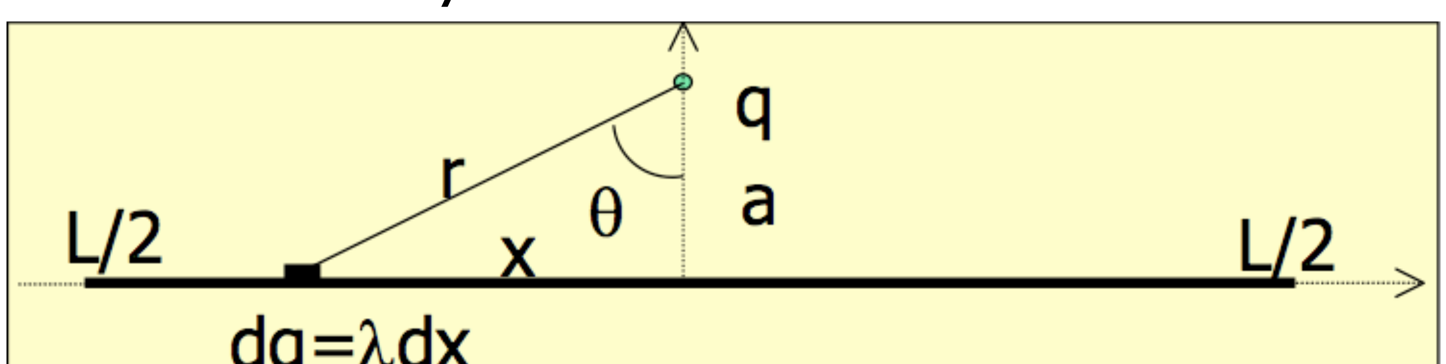




Answer: $$\vec{F} = \frac{Qq}{a\sqrt{a^2 + \left(\frac{L}{2}\right)^2}}\hat{y}$$

Solution: charged rod

Look at the symmetry of the problem and choose appropriate coordinate system: rod on x axis, symmetric wrt x=0; a on y axis:



Symmetry of the problem: F // y axis; define $\lambda = Q/L$ linear charge density

Trigonometric relations: $x/a = \tan\theta$; $a = r\cos\theta \rightarrow dx = a\,d\theta/\cos^2\theta$; $r = a/\cos\theta$ $dx = a \sec^2\theta\, d\theta$

Consider the infinitesimal charged F_y produced by the element dx:

$$dF_y = dF\cos\theta = \frac{\lambda dx}{r^2} q\cos\theta = \lambda q \frac{\frac{a d\theta}{\cos^2\theta}}{\frac{a^2}{\cos^2\theta}} \cos\theta = \frac{\lambda q}{a}\cos\theta d\theta$$

dF in direction along hypothenuse;
dF_y = vertical component

Now integrate between –L/2 and L/2:

$$\vec{F} = \hat{y}\int_{-L/2}^{L/2} \frac{\lambda q}{a}\cos\theta d\theta = \hat{y}\int_{-L/2}^{L/2} \frac{\lambda q}{a}\cos\theta d\theta = \hat{y}\frac{\lambda q}{a}\sin\theta \Bigg|_{\sin\theta = \frac{-L/2}{\sqrt{a^2+\left(\frac{L}{2}\right)^2}}}^{\sin\theta = \frac{L/2}{\sqrt{a^2+\left(\frac{L}{2}\right)^2}}} = \frac{Qq}{a\sqrt{a^2+\left(\frac{L}{2}\right)^2}}\hat{y}$$

Infinite rod? Taylor expansion!

What if the rod length is infinite?

What does "infinite" mean? For all practical purposes, infinite means >> than the other distances in the problem: L>>a: Let's look at the solution:

$$\vec{F} = \frac{Qq}{a\sqrt{a^2+\left(\frac{L}{2}\right)^2}}\hat{y} = \frac{Qq}{a\frac{L}{2}\sqrt{1+\left(\frac{2a}{L}\right)^2}}\hat{y}$$

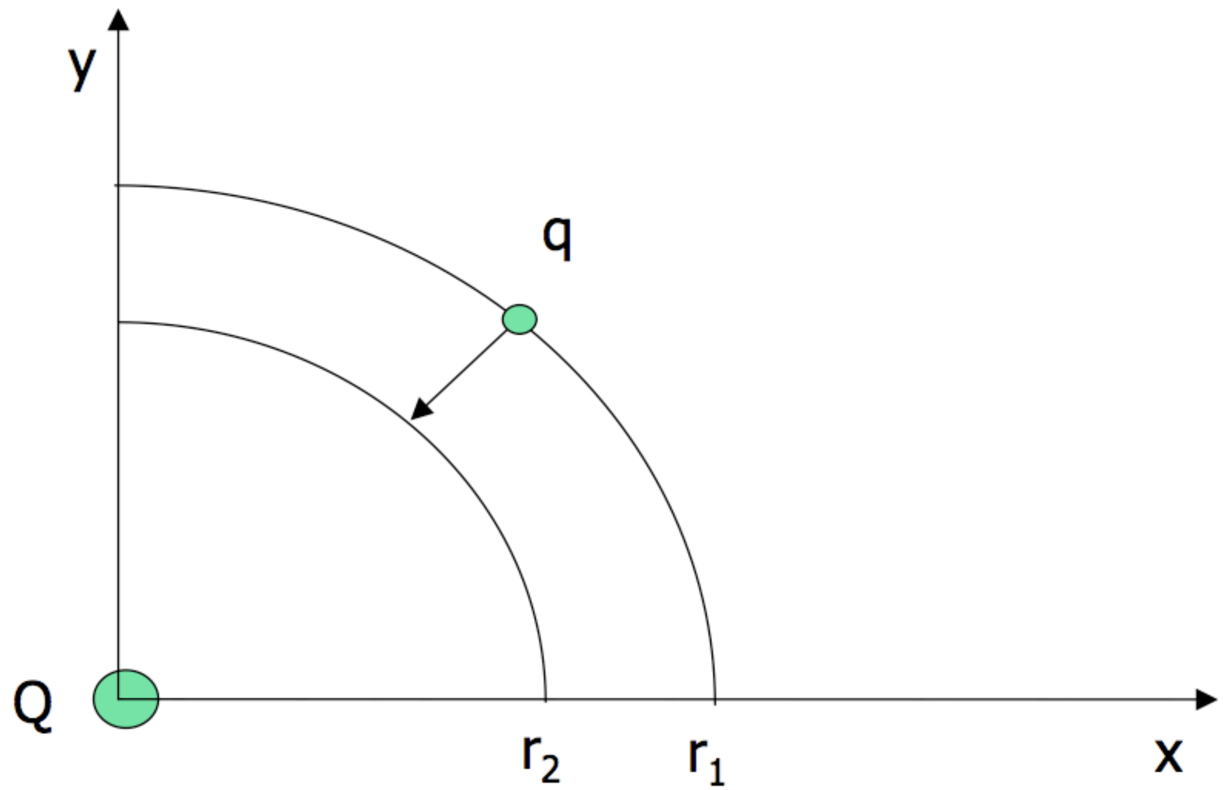
Taylor expand using $\left(\frac{2a}{L}\right)^2$ as expansion coefficient remembering that

$$(1\pm x)^n = 1 \pm \frac{nx}{1!} + \frac{n(n-1)}{2!}x^2 \pm \ldots\ldots \quad for \quad x^2 < 1$$

$$(1\pm x)^{-n} = 1 \mp \frac{nx}{1!} + \frac{n(n+1)}{2!}x^2 \mp \ldots\ldots \quad for \quad x^2 < 1$$

$$F=\frac{Qq}{a\sqrt{a^2+\left(\frac{L}{2}\right)^2}}=\frac{\frac{\lambda Lq}{a}}{\frac{L}{2}\left(1+\frac{2a}{L}\right)^{1/2}}=\frac{\lambda q}{2a}\left(1+\frac{2a}{L}\right)^{-1/2}=\frac{\lambda q}{2a}\left(1-\frac{1}{2}\left(\frac{2a}{L}\right)^2+....\right)\approx\frac{\lambda q}{2a}$$

Energy associated with Coulomb Force



$$Notation: d\vec{s} \ or \ d\vec{l} \rightarrow displacements$$

$$d\vec{S} \rightarrow area$$

Work done to move charges

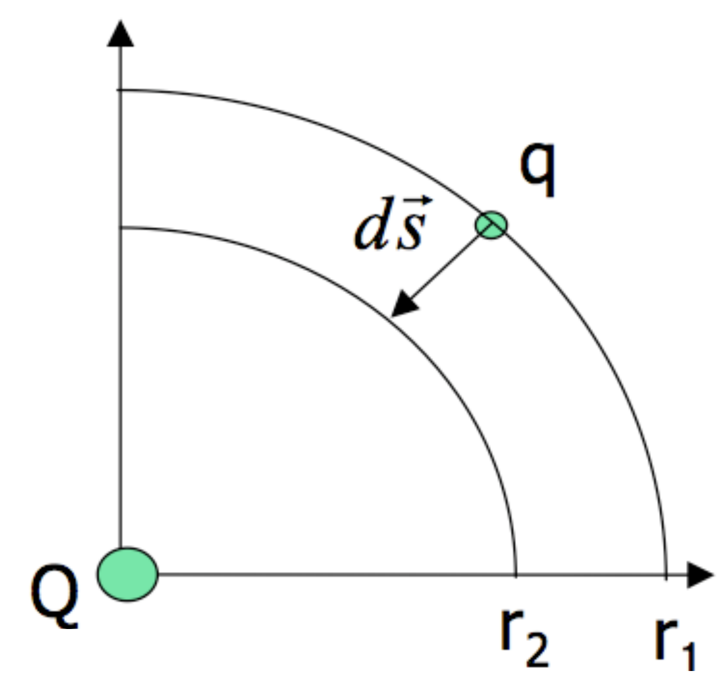
How much work do I have to do to move q from r_1 to r_2?

$$W=\int \vec{F}_I \cdot d\vec{s} \quad where \quad \vec{F}_I=-\vec{F}_{Coulomb}=-\frac{Qq}{r^2}\hat{r}$$

meaning?

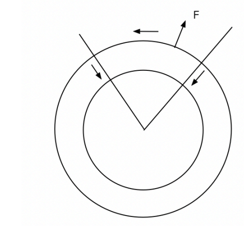
Let's assume a radial path:

$$W(r_1 \rightarrow r_2)=\int \vec{F}_I \cdot d\vec{s}=-\int_{r_1}^{r_2}\frac{Qq}{r^2}\hat{r}\cdot d\vec{r}=\frac{Qq}{r_2}-\frac{Qq}{r_1}$$



Does this result depend on the path chosen?

No! You can decompose any path in segments parallel to the radial direction and segments perpendicular to it. Since the component on the perpendicular to it is null the result does not change.

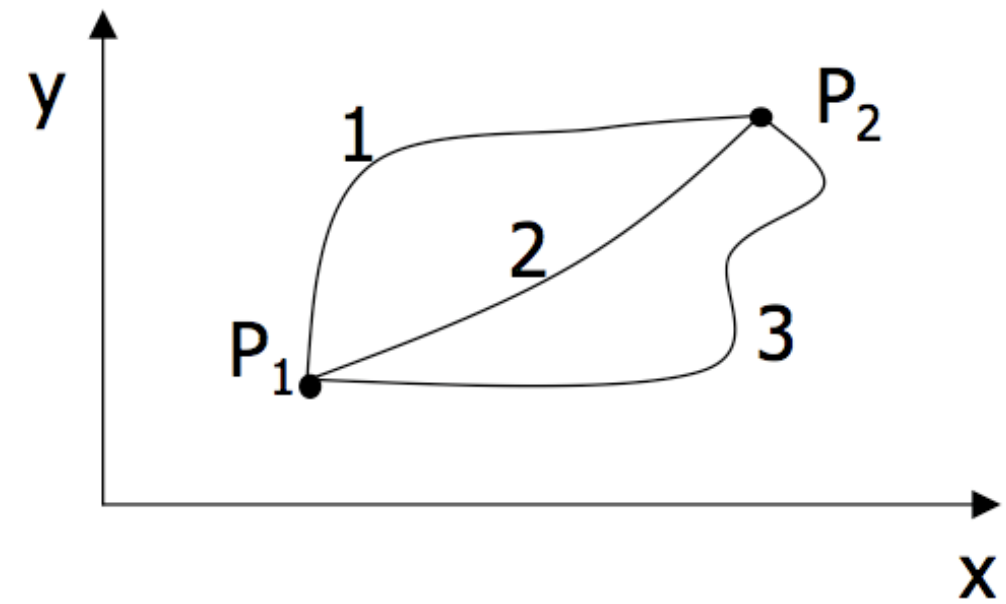


scalar product = 0

Corollaries

The work performed to move a charge between P_1 and P_2 is the same independently of the path chosen

$$W_{12} = \int_{Path\,1} \vec{F} \cdot d\vec{s} = \int_{Path\,2} \vec{F} \cdot d\vec{s} = \int_{Path\,3} \vec{F} \cdot d\vec{s}$$



This means that

The work to move a charge on a closed path is zero:

$$W_{11} = \oint_{Any} \vec{F} \cdot d\vec{s} = 0$$

In other words: the electrostatic force is conservative! This will allow us to introduce the concept of potential later.

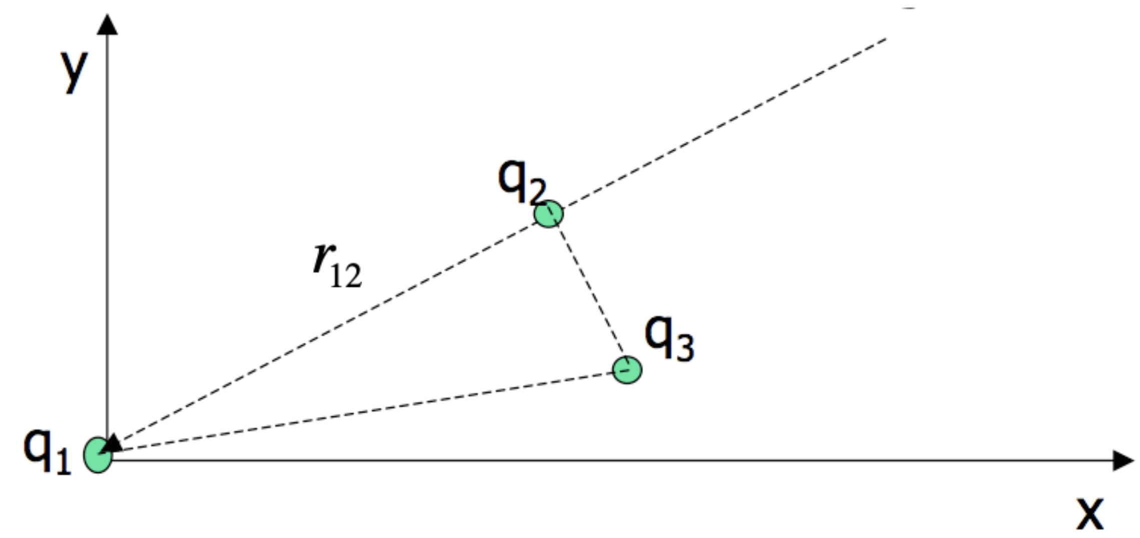
from mechanics!!

Energy of a system of charges

How much work does it take to assemble a certain configuration of charges?

$$W(q_1) = W_{q_1} = W_1 = \int_{\infty}^{P_1} \vec{F} \cdot d\vec{s} = 0 \quad \text{no other charges} \Rightarrow \vec{F} = 0$$

it is 1st change brought in !!



$$W_{1+2} = \oint_{Any} \vec{F}_I \cdot d\vec{s} = \frac{q_1 q_2}{r_{12}} \quad \text{source of field is only } q_1$$

$$W_{1+2+3} = W_{1+2} + W_{1+3} + W_{2+3} = \frac{q_1 q_2}{r_{12}} + \frac{q_1 q_3}{r_{13}} + \frac{q_2 q_3}{r_{23}}$$

Energy stored by N charges:

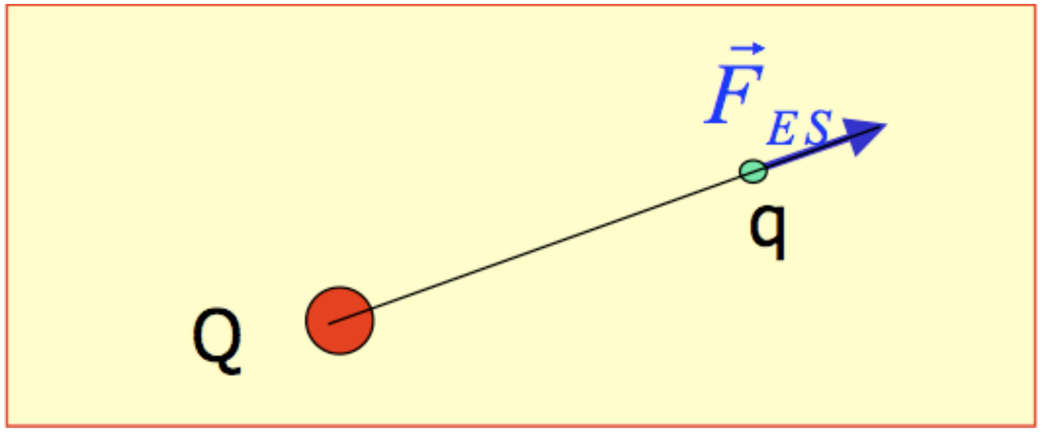
$$U = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \frac{q_i q_j}{r_{ij}}$$

The Electric Field: Question: what is the best way of describing the effect of charges?

1 charge in the Universe: no way

2 charges in the Universe:

$$\vec{F}_q = \frac{qQ}{r^2} \hat{r}$$



But: the force F depends on the test charge q… ☹

➙ define a quantity that describes the effect of the charge Q on the surroundings:

The Electric Field

$$\vec{E} = \frac{\vec{F}_q}{q} = \frac{Q}{r^2} \hat{r}$$

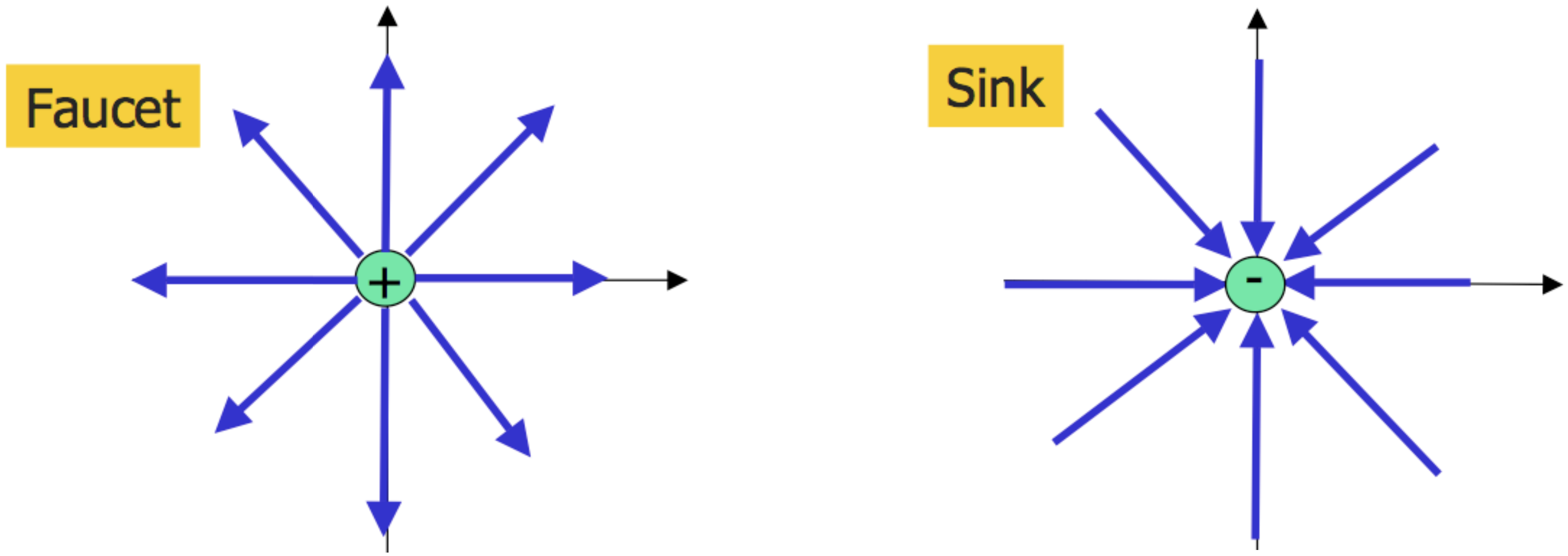
Units: dynes/esu

Electric field lines

Visualize the direction and strength of the Electric Field:

Direction: radial, , pointing towards – and away from +

Magnitude: the denser the lines - the stronger the field



Properties:

Field lines never cross (if so, that's where E=0)

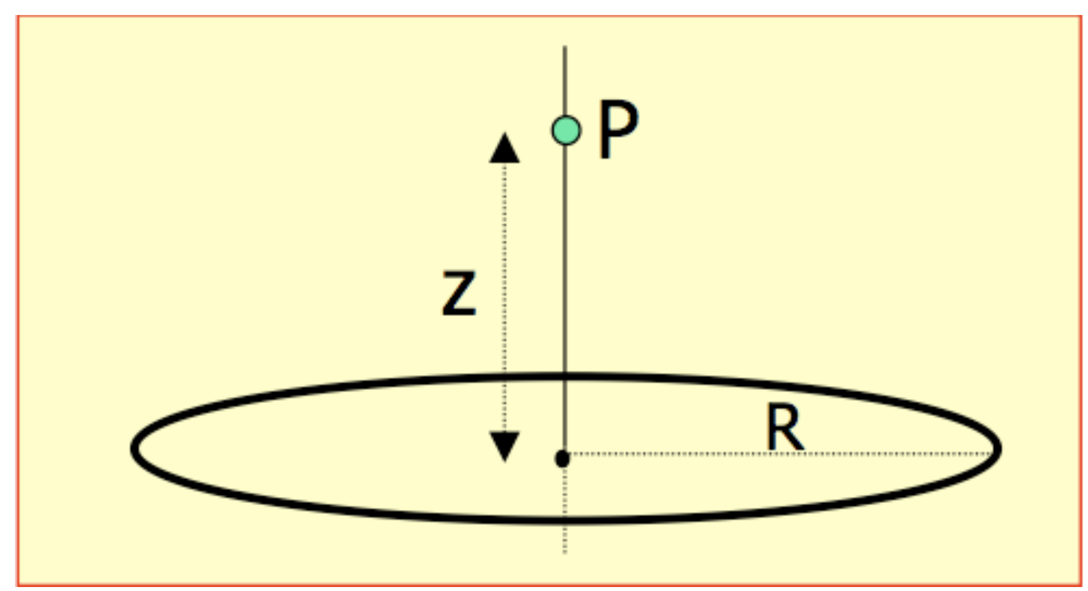
They are orthogonal to equipotential surfaces (will see this later).

Electric field of a ring of charge

Problem: Calculate the electric field created by a uniformly charged ring on its axis at P

Special case: center of the ring

General case: any point P on the axis



Answers:

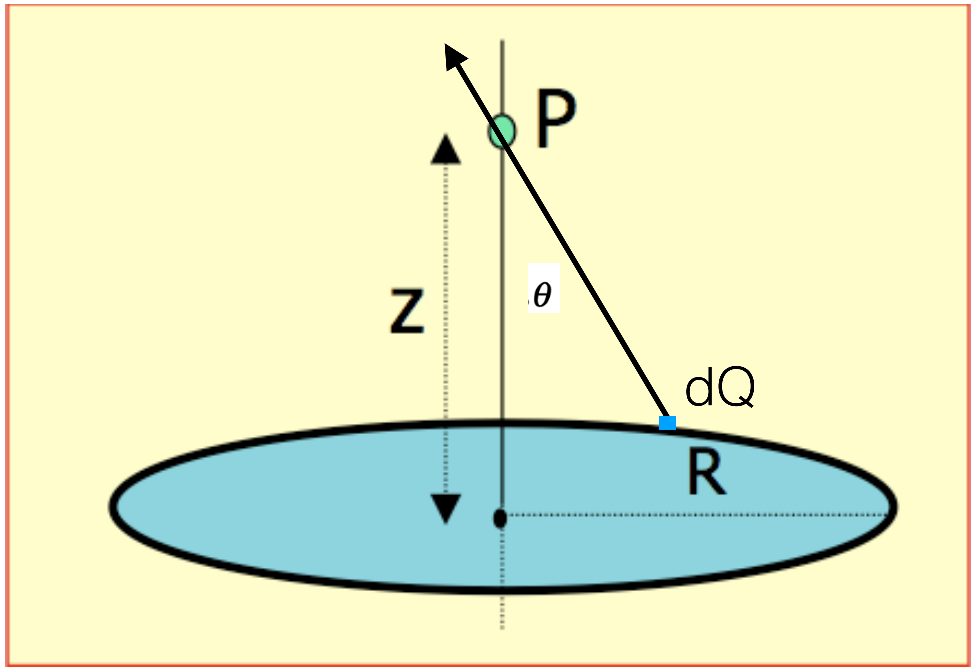
Center of the ring: E=0 by symmetry

General case:

$$\vec{E}_{ring} = \frac{Qz}{\left(R^2 + z^2\right)^{3/2}}\hat{z}$$

$$dE_z = \frac{dQ\cos\theta}{\sqrt{R^2 + z^2}} \quad , \quad \cos\theta = \frac{z}{\sqrt{R^2 + z^2}}$$

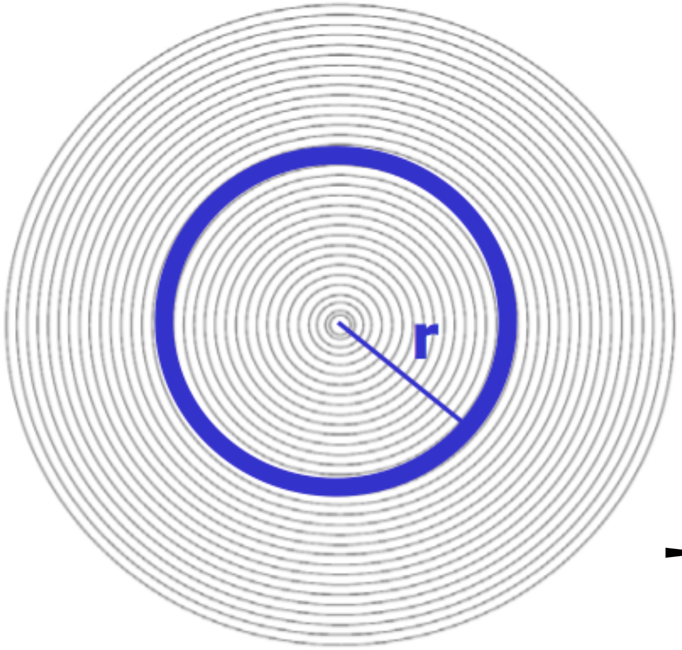
everything constant!!
—> just add up dQs



Electric field of disk of charge

Problem:
Find the electric field created by a disk of charges on the axis of the disk

Trick:
a disk is the sum of an infinite number of infinitely thin concentric rings.
And we know E_{ring}... (creative recycling is fair game in physics)



Electric field of a ring of radius r:

$$\vec{E}_{ring} = \frac{Qz}{\left(R^2 + z^2\right)^{3/2}}\hat{z}$$

If charge is uniformly spread:

$$dq = \sigma da = 2\pi r\sigma dr$$

→ Electric field created by the ring is:

$$d\vec{E} = \frac{2\pi r\sigma z dr}{\left(r^2 + z^2\right)^{3/2}}\hat{z}$$

Integrating on r: 0 ➞ R:

$$\vec{E} = \int_{r=0}^{r=R} d\vec{E} = \int_{r=0}^{r=R} \frac{2\pi r\sigma z dr}{\left(r^2+z^2\right)^{3/2}}\hat{z} = 2\pi\sigma z\hat{z}\left(\frac{1}{z} - \frac{1}{\sqrt{R^2+z^2}}\right)$$

Special case 1: R ➞ infinity: What if R infinity? E.g. what if R >> z?

Since $\lim_{R\to\infty} \frac{1}{\sqrt{R^2+z^2}} = 0$ we get $\vec{E} = 2\pi\sigma\hat{z}$

Conclusion:

Electric Field created by an infinite conductive plane:

Direction: perpendicular to the plane (+/- z)

Magnitude: $2\pi\sigma$ (constant!)

Special case 2: z >> R: What happens when z>>R? 2 approaches:

Physicist's approach:

The disk will look like a point charge with $Q = \sigma\pi R^2 \Rightarrow E = \frac{Q}{z^2}$

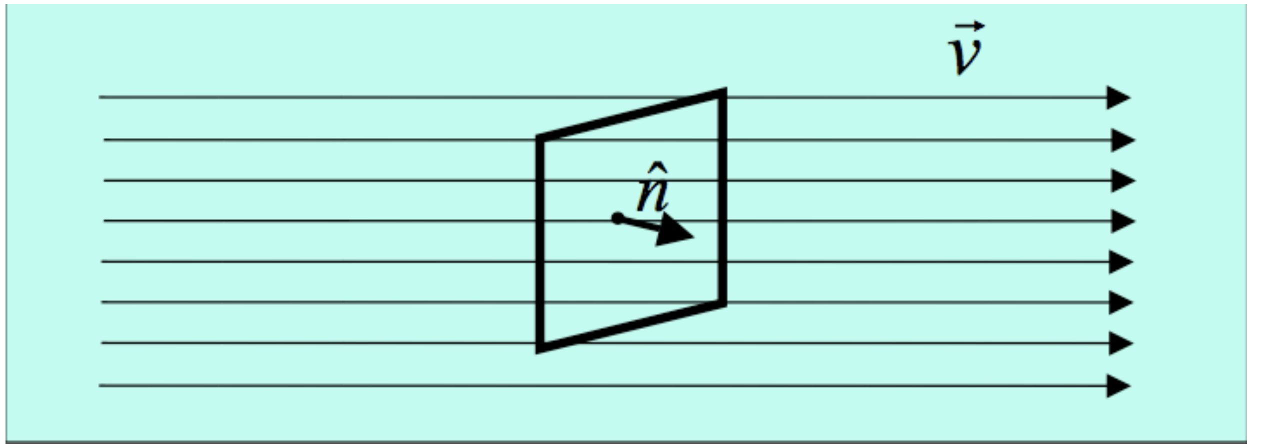
Mathematician's approach:

Calculate from the previous result for z >> R (Taylor expansion)

$$\vec{E} = 2\pi\sigma z\hat{z}\left(\frac{1}{z} - \frac{1}{\sqrt{R^2+z^2}}\right) = 2\pi\sigma z\hat{z}\frac{1}{z}\left(1 - \frac{z}{\sqrt{R^2+z^2}}\right) = 2\pi\sigma\hat{z}\left(1 - \frac{1}{\sqrt{1+\left(\frac{R}{z}\right)^2}}\right)$$

$$= 2\pi\sigma\hat{z}\left(1 - \left(1 - \frac{1}{2}\left(\frac{R}{z}\right)^2\right)\right) = \pi\sigma\hat{z}\left(\frac{R}{z}\right)^2 = \frac{Q}{z^2}\hat{z}$$

The concept of flux



Consider the flow of water in a river

The water velocity is described by

$$\vec{v}(x,y,z) = v_x\hat{x} + v_y\hat{y} + v_z\hat{z} = (v_x, v_y, v_z)$$

Immerse a square wire loop of area A in the water (surface S)

Define the loop area vector as $\vec{A} = A\hat{n}$

Q: how much water will flow through the loop? E.g.:

What is the "flux of the velocity" through the surface S?

It depends on how the loop is oriented wrt the water...

Assuming constant velocity and plane loop:

$$(1)\ if\ \vec{A} \perp \vec{v} \rightarrow \Phi_v(\text{flux of } \vec{v}) = 0$$

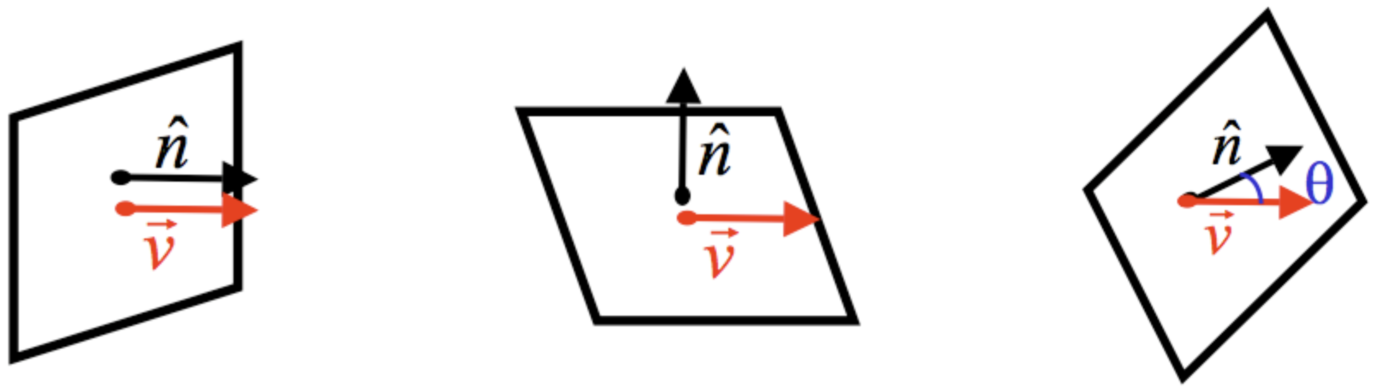
$$(2)\ if\ \vec{A} \parallel \vec{v} \rightarrow \Phi_v = vA$$

$$(3)\ if\ \vec{A} \measuredangle \vec{v} \rightarrow \Phi_v = vA\cos\theta = \vec{v}\cdot\vec{A}$$

Remember: area vector is orthogonal to area !

General case (definition of flux):

$$\Phi_v = \int_S \vec{v}\cdot d\vec{A}$$

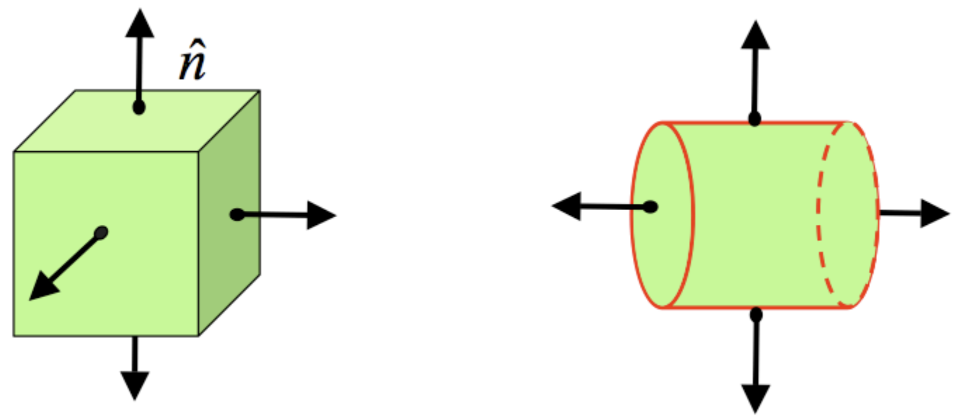


F.A.Q.: what is the direction of dA?

Defined unambiguously only for a 3d surface:

At any point in space, dA is perpendicular to the surface
It points towards the "outside" of the surface

Examples:



Intuitively: dA is oriented in such a way that if we have a hose inside the surface the flux through the surface will be positive

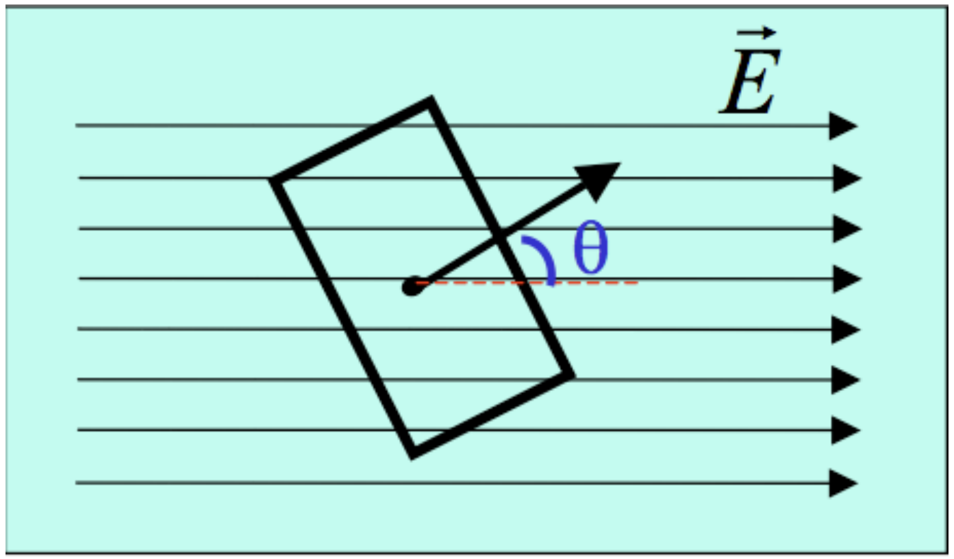
Flux of Electric Field

Definition:

$$\Phi_E = \Phi = \int_S \vec{E} \cdot d\vec{A}$$

Example: uniform electric field + flat surface.

Calculate the flux: $$\Phi = \int_S \vec{E} \cdot d\vec{A} = \vec{E} \cdot \vec{A} = EA\cos\theta$$



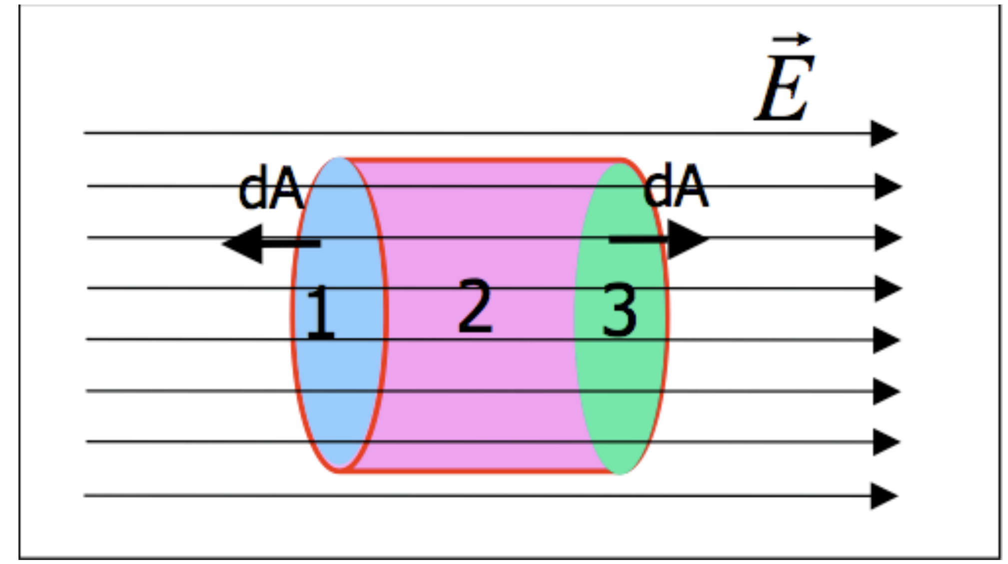
Interpretation: Represent E using field lines:
Φ_E is proportional to $N_{\text{field lines}}$ that go through the loop
NB: this interpretation is valid for any electric field and/or surface!

Φ_E through closed (3d) surface

Consider the total flux of E through a cylinder:

$$\Phi_{tot} = \Phi_1 + \Phi_2 + \Phi_3$$

Calculate Φ_1, Φ_2, Φ_3



Cylinder axis is // to field lines

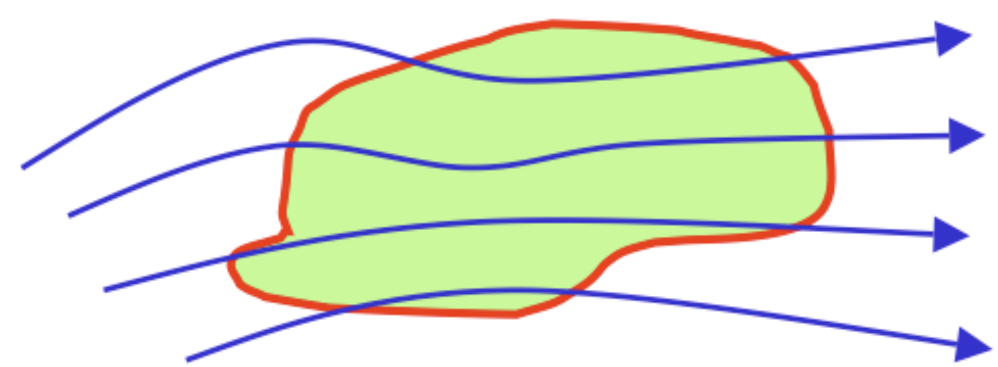
$\Phi_2 = 0$ because $\vec{E} \perp \hat{n}$

$|\Phi_1| = |\Phi_3|$ but opposite sign since $\Phi = \int_S \vec{E} \cdot d\vec{A} = \vec{E} \cdot \vec{A} = EA\cos\theta$

→ the total flux through the cylinder is zero!

Φ_E through closed empty surface - discussion:

Q1: Is this a coincidence due to shape/orientation of the cylinder?



Clue: Think about interpretation of Φ_E: proportional to # of field lines through a 3D surface …

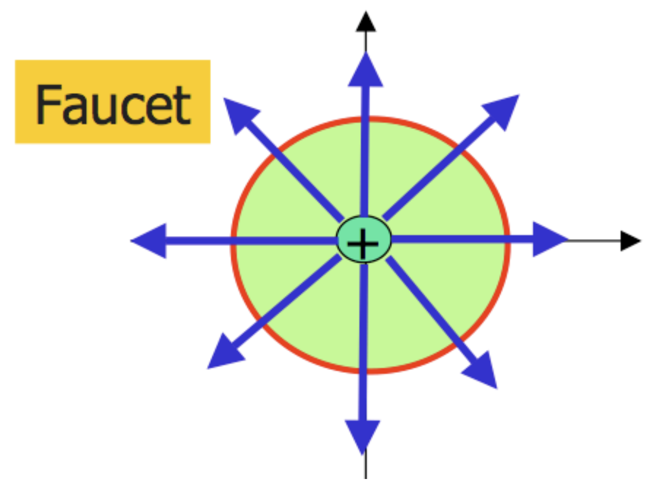
Conclusion: The electric flux through a closed surface that does not contain charges is zero.

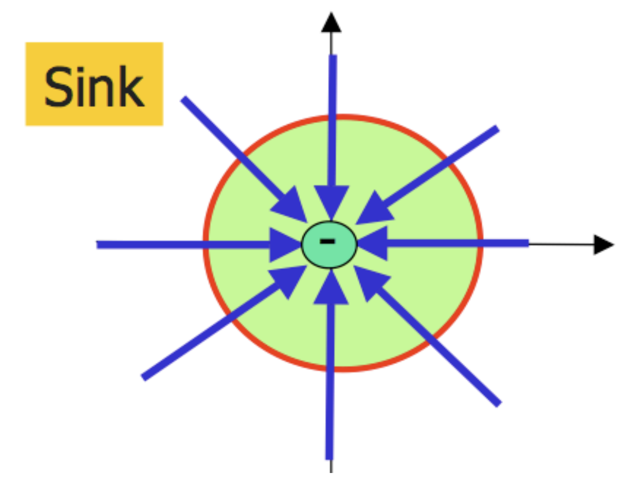
Φ_E through surface containing charge Q - discussion:

Q1: What if the surface contains charges?

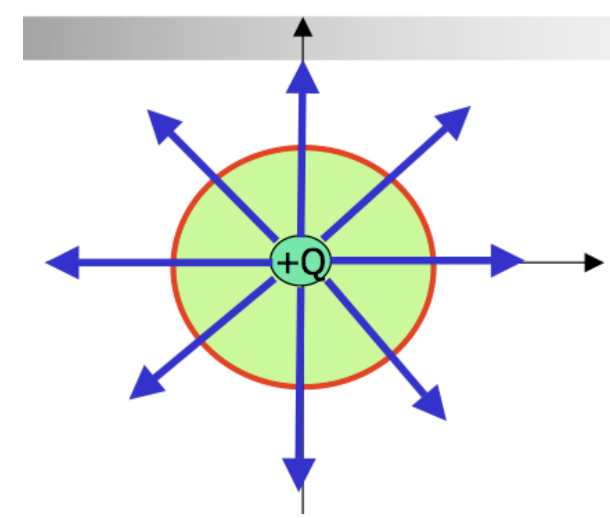
Clue: Think about interpretation of Φ_E: the lines will either originate inside the surface (positive flux) or terminate inside the surface (negative flux)

Conclusion: The electric flux through a closed surface that does contain a net charge is non zero.





Simple example: Φ_E of a charge at center of sphere



Problem: Calculate Φ_E for point charge +Q at the center of a sphere of radius R

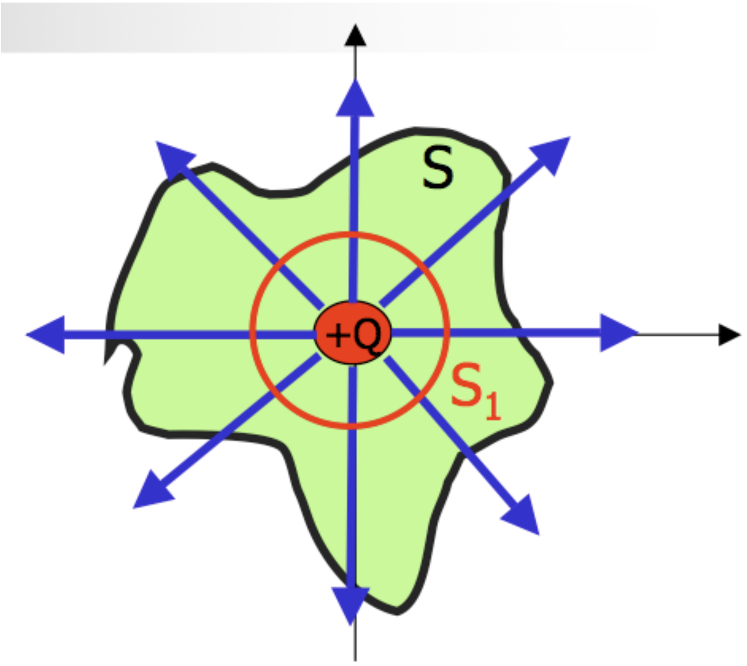
Solution :

$\vec{E} \parallel d\vec{A}$ everywhere on the sphere

Point charge at distance R: $\vec{E} = \dfrac{Q}{R^2}\hat{r}$

$$\Phi = \int_S \vec{E} \cdot d\vec{A} = \int_S \frac{Q}{R^2} dA = \frac{Q}{R^2}\int_S dA = \frac{Q}{R^2} 4\pi R^2 = 4\pi Q$$

Φ_E through a generic surface



What if the surface is not spherical ? Impossible integral?

Use intuition and interpretation of flux!

Consider the sphere S_1. Field lines are always continuous

$$\rightarrow \Phi_{S_1} = \Phi_S = 4\pi Q$$

(more rigorous demonstration later)

Conclusion: The electric flux Φ through any closed surface S containing a net charge Q is proportional to the charge enclosed:

$$\Phi = \oint_S \vec{E} \cdot d\vec{A} = 4\pi Q_{enclosed}$$ Gauss's law

Thoughts on Gauss's law: The form above is Gauss's law in integral form.

Why is Gauss's law so important? Because it relates the electric field E with its sources Q.

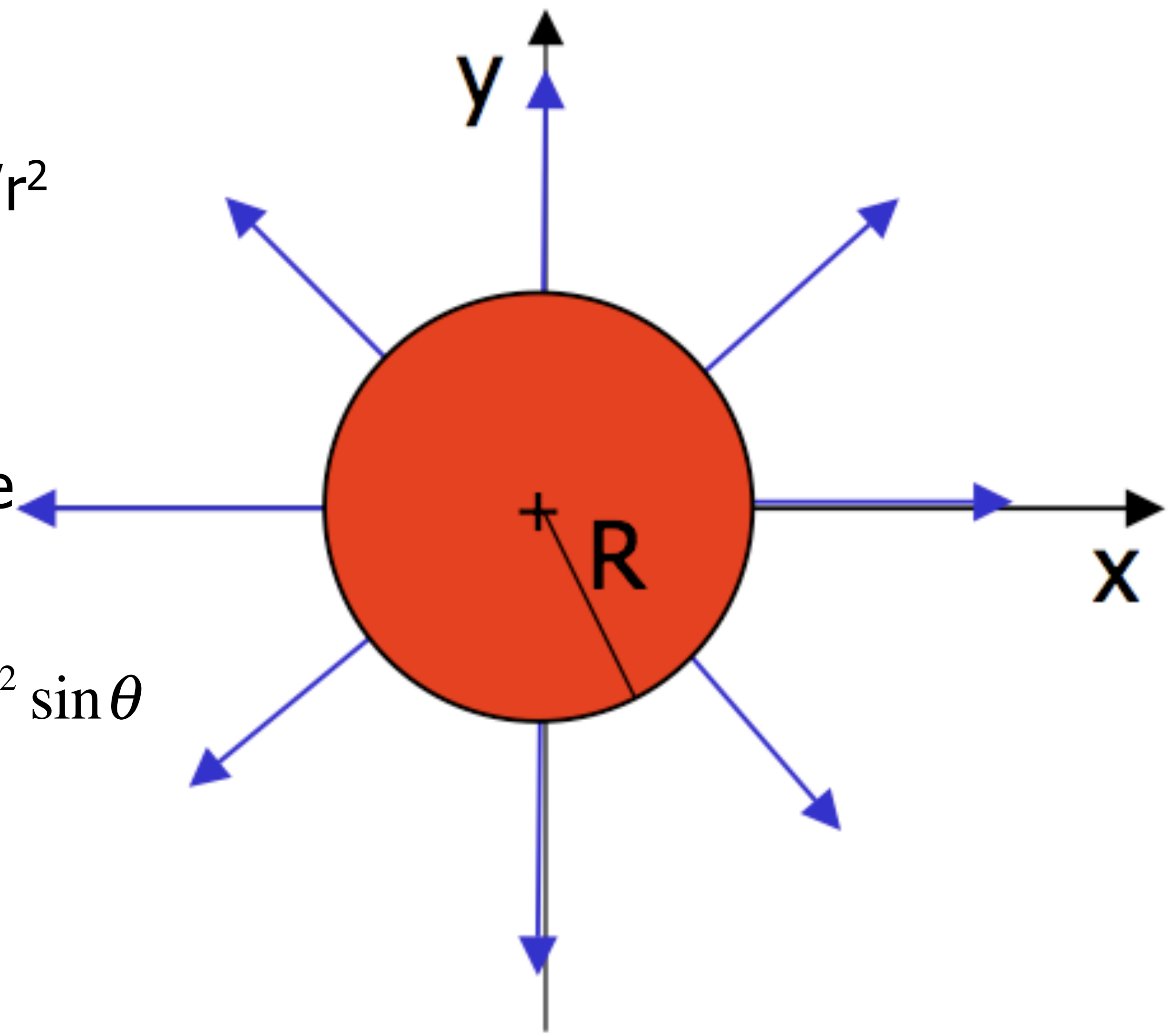
Given Q distribution → find E (integral form)
Given E → find Q (differential form - later)

Is Gauss's law always true? Yes, no matter what E or what S, the flux is always = $4\pi Q$

Is Gauss's law always useful? No, it's useful only when the problem has symmetries

Applications of Gauss's law:
Electric field of spherical distribution of charges

Problem: Calculate the electric field (everywhere in space) due to a spherical distribution of positive charges or radius R. (NB: solid sphere with volume charge density ρ)



Approach #1 (mathematician)

I know the E due to a point charge dq: $dE=dq/r^2$

I know how to integrate

Solve the integral inside and outside the sphere (e.g. r < R and r > R)

$$\int_{r'=0}^{r'=r} d\vec{E} = \int_{r'=0}^{r'=r} \frac{dq}{r'^2}\hat{r} = \int_{r'=0}^{r'=r} \frac{\rho dV}{r'^2}\hat{r}' = \int d\theta \int d\phi \int_{r'=0}^{r'=r} \frac{\rho}{r'^2}\hat{r}' r'^2 \sin\theta$$

Comment: correct but usually heavy on math!

Approach #2 (physicist)

Why would I ever solve an integral if somebody (Gauss) already did it for me?

Just use Gauss's law!

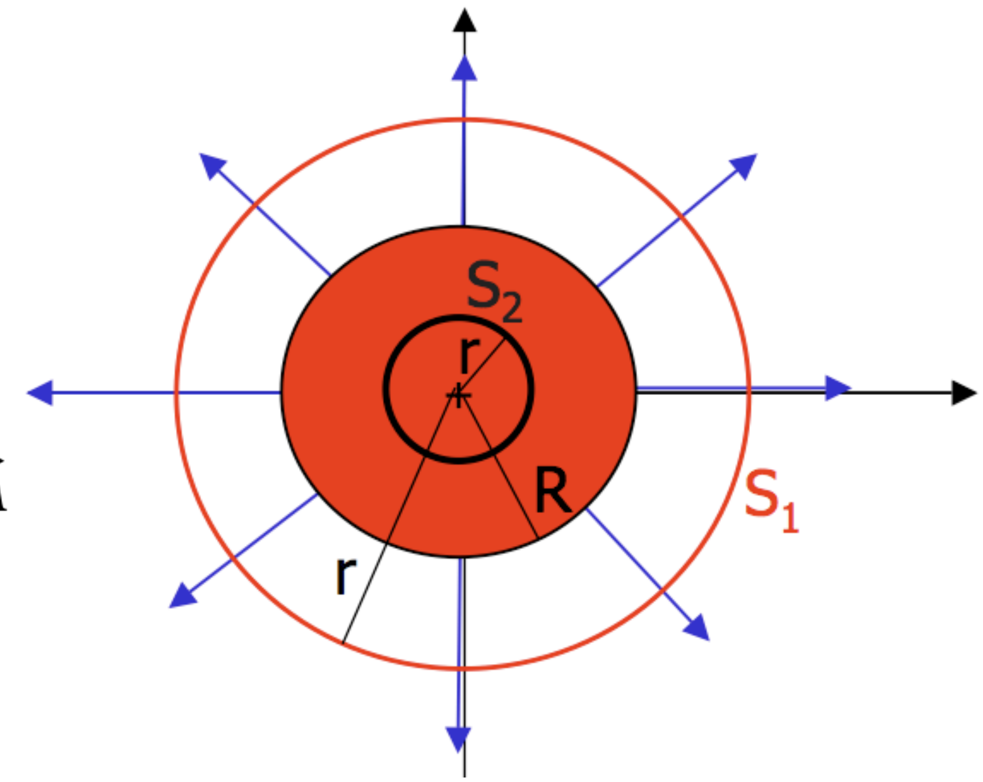
Comment: correct, much much less time consuming!

Physicist's solution:

Outside the sphere (r > R)

Apply Gauss on a sphere S_1 of radius r :

$$\Phi = \oint_{s_1} \vec{E} \cdot d\vec{A} = 4\pi Q_{enclosed}$$



Symmetry: E is constant on S and parallel to $d\vec{A}$

$$\oint_{s_1} \vec{E} \cdot d\vec{A} = E4\pi r^2 = 4\pi Q_{enclosed} \rightarrow E = \frac{Q}{r^2}$$

For r > R sphere looks like a point charge!

Inside the sphere (r < R)

Apply Gauss on a sphere S_2 of radius r :

Again: $\Phi = \oint_{s_2} \vec{E} \cdot d\vec{A} = 4\pi Q_{enclosed}$; symmetry: E is constant on S and || to $d\vec{A}$

$$\oint_{s_2} \vec{E} \cdot d\vec{A} = E4\pi r^2 \qquad Q_{enclosed} = \int \rho dV = \rho \frac{4}{3}\pi r^3$$

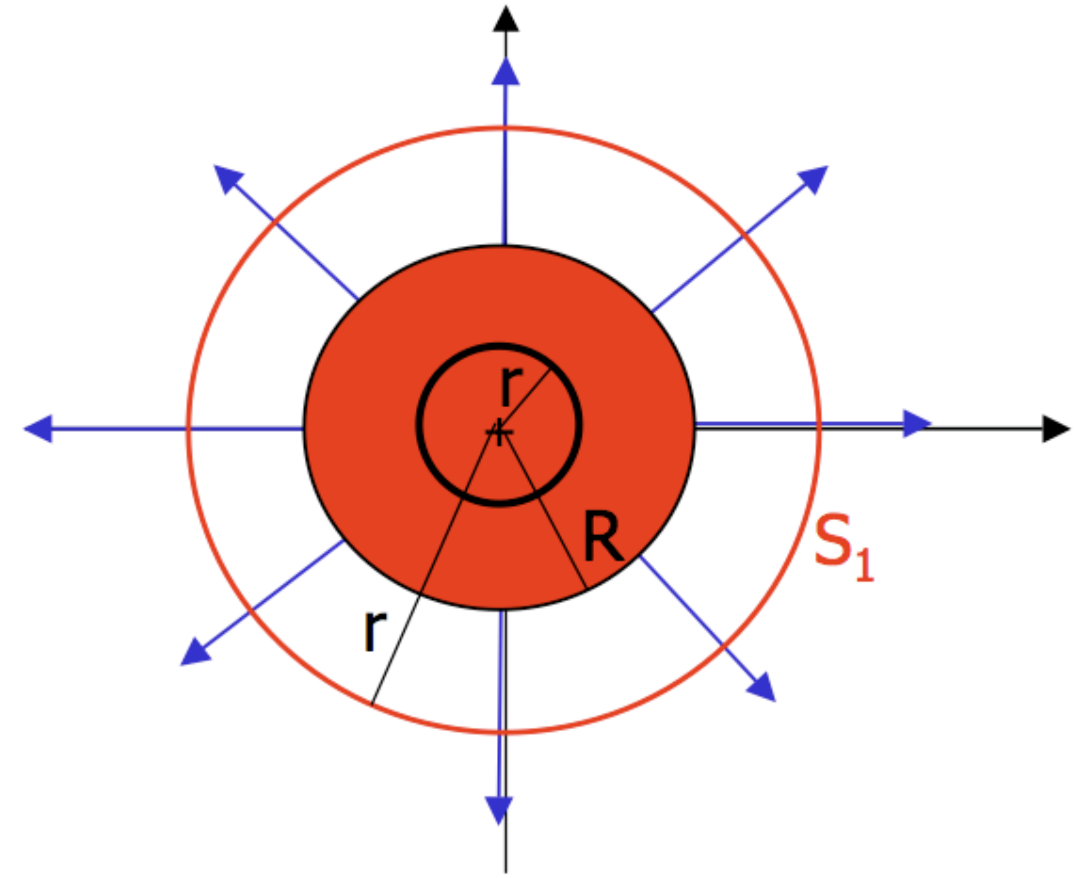
$$E4\pi r^2 = 4\pi Q_{enclosed} = 4\pi\rho \frac{4}{3}\pi r^3 \rightarrow E = \frac{4}{3}\pi\rho r$$

Do I get full credit for this solution?

Did I answer the question completely?

No! I was asked to determine the electric field.
The electric field is a vector

➙ both magnitude and direction needed



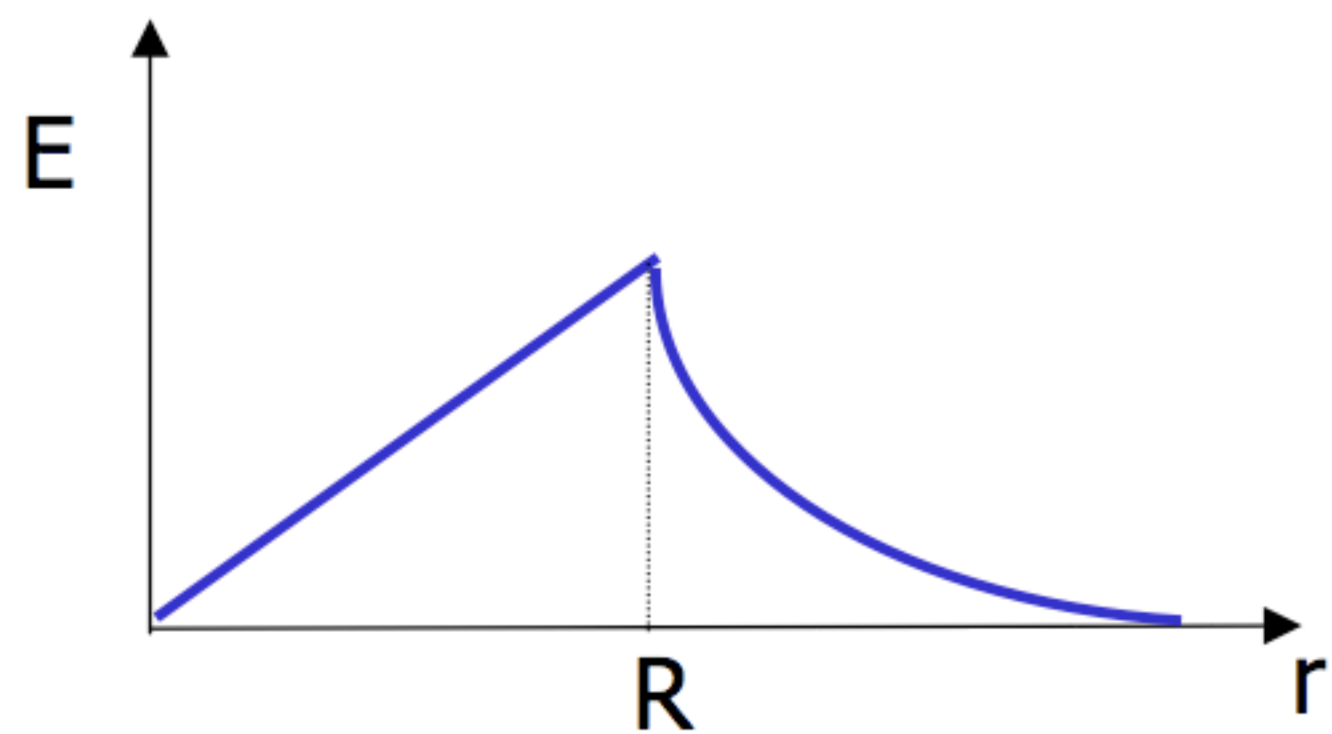
How to get the E direction?

Look at the symmetry of the problem:
Spherical symmetry ➙ E must point radially

Complete solution:

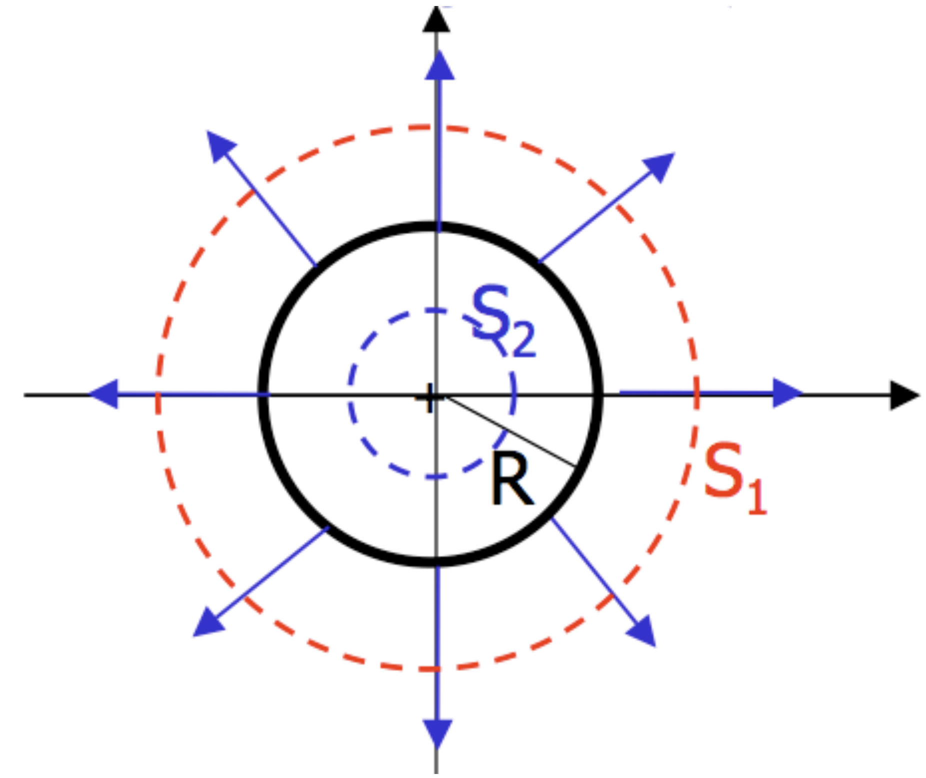
$$\vec{E} = \frac{Q}{r^2}\hat{r} \quad for \quad r > R$$

$$\vec{E} = \frac{4}{3}\pi\rho r\hat{r} \quad for \quad r < R$$



Another application of Gauss's law:
Electric field of spherical shell

Problem: Calculate the electric field (everywhere in space) due to a positively charged spherical shell or radius R (surface charge density σ)



spherical symmetry → E is radial

Physicist's approach: apply Gauss

Outside the sphere (r > R)

Apply Gauss on a sphere S_1 of radius r :

$$\Phi = \oint_{S_1} \vec{E} \cdot d\vec{A} = 4\pi Q_{enclosed}$$

Symmetry: E is constant on S_1 and || to d$\vec{A}$ (radial)

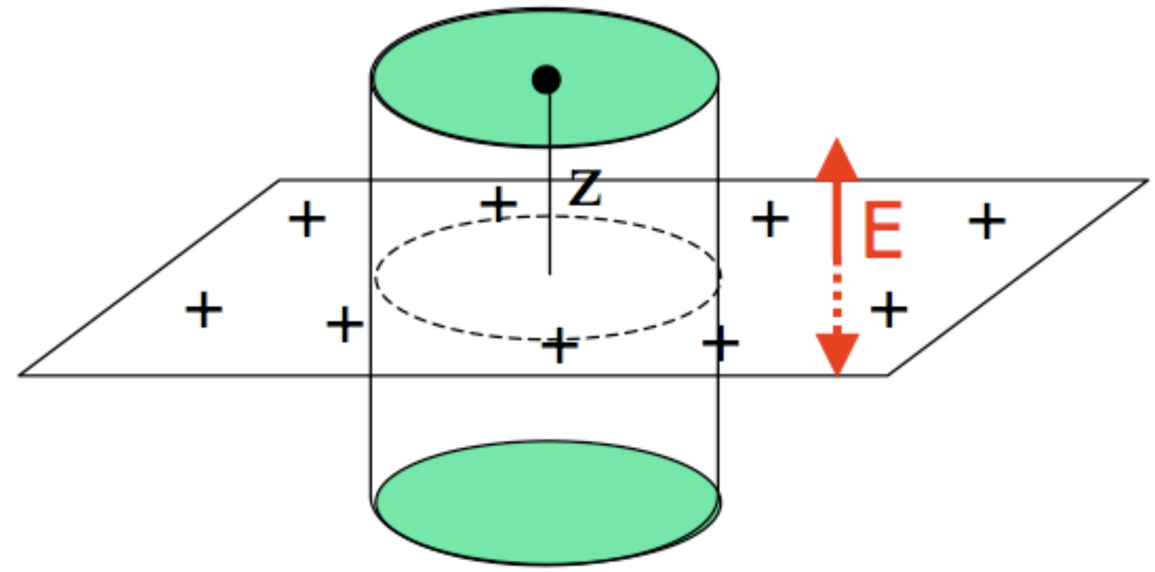
$$\Phi = \oint_{S_1} \vec{E} \cdot d\vec{A} = E 4\pi r^2 = 4\pi Q_{enclosed} = 4\pi\sigma 4\pi R^2$$

$$\rightarrow \vec{E} = \frac{4\pi\sigma R^2}{r^2}\hat{r} = \frac{Q}{r^2}\hat{r}$$ same as point charge!

Inside the sphere (r < R)

Apply Gauss on a sphere S_2 of radius r. But sphere is hollow → $Q_{enclosed} = 0$ → E = 0

Another application of Gauss's law:
Electric field of infinite sheet of charge



Problem: Calculate the electric field at a distance z from a positively charged infinite plane of surface charge density σ

Again apply Gauss
Trick #1: choose the right Gaussian surface!
Look at the symmetry of the problem
Choose a cylinder of area A and height +/- z
Trick #2: apply Gauss's theorem

$$\Phi_{tot} = \Phi_{side} + \Phi_{top} + \Phi_{bottom}$$

Symmetry: E || z axis $\Phi_{side} = 0$ and $\Phi_{top} = \Phi_{bottom}$

$$\Phi = \oint_{cylinder} \vec{E} \cdot d\vec{A} = 4\pi Q_{enclosed}$$

$$\oint_{cylinder} \vec{E} \cdot d\vec{A} = 2\int_{top} E dA = 2EA = 4\pi(\sigma A)$$

$$\rightarrow \vec{E} = 2\pi\sigma\hat{z}$$

same as earlier!

Checklist for solving problems

I told my class this when teaching…..

Read the problem (I am not joking!)

Look at the symmetries before choosing the best coordinate system

Look at the symmetries again and find out what cancels what and the direction of the vectors involved

Look for a way to avoid all complicated integration

Remember physicists are lazy: complicated integral you screwed up somewhere or there is an easier way out!

Turn the math crank…

Write down the complete solution(magnitudes and directions for all the different regions)

Box the solution: your graders will love you!

If you encounter expansions:

Find your expansion coefficient ($x<<1$) and “massage” the result until you get something that looks like $(1+x)^N$, $(1-x)^N$, or $\ln(1+x)$ or e^x

Don’t stop the expansion too early: Taylor expansions are more than limits…

Summary and outlook

What have we learned so far:

Energy of a system of charges

Concept of electric field E
To describe the effect of charges independently from the test charge

Gauss's theorem in integral form:

$$\Phi = \oint_s \vec{E} \cdot d\vec{A} = 4\pi Q_{enclosed}$$

Future lectures:

Derive Gauss's theorem in a more rigorous way (Purcell section 1.10)

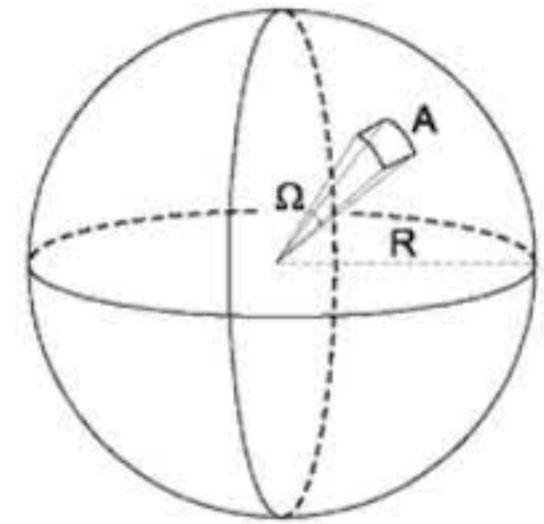
Gauss's law in differential form (more about vector calculus)
Useful to derive charge distribution given the electric fields

Energy associated with an electric field

We now discuss these topics:

- Electric potential
- Energy associated with an electric field
- Gauss's law in differential form
  - … and a lot of vector calculus… (yes, again!)

$$d\Omega = \frac{dA}{R^2}$$



What have we learned so far?

Energy of a system of charges $$U = \frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{q_i q_j}{r_{ij}}$$

Electric field $$\vec{E} = \frac{\vec{F}_q}{q} = \frac{Q}{r^2}\hat{r}$$

Gauss's law in integral form: $$\Phi = \oint_S \vec{E}\cdot d\vec{A} = 4\pi Q_{enclosed}$$

Derived last time, but not rigorously…

Just used electric field lines

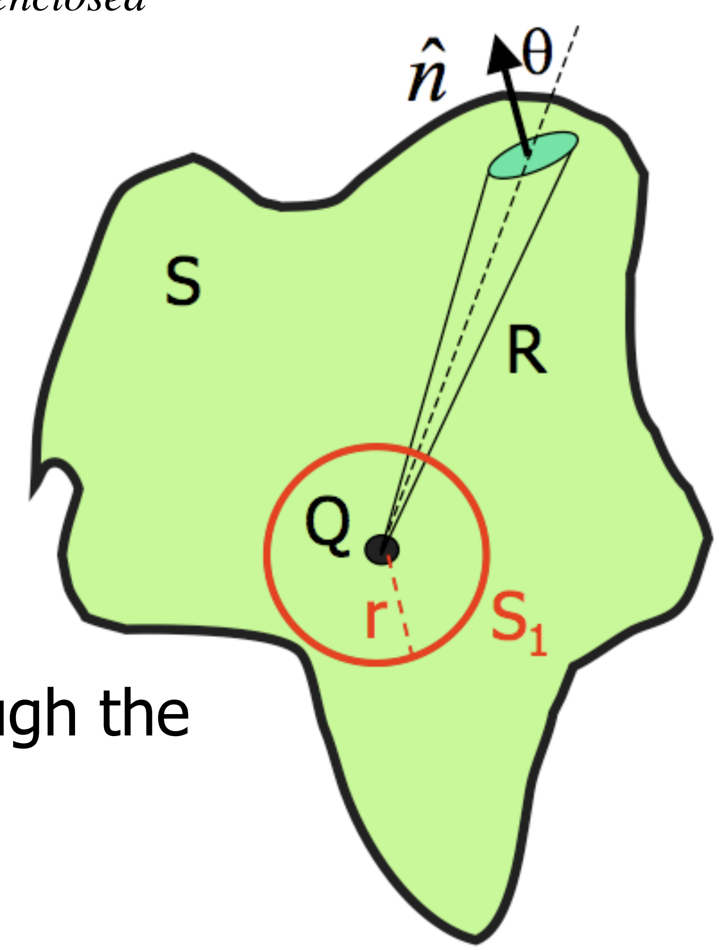
Gauss's law

Consider charge in a generic surface S

Surround charge with spherical surface S_1 concentric to charge

Consider cone of solid angle dΩ from charge to surface S through the little sphere

What is solid angle of hemisphere? of sphere?



Electric flux through little sphere:

$$d\Phi_S = \vec{E}\cdot d\vec{A} = \left(\frac{Q}{r^2}\hat{r}\right)\cdot\left(r^2 d\Omega\hat{r}\right) = Q d\Omega$$

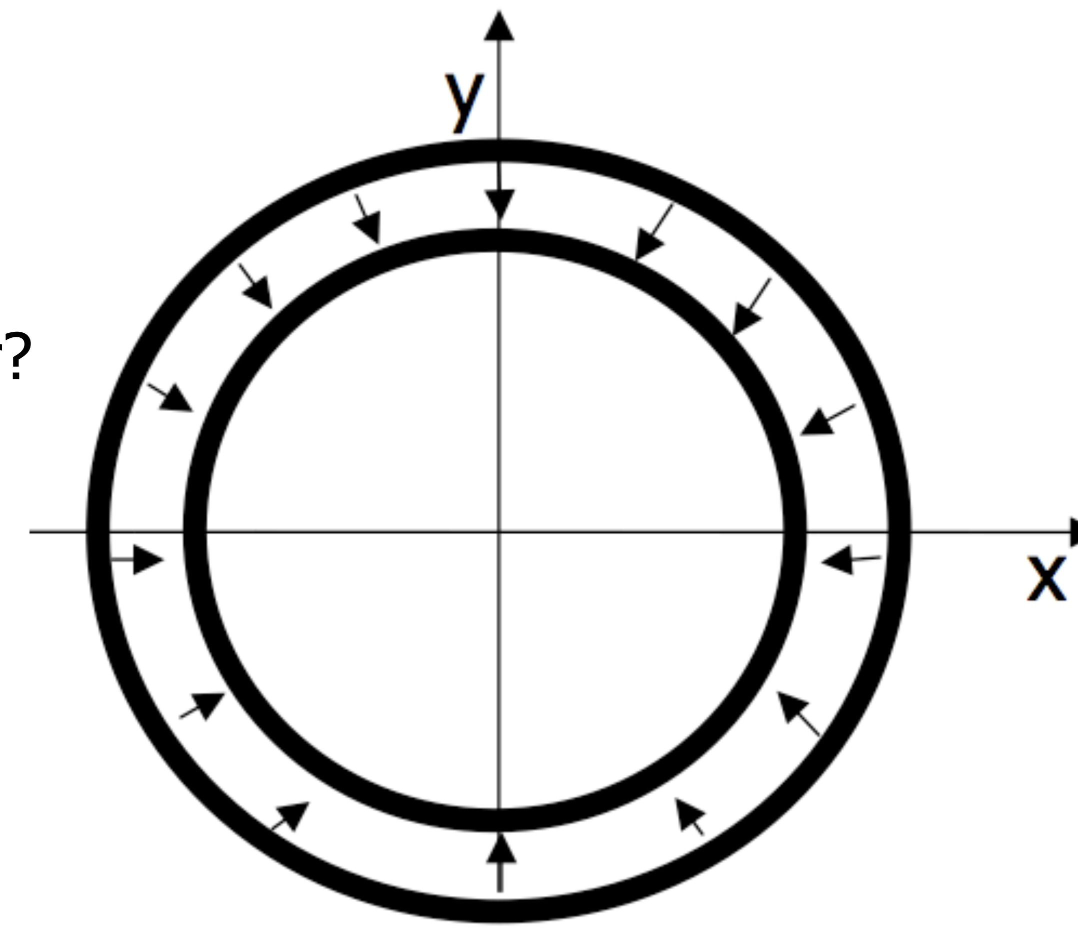
$$d\Phi_S = d\Phi_{S_1} \rightarrow \Phi_S = \Phi_{S_1} = 4\pi Q$$

$$\Phi = \oint_S \vec{E}\cdot d\vec{A} = 4\pi Q_{enclosed}$$ is valid for ANY shape S

Gauss's law is valid because E~1/r^2. If E ~ anything else, the r^2 would not cancel!!!

Energy stored in E field: Squeezing charges…

Consider a spherical shell of charge of radius r



How much work dW to "squeeze" it to a radius r-dr?

$$dW = Fdr = (PA)dr = P(Adr) = PdV$$

Calculate the pressure necessary to squeeze it:

$$P = \frac{F}{A} = \frac{QE}{A} = E\frac{Q}{A} = E\sigma$$

$$E_{outside} = \frac{Q}{r^2} \quad , \quad E_{inside} = 0 \rightarrow E_{surface} = \frac{1}{2}\frac{Q}{r^2}$$

$$\rightarrow P = E\sigma = \frac{1}{2}\frac{Q}{r^2}\sigma = \frac{\sigma}{2r^2}\left(4\pi r^2\sigma\right) = 2\pi\sigma^2$$

We can now calculate dW: $dW = PdV = 2\pi\sigma^2 dV$

Remembering that $E = 4\pi\sigma \quad \Rightarrow dW = \frac{E^2}{8\pi} dV$

Energy stored in the electric field

Work done on the system: $dW = \frac{E^2}{8\pi} dV$

We do work on the system (dW): same sign charges have been squeezed on a smaller surface, closer together and they do not like that…

Where does the energy go?

We created electric field where there was none (between r and r-dr)

The electric field we created must be storing the energy

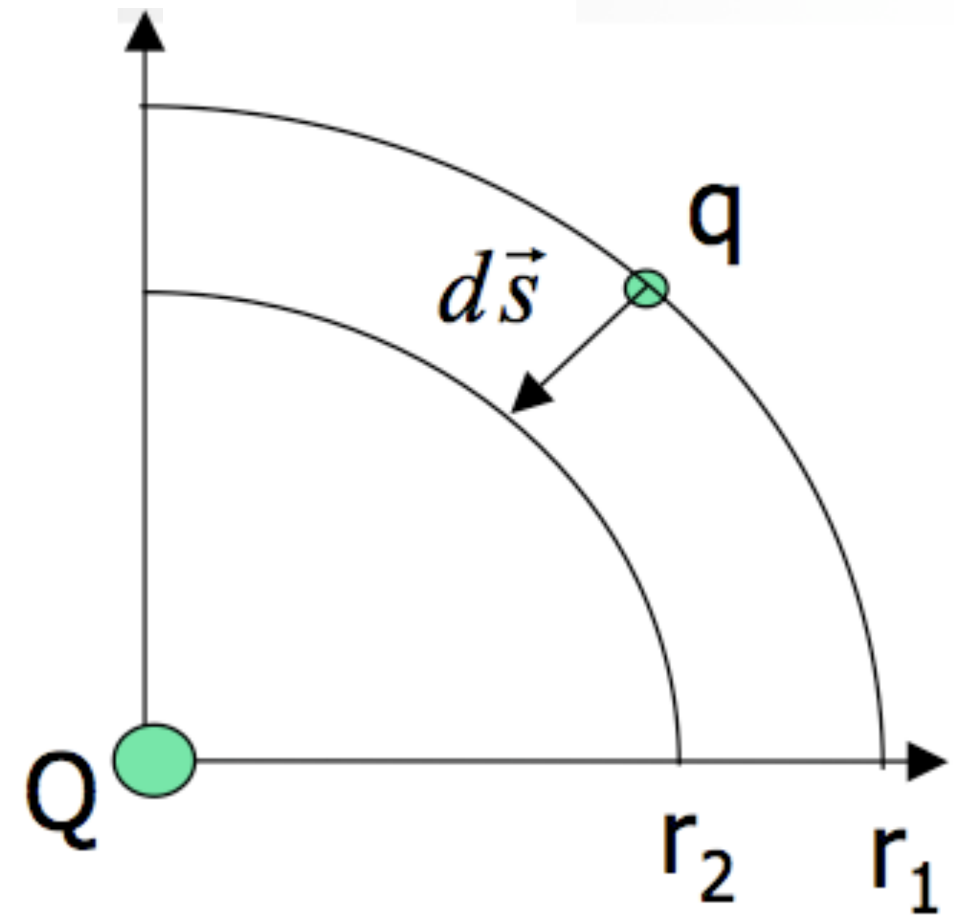
Energy is conserved → dU = dW

$u = \frac{E^2}{8\pi}$ is the energy density of the electric field E

Energy is stored in the E field: $U = \int_{\substack{Entire \\ space}} \frac{E^2}{8\pi} dV$

Note: One must integrate over entire space not only where charges are!

Electric potential difference



Work to move q from r_1 to r_2:

$$W_{12} = \int_1^2 \vec{F}_I \cdot d\vec{s} = -\int_1^2 \vec{F}_{Coulomb} \cdot d\vec{s} = -q\int_1^2 \vec{E} \cdot d\vec{s}$$

W_{12} depends on the test charge q ☹

➙ define a quantity that is independent of q and just describes the properties of the space:

$$\phi_{12} = \frac{W_{12}}{q} = -\int_1^2 \vec{E} \cdot d\vec{s}$$

Electric potential difference between P_1 and P_2

Physical interpretation: ϕ_{12} is work that I must do to move a unit charge from P_1 to P_2

Units: cgs: statvolts = erg/esu; SI: Volt = J/C; 1 statvolts = "3" 10^2 V

Electric potential

The electric potential difference ϕ_{12} is defined as the work to move a unit charge between P_1 to P_2: We need 2 points! ☺

Can we define similar concept describing the properties of the space?
Yes, just fix one of the points (e.g.: P_1 = infinity): ☺

$$\phi(\vec{r}) = -\int_\infty^{\vec{r}} \vec{E} \cdot d\vec{s} \Rightarrow Potential$$

Potential vs potential difference

Given a point charge q in the origin: Calculate $\phi(r)$

$$\phi(\vec{r}) = -\int_{\infty}^{\vec{r}} \vec{E} \cdot d\vec{s} = -\int_{\infty}^{r} \frac{q}{r^2} dr = \frac{q}{r}$$

Calculate potential difference between points P_1 and P_2:

$$\phi_{12} = -\int_{\vec{r}_1}^{\vec{r}_2} \vec{E} \cdot d\vec{s} = \frac{q}{r_2} - \frac{q}{r_1} = \phi(P_2) - \phi(P_1)$$

➙ Potential difference is really the difference of potentials!

Potentials of standard charge distributions

The potential created by a point charge is $\phi(\vec{r}) = \frac{q}{r}$

➙ Given this + superposition we can calculate anything!

Potential of N point charges: $\phi(\vec{r}) = \sum_{i=1}^{N} \frac{q_i}{r_1}$

Potential of charges in a volume V: $\phi(\vec{r}) = \int_V \frac{\rho dV}{r}$

Potential of charges on a surface S: $\phi(\vec{r}) = \int_S \frac{\sigma dA}{r}$

Potential of charges on a line L: $\phi(\vec{r}) = \int_L \frac{\lambda dl}{r}$

Some thoughts on potential

Why is potential useful? Isn't E good enough?

Potential is a scalar function → much easier to integrate than electric field or force that are vector functions

When is the potential defined?

Unless you set your reference somehow, the potential has no meaning

Usually we choose ϕ(infinity)=0

This does not work always: e.g.: potential created by a line of charges

Careful: do not confuse potential ϕ(x,y,z) with potential energy of a system of charges (U)

Potential energy of a system of charges:

= work done to assemble charge configuration $U = \frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{q_i q_j}{r_{ij}}$

Potential: work to move test charge from infinity to (x,y,z) $\phi(\vec{r}) = \sum_{i=1}^{N}\frac{q_i}{r_1}$

Energy of electric field revisited

Energy stored in a system of charges: $U = \frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{q_i q_j}{r_{ij}}$

This can be rewritten as follows:

$$U = \frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{q_i q_j}{r_{ij}} = \frac{1}{2}\sum_{j\neq i} q_j \sum_{i}\frac{q_i}{r_{ij}} = \frac{1}{2}\sum_{j\neq i} q_j \phi(r_j)$$

where $\phi(r_j)$ is the potential due to all charges excepted for the q_j at the location of q_j (r_j)

Taking a continuum limit: $$U = \frac{1}{2}\int_{\substack{Volume\\with\\charges}} \rho\phi(r)dV = \int_{\substack{Entire\\space}} \frac{E^2}{8\pi}dV$$

which works only when ϕ(infinity) = 0

Connection between ϕ and E

Consider potential difference between a point at r and r+dr :

$$d\phi = -\int_{\vec{r}}^{\vec{r}+d\vec{r}} \vec{E}\cdot d\vec{s} \sim \vec{E}(\vec{r})\cdot d\vec{r}$$

The infinitesimal change in potential can be written as:

$$d\phi = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy + \frac{\partial\phi}{\partial z}dz \equiv \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z}\right)\cdot\left(dx, dy, dz\right) \equiv \nabla\phi\cdot d\vec{r}$$

$$\vec{E} = -\nabla\phi$$

Useful info because it allows us to find E given ϕ

Good because ϕ is much easier to calculate than E

Getting familiar with gradients…

1d problem: $$\nabla f(x) = \frac{\partial f}{\partial x}\hat{x}$$

The derivative $\frac{\partial f}{\partial x}$ describes the function's slope

➙ the gradient describes the change of the function and the direction of the change

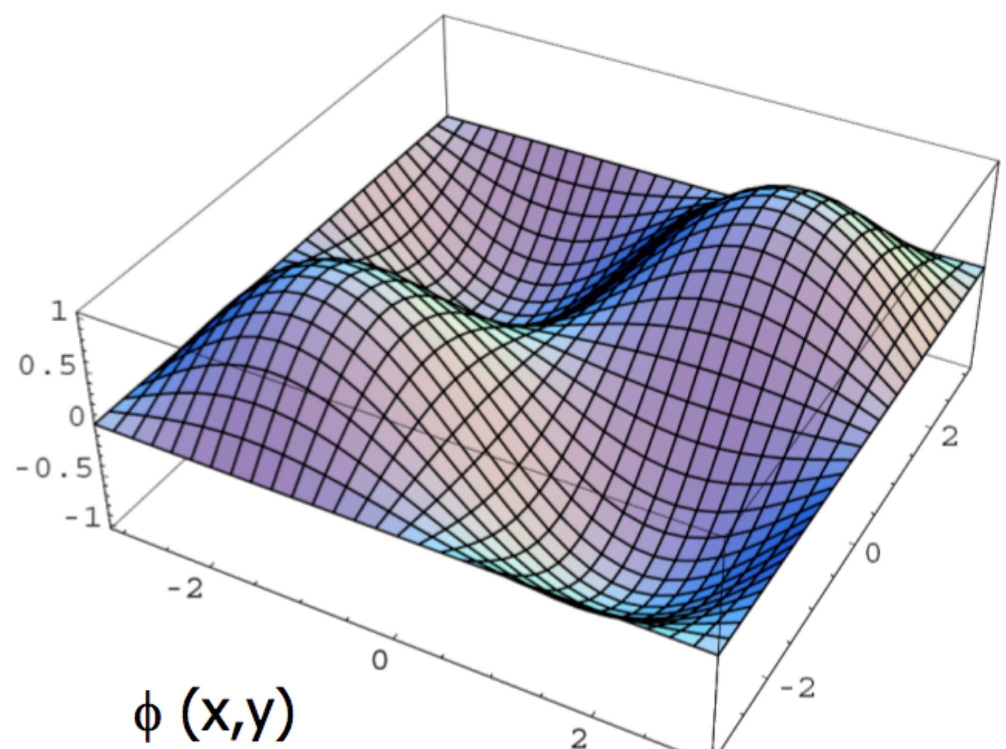
2d problem: $$\nabla f(x,y) = \frac{\partial f}{\partial x}\hat{x} + \frac{\partial f}{\partial y}\hat{y} \equiv \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

Interpretation is the same, but in both directions ➙ gradient points in direction where the slope is **steepest**

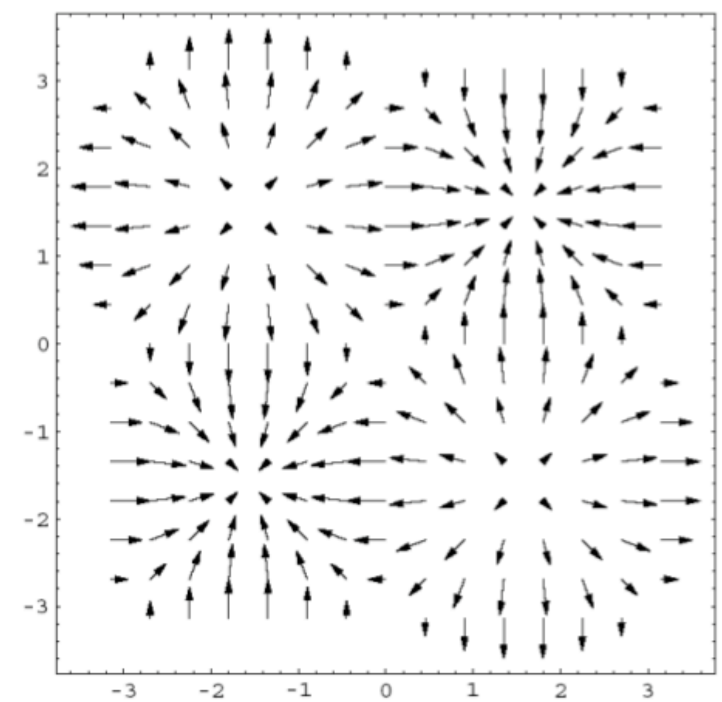
Visualization of gradients

Given the potential φ(x,y) = sin(x)sin(y), calculate its gradient.

$$\nabla \phi(x,y) = \frac{\partial \phi}{\partial x}\hat{x} + \frac{\partial \phi}{\partial y}\hat{y} = \cos(x)\sin(y)\hat{x} + \sin(x)\cos(y)\hat{y}$$



φ (x,y)

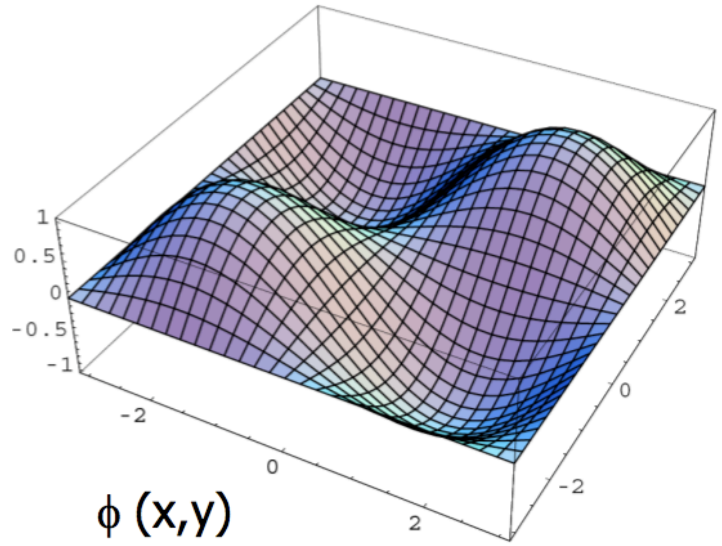


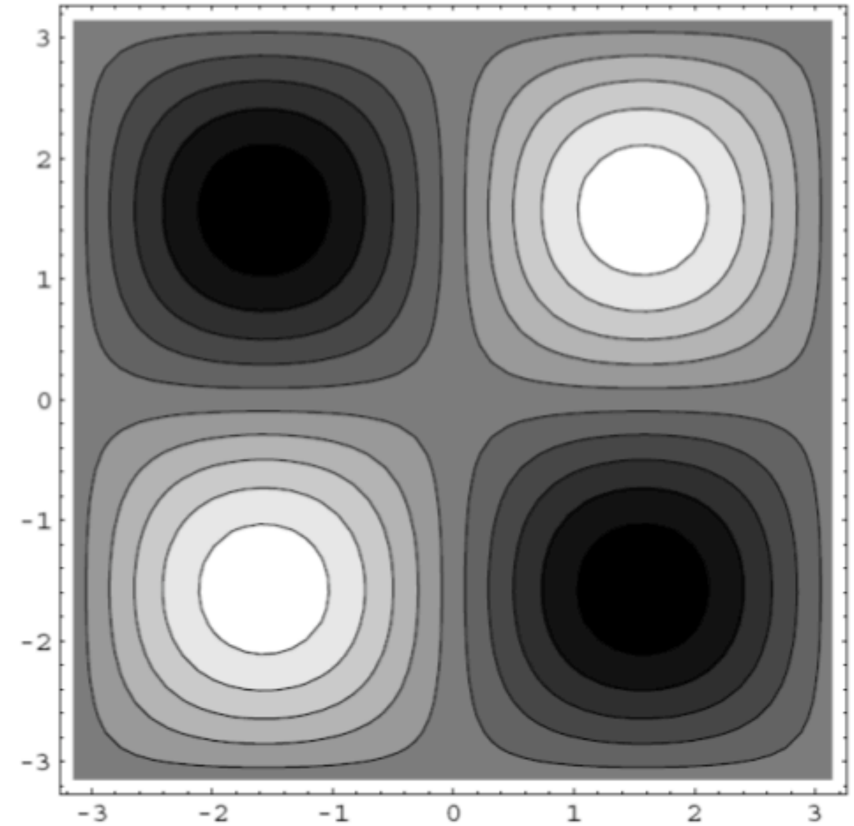
grad φ

The gradient always points uphill ➙ $\vec{E} = -grad\phi$ points downhill

Visualization of gradients: equipotential surfaces

Same potential $\phi(x,y) = \sin(x)\sin(y)$





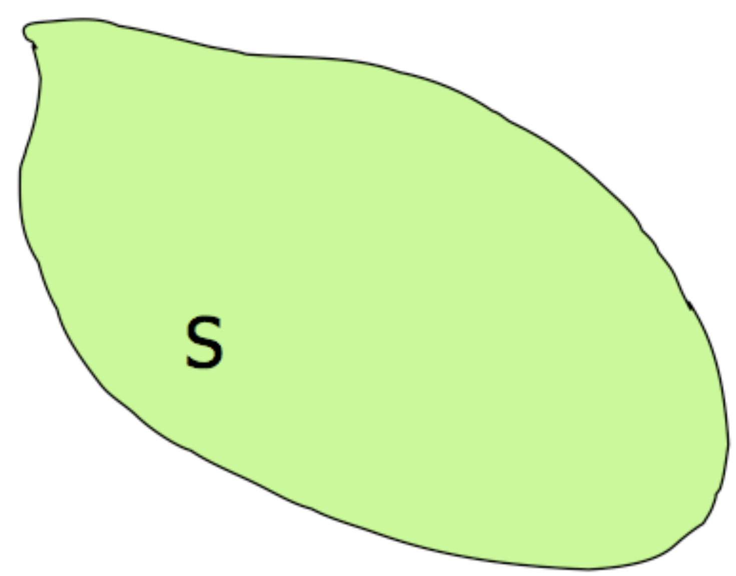
Equipotential lines are perpendicular to the gradient ➙ equipotential lines are always perpendicular to E

i.e., consider an equipotential line: displacement vector is tangent; $d\Phi=0 = \nabla\Phi\cdot dr = 0$

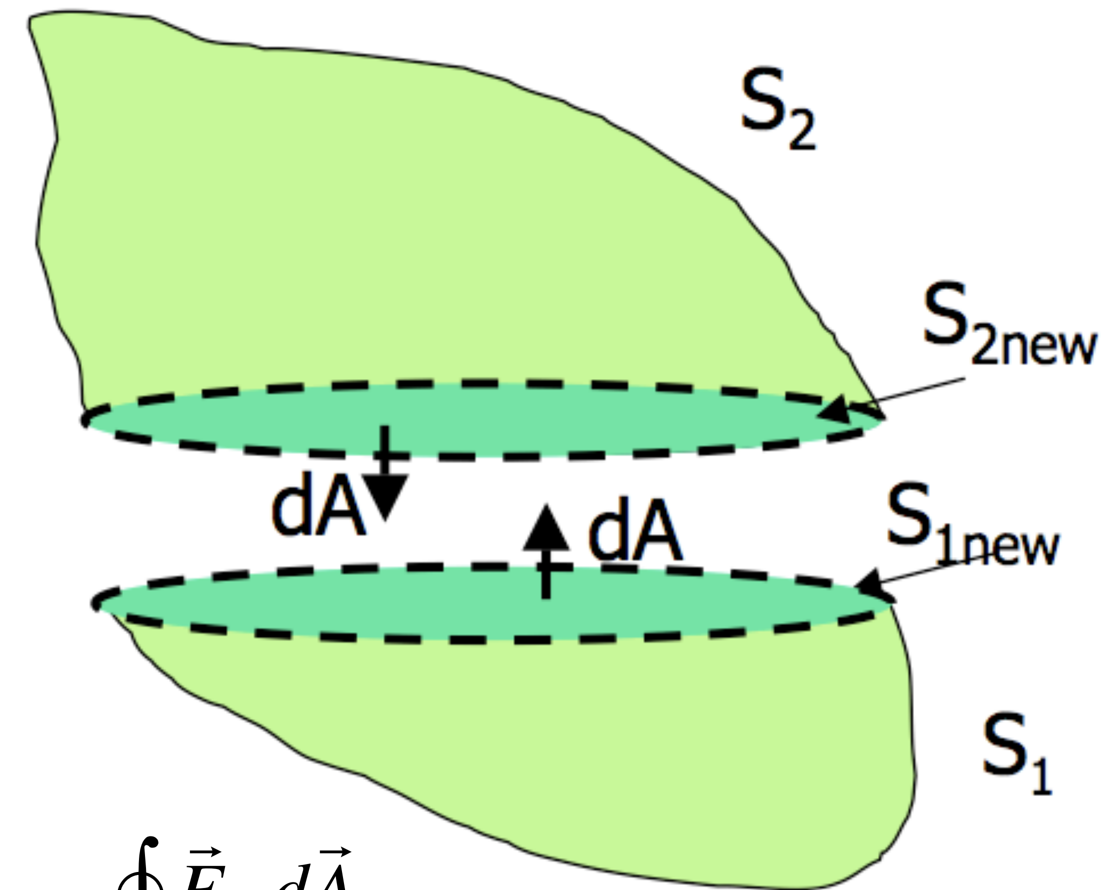
think about it!

Divergence in E&M (Part 1)

Consider flux of E through surface S:



Cut S into 2 surfaces: S_1 and S_2 with S_{new} the little surface in between



$$\phi = \oint_S \vec{E} \cdot d\vec{A}$$

$$= \oint_{S_1} \vec{E} \cdot d\vec{A} + \oint_{S_2} \vec{E} \cdot d\vec{A} + \underbrace{\oint_{S_{1new}} \vec{E} \cdot d\vec{A} + \oint_{S_{2new}} \vec{E} \cdot d\vec{A}}_{=0}$$

$$= \phi_1 + \phi_2$$

Divergence Theorem

Let's continue splitting into smaller volumes

$$\Phi = \sum_{i=1}^{\text{large N}} \Phi_i = \sum_{i=1}^{\text{large N}} \oint_{S_i} \vec{E} \cdot d\vec{A} = \sum_{i=1}^{\text{large N}} V_i \frac{\oint_{S_i} \vec{E} \cdot d\vec{A}}{V_i}$$

If we define the divergence of E as

$$\nabla \cdot \vec{E} \equiv \lim_{V \to 0} \frac{\oint_S \vec{E} \cdot d\vec{A}}{V} \quad \to \Phi = \sum_{i=1}^{\text{large N}} V_i \left(\nabla \cdot \vec{E} \right) \to \int_V \left(\nabla \cdot \vec{E} \right) dV$$

$$\to \oint_S \vec{E} \cdot d\vec{A} = \int_V \left(\nabla \cdot \vec{E} \right) dV$$ Divergence Theorem (Gauss's Theorem)

Gauss's law in differential form

Simple application of the divergence theorem:

$$\left.\begin{cases} \oint_S \vec{E}\cdot d\vec{A} = \int_V \left(\nabla\cdot\vec{E}\right) dV \\ \oint_S \vec{E}\cdot d\vec{A} = 4\pi Q = 4\pi \int_V \rho\, dV \end{cases}\right\} \rightarrow \int_V \left(\nabla\cdot\vec{E} - 4\pi\rho\right) dV = 0$$

This is valid for any surface V:

$$\nabla\cdot\vec{E} = 4\pi\rho$$

Gauss's law in differential form

Comments:

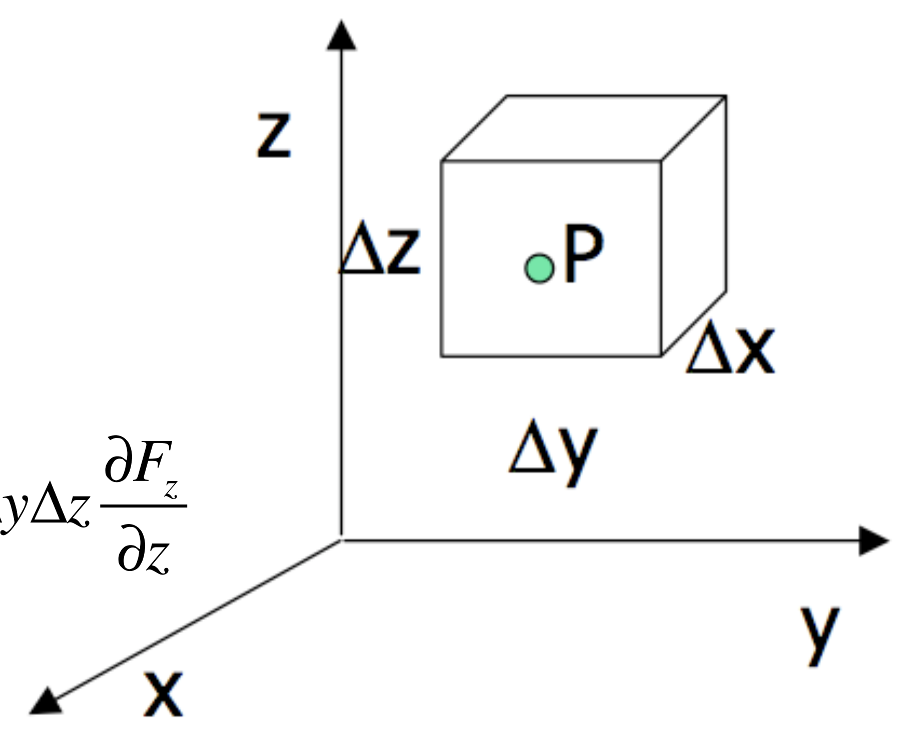
→ First of Maxwell's equations

Given E, allows us to easily extract charge distribution ρ

What's a divergence?

Consider infinitesimal cube centered at P=(x,y,z)

Flux of F through the cube in z direction:



$$\Delta\Phi_z = \int_{top+bottom} \vec{F}\cdot d\vec{A} \sim \Delta x\Delta y\left[F_z(x,y,z+\Delta z/2) - F_z(x,y,z-\Delta z/2)\right]$$

Since Δz → 0

$$\Delta\Phi_z = \Delta x\Delta y\Delta z \lim_{\Delta z\to 0}\left[\frac{F_z(x,y,z+\Delta z/2) - F_z(x,y,z-\Delta z/2)}{\Delta z}\right] = \Delta x\Delta y\Delta z\frac{\partial F_z}{\partial z}$$

Similarly for Φ_x and Φ_y

$$\Delta\Phi_x = \Delta x\Delta y\Delta z\frac{\partial F_x}{\partial x},\ \Delta\Phi_y = \Delta x\Delta y\Delta z\frac{\partial F_y}{\partial y}$$

Divergence in cartesian coordinates

We defined divergence as
$$\nabla \cdot \vec{E} \equiv \lim_{V \to 0} \frac{\oint_S \vec{E} \cdot d\vec{A}}{V}$$

But what does this really mean?

$$\nabla \cdot \vec{E} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0 \\ \Delta z \to 0}} \frac{\oint_S \vec{E} \cdot d\vec{A}}{V} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0 \\ \Delta z \to 0}} \frac{\Delta\Phi_x + \Delta\Phi_y + \Delta\Phi_z}{\Delta x \Delta y \Delta z} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0 \\ \Delta z \to 0}} \frac{\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}}{\Delta x \Delta y \Delta z} \Delta x \Delta y \Delta z$$

$$= \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}$$

This is the usable expression for the divergence: easy to calculate! ➙ $\nabla \cdot \vec{E} = 4\pi\rho$ is now useful!

Application of Gauss's law in differential form

Problem: given the electric field E(r), calculate the charge distribution that created it

$$\vec{E}(r) = \frac{4}{3}\pi K r \hat{r} \quad for \quad r < R \quad and \quad \vec{E}(r) = \frac{4\pi K}{3r^2} R^3 \hat{r} \quad for \quad r > R$$

Hint: what connects E and ρ? Gauss's law.

$$\oint_S \vec{E} \cdot d\vec{A} = 4\pi Q_{enclosed} \qquad \text{(integral form)}$$

$$\nabla \cdot \vec{E} = 4\pi\rho \qquad \text{(differential form)}$$

In cartesian coordinates(see next slide):

$$\nabla \cdot \vec{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \begin{cases} 4\pi K & when \quad r < R \\ 0 & when \quad r > R \end{cases}$$

➞ Sphere of radius R with uniform charge density K

Laplacian operator

What if we combine gradient and divergence?
Let's calculate the div grad f(Q: difference wrt grad div f ?)

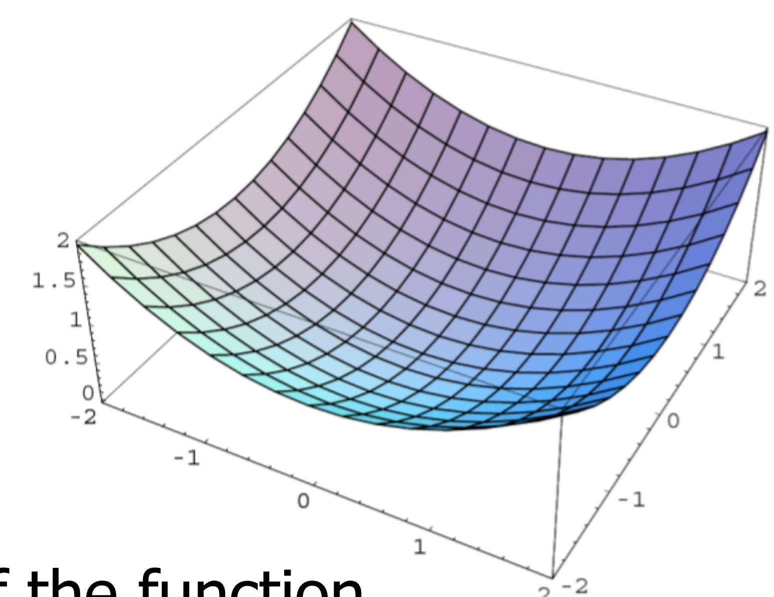
$$\nabla \bullet \nabla f = \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z} \right) \cdot \left(\frac{\partial f}{\partial x}\hat{x} + \frac{\partial f}{\partial y}\hat{y} + \frac{\partial f}{\partial z}\hat{z} \right)$$

$$= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) f \equiv \nabla^2 f$$

$$\nabla^2 f \equiv \nabla \bullet \nabla f \qquad \text{Laplacian Operator}$$

Interpretation of Laplacian

Given a 2d function $\phi(x,y) = a\frac{x^2 + y^2}{4}$ calculate the Laplacian



$$\nabla^2 \phi = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \phi = \frac{a}{4}(2+2) = a$$

As the second derivative, the Laplacian gives the curvature of the function

$r < R$

$$\vec{E} = \frac{4}{3}\pi K\left(x\hat{i} + y\hat{j} + z\hat{k}\right) \Rightarrow \nabla \bullet \vec{E} = \frac{4}{3}\pi K\left(\frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z}\right) = \frac{4}{3}\pi K(3) = 4\pi K$$

$r > R$

$$\vec{E} = \frac{4}{3}\pi K R^3\left(\frac{x\hat{i} + y\hat{j} + z\hat{k}}{r^3}\right) \Rightarrow \nabla \bullet \vec{E} = \frac{4}{3}\pi K R^3\left(\frac{\partial \frac{x}{r^3}}{\partial x} + \frac{\partial \frac{y}{r^3}}{\partial y} + \frac{\partial \frac{z}{r^3}}{\partial z}\right)$$

$$= \frac{4}{3}\pi K R^3\left[\frac{1}{r^3} + x\frac{\partial}{\partial x}\left(\frac{1}{r^3}\right) + \frac{1}{r^3} + y\frac{\partial}{\partial y}\left(\frac{1}{r^3}\right) + \frac{1}{r^3} + z\frac{\partial}{\partial z}\left(\frac{1}{r^3}\right)\right]$$

$$= \frac{4}{3}\pi K R^3\left[\frac{3}{r^3} + x\frac{\partial r}{\partial x}\frac{\partial}{\partial r}\left(\frac{1}{r^3}\right) + y\frac{\partial r}{\partial y}\frac{\partial}{\partial r}\left(\frac{1}{r^3}\right) + z\frac{\partial r}{\partial z}\frac{\partial}{\partial r}\left(\frac{1}{r^3}\right)\right]$$

$$= \frac{4}{3}\pi K R^3\left[\frac{3}{r^3} + x\frac{\partial\sqrt{x^2+y^2+z^2}}{\partial x}\frac{\partial}{\partial r}\left(\frac{1}{r^3}\right) + y\frac{\partial\sqrt{x^2+y^2+z^2}}{\partial y}\frac{\partial}{\partial r}\left(\frac{1}{r^3}\right) + z\frac{\partial\sqrt{x^2+y^2+z^2}}{\partial z}\frac{\partial}{\partial r}\left(\frac{1}{r^3}\right)\right]$$

$$= \frac{4}{3}\pi K R^3\left[\frac{3}{r^3} + x\frac{x}{r}\left(-\frac{3}{r^4}\right) + y\frac{y}{r}\left(-\frac{3}{r^4}\right) + z\frac{z}{r}\left(-\frac{3}{r^4}\right)\right]$$

$$= \frac{4}{3}\pi K R^3\left[\frac{3}{r^3} + \frac{x^2+y^2+z^2}{r}\left(-\frac{3}{r^4}\right)\right] = \frac{4}{3}\pi K R^3\left[\frac{3}{r^3} + \frac{r^2}{r}\left(-\frac{3}{r^4}\right)\right] = \frac{4}{3}\pi K R^3\left[\frac{3}{r^3} - \frac{3}{r^2}\right] = 0$$

There is a problem at origin; not relevant here. Return to later.

Poisson equation

Let's apply the concept of Laplacian to electrostatics.
Rewrite Gauss's law in terms of the potential

$$\left.\begin{cases}\nabla \cdot \vec{E} = 4\pi\rho \\ \nabla \cdot \vec{E} = \nabla \cdot (-\nabla\phi) = -\nabla^2\phi\end{cases}\right\} \rightarrow \nabla^2\phi = -4\pi\rho \qquad \text{Poisson Equation}$$

Useful to get the charge density ρ that originated the (known) potential ϕ

Laplace equation and Earnshaw's Theorem

What happens to Poisson's equation in vacuum?

$$\nabla^2\phi = -4\pi\rho \rightarrow \nabla^2\phi = 0 \qquad \text{Laplace Equation}$$

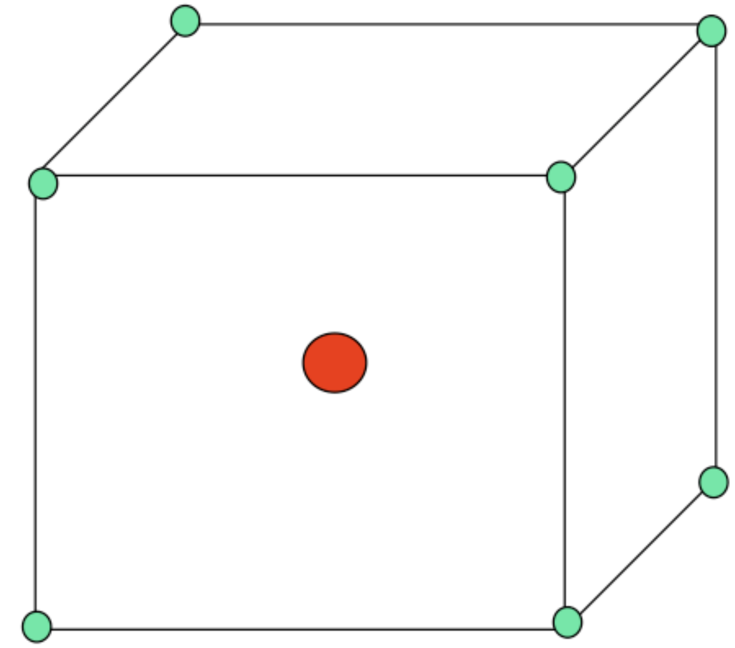
What does this teach us?

- In a region where ϕ satisfies Laplace's equation, then its curvature must be 0 everywhere in the region
- The potential has no local maxima or minima in that region

Important consequence for physics:

- Earnshaw'sTheorem:
  - It is impossible to hold a charge in stable equilibrium with electrostatic fields (no minima)

Application of Earnshaw's Theorem



8 charges on a cube and one free in the middle.
Is the equilibrium stable? No!

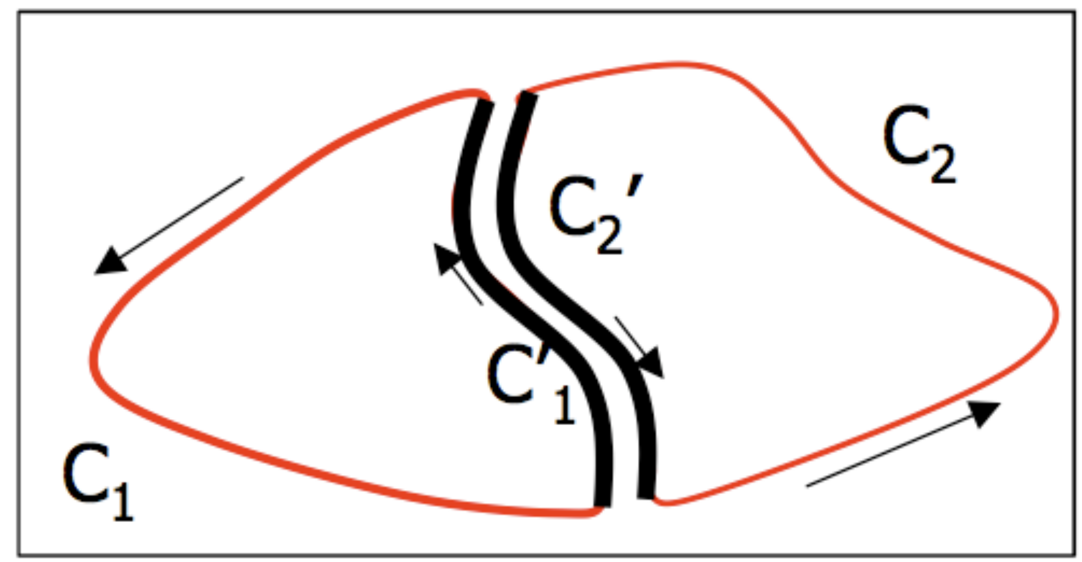
meaning? think about it

The Circulation

Consider the line integral of a vector function F over a closed path C:

$$\Gamma = \oint_C \vec{F} \cdot d\vec{s} \longrightarrow \text{Circulation}$$

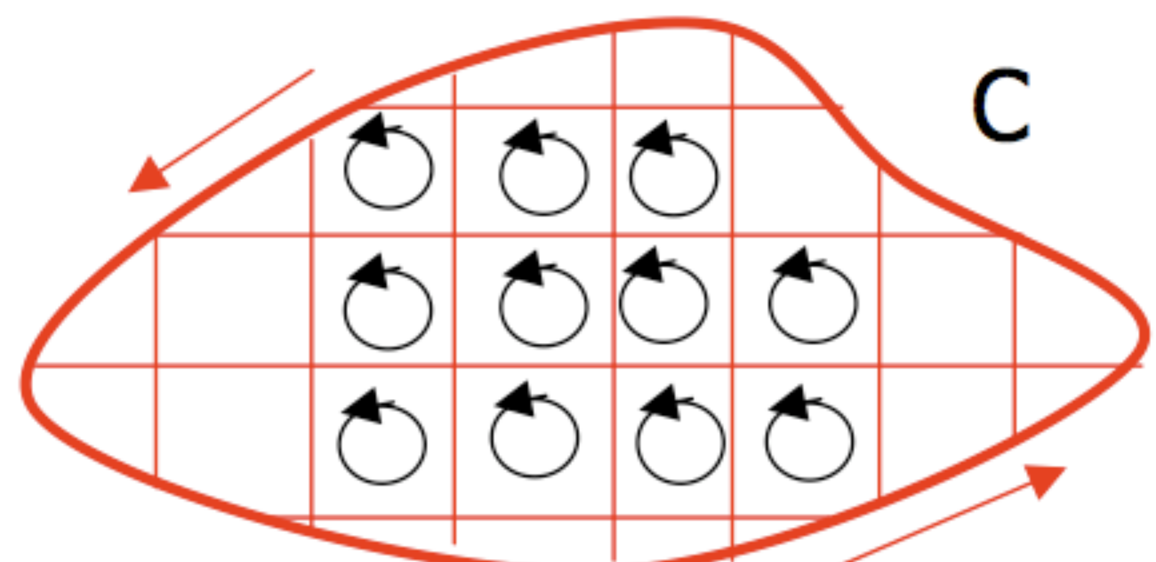
Let's now cut C into 2 smaller loops: C_1 and C_2 and write the circulation C in terms of the integral on C_1 and C_2



$$\Gamma = \oint_C \vec{F} \cdot d\vec{s} = \oint_{C_1} \vec{F} \cdot d\vec{s} + \oint_{C_2} \vec{F} \cdot d\vec{s} + \underbrace{\oint_{C'_1} \vec{F} \cdot d\vec{s} + \oint_{C'_2} \vec{F} \cdot d\vec{s}}_{=0} = \Gamma_1 + \Gamma_2$$

The curl of F

If we repeat the procedure N times:



$$\Gamma = \sum_{i=1}^{\text{large N}} \Gamma_i$$

Define the curl of F as circulation of F per unit area in the limit A → 0

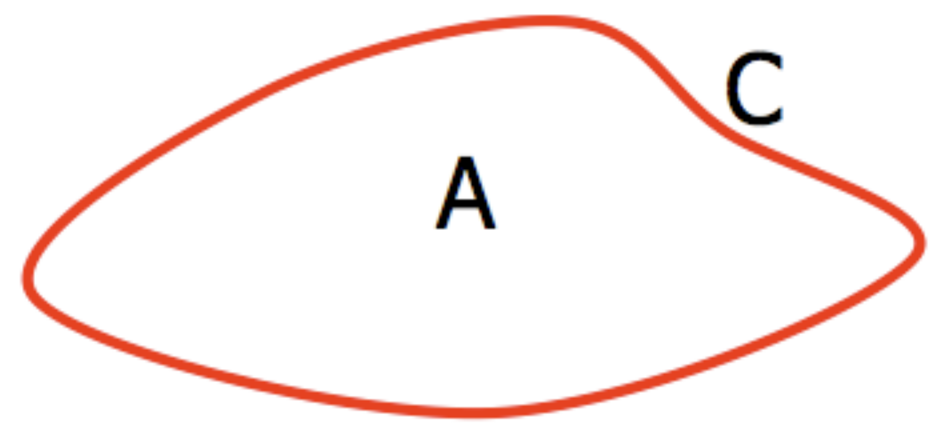
$$curl\left(\vec{F}\right)\cdot\hat{n} \equiv \lim_{A\to 0}\frac{\oint\limits_C \vec{F}\cdot d\vec{s}}{A}$$

where A is the area inside C

The curl is a vector normal to the surface A with direction given by the "right hand rule"

Stokes Theorem

$$\Gamma = \sum_{i=1}^{\text{large N}} \Gamma_i = \sum_{i=1}^{\text{large N}} \oint\limits_{C_i} \vec{F}\cdot d\vec{s} = \sum_{i=1}^{\text{large N}} A_i \frac{\oint\limits_{C_i} \vec{F}\cdot d\vec{s}}{A_i}$$



In the limit A → 0:

$$\frac{\oint\limits_C \vec{F}\cdot d\vec{s}}{A_i} \to curl\left(\vec{F}\right)\cdot\hat{n}$$

and

$$\sum_{i=1}^{\text{large N}} A_i \to \int\limits_A dA$$

$$\Gamma = \sum_{i=1}^{\text{large N}} A_i curl\left(\vec{F}\right)\cdot\hat{n} = \sum_{i=1}^{\text{large N}} curl\left(\vec{F}\right)\cdot A_i\hat{n} = \sum_{i=1}^{\text{large N}} curl\left(\vec{F}\right)\cdot\vec{A}_i \to \int\limits_A curl\left(\vec{F}\right)\cdot d\vec{A} \qquad \Gamma = \oint\limits_C \vec{F}\cdot d\vec{s}$$

$$\to \oint\limits_C \vec{F}\cdot d\vec{s} = \int\limits_A curl\left(\vec{F}\right)\cdot d\vec{A}$$ Stokes Theorem

Stokes relates the line integral of a vector function F over a closed line C and the surface integral of the curl of the function over the area enclosed by C

Application of Stoke's Theorem

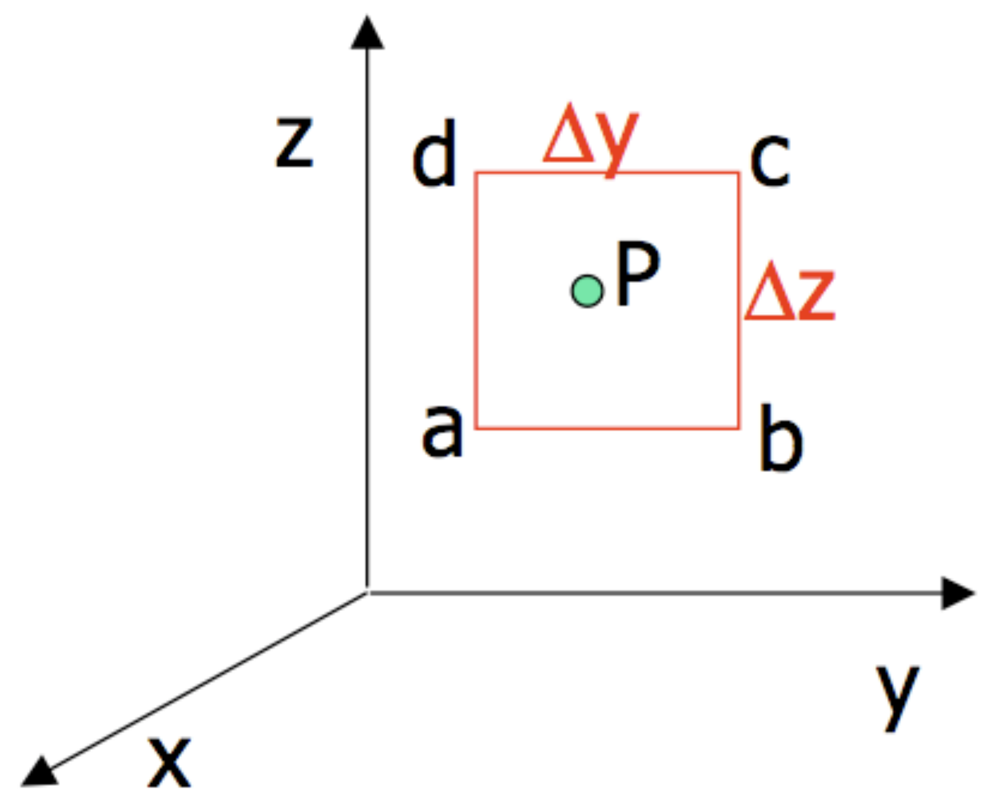
Stoke's theorem:

$$\oint_C \vec{F} \cdot d\vec{s} = \int_A curl\left(\vec{F}\right) \cdot d\vec{A}$$

The Electrostatics Force is conservative:

$$\oint_C \vec{E} \cdot d\vec{s} = 0 \Rightarrow \int_A curl\left(\vec{E}\right) \cdot d\vec{A} = 0 \text{ for any surface A } \Rightarrow curl\left(\vec{E}\right) = 0$$

The curl of an electrostatic field is zero.



Curl in cartesian coordinates

Consider infinitesimal rectangle in yz plane centered at P=(x,y,z) in a vector field F

Calculate circulation of F around the square:

$$\int_a^b \vec{F} \cdot d\vec{s} = F_y(x,y,z-\Delta z/2)\Delta y \sim \left[F_y(x,y,z) - \frac{\Delta z}{2}\frac{\partial F_y}{\partial z}\right]\Delta y$$

$$\int_b^c \vec{F} \cdot d\vec{s} = F_z(x,y+\Delta y/2,z)\Delta z \sim \left[F_z(x,y,z) + \frac{\Delta y}{2}\frac{\partial F_z}{\partial y}\right]\Delta z$$

$$\int_c^d \vec{F} \cdot d\vec{s} = F_y(x,y,z+\Delta z/2)(-\Delta y) \sim -\left[F_y(x,y,z) + \frac{\Delta z}{2}\frac{\partial F_y}{\partial z}\right]\Delta y$$

$$\int_d^a \vec{F} \cdot d\vec{s} = F_z(x,y-\Delta y/2,z)(-\Delta z) \sim -\left[F_z(x,y,z) - \frac{\Delta y}{2}\frac{\partial F_z}{\partial y}\right]\Delta z$$

Adding the 4 components:

$$\oint_{C(square)} \vec{F} \cdot d\vec{s} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\Delta y \Delta z$$

Combining this result with definition of curl:

$$curl\left(\vec{F}\right)\cdot\hat{n}\equiv\lim_{A\to 0}\frac{\oint_C \vec{F}\cdot d\vec{s}}{A}$$

$$\oint_{C(square)} \vec{F}\cdot d\vec{s} = \left(\frac{\partial F_z}{\partial y}-\frac{\partial F_y}{\partial z}\right)\Delta y\Delta z$$

$$\Rightarrow\left(curl\,\vec{F}\right)_x = \lim_{\substack{\Delta y\to 0\\ \Delta z\to 0}}\frac{\oint_{C(square)} \vec{F}\cdot d\vec{s}}{\Delta y\Delta z} = \left(\frac{\partial F_z}{\partial y}-\frac{\partial F_y}{\partial z}\right)$$

Similar results orienting the rectangles in || (xz) and (xy) planes ➙

$$\Rightarrow curl\,\vec{F} = \left(\frac{\partial F_z}{\partial y}-\frac{\partial F_y}{\partial z}\right)\hat{x}+\left(\frac{\partial F_x}{\partial z}-\frac{\partial F_z}{\partial x}\right)\hat{y}+\left(\frac{\partial F_y}{\partial x}-\frac{\partial F_x}{\partial y}\right)\hat{z} = \begin{vmatrix}\hat{x} & \hat{y} & \hat{z}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ F_x & F_y & F_z\end{vmatrix} = \nabla\times\vec{F}$$

This is the usable expression for the curl: easy to calculate!

Summary of vector calculus in electrostatics

Gradient: $\nabla\phi = \left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\phi$

In E&M: $\vec{E} = -\nabla\phi$

Divergence: $\nabla \cdot \vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$

Gauss's theorem: $\int_S \vec{E} \cdot d\vec{A} = \int_V (\nabla \cdot \vec{E}) dV$

In E&M: Gauss' law in differential form $\nabla \cdot \vec{E} = 4\pi\rho$

Curl: $curl\, \vec{F} = \nabla \times \vec{F}$

Stoke's theorem: $\oint_C \vec{F} \cdot d\vec{s} = \int_A (curl\, \vec{F}) \cdot d\vec{A}$

In Electrostatics: $\nabla \times \vec{E} = 0$

Laplacian: $\nabla^2 \phi \equiv \nabla \cdot \nabla \phi$

In E&M:

Poisson Equation: $\nabla^2 \phi = -4\pi\rho$

Laplace Equation: $\nabla^2 \phi = 0$ ➙ Earnshaw's theorem

This may look like a lot of math: it is!
Time and practice will help you to learn how to use it in E&M

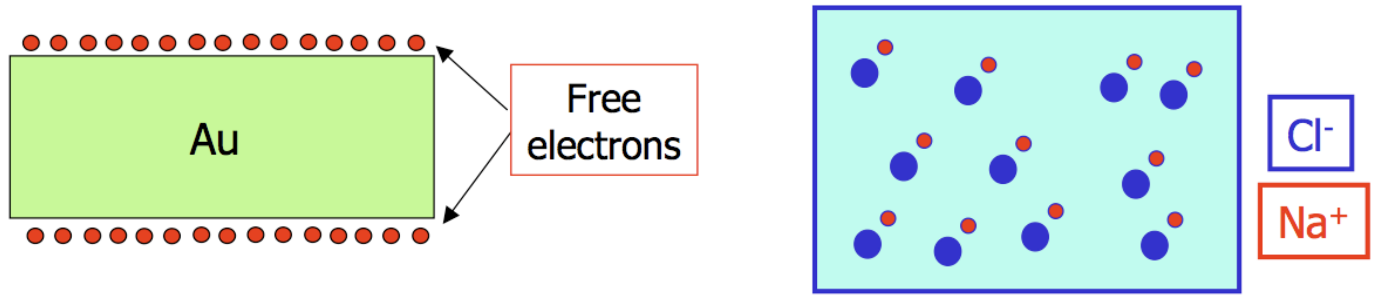
Conductors and Insulators

Let us get started…

Conductor: a material with free electrons

Excellent conductors: metals such as Au, Ag, Cu, Al,…

OK conductors: ionic solutions such as NaCl in H2O



Insulator: a material without free electrons

Organic materials: rubber, plastic,…

Inorganic materials: quartz, glass,…

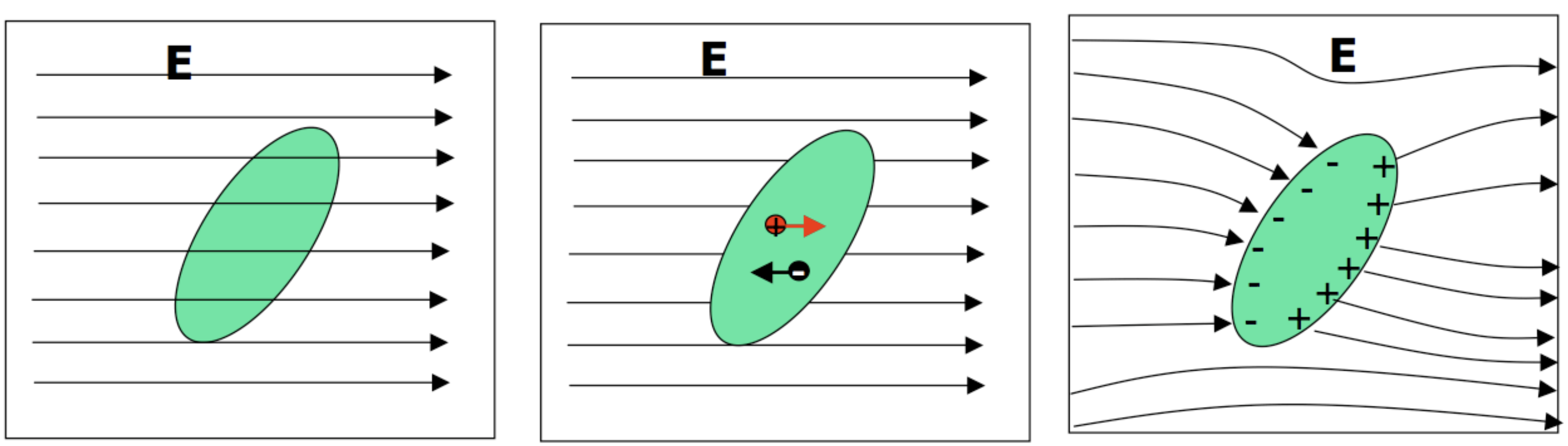
Properties of Conductors

A conductor is assumed to have an infinite supply of electric charges

Pretty good assumption…

Inside a conductor, E=0. Why? If E is not 0 charges will move from where the potential is higher to where the potential is lower; migration will stop only when E=0.

How long does it take? 10^{-17}-10^{-16}s (typical resistivity of metals)



see next slide

Electric potential inside a conductor is constant

Given 2 points inside the conductor P1 and P2 the Δϕ would be:

$$\Delta\phi = \int_{P_1}^{P_2} \vec{E} \cdot d\vec{s} = 0$$ since E=0 inside the conductor

Net charge can only reside on the surface

If net charge inside the conductor ➙ Electric Field ≠ 0 (Gauss's law)

External field lines are perpendicular to surface

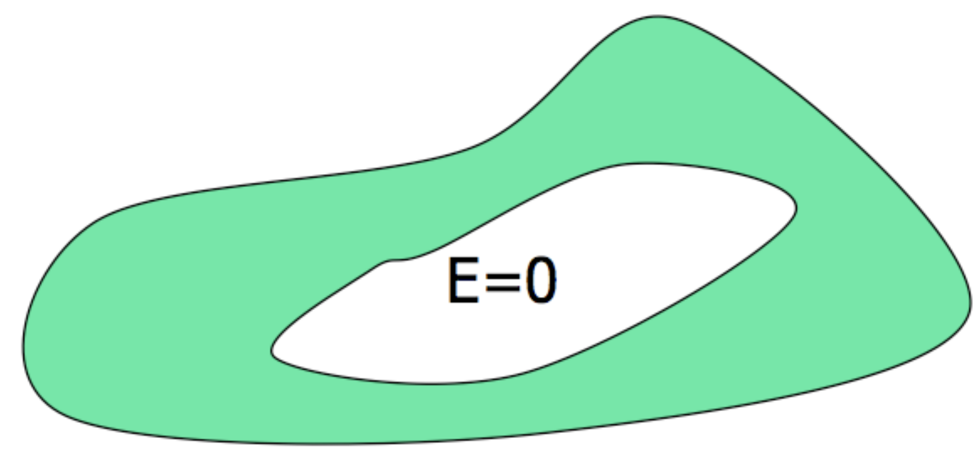
E || component would cause charge flow on the surface until Δϕ=0

Conductor's surface is an equipotential

Because it's perpendicular to field lines

Corollary 1:

In a hollow region inside conductor, ϕ=constant and E=0 if there aren't any charges in the cavity



Why?

Surface of conductor is equipotential

If no charge inside the cavity ➙ Laplace equation holds ➙ ϕ_{cavity} cannot have max or minima

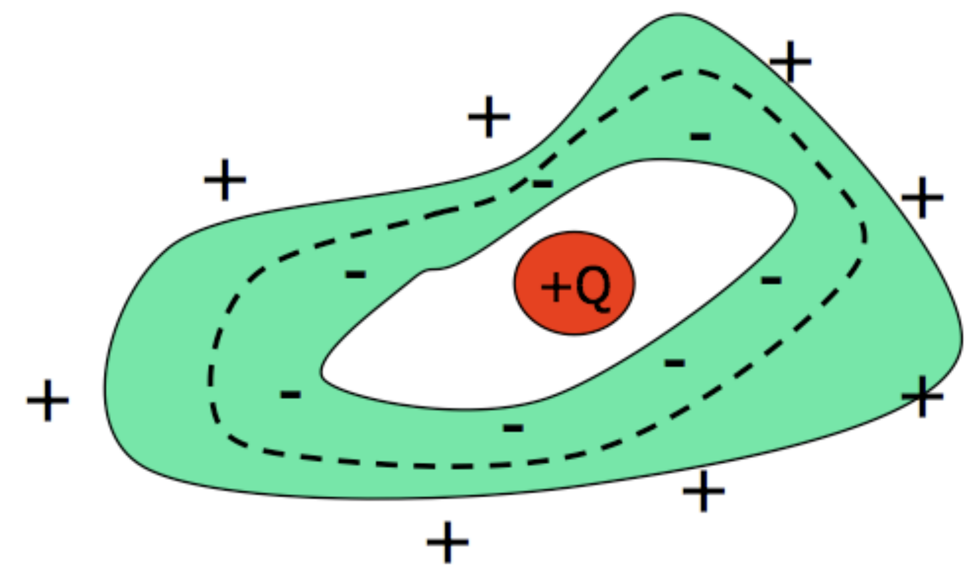
➙ ϕ must be constant ➙ E=0

Consequence:

Shielding of external electric fields: Faraday's cage

Corollary 2

A charge +Q in the cavity will induce a charge +Q on the outside of the conductor



Why?
Apply Gauss's law to surface ---inside the conductor+

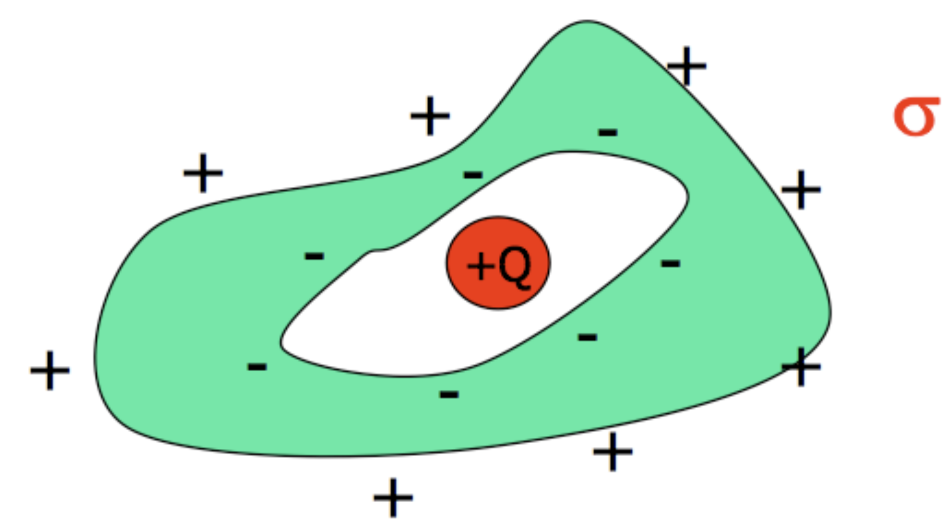
$\oint \vec{E} \cdot d\vec{A} = 0$ because E=0 inside a conductor

$\oint \vec{E} \cdot d\vec{A} = 4\pi(Q + Q_{inside})$ Gauss's law

$\Rightarrow Q_{inside} = -Q \Rightarrow Q_{outside} = -Q_{inside} = Q$ (conductor is overall neutral)

Corollary 3

The induced charge density on the surface of a conductor caused by a charge Q inside it is $\sigma_{induced} = E_{surface} / 4\pi$



Why?

For surface charge layer, Gauss tells us that $\Delta E = 4\pi\sigma$

Since $E_{inside} = 0$ $\rightarrow E_{surface} = 4\pi\sigma_{induced}$

Before continuing let us make a slight digression for fun to show you some of the techniques that one learns in a Mathematical methods class(advertisement).

Remember Laplace's equation

$$\text{Laplace Equation:} \qquad \nabla^2\phi = 0$$

Suppose that we set up a world in which the potential takes certain values on the "boundaries" of some region where the charge density is zero (no charges), then Laplaces's equation is satisfied in that region (within the boundaries).

In this case we can determine the potential everywhere in the region - inside using Laplace's equation and making sure that it always has the correct values on the boundaries. We can then plot the equipotential lines (surfaces).

Knowing that the electric field direction is always perpendicular to the equipotentials, we are then able to draw the electric flux lines.

First we will derive the equations for determining the potential from Laplace's equation.

This procedure turns out to simply be some approximations plus some algebra as shown on next slide.

Relaxation Methods

Cartesian Coordinates

$$\nabla^2 u(x,y,z) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

We approximate the derivatives by

$$\frac{\partial u}{\partial x}(x,y,z) = \frac{u(x+\Delta x/2,y,z) - u(x-\Delta x/2,y,z)}{\Delta x}$$

$$\frac{\partial^2 u}{\partial x^2}(x,y,z) = \frac{\frac{\partial u}{\partial x}(x+\Delta x/2,y,z) - \frac{\partial u}{\partial x}(x-\Delta x/2,y,z)}{\Delta x}$$

$$= \frac{u(x+\Delta x,y,z) - u(x,y,z)}{(\Delta x)^2} - \frac{u(x,y,z) - u(x-\Delta x,y,z)}{(\Delta x)^2}$$

$$= \frac{u(x+\Delta x,y,z) + u(x-\Delta x,y,z) - 2u(x,y,z)}{(\Delta x)^2}$$

We then have (choosing $\Delta x = \Delta y = \Delta z = \Delta$) the finite difference equation

$$\nabla^2 u(x,y,z) = 0$$

$$= \frac{u(x+\Delta,y,z) + u(x-\Delta,y,z) + u(x,y+\Delta,z) + u(x,y-\Delta,z) + u(x,y,z+\Delta) + u(x,y,z-\Delta) - 6u(x,y,z)}{\Delta^2}$$

or in 3 dimensions

$$u(x,y,z)=\frac{1}{6}\left[\begin{array}{c}u(x+\Delta,y,z)+u(x-\Delta,y,z)+u(x,y+\Delta,z)+u(x,y-\Delta,z)\\ +u(x,y,z+\Delta)+u(x,y,z-\Delta)\end{array}\right]$$

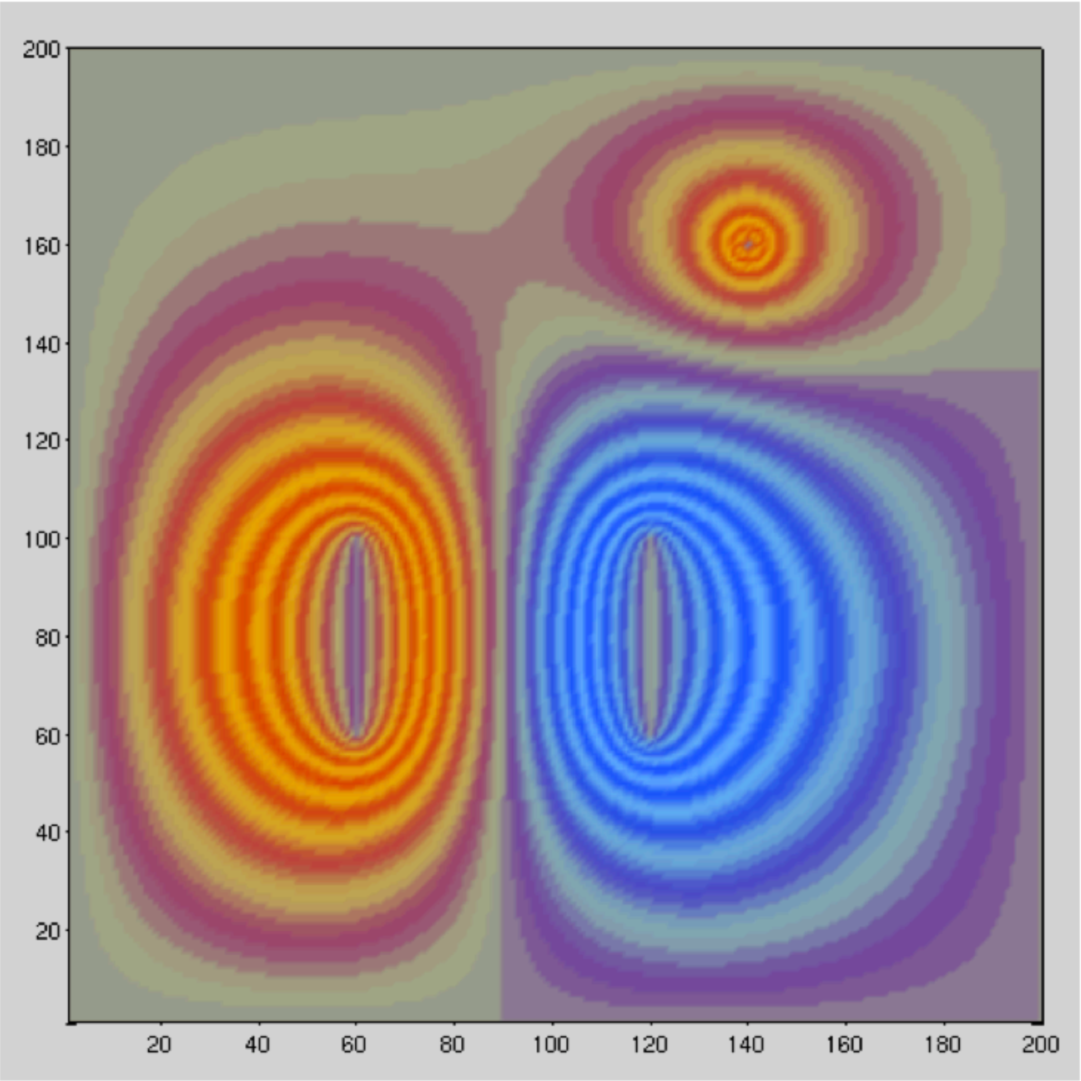
and in 2 dimensions

$$u(x,y)=\frac{1}{4}\left[u(x+\Delta,y)+u(x-\Delta,y)+u(x,y+\Delta)+u(x,y-\Delta)\right]$$

Process: Set boundary values; Set interior to zero; Reset potential values using above equation; Reset boundary values; Repeat until convergence.

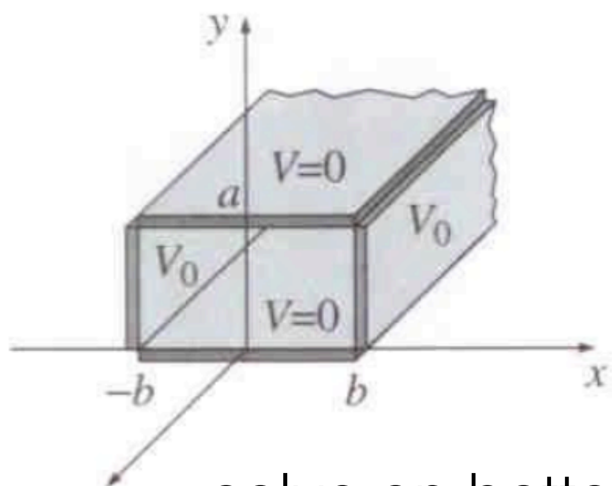
```
n = 200;
pin=zeros(n); p=pin; b=pin;
pin(16*n/20,14*n/20)=15 ;
pin(6*n/20:10*n/20,6*n/20)=15*ones(size(pin(6*n/20:10*n/20,6*n/20))) ;
pin(6*n/20:10*n/20,12*n/20)=-15*ones(size(pin(6*n/20:10*n/20,12*n/20)));
b(:,[1,n])=1*ones(size(b(:,[1,n])));
b([1,n],:)=1*ones(size(b([1,n],:)));
b=b+(pin ~= 0);
p=pin;
for i=1:5000
  p=0.25*(p(:,[n,[1:(n-1)]])+p(:,[[2:n],1]) ...
          +p([n,[1:(n-1)]],:)+p([[2:n],1],:));
  p=p.*(1-b)+pin;
end
figure('Position',[300,300,600,600]);
pcolor(p);
axis square;
colormap(waves);
shading interp;
```

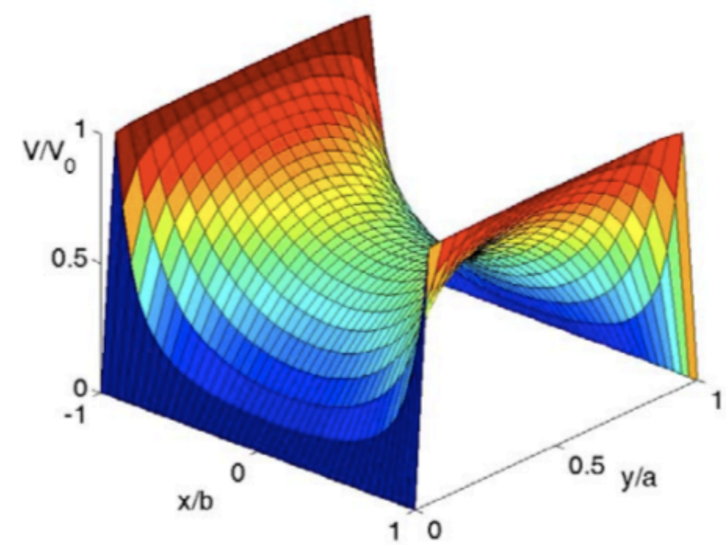
three03a.m , three03.m



Example of a solution to $\nabla^2\phi = 0$

Consider a square tube with two sides charged:

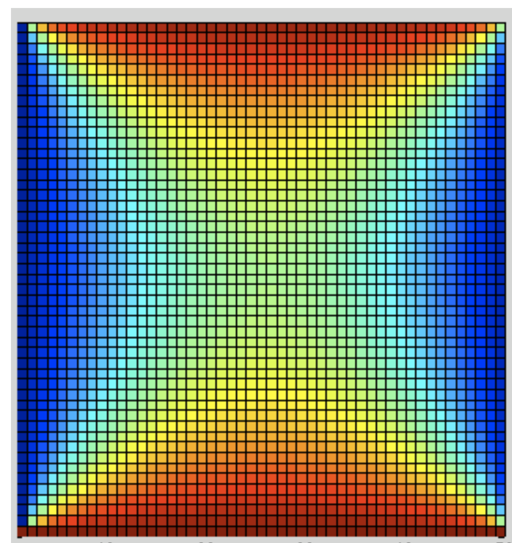


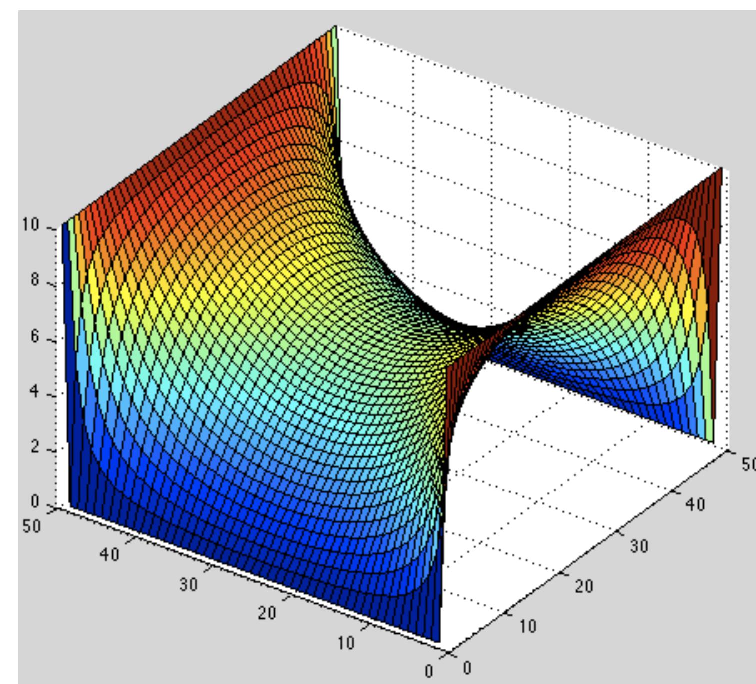


solve on bottom plane

three01.m

```
% relaxation method
n=50;
pin=zeros(n); p=pin; b=pin;
pin(1,1:n)=10;
pin(n,1:n)=10;
b(1,1:n)=ones(1,n);
b(n,1:n)=ones(1,n);
b(1:n,1)=ones(n,1);
b(1:n,n)=ones(n,1);
p=pin;
figure('Position',[50,50,400,400])
for i=1:5000
  p=0.25*(p(:,[n,[1:n-1]])+p(:,[[2:n],1])+p([n,[1:n-1]],:)+p([[2:n],1],:));
  p=p.*(1-b)+pin;
  if ((mod(i,10) == 0))
    i
    pcolor(p)
    pause(0.1)
  end
end
figure('Position',[150,150,400,400])
surf(p)
```

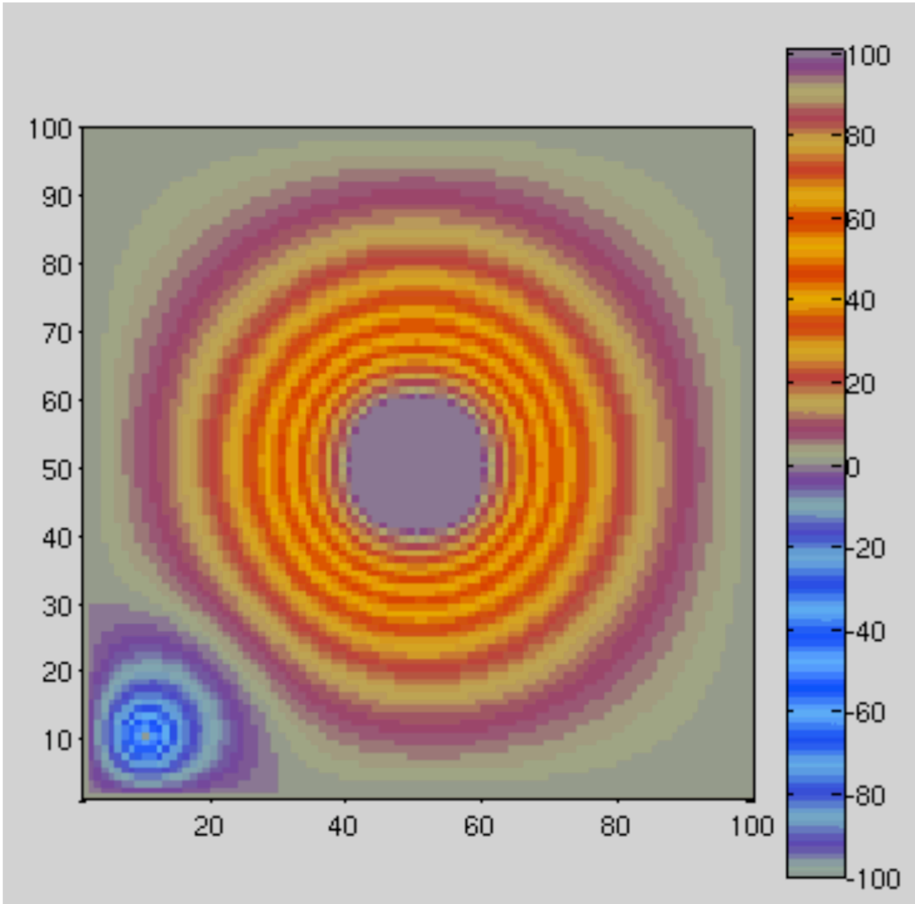




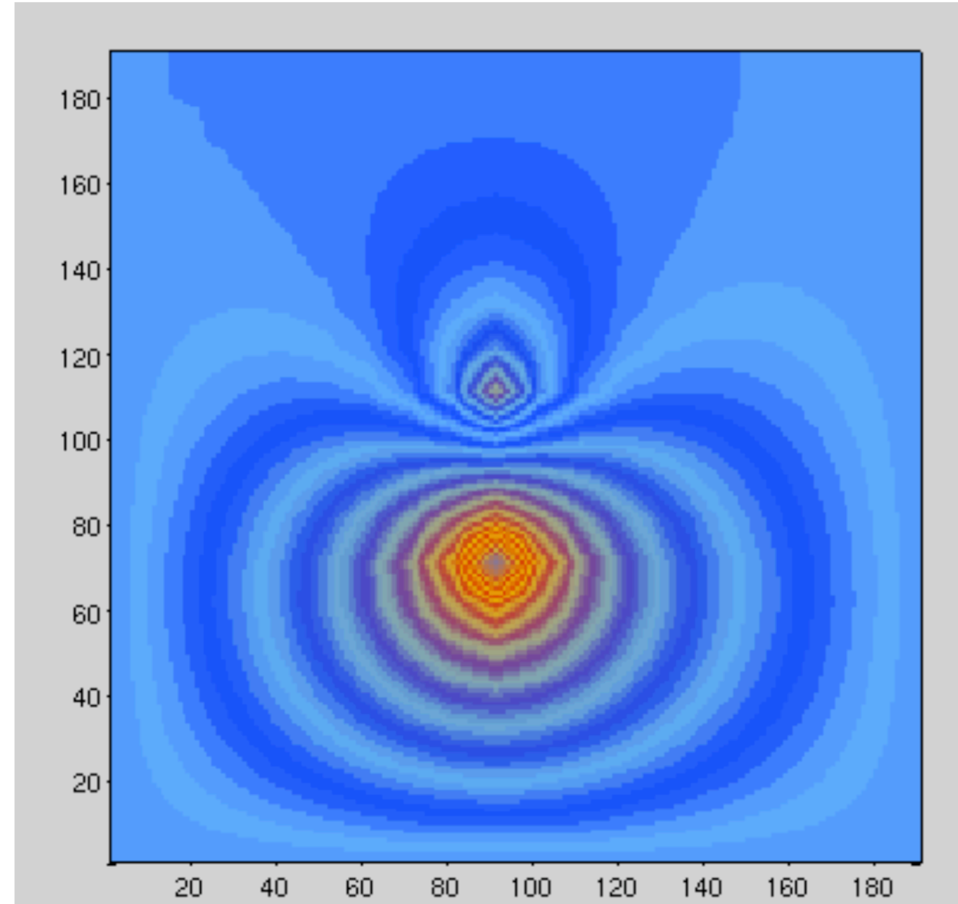
The solution is like a rubber sheath pulled to match the boundary conditions (note: No local max. or min.)

More examples:

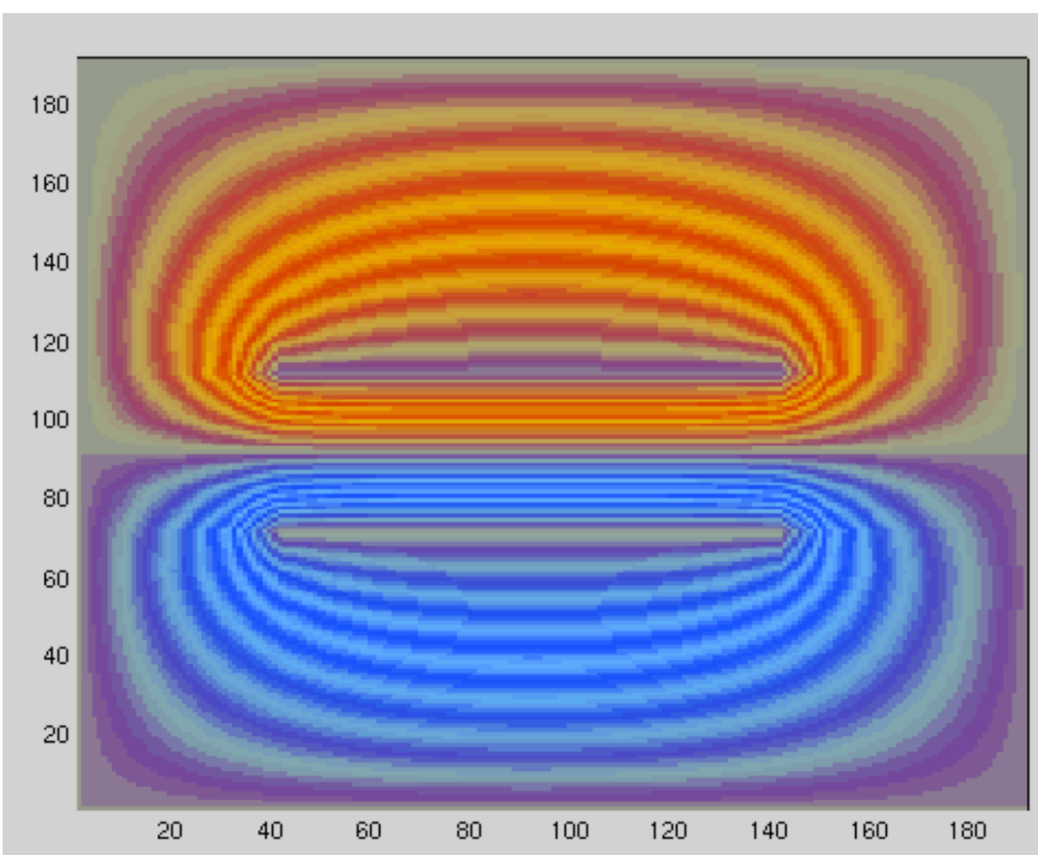
Point + sphere charges



2 Point charges

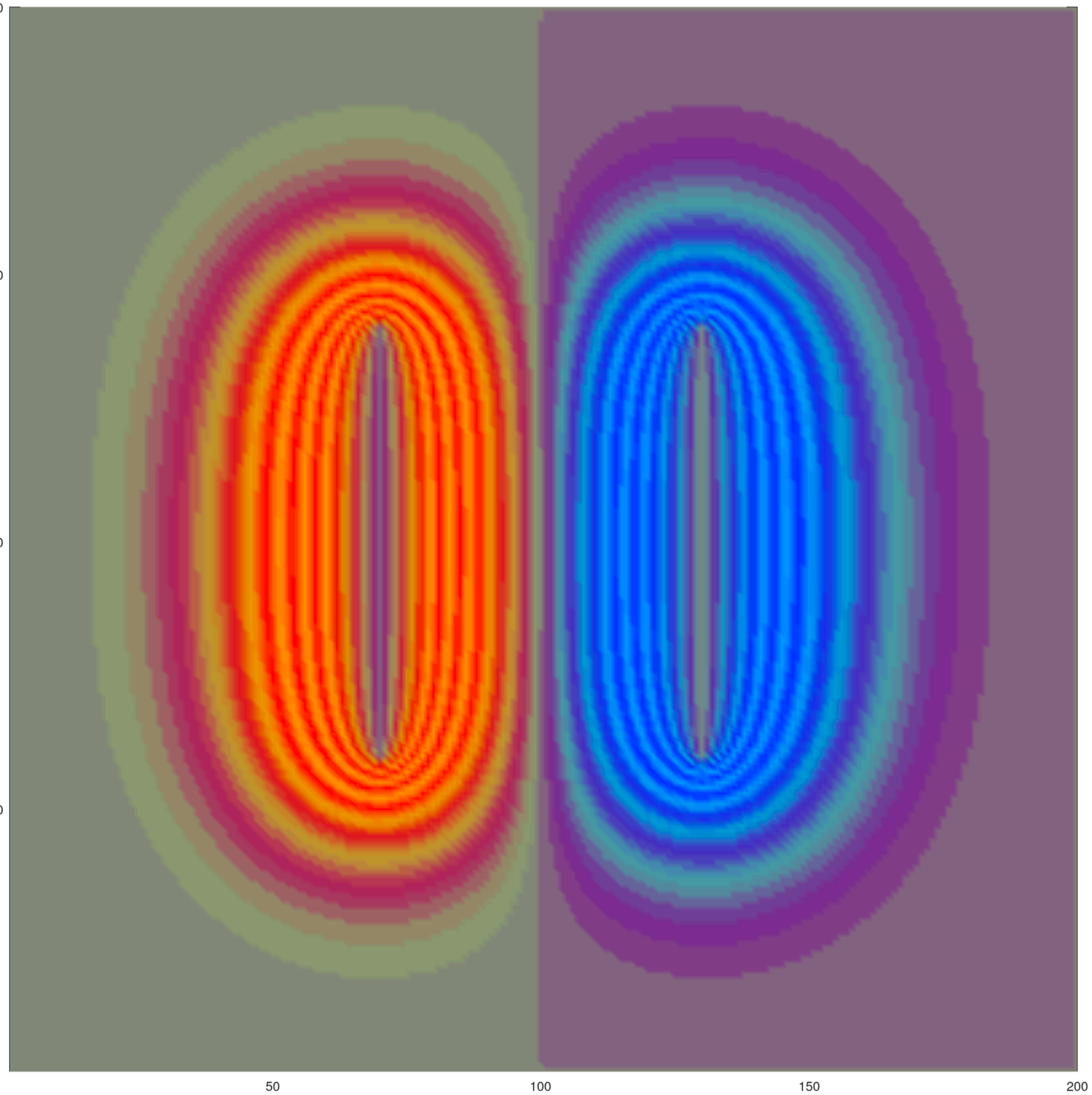


Parallel plates

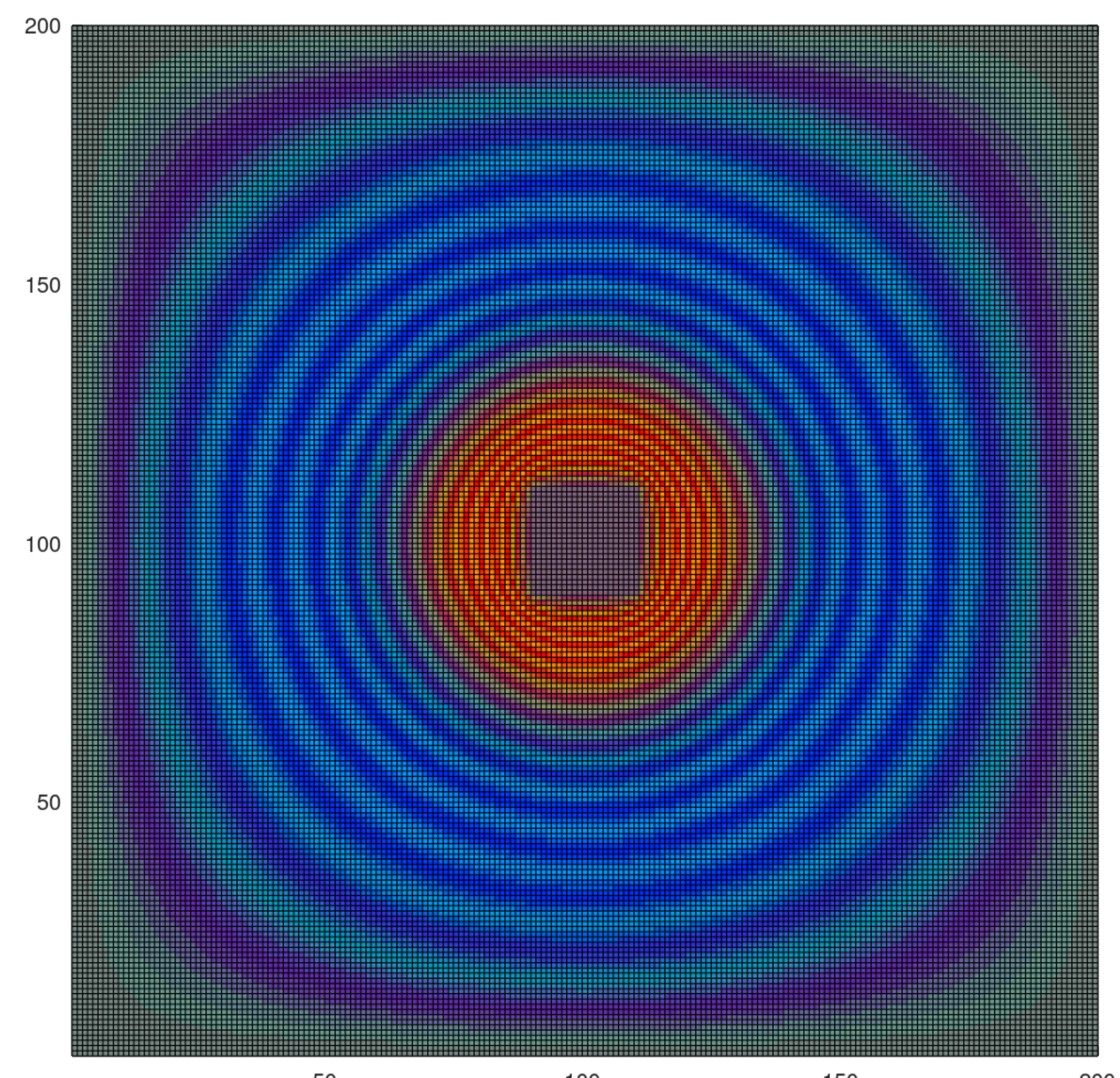


In class: relax1mov.m
relax1mova.m

In a Mathematical Methods course(next) one learns how to solve Laplace's equations and many other exactly and then with numerical methods.



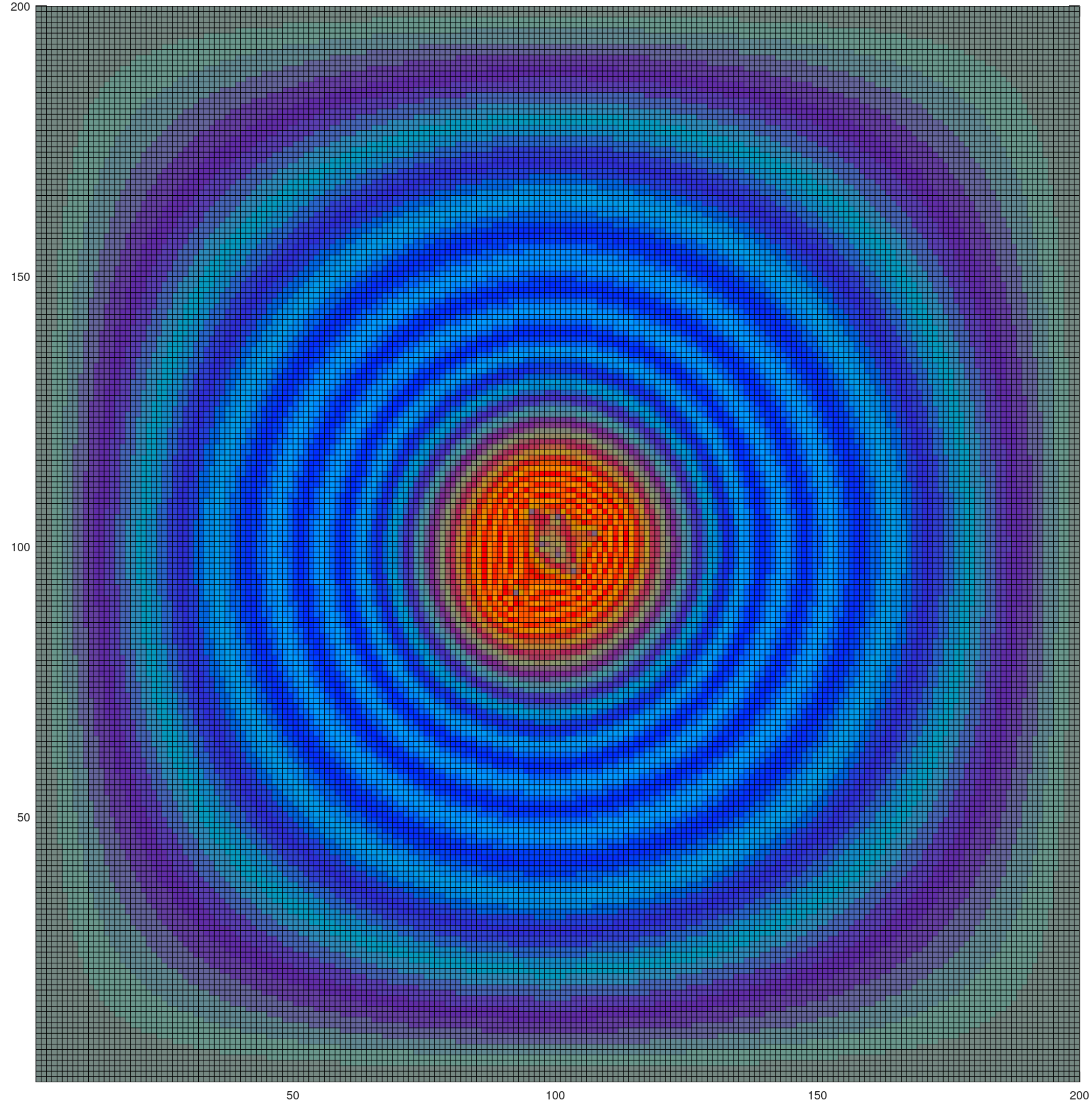
Even better of iterate longer!



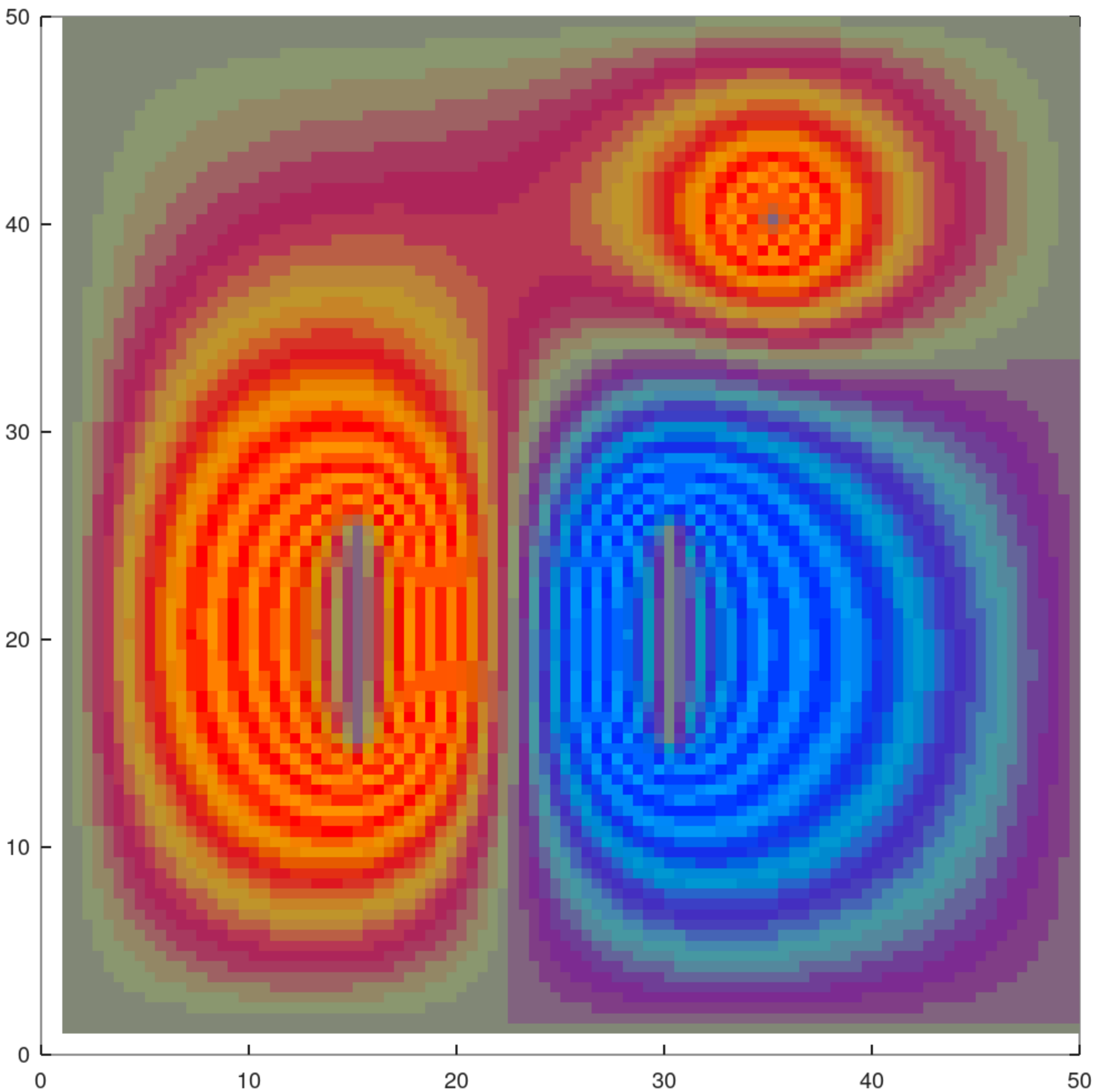
Far enough away square acts like point charge!

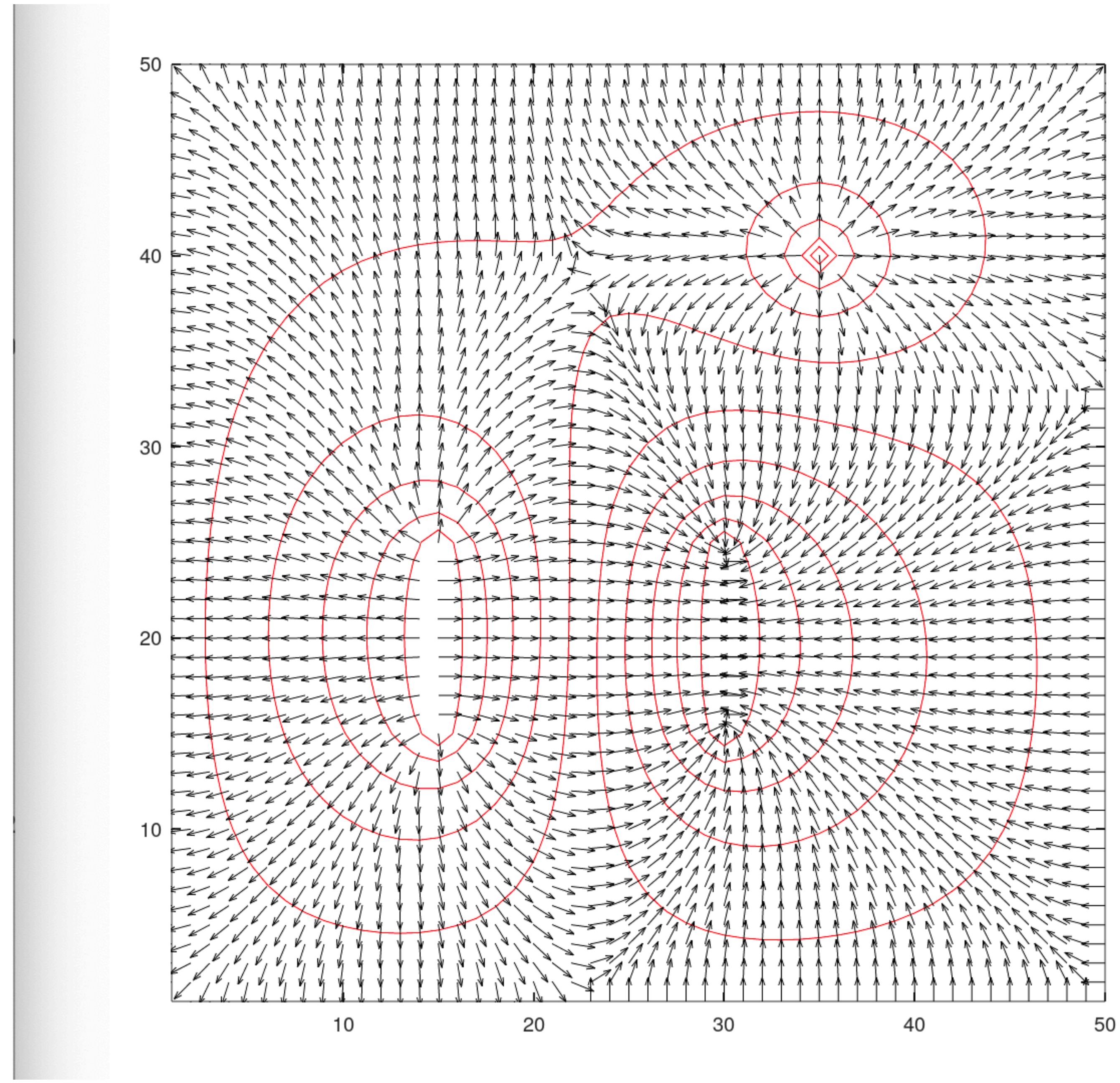
7 point charges

Far enough away square acts like point charge!

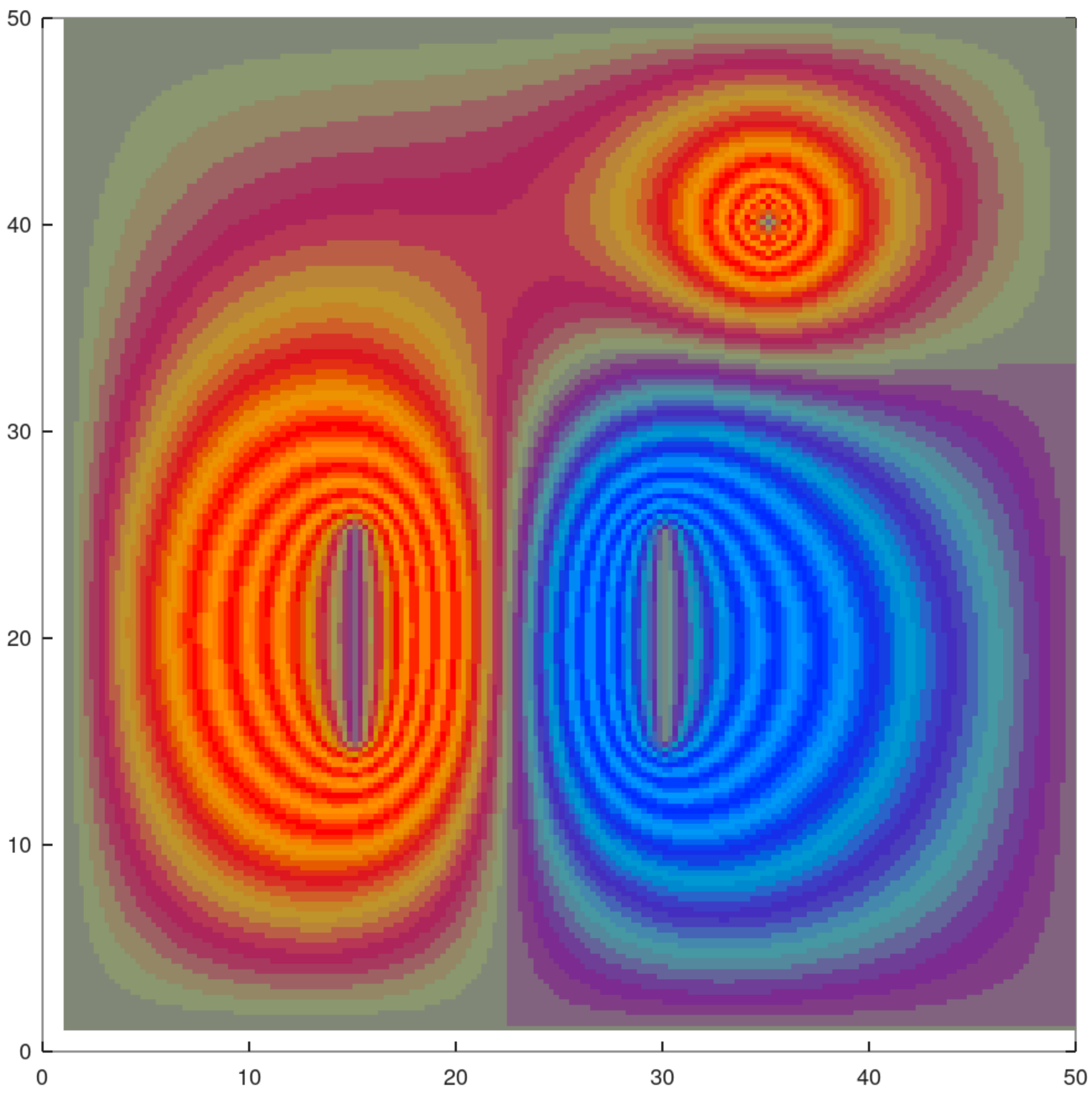


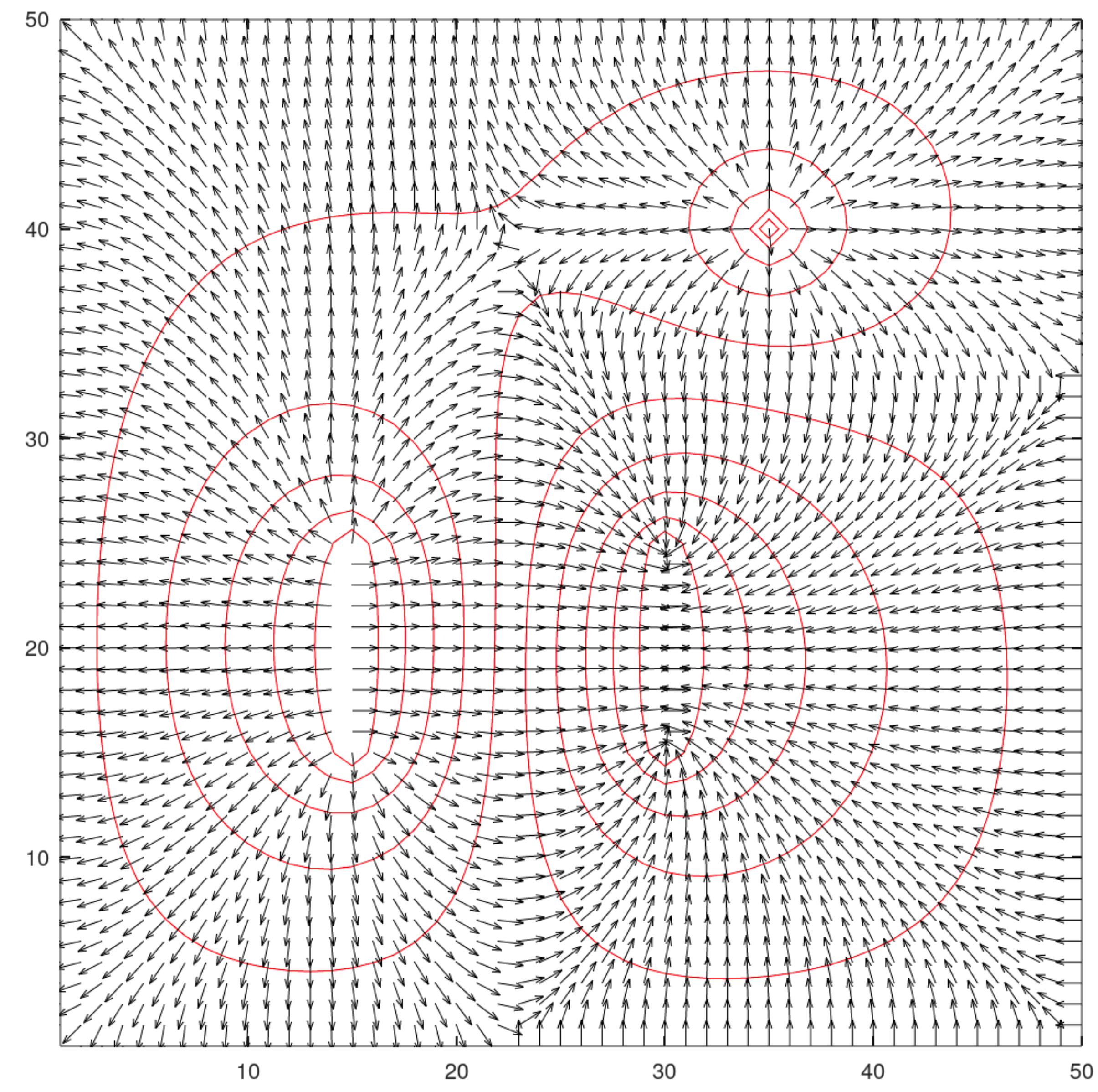
relaxbigcont.m

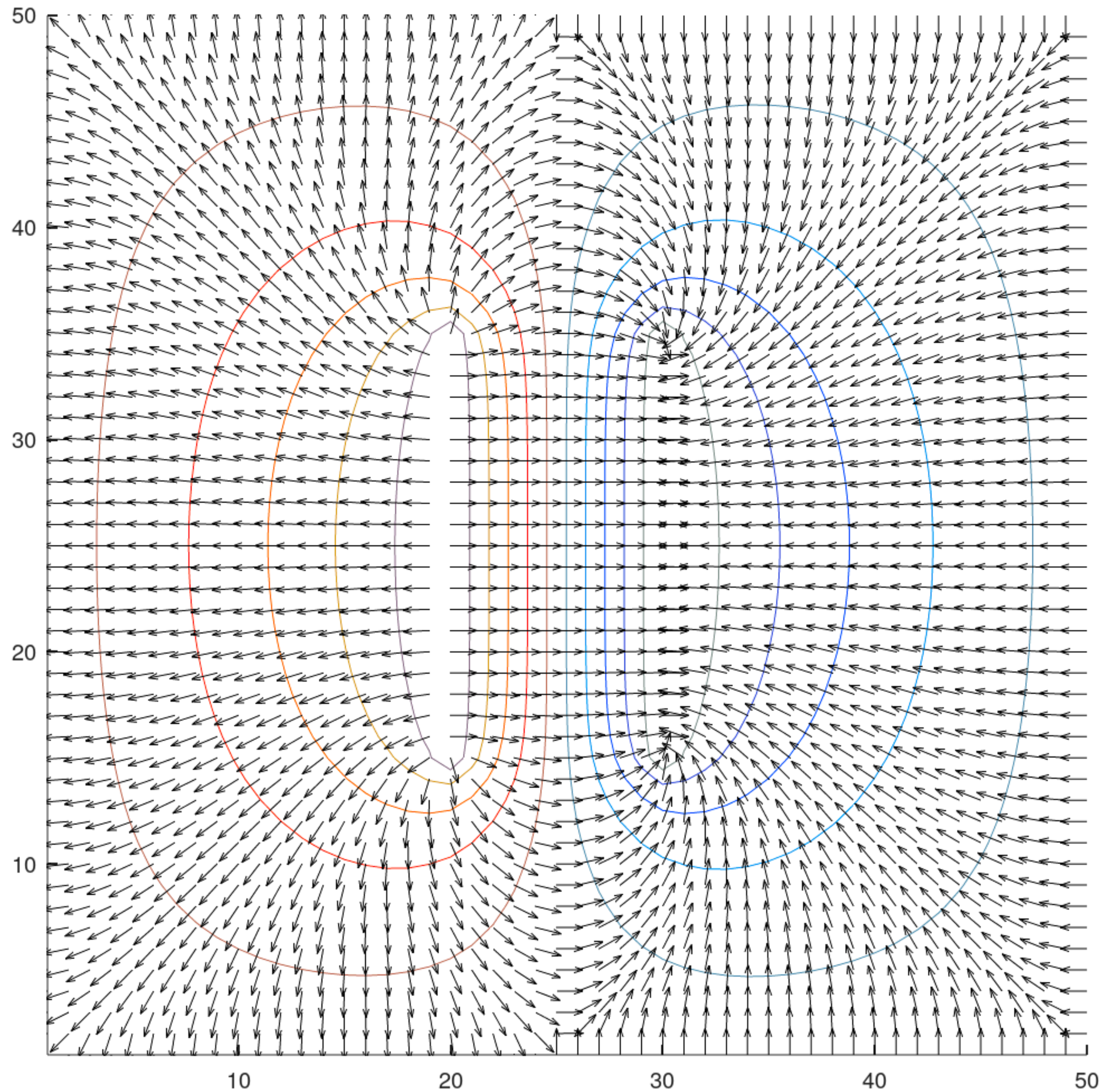




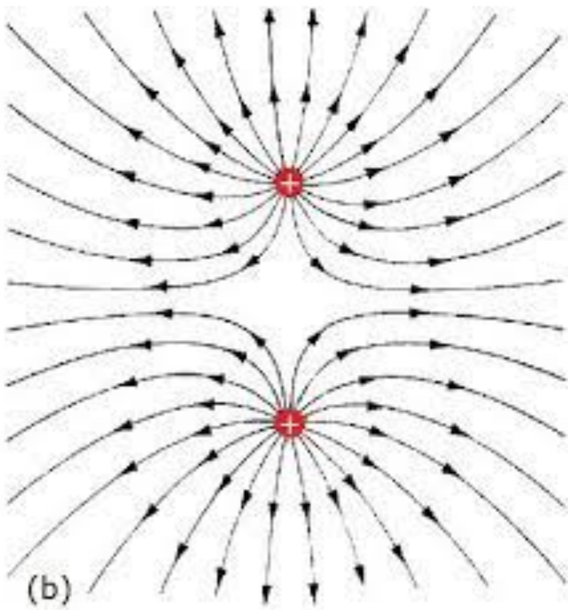
relaxbigcont.m

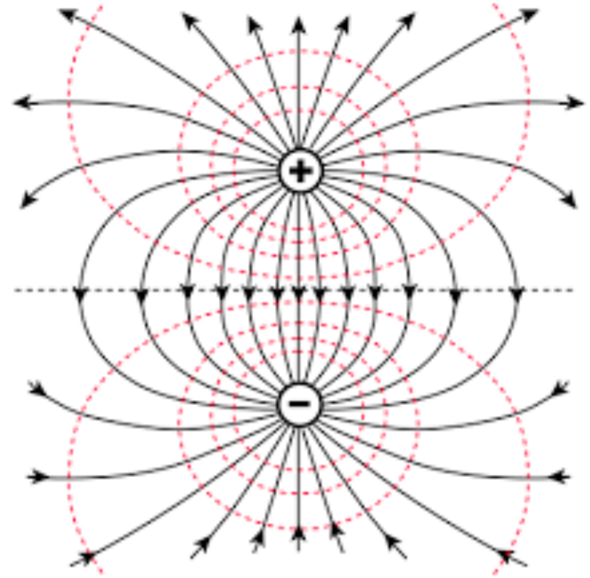


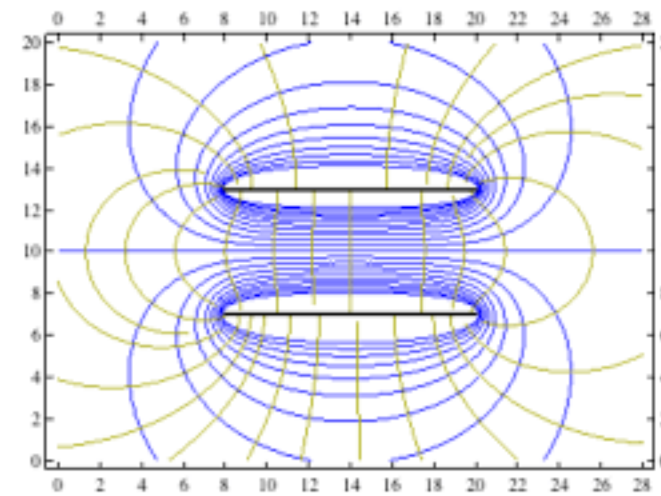


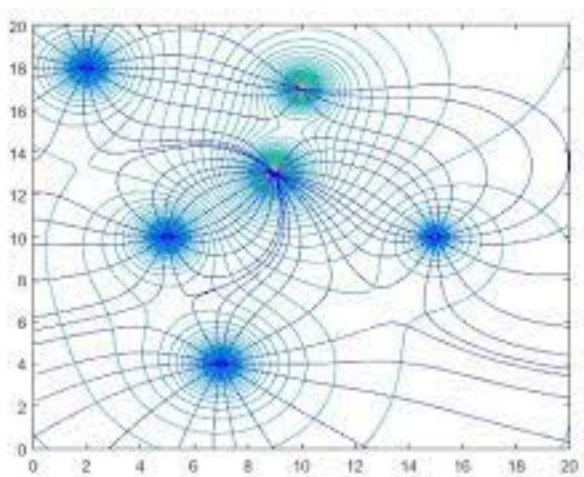


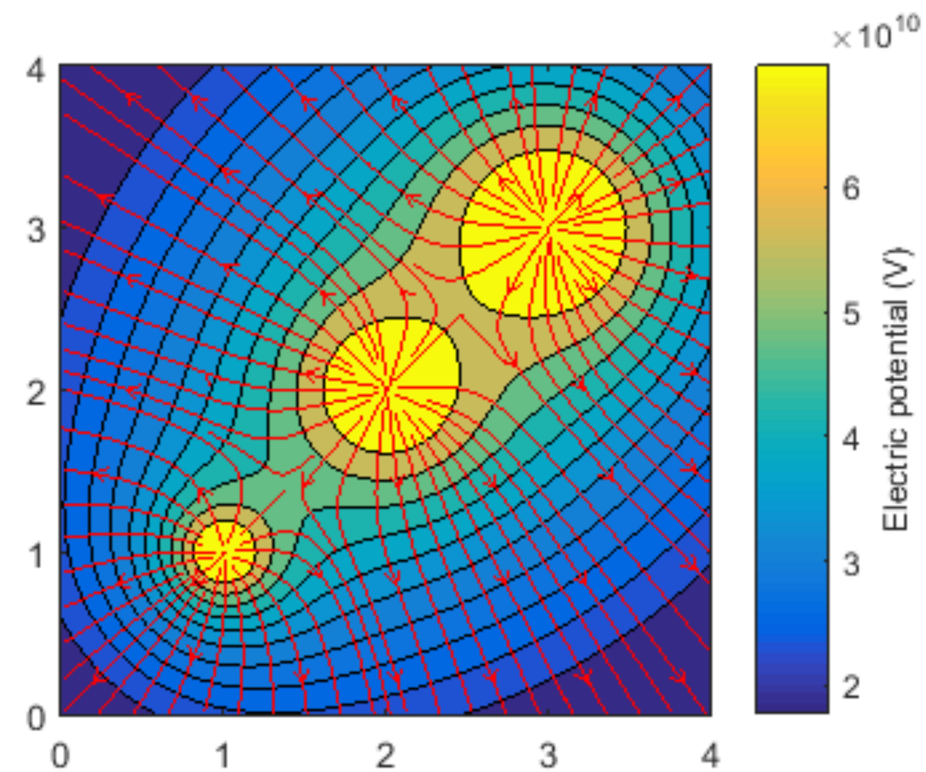
50
40
30
20
10
10
20
30
40
50











professionals…..

Now back to where we left off in last lecture.

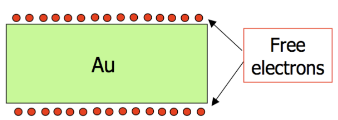
Review

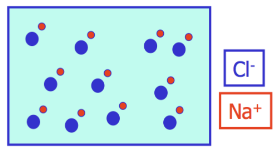
Conductors and Insulators

Conductor: a material with free electrons

Excellent conductors: metals such as Au, Ag, Cu, Al,...

OK conductors: ionic solutions such as NaCl in H2O





Insulator: a material without free electrons

Organic materials: rubber, plastic,...

Inorganic materials: quartz, glass,...

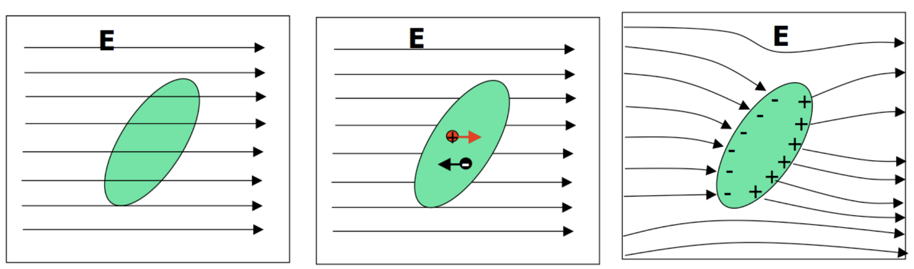
Properties of Conductors

A conductor is assumed to have an infinite supply of electric charges

Pretty good assumption...

Inside a conductor, E=0. Why? If E is not 0 charges will move from where the potential is higher to where the potential is lower; migration will stop only when E=0.

How long does it take? 10^{-17}-10^{-16}s (typical resistivity of metals)



Electric potential inside a conductor is constant

Given 2 points inside the conductor P1 and P2 the $\Delta\phi$ would be:

$$\Delta\phi = \int_{P_1}^{P_2} \vec{E}\cdot d\vec{s} = 0$$ since E=0 inside the conductor

Net charge can only reside on the surface

If net charge inside the conductor → Electric Field ≠ 0 (Gauss's law)

External field lines are perpendicular to surface

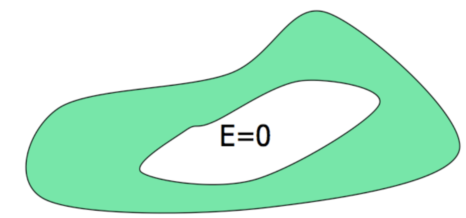
E || component would cause charge flow on the surface until $\Delta\phi=0$

Conductor's surface is an equipotential

Because it's perpendicular to field lines

Corollary 1:

In a hollow region inside conductor, ϕ=constant and E=0 if there aren't any charges in the cavity



Why?

Surface of conductor is equipotential

If no charge inside the cavity → Laplace equation holds → ϕ_{cavity} cannot have max or minima

→ ϕ must be constant → E=0

Consequence:

Shielding of external electric fields: Faraday's cage

Corollary 2

A charge +Q in the cavity will induce a charge +Q on the outside of the conductor

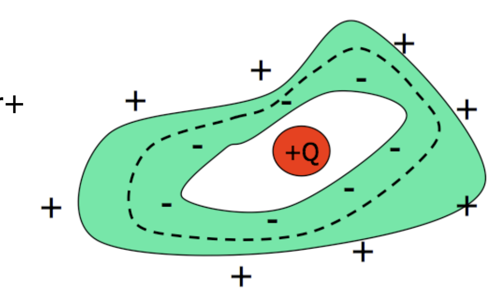
Why?

Apply Gauss's law to surface ---inside the conductor+

$\oint \vec{E}\cdot d\vec{A} = 0$ because E=0 inside a conductor

$\oint \vec{E}\cdot d\vec{A} = 4\pi(Q+Q_{inside})$ Gauss's law

$\Rightarrow Q_{inside} = -Q \Rightarrow Q_{outside} = -Q_{inside} = Q$ (conductor is overall neutral)



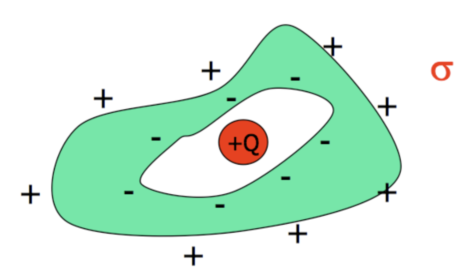
Corollary 3

The induced charge density on the surface of a conductor caused by a charge Q inside it is $\sigma_{induced} = E_{surface}/4\pi$

Why?

For surface charge layer, Gauss tells us that $\Delta E = 4\pi\sigma$

Since $E_{inside} = 0$ $\rightarrow E_{surface} = 4\pi\sigma_{induced}$



Uniqueness theorem

Given the charge density ρ(x,y,z) in a region and the value of the electrostatic potential ϕ(x,y,z) on the boundaries, there is only one function ϕ(x,y,z) which describes the potential in that region. As I just showed numerically for zero charge density!

Proof:

Assume there are 2 solutions: ϕ_1 and ϕ_2; they will satisfy the Poisson equation:

$$\nabla^2\phi_1 = 4\pi\rho(\vec{r}) \quad , \quad \nabla^2\phi_2 = 4\pi\rho(\vec{r})$$

Both ϕ_1 and ϕ_2 satisfy boundary conditions: on the boundary: $\phi_1 = \phi_2$ [On the boundary]

Superposition: any combination of ϕ_1 and ϕ_2 will be solution, including

$$\phi_3 = \phi_2 - \phi_1 : \quad \nabla^2\phi_3 = \nabla^2\phi_2 - \nabla^2\phi_1 = 4\pi\rho(\vec{r}) - 4\pi\rho(\vec{r}) = 0$$

So ϕ_3 satisfies the Laplace equation → no local maxima or minima inside the boundaries

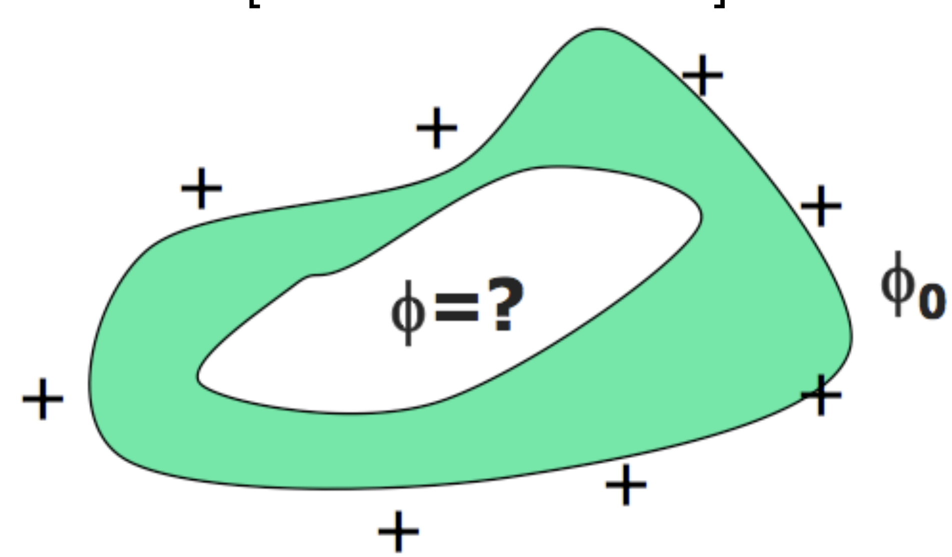
But on the boundaries $\phi_1 = \phi_2$ → $\phi_3 = 0$ → $\phi_3 = 0$ everywhere inside region
→ $\phi_1 = \phi_2$ everywhere inside region [Think relaxation method]

Uniqueness theorem: application 1

A hollow conductor is charged until its external surface reaches a potential (relative to infinity)

$\phi = \phi_0$

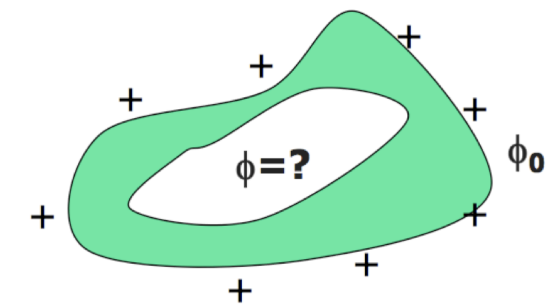
What is the potential inside the cavity?



Solution

$\phi=\phi_0$ everywhere inside the conductor's surface, including the cavity.

Why? think potential difference??



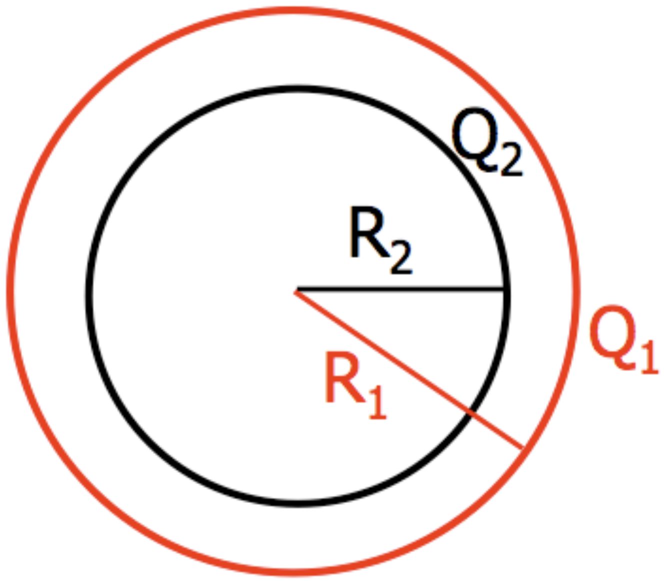
$\phi=\phi_0$ satisfies boundary conditions and the Laplace equation

The uniqueness theorem then tells me that it is THE solution.

That is very powerful!! There is only one! Think about that again!!!!!!!

Uniqueness theorem: application 2

Two concentric thin conductive spherical shells of radii R_1 and R_2 carry charges Q_1 and Q_2 respectively.



What is the potential of the outer sphere? ($\phi_{infinity}=0$)

What is the potential on the inner sphere?

What is the potential at r=0?

Solution

Outer sphere: $\phi_1=(Q_1+Q_2)/R_1$

Inner sphere

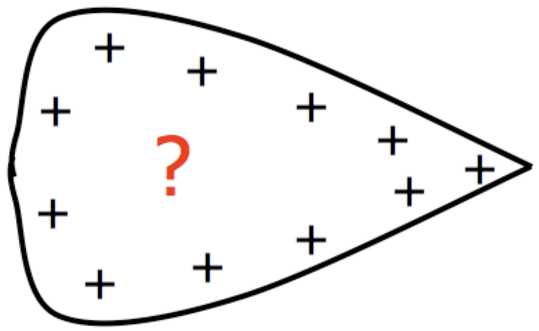
$$\phi_2 = \phi_1 - \int_{R_1}^{R_2} \vec{E}\cdot d\vec{s} = \phi_1 + \int_{R_1}^{R_2} \frac{Q_2}{r^2}dr = \phi_1 + \frac{Q_2}{R_2} - \frac{Q_2}{R_1}$$

$$\Rightarrow \phi_2 = \frac{Q_2}{R_2} - \frac{Q_2}{R_1} + \frac{Q_1}{R_1} + \frac{Q_2}{R_1} = \frac{Q_2}{R_2} + \frac{Q_1}{R_1}$$

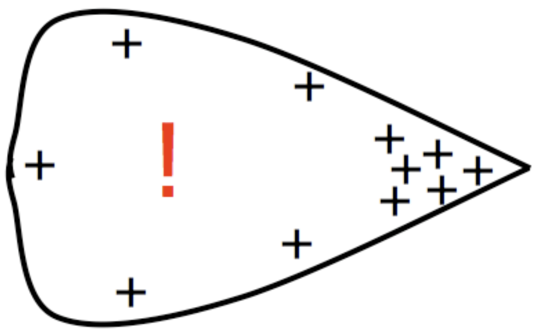
Because of uniqueness $\phi(r) = \phi_2 \quad \forall \quad r < R_2$ [if find solution it is it!]

Charge distribution on a conductor

Let's deposit a charge Q on a tear drop-shaped conductor
How will the charge distribute on the surface? Uniformly?

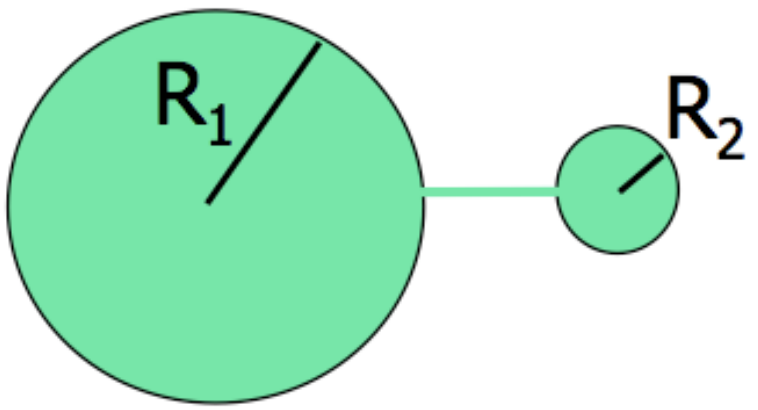


Experimental answer: NO!
$\sigma_{tip} >> \sigma_{flat}$
Important consequence
Although ϕ=constant and $E = 4\pi\sigma$
$E_{tip} >> E_{flat}$
Why?



Qualitative explanation

Consider 2 spherical conductors connected by a conductive wire
Radii: R_1 and R_2 with $R_1 >> R_2$
Deposit a charge Q on one of them
Charge redistributes itself until ϕ = constant

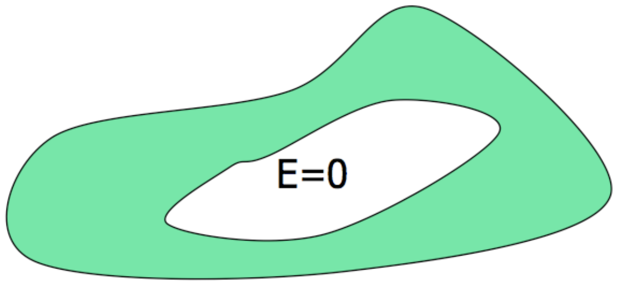


$$\phi_1 = \frac{Q_1}{R_1} = \frac{Q_2}{R_2} = \phi_2 \quad , \quad E_1 = \frac{Q_1}{R_1^2} = \frac{\phi_1}{R_1} \quad , \quad E_2 = \frac{Q_2}{R_2^2} = \frac{\phi_2}{R_2} \Rightarrow \frac{E_1}{E_2} = \frac{R_2}{R_1} \Rightarrow \frac{\sigma_1}{\sigma_2} = \frac{R_2}{R_1}$$

Conclusion: Electric field is strong where radius of curvature (R) is small

Shielding

We proved that in a hollow region inside a conductor E=0

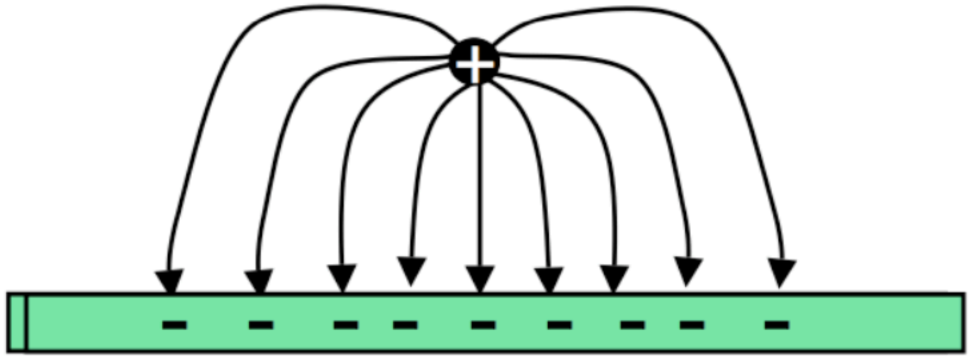


This is the principle of shielding - Example is a Faraday Cage

Next slide for video

Application of Uniqueness Theorem: Method of images

What is the electric potential created by a point charge +Q at a distance y from an infinite conductive plane?



Consider field lines:

Radial around the charge

Perpendicular to the surface conductor

——> The point charge +Q induces – charges on the conductor

Method of images

Apply the uniqueness theorem

It does not matter how you find the potential ϕ as long as the boundary conditions are satisfied. The solution is unique.

In our case: on the conductor surface: $\phi=0$ and E always perpendicular!


SPARK
SPARK

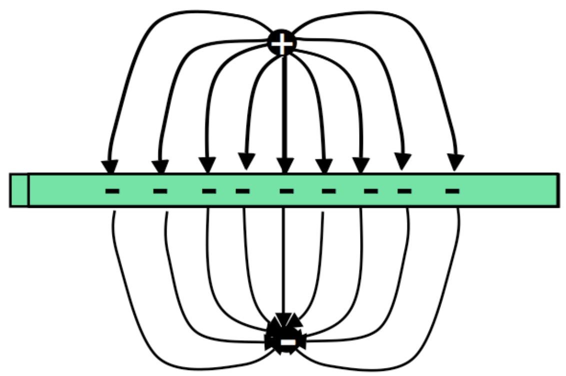
Can we find an easier configuration of charges that will create the same field lines above the conductor surface? YES!

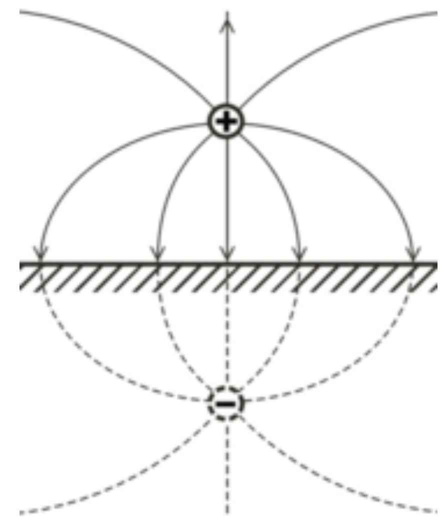
For this system of point charges we can calculate $\phi(x,y,z)$ anywhere

This is THE solution (uniqueness)

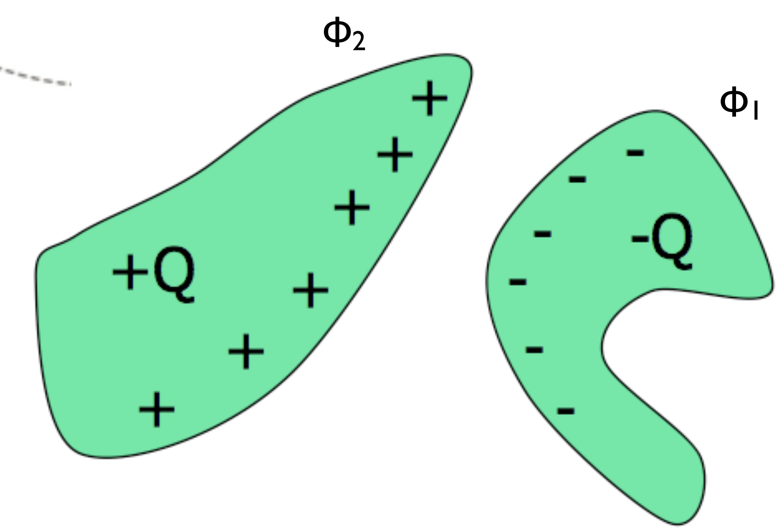
Note: we do not care what happens below the surface of the conductor:

That is not the region under study





Capacitance



Consider 2 conductors at a certain distance apart

Deposit charge +Q on one and –Q on the other

They are conductors

➙ each surface is equipotential

What is the $\Delta\phi$ between the two conductors?

Let's try to calculate:

$$V = \phi_2 - \phi_1 = -\int_1^2 \vec{E}\cdot d\vec{s} = Q \times (\text{constant depending on geometry})$$

Naming the proportionality constant 1/C: $\Rightarrow Q = CV$

Definitions: C = capacitance of the system

Capacitor: system of 2 oppositely charged conductors

Units of capacitance

Definition of capacitance: $$Q = CV \Rightarrow C = \frac{Q}{V}$$

Units:

SI: Farad (F) = Coulomb/Volt

cgs: cm = esu/(esu/cm)

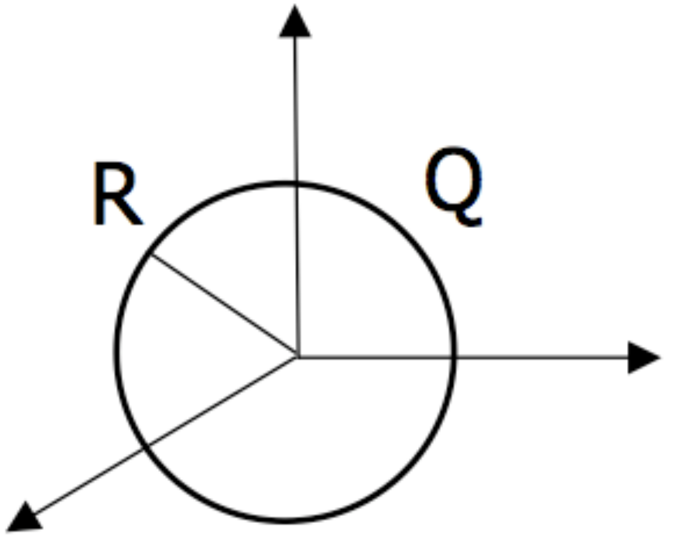
Conversion: 1 cm = 1.11 x 10^{-12}F ~ 1 pF

Remember:

1 Coulomb is a BIG charge: 1 F is a BIG capacitance

Usual C ~ pF - μF

Simple capacitors: Isolated Sphere



Conductive sphere of radius R origin (0,0,0) with a charge Q

Review questions:

Where is the charge located? ---- surface: E=0 inside

Hollow sphere? Solid sphere? Why? ----- same

What is the E everywhere in space? ------ point charge

Is this a capacitor?

Yes! The second conductor is a virtual one at infinity

Calculate the capacitance:

$$V = \phi_R - \phi_\infty = \frac{Q}{R} \quad , \quad Q = CV \Rightarrow C_{sphere} = R$$

Capacitors are everywhere!

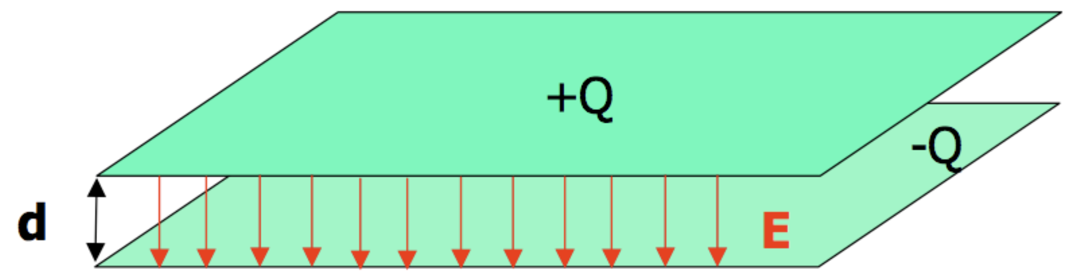
The prototypical capacitor: Parallel plates

Physical configuration:

2 parallel plates, each of area A, at a separation distance d

Note: if $d^2 << A$ ~ infinite parallel planes

Deposit +Q on top plate and –Q on bottom plate



Think 2 charge sheets + and - net field only inside = $4\pi\sigma$

Capacitance:

$$V = -\int_{bottom}^{top} \vec{E} \cdot d\vec{s} = \int_{bottom}^{top} 4\pi\sigma \hat{n} \cdot d\vec{s} = 4\pi \frac{Q}{A} d$$

$\sigma = Q/A$

$$\Rightarrow C = \frac{Q}{V} = \frac{A}{4\pi d}$$

Observations:

C depends only on the geometry of the arrangement

As it should, not on Q deposited or V between the plates!

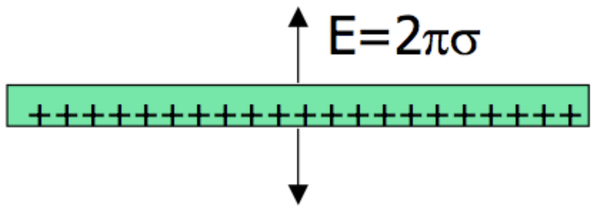
Electric field on surface of conductor: $2\pi\sigma$ or $4\pi\sigma$???

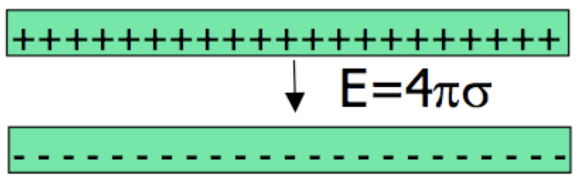
Infinite plane of charges: $E=2\pi\sigma$

With $\sigma=Q/A$

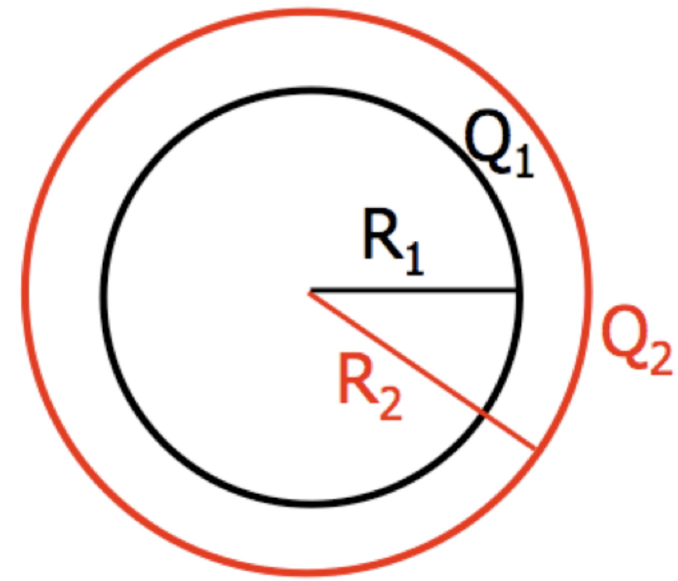
Always $\Delta E=4\pi\sigma$!

What is the E outside the capacitor? Zero!





E in Nested Spherical Shells



Configuration:

2 concentric spherical shells

Charge: +Q (–Q) on inner (outer) sphere

Calculate E in the following regions:

$r < R_1$, $R_1 < r < R_2$, $r > R_2$

Gauss's law is the key.

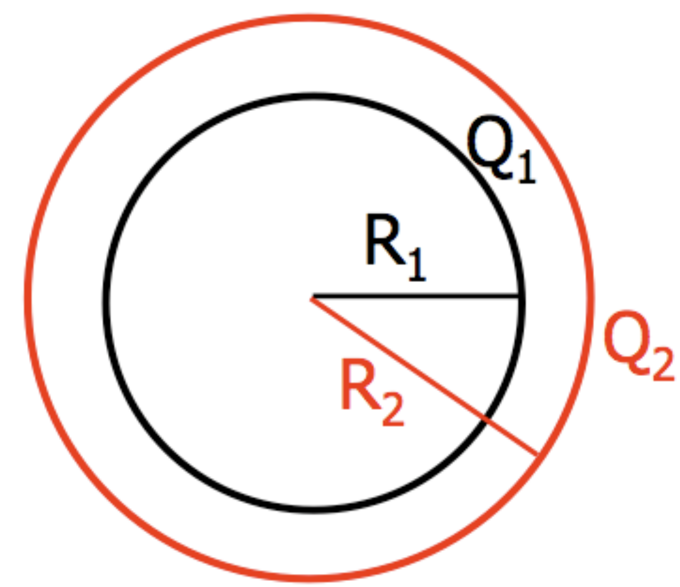
On spherical surface with $r < R_1$ $Q_{enc} = 0 \rightarrow E=0$

On spherical surface with $r > R_2$ $Q_{enc} = +Q - Q = 0 \rightarrow E=0$

On spherical surface with $R_1 < r < R_2$ $Q_{enc} = +Q$

$$\int \vec{E} \cdot d\vec{A} = E 4\pi r^2 = 4\pi Q \Rightarrow \vec{E} = \frac{Q}{r^2}\hat{r}$$

More capacitors: Nested Spherical Shells



Same configuration: 2 concentric spherical shells

Charge: +Q (–Q) on inner (outer) sphere

Capacitance:

Key: find the potential difference V

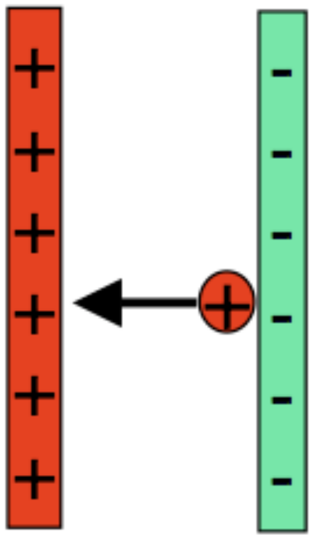
$$V = \phi_1 - \phi_2 = \int_{R_2}^{R_1} \vec{E} \cdot d\vec{s} = -\int_{R_2}^{R_1} \frac{Q}{r^2} dr = \frac{Q}{R_1} - \frac{Q}{R_2} \Rightarrow C = \frac{Q}{V} = \frac{R_1 R_2}{R_2 - R_1}$$

If $R_2 - R_1 = d << 0 : R_2 \sim R_1$

$$C = \frac{R_1 R_2}{R_2 - R_1} \sim \frac{R_1^2}{d} = \frac{4\pi R_1^2}{4\pi d} = \frac{A_{sphere}}{4\pi d}$$

same as plane capacitor!

Energy stored in a capacitor



Consider a capacitor with charge +/- q

How much work is needed to bring a positive charge dq from the negative plate to the positive plate?

We are charging the capacitor!

$$dW = V(q)dq = \frac{q}{C}dq$$

How much work is needed to charge the capacitor from scratch?

$$W = \int_0^Q dW = \int_0^Q \frac{q}{C}dq = \frac{Q^2}{2C}$$

Energy stored in the capacitor: $U = \frac{Q^2}{2C} = \frac{1}{2}CV^2$

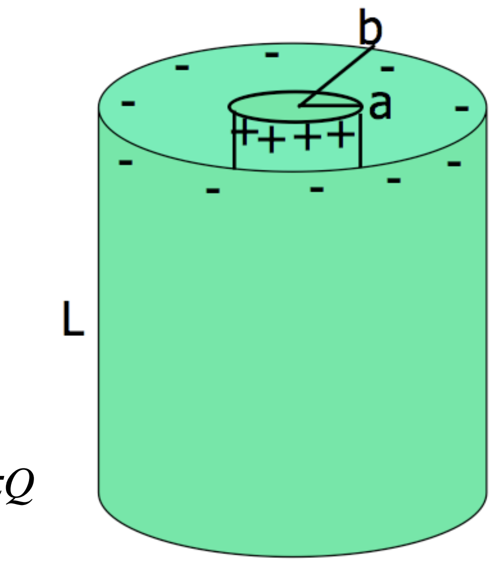
Is this result consistent with what we found earlier?

Example: parallel plate capacitor

$$U = \frac{1}{8\pi}\int E^2 dV = \frac{1}{8\pi}E^2 Ad = \frac{1}{8\pi}(4\pi\sigma)^2 Ad\frac{A}{A} = \frac{Q^2}{2}\left(4\pi\frac{d}{A}\right) = \frac{Q^2}{2C}$$

Yes!

Cylindrical Capacitor



Concentric cylindrical shells with charge +/- Q. Calculate:

Electric Field in between plates

r < a and r > b: E=0 (using Gauss)

a < r < b: Gauss's law on cylinder of radius r

$$\vec{E}(r) = \frac{2Q}{L}\frac{\hat{r}}{r}$$

$$A = 2\pi rL \quad , \quad E \quad radial$$
$$EA = E2\pi rL = 4\pi Q_{enc} = 4\pi Q$$
$$\vec{E} = \frac{2Q}{L}\frac{\hat{r}}{r}$$

V between plates:

$$V = \int_a^b \vec{E}\cdot d\vec{r} = \int_a^b \frac{2Q}{L}\frac{dr}{r} = \frac{2Q}{L}\ln\frac{b}{a}$$

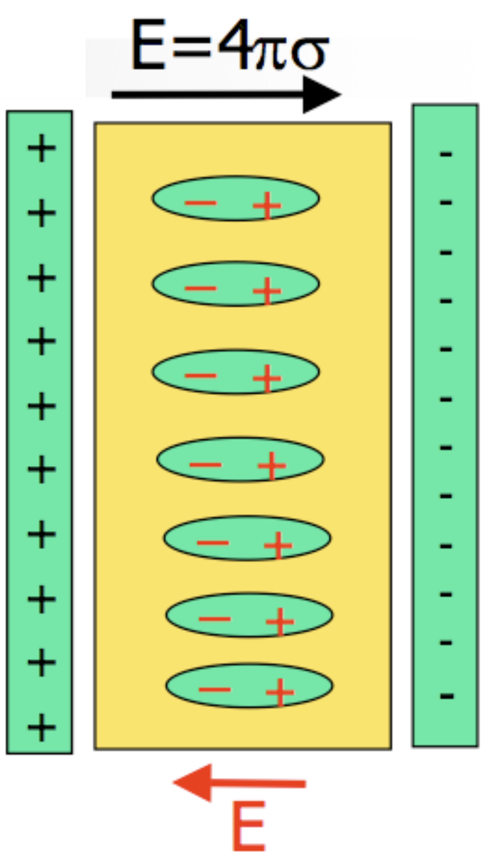
Capacitance C:

$$C = \frac{Q}{V} = \frac{L}{2\ln\frac{b}{a}}$$

Calculate energy stored in capacitor:

$$U = \frac{1}{2}CV^2 \frac{1}{2}\frac{L}{2\ln\frac{b}{a}}\left(\frac{2Q}{L}\ln\frac{b}{a}\right)^2 = \frac{Q^2}{L}\ln\frac{b}{a}$$

Capacitors and dielectrics



Parallel plates capacitor: $$C = \frac{Q}{V} = \frac{Q}{Ed} = \frac{A}{4\pi d}$$

Add a dielectric(non-conducting(insulator) between the plates
- Dielectric's molecules are not spherically symmetric
- Electric charges are not free to move

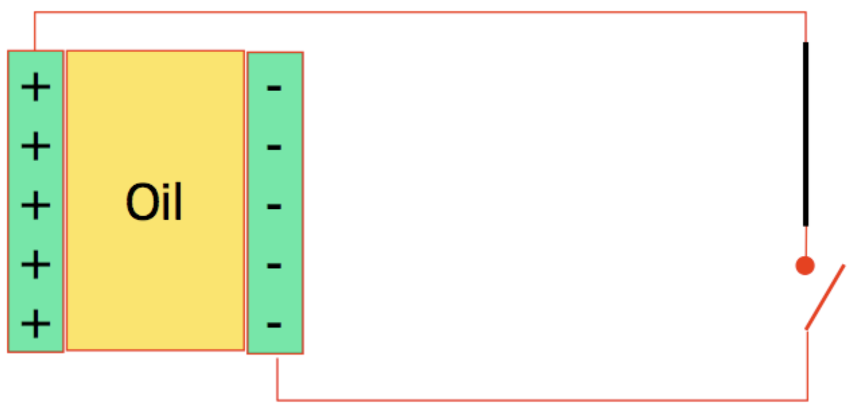
E will pull + and – charges apart and orient them || E
$E_{dielectric}$ (extra new field due to charge separation) is opposite to $E_{capacitor}$
- Given Q → V decreases
- Given V → Q increases
  - either way → C is increased!

Energy is stored in capacitors

A 100 μF oil filled capacitor is charged to 4 kV.
What happens if we discharge it through a 12" long iron wire?

How much energy is stored in the capacitor?

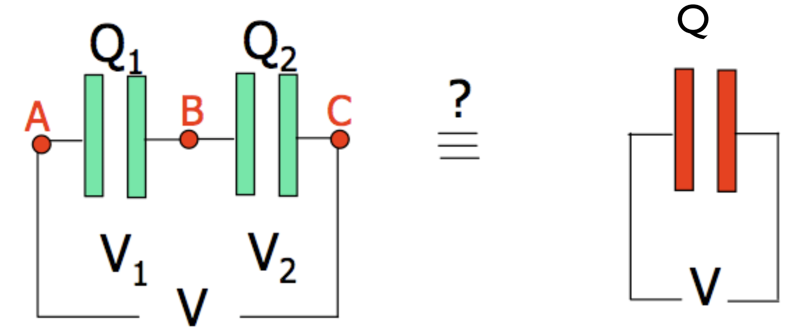
$U = \frac{1}{2} CV^2 = 800$ J Big!!!!!



Resistance of iron wire: very small, but >> than the rest of the circuit
All the energy is dumped on the wire in a small time
Huge currents! Huge temperatures! The wire will explode!

Capacitors in series

Let's connect 2 capacitors C_1 and C_2 in the following way:

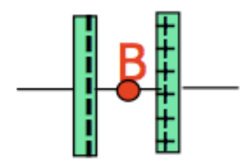


What is the total capacitance C of the new system?

follow path A to C

$$V_1 + V_2 = V \quad , \quad Q_1 = Q_2 = Q$$

$$\frac{1}{C} = \frac{V}{Q} = \frac{V_1 + V_2}{Q} = \frac{1}{C_1} + \frac{1}{C_2}$$



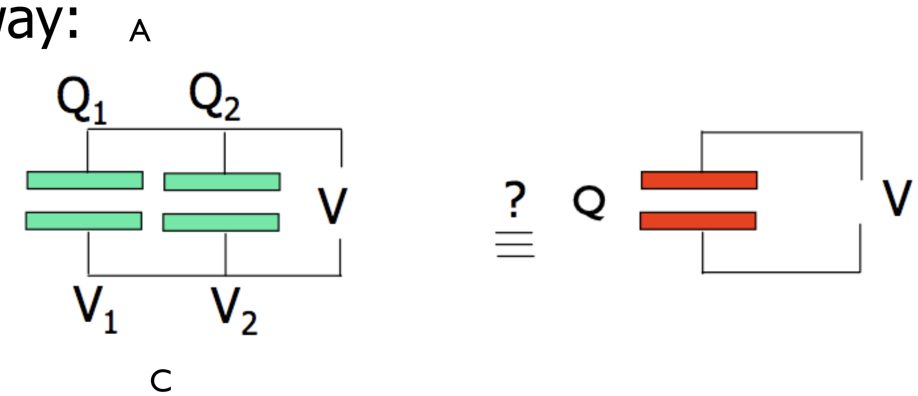
This is 1 conductor that starts electrically neutral $Q_1 = Q_2$

follow charge movement

$$C = \left(\sum_i \frac{1}{C_i} \right)^{-1}$$

Capacitors in parallel

Let's connect 2 capacitors C_1 and C_2 in the following way:



What is the total capacitance C of the new system?

$$V_1 = V_2 = V \quad , \quad Q_1 + Q_2 = Q$$

follow two paths A to C

$$C = \frac{Q_1 + Q_2}{V} = \frac{Q_1}{V_1} + \frac{Q_2}{V_2} = C_1 + C_2$$

$$C = \sum_i C_i$$

Why are capacitors useful?

…among other things…

They can store large amount of energy and release it in very short time

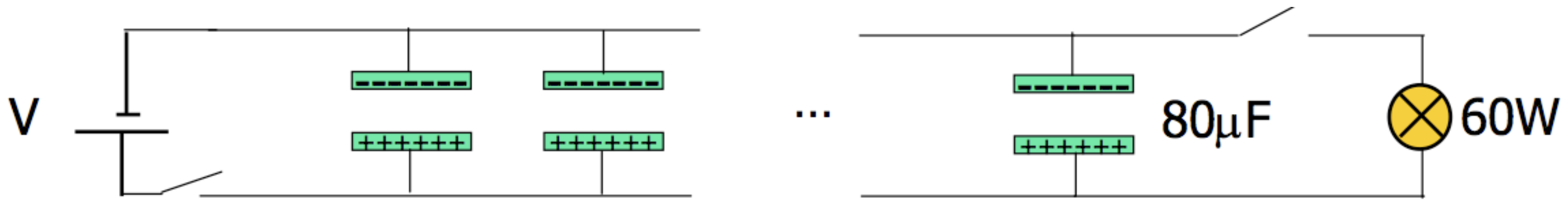
Energy stored: $U = \frac{1}{2} CV^2$

The larger the capacitance, the larger the energy stored at a given V

How to increase the capacitance?

- Modify geometry
  - For parallel plates capacitors $C=A/(4\pi d)$: increase A or decrease d
- Add a dielectric in between the plates
- Add capacitors in parallel

Bank of capacitors: Bank of 12 x 80 μF capacitors is parallel



Total capacitance: 960 μF

Discharged on a 60 W light bulb when capacitors are charged at:

V = 100 V, 200 V, V = 300 V

What happens?

Energy stored in capacitor is $U = \frac{1}{2} CV^2$

→ $V = V_0 : 2xV_0 : 3xV_0$ → $U = U_0 : 4xU_0 : 9xU_0$

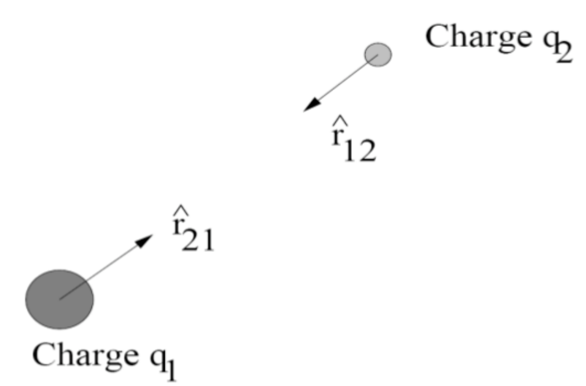
R is the same →time of discharge will not change with V

The power will increase by a factor 9! ($P=RI^2$ and $I=V/R$)

Will the bulb survive?

Current very high, T very high →wire will "explode again"

Some final thoughts before proceeding to currents and circuits

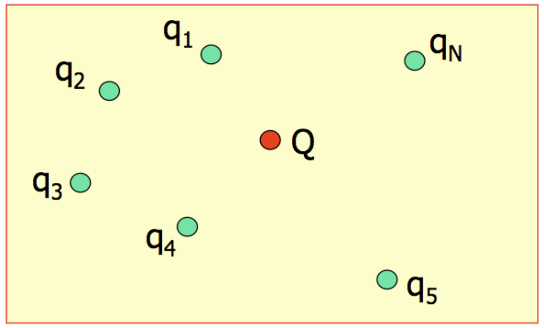


Coulomb's law $$\vec{F}_2 = \frac{q_1 q_2}{|r_{21}|^2}\hat{r}_{21}$$

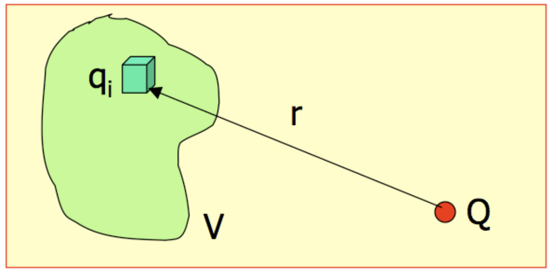
where F_2 is the force that the charge q_2 feels due to q_1

In principle this is the only thing you have to remember:
all the rest follows from this and the superposition principle!!

Superposition principle $$\vec{F}_Q = \sum_{i=1}^{N} \frac{q_i Q}{|r_i|^2}\hat{r}_i$$



$$\vec{F}_Q = \int_V \frac{Q\,dq}{|r|^2}\hat{r} = \int_V \frac{Q\rho\,dV}{|r|^2}\hat{r}$$

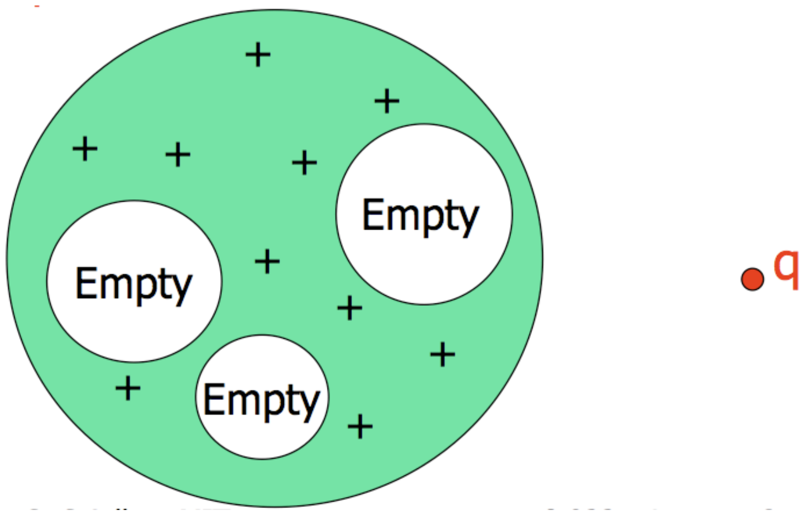


The Importance of Superposition

Extremely important because it allows us to transform complicated problems into sum of small, simple problems that we know how to solve.

Example:

Calculate force F exerted by this distribution of charges on the test charge q



Solving problems in terms of $F_{coulomb}$ is not always convenient; F depends on probe charge q

Remove dependence by introducing the Electric Field

E describes the properties of space due to the presence of charge Q
It's a vector→hard integrals when applying superposition.

Introduce Electric Potential

Superposition still holds but simpler calculation (scalar)

Energy stored in the electric field

$$U = \frac{1}{2}\int_{\substack{\text{Volume}\\\text{with}\\\text{charges}}} \rho\phi dV = \int_{\substack{\text{Entire}\\\text{space}}} \frac{E^2}{8\pi} dV$$

In electrostatics there are 3 different ways of describing a problem:

From ρ→E

Point charge
$$\vec{E} = \frac{q}{|r|^2}\hat{r}$$

Superposition principle
$$\vec{E} = \int_V d\vec{E} = \int_V \frac{dq}{|r|^2}\hat{r}$$

Look for symmetry and thank Mr. Gauss who solved the integrals for you

Gauss's Law:
$$\oint_S \vec{E} \bullet d\vec{A} = 4\pi \int_V \rho\, dV$$

Gauss's law is always true but not always useful: Symmetry is needed!
Main step: choose the "right" gaussian surface so that E is constant on the surface of integration

From $\rho \to \phi$

Point charge: $\phi = \frac{q}{r}$

Superposition principle: $\phi = \int_V \frac{dq}{r}$

implicit hypothesis: $\phi(\text{infinity}) = 0$

The problem is simpler than for E(only scalars involved) but not trivial…

Special cases:

If symmetry allows, use Gauss's law to extract E and then integrate E to get ϕ:

$$\phi_2 - \phi_1 = -\int_1^2 \vec{E} \bullet d\vec{s}$$

The force is conservative $\to$ the result is the same for any path, but choosing a simple one makes your life much easier….

From ϕ to E and ρ Easy! No integration needed!

From ϕ to E $\vec{E} = -\nabla \phi$

One derivative is all it takes but… make sure you choose the best coordinate system

From ϕ to ρ

Poisson tells you how to get from potential to charge distributions directly:

$$\nabla^2 \phi = -4\pi\rho$$

Uncomfortable with Laplacian? Get there in 2 steps:

First calculate E: $\vec{E} = -\nabla \phi$

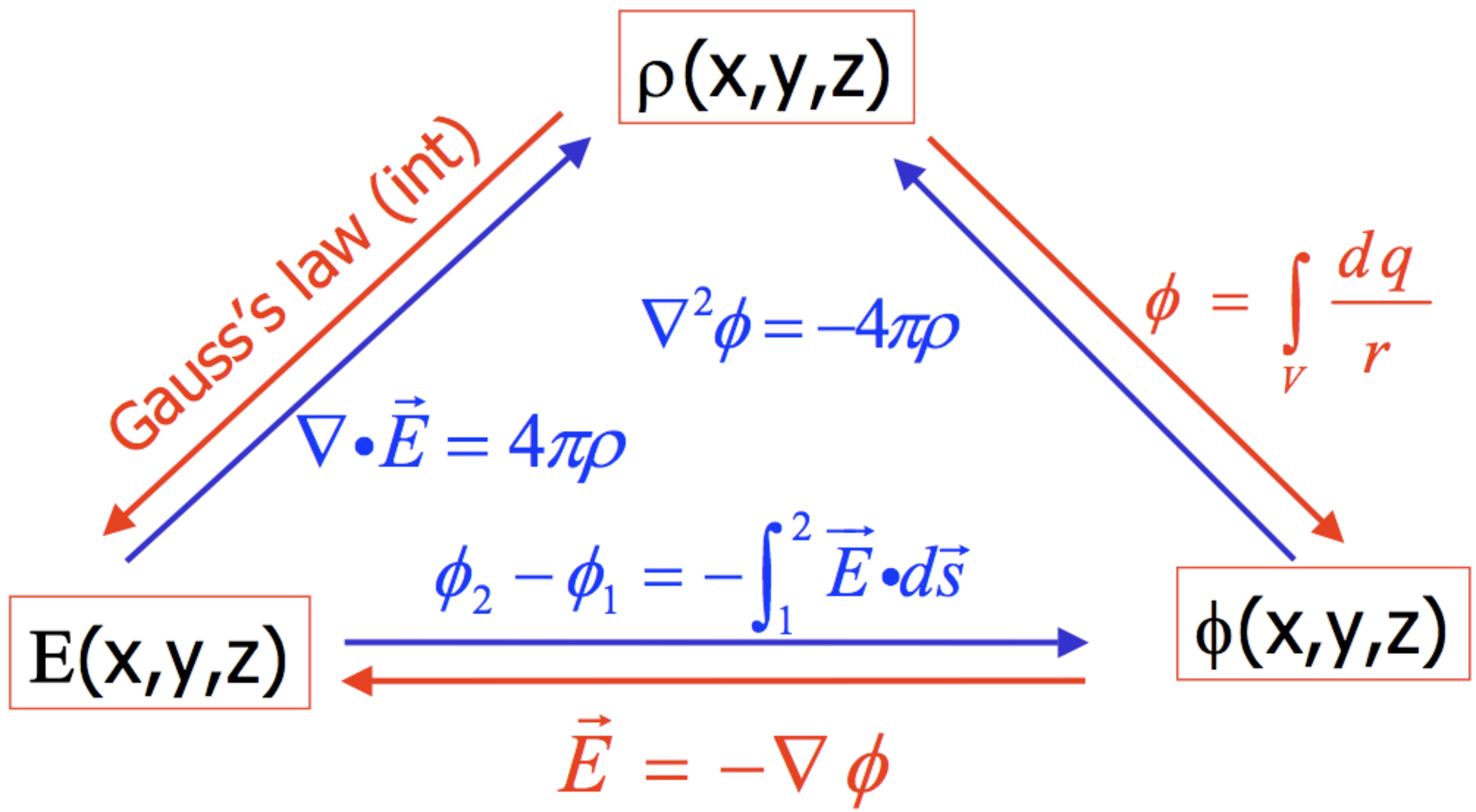
Then use differential form of Gauss's law: $\nabla \bullet \vec{E} = 4\pi\rho$

Some Thoughts

The potential ϕ is always continuous

E is not always continuous: it can "jump"
When we have surface charge distributions

Summary



Conductors

Properties:

Surface of conductors are equipotential

E (field lines) always perpendicular to the surface

$E_{inside} = 0$

$E_{surface} = 4\pi\sigma$

What's the most useful info?

$E_{inside} = 0$ because it comes handy in conjunction with Gauss's law to solve problems of charge distributions inside conductors.

Example: concentric cylindrical shells

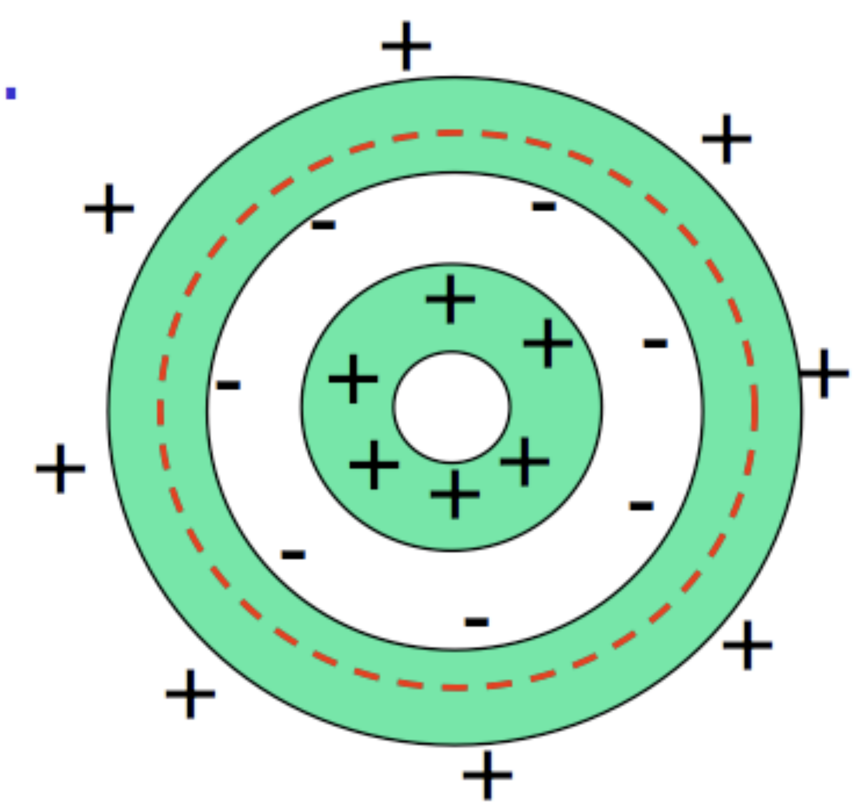
Charge +Q deposited in inner shell

No charge deposited on external shell

What is E between the 2 shells?

-Q induced on inner surface of inner cylinder

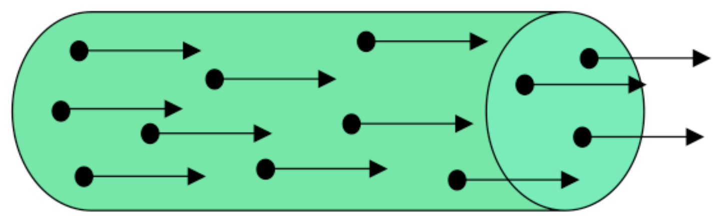
+Q induced on outer surface of outer cylinder



already solved……

Electric current I

Consider a region in which there is a flow of charges:
cylindrical conductor



We define the current I as the charge/unit time flowing through a certain surface

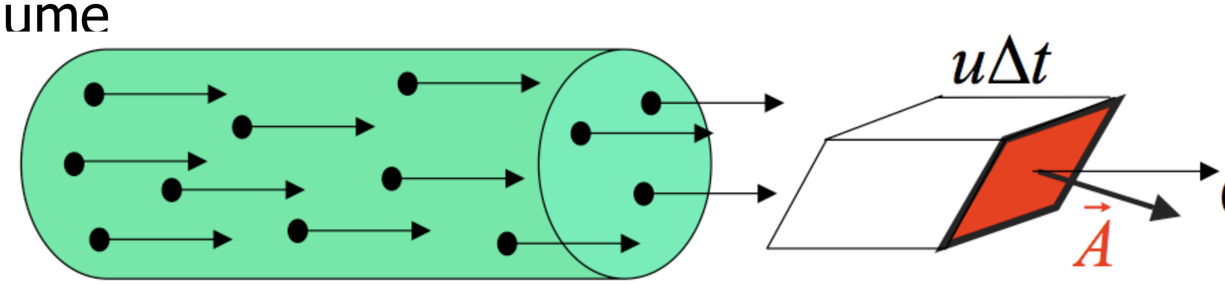
$$I = \frac{dQ}{dt}$$

Units:
cgs: esu/s
SI: C/s = Ampere (A)
Conversion: 1 A = 2.998 x 10^9 esu/s

Current density J

Number density: n= #charges/unit volume

Velocity of each charge: u



Current flowing through area A: I = ΔQ/Δt

where ΔQ = q x number of charges in the prism

$$\rightarrow I = \frac{\Delta Q}{\Delta t} = \frac{q\Delta N}{\Delta t} = \frac{qnV_{prism}}{\Delta t} = \frac{qnA\cos\theta u\Delta t}{\Delta t} = qn\vec{u}\cdot\vec{A} = \vec{J}\cdot\vec{A}$$

Where we defined the current density J as: $\vec{J} = qn\vec{u} = \rho\vec{u}$

More realistic case…

We made a number of unrealistic assumptions:

- Only 1 kind of charge carriers: we could have several, e.g.: + and –ions
- u assumed to be the same for all particles: unrealistic!
- Regular surface with J constant on it

Multiple charge carriers: $\vec{J} = \sum_k q_k n_k \vec{u}_k = \sum_k \rho_k \vec{u}_k$

E.g.: solution with different kind of ions

Note: + ion with velocity u_k is equivalent to –ion with velocity $-u_k$

Velocity:

Not all charges have the same velocity→average velocity $\langle \vec{u}_k \rangle = \frac{1}{N_k} \sum_i (\vec{u}_k)_i$

$$\vec{J} = \sum_k q_k n_k \langle \vec{u}_k \rangle = \sum_k \rho_k \langle \vec{u}_k \rangle$$

Arbitrary surface S, arbitrary J: $I = \int_S \vec{J} \cdot d\vec{A}$

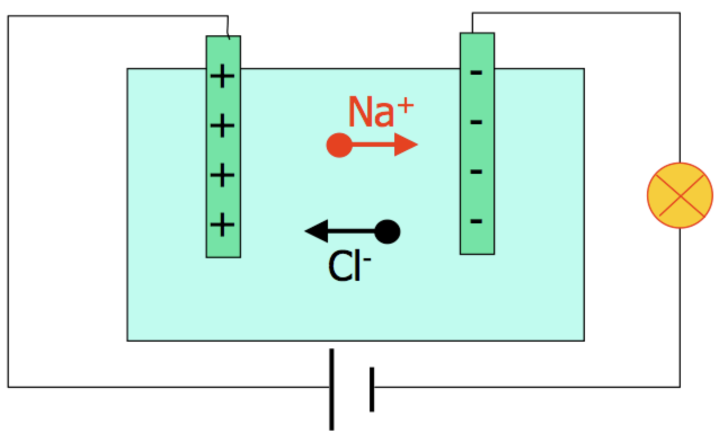
Non standard currents

Usually think of currents as electrons moving inside conductor

This is only one of the many examples!

Other kinds of currents

Ions in solution such as Salt (NaCl) in water



Note: more salt→higher I

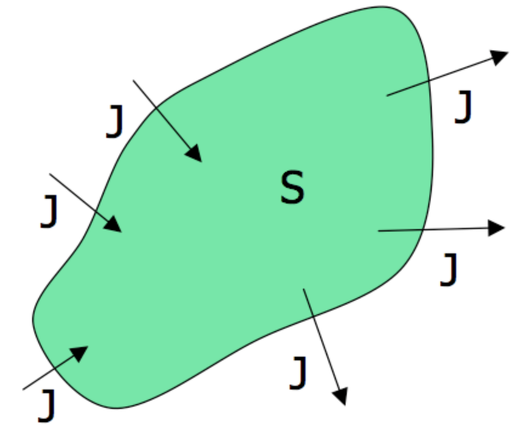
Why? J depends on n!

The continuity equation

A current I flows through the closed surface S:

Some charge enters

Some charge exits



What happens to the charge after it enters?

Piles up inside

Leaves the surface

$$\oint_S \vec{J} \cdot d\vec{A} = -\frac{\partial Q_{inside}}{\partial t}$$

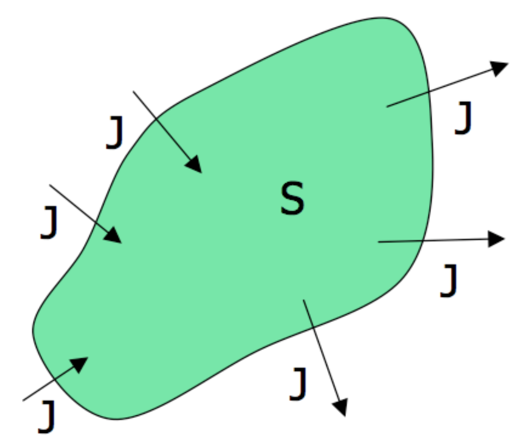
Note: - sign because dA points outside the surface

Apply Gauss's theorem and obtain continuity equation:

$$\left\{ \begin{array}{l} \oint_S \vec{J} \cdot d\vec{A} = \oint_V \nabla \cdot \vec{J} dV \\ -\frac{\partial Q_{inside}}{\partial t} = -\frac{\partial}{\partial t} \int_V \rho dV \end{array} \right\} \Rightarrow \int_V \left(\nabla \cdot \vec{J} + \frac{\partial \rho}{\partial t} \right) dV = 0 \Rightarrow \nabla \cdot \vec{J} + \frac{\partial \rho}{\partial t} = 0$$

Thoughts on continuity equation

$$\nabla \cdot \vec{J} + \frac{\partial \rho}{\partial t} = 0$$



What does it teach us?

Conservation of electric charges in presence of currents

From microscopic point of view

For steady currents:

no accumulation of charges inside the surface: dρ/dt = 0 $\rightarrow \nabla \cdot \vec{J} = 0$

Microscopic Ohm's law

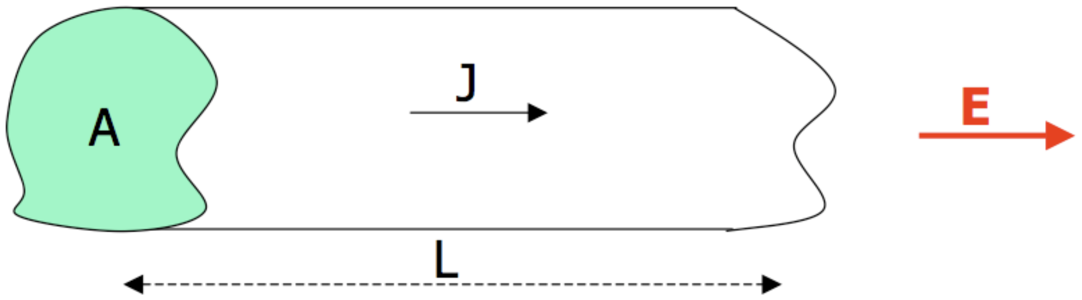
- Electric fields cause charges to move
  - Not too surprising…
- Experimentally, it was observed by Ohm (19th century) that for large number of materials:

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

- Microscopic version of Ohm's law:
  - Proportionality between E and J in each point in space
- Proportionality constant: conductivity σ
  - Typical of material considered at a certain temperature

Macroscopic Ohm's law

- Current is flowing in a uniform material of length L in uniform electric field E || L



- Potential difference between two ends: V = EL
- Ohm's law J=σE holds in every point:

$$\vec{J} = \sigma \vec{E} \Rightarrow \frac{I}{A} = \sigma \frac{V}{L} \Rightarrow V = IR \quad where \quad R = \frac{L}{\sigma A}$$

Resistance R

- Proportionality constant between V and I in Ohm's law

$$R = \frac{L}{\sigma A} = \frac{\rho L}{A} \quad with \quad \rho = \frac{1}{\sigma} \rightarrow resistivity$$

Units: [V]=[R][I]
SI: Ohm (Ω) = V/A
cgs: [esu/cm]/[esu/s]= s/cm

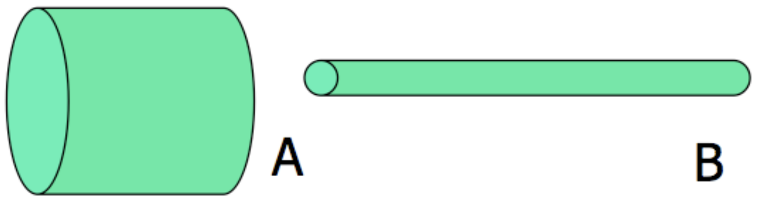
Dependence on the geometry:

Inversely proportional to A and proportional to L

Dependence on the property of the material:

Inversely proportional to conductivity

What wire has the largest R?



Resistivity

Resistivity $\rho = 1/\sigma$

Describes how fast electrons can travel in the material

Units: in SI: Ω-m; in cgs: s

Material	Resistivity (Ω-m)	Resistivity (sec)
Silver	1.6×10^{-8}	1.8×10^{-17}
Copper	1.7×10^{-8}	1.9×10^{-17}
Gold	2.4×10^{-8}	2.6×10^{-17}
Iron	1.0×10^{-7}	1.1×10^{-16}
Sea water	0.2	2.2×10^{-10}
Polyethylene	2.0×10^{11}	220
Glass	$\sim 10^{12}$	$\sim 10^{3}$
Fused quartz	7.5×10^{17}	8.3×10^{8}

Depends on chemistry of material, temperature,…

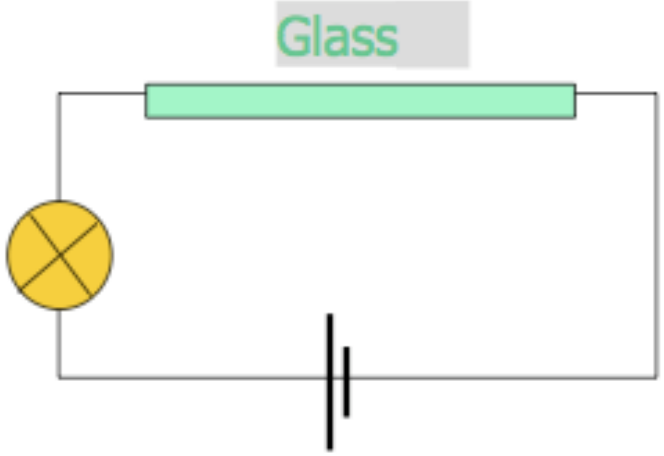
Resistivity vs. Temperature

Does resistivity depend on T?

Does glass conduct electricity?

Room T: no

What if we increase T? Yes



Effect of T on R $R = \frac{\rho L}{A}$

What is I?

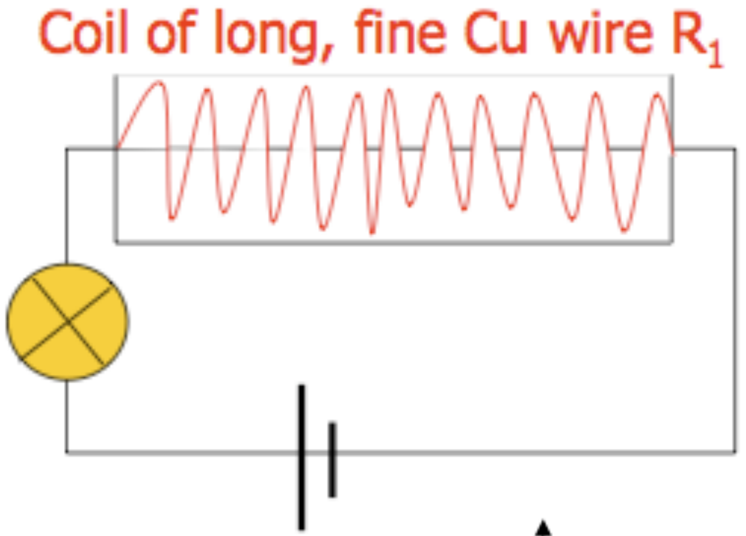
$I = V/(R_1 + r_2)$

Cool R_1 with liquid N

r_2 is small…

R lower, I higher

Power = RI^2 increases!



Large dependence of ρ vs T!

Room temperature:

ρ depends upon collisional processes:

T increases

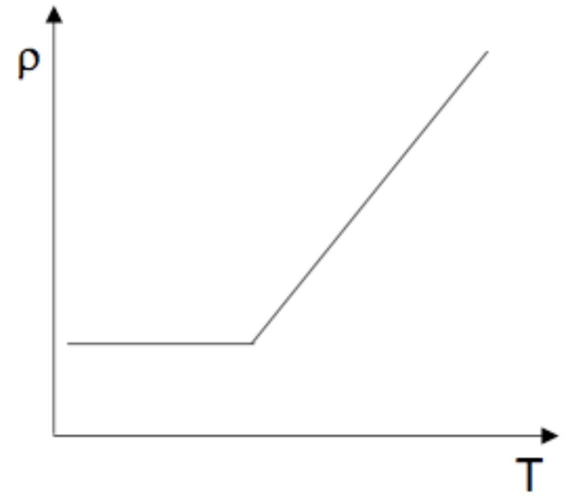
→more collisions→ρ increases

Very low temperature:

Mean free path dominated by impurities or defects in the material→~constant with temperature.

With sufficient purity, some metals become superconductors

ρ(Hg)~0 at T < 4 Kelvin: you can keep I in a ring for years!



Thoughts on Ohm's law

Ohm's law in microscopic formulation: $\vec{J} = \sigma \vec{E}$

In plain English:

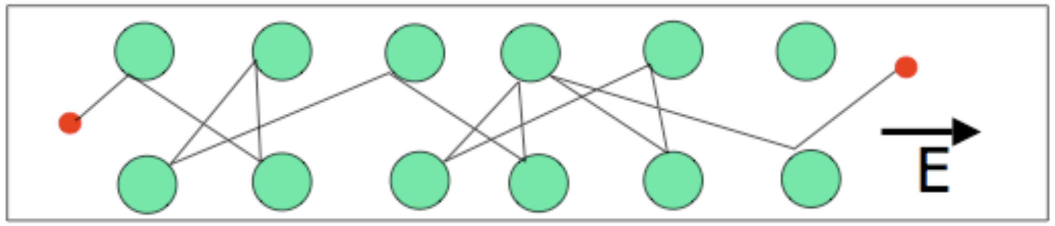
A constant electric field creates a steady current: $\vec{E} \propto \vec{v}$

Does this make sense? $\vec{F} = m\vec{a} \Rightarrow \vec{E} \propto \vec{a}$

Charges are moving in an effectively viscous medium

As sky diver in free fall: first accelerate, then reach constant v

Why? Charges are accelerated by E but then bump into nuclei and are scattered→the average behavior is a uniform drift



Motion of electrons in conductor

N electrons are moving in a material immersed in E

Momentum of each electron has 2 components:

Random collision velocity u_0: $\vec{p}_{random} = m\vec{u}_0$

Impulse due to electric field: $\vec{p}_E = q\vec{E}t$

The average momentum of each electron is:

$$\langle p\rangle = m\langle u\rangle = \frac{1}{N}\sum_{i=1}^{N}\left(m\vec{u}_i + q\vec{E}t_i\right) = m\frac{1}{N}\sum_{i=1}^{N}\vec{u}_i + q\vec{E}\frac{1}{N}\sum_{i=1}^{N}t_i$$

For large N: $$\sum_{i=1}^{N}\vec{u}_i \to 0 \Rightarrow m\langle u\rangle = q\vec{E}\frac{1}{N}\sum_{i=1}^{N}t_i = q\vec{E}\tau$$

Where $\tau = \frac{1}{N}\sum_{i=1}^{N}t_i$ is the average time between 2 collisions

Property of the material

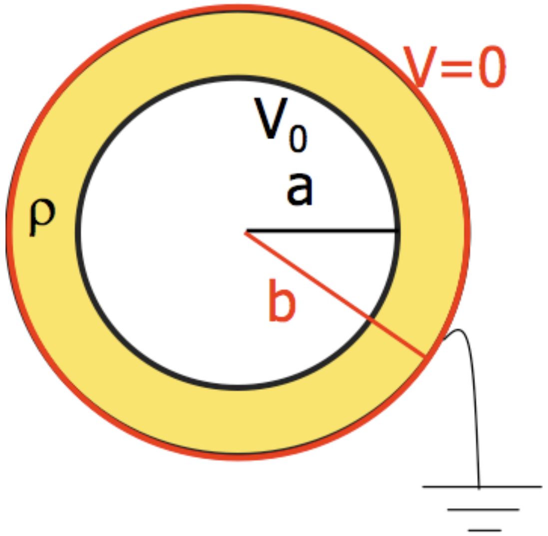
What is the conductivity?

From this derivation + definition of J we derive Ohm's law and read off the conductivity

$$\left\{\begin{array}{l}\vec{J} = nq\langle\vec{u}\rangle \\ m\langle\vec{u}\rangle = q\vec{E}\tau\end{array}\right\} \Rightarrow \vec{J} = nq\frac{q\vec{E}\tau}{m} = \sigma\vec{E} \Rightarrow \sigma = \frac{nq^2\tau}{m}$$

For multiple carriers:

$$\sigma = \sum_{k=1}^{N}\frac{n_k q_k^2 \tau_k}{m_k}$$



Application 1: Resistance of a spherical shell

2 concentric spheres; material in between has resistivity ρ

Difference in potential V→current

$\phi_{inner} = V$; $\phi_{outer} = 0$

Q: what is the resistance R?

Microscopic Ohm will hold: J = σE

Spherical symmetry→spherical potential:

Boundary conditions: $\phi(a) = V_0$ and $\phi(b) = 0$

$$\phi(r) = A + \frac{B}{r}$$

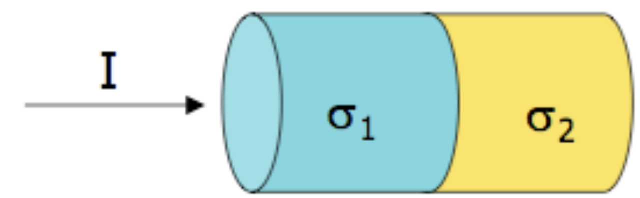
$$\rightarrow \phi(r) = V_0\left(\frac{ab}{b-a}\frac{1}{r} - \frac{a}{b-a}\right)$$

$$\vec{E} = -grad(\phi): \quad \vec{E}(r) = V_0\frac{ab}{b-a}\frac{1}{r^2}\hat{r} \Rightarrow J = \sigma V_0\frac{ab}{b-a}\frac{1}{r^2}$$

$$I = \int_{sphere}\vec{J}\cdot d\vec{A} = \vec{J}\cdot\vec{A} = \sigma V_0\frac{ab}{b-a}\frac{1}{r^2}4\pi r^2 = 4\pi\sigma V_0\frac{ab}{b-a} \Rightarrow R = \frac{\Delta V}{I} = \frac{V_0}{4\pi\sigma V_0\frac{ab}{b-a}} = \frac{b-a}{4\pi\sigma ab}$$

Application 2: What if σ is not constant?

Cylindrical wire made of 2 conductors with conductivity σ_1 and σ_2



What is the consequence?

Current flowing must be the same in the whole cylinder

$$I = A\sigma_1 E_1 = A\sigma_2 E_2$$

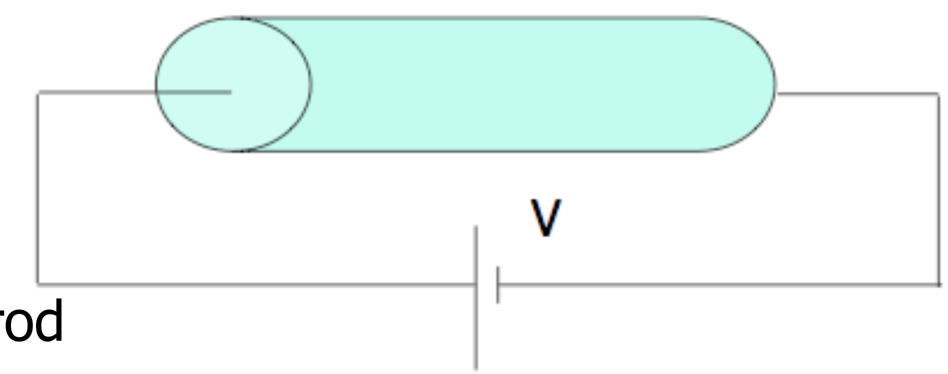
→Electric fields are different in the 2 regions

→E discontinuous→surface layer σ_q at the boundary

$$\sigma_q = \frac{E_{surface}}{4\pi} = \frac{E_2 - E_1}{4\pi} = \frac{I(\rho_2 - \rho_1)}{4\pi}$$

When conductivity changes there is the possibility that some charge accumulates somewhere. This is necessary to maintain steady flow.

One more application: Conducting Glass



Cylindrical glass rod of radius r and length L

Heat the center and conductivity becomes

$\sigma = \sigma_0 L^4/x^4$ with x =distance from center of the rod

Questions:

What is the resistance R of the rod?

When voltage V is applied, what is J(x) and what the steady E(x)?

In steady state, what is the volume charge density ρ(x)?

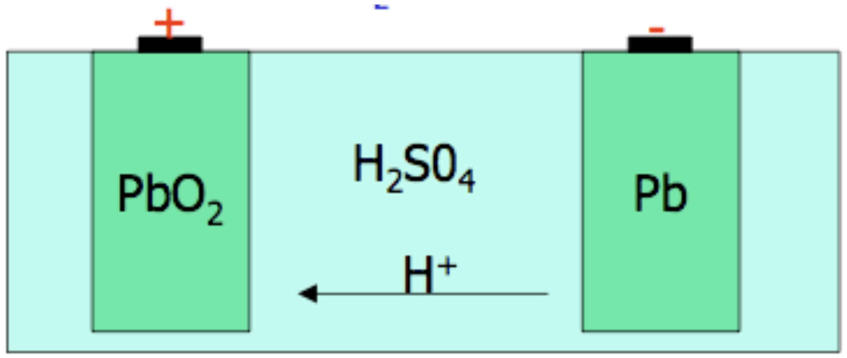
you can figure out!!

EMF: Electromotive force

- What is needed to have charges flowing in circuits?
  - Electric field created by a potential difference ΔV
  - Source of charges
- This is what the EMF provides
  - Note: EMF = Electromotive force but it's not a force!!!
- Example of EMF: battery
  - Device that maintains separation of charges between 2 electrodes
  - Current flows inside via electrochemical reactions that produce ΔV

Car Battery

- Two terminals (lead oxide $Pb0_2$ and porous lead Pb) in sulfuric acid (H_2S0_4)



- At Pb terminal: when immersed in acid, Pb provides free electrons:

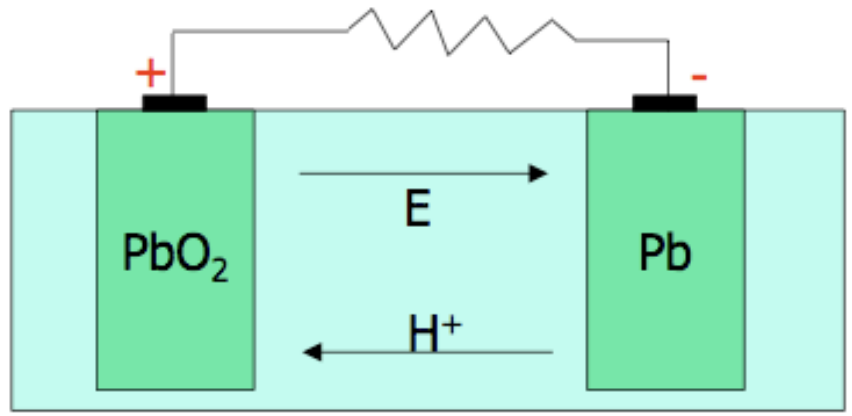
$$Pb + HSO_4^- \rightarrow PbSO_4 + H^+ + 2e^-$$

- At PbO_2 terminal: this reaction is energetically favored:

$$PbO_2 + 3H^+ + H_2SO_4 + 2e^- \rightarrow PbSO_4 + 2H_2O$$

- If it is possible for both e^- and H^+ to travel from one terminal to the other:

$$Pb + PbO_2 + H_2SO_4 \rightarrow 2PbSO_4 + 2H_2O + 4.4\ eV$$



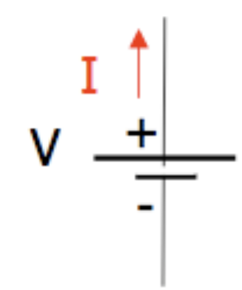
When terminals are not connected: no flow of e^-
E in battery does not allow flow of H^+ → inhibits reaction
When terminals are connected: electrons start flowing freely
Electric field is reduced → H^+ can flow → reaction occurs

EMF of battery: ϕ(+ terminal) – ϕ(- terminal): ΔV available to drive circuit

$$EMF = \int_{\text{- terminal}}^{\text{+ terminal}} E \cdot d\vec{s}$$

Convention

We indicate EMF with this symbol:
Long side: + terminal
Short side: - terminal
The current flows from + to –
Counterintuitive if you think about it in terms of electrons....

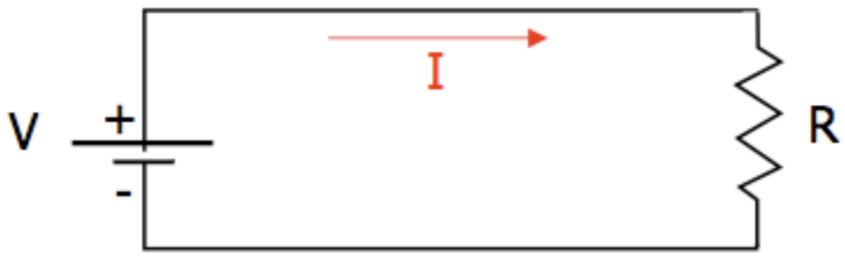


chemical reaction....

Kirchhoff's second rule

Close a battery on a resistor: simplest circuit!

How much current flows in the circuit? Ohm's law:



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

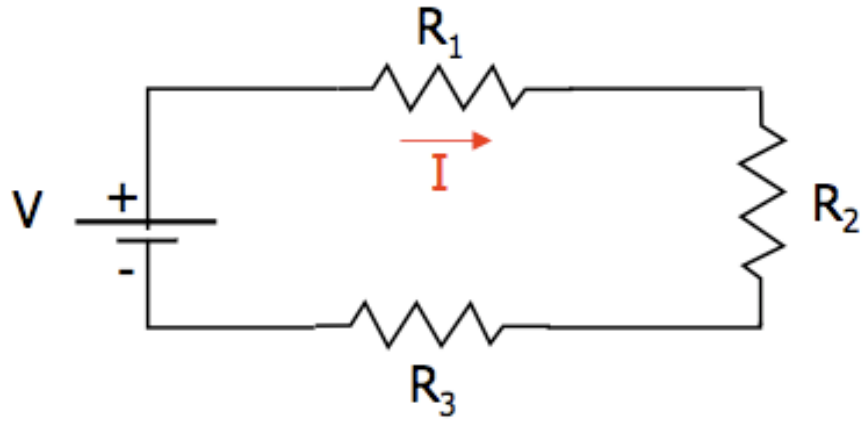
When the current flows in a resistor there is a voltage drop $\Delta V = -IR$

Kirchhoff's second law:
Around any closed loops, the sum of EMF and potential drops is 0

Equivalent to say that Electrostatic filed is **conservative** $\oint \vec{E} \cdot d\vec{s} = 0$

Solving circuits

If we have more than 1 resistor:



Solve the circuit:
determine currents and voltages everywhere

What we know:
- Current flowing in the circuit must be the same everywhere, or Q would accumulate somewhere
- Voltage drop in the i^{th} resistor: $\Delta V_i = -IR_i$
- Second Kirchhoff rule:

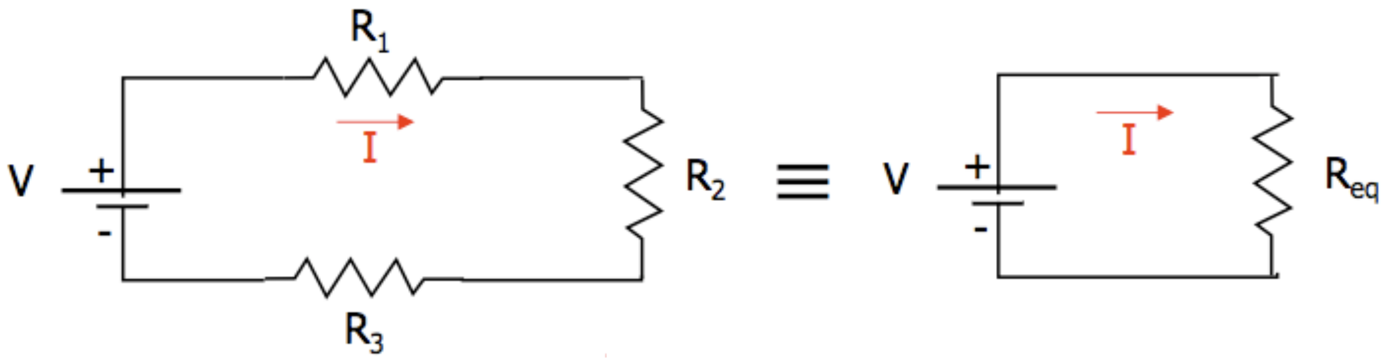
$$V + \sum_i V_i = V - I\sum_i R_i = 0 \Rightarrow I = \frac{V}{R_1 + R_2 + R_3}$$

Resistors in series

We implicitly derived an important result. We wrote:

$$V + \sum_i V_i = V - I \sum_i R_i = 0 \Rightarrow I = \frac{V}{R_1 + R_2 + R_3}$$

What does it mean? Same current flowing in these two circuits:



$$R_{eq} = \sum_i R_i$$

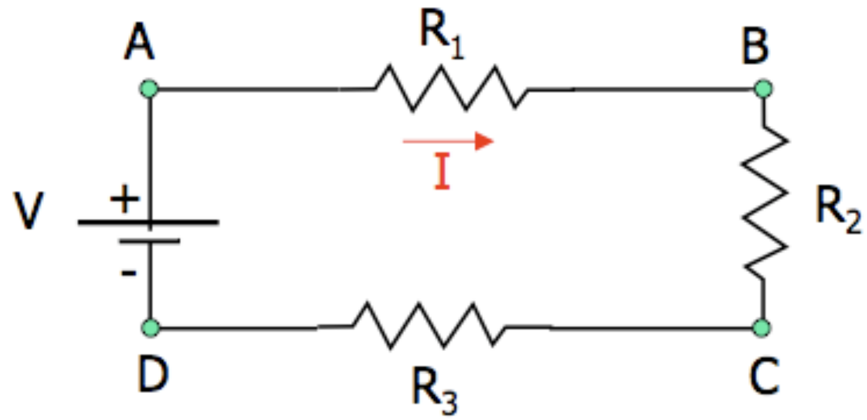
Solving circuits

Solve the circuit: determine currents and voltages everywhere

Calculate V_{AB}, V_{BC}, V_{CD}, V_{DA}, V_{AC}, ...

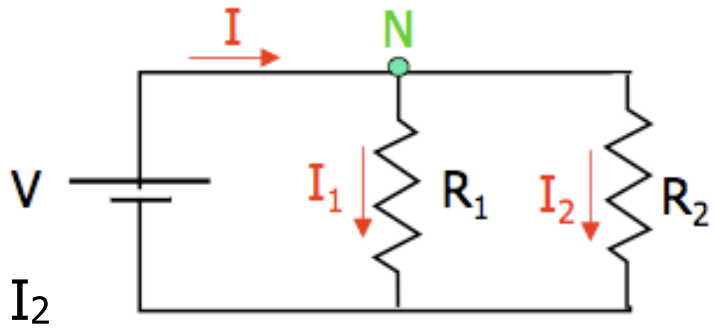
$$V_{AB} = IR_1 = \frac{VR_1}{R_1 + R_2 + R_3} \quad , \quad V_{BC} = IR_2 = \frac{VR_2}{R_1 + R_2 + R_3}$$

$$V_{CD} = IR_3 = \frac{VR_3}{R_1 + R_2 + R_3} \quad , \quad V_{AC} = I\left(R_1 + R_2\right) = \frac{V\left(R_1 + R_2\right)}{R_1 + R_2 + R_3}$$



Kirchhoff's first rule

Let's now connect resistors in parallel:

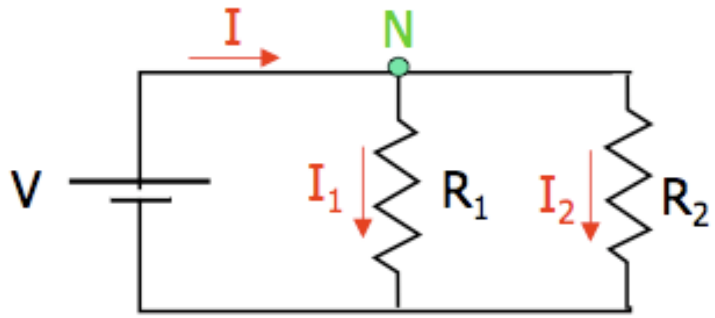


At the node N the current I divides up into 2 pieces: I_1 and I_2

Kirchhoff's first law:
At any node, sum of the currents in = sum of the currents out

In other words: there is no accumulation of charges in the circuit

Kirchhoff's first rule: application



Solve the circuit:

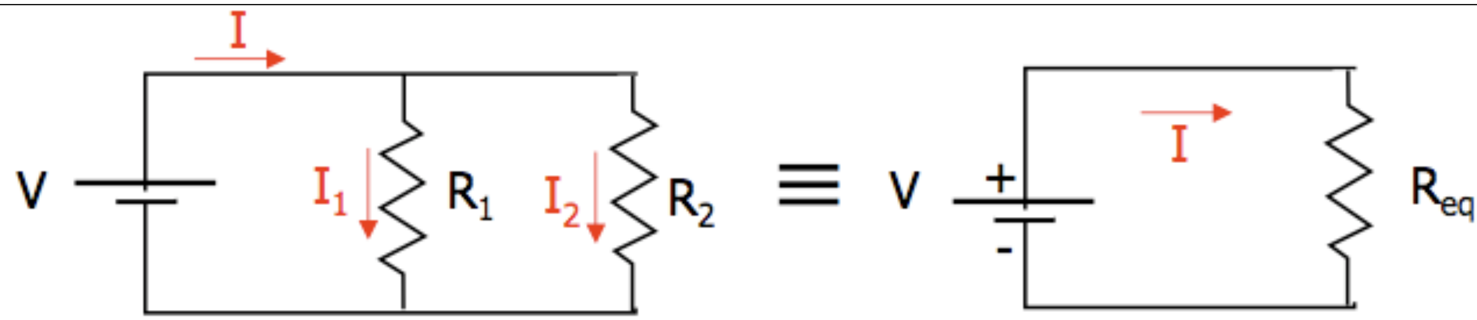
$$V = I_1 R_1 \Rightarrow I_1 = \frac{V}{R_1} \quad , \quad V = I_2 R_2 \Rightarrow I_2 = \frac{V}{R_2}$$

Apply Kirchhoff's first law: $I = I_1 + I_2$

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

Resistors in parallel

This is an important result. We have:



$$I = I_1 + I_2 = V\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \qquad\qquad I = V\frac{1}{R_{eq}}$$

What does it mean?

$$\frac{1}{R_{eq}} = \sum_i \frac{1}{R_i}$$

Resistors in parallel vs. in series

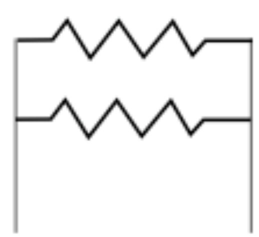
Resistors in series:

The current flowing is same→add resistors→make the path harder →I decreases→R_{eq} increases→R_{eq} larger than any single resistor

$$R_{eq} = \sum_i R_i$$

Resistors in parallel:

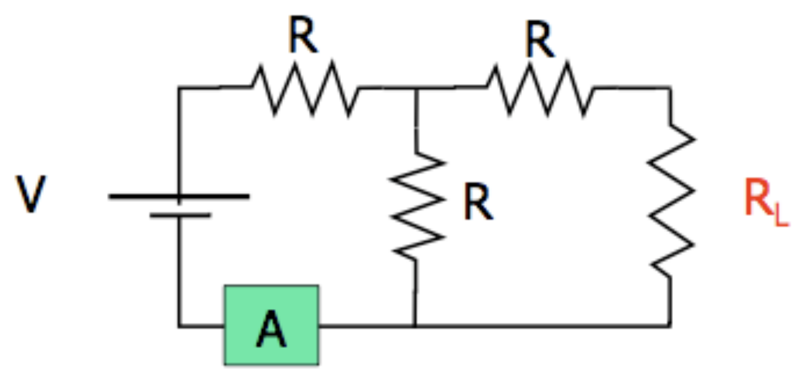
The current flows in many resistors→add resistors→make path easier →I increases→R_{eq} is smaller than any single resistor

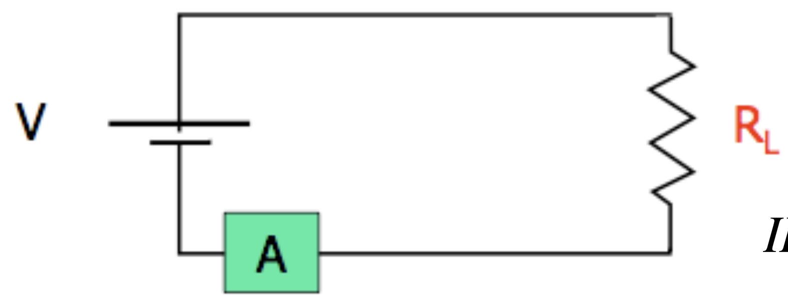


$$\frac{1}{R_{eq}} = \sum_i \frac{1}{R_i}$$

Solving circuits: application

Consider the circuit:





V = 1.5 V
R = 912 Ω
R_L = Adjustable resistor
Ammeter reading: 0.94 mA

R_L was chosen so that taking out the resistors R the same current will flow in the circuit.

What is R_L? R= 1580 Ω

$$IR + I_1 R = V \quad , \quad I_2 R + I_2 R_L - I_1 R = 0 \quad , \quad I = I_1 + I_2 \quad , \quad V = IR_L$$

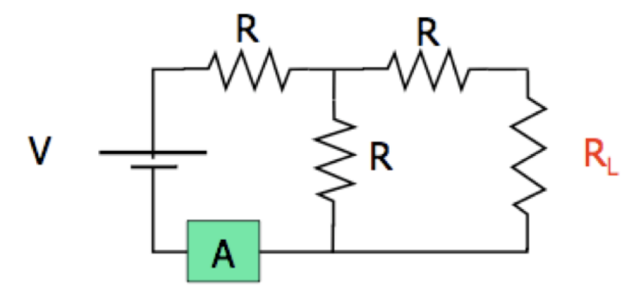
$$I + I_1 = \frac{V}{R} \quad , \quad I_2 = \frac{I_1 R}{R + R_L} \quad , \quad I = I_1 + \frac{I_1 R}{R + R_L} \Rightarrow I_1 = \frac{I}{1 + \frac{R}{R + R_L}}$$

$$I_1 = I\frac{R + R_L}{2R + R_L} \Rightarrow I_2 = I\frac{R}{2R + R_L} \Rightarrow I + I\frac{R + R_L}{2R + R_L} = \frac{V}{R} \Rightarrow I = \frac{V}{R}\left(\frac{2R + R_L}{3R + 2R_L}\right) \Rightarrow R_L = \frac{V}{I} = R\frac{3R + 2R_L}{2R + R_L}$$

$$\Rightarrow R_L\left(2R + R_L\right) = R\left(3R + 2R_L\right) \Rightarrow R_L^2 = 3R^2 \Rightarrow R_L = \sqrt{3}R = 1580\,\Omega$$

$$V = 1.5V \Rightarrow I = \frac{V}{R_L} = 0.94\ mA \text{ (as stated)}$$

We could also do the solution by reducing the circuit:



R and R_L in series ➙ R_{total} = R+R_L
In parallel with R➙

$$\frac{1}{R'_{total}} = \frac{1}{R_{total}} + \frac{1}{R} \Rightarrow R'_{total} = \frac{RR_{total}}{R+R_{total}} = \frac{R(R+R_L)}{2R+R_L}$$

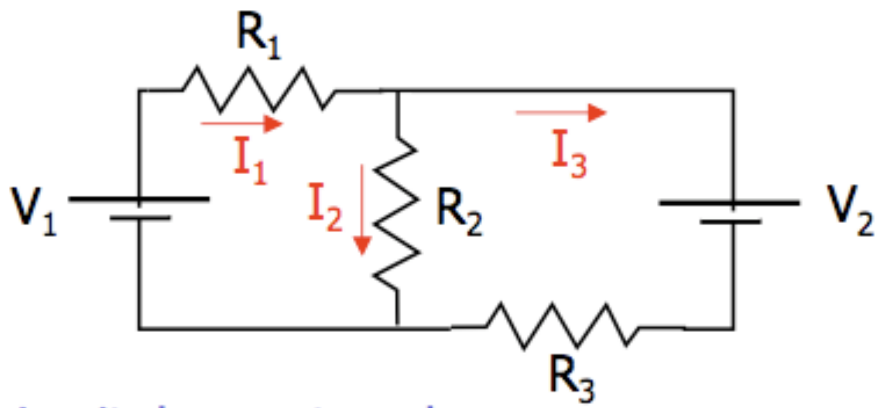
R'_{total} and R in series➙R_{eq}=R+ R'_{total} = R_L

$$R_L = \frac{R(R+R_L)}{2R+R_L} + R = R\left(\frac{3R+2R_L}{2R+R_L}\right)$$

which is the same equation as earlier!

Slightly harder circuits

How do we solve this?



Reducing the circuit does not work:
- Series and parallels won't work
- Because of second EMF

But Kirchhoff still holds so:
- Apply First Kirchhoff law to each node
- Apply Second Kirchhoff law to each loop

Solution:

Left loop:

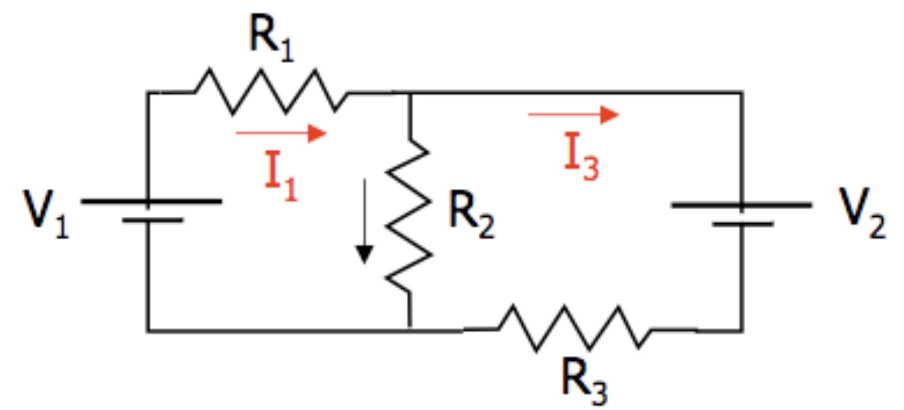
$V_1 - I_1R_1 - (I_1 - I_3)R_2 = 0$

Right loop:

$-V_2 - I_3R_3 - (I_3 - I_1)R_2 = 0$

Node:

$I(\text{in } R_2) = I_1 - I_3$



Solving the system:

$$I_1 = \frac{V_1R_3 + (V_1 - V_2)R_2}{R_1R_2 + R_1R_3 + R_2R_3}$$

$$I_3 = \frac{(V_1 - V_2)R_2 + V_2R_1}{R_1R_2 + R_1R_3 + R_2R_3}$$

How to go through a loop?

Assign current direction (arbitrary)

Choose a path (CW or CCW)

EMF: > 0 when - ➙ + and < 0 when + ➙ -

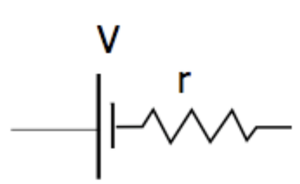
V drops on R when moving || to I direction

Internal resistance

When battery delivers current to a circuit there is a flow of current in the batteryitself

In previous example this current comes from flux of H^+ ions.

The chemical reaction will dissipate energy, battery gets hot, energy that could have gone the circuit is lost. This is equivalent to having a resistor r inside the battery:



Corollary: There is a max current the battery can generate: $I = V/r$

Power dissipated in resistors

When current flows in circuit, it moves charges through $\Delta V \rightarrow$ work
Work done to drive a charge dq through a potential difference V:

$$dW = Vdq$$

If work is done in time dt $\rightarrow$ the power dissipated is:

$$P = \frac{dW}{dt} = V\frac{dq}{dt} = VI$$

When Ohm's law holds:

$$P = VI = RI^2$$

Units: [P]=[Energy]/[time] cgs: erg/s SI: J/s

Power is important: it's what does work in a circuit: how much light is produced, how much heat, etc.

Dependence of R on T

Ohm's law tells us that V = RI
This is valid in any resistor
Does it mean that given a voltage I is constant over time?
Not necessarily!
- When I goes through R it dissipates power = RI^2
- R = f(T) $\rightarrow$ R is not constant while resistor heats up!

Increase T $\rightarrow$ increase R $\rightarrow$ decrease I

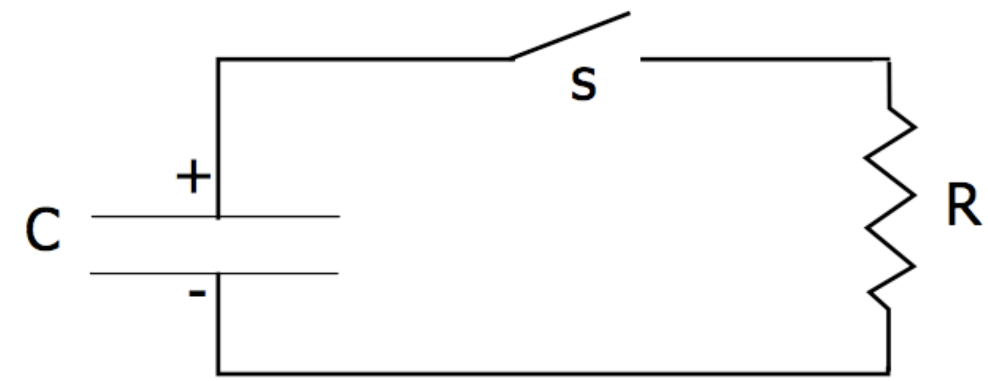
So far we considered only emf and R in the circuits. What if we use capacitors?

Capacitors in circuits ← **that is enough to do next homework**

A new way of looking at problems:

Until now: charges at rest or constant currents

When capacitors present: currents vary over time



Consider the following situation:

A capacitor C with charge $Q_0 \rightarrow V_0=Q_0/C$

A resistor R in series connected by switch s

What happens when switch s is closed?

Discharging capacitors: qualitative

Before switch s is closed:

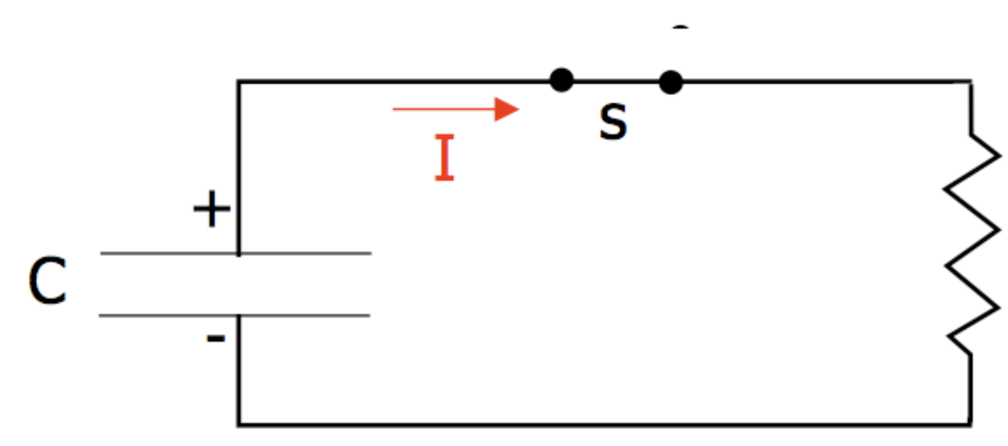
Difference in potential between C plates: V_0

No current circulating in the circuit (open)

After switch s is closed:

Difference in potential between capacitor plates will induce current I

As I flows, charge difference on capacitor decreases $\rightarrow$ VC decreases $\rightarrow$ I decreases over time



Discharging capacitors: quantitative

Apply second Kirchhoff's law:

EMF supplied by capacitor C: V=Q/C

Note: this is true at any moment in time ➙Q(t) ➙V(t)

Voltage drop across the resistor: -IR

$$\frac{Q}{C} - IR = 0$$

Not useful in this form since I=I(Q) ⟵ two changing variables!!

I=-dQ/dt (- sign because C is losing charge)

$$\frac{Q}{C} + R\frac{dQ}{dt} = 0$$

simple 1st order differential equation that can be solved by separation of variables

Easy integral yields to exponential decay of the charge:

$$\frac{Q}{C} + R\frac{dQ}{dt} = 0 \Rightarrow \frac{dQ}{Q} = -\frac{1}{RC}dt \Rightarrow \int_0^t \frac{dQ}{Q} = -\frac{1}{RC}\int_0^t dt$$

$$\Rightarrow \ln Q(t) - \ln Q(0) = -\frac{t}{RC} \Rightarrow \ln\frac{Q(t)}{Q_0} = -\frac{t}{RC}$$

$$\Rightarrow \frac{Q(t)}{Q_0} = e^{-\frac{t}{RC}} \Rightarrow Q(t) = Q_0 e^{-\frac{t}{RC}} = Q_0 e^{-\frac{t}{\tau}}$$

τ = RC is called "decay constant" of the circuit

Solution of RC circuit

Solution:

$$Q(t) = Q_0 e^{-\frac{t}{RC}}$$

Exponential decay of charge stored in capacitor

What are the units of RC?

cgs: [R]=statvolts /esu; [C]=esu/statvolt ➙[RC]=s

SI: [R]=V/A; [C]=C/V; A=C/s ➙[RC]=s

τ=RC is called "decay constant" of the circuit

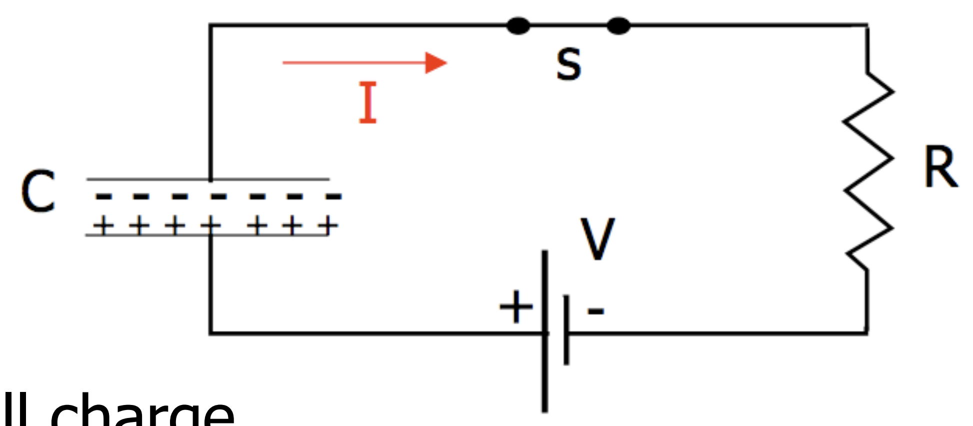
After a time RC, the charge decreased by 1/e wrt original value

Derive the current:

$$I(t) = -\frac{dQ}{dt} = Q_0 \frac{d}{dt}\left(e^{-\frac{t}{RC}}\right) = \frac{Q_0}{RC} e^{-\frac{t}{RC}}$$

Same exponential decay as for Q(t)

Charging capacitors



Now 3 elements in circuit: EMF, capacitor and resistor

Capacitor starts uncharged

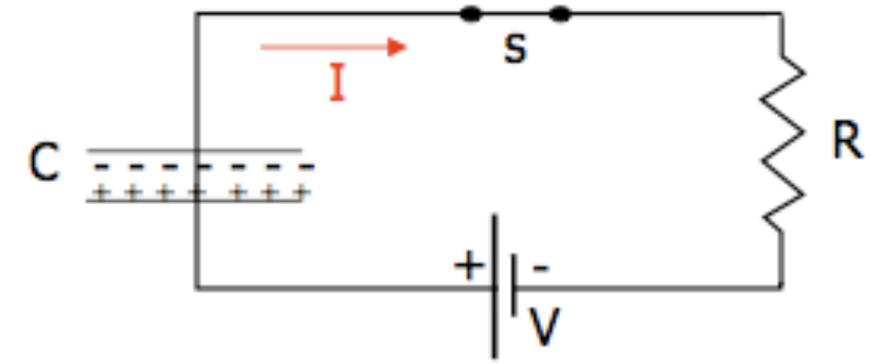
What happens when switch s is closed?

When s is closed, current will suddenly flow and C will charge

As C charges, E opposite to EMF builds up and slows down current

I(t) stops when V_C reaches V

Charging capacitor: solve the circuit $V - \frac{Q}{C} - IR = 0$



Solve using Kirchhoff's second law:

I(t) = +dQ/dt

+ because the capacitor is now charging!

First order differential equation

$$R\frac{dQ}{dt} + \frac{Q}{C} - V = 0$$

simple 1st order differential equation
that can be solved by separation of variables

Details of integration

$$R\frac{dQ}{dt} + \frac{Q}{C} - V = 0 \Rightarrow \frac{dQ}{dt} = -\frac{Q - VC}{RC}$$

$$Q' = Q - CV \Rightarrow \frac{dQ'}{dt} = -\frac{Q'}{RC} \Rightarrow \frac{dQ'}{Q'} = -\frac{dt}{RC}$$

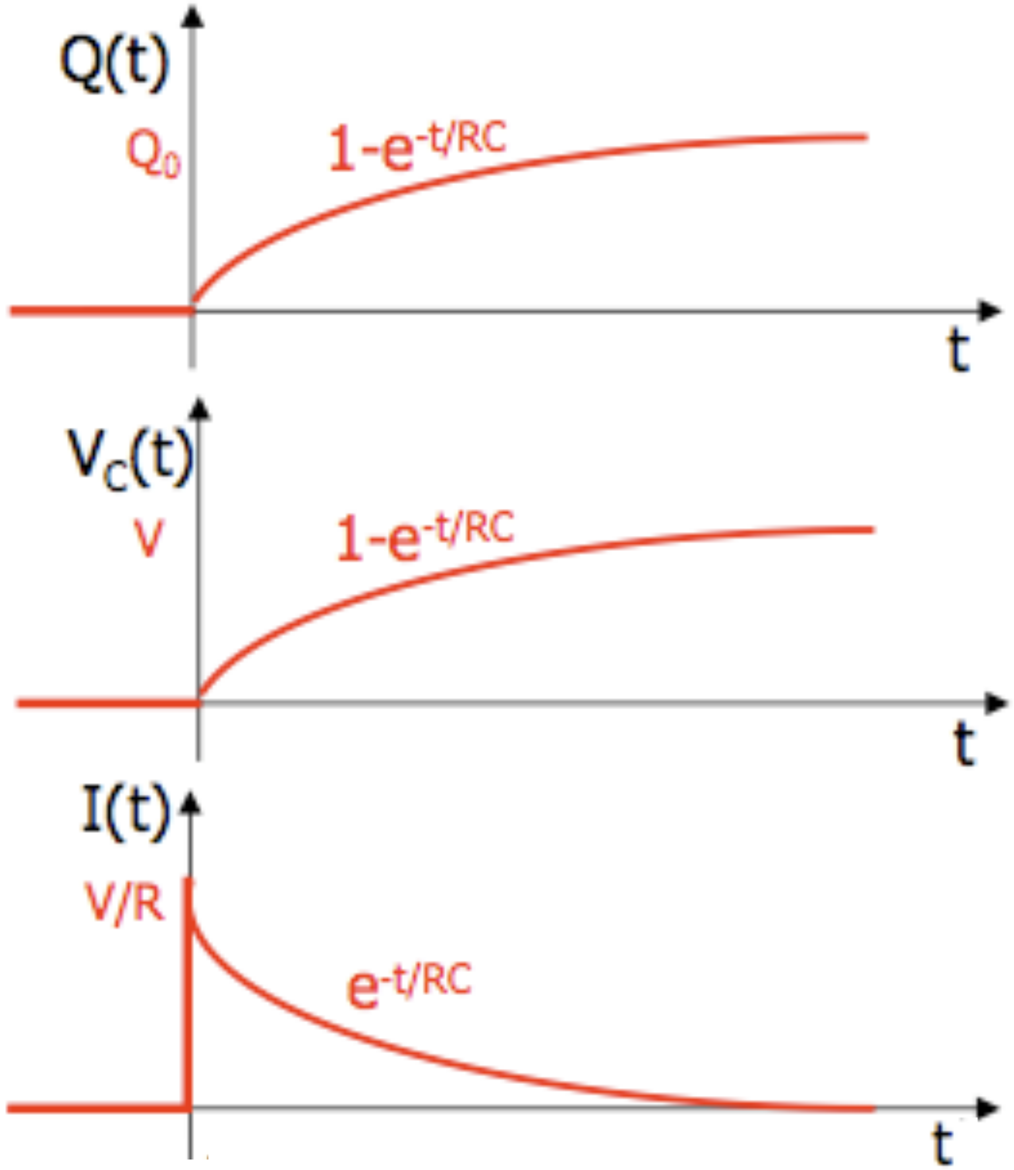
Integrating between t = 0 and t :

$$\int_{Q=0}^{Q=Q(t)} \frac{dQ'}{Q'} = -\frac{1}{RC}\int_{t=0}^{t=t} dt \Rightarrow \ln\frac{Q'(t)}{Q'(0)} = \ln\frac{Q(t) - CV}{-CV} = -\frac{t}{RC}$$

$$\Rightarrow \frac{Q(t) - CV}{-CV} = e^{-\frac{t}{RC}} \Rightarrow Q(t) = CV\left(1 - e^{-\frac{t}{RC}}\right)$$

Solution

Graphical solution



$$Q(t) = CV\left(1 - e^{-\frac{t}{RC}}\right)$$

$$V_C(t) = \frac{Q(t)}{C} = V\left(1 - e^{-\frac{t}{RC}}\right)$$

$$I(t) = \frac{dQ(t)}{dt} = \frac{V}{R} e^{-\frac{t}{RC}}$$

Important comments

Solution of RC circuit: $V_C(t) = \frac{Q(t)}{C} = V\left(1 - e^{-\frac{t}{RC}}\right)$ $\quad I(t) = \frac{dQ(t)}{dt} = \frac{V}{R} e^{-\frac{t}{RC}}$

Are Kirchhoff's laws valid at any moment in time?

$$V - \frac{Q}{C} - IR = V - V\left(1 - e^{-\frac{t}{RC}}\right) - R\frac{V}{R} e^{-\frac{t}{RC}} = 0 \qquad \text{OK!}$$

Asymptotic behavior of the capacitor:

At t = 0: I = V/R as if C were a short circuit

At t = infinity, I = 0 as if C were an open circuit

Conclusion: no need to solve the differential equation!

Solution is an exponential with time constant RC

Asymptotic behavior of C gives initial/final values for V(t) and I(t)

Time constant of RC circuit

Simple RC circuit with

V_{emf} = 3 V

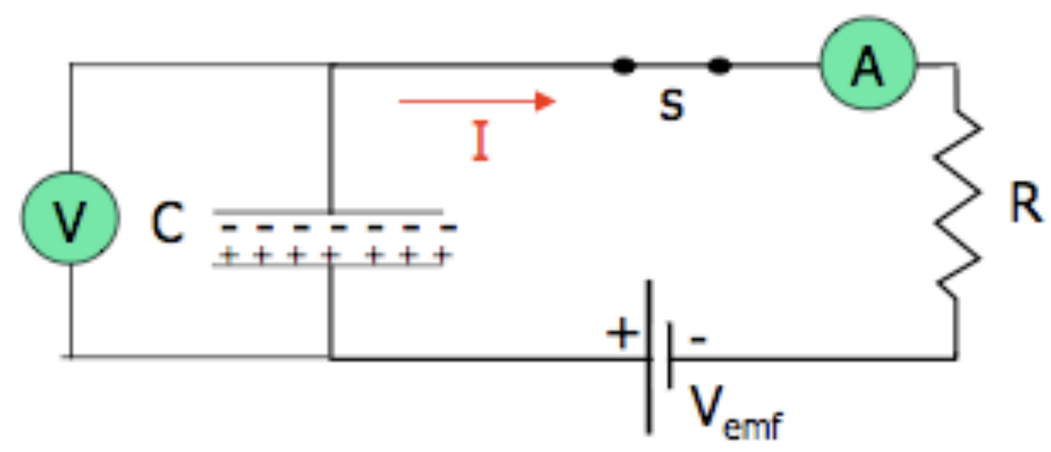
C = 1.3 F

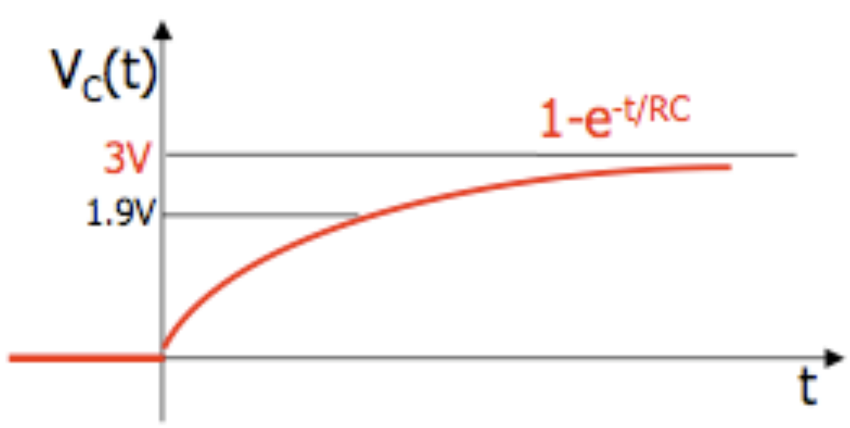
R = 11.7 Ω

Questions:

What are V_C and I?

Verify that time constant is RC: how long does it take to charge C ?





$$V_C(t) = V_{emf}\left(1 - e^{-\frac{t}{RC}}\right)$$

Note: R and C VERY large!

$$RC = 15.2\,s$$

If formula is correct →

V_C = V (1 - 1/e)=1.9 V when t=15.2

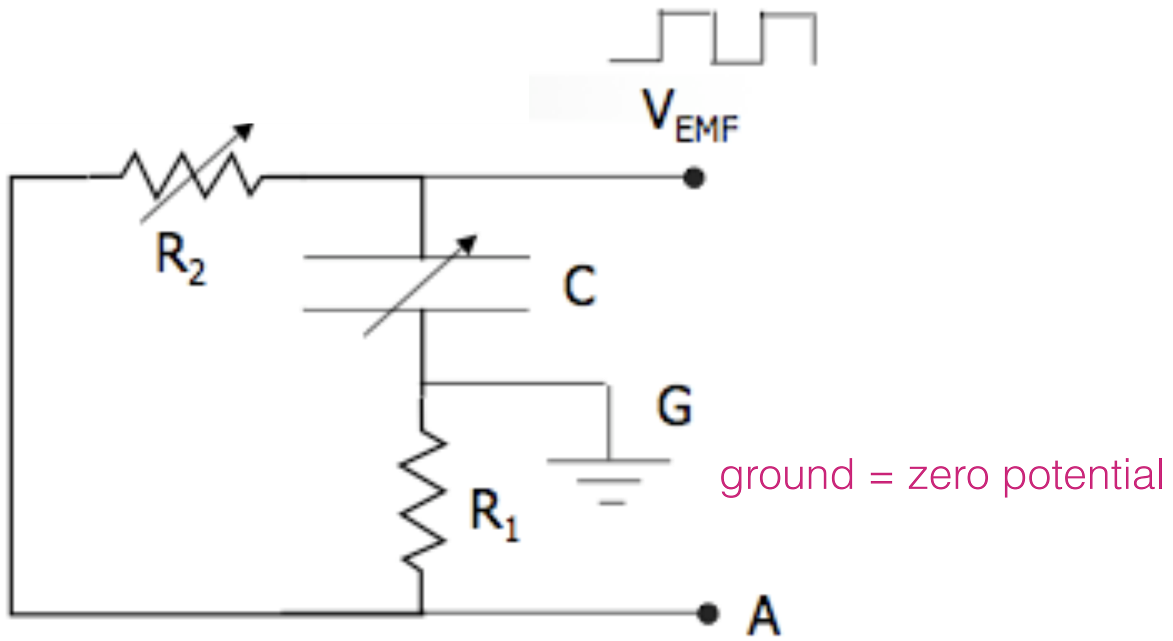
Verify time constant

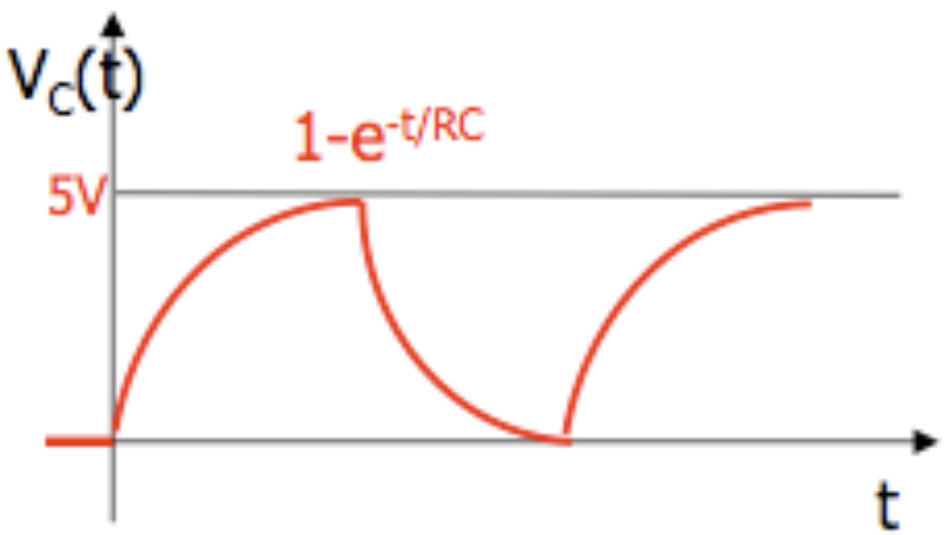
RC circuit with

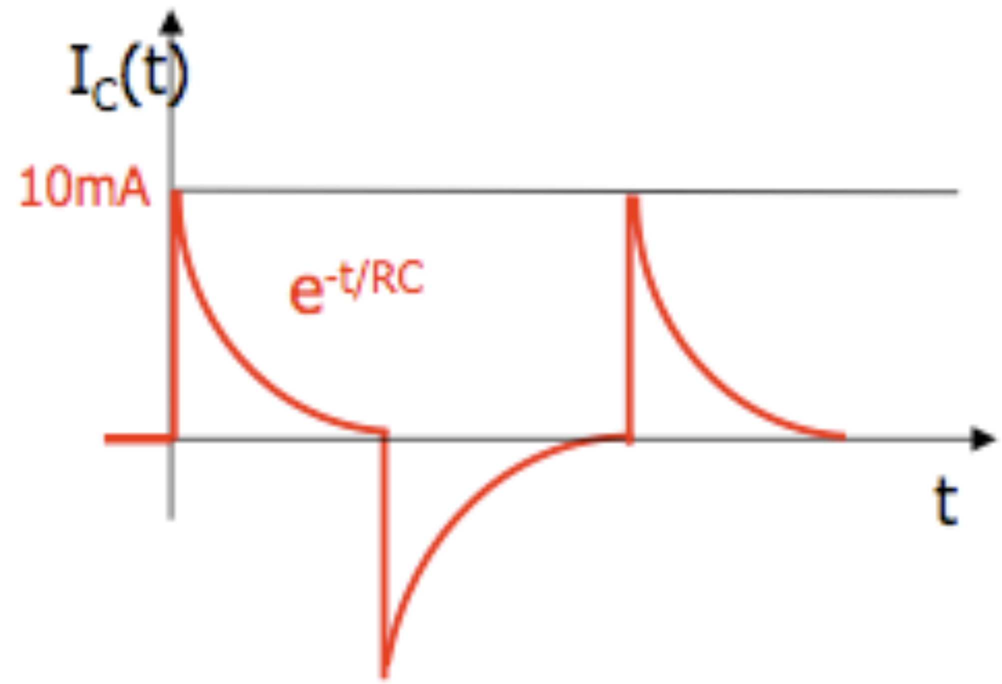
- V_{EMF}= squared 5 V pulses
- Variable C initially = 0.3 μF
- Variable R_2 initially = 400 Ω
- R_1= 100 Ω

Display on scope V_C and $I(R_1)$

- Verify RC=150 μs

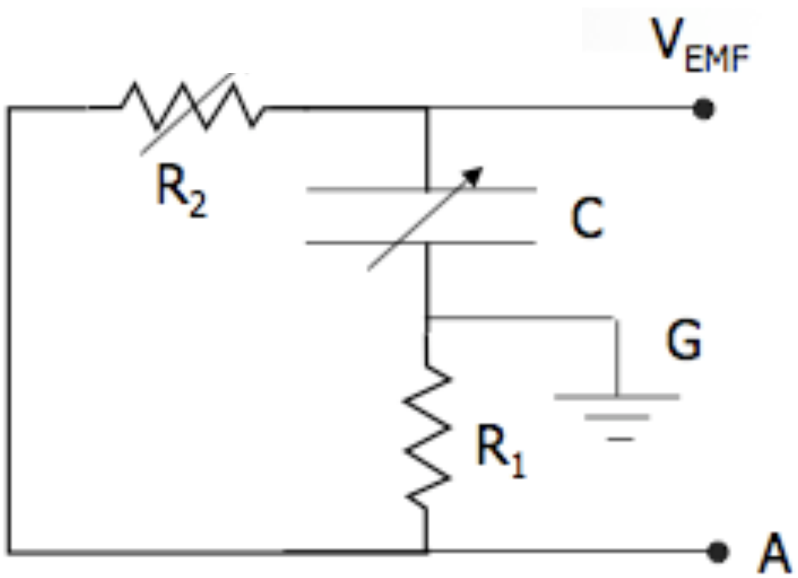






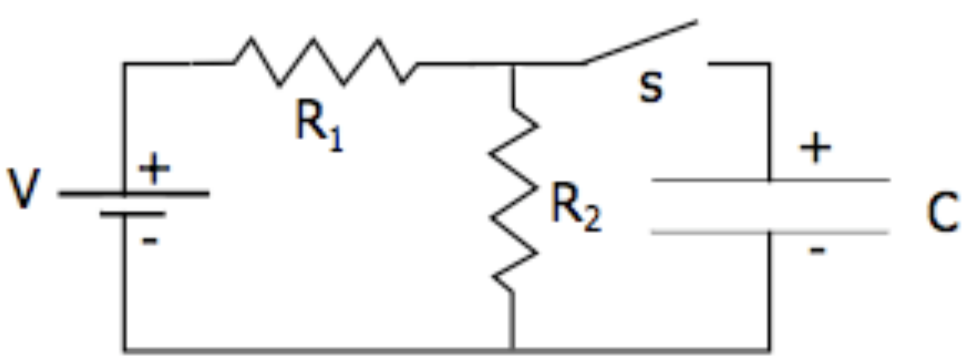
Verify time constant

- RC circuit with
  - V_{EMF}= squared 5 V pulses
  - Variable C initially = 0.3 μF
  - Variable R_2 initially = 400 Ω
  - R1= 100 Ω
- Let's now change the settings!
  - What happens when we double C?
    - $\tau_1=RC'=2RC=2\tau_0 \rightarrow V\ (I_{AG})$ raises (falls) twice as fast
  - How should we change R_2 to have the same effect?
    - $R' = 2R = 2(R_1+R_2') \rightarrow R_2'$: 400→900 Ω



More complicated RC circuits

- What if the RC circuit is more than just a series of R and C?
- Consider the following circuit:



- Calculate Q(t) on the capacitor
- Solution:
  - Kirchoff's laws will solve it: TEDIOUS!
  - Use Thevenin's Theorem!

Thevenin equivalence

Lots of words first - can be confusing

Thevenin's theorem:

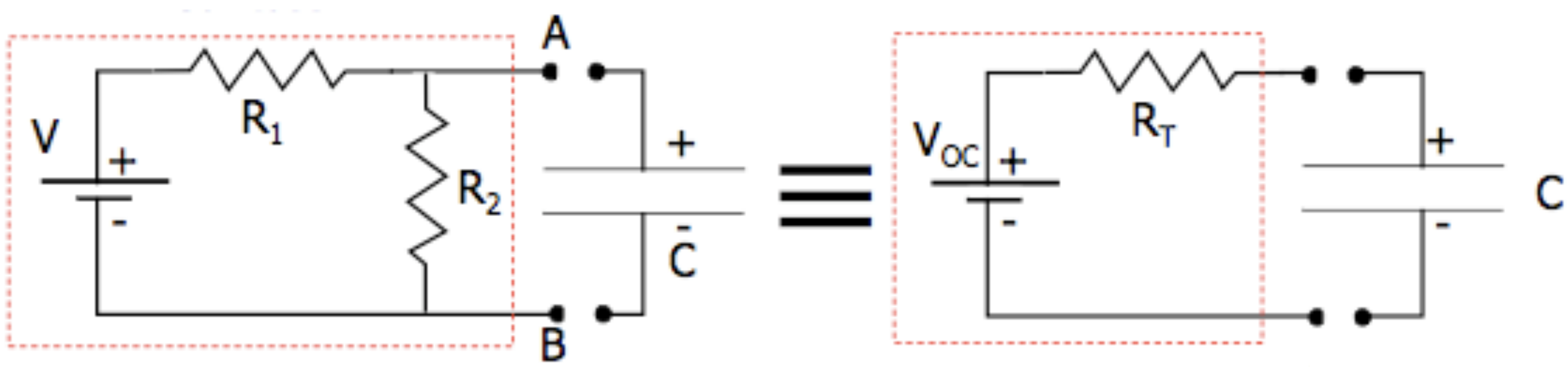
Any combination of resistors and EMFs with 2 terminals can be replaced with a series circuit of an emf V_{OC} and a resistor R_T where

V_{OC} is the open circuit voltage

$R_T=V_{OC}/I_{short}$ where I_{short} is the current going through the shorted terminals

or $R_T = R_{eq}$ with all the EMF shorted

In our case:



Once the circuit is reduced, the solution is known: $Q(t) = CV_{OC}\left(1 - e^{-\frac{t}{R_T C}}\right)$

Thevenin's demonstration

Prove that V_{OC} is the open circuit voltage

Since

$$Q(t) = CV_{OC}\left(1 - e^{-\frac{t}{R_T C}}\right) \Rightarrow V_C(t) = V_{OC}\left(1 - e^{-\frac{t}{R_T C}}\right)$$

So V_{OC} is the asymptotic V for the capacitor

Since for t → infinity, C → open circuit: V_{OC} = V of the open circuit

Prove that $R_T = V_{OC}/I_{short}$ with I_{short} = current through shorted terminals

There is only one current going through the reduced circuit

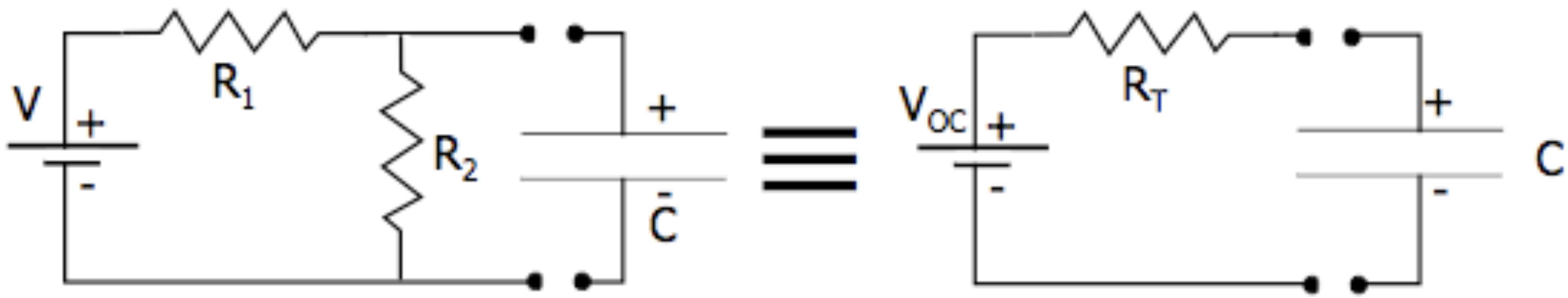
At t = 0, C behaves like a short → At t = 0 $I_{short} = V_{OC}/R_T$

→ $R_T=V_{OC}/I_{short}$

Solve the actual problem

clears up confusions as always.....

Calculate V_{OC} and $R_T = V_{OC}/I_{short}$ for our problem:



$$V_{OC} = \frac{V}{R_1 + R_2} R_2$$

Shorting C is makes R_2 irrelevant in the circuit:

$$I_{short} = \frac{V}{R_1}$$

$$\Rightarrow Q(t) = C\frac{VR_2}{R_1 + R_2}\left(1 - e^{-\frac{t(R_1+R_2)}{R_1R_2C}}\right)$$

$$R_{Thevenin} = \frac{V_{OC}}{I_{short}} = \frac{R_1R_2}{R_1 + R_2}$$

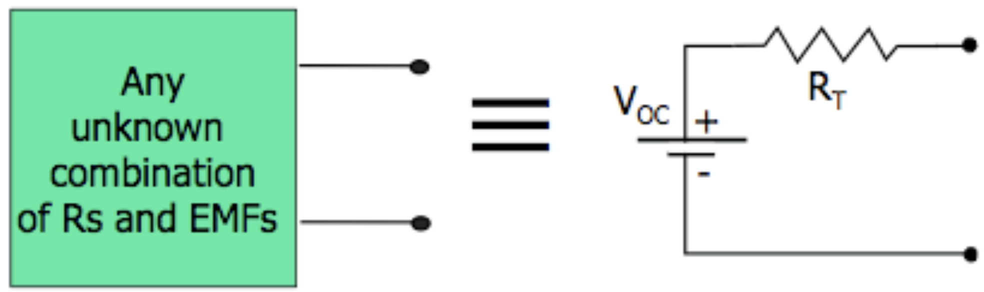
$$\Rightarrow I(t) = \frac{V}{R_1}e^{-\frac{t(R_1+R_2)}{R_1R_2C}}$$

Note : This is R1 || R2, same resistance we would get if we shorted EMF!

Thoughts on Thevenin

The importance of Thevenin:

When we have a messy system of resistors and EMFs, we can reduce it to a simple R+EMF in series just measuring I_{short} and V_{open}:



Careful:

Thevenin works only when the elements in the box follow Ohm's law, i.e. linear relation between V and I

Oscillating circuit

RC circuit with:

V_{EMF}= 1 kV

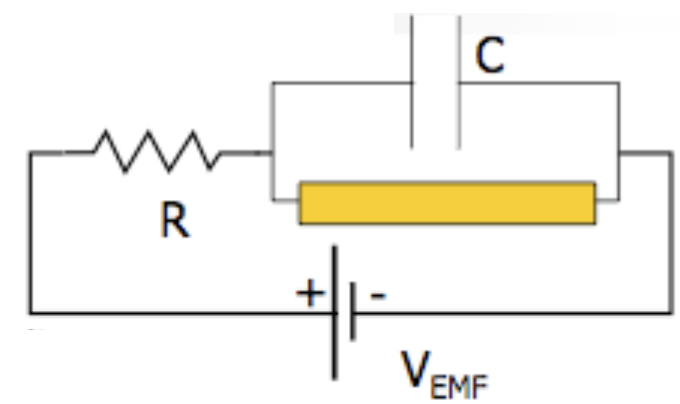
C = 0.1 μF

R = 2.5 MΩ

Fluorescent light in parallel with capacitor

(R_{FL} <<< R when current flows; ~infinite otherwise)

what is fluorescent flashing frequency?



what is fluorescent flashing frequency?

Initially the capacitor will start charging (no current thru lamp)

When V_C>certain value ~ 1kV current flows thru fluorescent light discharging the capacitor very quickly

The process will start again

$f \sim 1/\tau = 1/RC = 4$ Hz

Note: charging and discharging time constants are very different!

Charging: fluorescent light is ~ open circuit: $\tau_{charge} = RC$

Discharge: fluorescent light has a (very small) resistance R_{FL}

Thevenin: $R_T = R \,||\, R_{FL} \sim R_{FL}$

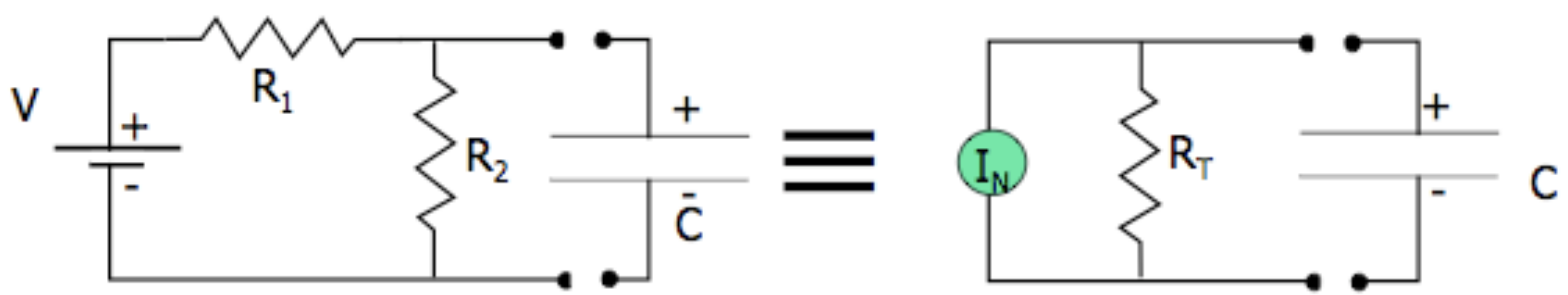
$\tau_{discharge} = R_T C \sim R_{FL} C << \tau_{charge}$

Norton's theorem

Any combination of resistors and EMFs with 2 terminals can be replaced with a parallel combination of a current generator I_N and a resistor R_T where

R_T is the equivalent resistance of the circuit with all the EMF shorted and all the current sources open (same as Thevenin!)

$I_N = V_{OC}/R_T$



$$R_T = R_1 || R_2 = \frac{R_1 R_2}{R_1 + R_2}$$

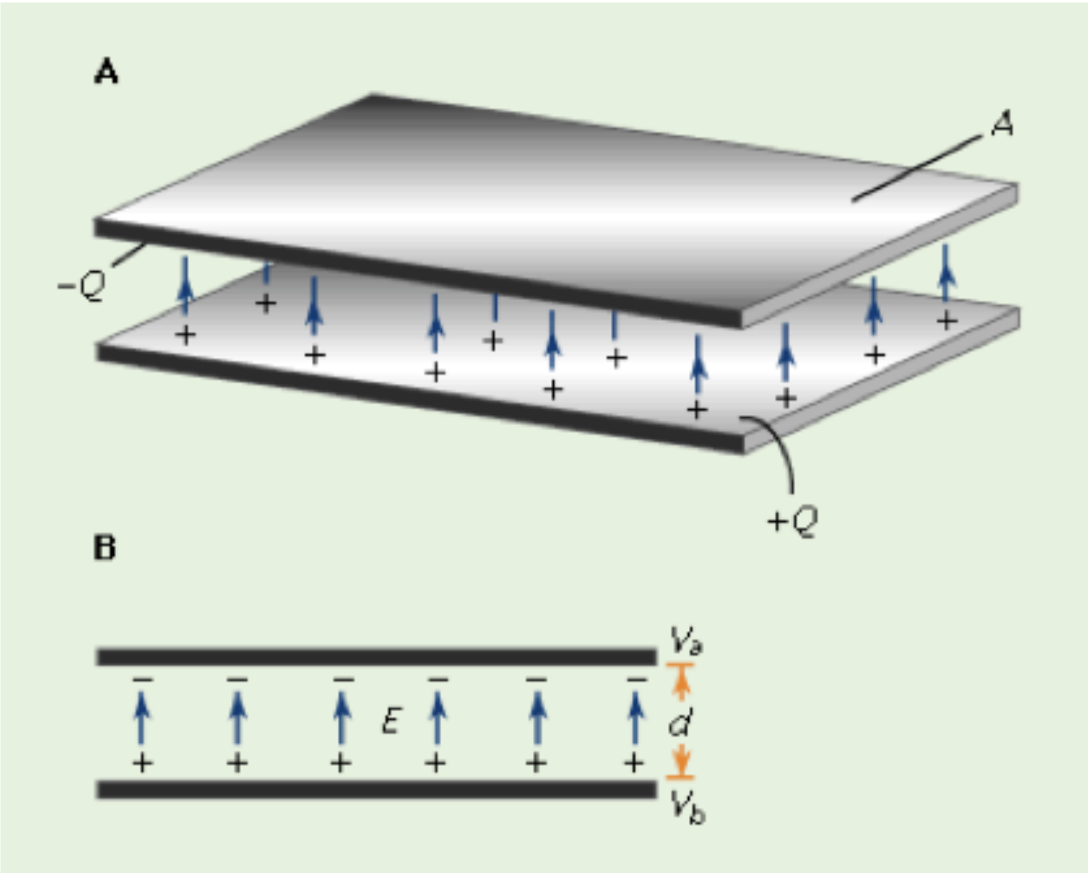
$$I_N = \frac{V_{OC}}{R_T} = \frac{VR_2}{R_1 + R_2}\frac{R_1 + R_2}{R_1 R_2} = \frac{V}{R_1}$$

Thevinin and Norton do not become clear until you do some problems!!!!!!

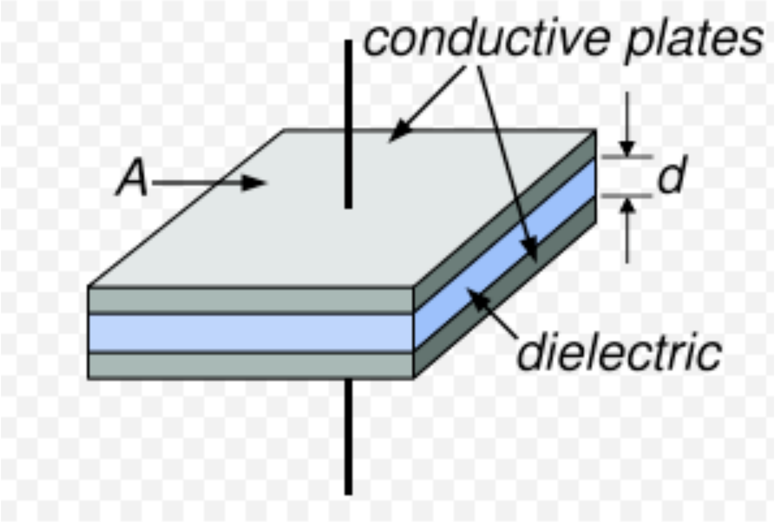
Some topics in Electric Fields in Matter

Dielectrics

We studied a capacitor earlier ---- 2 parallel insulated conducting plates with empty space between them, separation d, ±Q on plates and held at potential difference V. We defined capacitance = C = Q/V. For a parallel-plate capacitor we found $C=A/4\pi d$. We also solved some circuit problems with resistors and capacitors.



Suppose that we fill the empty space with a slab of material called a dielectric. We still have the relationship C = Q/V but the capacitance increases dramatically. This means that there is more charge on the plates for the same V.

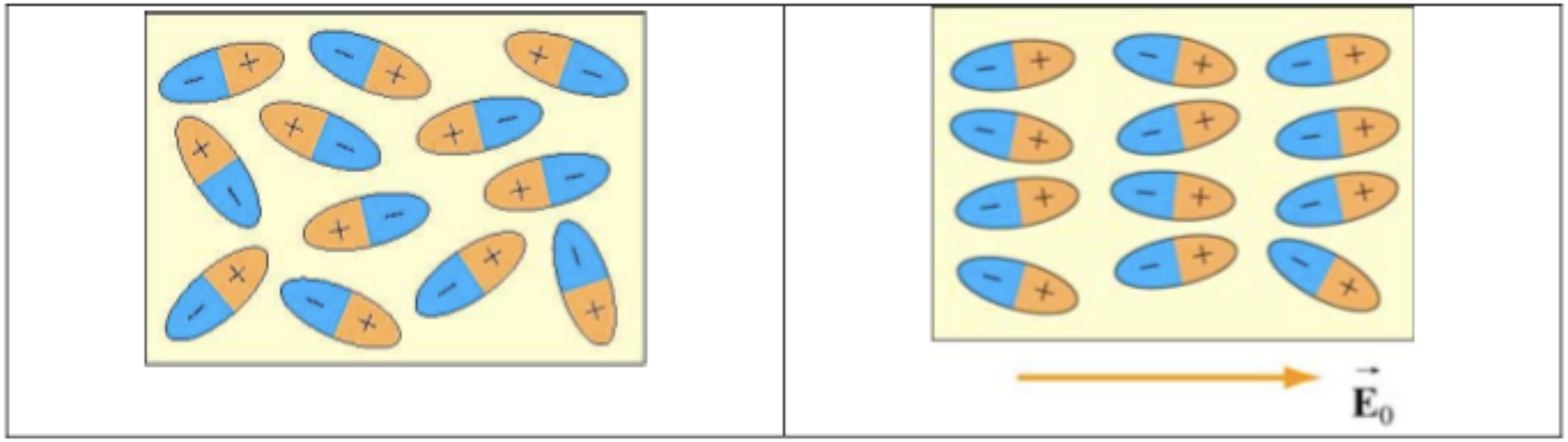


$$C = \kappa_e C_0$$

where κ_e is called the dielectric constant. In the Table below, we show some dielectric materials with their dielectric constant. Experiments indicate that all dielectric materials have $\kappa_e > 1$. Note that every dielectric material has a characteristic dielectric strength which is the maximum value of electric field before breakdown occurs and charges begin to flow.

Material	κ_e	Dielectric strength $\left(10^6\,\mathrm{V/m}\right)$
Air	1.00059	3
Paper	3.7	16
Glass	4–6	9
Water	80	–

The fact that capacitance increases in the presence of a dielectric can be explained from a molecular point of view. We shall show that κ_e is a measure of the dielectric response to an external electric field. There are two types of dielectrics. The first type is polar dielectrics, which are dielectrics that have permanent electric dipole moments. An example of this type of dielectric is water.

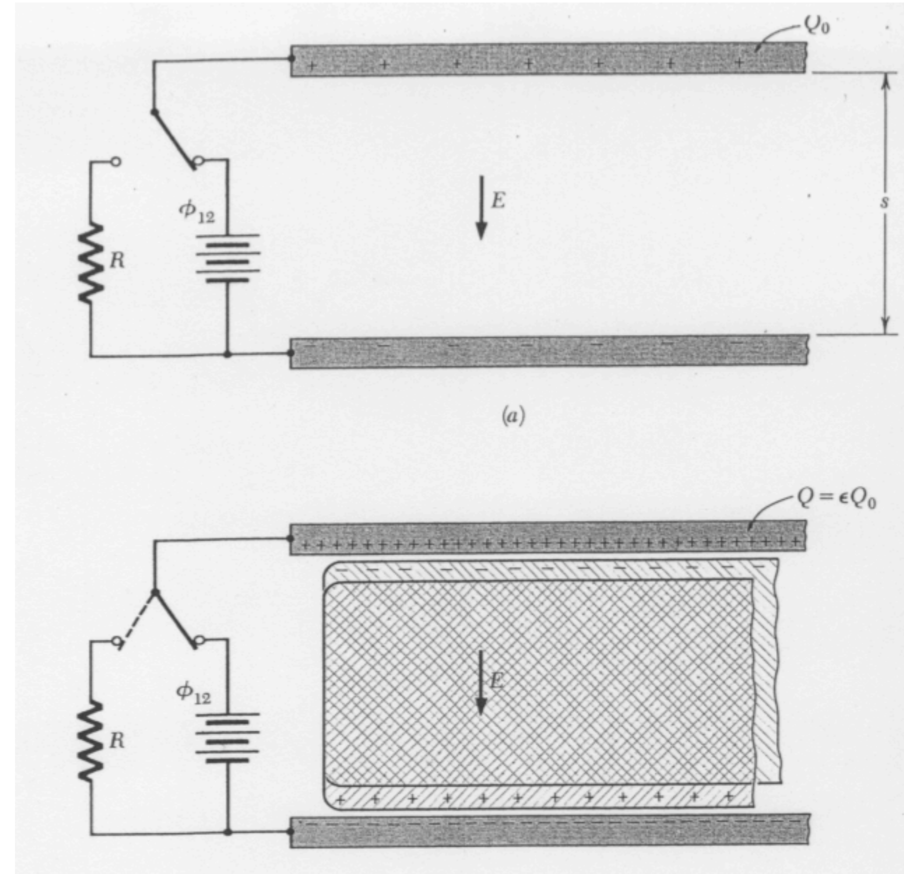


Orientations of polar molecules when (a) $\vec{\mathbf{E}}_0 = \vec{\mathbf{0}}$ and (b) $\vec{\mathbf{E}}_0 \neq 0$.

As depicted in Figure the orientation of polar molecules is random in the absence of an external field. When an external electric field $\vec{\mathbf{E}}_0$ is present, a torque is set up and causes the molecules to align with $\vec{\mathbf{E}}_0$. However, the alignment is not complete due to random thermal motion. The aligned molecules then generate an electric field that is opposite to the applied field but smaller in magnitude.

The second type of dielectrics is the non-polar dielectrics, which are dielectrics that do not possess permanent electric dipole moment. Electric dipole moments can be induced by placing the materials in an externally applied electric field.

As shown in figure:



The dielectric increases the charge on the plates of the capacitor. No dielectric

$$Q_0 = C_0\varphi_{12}$$

Add dielectric; charges rearrange inside dielectric(due to electric field). Total charge at top remains the same, i.e.,

$$Q_0 = Q - Q\text{dielectric}$$

$$Q = \varepsilon Q_0 > Q_0$$

Q is now amount of charge available to flow through any circuit.

Let us look at a microscopic or atomic level theory of what is going on inside the dielectric.

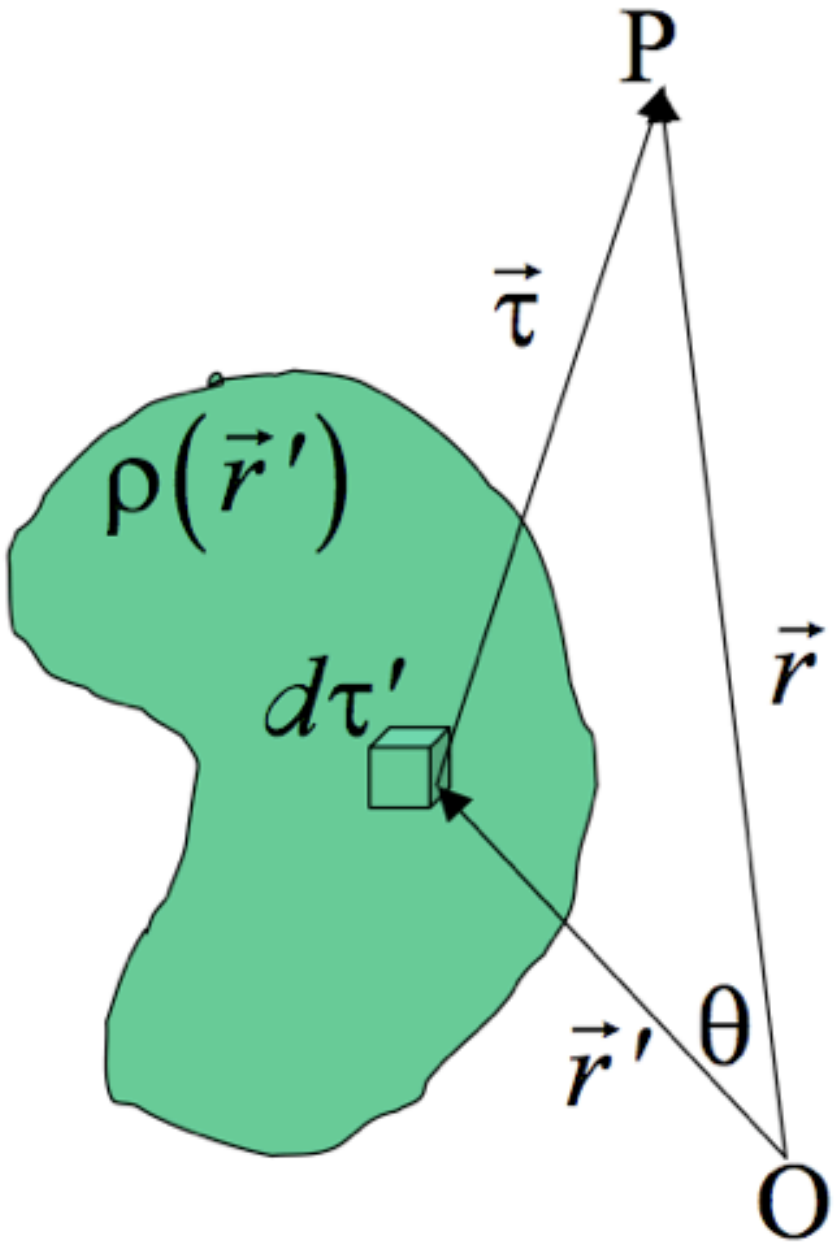
The Moments of a Charge Distribution

Multipole Expansions

higher level for the moment!

- Multipole expansions give yet another way of calculating V **outside** a region where charges reside. They provide an **approximate** answer where an exact answer is not needed
- One reason to use **Multipole expansions** is that they are often simpler to use than other techniques you have learned.
- **Multipole expansions** are just a particular form of **series expansion** in a **small parameter**. Let us proceed by looking from far away at a charge distribution:

Integrate only over volume (r')



$$V(r) = \int_{Vol} \frac{\rho(\vec{r}\,')\,d\tau'}{\tau}$$

Law of Cosines

$$\tau^2 = r^2 + (r')^2 - 2rr'\cos\theta$$

$$= r^2\left(1 + \left[\frac{r'}{r}\right]^2 - 2\frac{r'}{r}\cos\theta\right)$$

$$\Rightarrow \tau^2 = r^2(1+\varepsilon)$$

• It will be assumed that point P is far enough away from the charge distribution that $\varepsilon << 1$ can be regarded as a **small parameter**

$$\frac{r'}{r} << 1$$

$$\frac{1}{\tau}=\frac{1}{r}\frac{1}{\sqrt{1+\left[\frac{r'}{r}\right]^2-2\frac{r'}{r}\cos\theta}}=\frac{1}{r}\frac{1}{\sqrt{1+\varepsilon}}$$

$$\Rightarrow \frac{1}{\tau}=\frac{1}{r}\left(1-\frac{1}{2}\varepsilon+\frac{3}{8}\varepsilon^2-\frac{5}{16}\varepsilon^3+.....\right)$$

$$\frac{1}{\tau}=\frac{1}{r}\left(1-\frac{1}{2}\frac{r'}{r}\left(\frac{r'}{r}-2\cos\theta\right)+\frac{3}{8}\left(\frac{r'}{r}\right)^2\left(\frac{r'}{r}-2\cos\theta\right)^2-...\right)$$

$$=\frac{1}{r}\left(1+\left(\frac{r'}{r}\right)\cos\theta+\left(\frac{r'}{r}\right)^2\left(\frac{3}{2}\cos^2\theta-\frac{1}{2}\right)-......\right)$$

P_0 P_1 P_2 Legendre polynomials P_n

$$\Rightarrow \frac{1}{\tau}=\frac{1}{r}\sum_{n=0}^{\infty}\left(\frac{r'}{r}\right)^n P_n(\cos\theta)$$

- The electric potential **outside** the charge distribution at point P is thus given by:

$$V(r) = \int_{Vol} \frac{\rho(\vec{r}\,')\,d\tau'}{\tau}$$

$$V(r) = \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \int_{Vol} \rho(\vec{r}\,')(r')^n P_n(\cos\theta)\,d\tau'$$

$$V(r) = \frac{1}{r}\int_{Vol} \rho(\vec{r}\,')\,d\tau' + \frac{1}{r^2}\int_{Vol} \rho(\vec{r}\,')\,r'\cos\theta\,d\tau' + \frac{1}{r^3}\int_{Vol} \rho(\vec{r}\,')(r')^2\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)d\tau' + \dots$$

- The term proportional to $1/r$ is the **monopole** term; the term proportional to $1/r^2$ is the **dipole** term; the term proportional to $1/r^3$ is the **quadrupole** term and so on.

- If the **total charge** Q of the charge distribution is **non-zero**, the electrostatic field at large distances is dominated by the **monopole** term:

$$V(r) = \frac{1}{r}\int_{Vol} \rho(\vec{r}\,')d\tau' = \frac{Q}{r}$$

$$\Rightarrow \vec{E} = -\nabla V = -Q\nabla\left(\frac{1}{r}\right) = \frac{Q}{r^2}\hat{r}$$

- If the **total charge** Q of the charge distribution is **zero**, the monopole field is zero and the field at large distances is dominated by the **dipole** term:

$$V(r) = \frac{1}{r^2}\int_{Vol} \rho(\vec{r}\,')r'\cos\theta d\tau'$$

$$r'\cos\theta = \hat{r}\cdot\vec{r}\,'$$

$$V(r) = \frac{\hat{r}\cdot}{r^2}\int_{Vol}\rho(\vec{r}\,')\vec{r}\,'d\tau' = \frac{\hat{r}\cdot\vec{p}}{r^2}$$

In the above expression, $\vec{p}$ is the dipole moment of the charge distribution

$$\vec{p} = \int_{Vol}\rho(\vec{r}\,')\vec{r}\,'d\tau' = \int_{Vol}\rho(\vec{r})\vec{r}d\tau$$

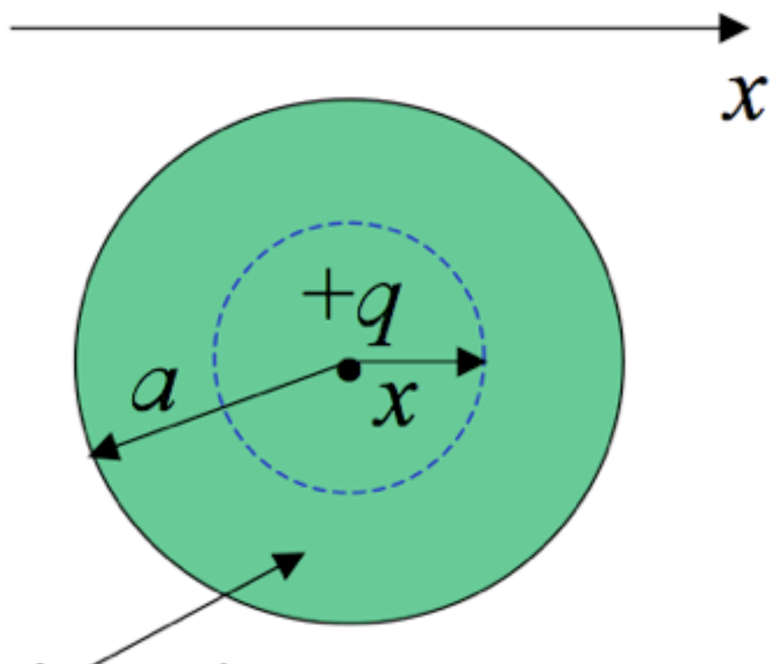
The electric field components can then be computed:

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

$$E_\theta = -\frac{1}{r}\frac{\partial V}{\partial\theta} = \frac{p\sin\theta}{r^3} \quad , \quad E_\varphi = 0$$

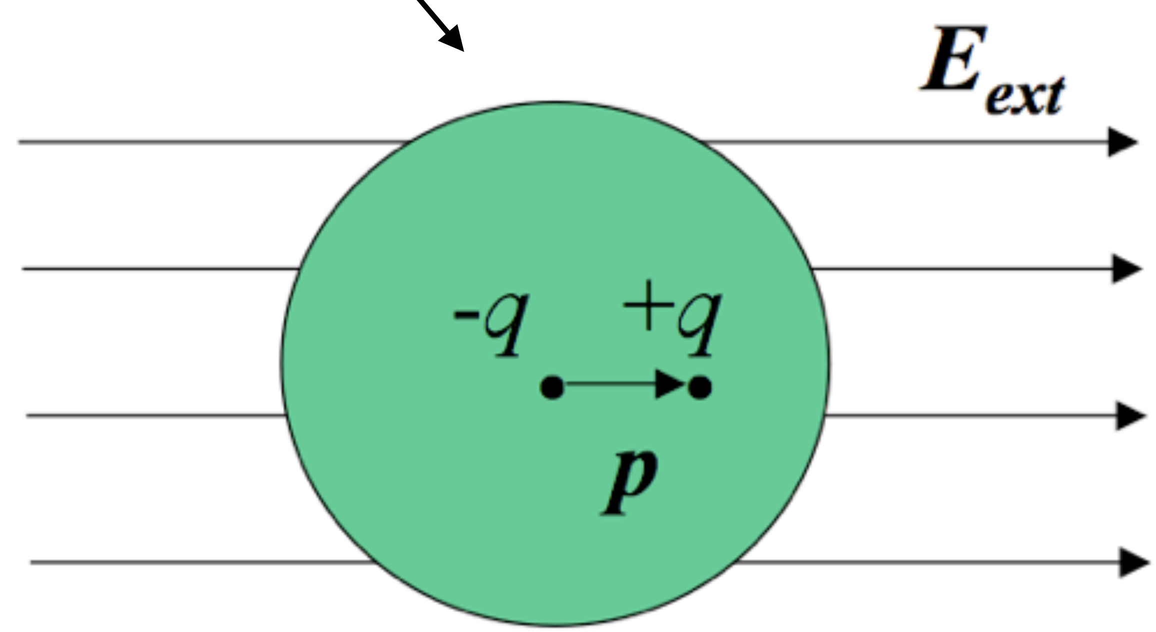
Permanent and Induced Dipole Moments

- Many atoms or molecules have **permanent** electric dipole moments, or in **external E**-fields they may have **induced** electric dipole moments. Consider the H atom:



Negative electron cloud (charge $-q$)

$$\boldsymbol{E}_e(x) = -\frac{q\boldsymbol{x}}{a^3} = -\frac{\boldsymbol{p}}{a^3}$$



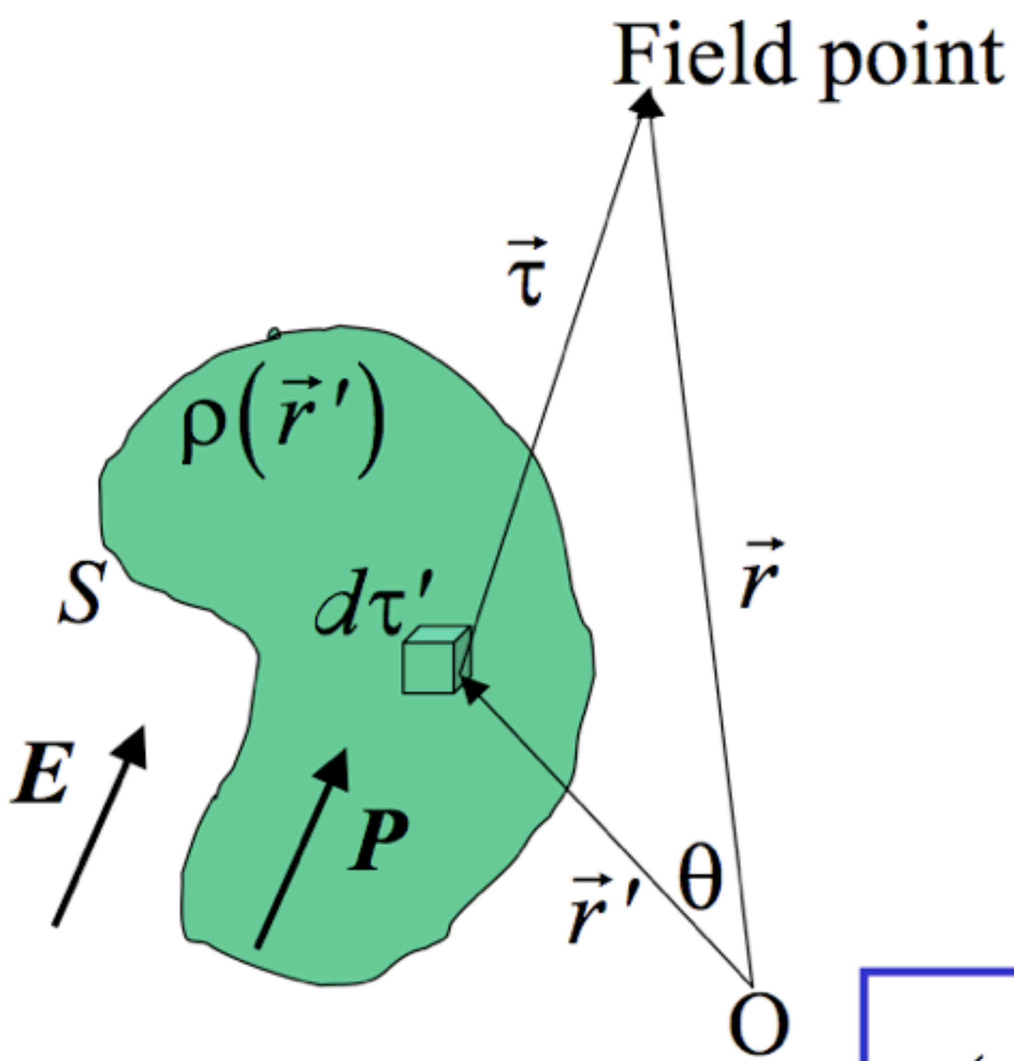
- Electron is effectively at center. Equilibrium when:

$$\boldsymbol{E}_e(x) = -\frac{\boldsymbol{p}}{a^3} = -\boldsymbol{E}_{ext}$$

Dielectric Materials

- Many **non-conducting solids** have permanent dipole moments, or become **polarized** when immersed in an external electric field. Materials such as these are known as **dielectrics**:
  - Normally, the **dipole moment** is zero on large scales since atomic dipoles are oriented in random directions
  - Immersion of a dielectric in an electric field **polarizes** atoms and tends to **align** the atomic dipoles
  - The **induced polarization** is defined in terms of a **dipole moment per unit volume** in the direction of the applied field (as in the example of the H atom)

Polarization and Bound Charge



$$\boldsymbol{P} = \frac{d\boldsymbol{p}}{d\tau} \longrightarrow$$ **Polarization** or dipole moment per unit volume

$$\boldsymbol{p} = \int_{Vol} \boldsymbol{P}(\boldsymbol{r}')\, d\tau'$$

→ Dipole moment charge distribution

$$V(r) = \int_{Vol} \frac{\hat{\boldsymbol{\tau}} \bullet [\boldsymbol{P}(\boldsymbol{r}')\, d\tau']}{\tau^2}$$

We have not used any monopole expansion here

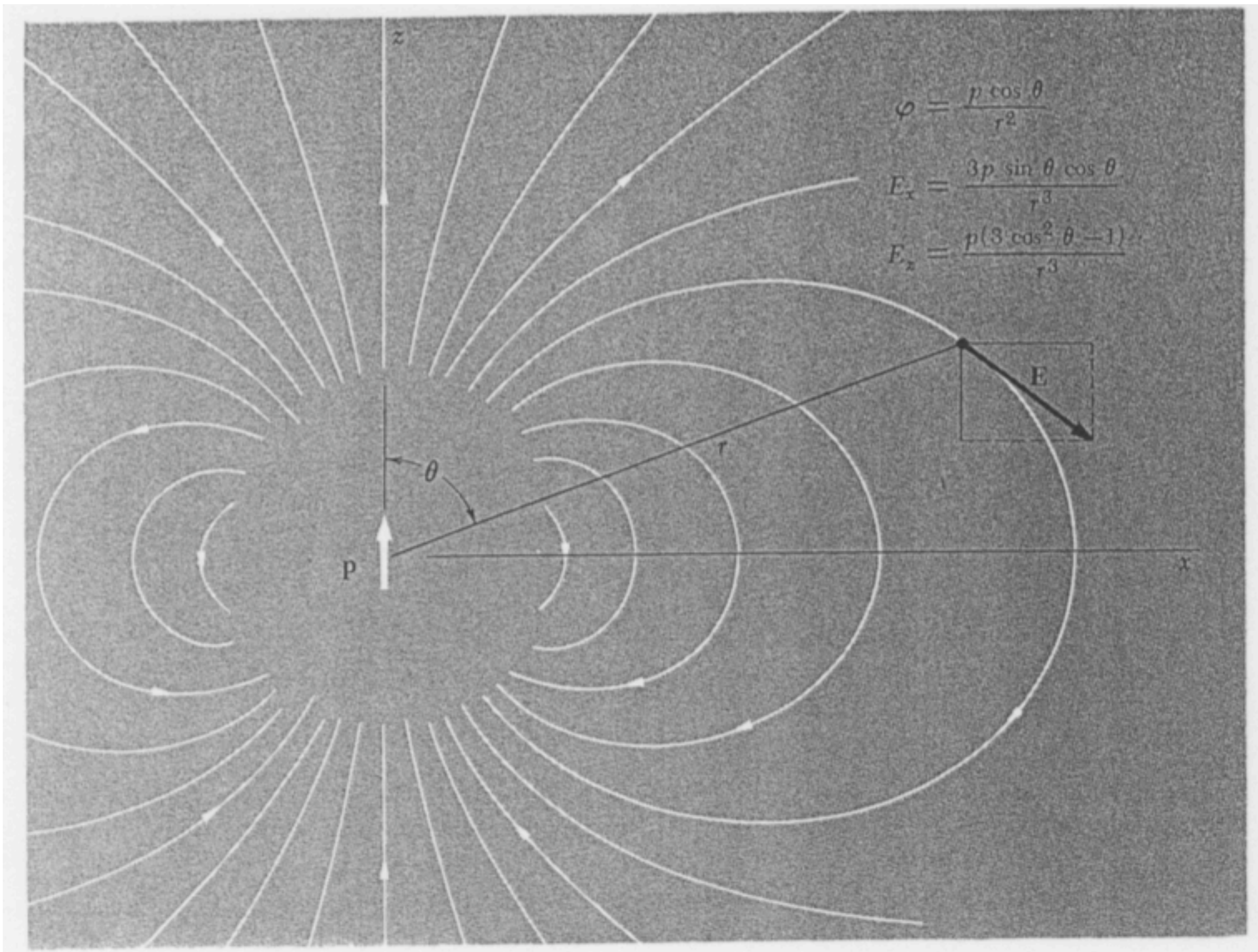
Enough for all the technical details which we are not going to use at this level anyway!!

Let us look at the dipole again:

We found earlier that the "dipole" contribution to the potential is given by

$$V(\vec{r}) = \frac{p\cos\theta}{r^2}$$

with r and Θ defined as in the diagram below:



We then have:

$$\cos\theta = \frac{z}{r} \quad , \quad r^2 = x^2 + z^2 \Rightarrow V(\vec{r}) = \frac{pz}{\left(x^2 + z^2\right)^{3/2}}$$

The Cartesian components of the electric field are readily derived:

$$E_x = \left(-\nabla V(\vec{r})\right)_x = \frac{\partial V}{\partial x} = \frac{3pxz}{\left(x^2 + z^2\right)^{5/2}} = \frac{3pr\sin\theta r\cos\theta}{r^5} = \frac{3p\sin\theta\cos\theta}{r^3}$$

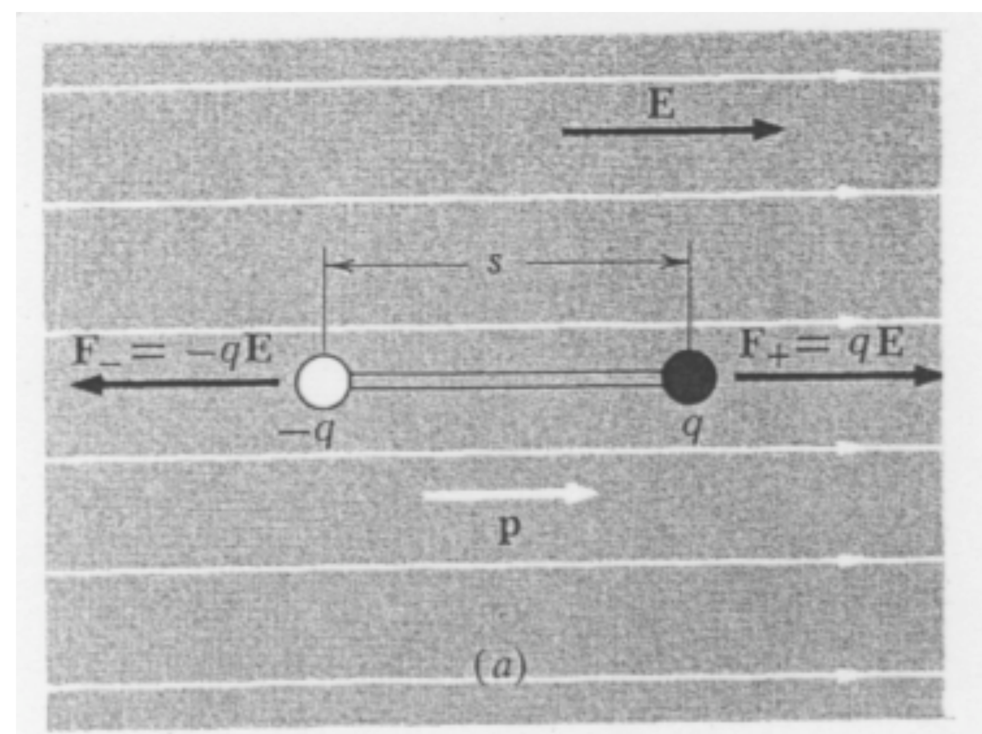
$$E_z = \left(-\nabla V(\vec{r})\right)_z = \frac{\partial V}{\partial z} = p\left[\frac{3z^2}{\left(x^2 + z^2\right)^{5/2}} - \frac{1}{\left(x^2 + z^2\right)^{3/2}}\right] = \frac{p\left(3\cos^2\theta - 1\right)}{r^3}$$

As we saw earlier, the plane-polar components of the electric field are:

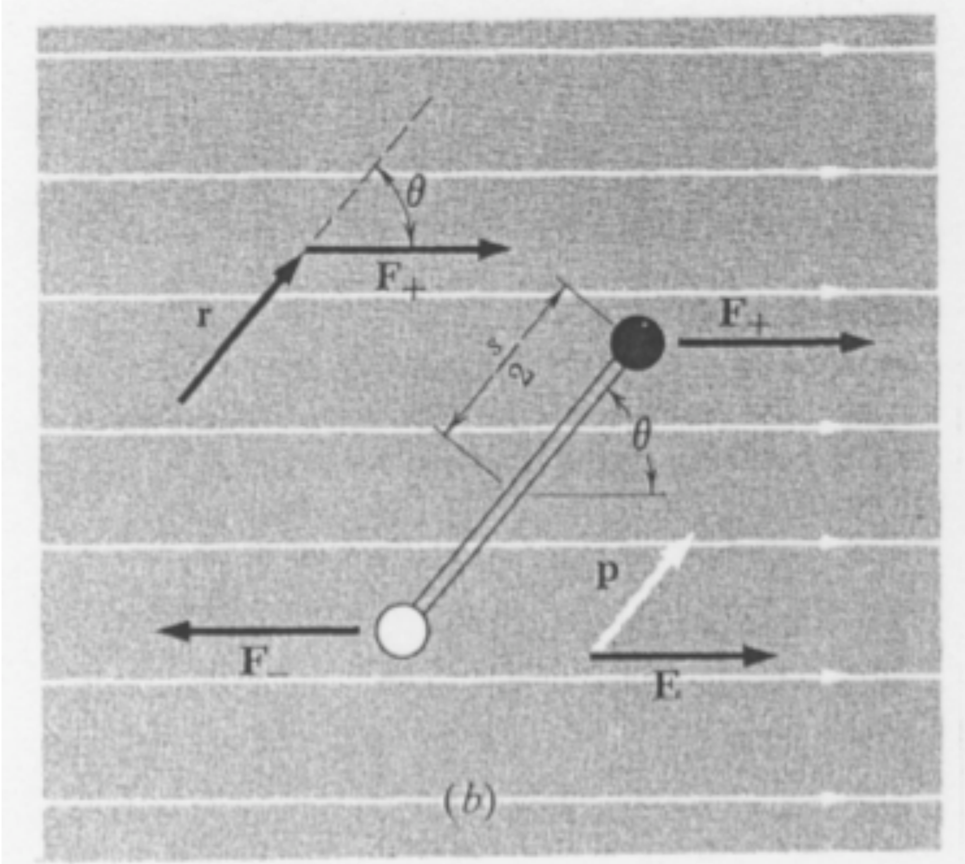
$$E_r = \frac{2p\cos\theta}{r^3} \quad , \quad E_\theta = \frac{p\sin\theta}{r^3}$$

Torque and Force on a Dipole in an External Field

Consider 2 charges ±q connected so that s, the distance between them, is fixed = dipole. Its dipole moment = p = qs. Now insert the dipole into an uniform external electric field as shown in (a). The positive end of the dipole is pulled to the right, the negative end to the left by force qE. Net force on dipole = 0 and net torque = 0 in this position. A dipole at an angle Θ has net force = 0, but net



torque ≠ 0. In general, the torque N around an axis through some origin is r x F (F=force applied at distance r from origin). Let origin = center of dipole (r=s/2) (b). Then



$$\vec{N} = \vec{r} \times \vec{F}_+ + (-\vec{r}) \times \vec{F}_-$$

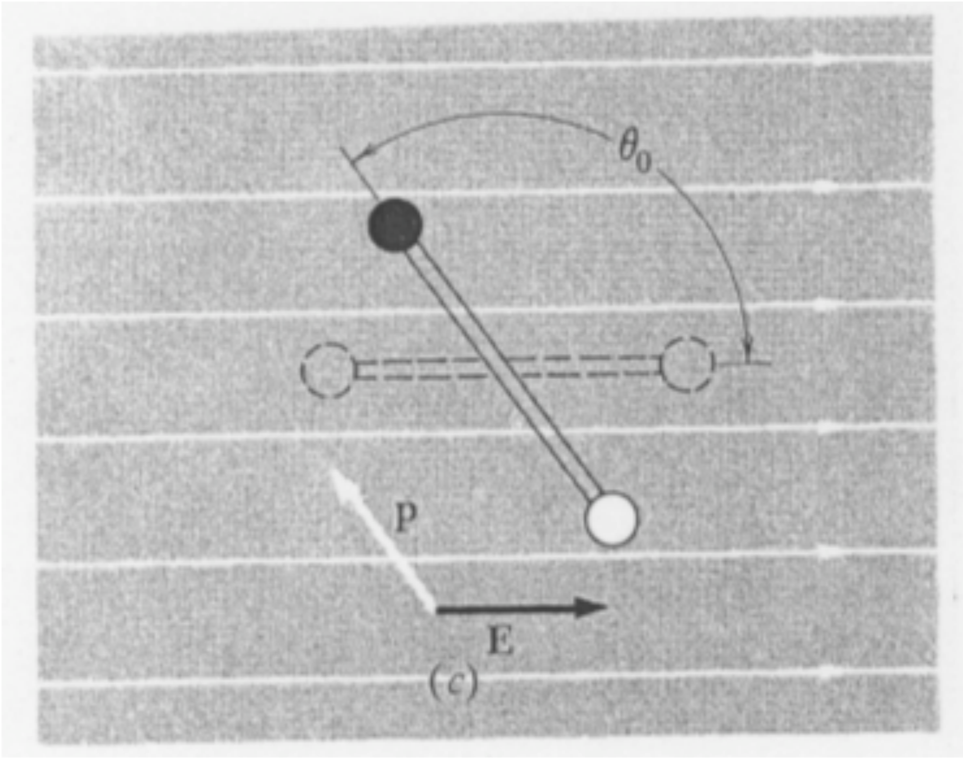
$\vec{N} =$ vector perpendicular to page with magnitude

$$N = \frac{s}{2} qE\sin\theta + \frac{s}{2} qE\sin\theta = sqE\sin\theta = pE\sin\theta$$

or

$$\vec{N} = \vec{p} \times \vec{E}$$

This is independent of origin when the net force = 0. Orientation of dipole in (a) is minimum energy configuration. Work must be done to rotate it to other positions. We calculate the work required to rotate the dipole from a position parallel to the field through an angle Θ_0 as shown in (c) as follows:



$$\int_0^{\theta_0} N\,d\theta = \int_0^{\theta_0} pE\sin\theta\,d\theta = pE(1-\cos\theta_0)$$

To reverse the dipole ($\Theta_0 = \pi$) work = 2pE.

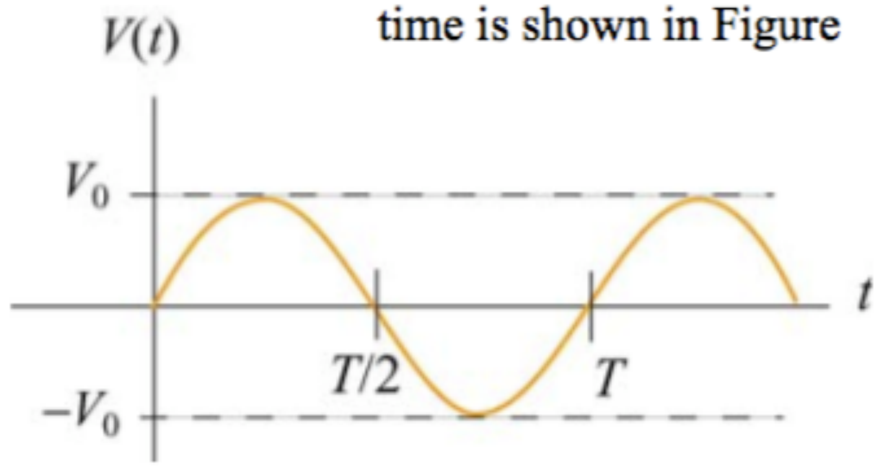
As you can see materials are messy!

First Thoughts about AC Circuits

An example of an AC source is

$$V(t) = V_0 \sin \omega t$$

where the maximum value V_0 is called the *amplitude*. The voltage varies between V_0 and $-V_0$ since a sine function varies between +1 and −1. A graph of voltage as a function of time is shown in Figure



Sinusoidal voltage source

The sine function is periodic in time. This means that the value of the voltage at time t will be exactly the same at a later time $t' = t + T$ where T is the *period*. The *frequency*, f, defined as $f = 1/T$, has the unit of inverse seconds (s^{-1}), or hertz (Hz). The angular frequency is defined to be $\omega = 2\pi f$.

as we will see

When a voltage source is connected to an *RLC* circuit, energy is provided to compensate the energy dissipation in the resistor, and the oscillation will no longer damp out. The oscillations of charge, current and potential difference are called driven or forced oscillations.

After an initial "transient time," an AC current will flow in the circuit as a response to the driving voltage source. The current, written as

$$I(t) = I_0 \sin(\omega t - \phi)$$

will oscillate with the same frequency as the voltage source, with an amplitude I_0 and phase ϕ that depends on the driving frequency.

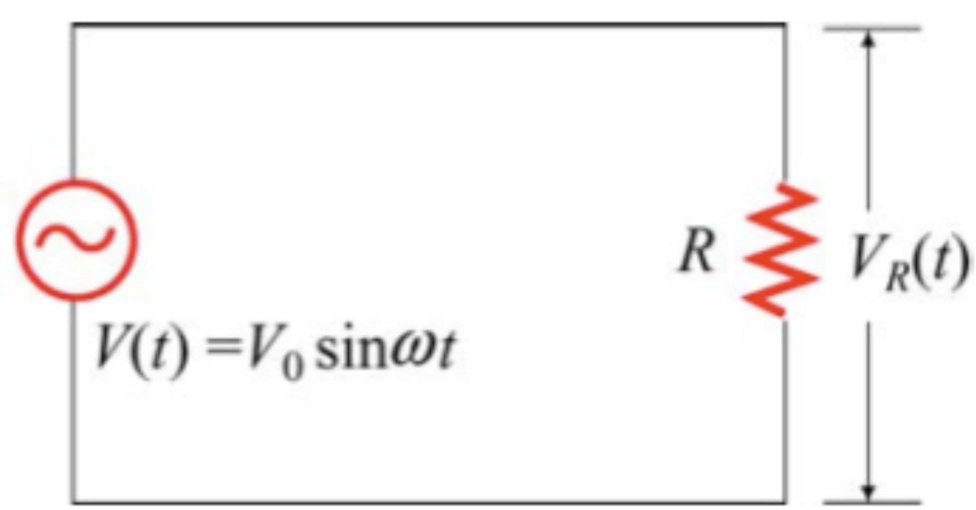
it might not be in phase with the voltage!

Simple AC circuits

Before examining the driven *RLC* circuit, let's first consider the simple cases where only one circuit element (a resistor, an inductor or a capacitor) is connected to a sinusoidal voltage source.

Purely Resistive load

Consider a purely resistive circuit with a resistor connected to an AC generator, as shown in Figure (As we shall see, a purely resistive circuit corresponds to infinite capacitance $C = \infty$ and zero inductance $L = 0$.)



A purely resistive circuit

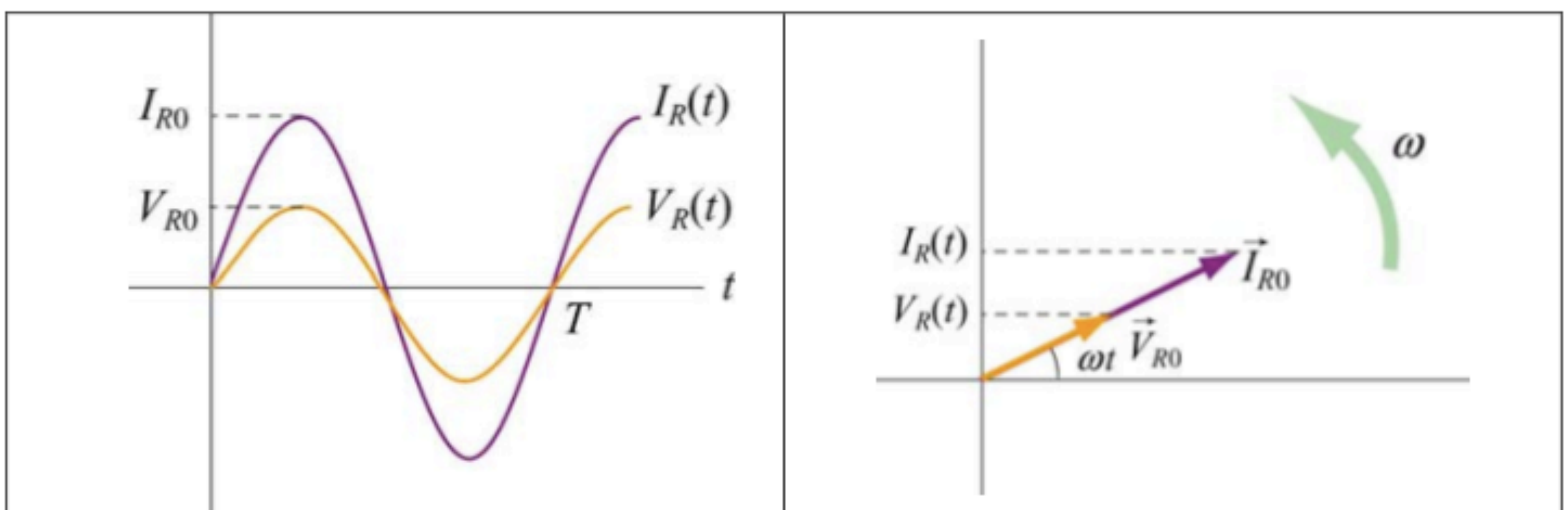
Applying Kirchhoff's loop rule yields

$$V(t)-V_R(t)=V(t)-I_R(t)R=0$$

where $V_R(t)=I_R(t)R$ is the instantaneous voltage drop across the resistor. The instantaneous current in the resistor is given by

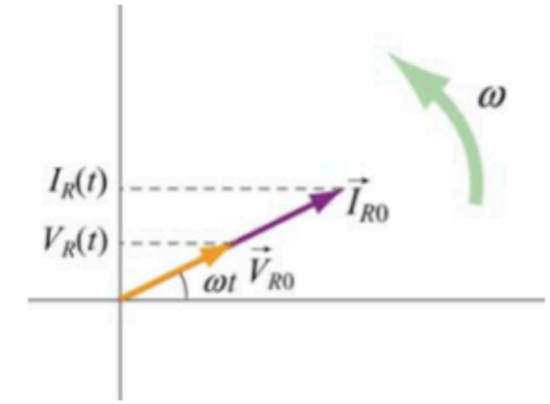
$$I_R(t)=\frac{V_R(t)}{R}=\frac{V_{R0}\sin\omega t}{R}=I_{R0}\sin\omega t$$

where $V_{R0}=V_0$, and $I_{R0}=V_{R0}/R$ is the maximum current.
we find $\phi=0$, which means that $I_R(t)$ and $V_R(t)$ are in phase with each other, meaning that they reach their maximum or minimum values at the same time. The time dependence of the current and the voltage across the resistor is depicted in Figure



(a) Time dependence of $I_R(t)$ and $V_R(t)$ across the resistor. (b) Phasor diagram for the resistive circuit.

The behavior of $I_R(t)$ and $V_R(t)$ can also be represented with a phasor diagram, as shown



A phasor is a rotating vector having the following properties:

(i) length: the length corresponds to the amplitude.

(ii) angular speed: the vector rotates counterclockwise with an angular speed ω.

(iii) projection: the projection of the vector along the vertical axis corresponds to the value of the alternating quantity at time t.

We shall denote a phasor with an arrow above it. The phasor $\vec{V}_{R0}$ has a constant magnitude of V_{R0}. Its projection along the vertical direction is $V_{R0}\sin\omega t$, which is equal to $V_R(t)$, the voltage drop across the resistor at time t. A similar interpretation applies to $\vec{I}_{R0}$ for the current passing through the resistor. From the phasor diagram, we readily see that both the current and the voltage are in phase with each other.

The average value of current over one period can be obtained as:

The average value of current over one period can be obtained as:

$$\langle I_R(t)\rangle = \frac{1}{T}\int_0^T I_R(t)dt = \frac{1}{T}\int_0^T I_{R0}\sin\omega t\; dt = \frac{I_{R0}}{T}\int_0^T \sin\frac{2\pi t}{T}\; dt = 0$$

This average vanishes because

$$\langle \sin\omega t\rangle = \frac{1}{T}\int_0^T \sin\omega t\; dt = 0$$

Similarly, one may find the following relations useful when averaging over one period:

$$\langle \cos\omega t\rangle = \frac{1}{T}\int_0^T \cos\omega t\; dt = 0$$

$$\langle \sin\omega t\cos\omega t\rangle = \frac{1}{T}\int_0^T \sin\omega t\cos\omega t\; dt = 0$$

$$\langle \sin^2\omega t\rangle = \frac{1}{T}\int_0^T \sin^2\omega t\; dt = \frac{1}{T}\int_0^T \sin^2\left(\frac{2\pi t}{T}\right) dt = \frac{1}{2}$$

$$\langle \cos^2\omega t\rangle = \frac{1}{T}\int_0^T \cos^2\omega t\; dt = \frac{1}{T}\int_0^T \cos^2\left(\frac{2\pi t}{T}\right) dt = \frac{1}{2}$$

From the above, we see that the average of the square of the current is non-vanishing:

$$\langle I_R^2(t)\rangle = \frac{1}{T}\int_0^T I_R^2(t)dt = \frac{1}{T}\int_0^T I_{R0}^2 \sin^2 \omega t \, dt = I_{R0}^2 \frac{1}{T}\int_0^T \sin^2\left(\frac{2\pi t}{T}\right) dt = \frac{1}{2} I_{R0}^2$$

It is convenient to define the root-mean-square (rms) current as

$$I_{\text{rms}} = \sqrt{\langle I_R^2(t)\rangle} = \frac{I_{R0}}{\sqrt{2}}$$

In a similar manner, the rms voltage can be defined as

$$V_{\text{rms}} = \sqrt{\langle V_R^2(t)\rangle} = \frac{V_{R0}}{\sqrt{2}}$$

The rms voltage supplied to the domestic wall outlets in the United States is $V_{\text{rms}} = 120$ V at a frequency $f = 60$ Hz.

The power dissipated in the resistor is

$$P_R(t) = I_R(t) V_R(t) = I_R^2(t) R$$

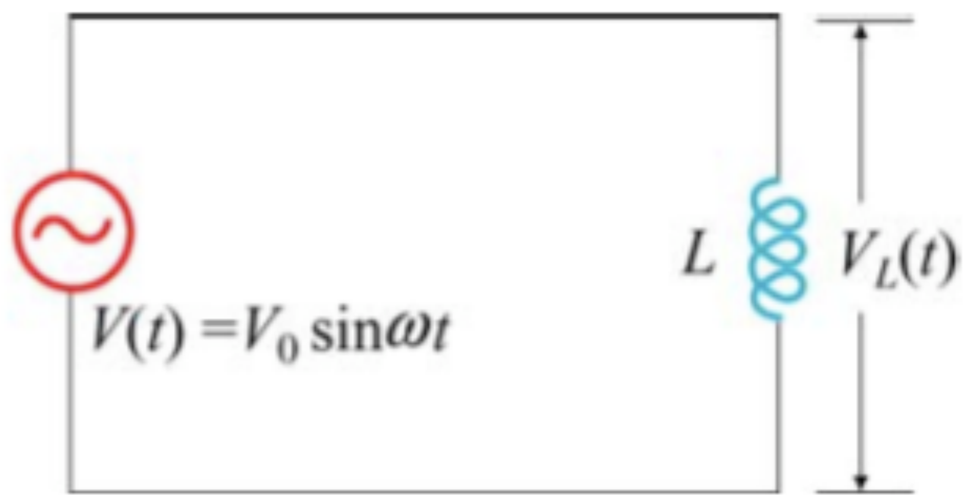
from which the average over one period is obtained as:

$$\langle P_R(t)\rangle = \langle I_R^2(t) R\rangle = \frac{1}{2} I_{R0}^2 R = I_{\text{rms}}^2 R = I_{\text{rms}} V_{\text{rms}} = \frac{V_{\text{rms}}^2}{R}$$

Purely Inductive Load

Introduce inductors on the fly!!

Consider now a purely inductive circuit with an inductor connected to an AC generator, as shown in Figure

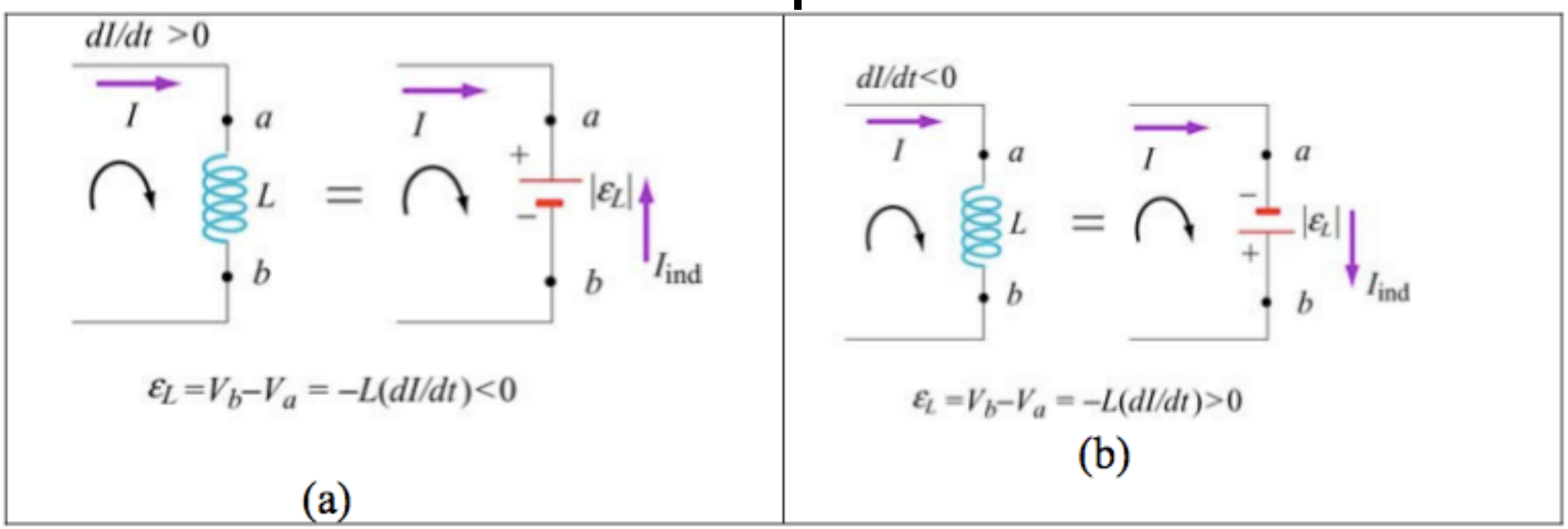


A purely inductive circuit

As we shall see below, a purely inductive circuit corresponds to infinite capacitance $C = \infty$ and zero resistance $R = 0$. Applying the modified Kirchhoff's rule for inductors, the circuit equation reads

we will prove this result later

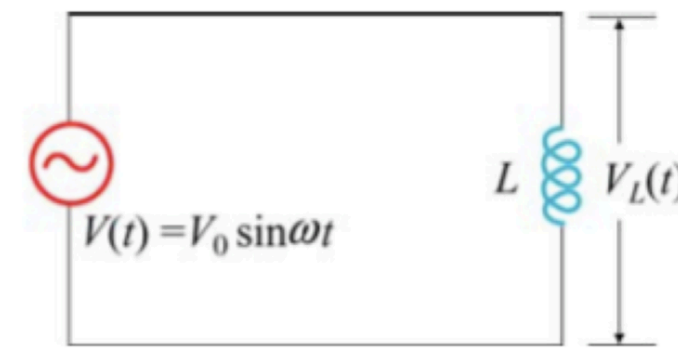
(Lenz's law)



2 Modified Kirchhoff's rule for inductors (a) with increasing current, and (b) with decreasing current.

As we will see later, Lenz's law will say that an inductor (think of a coil of wire carrying a current), will always try to opposing any increase in current in the circuit. Thus, an emf develops across the inductor such that it appear to be a battery trying to reduce the current as shown in the above diagrams.

$$V(t) - V_L(t) = V(t) - L\frac{dI_L}{dt} = 0$$



which implies

$$\frac{dI_L}{dt} = \frac{V(t)}{L} = \frac{V_{L0}}{L}\sin\omega t$$

where $V_{L0} = V_0$. Integrating over the above equation, we find

$$I_L(t) = \int dI_L = \frac{V_{L0}}{L}\int \sin\omega t\, dt = -\left(\frac{V_{L0}}{\omega L}\right)\cos\omega t = \left(\frac{V_{L0}}{\omega L}\right)\sin\left(\omega t - \frac{\pi}{2}\right)$$

where we have used the trigonometric identity

$$-\cos\omega t = \sin\left(\omega t - \frac{\pi}{2}\right)$$

for rewriting the last expression. Comparing we see that

the amplitude of the current through the inductor is

$$I_{L0} = \frac{V_{L0}}{\omega L} = \frac{V_{L0}}{X_L}$$

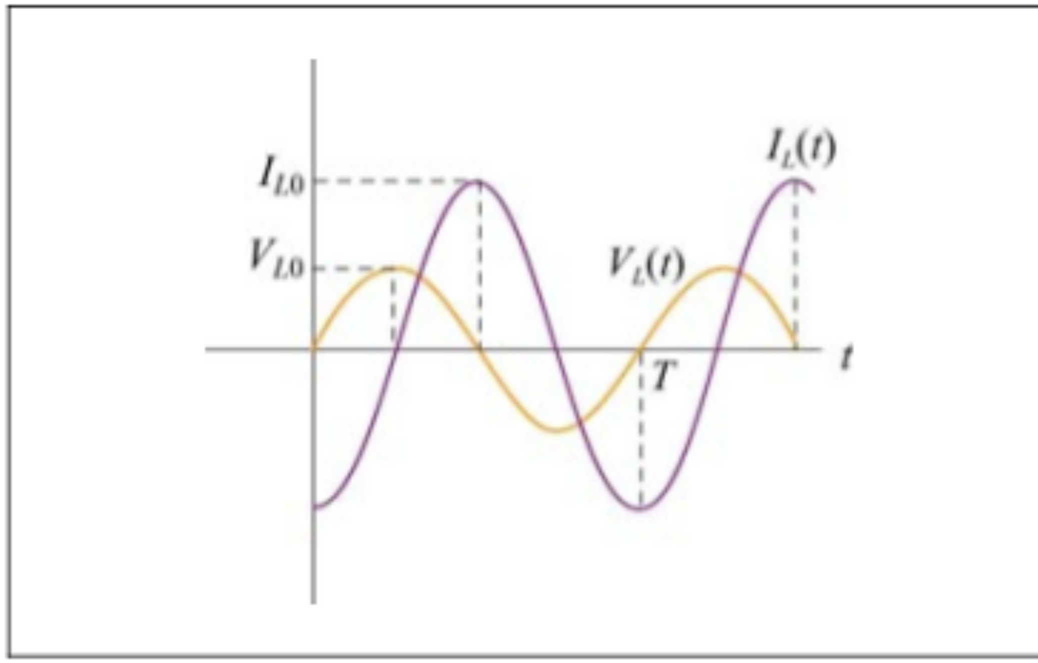
where

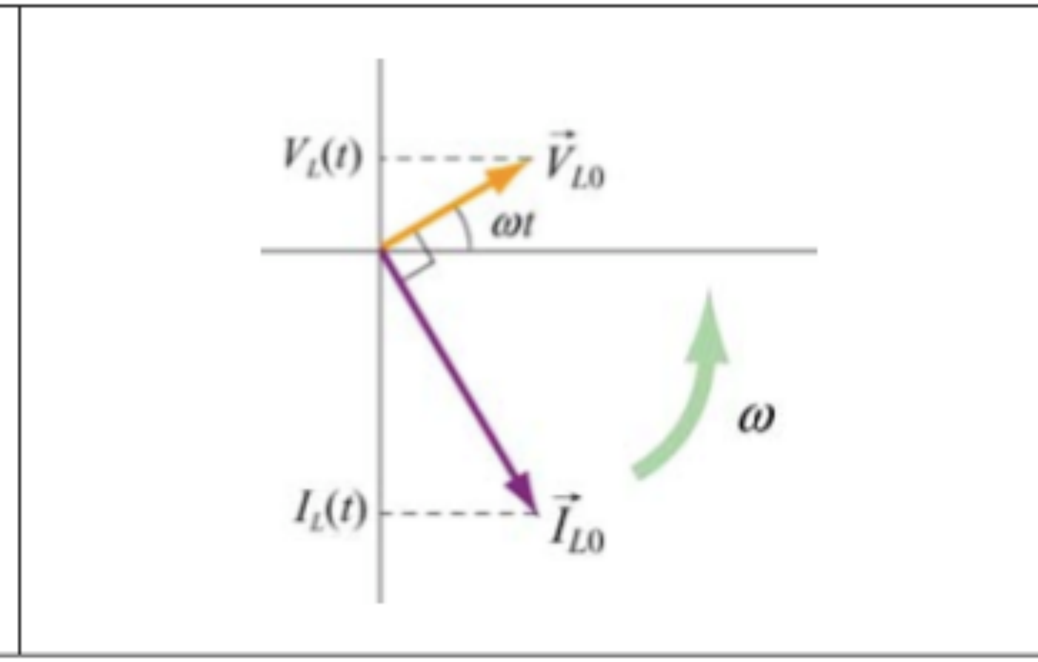
$$X_L = \omega L$$

is called the *inductive reactance*. It has SI units of ohms (Ω), just like resistance. However, unlike resistance, X_L depends linearly on the angular frequency ω. Thus, the resistance to current flow increases with frequency. This is due to the fact that at higher frequencies the current changes more rapidly than it does at lower frequencies. On the other hand, the inductive reactance vanishes as ω approaches zero.

we also find the phase constant to be $\phi = +\frac{\pi}{2}$

The current and voltage plots and the corresponding phasor diagram are shown in the Figure





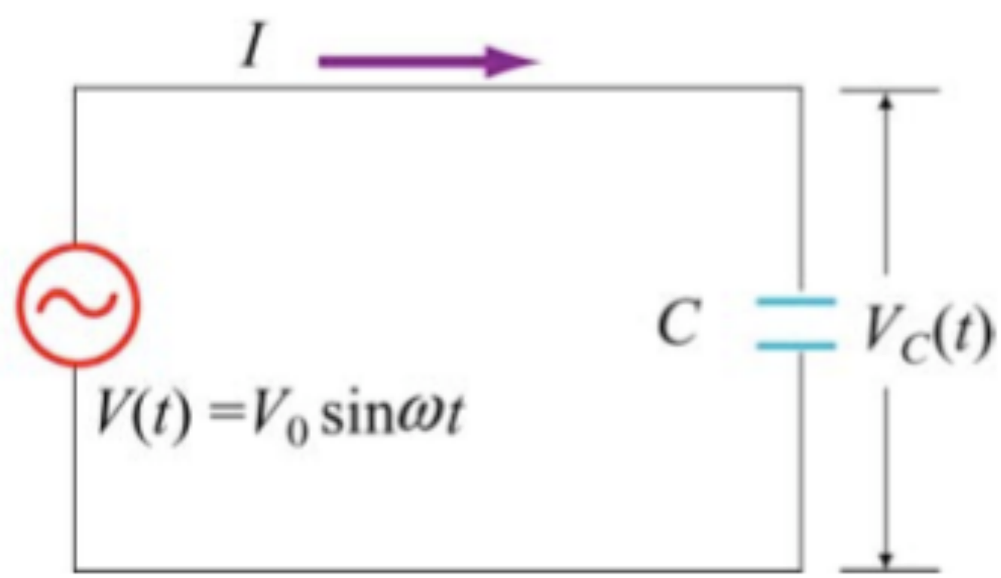
(a) Time dependence of $I_L(t)$ and $V_L(t)$ across the inductor. (b) Phasor diagram for the inductive circuit.

As can be seen from the figures, the current $I_L(t)$ is out of phase with $V_L(t)$ by $\phi = \pi / 2$; it reaches its maximum value after $V_L(t)$ does by one quarter of a cycle. Thus, we say that

The current lags voltage by $\pi / 2$ in a purely inductive circuit

Purely Capacitive Load

In the purely capacitive case, both resistance R and inductance L are zero. The circuit diagram is shown in Figure



A purely capacitive circuit

Again, Kirchhoff's voltage rule implies

$$V(t) - V_C(t) = V(t) - \frac{Q(t)}{C} = 0$$

which yields

$$Q(t) = CV(t) = CV_C(t) = CV_{C0} \sin \omega t$$

where $V_{C0} = V_0$. On the other hand, the current is

$$I_C(t) = +\frac{dQ}{dt} = \omega C V_{C0} \cos \omega t = \omega C V_{C0} \sin\left(\omega t + \frac{\pi}{2}\right)$$

where we have used the trigonometric identity

$$\cos \omega t = \sin\left(\omega t + \frac{\pi}{2}\right)$$

The above equation indicates that the maximum value of the current is

$$I_{C0} = \omega C V_{C0} = \frac{V_{C0}}{X_C}$$

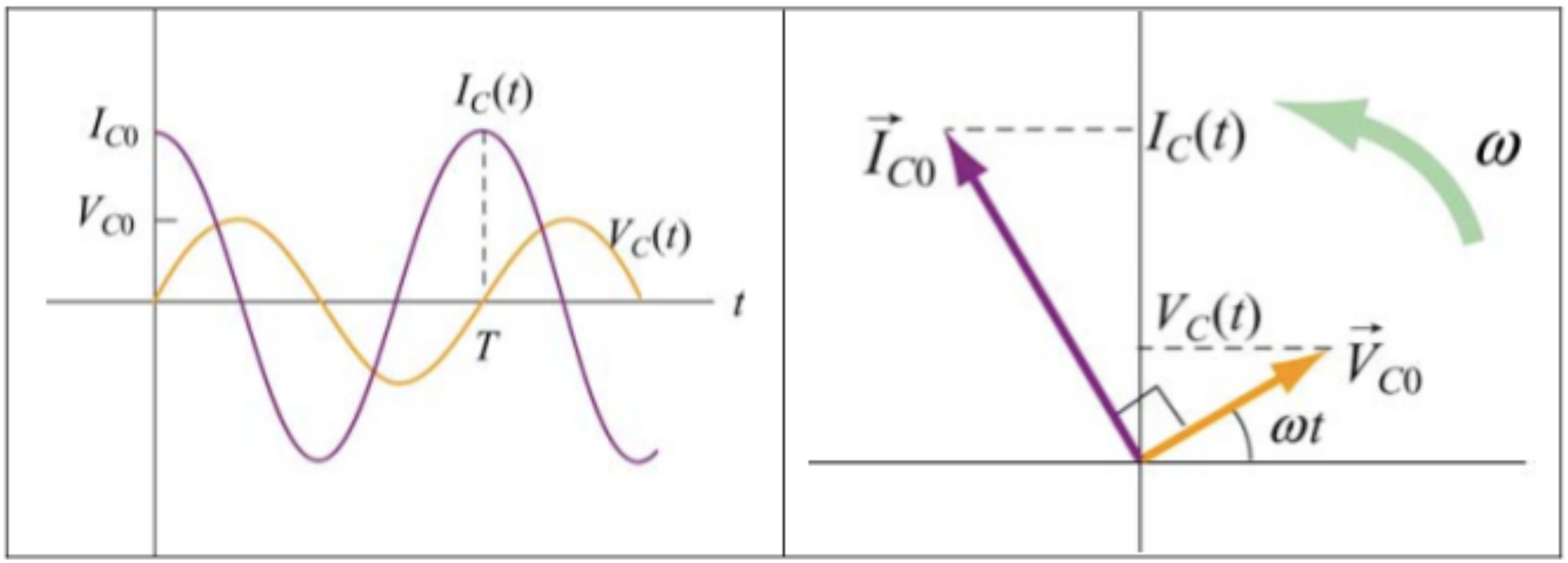
where

$$X_C = \frac{1}{\omega C}$$

is called the *capacitance reactance*. It also has SI units of ohms and represents the effective resistance for a purely capacitive circuit. Note that X_C is inversely proportional to both C and ω, and diverges as ω approaches zero.

the phase constant is given by $$\phi = -\frac{\pi}{2}$$

The current and voltage plots and the corresponding phasor diagram are shown in the Figure



(a) Time dependence of $I_C(t)$ and $V_C(t)$ across the capacitor. (b) Phasor diagram for the capacitive circuit.

Notice that at $t = 0$, the voltage across the capacitor is zero while the current in the circuit is at a maximum. In fact, $I_C(t)$ reaches its maximum before $V_C(t)$ by one quarter of a cycle ($\phi = \pi / 2$). Thus, we say that

The current leads the voltage by $\pi/2$ in a capacitive circuit

RL circuits: **qualitative** description

At t=0, close S1 :

Lenz's law -> L opposes change in current through L by developing an opposing emf, initially L will impede current flow and initially I(0)=0 (dI/dt>0)

As time passes, I will start flowing saturating at I=V/R

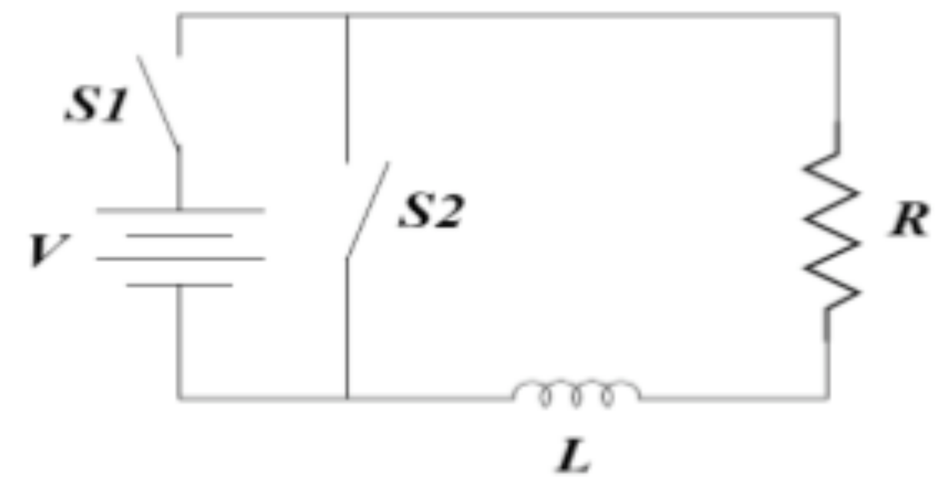
After a long time, simultaneously open S_1 and close S_2:

Lenz's law opposes change in current through L

Now wants to stop current decrease(dI/dt<0).

Back emf will keep current flowing for a while

R dissipates power → the current will die exponentially



RL circuits: **quantitative** description

At t = 0: close S_1 and open S_2

Kirchoff's rule #2:(series circuit)

$$V - IR - L\frac{dI}{dt} = 0$$

$$V - IR - L\frac{dI}{dt} = 0$$

$$-I + \frac{V}{R} = \frac{L}{R}\frac{dI}{dt} \Rightarrow \frac{dI}{I - \frac{V}{R}} = -\frac{R}{L}dt$$

Rewrite as:

$$\int_{I(0)}^{I(t)} \frac{dI}{I - \frac{V}{R}} = -\frac{R}{L}\int_0^t dt \Rightarrow \ln\left(\frac{I(t) - \frac{V}{R}}{I(0) - \frac{V}{R}}\right) = -\frac{R}{L}t$$

$$\Rightarrow \frac{I(t) - \frac{V}{R}}{-\frac{V}{R}} = e^{-\frac{R}{L}t} \Rightarrow I(t) = \frac{V}{R}\left(1 - e^{-\frac{R}{L}t}\right)$$

-> I(0)=0; I(∞)=V/R->saturation)

At t = t': open S_1 and close S_2

no emf,but energy stored in inductor

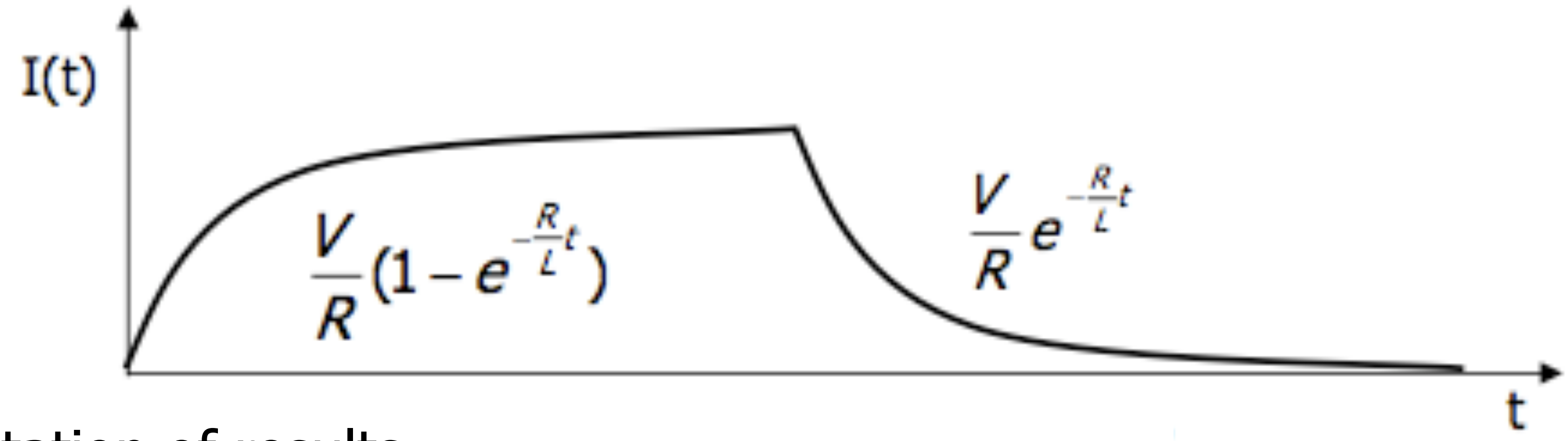
Kirchoff's rule #2: (series circuit)

$$-IR - L\frac{dI}{dt} = 0$$

$$-I = \frac{L}{R}\frac{dI}{dt} \Rightarrow \int_{I=I_0}^{I=I(t)} \frac{dI}{I} = -\int_{t=0}^{t} \frac{R}{L}dt \Rightarrow \ln\frac{I}{I_0} = -\frac{R}{L}t$$

$$\Rightarrow I(t) = I_0 e^{-\frac{R}{L}t} = \frac{V}{R}e^{-\frac{R}{L}t}$$

Graphically:



RL circuits: interpretation of results

How do we interpret these results?

Inductors cause currents to have an "inertia"

- If no current flowing: L forces I to build up gradually
- If current is flowing: L will do what it takes to make it continue (back- emf)

Asymptotic behavior when "charging" L

- At t=0, I=0, as if L were an open circuit   t=0: L ➙ open circuit
- At t=infinity, I=V/R, as if L did not exist   t=∞: L ➙ short circuit

RL circuits: time constant

Results of RL circuit are exponentials, as in RC circuits

- RC circuit: time constant $\tau=RC$
- RL circuits: time constant $\tau=L/R$

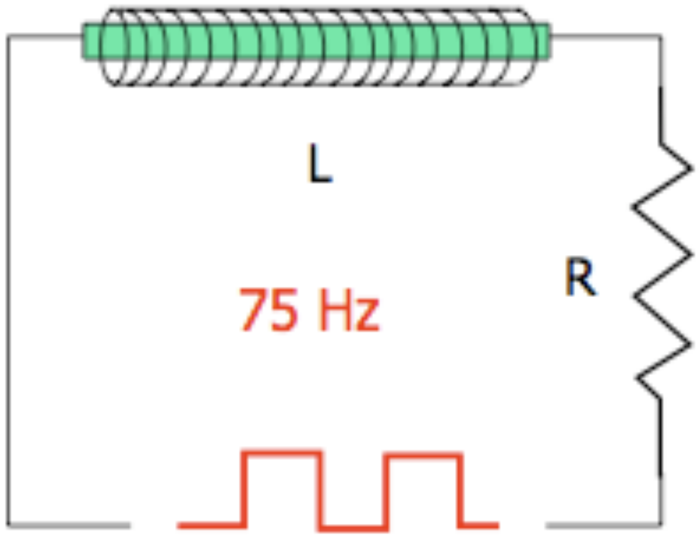
Note: time constant is the time it takes the exponential function to decrease (increase) to 1/e (1-1/e) of its original (final) value

Check units

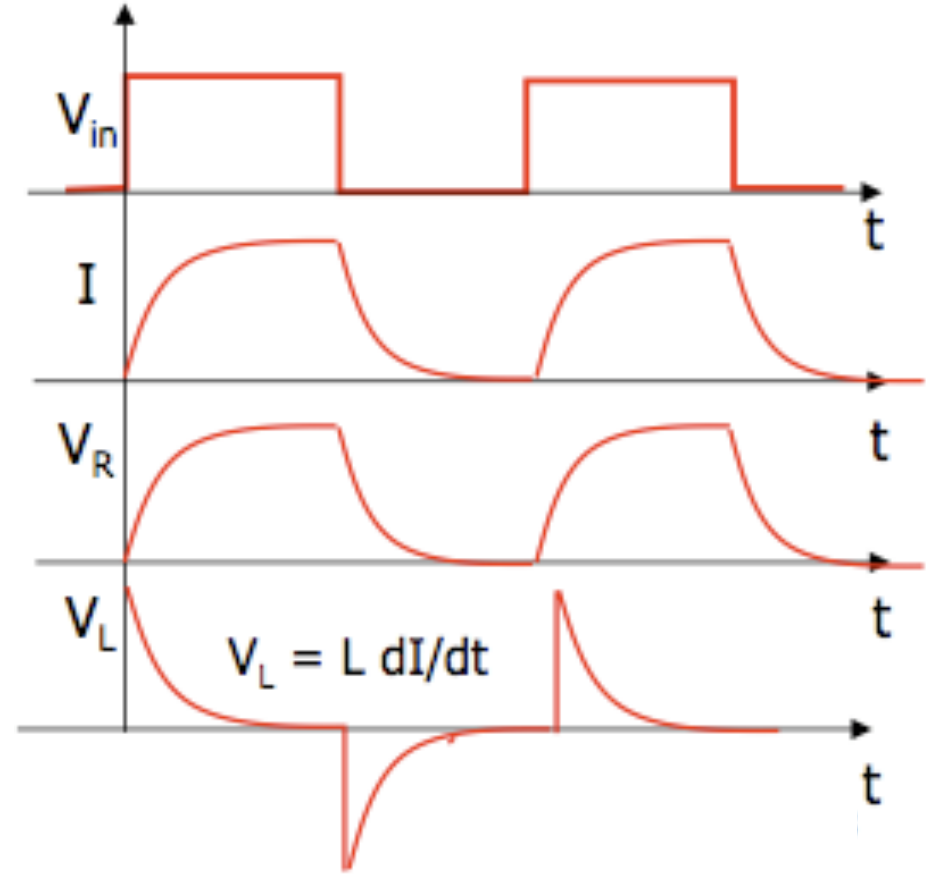
- cgs: $[L]/[R]=(sec^2/cm)/(sec/cm)=sec$
- SI: $[L]/[R]= H/\Omega= (V\ sec/A)/(V/A) = sec$

LR and square pulses…

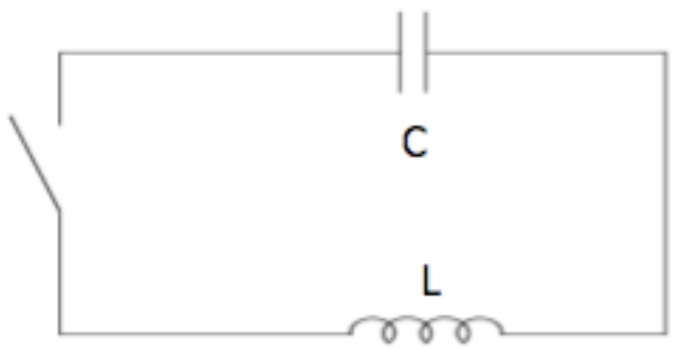
Consider the following circuit



On the oscilloscope: V_{input}, V_L , V_R , I in the circuit



LC circuits



Start with charged capacitor and close switch at t = 0:

Kirchoff's second rule: $$\frac{Q}{C} - L\frac{dI}{dt} = 0$$

Since

$$I = -\frac{dQ}{dt} \Rightarrow \frac{d^2Q}{dt^2} + \frac{Q}{LC} = 0$$

How to solve this? Educated guess(always 1st try of a physicist):

$$Q(t) = A\cos\omega_0 t + B\sin\omega_0 t$$

See if there exists ANY values of A,B, ω_0--> solution exists

LC circuits: solution

Plug this in the differential equation:

$$\frac{d^2Q(t)}{dt^2} = -\frac{Q(t)}{LC} \Rightarrow -\omega_0^2 Q(t) = -\frac{Q(t)}{LC} \Rightarrow \omega_0 = \frac{1}{\sqrt{LC}}$$

if,then --->solution

Determine constants A and B from initial conditions:

$Q(t=0)=Q_0= A\cos(0) + B\sin(0)$ → $A=Q_0$

$I(t=0)=0 = -\omega_0 A\sin(0) + \omega_0 B\cos(0)$ → $B=0$

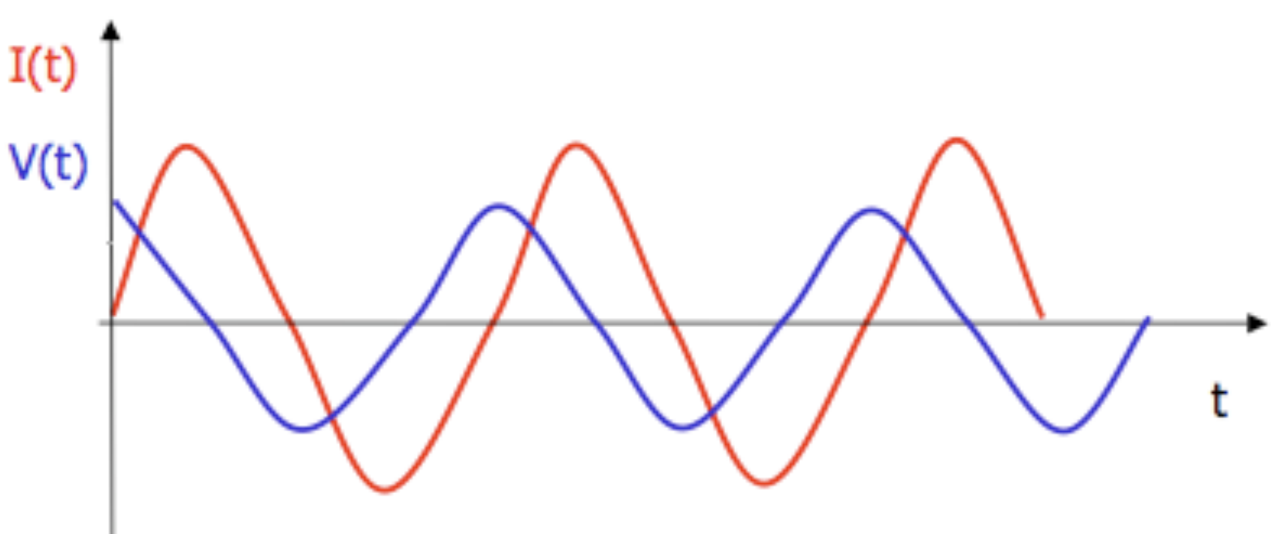
Complete solution:

$$Q(t) = Q_0\cos\omega_0 t \Rightarrow V_C(t) = \frac{Q(t)}{C} = \frac{Q_0}{C}\cos\omega_0 t$$

$$I(t) = -\frac{dQ}{dt} = \frac{Q_0}{\sqrt{LC}}\sin\omega_0 t = \frac{Q_0}{\sqrt{LC}}\cos\left(\omega_0 t - \pi/2\right)$$

Note: current and voltages are out of phase by 90 degrees(π/2)

Graphical representation of the solution:



Digression: Solving Simple Ordinary Differential Equations(ODEs)
Guessing is not good in general

First-Order Equations

Consider an equation of the form

$$\frac{dy}{dx} + Ay = 0 \qquad \text{a homogeneous equation} \qquad (01)$$

where A is a constant.

This equation is called first-order (highest derivative is a first derivative) ordinary linear (all functions and derivatives appear to the first power or less) differential equation with constant coefficients.

We solve this type of ODE using an exponential substitution method which converts the differential equation into a solvable algebraic equation.

Exponential Substitution Method

We make the exponential substitution

$$y = e^{\alpha x} \qquad (02)$$

into the ODE. This will convert the ODE into an **algebraic** equation for α. We have

$$\frac{dy}{dx} = \frac{de^{\alpha x}}{dx} = \alpha e^{\alpha x}$$

$$y = e^{\alpha x}$$

which upon substitution into (01) gives the result

$$(\alpha + A)e^{\alpha x} = 0 \rightarrow \alpha + A = 0 \tag{03}$$

since $e^{\alpha x} \neq 0$.

The solution(s) of this algebraic equation tell us the **allowed** values of α that give **valid** solutions to the ODE. In particular in this case we get only one allowed value, namely,

$$\alpha = -A \tag{04}$$

as a solution to the linear algebraic equation in (03). This result means that the function

$$y = e^{-Ax} \tag{04}$$

satisfy the original ODE and therefore a solution.

The general solution to the equation is then written

$$y(x) = Ce^{-Ax} \tag{05}$$

where C is a constant that is determined by the initial conditions as shown below:

Suppose we have the initial condition(s) (number = order of ODE, which is one in this case)

$$y(0) = 7 \tag{06}$$

Then the value of the unknown constant C is determined by making sure that the general solution agrees with (or satisfies) the initial condition(s). We have

$$y(0) = 7 = C \tag{07}$$

Therefore the solution satisfying the initial condition is

$$y(x) = 7e^{-Ax} \tag{08}$$

Check:

$$\frac{dy}{dx} = \frac{d(7e^{-Ax})}{dx} = -7Ae^{-Ax}$$

$$y = 7e^{-Ax}$$

$$\frac{dy}{dx} + Ay = -7Ae^{-Ax} + 7Ae^{-Ax} = 0$$

$$y(0) = 7e^{0} = 7$$

Therefore, for this type of ODE, this exponential substitution method is able to generate a solution.

Example:

Remember RL circuit discussion...

$$IR + L\frac{dI}{dt} = V \Rightarrow I(t) = I_h(t) + I_p(t) \quad where \quad I_h R + L\frac{dI_h}{dt} = 0 \quad \text{homogeneous equation}$$

$$and \quad I_p R + L\frac{dI_p}{dt} = V \quad \text{particular equation}$$

$$I_h(t) = Ae^{\alpha t} \Rightarrow Ae^{\alpha t}(R + \alpha L) = 0 \Rightarrow \alpha = -R / L \Rightarrow I_h(t) = Ae^{-Rt/L}$$

$$I_p(t) = V / R \quad \text{any solution that works is the one by uniqueness!!!}$$

$$I(t) = I_h(t) + I_p(t) = \frac{V}{R} + Ae^{\alpha t}$$

$$I(0) = 0 = \frac{V}{R} + A \Rightarrow A = -\frac{V}{R} \Rightarrow I(t) = I_h(t) + I_p(t) = \frac{V}{R}(1 - e^{\alpha t}) \quad \text{as earlier....}$$

Second-Order Equations

Consider an equation of the form

$$A\frac{d^2y}{dt^2}+B\frac{dy}{dt}+Cy=0 \tag{09}$$

where $A, B, and\, C$ are constants.

Many equations derived from Newton's second law in mechanics and from Schrodinger's equation in quantum mechanics take this form.

This equation is called second-order (highest derivative is a second derivative) ordinary linear (all functions and derivatives appear to the first power or less) differential equation with constant coefficients.

We solve this type of ODE using an exponential substitution method which converts the differential equation into a solvable algebraic equation.

Exponential Substitution Method

The Method: consider a typical equation of the form

$$\frac{d^2y}{dt^2} + 3\frac{dy}{dt} + 2y = 0 \tag{10}$$

We make the exponential substitution

$$y = e^{\alpha t} \tag{11}$$

into the ODE. This will convert the ODE into an **algebraic** equation for α. We have

$$\frac{d^2y}{dt^2} = \frac{d^2e^{\alpha t}}{dt^2} = \alpha\frac{de^{\alpha t}}{dt} = \alpha^2 e^{\alpha t}$$

$$\frac{dy}{dt} = \frac{de^{\alpha t}}{dt} = \alpha e^{\alpha t}$$

$$y = e^{\alpha t}$$

which gives the result

$$(\alpha^2 + 3\alpha + 2)e^{\alpha t} = 0 \rightarrow \alpha^2 + 3\alpha + 2 = 0 \tag{12}$$

since $e^{\alpha t} \neq 0$.

The solutions of this algebraic equation tell us the **allowed** values of α that give **valid** solutions to the ODE. In particular, in this case we get

$$\alpha = -1 \; and \; -2 \tag{13}$$

as solutions to the quadratic equation. This result means that the functions $y = e^{-t}$ and $y = e^{-2t}$ **each** satisfy the original ODE.

If there is more than one allowed value of α (as in this case), then the most general solution will be a linear combination of all possible solutions(because this is a linear ODE, that is, all derivative and functions enter in the equation to the first-power). Since, in this case, the allowed values of α are

$$\alpha = -1 \; and \; -2 \tag{14}$$

the most general solution of the ODE is

$$y(t) = ae^{-t} + be^{-2t} \tag{15}$$

where a and b are constants to be determined by the **initial conditions.**

Again, the number of arbitrary constants that need to be determined by the initial conditions is equal to the order(highest derivative → 2 in this case) of the ODE.

Suppose the initial conditions are $y(0)=0$ and $\left.\frac{dy}{dt}\right|_{t=0}=1$ (at $t=0$). Then we have

$$\begin{aligned} y(t) &= ae^{-t}+be^{-2t} \\ y(0) &= 0 = a+b \\ \frac{dy}{dt} &= -ae^{-t}-2be^{-2t} \\ \frac{dy}{dt}(0) &= -a-2b=1 \end{aligned} \tag{16}$$

which gives $a=-b=1$ and

$$y(t)=e^{-t}-e^{-2t} \tag{17}$$

If we substitute this solution into the original equation we have

$$\begin{aligned} y(t) &= e^{-t}-e^{-2t} \Rightarrow y(0)=0 \\ \frac{dy}{dt} &= -e^{-t}+2e^{-2t} \Rightarrow \frac{dy}{dt}(0)=1 \end{aligned}$$

as required and

$$y(t) = e^{-t} - e^{-2t}$$

$$\frac{dy}{dt} = -e^{-t} + 2e^{-2t}$$

$$\frac{d^2y}{dt^2} = e^{-t} - 4e^{-2t}$$

$$\frac{d^2y}{dt^2} + 3\frac{dy}{dt} + 2y = (e^{-t} - 4e^{-2t}) + 3(-e^{-t} + 2e^{-2t}) + 2(e^{-t} - e^{-2t}) = 0$$

as required. Therefore, we see that the method works and generates a solution with the correct initial conditions!!

Although this method is very powerful as described, we can make it even more powerful by using the new mathematical quantity called the **complex exponential** as defined in the notes. This will allow us to use the method for the Schrodinger equation case.

Reference - if there are any questions

Digression on complex numbers

$i^2 = -1$, $i^3 = -i$, $i^4 = +1$, $i^5 = i$, $i^6 = -1$, and so on **definition** of i

$z = a + bi$ **definition** of complex number a = **real** part b= **imaginary** part

$$(7 + 4i) + (-2 + 9i) = 5 + 13i$$

definition of addition

add real part and imaginary parts separately

$$(7 + 4i)(-2 + 9i) = (7)(-2) + (7)(9i) + (4i)(-2) + (4i)(9i)$$
$$= -14 + 63i - 8i - 36 = -50 + 55i$$

using $i^2 = -1$

definition of multiplication

definition complex conjugate $z^* = a - bi$

$$|z|^2 = z^* z = (a - bi)(a + bi) = a^2 + b^2$$

$|z| = \sqrt{a^2 + b^2}$ absolute value

Note that z is real is $z^* = z$

Reference - if there are any questions

A power series representation of a function:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots = \sum_{k=0}^{\infty} a_k x^k$$

Special cases

$$e^{\alpha x} = 1 + \alpha + \frac{1}{2}\alpha^2 + \frac{1}{6}\alpha^3 + ... = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} x^n$$

$$\sin \alpha x = \alpha x - \frac{1}{6}(\alpha x)^3 + \frac{1}{24}(\alpha x)^5 + ... = \sum_{n=0}^{\infty} (-1)^n \frac{\alpha^{2n+1}}{(2n+1)!} x^{2n+1}$$

$$\cos \alpha x = 1 - \frac{1}{2}(\alpha x)^2 + \frac{1}{16}(\alpha x)^4 + ... = \sum_{n=0}^{\infty} (-1)^n \frac{\alpha^{2n}}{(2n)!} x^{2n}$$

Expansions are still valid if α is a complex number

—> important mathematical result for QM

Now

$$e^{i\alpha x} = \sum_{n=0}^{\infty} \frac{i^n \alpha^n}{n!} x^n = 1 + i\alpha x - \frac{\alpha^2}{2!}x^2 - i\frac{\alpha^3}{3!}x^3 + \frac{\alpha^4}{4!}x^4 + i\frac{\alpha^5}{5!}x^5 - \ldots\ldots$$

$$= \left(1 - \frac{\alpha^2}{2!}x^2 + \frac{\alpha^4}{4!}x^4 - \ldots\ldots\right) + \left(i\alpha x - i\frac{\alpha^3}{3!}x^3 + i\frac{\alpha^5}{5!}x^5 - \ldots\ldots\right)$$

$$= \cos\alpha x + i\sin\alpha x$$ —> Euler relation

Reference - if there are any questions

Then

$$\sin\alpha x = \frac{e^{i\alpha x} - e^{-i\alpha x}}{2i} \quad , \quad \cos\alpha x = \frac{e^{i\alpha x} + e^{-i\alpha x}}{2}$$

Euler relation —> define i without using square root.

$$e^{i\pi} = \cos\pi + i\sin\pi = -1 \qquad e^{i\pi/2} = \cos\pi/2 + i\sin\pi/2 = i$$

Since $e^{a+b} = e^a e^b$ $\quad e^a = e^{a/2}e^{a/2} \to e^{a/2} = \sqrt{e^a}$ and $(e^a)^n = e^{na}$

$$\sqrt{e^{i\pi}} = e^{i\pi/2} = \cos\pi/2 + i\sin\pi/2 = i$$

Now from earlier we had $a = R\cos\theta \;,\; b = R\sin\theta \to \dfrac{b}{a} = \tan\theta \;,\; R^2 = a^2 + b^2$

or

$$z = a + ib = R\cos\theta + iR\sin\theta = R(\cos\theta + i\sin\theta) = Re^{i\theta}$$

some repetition

Complex Exponentials - Alternative Very Powerful Method

Remember our discussion earlier in class about power series expansions of a function around some point

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \qquad (18)$$

the point is $x = 0$ in this case.

If we apply this to the exponential function we get

$$f(x) = e^{\alpha x}$$
$$f^{(0)}(0) = f(0) = 1$$
$$f^{(1)}(0) = \left.\frac{df}{dx}\right|_{x=0} = \alpha e^{\alpha x}\Big|_{x=0} = \alpha$$
$$f^{(2)}(0) = \left.\frac{d^2 f}{dx^2}\right|_{x=0} = \alpha^2 e^{\alpha x}\Big|_{x=0} = \alpha^2$$

and so on ....

or

$$e^{\alpha x} = 1 + \alpha x + \frac{\alpha^2}{2!}x^2 + \frac{\alpha^3}{3!}x^3 + \frac{\alpha^4}{4!}x^4 + ... = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!}x^n \qquad (19)$$

If we apply this to the sine and cosine functions in the same manner we get (you should do this for at least one of these functions)

$$\sin \alpha x = \sum_{n=0}^{\infty} (-1)^n \frac{\alpha^{2n+1}}{(2n+1)!}x^{2n+1} = \alpha x - \frac{\alpha^3}{3!}x^3 + \frac{\alpha^5}{5!}x^5 + ... \qquad (20)$$

$$\cos \alpha x = \sum_{n=0}^{\infty} (-1)^n \frac{\alpha^{2n}}{(2n)!}x^{2n} = 1 - \frac{\alpha^2}{2!}x^{24} + \frac{\alpha^4}{4!}x^4 + ... \qquad (21)$$

We can then show the neat result that

$$e^{\pm i\alpha t} = \cos\alpha t \pm i\sin\alpha t$$

which will be very useful throughout the course. It is called Euler's formula.

Proof:

$$e^{i\alpha t} = 1 + i\alpha t + \frac{(i\alpha t)^2}{2!} + \frac{(i\alpha t)^3}{3!} + \frac{(i\alpha t)^4}{4!} + ...$$

$$= \left(1 - \frac{(\alpha t)^2}{2!} + \frac{(\alpha t)^4}{4!} - ....\right) + i\left(\alpha t - \frac{(\alpha t)^3}{3!} + \frac{(\alpha t)^5}{5!} - ....\right) \quad (22)$$

$$= \cos\alpha t + i\sin\alpha t$$

and similarly

$$e^{-i\alpha t} = \cos\alpha t - i\sin\alpha t \quad (23)$$

From these results we can also derive the relations

$$\frac{e^{i\alpha t} - e^{-i\alpha t}}{2i} = \frac{\cos\alpha t + i\sin\alpha t - \cos\alpha t + i\sin\alpha t}{2i} = \sin\alpha t$$

$$\frac{e^{i\alpha t} + e^{-i\alpha t}}{2} = \frac{\cos\alpha t + i\sin\alpha t + \cos\alpha t - i\sin\alpha t}{2} = \cos\alpha t \quad (24)$$

Finally, we use these results to solve the equation which results from applying Newton's second law to a simple harmonic oscillator (a spring for example). We have

$$F = -ky \rightarrow M\frac{d^2y}{dt^2} + ky = 0 \rightarrow \frac{d^2y}{dt^2} + \omega^2 y = 0 \quad , \quad \omega^2 = \frac{k}{M}$$

Now we use the exponential substitution method to solve the equation

$$\frac{d^2y}{dt^2} + \omega^2 y = 0 \tag{25}$$

Substituting $y = e^{\alpha t}$ we get the algebraic equation

$$\alpha^2 + \omega^2 = 0 \tag{26}$$

which has solutions (allowed values of α) of

$$\alpha = \pm i\omega \tag{27}$$

so that the most general solution takes the form

$$y(t) = Ae^{i\omega t} + Be^{-i\omega t}$$

This is same equation as for LC circuit earlier

Suppose now that the initial conditions are $y = y_0$ and $dy/dt = 0$ at $t = 0$. Then we have

$$y(t) = Ae^{i\omega t} + Be^{-i\omega t} \Rightarrow y(0) = y_0 = A + B$$

$$\frac{dy}{dt} = i\omega Ae^{i\omega t} - i\omega Be^{-i\omega t} \Rightarrow \frac{dy}{dt}(0) = i\omega A - i\omega B = 0 \rightarrow A - B = 0$$

or

$$A = B = \frac{y_0}{2}$$

and

$$y(t) = y_0 \frac{e^{i\omega t} + e^{-i\omega t}}{2} = y_0 \cos\omega t \tag{28}$$

which correctly corresponds to the motion of a mass M that is released from rest while attached to a spring with spring constant k.

Let us return to the ODE we solved before: $I = -\frac{dQ}{dt} \Rightarrow \frac{d^2Q}{dt^2} + \frac{Q}{LC} = 0$

We have: $Q(t) = e^{\alpha t} \Rightarrow \alpha^2 = -\frac{1}{LC} = -\omega_0^2 \Rightarrow \alpha = \pm i\omega_0$

We then choose solution: $Q(t) = Ae^{i\omega_0 t} + Be^{-i\omega_0 t}$

Using initial conditions: $Q(0) = Q_0 = A + B \quad , \quad \frac{dQ(0)}{dt} = I(0) = 0 = i\omega_0(A - B)$

$$\Rightarrow A = B = \frac{Q_0}{2} \Rightarrow Q(t) = \frac{Q_0}{2}\left(e^{i\omega_0 t} + e^{-i\omega_0 t}\right) = Q_0 \cos\omega_0 t$$

in agreement with our earlier result.

Energy conservation

Energy stored in the capacitor over time:

$$U_C(t) = \frac{Q^2(t)}{2C} = \frac{Q_0^2}{2C}\cos^2\omega_0 t$$

Energy stored in the inductor:

$$U_L(t) = \frac{1}{2}LI^2(t) = \frac{1}{2}L\frac{Q_0^2}{LC}\sin^2\omega_0 t = \frac{Q_0^2}{2C}\sin^2\omega_0 t$$

Total energy:

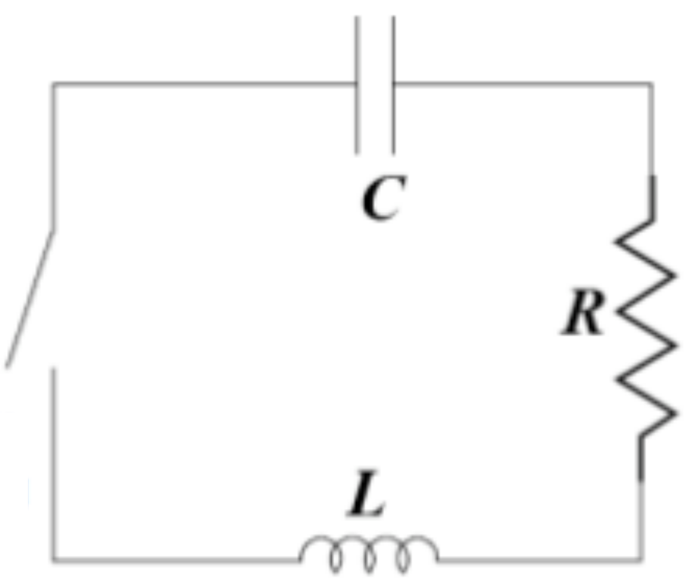
$$U(t) = U_L(t) + U_C(t) = \frac{Q_0^2}{2C}\left(\cos^2\omega_0 t + \sin^2\omega_0 t\right) = \frac{Q_0^2}{2C}$$

What is happening over time?

Energy swings back and forth between C and L but at any moment in time the total energy is equal to the energy initially stored in the capacitor:

Energy is conserved! no dissipative elements in circuit

RCL circuits(1st pass)



LC circuits don't belong to this world:

R is never exactly 0!

So let's concentrate on RCLs

Start with a charged C

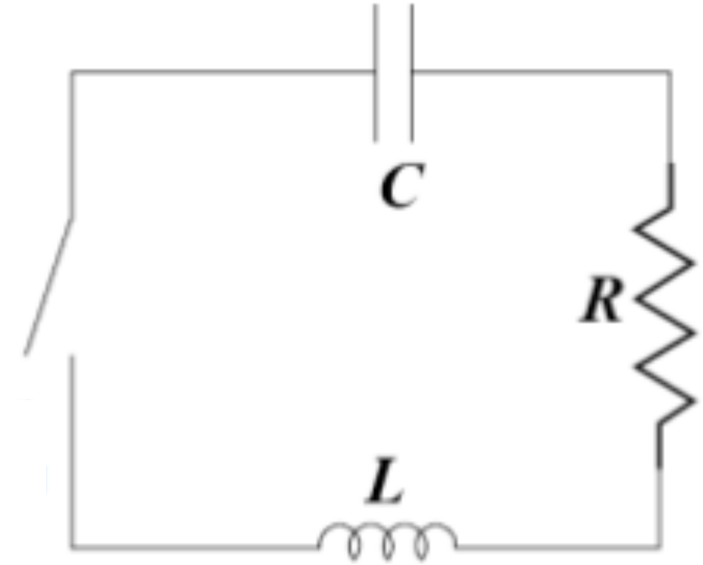
Intuitively:

LC ➙ oscillatory part: sin and cos solution

R ➙ dissipative part: exponential damping

expect damped oscillatory time behavior

RCL circuits: rigorous solution



Use Kirchoff: $$\frac{Q}{C} - IR - L\frac{dI}{dt} = 0$$

Since $$I(t) = -\frac{dQ}{dt} \Rightarrow \frac{d^2Q}{dt^2} + \frac{R}{L}\frac{dQ}{dt} + \frac{Q}{LC} = 0$$

How to solve this equation?

Educated guess!

Intuition tells us that the solution must have an oscillatory term and a damping term

Strategy #1: exponential * sin/cos functions:

$$Q(t) = e^{-t/\tau}\left(A\cos\omega_0 t + B\sin\omega_0 t\right)$$

Very heavy on algebra!!!

Strategy #2: complex exponentials

Idea: the solution is the real part of a complex solution

$$\tilde{Q}(t) = Ae^{i\phi_0}e^{i\alpha\tau} \Rightarrow Q(t) = \mathrm{Re}\left(\tilde{Q}(t)\right)$$

Much easier algebra!!!

Note: α can be complex!

Repeat our earlier work --- more times = better understanding !!!!!!!

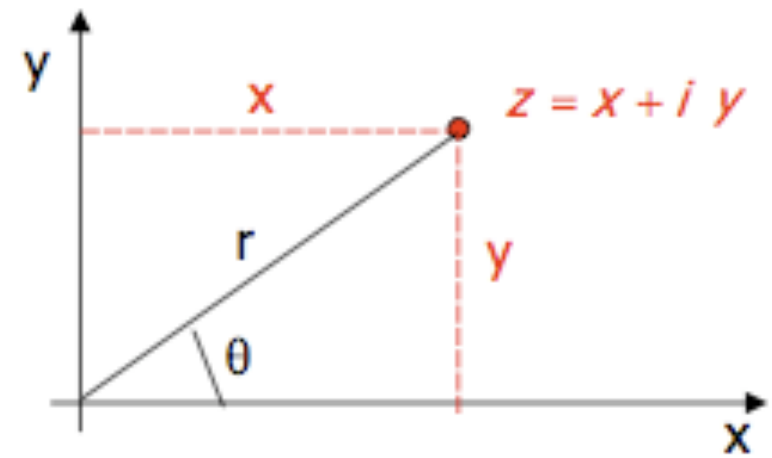
Technique is the same substitution method as before; we substitute the above mathematical form and let the differential equation determine the possible values of the unknown parameters corresponding to a valid solution

Complex number notation (Reminder)

Complex number: number with both a real and an imaginary part

$$z = x + iy \quad with \quad i^2 = -1$$

Complex plane representation z=(x,y)→



Another useful representation

Set magnitude $r = \sqrt{x^2 + y^2}$ and phase $\theta = \tan^{-1} y/x \quad \Rightarrow z = r(\cos\theta + i\sin\theta)$

Given Euler's relation: $e^{i\theta} = \cos\theta + i\sin\theta$ we have

$$z = re^{i\theta}$$ (= Phasor representation)

Continuing with the solution: substitute $\tilde{Q}(t) = Ae^{i\phi_0}e^{i\alpha\tau}$ into the differential equation

$$\frac{d^2\tilde{Q}}{dt^2} + \frac{R}{L}\frac{d\tilde{Q}}{dt} + \frac{\tilde{Q}}{LC} = 0$$

$$\frac{d\tilde{Q}}{dt} = i\alpha\tilde{Q} \quad , \quad \frac{d^2\tilde{Q}}{dt^2} = -\alpha^2\tilde{Q} \Rightarrow \tilde{Q}\left(-\alpha^2 + i\alpha\frac{R}{L} + \frac{1}{LC}\right) = 0$$

$$\Rightarrow -\alpha^2 + i\alpha\frac{R}{L} + \frac{1}{LC} = 0$$

The differential equation has been converted into an algebraic equation!!

The allowed values of α corresponding to valid solutions are given by

$$\alpha_{\pm} = i\frac{R}{2L} \pm \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}$$

Both solutions have the same real part, i.e.,

$$\tilde{Q}_{\pm}(t) = Ae^{i\phi_0} e^{i\left(i\frac{R}{2L} \pm \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}\right)t} = Ae^{i\phi_0} e^{-\frac{R}{2L}t} e^{\pm it\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}}$$

$$Q(t) = \mathrm{Re}\left(\tilde{Q}_{\pm}(t)\right) = Ae^{-\frac{R}{2L}t} \cos\left(\omega t + \phi_0\right) \quad , \quad \omega = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}$$

The weak damping(small R) limit

Weak damping limit: small R➙the damping is small➙several oscillations occur before amplitude starts decreasing in sizable way

$$I(t) = -\frac{dQ}{dt} = Ae^{-\frac{R}{2L}t}\left(\omega \sin\left(\omega t + \phi_0\right) + \frac{R}{2L}\cos\left(\omega t + \phi_0\right)\right)$$

When $\omega >> R/2L$ (weak damping limit), the second term can be ignored and

$$I(t) \approx A\omega e^{-\frac{R}{2L}t} \sin\left(\omega t + \phi_0\right) \qquad \omega = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}} \approx \frac{1}{\sqrt{LC}} = \omega_0$$

➙final solution for "weak damping":

$$Q(t) \approx Ae^{-\frac{R}{2L}t} \cos\left(\omega_0 t + \phi_0\right)$$

$$I(t) \approx A\omega_0 e^{-\frac{R}{2L}t} \sin\left(\omega_0 t + \phi_0\right)$$

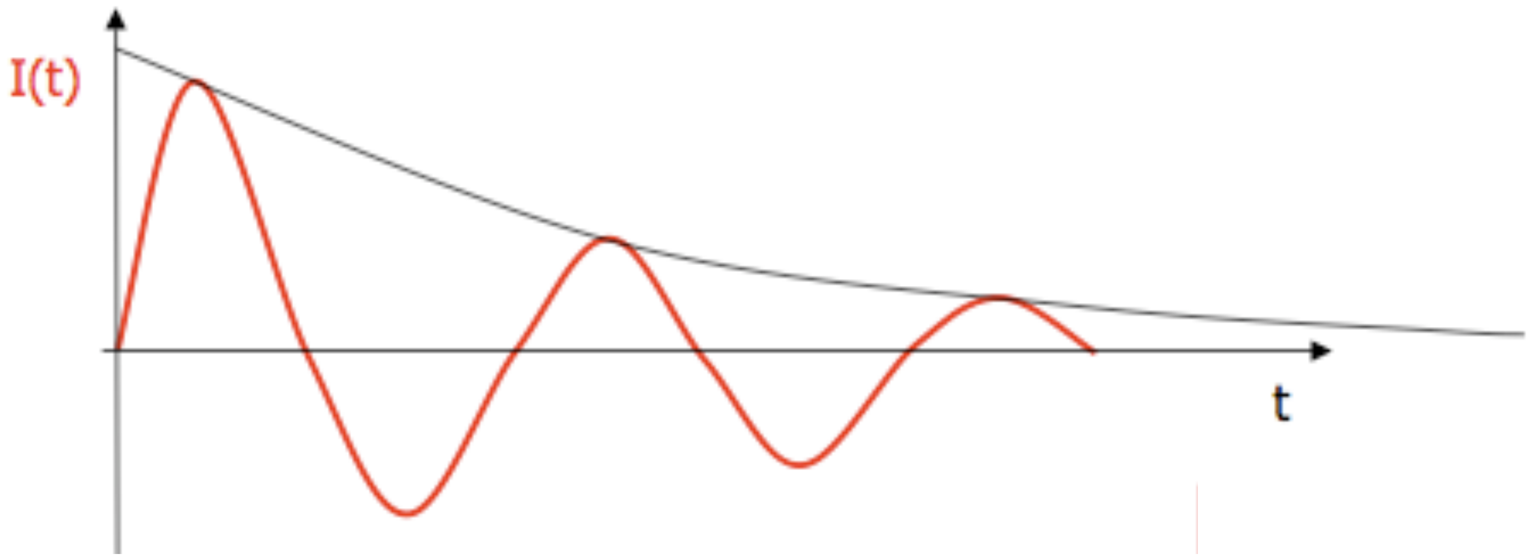
$$\omega_0 = \frac{1}{\sqrt{LC}}$$

RCL in weak damping limit

Initial conditions: $Q(0) = Q_0 = A\cos\phi_0 \quad , \quad I(0) = 0 = A\omega_0 \sin\phi_0$

$$\Rightarrow A = Q_0 \quad , \quad \phi_0 = 0$$

Graphical representation of solution:



What happened when we put inductors into circuits?

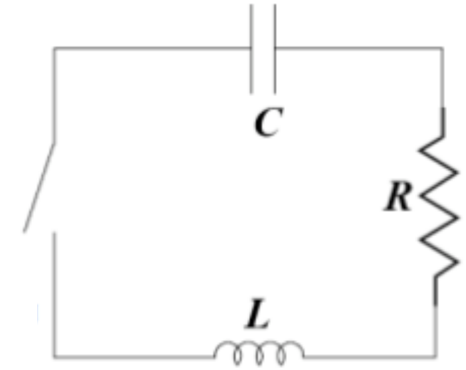
RL circuits:exponential solutions 

LC circuits:oscillatory solution 

RCL circuits:damped oscillation 

RCL circuits are particularly interesting. Let’s see them in some more detail…

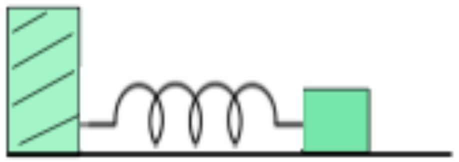
Undriven RCL circuits: recap



Kirchoff's second rule:

$$L\frac{d^2Q}{dt^2}+R\frac{dQ}{dt}+\frac{1}{C}Q=0$$

Does it look familiar?



$$m\frac{d^2x}{dt^2}+k_f\frac{dx}{dt}+k_e x=0$$

Mechanics: damped harmonic oscillator!

RCL	Mechanics	Interpretation
$L\ d^2Q/dt^2$	$ma=m\ d^2x/dt^2$	L ~ m: inertia term
$R\ dQ/dt$	$k_f\ v = k_f\ dx/dt$	R ~ k_f → friction (damping) term
$1/C\ Q$	$k_e\ x$	1/C ~k_e → elastic term due to spring

Undriven RCLs: solution

$$\tilde{Q}_\pm(t)=\tilde{Q}_\pm(0)e^{-\frac{R}{2L}t}e^{\pm it\sqrt{\frac{1}{LC}-\frac{R^2}{4L^2}}}=\tilde{Q}_\pm(0)e^{-\frac{R}{2L}t}e^{\pm t\sqrt{\frac{R^2}{4L^2}-\frac{1}{LC}}}=\tilde{Q}_\pm(0)e^{\beta t}$$

$$\beta=-\frac{R}{2L}\pm\sqrt{\frac{R^2}{4L^2}-\frac{1}{LC}}$$

β *purely real* : $\frac{R^2}{4L^2} - \frac{1}{LC} > 0 \Rightarrow R > 2\sqrt{\frac{L}{C}} \Rightarrow$ *overdamped motion*

β *purely imaginary* $:\Rightarrow R = 0 \quad \Rightarrow$ *undamped LC*

β *truly complex* $: R > 0$ *and* $\frac{R^2}{4L^2} - \frac{1}{LC} < 0 \Rightarrow \tilde{Q}(t) = e^{\beta t} = e^{-\alpha t} e^{i\omega t}$

$$\alpha = \frac{R}{2L} \quad and \quad \omega = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}} \Rightarrow underdamped\ motion$$

When $\frac{R^2}{4L^2} - \frac{1}{LC} = 0$ = critical damping (fastest way to damp an oscillator).

Energy

Energy of the circuit in the weak damping limit:

$$U_C(t) = \frac{Q^2(t)}{2C} = \frac{Q_0^2}{2C} e^{-\frac{Rt}{L}} \cos^2 \omega_0 t$$

$$U_L(t) = \frac{1}{2} L I^2(t) = \frac{1}{2} \omega_0^2 L Q_0^2 e^{-\frac{Rt}{L}} \sin^2 \omega_0 t = \frac{Q_0^2}{2C} e^{-\frac{Rt}{L}} \sin^2 \omega_0 t$$

$$U(t) = U_L(t) + U_C(t) = \frac{Q_0^2}{2C} e^{-\frac{Rt}{L}} \left(\cos^2 \omega_0 t + \sin^2 \omega_0 t\right) = \frac{Q_0^2}{2C} e^{-\frac{Rt}{L}}$$

Since $Q_0^2/2C$ = total energy stored initially in the system

➙U decreases exponentially over time: as expected!

Quality Factor Q (same letter!!)

Definition 1: the quality factor measures how many times the circuit oscillates before it loses a certain amount of energy

In the time $\tau = L/R$ the energy decreases to $e^{-1}U(t=0)$

$$U(t) = \frac{Q_0^2}{2C} e^{-\frac{Rt}{L}}$$

The oscillation = $\omega\tau$ radians $\Rightarrow Q = \omega\tau = \frac{\omega L}{R}$

Definition 2: the quality factor measures the ratio between energy stored (in C and L) and average power dissipated (in R)*ω(dimensionless!)

For an oscillation with frequency ω $\Rightarrow Q = \omega \frac{Energy\ stored}{\langle power \rangle} = \omega \frac{LI_0^2/2}{RI_0^2/2} = \frac{\omega L}{R}$

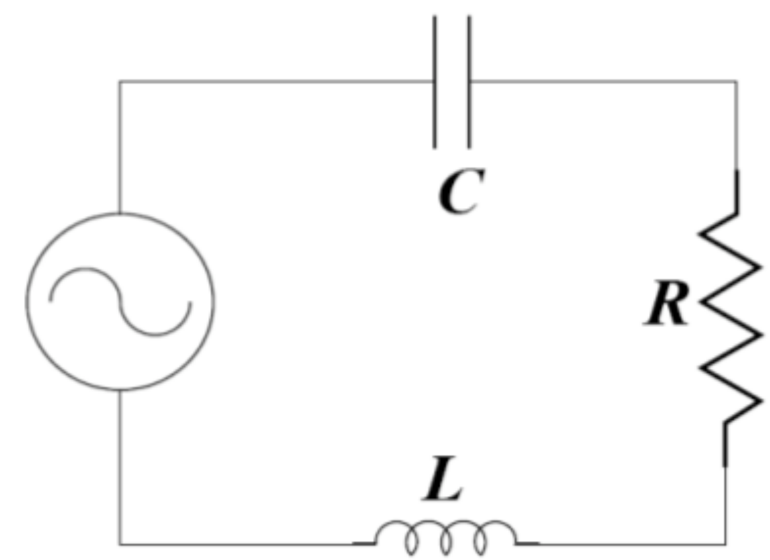
Q factor can be defined for any system that creates vibrations.

Acoustics: Q of a tuning fork is much higher than the Q of a table…

Table	Q ~ 1
Tuning Fork	Q ~ 10^3
Laser	Q ~ 10^{14}

Driven RCL circuits

 is an AC emf



AC voltage supplied to the circuit: $emf(t) = V_0 \cos\omega t$

Convenient assumption: $V(t) = \mathrm{Re}\left(\tilde{V}(t)\right) \quad with \quad \tilde{V}(t) = V_0 e^{i\omega t}$

Note: V_0 is purely real!

How to solve this? Just generalize what we used for DC!

Sum of voltage drops in loop is equal to emf (Kirchoff #2)

$$V_{emf}(t) = V_R(t) + V_C(t) + V_L(t) \Leftrightarrow \tilde{V}_{emf}(t) = \tilde{V}_R(t) + \tilde{V}_C(t) + \tilde{V}_L(t)$$

The same current must pass through every circuit element

$$I(t) = I_R(t) = I_C(t) = I_L(t) \Leftrightarrow \tilde{I}(t) = \tilde{I}_R(t) = \tilde{I}_C(t) = \tilde{I}_L(t)$$

AC current is generated using Faraday's law which we will learn later.

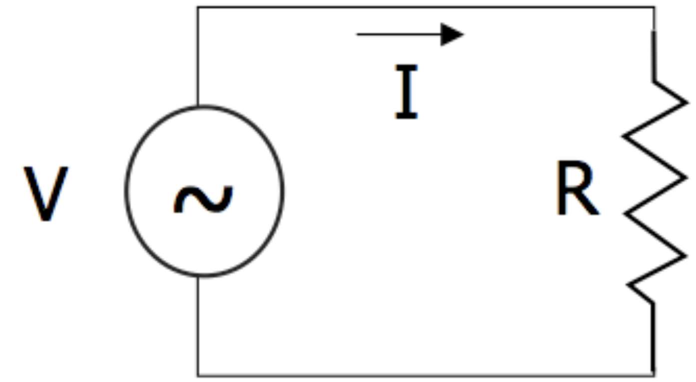
In U.S.: v=60 Hz, ω=377

Let us go back and review some facts we talked about earlier.

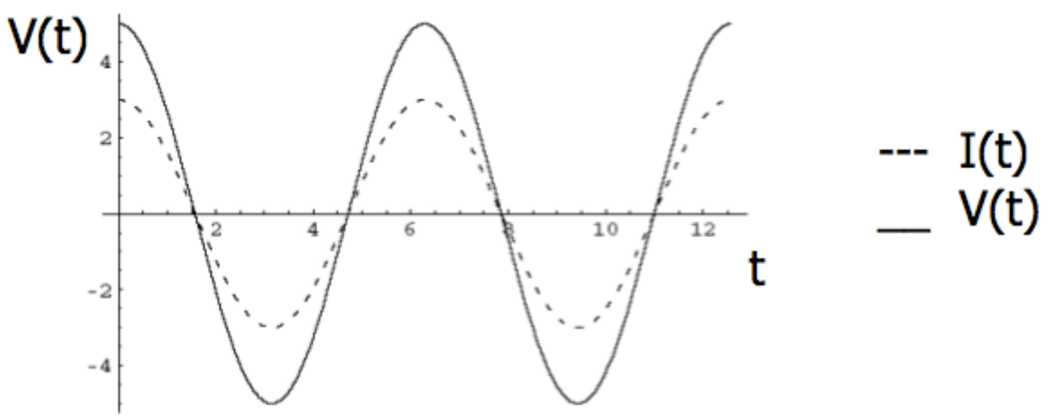
AC emf + resistor R

impedance

Ohm's law holds for AC too: $V(t) = V_R(t) = I(t)R \Leftrightarrow \tilde{V}(t) = \tilde{I}(t)Z_R \Rightarrow Z_R = R(real)$



Let's plot I(t) and V(t) on the same graph:



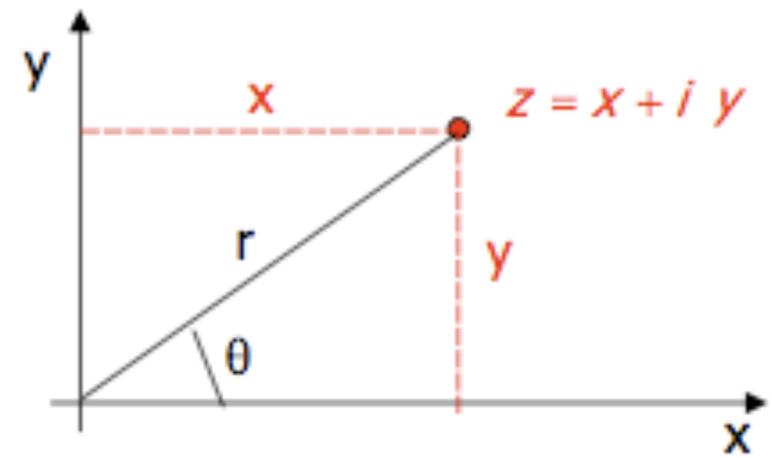
→In a resistor the voltage and the current are in phase
(peak voltage occurs at the same time as peak current)

Reminder: phasor notation

Complex number: number with both a real and an imaginary part

$$z = x + iy \quad with \quad i^2 = -1$$

Complex plane representation z=(x,y)→



Another useful representation

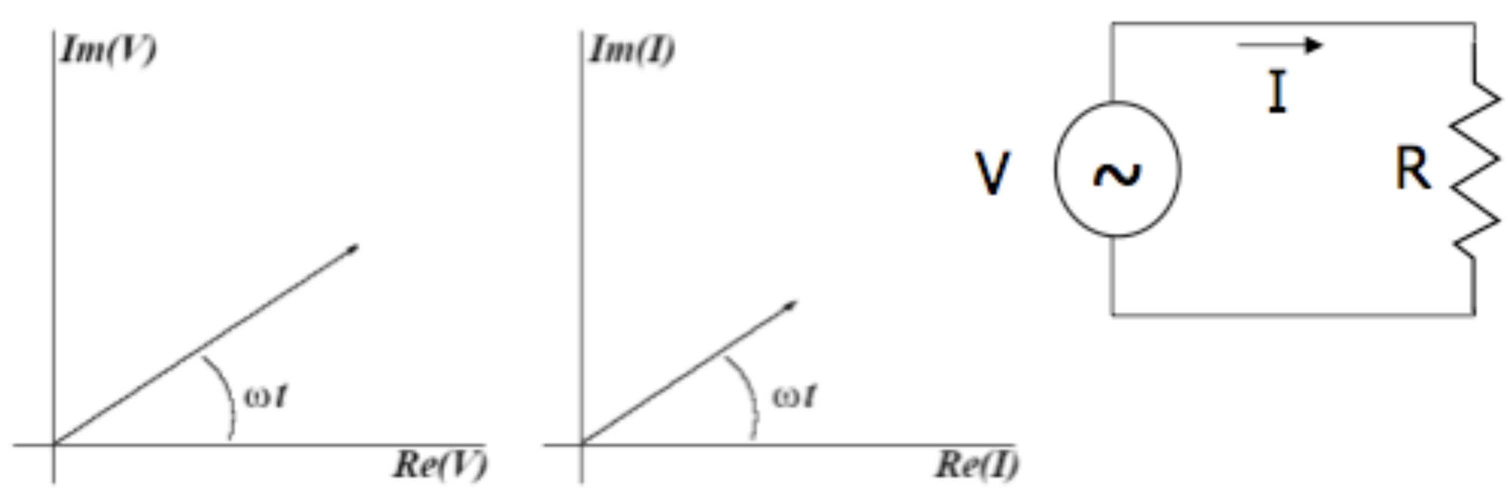
Set magnitude $r = \sqrt{x^2 + y^2}$ and phase $\theta = \tan^{-1} y/x \quad \Rightarrow z = r(\cos\theta + i\sin\theta)$

Given Euler's relation: $e^{i\theta} = \cos\theta + i\sin\theta$ we have

$$z = re^{i\theta}$$ (Phasor representation) $\tilde{V}(t) = phasor = V_0 e^{i\omega t}$

AC emf + R with phasors

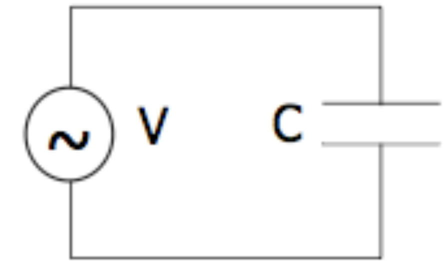
The same information can be represented with phasors in the complex plane:



→In a resistor the voltage and the current are in phase

In phase means that both phasors are at the same angle

AC emf + capacitor C

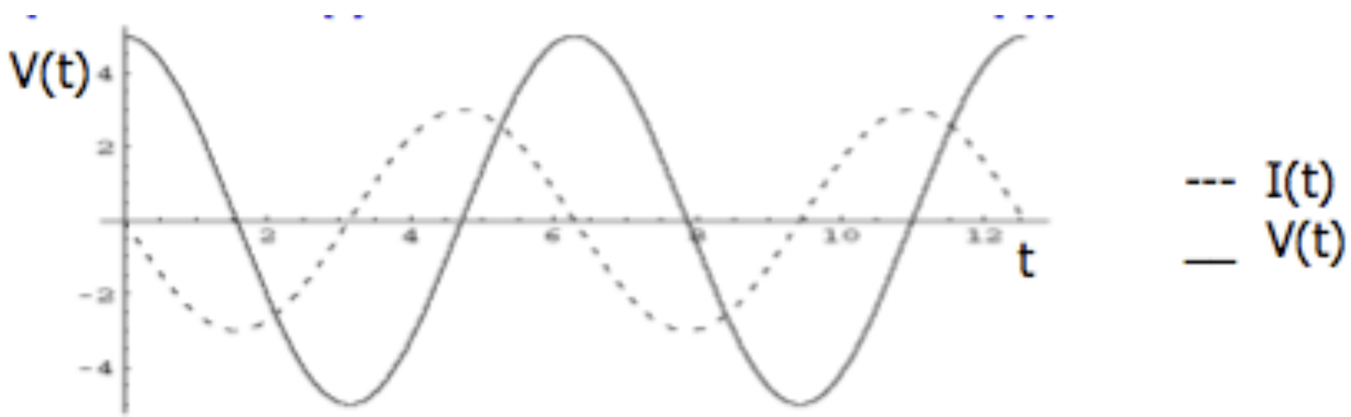


Connect AC emf across a capacitor C: $V(t) = V_C(t) = \dfrac{Q(t)}{C}$

Since V(t) = V_0cosωt and I(t) = dQ/dt:

$$I(t) = \frac{dQ(t)}{dt} = -\omega C V_0 \sin \omega t = \omega C V_0 \cos\left(\omega t + \pi / 2\right)$$

➙I(t) LEADS V(t) by 90 deg or V(t) lags I(t) by 90 deg
(maxima in I(t) occur before maxima in V(t))



Ohm’s law revisited and Impedance

Relation between I(t) and V(t) becomes more obvious when using phasor notation:

$$V_C(t) = V_0 \cos \omega t = \mathrm{Re}\left(\tilde{V}_C(t)\right) \quad with \quad \tilde{V}(t) = V_0 e^{i\omega t}$$

For the current:

$$I_C(t) = \omega C V_0 \cos\left(\omega t + \pi / 2\right) = \mathrm{Re}\left(\tilde{I}_C(t)\right)$$

$$with \quad \tilde{I}(t) = \omega C V_0 e^{i(\omega t + \pi/2)} = i\omega C V_0 e^{i\omega t} \quad (remember \quad e^{i\pi/2} = i)$$

$$\cos\frac{\pi}{2} + i \sin\frac{\pi}{2} = 0 + i(1) = i$$

Combining complex currents and voltages we can write:

$\tilde{V}(t) = \tilde{I}(t) Z_C$ (complex equivalent of Ohm's law)

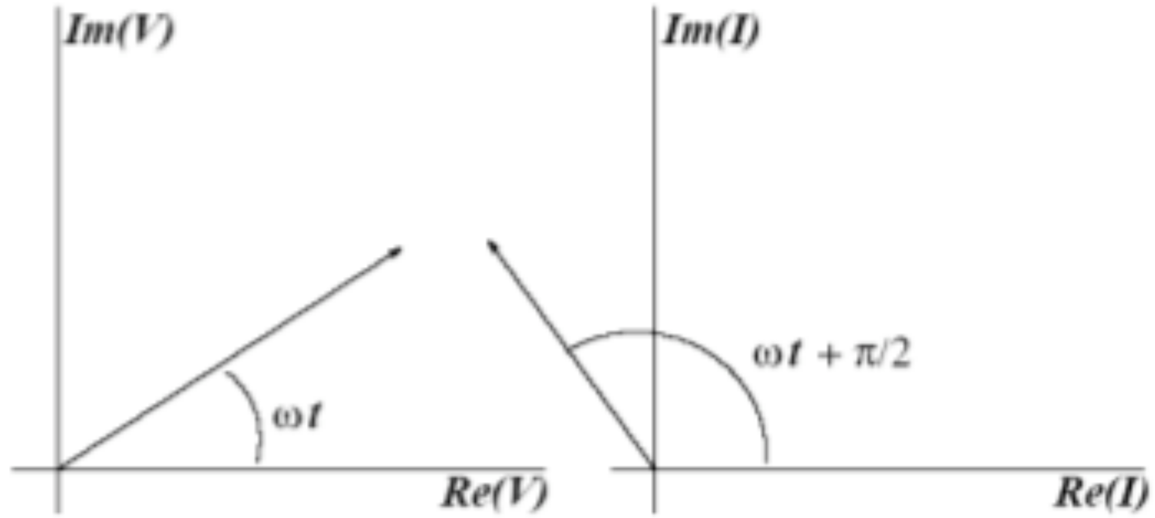
where Z_C is the impedance of a capacitor:

$$Z_C = \frac{1}{i\omega C}$$

AC emf + C: phasor representation

Given $\tilde{V}(t) = V_0 e^{i\omega t}$ *and* $\tilde{I}(t) = Z_C V_0 e^{i\omega t} = i\omega C V_0 e^{i\omega t}$

V(t) and I(t) can easily be represented in the complex plane:

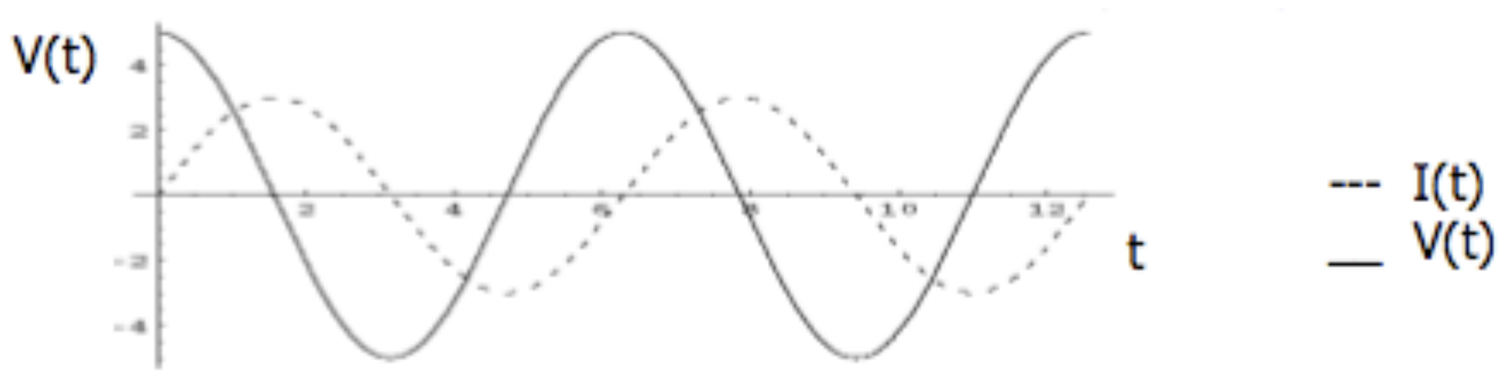


Note: I(t) is ahead of V(t) by 90 degrees: I(t) leads V(t) by 90 degrees

AC emf + inductor L

Connect AC emf across an inductor L: $V(t) = V_L(t) = L\dfrac{dI}{dt}$

Since $V(t)=V_0\cos\omega t$: $\dfrac{dI}{dt} = \dfrac{V_0}{L}\cos\omega t \Rightarrow I(t) = \dfrac{V_0}{\omega L}\sin\omega t = \dfrac{V_0}{\omega L}\cos(\omega t - \pi/2)$



→I(t) LAGS V(t) by 90 degrees, or V(t) LEADS I(t) by 90 degrees
(maxima in I(t) occur before maxima in V(t))

Impedance of inductors

Using phasor notation: $V_L(t) = V_0 \cos\omega t = \mathrm{Re}\left(\tilde{V}_L(t)\right) \quad with \quad \tilde{V}(t) = V_0 e^{i\omega t}$

The current is:

$$I(t) = \frac{V_0}{\omega L}\cos(\omega t - \pi/2) = \mathrm{Re}\left(\tilde{I}(t)\right)$$

$$with \quad \tilde{I}(t) = \frac{V_0}{\omega L} V_0 e^{i(\omega t - \pi/2)} = \frac{V_0}{i\omega L} e^{i\omega t} \quad (remember \quad e^{-i\pi/2} = -i)$$

Combining complex currents and voltages we can write:

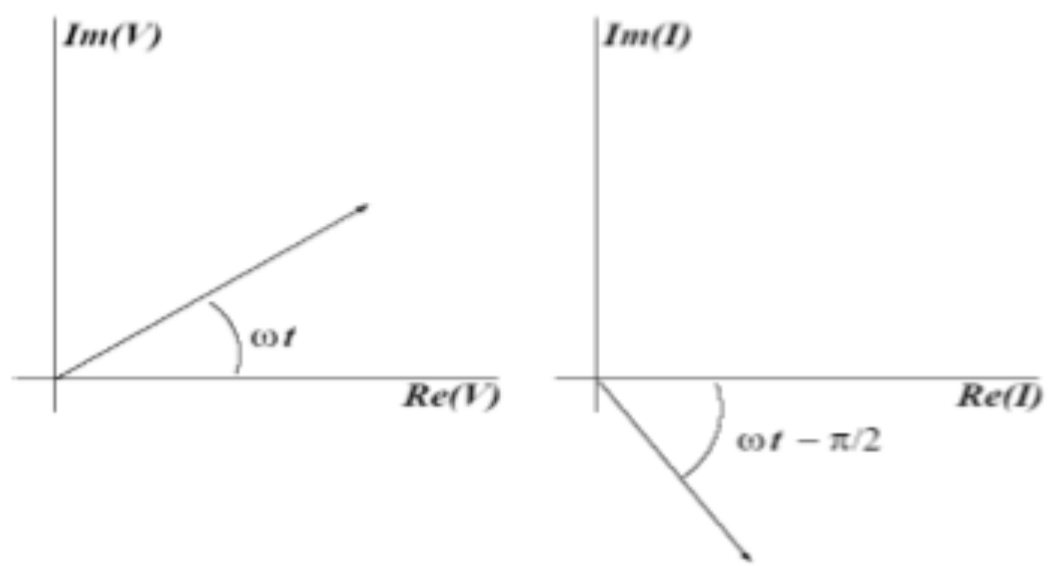
$\tilde{V}(t) = \tilde{I}(t) Z_L$ (complex equivalent of Ohm's law)

where Z_L is the impedance of an inductor: $Z_L = i\omega L$

AC emf + L: phasor representation

Given $\tilde{V}(t) = V_0 e^{i\omega t} \quad and \quad \tilde{I}(t) = Z_L V_0 e^{i\omega t} = \frac{V_0}{i\omega L} e^{i\omega t}$

V(t) and I(t) can easily be represented in the complex plane:



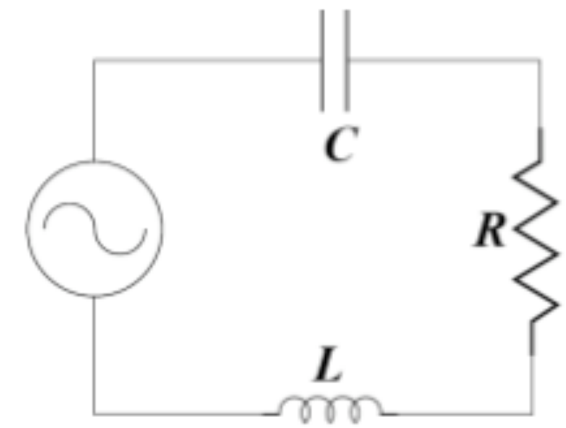
Note: I(t) is 90 degrees behind V(t): I(t) lags V(t) by 90 degrees

Driven RCLs using impedance

Impedance simplifies the study of driven RCL circuits

Let's work with complex numbers and use Ohm's and Kirchoff's extensions

$$\tilde{V}_{emf}(t) = \tilde{V}_R(t) + \tilde{V}_C(t) + \tilde{V}_L(t)$$



Since
$$\begin{cases} \tilde{V}_R(t) = R\tilde{I}(t) \\ \tilde{V}_C(t) = Z_C\tilde{I}(t) = \dfrac{1}{i\omega C}\tilde{I}(t) \\ \tilde{V}_L(t) = Z_L\tilde{I}(t) = i\omega L\tilde{I}(t) \end{cases}$$

$$\Rightarrow \tilde{V}_{emf}(t) = \tilde{I}(t)\left(R + i\left(\omega L - \frac{1}{\omega C}\right)\right) = \tilde{I}(t)\tilde{Z}_{tot}$$

where the total impedance of the circuit is $\tilde{Z}_{tot} = R + i\left(\omega L - \dfrac{1}{\omega C}\right)$

Driven RCLs: phasor notation

The complex current can be written as $\tilde{I}(t) = \dfrac{\tilde{V}_{emf}(t)}{\tilde{Z}_{tot}} = \dfrac{V_0 e^{i\omega t}}{R + i\left(\omega L - \dfrac{1}{\omega C}\right)}$

This can be written as:

$$\tilde{I}(t) = \frac{\tilde{V}_{emf}(t)}{\tilde{Z}_{tot}} = \frac{V_0 e^{i\omega t}}{\tilde{Z}_{tot}\tilde{Z}^*_{tot}}\tilde{Z}^*_{tot} = \frac{V_0 e^{i\omega t}}{R^2 + \left(\omega L - \dfrac{1}{\omega C}\right)^2}\left(R - i\left(\omega L - \frac{1}{\omega C}\right)\right) = I_0 e^{i\omega t} e^{-i\phi}$$

where remembering that $e^{i\theta} = \cos\theta + i\sin\theta$ we have $I_0 = \dfrac{V_0}{\sqrt{R^2 + \left(\omega L - \dfrac{1}{\omega C}\right)^2}}$ and

$$\tan\phi = \frac{\omega L - \dfrac{1}{\omega C}}{R} = \frac{\omega L}{R} - \frac{1}{\omega RC}$$

Derivation on next page....

$$\tilde{I}(t)=\frac{\tilde{V}_{emf}(t)}{\tilde{Z}_{tot}}=\frac{V_0e^{i\omega t}}{R+i\left(\omega L-\frac{1}{\omega C}\right)}=\frac{V_0e^{i\omega t}}{R+i\left(\omega L-\frac{1}{\omega C}\right)}\frac{R-i\left(\omega L-\frac{1}{\omega C}\right)}{R-i\left(\omega L-\frac{1}{\omega C}\right)}$$

$$\tilde{I}(t)=\frac{V_0e^{i\omega t}}{R^2+\left(\omega L-\frac{1}{\omega C}\right)^2}\left(R-i\left(\omega L-\frac{1}{\omega C}\right)\right)=I_0e^{i\omega t}e^{-i\phi}=ze^{i\omega t}$$

$$z=x-iy=I_0e^{-i\phi}=\frac{V_0}{R+i\left(\omega L-\frac{1}{\omega C}\right)}\left(R-i\left(\omega L-\frac{1}{\omega C}\right)\right)$$

$$\Rightarrow x=\frac{V_0}{R^2+\left(\omega L-\frac{1}{\omega C}\right)^2}R=I_0\cos\phi \quad , \quad y=\left(\omega L-\frac{1}{\omega C}\right)\frac{V_0}{R^2+\left(\omega L-\frac{1}{\omega C}\right)^2}=I_0\sin\phi$$

$$\Rightarrow I_0=\sqrt{R^2+\left(\omega L-\frac{1}{\omega C}\right)^2} \quad , \quad \tan\phi=\frac{\omega L-\frac{1}{\omega C}}{R}$$

Full blown ODE derivation with ODE

$$V_{emf}(t) = V_0 e^{i\omega t} = L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C}Q$$

$$Q(t) = Q_h(t) + Q_p(t)$$

$$\frac{d^2Q_h}{dt^2} + \frac{R}{L}\frac{dQ_h}{dt} + \frac{1}{LC}Q_h = 0 \qquad \text{Homogeneous solution}$$

$$\frac{d^2Q_p}{dt^2} + \frac{R}{L}\frac{dQ_p}{dt} + \frac{1}{LC}Q_p = \frac{V_0}{L}e^{i\omega t} \qquad \text{Particular solution}$$

$$Q_h(t) = A_+ e^{-\beta_+ t} + A_- e^{-\beta_- t} \quad , \quad \beta_\pm = -\frac{R}{2L} \pm \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} \quad , \quad \frac{1}{LC} = \omega_0^2$$

$$Q_p(t) = A(\omega)e^{i\omega t}$$

Particular solution oscillates with driving frequency after Homogeneous solution damps out.

$$\frac{d^2Q_p}{dt^2} + \frac{R}{L}\frac{dQ_p}{dt} + \frac{1}{LC}Q_p = \frac{V_0}{L}e^{i\omega t} \Rightarrow \frac{V_0}{L} = \left(-\omega^2 + i\omega\frac{R}{L} + \frac{1}{LC}\right)A(\omega)$$

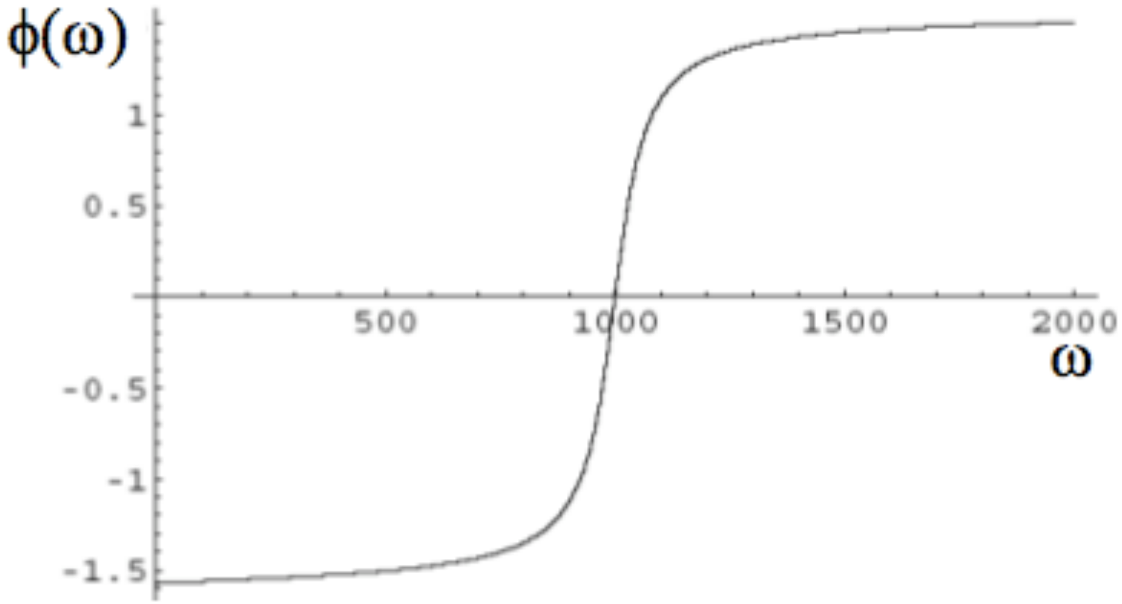
$$\Rightarrow A(\omega) = \frac{V_0}{L}\frac{1}{-\omega^2 + i\omega\frac{R}{L} + \frac{1}{LC}} = \frac{V_0}{L}\frac{1}{-\omega^2 + i\omega\frac{R}{L} + \frac{1}{LC}}\frac{-\omega^2 - i\omega\frac{R}{L} + \frac{1}{LC}}{-\omega^2 - i\omega\frac{R}{L} + \frac{1}{LC}}$$

$$A(\omega) = \frac{V_0}{L}\frac{-\omega^2 - i\omega\frac{R}{L} + \omega_0^2}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + \frac{\omega^2 R^2}{L^2}}}\frac{1}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + \frac{\omega^2 R^2}{L^2}}}$$

Now $$\frac{-\omega^2 - i\omega\frac{R}{L} + \omega_0^2}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + \frac{\omega^2 R^2}{L^2}}} = e^{i\delta} \Rightarrow A(\omega) = \frac{V_0}{L}e^{i\delta}\frac{1}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + \frac{\omega^2 R^2}{L^2}}}$$

Now $$\frac{1}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + \frac{\omega^2 R^2}{L^2}}} = \frac{1}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}} \Rightarrow \text{Same as earlier......,}$$

Dependence of ϕ on ω



$$\tan\phi = \frac{\omega L}{R} - \frac{1}{\omega RC}$$

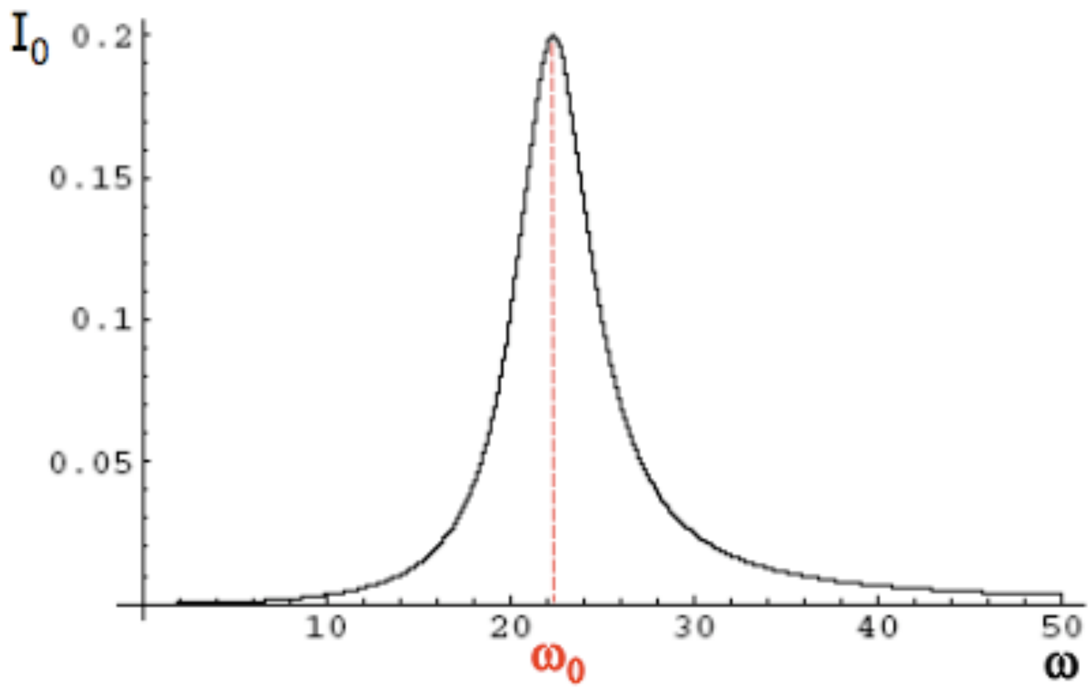
Note:

$$\tilde{I}(t) = I_0 e^{i\omega t} e^{-i\phi} \qquad I(t) = \text{Re}(\tilde{I}(t)) = I_0 \cos(\omega t - \phi)$$

➙high ω: I lags voltage

➙low ω: I leads voltage

Dependence of I_0 on ω



$$I_0 = \frac{V_0}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}$$

Maximum current when $\omega L = \frac{1}{\omega C}$

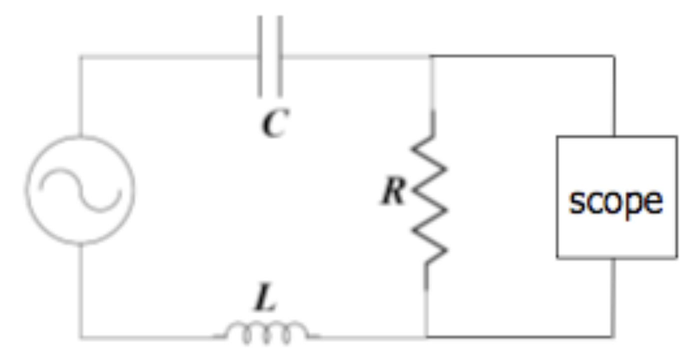
$$\Rightarrow \omega_0 = \frac{1}{\sqrt{LC}}$$ resonance frequency

RCL resonance

RCL circuit driven with variable frequency ω

L=60 mH

C=0.3 μF



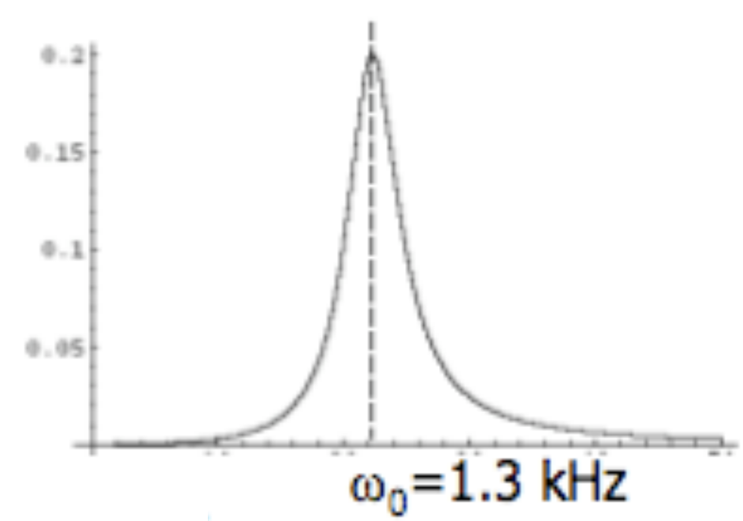
Tune frequency to maximize V_R on scope

What is the expect resonance frequency?

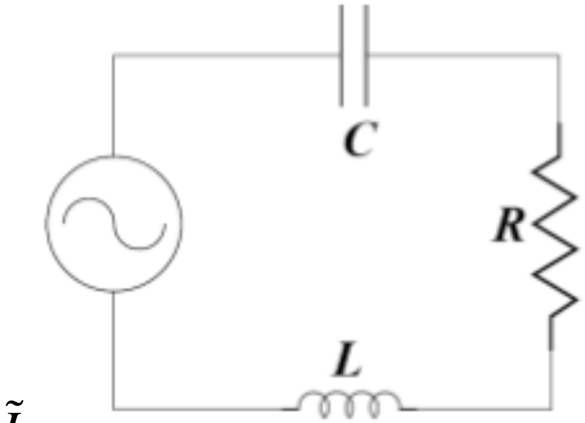
$$\omega_0 = \frac{1}{\sqrt{LC}} = 7 \times 10^3 \Rightarrow \nu = 1.2\,kHz$$

Display V_R vs ω on the scope while sweeping

What do you expect to see?



RCL circuits(AGAIN): More is better.



Reminder: AC driven RCLs

Simple solution when introducing following rules:

Work with complex V and I

Real currents and voltages are just the real part of the $\tilde{V}$ and $\tilde{I}$

Generalization of Ohm's law to complex V and I:

$$\tilde{V}(t) = Z_X \tilde{I}(t)$$

where Z_X is the impedance of component X:

$$\begin{cases} Z_R = R \\ Z_C = \dfrac{1}{i\omega C} \\ Z_L = i\omega L \end{cases}$$

Analyze circuit as if it were DC with only resistors

Take the real part of I(t) and V(t)

The End.

"Analyze as DC with only resistors"

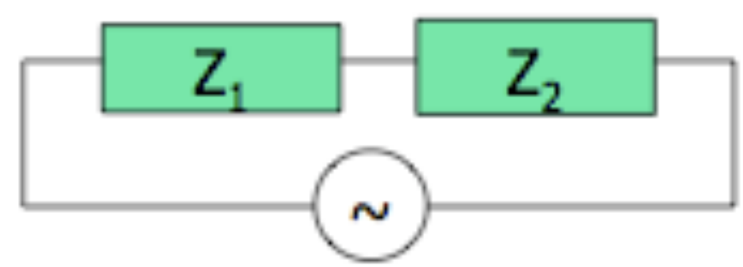
What do I mean with this statement?

Impedances in series

Same current flowing in each element

$I_1 Z_1 = V_1$; $I_2 Z_2 = V_2$; $V_1 + V_2 = V$; $V = ZI$

→$Z_{eq} = Z_1 + Z_2$

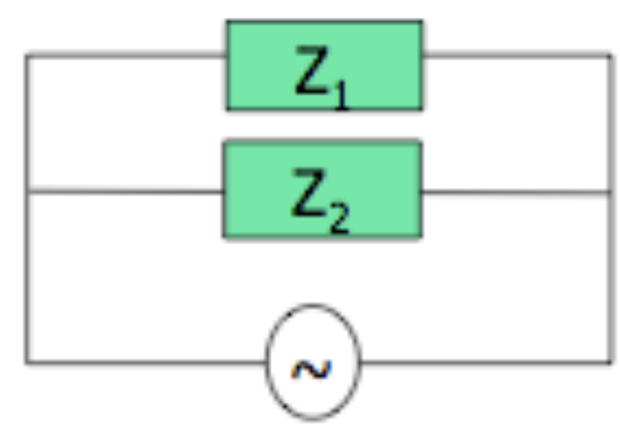


Impedances in parallel

Same voltage drop across each element

$V_1/Z_1 = V_2/Z_2 = V/Z_{eq}$; $V_1 = V_2 = V$

→$1/Z_{eq} = 1/Z_1 + 1/Z_2$

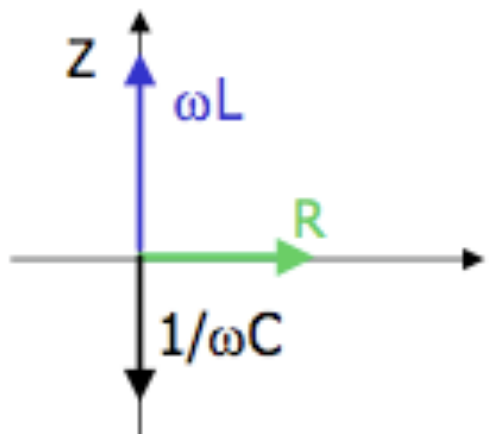


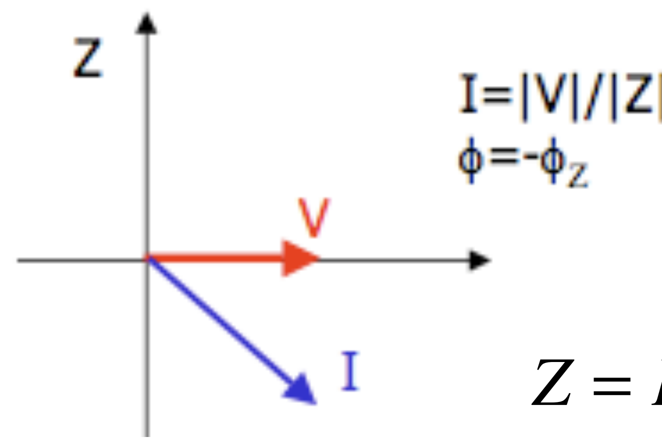
→ Same rules as resistors in series and parallel!

(but are complex impedances)

Is the current leading or lagging?

- Instead of thinking of the problems in terms of complex currents, think in terms of complex impedance!
  - Generalized Ohm's law:
    - All that we really care about is amplitude of I and relative phase between I and V
  - Trick: let's choose V real (no law against it!) and draw the complex I, V and Z in the complex plane





$$Z = R + i\left(\omega L - \frac{1}{\omega C}\right) = |Z|\, e^{i\phi_z}$$

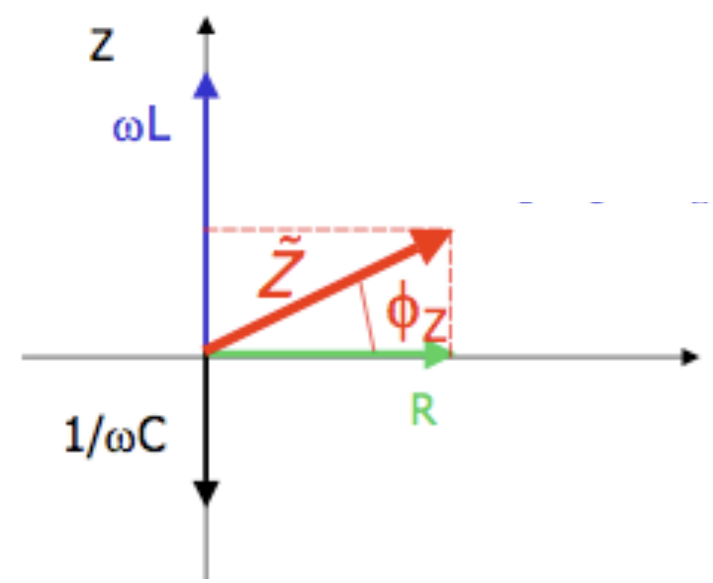
- Consider the complex impedance:
  - Real part: only R contributes
  - Imaginary part: Z_L "pulls up" by ωL and Z_C pulls down by 1/ωC
    - ➞The phase of Z will depend on who prevails:

$Z_L > Z_C$ ➞ $\phi_Z > 0$

$Z_L < Z_C$ ➞ $\phi_Z < 0$

$$|Z| = \sqrt{(\operatorname{Re} Z)^2 + (\operatorname{Im} Z)^2} = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}$$

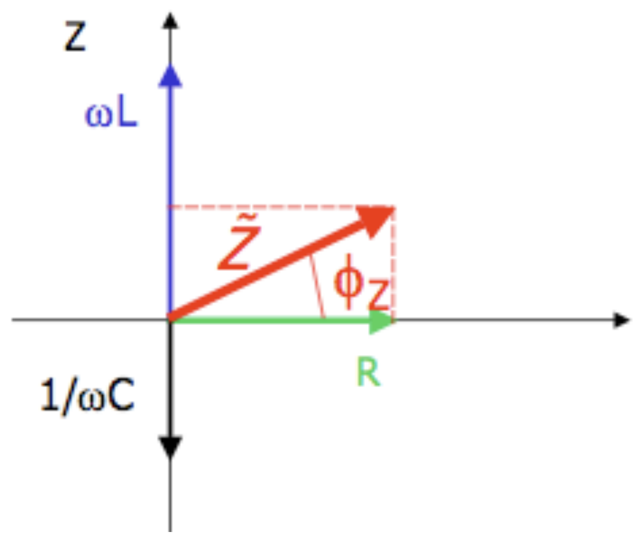
$$\tan\phi_Z = \frac{\operatorname{Im} Z}{\operatorname{Re} Z} = \frac{\omega L - \frac{1}{\omega C}}{R}$$

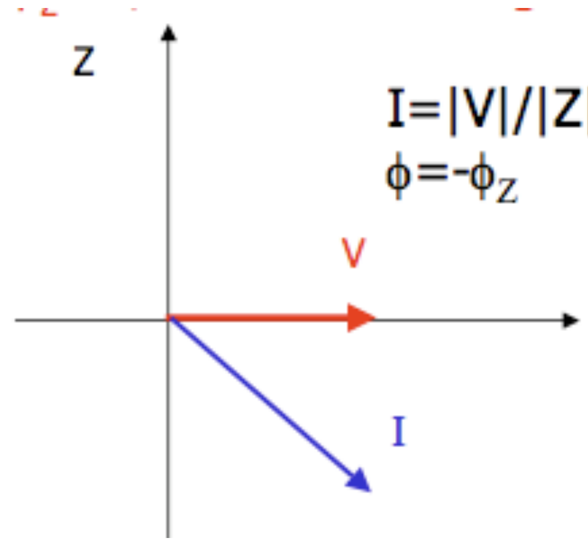


Now remember that $\tilde{V}(t)=\tilde{I}(t)\tilde{Z}_C$ and that we chose a real V:

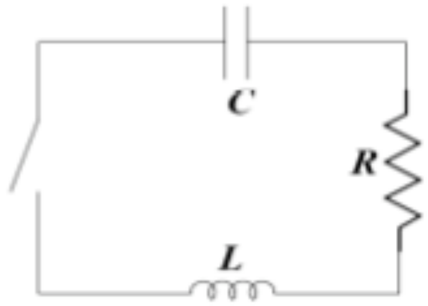
$$\tilde{I}(t)=\frac{\tilde{V}(t)}{\tilde{Z}_C}=\frac{V(t)}{|\tilde{Z}_C|}e^{-i\phi_Z}\Rightarrow$$

if $\phi_Z > 0$, I will be lagging V
if $\phi_Z < 0$, I will be leading V





Power in RCL circuits



Power delivered in a circuit is $P(t)=V(t)I(t)$

Given $V(t)=V_0\cos\omega t$, $I(t)=I_0\cos(\omega t-\phi)$

The average power over a period T will be

$$\langle P\rangle=\frac{1}{T}\int_T V(t)I(t)dt=\frac{\omega}{2\pi}\int_T V_0\cos\omega t I_0\cos(\omega t-\phi)dt$$

$$=\frac{\omega}{2\pi}\frac{V_0^2}{|Z|}\int_T\cos\omega t\cos(\omega t-\phi)dt$$

Note: when we say light bulb has a P of 100W we are referring to <P>

Using the identity: cos(α–β)=cosαcosβ+sinαsinβ we obtain:

$$\langle P\rangle=\frac{\omega}{2\pi}\frac{V_0^2}{|Z|}\left[\int_0^{2\pi/\omega}\cos^2\omega t\cos\phi dt+\int_0^{2\pi/\omega}\cos\omega t\sin\omega t\sin\phi dt\right]$$

Since: $\begin{cases} \dfrac{\omega}{2\pi}\displaystyle\int_0^{2\pi/\omega} \cos^2 \omega t \cos\phi dt = \dfrac{1}{2}\cos\phi \\ \dfrac{\omega}{2\pi}\displaystyle\int_0^{2\pi/\omega} \cos\omega t \sin\omega t \sin\phi dt = 0 \end{cases}$ $\Rightarrow \langle P \rangle = \dfrac{1}{2}\dfrac{V_0^2}{|Z(\omega)|}\cos\phi$

Note: Power depends on relative phase between I and V

cosϕ=0 ➙ no power dissipated in the circuit no work done!

cosϕ=0 when ϕ= 90° ➙ when Z is purely imaginary: R needed!

Introducing: RMS (root mean squared) voltage and currents:

$$V_{RMS} = \frac{V_0}{\sqrt{2}} \quad and \quad I_{RMS} = \frac{I_0}{\sqrt{2}}$$

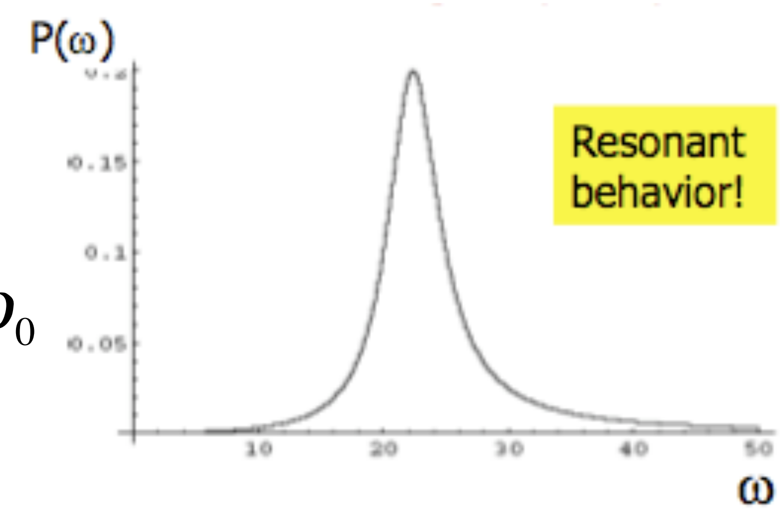
Note: in the US: outlet voltage is 120 V. This is the RMS voltage: V_{max}=170

$$\Rightarrow \langle P \rangle = \frac{V_{RMS}^2}{|Z(\omega)|}\cos\phi = RI_{RMS}^2(\omega)$$ remembering that $$\cos\phi = \frac{R}{|Z(\omega)|}$$

Power vs. frequency

Note: Z depends on ω ➙ power dissipated depends on driving frequency!

$$\langle P \rangle = \frac{V_{RMS}^2}{|Z(\omega)|^2}R = \frac{V_{RMS}^2}{R^2 + \left(\omega L - \dfrac{1}{\omega C}\right)^2}R$$



Resonant behavior!

At what ω is P is max? $\omega L - \dfrac{1}{\omega C} = 0 \Rightarrow \omega = \dfrac{1}{\sqrt{LC}} = \omega_0$

What is the max P? $P_{max} = \dfrac{V_{RMS}^2}{R}$

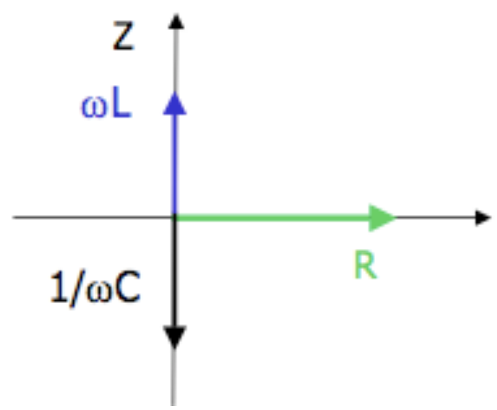
What is the corresponding phase? Zero: imaginary part due to C and L exactly cancel out!

ω_0 in term of L and C

What does $\omega=\omega_0$ mean in terms of L and C?

Remember: $$\omega_0 = \frac{1}{\sqrt{LC}} \Leftrightarrow \omega L = \frac{1}{\omega C}$$

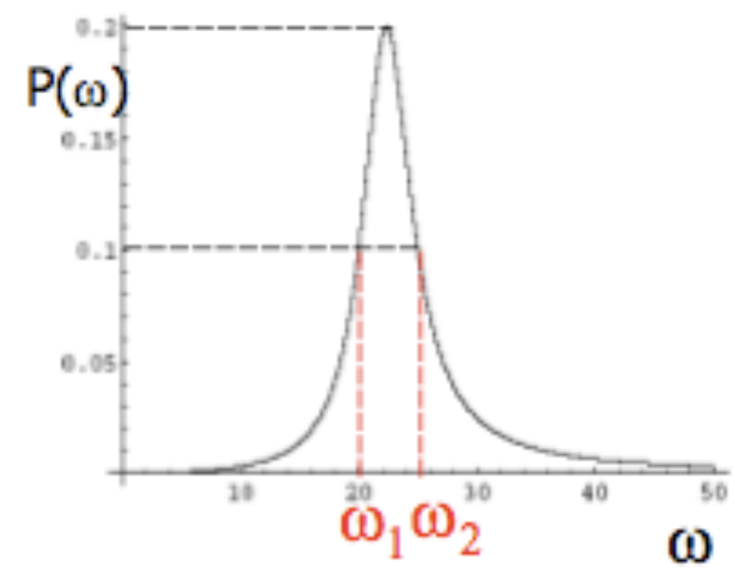
Back to the phasor representation for Z



The imaginary part due to C exactly compensates the one due to L
➙Z is purely real!

How good is the resonant system?

Definition: width of resonance wrt the height
Width: $\Delta\omega$ between the points where the power goes to $P_{max}/2$: ω_1 and ω_2

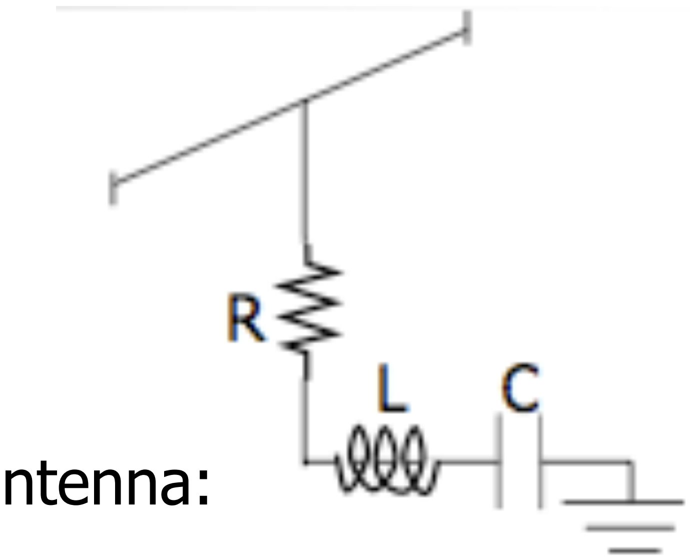


$$\frac{V_{RMS}^2}{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2} R = \frac{V_{RMS}^2}{2R} \Rightarrow \left|\omega L - \frac{1}{\omega C}\right| = \pm R$$

$$\begin{cases} \omega_1 L - \frac{1}{\omega_1 C} = -R \\ \omega_2 L - \frac{1}{\omega_2 C} = +R \end{cases} \Rightarrow \begin{cases} \omega_1^2 LC + RC\omega_1 - 1 = 0 \\ \omega_2^2 LC - RC\omega_2 - 1 = 0 \end{cases}$$

$$\left\{\begin{aligned} \omega_1 &= \frac{-RC+\sqrt{R^2C^2+4LC}}{2LC} \\ \omega_2 &= \frac{RC+\sqrt{R^2C^2+4LC}}{2LC} \end{aligned}\right\} \Rightarrow \Delta\omega = \omega_2 - \omega_1 = \frac{R}{L} \Rightarrow Q = \frac{\omega_{res}}{\Delta\omega} = \frac{L\omega_0}{R}$$

Application: FM antenna: LC Resonant circuit



Consider the following circuit:
L=8.22 μH
C=0.27 pF=0.27x10-12F
R=377 Ω(match air intrinsic impedance))

The radio signal in the air induces an alternating emf in the antenna:

V_{RMS}=9.13 μV

Find frequency of incoming wave for which antenna is in tune

Resonance frequency: $\omega_0 = \frac{1}{\sqrt{LC}} = 6.7\times 10^8$

$$\omega_0 = 2\pi\nu \Rightarrow \nu_0 = \frac{\omega_0}{2\pi} = 106\,MHz \rightarrow FM\ \ Radio$$

Calculate I_{RMS}

$$I_{RMS} = \frac{I_0}{\sqrt{2}} = \frac{V_{RMS}}{|Z_0|} = \frac{V_{RMS}}{R}$$

Note: at resonance $|Z_0|$=R

ΔV_{RMS} across C $\quad V_C = I_{RMS}Z_C = \frac{1}{\omega C}\frac{V_{RMS}}{R} = 0.66\,mV$

Question: V_C = 0.66 mV while V = 9 μ V. How can this happen?
L and C cancel almost perfectly ➙ Z can be small while C and L
are large and Z ~ real. Note: all circuits with good Q value have this feature!

Calculate width of resonance

$$\Delta\omega = \frac{R}{L} = 9\times 10^6 \Rightarrow \Delta\nu = \frac{\Delta\omega}{2\pi} = 1.4\ MHz$$

Question: is this a good antenna?

No, since separation between stations is~ 0.2 MHz

Q factor

$$Q = \frac{L\omega_0}{R}$$ = 73 = good, but not enough for a radio.

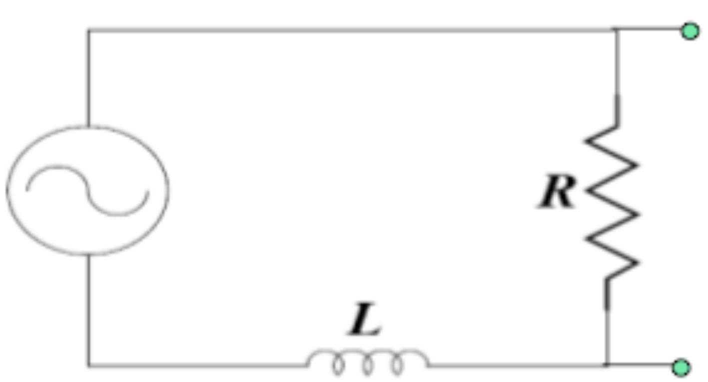
How can this be improved?

Can we increase L? No, it would change frequency...

Can we decrease R? No: R = 377 Ohms (match intrinsic impedance)

--> increase L and decrease C to keep LC constant!

Low pass RL filter



RCL circuits have a frequency dependent response:

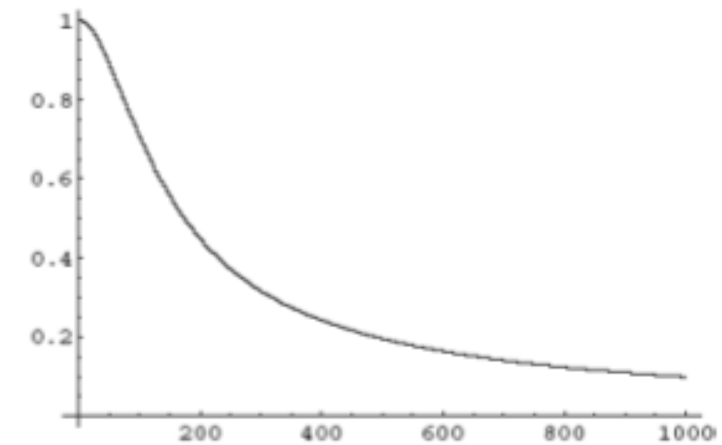
they can act as filters (select only certain frequencies)

Example: RL circuit

Calculate the complex current

$$\tilde{I} = \frac{\tilde{V}}{\tilde{Z}} = \frac{\tilde{V}}{R+i\omega L} \Rightarrow |I| = \frac{|\tilde{V}|}{|R+i\omega L|} = \frac{V_0}{\sqrt{R^2+\omega^2 L^2}} \Rightarrow V_R = |I|R = \frac{V_0 R}{\sqrt{R^2+\omega^2 L^2}}$$

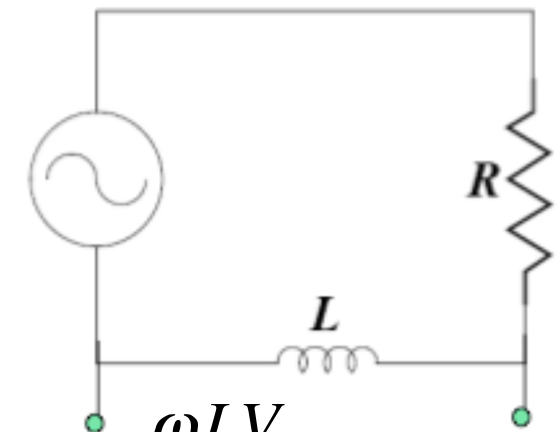
$$\Rightarrow \begin{Bmatrix} \omega\to 0 : V_R \to V_0 \\ \omega\to\infty : V_R \to 0 \end{Bmatrix} \Rightarrow low\ pass\ filter$$



High pass RL filter

What if we take the voltage V_L across the inductor?

Same complex current



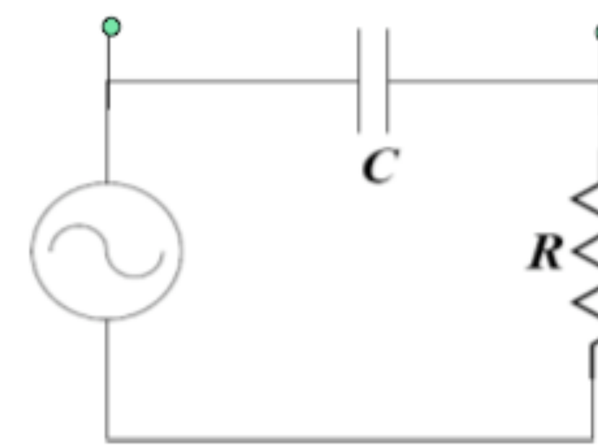
$$\tilde{I} = \frac{\tilde{V}}{\tilde{Z}} = \frac{\tilde{V}}{R + i\omega L} \Rightarrow |I| = \frac{|\tilde{V}|}{|R + i\omega L|} = \frac{V_0}{\sqrt{R^2 + \omega^2 L^2}} \Rightarrow |V_L| = \omega L |I| = \frac{\omega L V_0}{\sqrt{R^2 + \omega^2 L^2}}$$

$$\Rightarrow \left\{ \begin{array}{c} \omega \to 0 : V_L \to \frac{\omega L V_0}{R} \to 0 \\ \omega \to \infty : V_L \to \frac{\omega L V_0}{\omega \sqrt{\left(\frac{R}{\omega}\right)^2 + L^2}} = \frac{L V_0}{L} = V_0 \end{array} \right\} \Rightarrow high\ pass\ filter$$

Low pass RC filter

Let's now study the voltage across a capacitor of a driven RC circuit

The complex current is now:



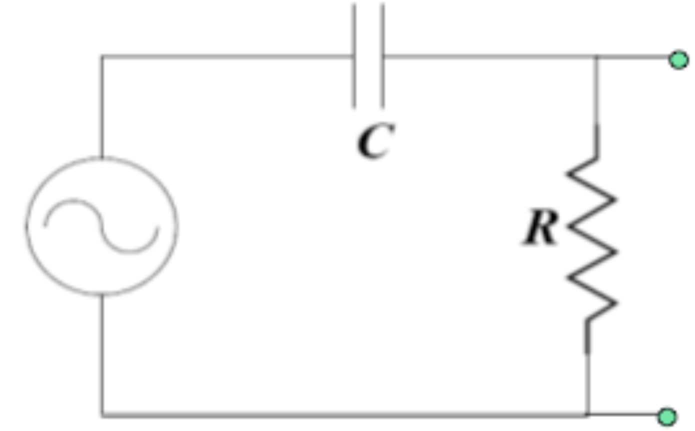
$$\tilde{I} = \frac{\tilde{V}}{\tilde{Z}} = \frac{\tilde{V}}{R - \frac{i}{\omega C}} \Rightarrow |I| = \frac{|\tilde{V}|}{\left|R - \frac{i}{\omega C}\right|} = \frac{V_0}{\sqrt{R^2 + \frac{1}{\omega^2 C^2}}} \Rightarrow V_C = \frac{|I|}{\omega C} = \frac{\frac{V_0}{\omega C}\omega C}{\sqrt{\omega^2 C^2 R^2 + 1^2}} = \frac{V_0}{\sqrt{\omega^2 C^2 R^2 + 1^2}}$$

$$\Rightarrow \left\{ \begin{array}{c} \omega \to 0 : V_C \to V_0 \\ \omega \to \infty : V_C \to 0 \end{array} \right\} \Rightarrow low\ pass\ filter$$

High pass RC filter

What if we take the voltage V_R across the resistor?

Same complex current

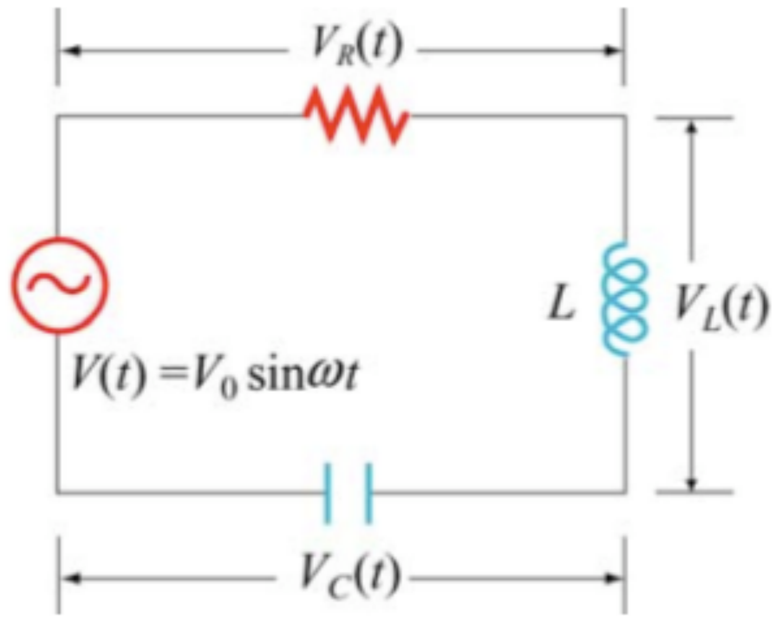


$$\tilde{I} = \frac{\tilde{V}}{\tilde{Z}} = \frac{\tilde{V}}{R - \frac{i}{\omega C}} \Rightarrow |I| = \frac{|\tilde{V}|}{\left|R - \frac{i}{\omega C}\right|} = \frac{V_0}{\sqrt{R^2 + \frac{1}{\omega^2 C^2}}} \Rightarrow V_R = R|I| = \frac{V_0}{\sqrt{R^2 + \frac{1}{\omega^2 C^2}}} = \frac{\omega C R V_0}{\sqrt{\omega^2 C^2 R^2 + 1^2}}$$

$$\Rightarrow \left\{ \begin{array}{c} \omega \to 0 : V_R \to 0 \\ \omega \to \infty : V_R \to V_0 \end{array} \right\} \Rightarrow high\ pass\ filter$$

The *RLC* Series Circuit

Consider now the driven series *RLC* circuit shown in Figure



Driven series *RLC* Circuit

Applying Kirchhoff's loop rule, we obtain

$$V(t) - V_R(t) - V_L(t) - V_C(t) = V(t) - IR - L\frac{dI}{dt} - \frac{Q}{C} = 0$$

which leads to the following differential equation:

$$L\frac{dI}{dt} + IR + \frac{Q}{C} = V_0 \sin\omega t$$

Assuming that the capacitor is initially uncharged so that $I = +dQ/dt$ is proportional to the *increase* of charge in the capacitor, the above equation can be rewritten as

$$\boxed{L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{Q}{C} = V_0 \sin\omega t}$$

One possible solution to Eq. (12.3.3) is

show how shortly

$$Q(t) = Q_0 \cos(\omega t - \phi)$$

where the amplitude and the phase are, respectively,

$$Q_0 = \frac{V_0 / L}{\sqrt{(R\omega / L)^2 + (\omega^2 - 1/LC)^2}} = \frac{V_0}{\omega\sqrt{R^2 + (\omega L - 1/\omega C)^2}}$$

$$= \frac{V_0}{\omega\sqrt{R^2 + (X_L - X_C)^2}}$$

and

$$\tan\phi = \frac{1}{R}\left(\omega L - \frac{1}{\omega C}\right) = \frac{X_L - X_C}{R}$$

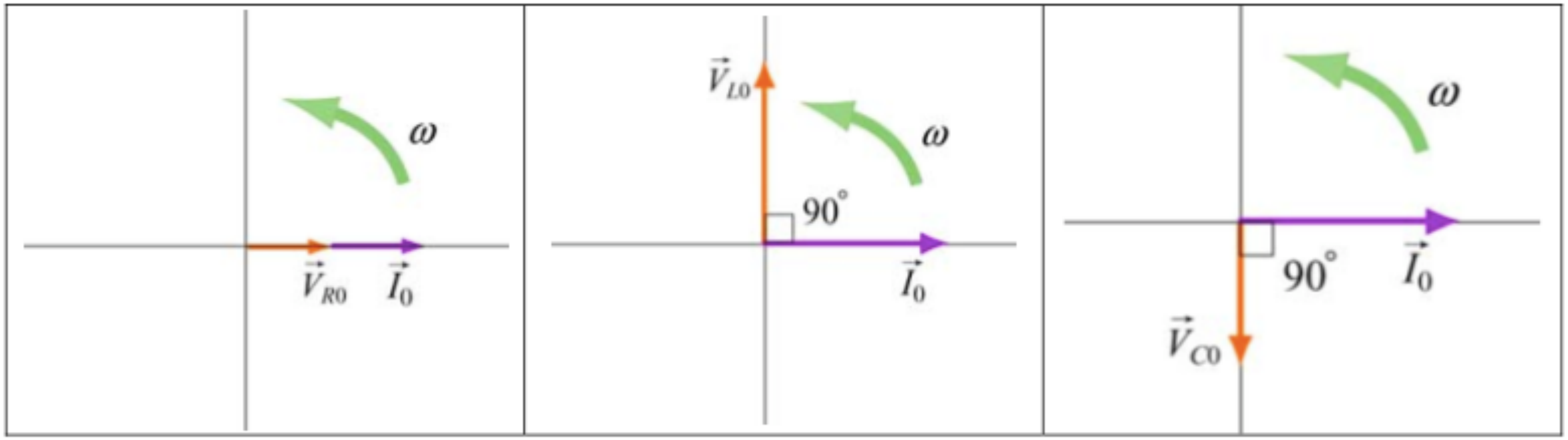
The corresponding current is

$$I(t) = +\frac{dQ}{dt} = I_0 \sin(\omega t - \phi)$$

with an amplitude

$$I_0 = -Q_0\omega = -\frac{V_0}{\sqrt{R^2 + (X_L - X_C)^2}}$$

Notice that the current has the same amplitude and phase at all points in the series *RLC* circuit. On the other hand, the instantaneous voltage across each of the three circuit elements *R*, *L* and *C* has a different amplitude and phase relationship with the current, as can be seen from the phasor diagrams shown in Figure



Phasor diagrams for the relationships between current and voltage in (a) the resistor, (b) the inductor, and (c) the capacitor, of a series *RLC* circuit.

From Figure the instantaneous voltages can be obtained as:

$$V_R(t) = I_0 R \sin\omega t = V_{R0} \sin\omega t$$

$$V_L(t) = I_0 X_L \sin\left(\omega t + \frac{\pi}{2}\right) = V_{L0} \cos\omega t$$

$$V_C(t) = I_0 X_C \sin\left(\omega t - \frac{\pi}{2}\right) = -V_{C0} \cos\omega t$$

where

$$V_{R0} = I_0 R, \quad V_{L0} = I_0 X_L, \quad V_{C0} = I_0 X_C$$

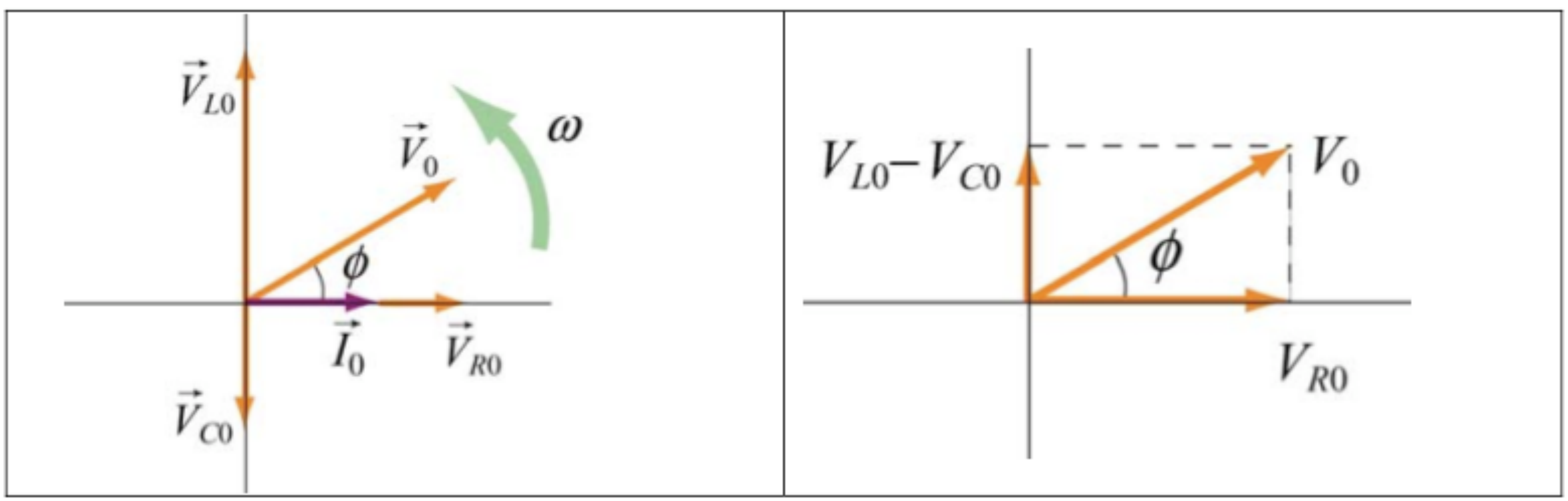
are the amplitudes of the voltages across the circuit elements. The sum of all three voltages is equal to the instantaneous voltage supplied by the AC source:

$$V(t)=V_R(t)+V_L(t)+V_C(t)$$

Using the phasor representation, the above expression can also be written as

$$\vec{V}_0=\vec{V}_{R0}+\vec{V}_{L0}+\vec{V}_{C0}$$

as shown in Figure (a). Again we see that current phasor $\vec{I}_0$ leads the capacitive voltage phasor $\vec{V}_{C0}$ by $\pi/2$ but lags the inductive voltage phasor $\vec{V}_{L0}$ by $\pi/2$. The three voltage phasors rotate counterclockwise as time passes, with their relative positions fixed.



(a) Phasor diagram for the series *RLC* circuit. (b) voltage relationship

The relationship between different voltage amplitudes is depicted in Figure (b).

From the Figure, we see that

$$V_0 = |\vec{V}_0| = |\vec{V}_{R0} + \vec{V}_{L0} + \vec{V}_{C0}| = \sqrt{V_{R0}^2 + (V_{L0} - V_{C0})^2}$$
$$= \sqrt{(I_0 R)^2 + (I_0 X_L - I_0 X_C)^2}$$
$$= I_0 \sqrt{R^2 + (X_L - X_C)^2}$$

which leads to the same expression for I_0

It is crucial to note that the maximum amplitude of the AC voltage source V_0 is not equal to the sum of the maximum voltage amplitudes across the three circuit elements:

$$V_0 \neq V_{R0} + V_{L0} + V_{C0}$$

This is due to the fact that the voltages are not in phase with one another, and they reach their maxima at different times.

Actually solving the equation:

$$L\frac{d^2Q}{dt^2}+R\frac{dQ}{dt}+\frac{Q}{C}=V_0\sin\omega t$$

rearranging:

$$\frac{d^2Q}{dt^2}+\frac{R}{L}\frac{dQ}{dt}+\frac{Q}{LC}=\frac{V_0}{L}\sin\omega t$$

or

$$\frac{d^2Q}{dt^2}+\gamma\frac{dQ}{dt}+\omega_0^2 Q=F_0\sin\omega t$$

The solution of this equation is the sum of two solutions

$$Q(t)=Q_h(t)+Q_p(t)$$

where

$$\frac{d^2Q_h}{dt^2}+\gamma\frac{dQ_h}{dt}+\omega_0^2 Q_h=0$$ homogeneous solution

and

$$\frac{d^2Q_p}{dt^2}+\gamma\frac{dQ_p}{dt}+\omega_0^2 Q_p=F_0\sin\omega t$$ particular solution

Exponential substitution method:

$$Q_h(t) = e^{i\alpha t}$$

This gives

$$(\alpha^2 + \gamma\alpha + \omega_0^2)e^{i\alpha t} = 0$$

or

$$\alpha^2 + \gamma\alpha + \omega_0^2 = 0$$

The solutions of this equation tell us the allowed values of α that give solutions to the diffEQ.

If there is more than one allowed value of α , then the most general solution will be a linear combination of all possible solutions.

In this case, the allowed values of α are

$$\alpha = i\frac{\gamma}{2} \pm \sqrt{\omega_0^2 - \frac{\gamma^2}{4}}$$

from quadratic formula

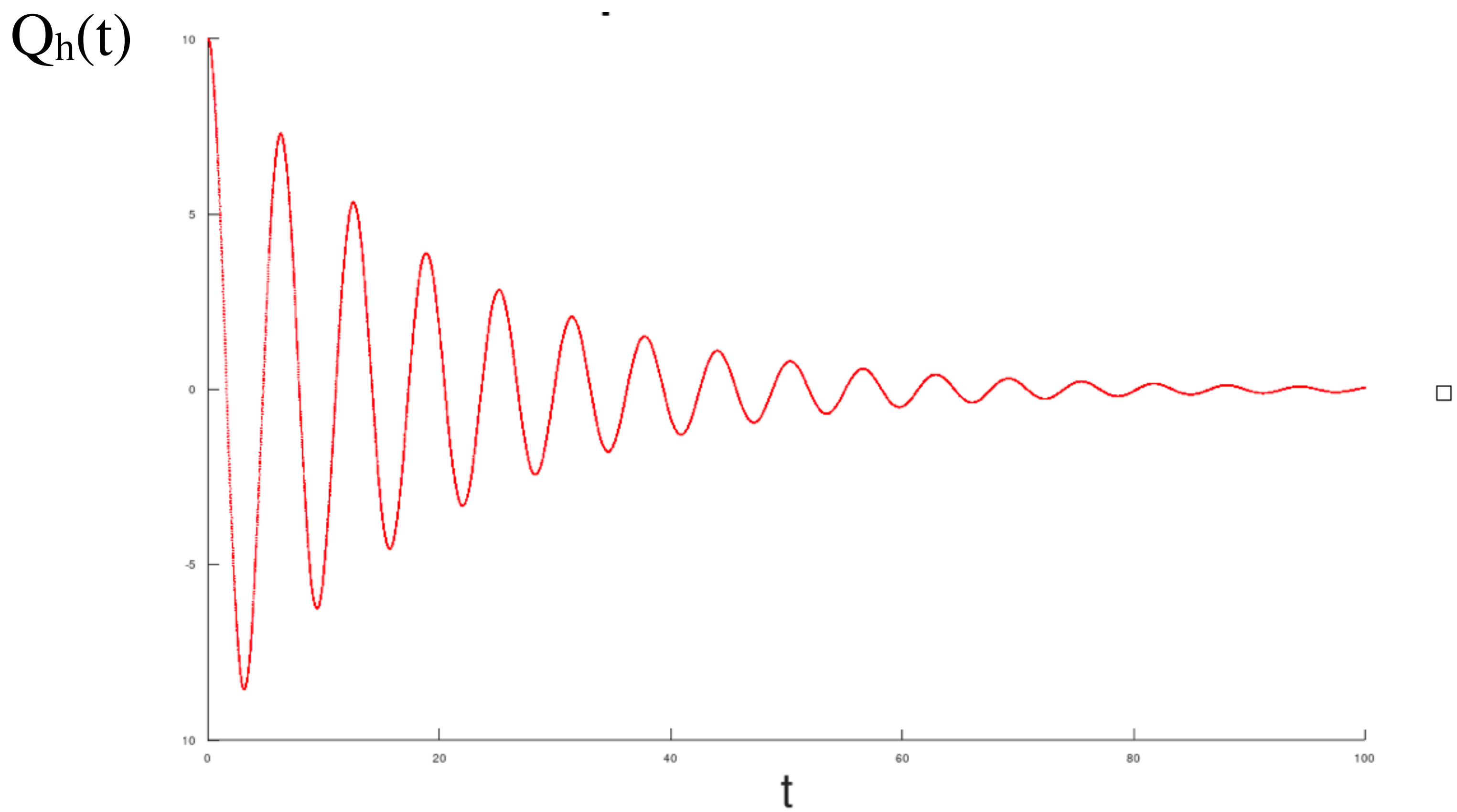
Therefore the most general solution is

$$x(t) = e^{-\frac{\gamma}{2}t}(Be^{i\omega_1 t} + Ce{-}i\omega_1 t) \quad , \quad \omega_1 = \sqrt{\omega_0^2 - \frac{\gamma^2}{4}}$$

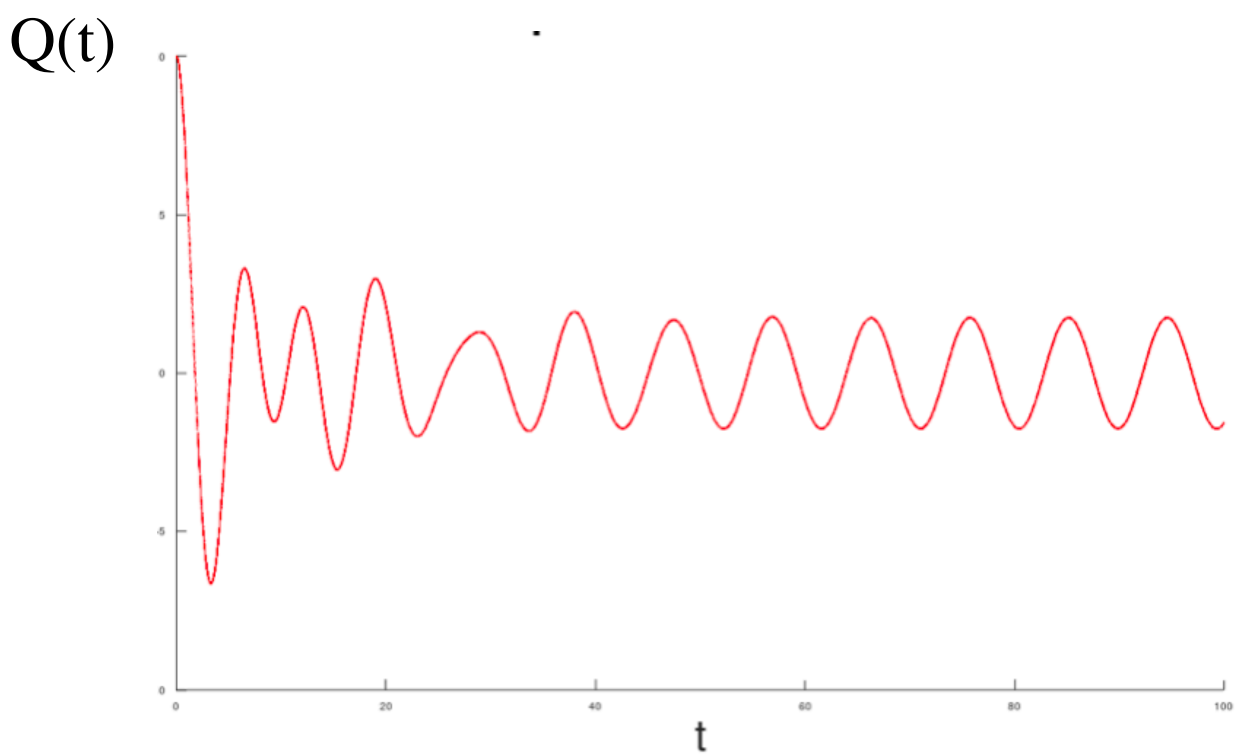
or

$$x(t) = Ae^{-\frac{\gamma}{2}t}\cos(\omega_1 t + \phi)$$

—> a “transient” solution



If one observes this type of system, then its time behavior looks like



We initially observe erratic behavior when both solutions are present.

As transient solution decays away, system settles into steady-state solution as shown with frequency ω, same as frequency of driving force. **This suggests that we assume a steady-state solution of the form**

$$Q_p(t) = Q_0 \cos(\omega t - \phi)$$

which is what we used earlier in this discussion!

Direct substitution and lots of algebra gives

$$Q_0 = \frac{V_0 / L}{\sqrt{(R\omega / L)^2 + (\omega^2 - 1/LC)^2}} = \frac{V_0}{\omega\sqrt{R^2 + (\omega L - 1/\omega C)^2}}$$

$$= \frac{V_0}{\omega\sqrt{R^2 + (X_L - X_C)^2}}$$

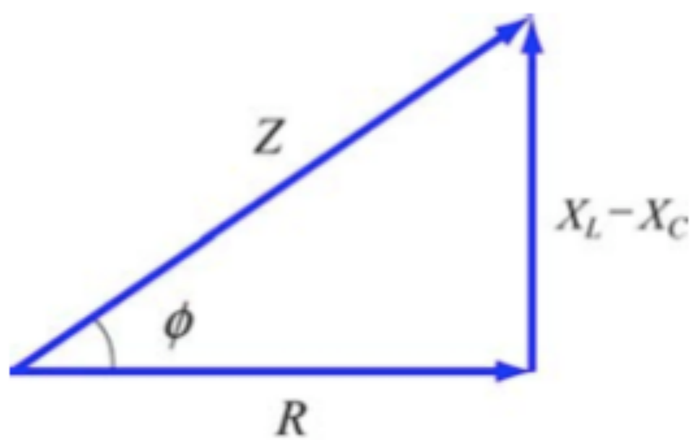
$$\tan\phi = \frac{1}{R}\left(\omega L - \frac{1}{\omega C}\right) = \frac{X_L - X_C}{R}$$

Impedance

We have already seen that the inductive reactance $X_L = \omega L$ and capacitance reactance $X_C = 1/\omega C$ play the role of an effective resistance in the purely inductive and capacitive circuits, respectively. In the series *RLC* circuit, the effective resistance is the *impedance*, defined as

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

The relationship between Z, X_L and X_C can be represented by the diagram shown in Figure



Diagrammatic representation of the relationship between Z, X_L and X_C

The impedance also has SI units of ohms. In terms of Z, the current may be rewritten as

$$I(t) = \frac{V_0}{Z} \sin(\omega t - \phi)$$

Notice that the impedance Z also depends on the angular frequency ω, as do X_L and X_C.

Simple Circuit	R	L	C	$X_L = \omega L$	$X_C = \frac{1}{\omega C}$	$\phi = \tan^{-1}\left(\frac{X_L - X_C}{R}\right)$	$Z = \sqrt{R^2 + (X_L - X_C)^2}$
purely resistive	R	0	∞	0	0	0	R
purely inductive	0	L	∞	X_L	0	$\pi/2$	X_L
purely capacitive	0	0	C	0	X_C	$-\pi/2$	X_C

Simple-circuit limits of the series *RLC* circuit

Resonance

the amplitude of the current $I_0 = V_0 / Z$ reaches a maximum

when Z is at a minimum. This occurs when $X_L = X_C$, or $\omega L = 1/\omega C$, leading to

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

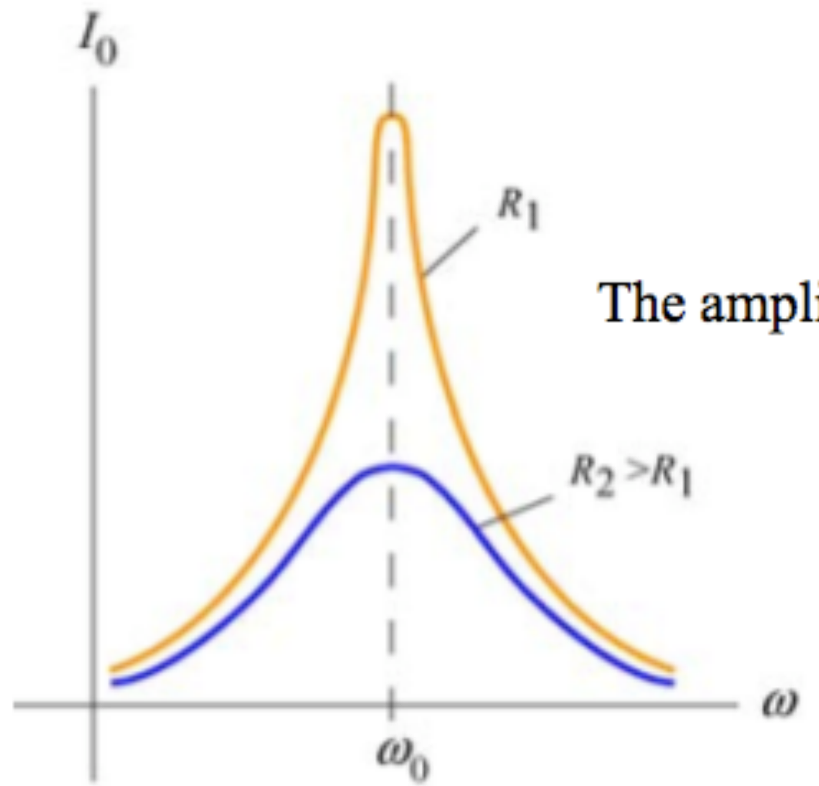
The phenomenon at which I_0 reaches a maximum is called a resonance, and the frequency ω_0 is called the resonant frequency. At resonance, the impedance becomes $Z = R$, the amplitude of the current is

$$I_0 = \frac{V_0}{R}$$

and the phase is

$$\phi = 0$$

The qualitative behavior is illustrated in Figure



The amplitude of the current as a function of ω in the driven RLC circuit.

Power in an AC circuit

In the series RLC circuit, the instantaneous power delivered by the AC generator is given by

$$P(t) = I(t)V(t) = \frac{V_0}{Z}\sin(\omega t - \phi)\cdot V_0 \sin\omega t = \frac{V_0^2}{Z}\sin(\omega t - \phi)\sin\omega t$$

$$= \frac{V_0^2}{Z}\left(\sin^2\omega t\cos\phi - \sin\omega t\cos\omega t\sin\phi\right)$$

where we have used the trigonometric identity

$$\sin(\omega t - \phi) = \sin\omega t\cos\phi - \cos\omega t\sin\phi$$

The time average of the power is

$$\langle P(t)\rangle = \frac{1}{T}\int_0^T \frac{V_0^{\,2}}{Z}\sin^2\omega t\cos\phi\; dt - \frac{1}{T}\int_0^T \frac{V_0^{\,2}}{Z}\sin\omega t\cos\omega t\sin\phi\; dt$$

$$= \frac{V_0^{\,2}}{Z}\cos\phi\langle\sin^2\omega t\rangle - \frac{V_0^{\,2}}{Z}\sin\phi\langle\sin\omega t\cos\omega t\rangle$$

$$= \frac{1}{2}\frac{V_0^{\,2}}{Z}\cos\phi$$

In terms of the rms quantities, the average power can be rewritten as

$$\langle P(t)\rangle = \frac{1}{2}\frac{V_0^{\,2}}{Z}\cos\phi = \frac{V_{\mathrm{rms}}^{\;\;2}}{Z}\cos\phi = I_{\mathrm{rms}}V_{\mathrm{rms}}\cos\phi$$

The quantity $\cos\phi$ is called the *power factor*.

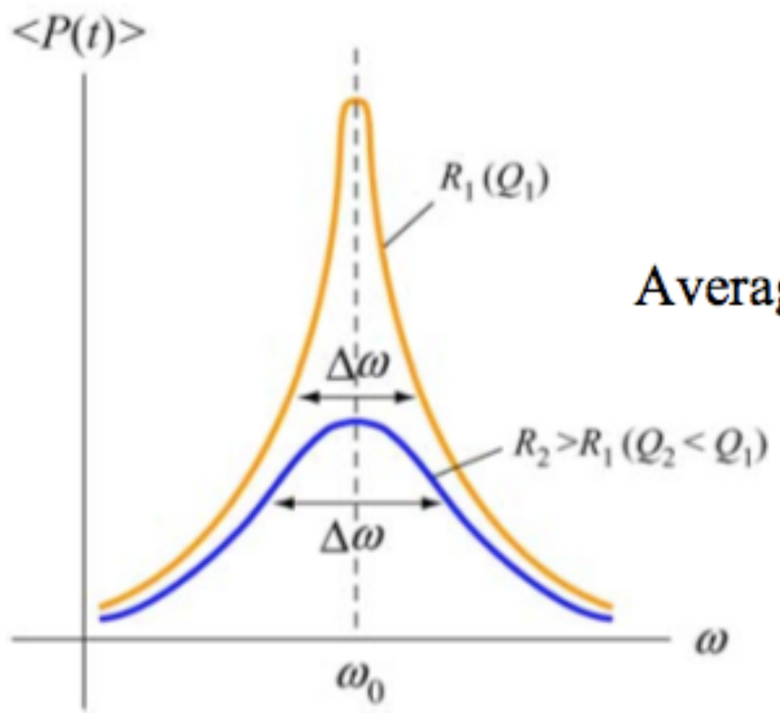
one can readily show

$$\cos\phi = \frac{R}{Z}$$

Thus, we may rewrite $\langle P(t)\rangle$ as

$$\langle P(t)\rangle = I_{\mathrm{rms}}V_{\mathrm{rms}}\left(\frac{R}{Z}\right) = I_{\mathrm{rms}}\left(\frac{V_{\mathrm{rms}}}{Z}\right)R = I_{\mathrm{rms}}^2 R$$

we plot the average power as a function of the driving angular frequency ω.



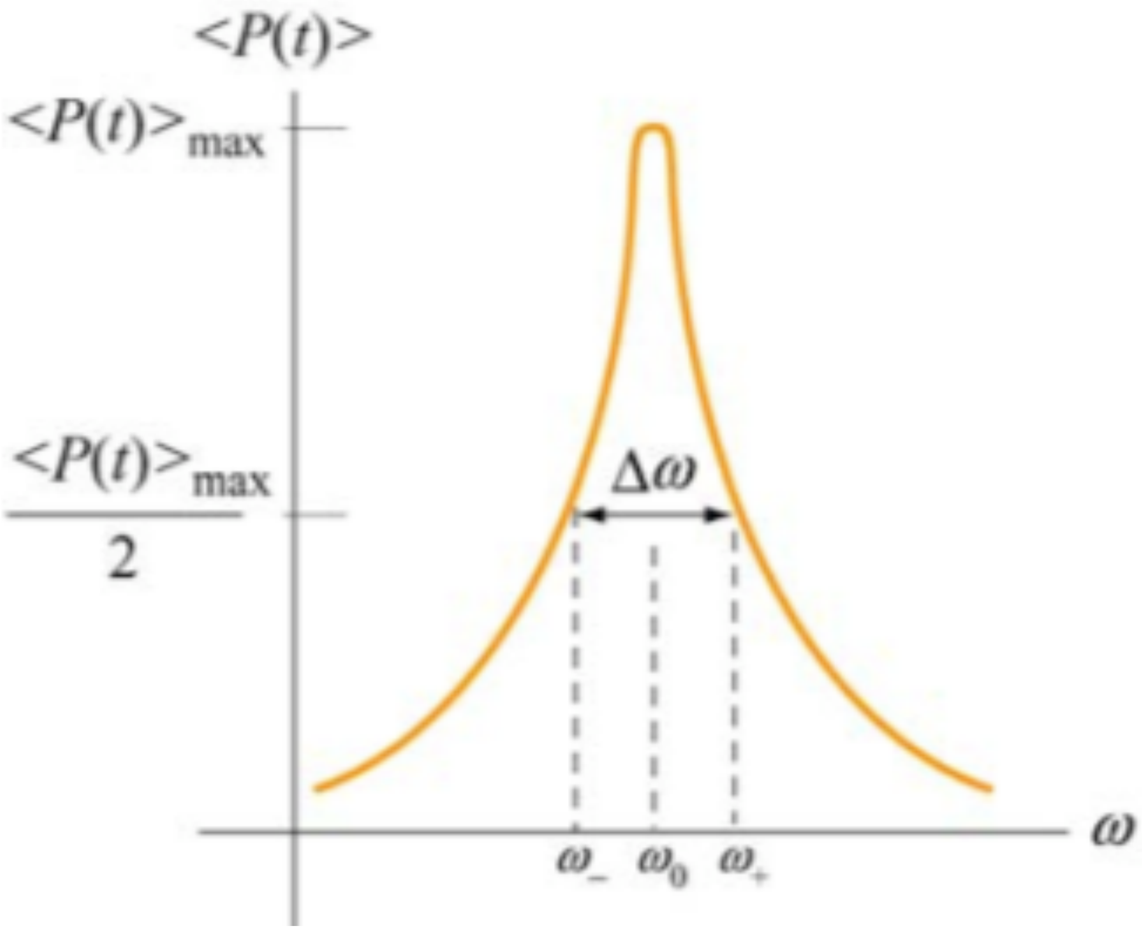
Average power as a function of frequency in a driven series *RLC* circuit.

We see that $\langle P(t) \rangle$ attains the maximum when $\cos\phi = 1$, or $Z = R$, which is the resonance condition. At resonance, we have

$$\langle P \rangle_{\text{max}} = I_{\text{rms}} V_{\text{rms}} = \frac{V_{\text{rms}}^2}{R}$$

Width of the Peak

The peak has a line width. One way to characterize the width is to define $\Delta\omega = \omega_+ - \omega_-$, where $\omega_\pm$ are the values of the driving angular frequency such that the power is equal to half its maximum power at resonance. This is called *full width at half maximum*, as illustrated in Figure The width $\Delta\omega$ increases with resistance R.



Width of the peak

To find $\Delta\omega$, it is instructive to first rewrite the average power $\langle P(t)\rangle$ as

$$\langle P(t)\rangle = \frac{1}{2}\frac{V_0^{\,2}R}{R^2+(\omega L-1/\omega C)^2} = \frac{1}{2}\frac{V_0^{\,2}R\omega^2}{\omega^2R^2+L^2(\omega^2-\omega_0^2)^2}$$

with $\langle P(t)\rangle_{\max} = V_0^2/2R$. The condition for finding $\omega_\pm$ is

$$\frac{1}{2}\langle P(t)\rangle_{\max} = \langle P(t)\rangle\Big|_{\omega_\pm} \quad \Rightarrow \quad = \frac{V_0^{\,2}}{4R} = \frac{1}{2}\frac{V_0^{\,2}R\omega^2}{\omega^2R^2+L^2(\omega^2-\omega_0^2)^2}\Bigg|_{\omega_\pm}$$

which gives

$$(\omega^2-\omega_0^2)^2 = \left(\frac{R\omega}{L}\right)^2$$

Taking square roots yields two solutions, which we analyze separately.

case 1: Taking the positive root leads to

$$\omega_+^2 - \omega_0^2 = +\frac{R\omega_+}{L}$$

Solving the quadratic equation, the solution with positive root is

$$\omega_+ = \frac{R}{2L} + \sqrt{\left(\frac{R}{4L}\right)^2 + \omega_0^{\ 2}}$$

Case 2: Taking the negative root

The solution to this quadratic equation with positive root is

$$\omega_- = -\frac{R}{2L} + \sqrt{\left(\frac{R}{4L}\right)^2 + \omega_0^{\ 2}}$$

The width at half maximum is then

$$\Delta\omega = \omega_+ - \omega_- = \frac{R}{L}$$

Once the width $\Delta\omega$ is known, the quality factor Q (not to be confused with charge) can be obtained as

$$Q = \frac{\omega_0}{\Delta\omega} = \frac{\omega_0 L}{R}$$

Comparing we see that both expressions agree

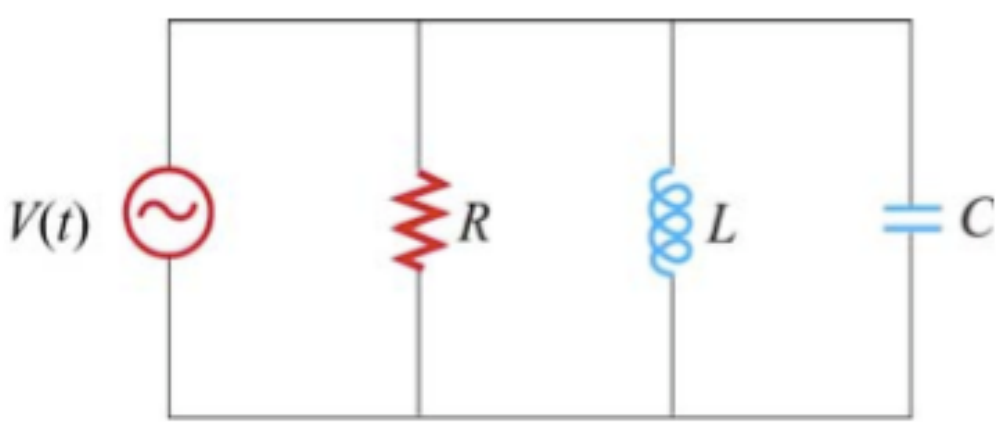
with each other in the limit where the resistance is small, and $\omega' = \sqrt{\omega_0^2 - (R/2L)^2} \approx \omega_0$.

Parallel *RLC* Circuit

Consider the parallel *RLC* circuit illustrated in Figure

The AC voltage source is

$$V(t) = V_0 \sin \omega t \ .$$



Parallel *RLC* circuit.

Unlike the series *RLC* circuit, the instantaneous voltages across all three circuit elements *R*, *L*, and *C* are the same, and each voltage is in phase with the current through the resistor. However, the currents through each element will be different.

The current in the resistor is

$$I_R(t) = \frac{V(t)}{R} = \frac{V_0}{R}\sin\omega t = I_{R0}\sin\omega t$$

where $I_{R0} = V_0 / R$. The voltage across the inductor is

$$V_L(t) = V(t) = V_0 \sin\omega t = L\frac{dI_L}{dt}$$

which gives

$$I_L(t) = \int_0^t \frac{V_0}{L}\sin\omega t' dt' = -\frac{V_0}{\omega L}\cos\omega t = \frac{V_0}{X_L}\sin\left(\omega t - \frac{\pi}{2}\right) = I_{L0}\sin\left(\omega t - \frac{\pi}{2}\right)$$

where $I_{L0} = V_0 / X_L$ and $X_L = \omega L$ is the inductive reactance.

Similarly, the voltage across the capacitor is $V_C(t) = V_0 \sin\omega t = Q(t)/C$, which implies

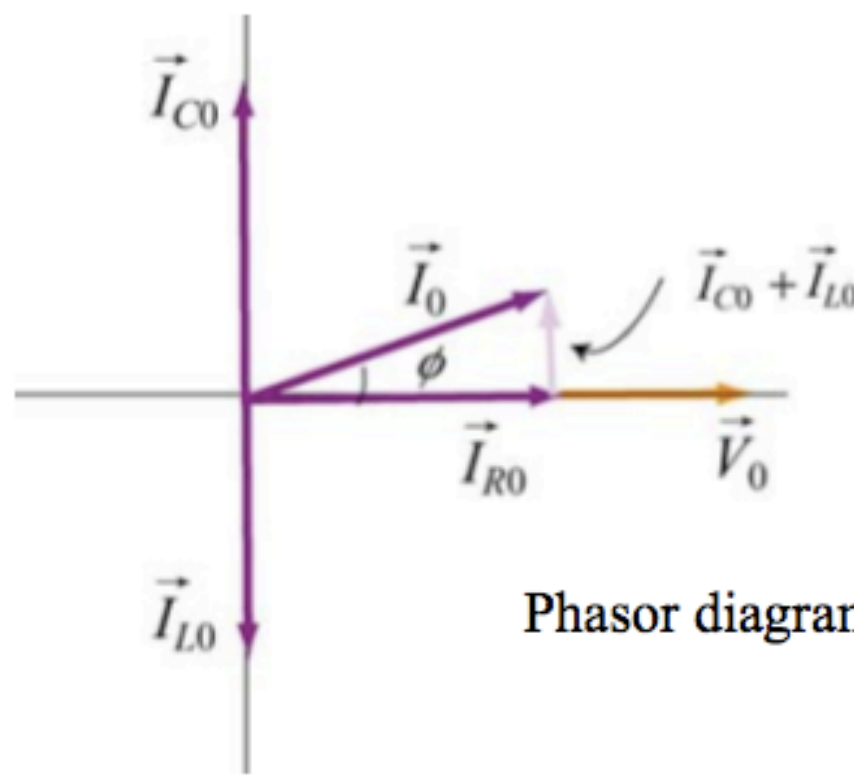
$$I_C(t) = \frac{dQ}{dt} = \omega C V_0 \cos\omega t = \frac{V_0}{X_C}\sin\left(\omega t + \frac{\pi}{2}\right) = I_{C0}\sin\left(\omega t + \frac{\pi}{2}\right)$$

where $I_{C0} = V_0 / X_C$ and $X_C = 1/\omega C$ is the capacitive reactance.

Using Kirchhoff's junction rule, the total current in the circuit is simply the sum of all three currents.

$$\begin{aligned} I(t) &= I_R(t) + I_L(t) + I_C(t) \\ &= I_{R0}\sin\omega t + I_{L0}\sin\left(\omega t - \frac{\pi}{2}\right) + I_{C0}\sin\left(\omega t + \frac{\pi}{2}\right) \end{aligned}$$

The currents can be represented with the phasor diagram shown in Figure



Phasor diagram for the parallel *RLC* circuit

From the phasor diagram, we see that

$$\vec{I}_0 = \vec{I}_{R0} + \vec{I}_{L0} + \vec{I}_{C0}$$

and the maximum amplitude of the total current, I_0, can be obtained as

$$I_0 = |\vec{I}_0| = |\vec{I}_{R0} + \vec{I}_{L0} + \vec{I}_{C0}| = \sqrt{I_{R0}^2 + (I_{C0} - I_{L0})^2}$$
$$= V_0 \sqrt{\frac{1}{R^2} + \left(\omega C - \frac{1}{\omega L}\right)^2} = V_0 \sqrt{\frac{1}{R^2} + \left(\frac{1}{X_C} - \frac{1}{X_L}\right)^2}$$

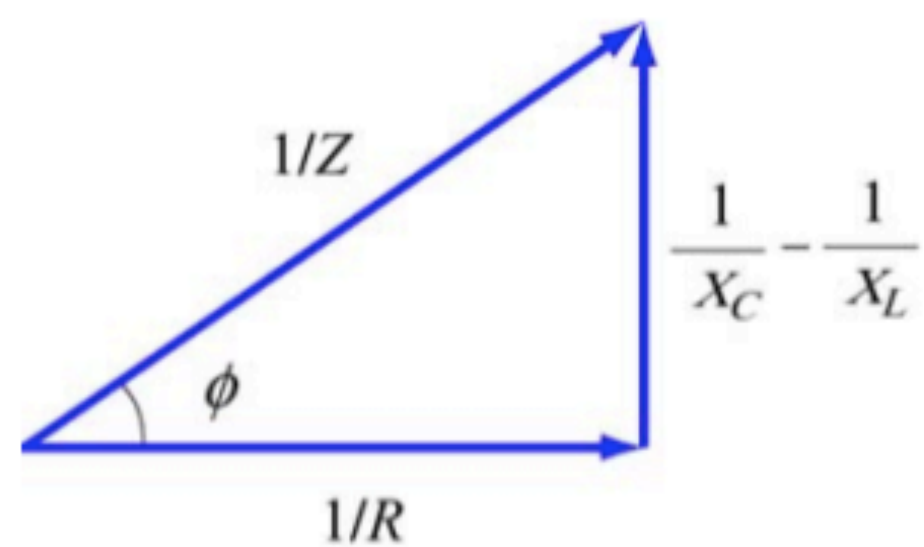
Note however, since $I_R(t)$, $I_L(t)$ and $I_C(t)$ are not in phase with one another, I_0 is not equal to the sum of the maximum amplitudes of the three currents:

$$I_0 \neq I_{R0} + I_{L0} + I_{C0}$$

With $I_0 = V_0 / Z$, the (inverse) impedance of the circuit is given by

$$\frac{1}{Z} = \sqrt{\frac{1}{R^2} + \left(\omega C - \frac{1}{\omega L}\right)^2} = \sqrt{\frac{1}{R^2} + \left(\frac{1}{X_C} - \frac{1}{X_L}\right)^2}$$

The relationship between Z, R, X_L and X_C is shown in Figure



Relationship between Z, R, X_L and X_C in a parallel *RLC* circuit.

From the figure or the phasor diagram shown we see that the phase can be obtained as

$$\tan\phi = \left(\frac{I_{C0}-I_{L0}}{I_{R0}}\right) = \frac{\dfrac{V_0}{X_C}-\dfrac{V_0}{X_L}}{\dfrac{V_0}{R}} = R\left(\frac{1}{X_C}-\frac{1}{X_L}\right) = R\left(\omega C - \frac{1}{\omega L}\right)$$

The resonance condition for the parallel *RLC* circuit is given by $\phi = 0$, which implies

$$\frac{1}{X_C} = \frac{1}{X_L}$$

The resonant frequency is

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

which is the same as for the series *RLC* circuit. we readily see that $1/Z$ is minimum (or Z is maximum) at resonance. The current in the inductor exactly cancels out the current in the capacitor, so that the total current in the circuit reaches a minimum, and is equal to the current in the resistor:

$$I_0 = \frac{V_0}{R}$$

As in the series *RLC* circuit, power is dissipated only through the resistor. The average power is

$$\left\langle P(t)\right\rangle=\left\langle I_R(t)V(t)\right\rangle=\left\langle I_R^2(t)R\right\rangle=\frac{V_0^2}{R}\left\langle \sin^2\omega t\right\rangle=\frac{V_0^2}{2R}=\frac{V_0^2}{2Z}\left(\frac{Z}{R}\right)$$

Thus, the power factor in this case is

$$\text{power factor}=\frac{\left\langle P(t)\right\rangle}{V_0^2/2Z}=\frac{Z}{R}=\frac{1}{\sqrt{1+\left(R\omega C-\dfrac{R}{\omega L}\right)^2}}=\cos\phi$$

Examples

RLC Series Circuit

A series *RLC* circuit with $L=160\text{ mH}$, $C=100\,\mu\text{F}$, and $R=40.0\,\Omega$ is connected to a sinusoidal voltage $V(t)=(40.0\text{ V})\sin\omega t$, with $\omega=200\text{ rad/s}$.

(a) What is the impedance of the circuit?

(b) Let the current at any instant in the circuit be $I(t)=I_0\sin(\omega t-\phi)$. Find I_0.

(c) What is the phase ϕ?

(a) The impedance of a series *RLC* circuit is given by

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

where

$$X_L = \omega L$$

and

$$X_C = \frac{1}{\omega C}$$

are the inductive reactance and the capacitive reactance, respectively. Since the general expression of the voltage source is $V(t) = V_0 \sin(\omega t)$, where V_0 is the maximum output voltage and ω is the angular frequency, we have $V_0 = 40$ V and $\omega = 200$ rad/s . Thus, the impedance Z becomes

$$Z = \sqrt{(40.0\ \Omega)^2 + \left((200\text{ rad/s})(0.160\text{ H}) - \frac{1}{(200\text{ rad/s})(100\times 10^{-6}\text{ F})}\right)^2}$$
$$= 43.9\Omega$$

(b) With $V_0 = 40.0\,\text{V}$, the amplitude of the current is given by

$$I_0 = \frac{V_0}{Z} = \frac{40.0\,\text{V}}{43.9\Omega} = 0.911\,\text{A}$$

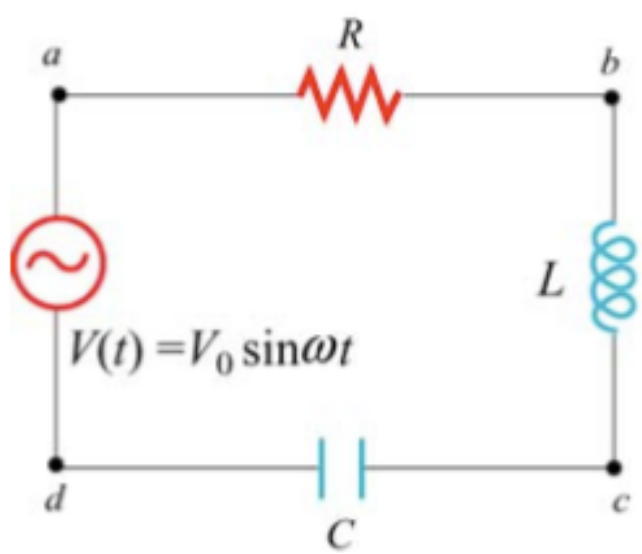
(c) The phase between the current and the voltage is determined by

$$\phi = \tan^{-1}\left(\frac{X_L - X_C}{R}\right) = \tan^{-1}\left(\frac{\omega L - \dfrac{1}{\omega C}}{R}\right)$$

$$= \tan^{-1}\left(\frac{(200\ \text{rad/s})(0.160\ \text{H}) - \dfrac{1}{(200\ \text{rad/s})(100\times10^{-6}\ \text{F})}}{40.0\ \Omega}\right) = -24.2°$$

RLC Series Circuit

Suppose an AC generator with $V(t) = (150\,\text{V})\sin(100t)$ is connected to a series *RLC* circuit with $R = 40.0\ \Omega$, $L = 80.0\ \text{mH}$, and $C = 50.0\ \mu\text{F}$, as shown in Figure



RLC series circuit

(a) Calculate V_{R0}, V_{L0} and V_{C0}, the maximum of the voltage drops across each circuit element.

(b) Calculate the maximum potential difference across the inductor and the capacitor between points *b* and *d* shown in Figure

Solutions:

(a) The inductive reactance, capacitive reactance and the impedance of the circuit are given by

$$X_C = \frac{1}{\omega C} = \frac{1}{(100 \text{ rad/s})(50.0\times 10^{-6} \text{ F})} = 200\ \Omega$$

$$X_L = \omega L = (100 \text{ rad/s})(80.0\times 10^{-3} \text{ H}) = 8.00\ \Omega$$

and

$$Z = \sqrt{R^2 + (X_L - X_C)^2} = \sqrt{(40.0\ \Omega)^2 + (8.00\ \Omega - 200\ \Omega)^2} = 196\ \Omega$$

respectively. Therefore, the corresponding maximum current amplitude is

$$I_0 = \frac{V_0}{Z} = \frac{150 \text{ V}}{196\ \Omega} = 0.765 \text{ A}$$

The maximum voltage across the resistance would be just the product of maximum current and the resistance:

$$V_{R0} = I_0 R = (0.765\ \text{A})(40.0\ \Omega) = 30.6\ \text{V}$$

Similarly, the maximum voltage across the inductor is

$$V_{L0} = I_0 X_L = (0.765\ \text{A})(8.00\ \Omega) = 6.12\ \text{V}$$

and the maximum voltage across the capacitor is

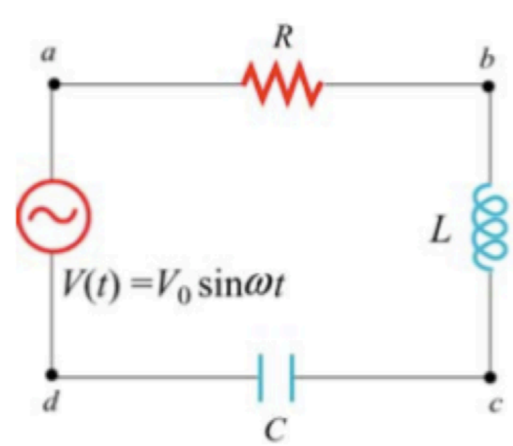
$$V_{C0} = I_0 X_C = (0.765\ \text{A})(200\ \Omega) = 153\ \text{V}$$

Note that the maximum input voltage V_0 is related to V_{R0}, V_{L0} and V_{C0} by

$$V_0 = \sqrt{V_{R0}^2 + (V_{L0} - V_{C0})^2}$$

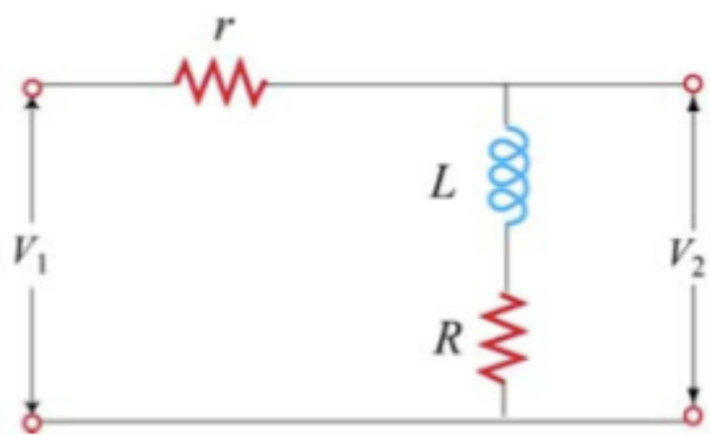
(b) From b to d, the maximum voltage would be the difference between V_{L0} and V_{C0}:

$$|V_{bd}| = |\vec{V}_{L0} + \vec{V}_{C0}| = |V_{L0} - V_{C0}| = |6.12\ \text{V} - 153\ \text{V}| = 147\ \text{V}$$



RL High-Pass Filter

An *RL* high-pass filter (circuit that filters out low-frequency AC currents) can be represented by the circuit in Figure where R is the internal resistance of the inductor.



RL filter

(a) Find V_{20}/V_{10}, the ratio of the maximum output voltage V_{20} to the maximum input voltage V_{10}.

(b) Suppose $r = 15.0\,\Omega$, $R = 10\Omega$ and $L = 250$ mH. Find the frequency at which $V_{20}/V_{10} = 1/2$.

Solution:

(a) The impedance for the input circuit is $Z_1 = \sqrt{(R+r)^2 + X_L^2}$ where $X_L = \omega L$ and $Z_2 = \sqrt{R^2 + X_L^2}$ for the output circuit. The maximum current is given by

$$I_0 = \frac{V_{10}}{Z_1} = \frac{V_0}{\sqrt{(R+r)^2 + X_L^2}}$$

Similarly, the maximum output voltage is related to the output impedance by

$$V_{20} = I_0 Z_2 = I_0 \sqrt{R^2 + X_L^2}$$

This implies

$$\frac{V_{20}}{V_{10}} = \frac{\sqrt{R^2 + X_L^2}}{\sqrt{(R+r)^2 + X_L^2}}$$

(b) For $V_{20}/V_{10} = 1/2$, we have

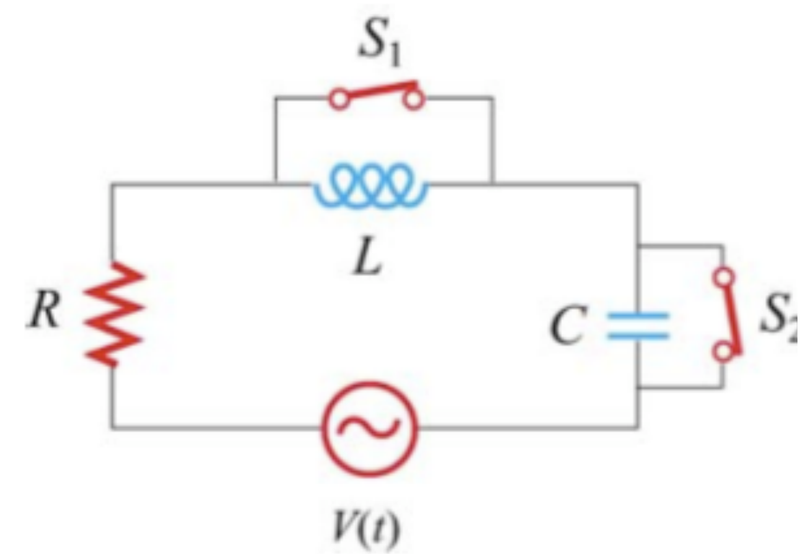
$$\frac{R^2 + X_L^2}{(R+r)^2 + X_L^2} = \frac{1}{4} \quad \Rightarrow \quad X_L = \sqrt{\frac{(R+r)^2 - 4R^2}{3}}$$

Since $X_L = \omega L = 2\pi fL$, the frequency which yields this ratio is

$$f = \frac{X_L}{2\pi L} = \frac{1}{2\pi(0.250\ \text{H})}\sqrt{\frac{(10.0\ \Omega + 15.0\ \Omega)^2 - 4(10.0\ \Omega)^2}{3}} = 5.51\,\text{Hz}$$

RLC Circuit

Consider the circuit shown in Figure The sinusoidal voltage source is $V(t)=V_0 \sin\omega t$. If both switches S_1 and S_2 are closed initially, find the following quantities, ignoring the transient effect and assuming that R , L ,V_0 and ω are known:



(a) the current $I(t)$ as a function of time,

(b) the average power delivered to the circuit,

(c) the current as a function of time a long time after only S_1 is opened.

(d) the capacitance C if both S_1 and S_2 are opened for a long time, with the current and voltage in phase.

(e) the impedance of the circuit when both S_1 and S_2 are opened.

(f) the maximum energy stored in the capacitor during oscillations.

(g) the maximum energy stored in the inductor during oscillations.

(h) the phase difference between the current and the voltage if the frequency of $V(t)$ is doubled.

(i) the frequency at which the inductive reactance X_L is equal to half the capacitive reactance X_C.

Solutions:

(a) When both switches S_1 and S_2 are closed, the current goes through only the generator and the resistor, so the total impedance of the circuit is R and the current is

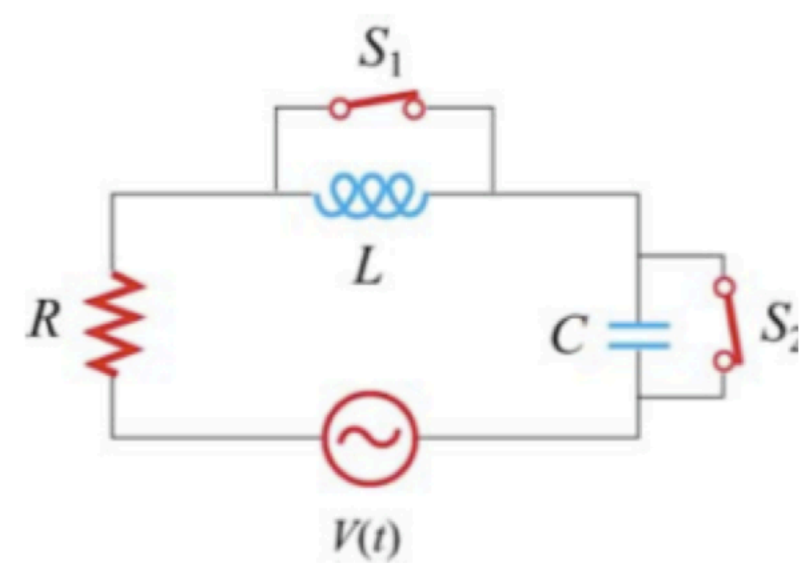
$$I_R(t) = \frac{V_0}{R}\sin\omega t$$

(b) The average power is given by

$$\langle P(t)\rangle = \langle I_R(t)V(t)\rangle = \frac{V_0^2}{R}\langle \sin^2\omega t\rangle = \frac{V_0^2}{2R}$$

(c) If only S_1 is opened, after a long time the current will pass through the generator, the resistor and the inductor. For this RL circuit, the impedance becomes

$$Z = \frac{1}{\sqrt{R^2 + X_L^2}} = \frac{1}{\sqrt{R^2 + \omega^2 L^2}}$$



and the phase angle ϕ is

$$\phi = \tan^{-1}\left(\frac{\omega L}{R}\right)$$

Thus, the current as a function of time is

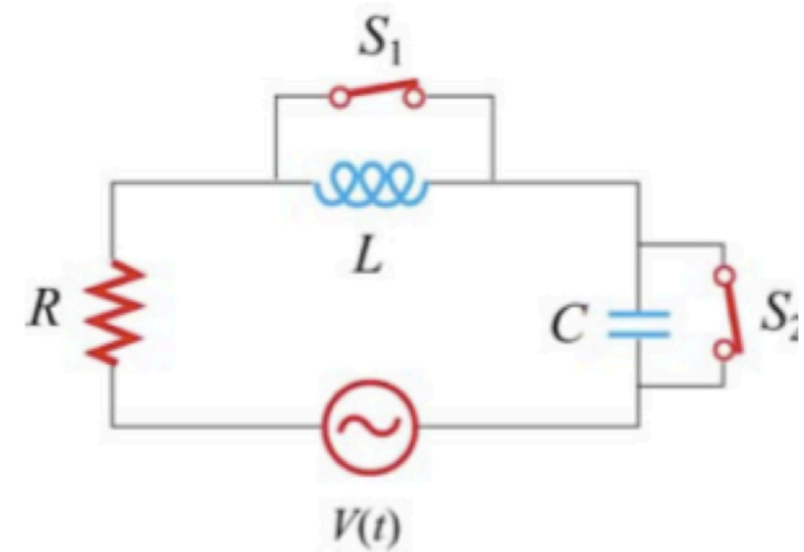
$$I(t) = I_0\sin(\omega t - \phi) = \frac{V_0}{\sqrt{R^2 + \omega^2 L^2}}\sin\left(\omega t - \tan^{-1}\frac{\omega L}{R}\right)$$

Note that in the limit of vanishing resistance $R=0$, $\phi=\pi/2$, and we recover the expected result for a purely inductive circuit.

(d) If both switches are opened, then this would be a driven *RLC* circuit, with the phase angle ϕ given by

$$\tan\phi = \frac{X_L - X_C}{R} = \frac{\omega L - \dfrac{1}{\omega C}}{R}$$

If the current and the voltage are in phase, then $\phi=0$, implying $\tan\phi=0$. Let the corresponding angular frequency be ω_0; we then obtain



$$\omega_0 L = \frac{1}{\omega_0 C}$$

and the capacitance is

$$C = \frac{1}{\omega_0^2 L}$$

(e) From (d), we see that when both switches are opened, the circuit is at resonance with $X_L = X_C$. Thus, the impedance of the circuit becomes

$$Z = \sqrt{R^2 + (X_L - X_C)^2} = R$$

(f) The electric energy stored in the capacitor is

$$U_E = \frac{1}{2}CV_C^2 = \frac{1}{2}C(IX_C)^2$$

It attains maximum when the current is at its maximum I_0:

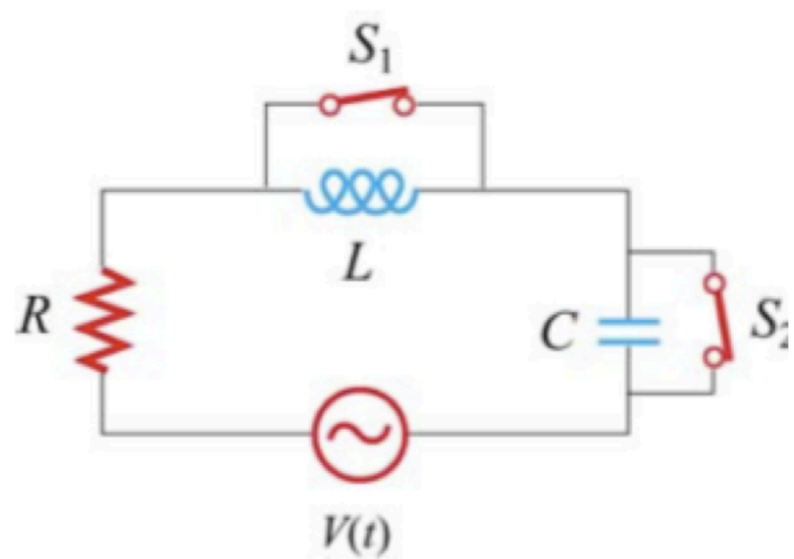
$$U_{C,\max} = \frac{1}{2}CI_0^2X_C^2 = \frac{1}{2}C\left(\frac{V_0}{R}\right)^2\frac{1}{\omega_0{}^2C^2} = \frac{V_0^2L}{2R^2}$$

where we have used $\omega_0^2 = 1/LC$.

(g) The maximum energy stored in the inductor is given by

$$U_{L,\max} = \frac{1}{2}LI_0^2 = \frac{LV_0^2}{2R^2}$$

(h) If the frequency of the voltage source is doubled, i.e., $\omega = 2\omega_0 = 1/\sqrt{LC}$, then the phase becomes



$$\phi = \tan^{-1}\left(\frac{\omega L - 1/\omega C}{R}\right) = \tan^{-1}\left(\frac{\left(2/\sqrt{LC}\right)L - \left(\sqrt{LC}/2C\right)}{R}\right) = \tan^{-1}\left(\frac{3}{2R}\sqrt{\frac{L}{C}}\right)$$

(i) If the inductive reactance is one-half the capacitive reactance,

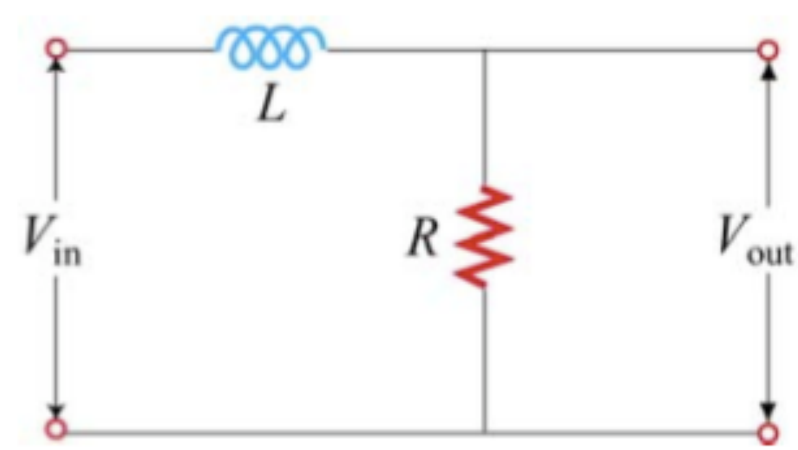
$$X_L = \frac{1}{2}X_C \quad \Rightarrow \quad \omega L = \frac{1}{2}\left(\frac{1}{\omega C}\right)$$

then

$$\omega = \frac{1}{\sqrt{2LC}} = \frac{\omega_0}{\sqrt{2}}$$

RL **Filter**

The circuit shown in Figure represents an *RL* filter.



Let the inductance be L = 400 mH, and the input voltage $V_{\text{in}} = (20.0\,\text{V})\sin\omega t$, where $\omega = 200$ rad/s.

(a) What is the value of R such that the output voltage lags behind the input voltage by $30.0°$?

(b) Find the ratio of the amplitude of the output and the input voltages. What type of filter is this circuit, high-pass or low-pass?

(c) If the positions of the resistor and the inductor are switched, would the circuit be a high-pass or a low-pass filter?

Solutions:

(a) The phase relationship between V_L and V_R is given by

$$\tan\phi = \frac{V_L}{V_R} = \frac{IX_L}{IX_R} = \frac{\omega L}{R}$$

Thus, we have

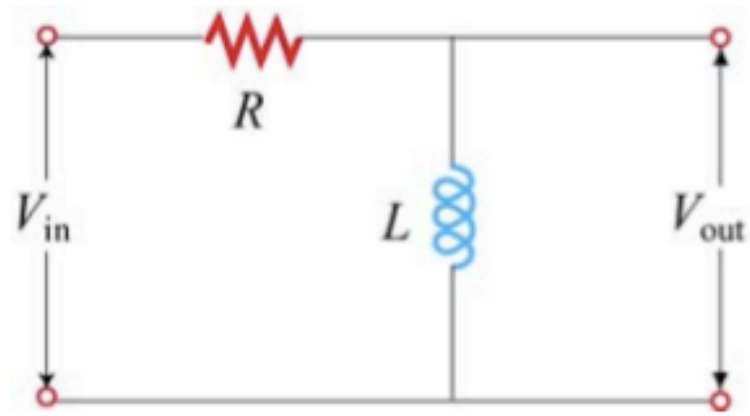
$$R = \frac{\omega L}{\tan\phi} = \frac{(200\,\text{rad/s})(0.400\,\text{H})}{\tan 30.0°} = 139\,\Omega$$

(b) The ratio is given by

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{V_R}{V_{\text{in}}} = \frac{R}{\sqrt{R^2 + X_L^2}} = \cos\phi = \cos 30.0° = 0.866.$$

The circuit is a low-pass filter, since the ratio V_{out} / V_{in} decreases with increasing ω.

(c) In this case, the circuit diagram is



RL high-pass filter

The ratio of the output voltage to the input voltage would be

$$\frac{V_{out}}{V_{in}} = \frac{V_L}{V_{in}} = \frac{X_L}{\sqrt{R^2 + X_L^2}} = \frac{\omega^2 L^2}{\sqrt{R^2 + \omega^2 L^2}} = \left[1 + \left(\frac{R}{\omega L}\right)^2\right]^{-1/2}$$

The circuit is a high-pass filter, since the ratio V_{out} / V_{in} approaches one in the large-ω limit.

Special relativity **some different approaches!**

- Ready for the challenge?
  - Special relativity seems easy but it’s not!
  - A new way of thinking that often goes against intuition
  - It will take some time to “digest it”, but believe me: it’s worth the effort!
- Why do we need it in Physics 008?
  - Weren't you frustrated last time when magnetic forces came out of nowhere?
  - Special relativity naturally explains them in terms of electric forces seen from in a reference frame in motion
- This is important for everybody
  - Physics majors: first of many iterations on a crucial topic
  - Non Physics majors: chance to realize what you are missing
    - Don’t forget: you are still in time… ☺

this is what I said to students!!

The principles of special relativity

- Formulated in 1905 by A. Einstein
  - Incredible but true: no Nobel Prize for this!
- Based upon 2 postulates
  - The laws of physics are the same for all reference frames
  - The speed of light is the same (c) in all reference frames
- (Inertial) Reference frame
  - System of coordinates in which the observer is non accelerating
    - (inertial = non accelerating)

Special Relativity — Theoretical Approach

We have the following **postulates**:

(1) All the laws of nature (not just mechanics) must be the same for all inertial observers moving with constant velocity relative to each other.

NOTE: If we were to write "All the laws of nature must be the same for all observers" for any pair of frames(could be accelerating), then we could derive General Relativity-much harder.

This is the **Principle of Relativity** and restricts the **form** of the laws in each frame.

(2) The speed of light is an invariant.

(3) The motion of a particle observed to be linear in one inertial frame must be linear in all inertial frames.

→ implies that the Lorentz transformations must be linear.

transformations between frames

We now do a **thought experiment.** Only kind that theorists do!

We consider two inertial frames K and K' moving relative to each other with speed v.

At the instant that the two origins coincide, we set both clocks to zero, i.e., their worldlines cross at the event $(x = 0,\ ct = 0)$, $(x' = 0,\ ct' = 0)$ and a light pulse is emitted.

The equations that describe the propagation of the light pulse must be of the same form in each frame (Postulate 1).

We have, if both observers describe the light worldline:

same c

$$c^2t^2 - x^2 - y^2 - z^2 = s^2 = 0 \quad \text{in } K$$

$$c^2t'^2 - x'^2 - y'^2 - z'^2 = s'^2 = 0 \quad \text{in } K'$$

We have explicitly used the second postulate at this point.

These equations state that the vanishing of the spacetime interval between two events in any inertial frame implies the vanishing of the interval between the same two events in any other inertial frame.

However, we want to prove a more powerful statement, namely, that

$$s^2 = s'^2 \qquad \text{in general!}$$

We now use the third postulate. But first, mathematics to the rescue…..

linear transformations only

A general theorem from the mathematics of quadratic forms or a lot of messy algebra then says that the quadratic forms s^2 and s'^2 can be connected, at most, by a proportionality factor

$$s'^2 = k(x, y, z, t, \vec{v})s^2$$

does not rule out much as yet!!

Now all physical theories assume that for a free particle

(1) the laws of motion are independent of the choice of origin for the coordinate system

(2) the laws of motion are independent of the orientation of the coordinate system

(3) its velocity during any time interval is the same

These are rules that correspond to the statement spacetime is **homogeneous**.

on large scale

This implies that the proportionality factor can only depend on $\vec{v}$, i.e.,

$$s'^2 = k(\vec{v})s^2$$

Physicists also assume that space is **isotropic**, which means we cannot have a dependence on the direction of $\vec{v}$. Thus, we must have

on large scale

$$s'^2 = k(v)s^2$$ where v is the magnitude of $\vec{v}$.

Now, if we transform from K' back to K we must have the result

$$s^2 = k(v)s'^2$$

since $-\vec{v}$ has the **same** magnitude as $\vec{v}$.

Putting these two results together we have $k^2 = 1 \rightarrow k = \pm 1$.

k is a constant, but which one?

If we let $v \rightarrow 0$, then the systems K and K' become identical and hence $k(0) = 1$ and thus $k = +1$ for all v.

We have thus proved that $s^2 = s'^2$ in general. From first principles!!

Once we have invariance of spacetime interval and linearity of transformation equations between frames it is straightforward to derive Lorentz transformations and all other results follow.

That was an example of how a theorist works........

$$ct' = \gamma(ct - \beta x) \qquad x' = \gamma(x - \beta ct) \qquad y' = y \qquad z' = z$$

$$\beta = \frac{v}{c} \quad , \quad \gamma = \sqrt{\frac{1}{1-\beta^2}}$$

$$(\Delta s)^2 = c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$$

$$(\Delta s_A)^2 = c^2(t_{A2} - t_{A1})^2 - (x_{A2} - x_{A1})^2 - (y_{A2} - y_{A1})^2 - (z_{A2} - z_{A1})^2$$

$$(\Delta s_B)^2 = c^2(t_{B2} - t_{B1})^2 - (x_{B2} - x_{B1})^2 - (y_{B2} - y_{B1})^2 - (z_{B2} - z_{B1})^2$$

$$(\Delta s_A)^2 = (\Delta s_B)^2$$

Observer A: $(ct_{A1}, x_{A1}, y_{A1}, z_{A1})$ and $(ct_{A2}, x_{A2}, y_{A2}, z_{A2})$

Observer B: $(ct_{B1}, x_{B1}, y_{B1}, z_{B1})$ and $(ct_{B2}, x_{B2}, y_{B2}, z_{B2})$

$$ct_{B1} = \gamma(ct_{A1} - \beta x_{A1}) \ , \ x_{B1} = \gamma(x_{A1} - \beta ct_{A1}) \ , \ y_{B1} = y_{A1} \ , \ z_{B1} = z_{A1}$$

$$ct_{B2} = \gamma(ct_{A2} - \beta x_{A2}) \ , \ x_{B2} = \gamma(x_{A2} - \beta ct_{A2}) \ , \ y_{B2} = y_{A2} \ , \ z_{B2} = z_{A2}$$

Now back to being “physical”!

Reference frames: examples

Situation

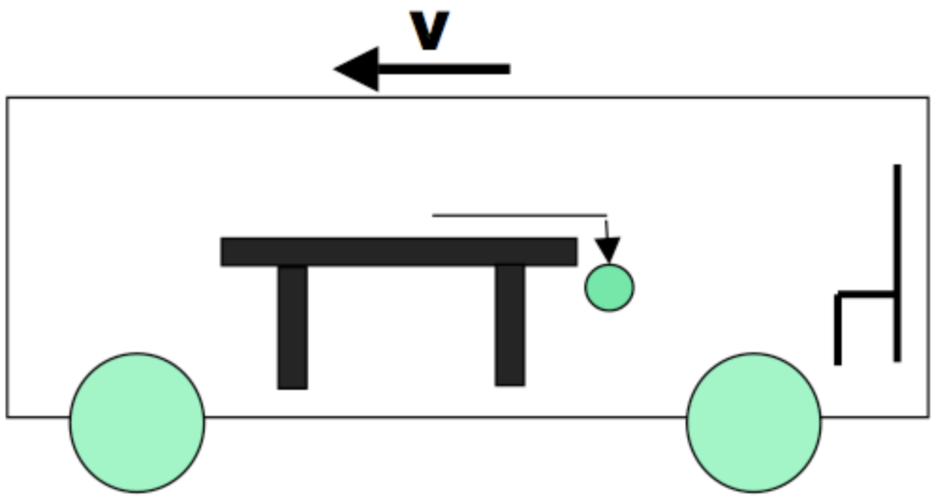
- A train is moving with velocity v wrt to a station
- A table is anchored to the train
- A ball is falling from the table

We can identify 3 systems of reference and 3 observers:

- Observer 1: sitting on a bench at the station
- Observer 2: sitting on a chair on the train
- Observer 3: a bug sitting on top of the falling ball

Who are the observers in an inertial reference frame?

- Observers 1 and 2
- Observer 3 is not: the ball is falling with acceleration g



Is time the same in all reference frames?

The (apparently) innocent S.R. assumptions have amazing consequences such as time is not absolute!

Problem

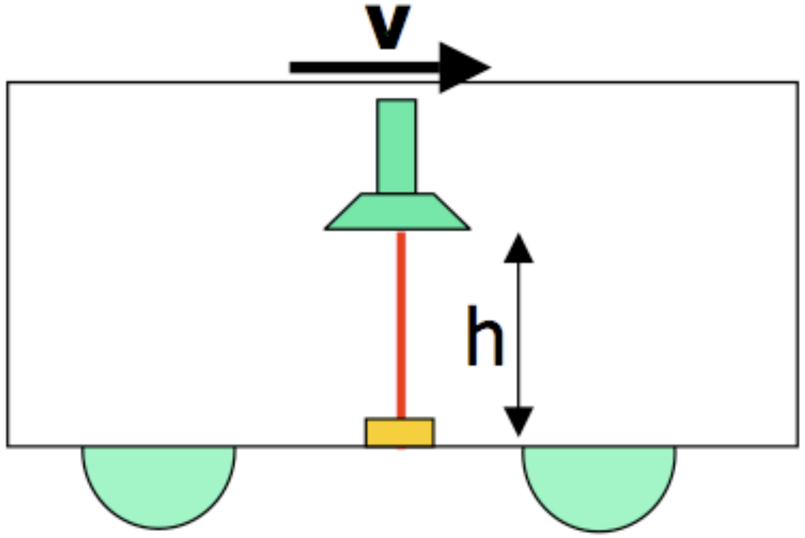
The train is moving with velocity v || x-axis

- Observer 1: standing in the train
- Observer 2: at the station

Observer 1 flashes a pulse of light vertically to a photosensor mounted on the floor of the train

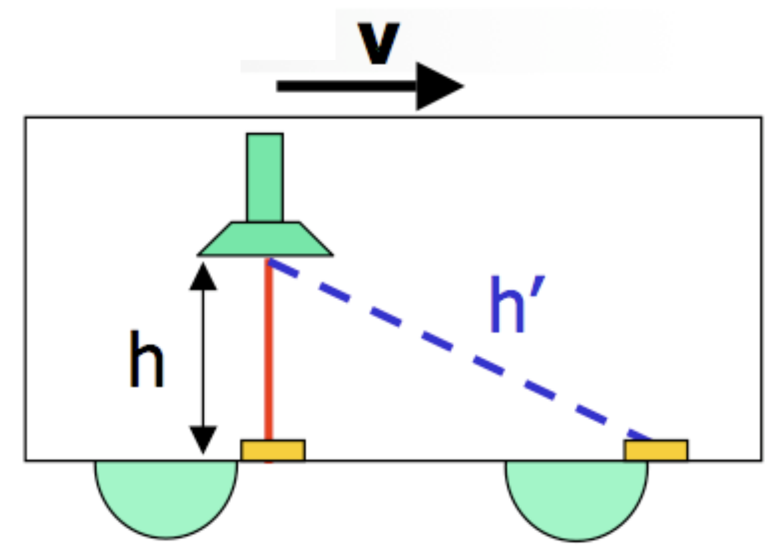
Both observers measure the time between when the light is emitted and when the light reaches the sensor

Will the 2 observers measure the same time?



Time in different reference frames

Let's calculate time measured by the 2 observers



Train reference frame (observer 1)

Distance traveled by light: h

Velocity of light: c

$$\Rightarrow \Delta t' = \Delta t_1 = \frac{h}{c}$$

Station reference frame (observer 2)

Distance traveled by light: $h' = \sqrt{h^2 + (v\Delta t_2)^2}$

Velocity of light: c

$$\Rightarrow \Delta t' = \Delta t_2 = \frac{h'}{c}$$

$$(\Delta t_2)^2 = \left(\frac{h'}{c}\right)^2 = \frac{h^2 + (v\Delta t_2)^2}{c^2} = (\Delta t_1)^2 + \frac{v^2}{c^2}(\Delta t_2)^2 \Rightarrow \Delta t_1 = \Delta t_2\sqrt{1 - \frac{v^2}{c^2}}$$

Defining

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \Rightarrow \Delta t' = \gamma \Delta t$$

Time dilation

We just derived a very important result!

Gamma factor:

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \beta^2}} > \quad \text{with } \beta = \beta_v = \frac{v}{c}$$

Since Δt' = γΔt → Δt' is always larger than Δt

Δt' = time measured by the observer in the station who sees the clock in motion

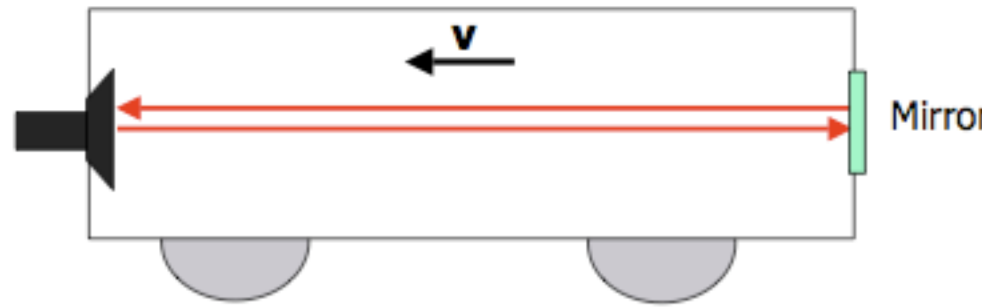
Δt = time measured by the observer on the train, at rest wrt the clock

Conclusion: Clocks in motion run slower (time dilation) $\Delta t' = \gamma \Delta t$

Length in different reference frames

Problem 2

- Now observer 1 flashes a pulse of light horizontally from left end of the train
- The light is reflected by a mirror on the right end wall and detected by a photosensor on the left wall

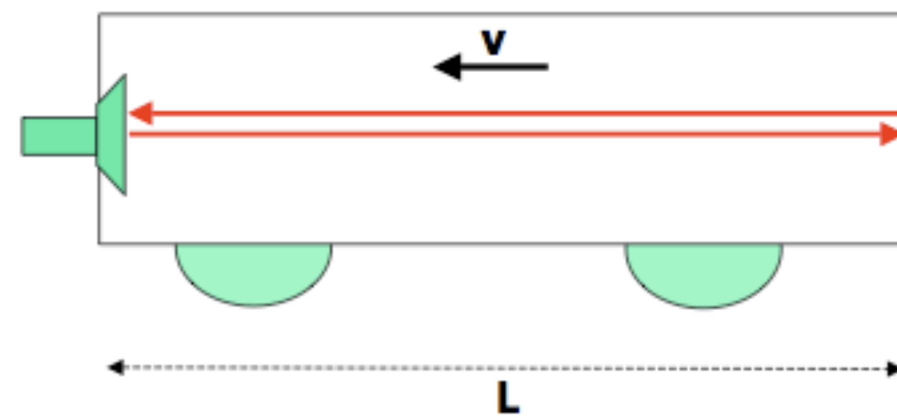


What is the length of the train measured by each observer?

For observer in train reference frame

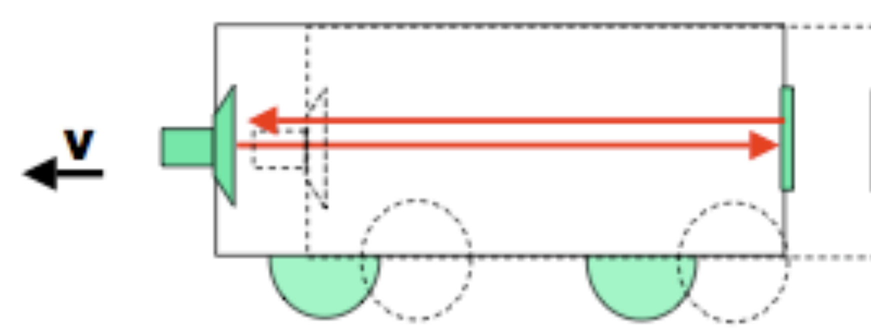
- Events we are interested in: emission and reception of light
- Time in between the two: $\Delta t = \Delta t_{train}$
- Length of the train:

$$L = \frac{c\Delta t}{2} \Rightarrow \Delta t = \frac{2L}{c}$$



Length in the station reference frame

Calculate separately Δx_1(L→R) and Δx_2(R→L)



$$\Delta t'_1 = (L' - v\Delta t'_1)/c$$

$$\Delta t'_2 = (L' + v\Delta t'_2)/c$$

Δt_1 is shorter because train and light move in opposite directions
Δt_2 is longer because train and light move in the same direction
$L'(t')$ = length (time) measured from station reference frame

Rearrange terms: $$\Delta t'_1 = \frac{L'}{c+v} \quad , \quad \Delta t'_2 = \frac{L'}{c-v}$$

Length contraction

Total time in the station reference frame = sum of $\Delta t'_1$ and $\Delta t'_2$:

$$\Delta t' = \Delta t'_1 + \Delta t'_2 = \frac{L'}{c+v} + \frac{L'}{c-v} = L'\frac{2c}{c^2 - v^2} = \frac{2L'\gamma^2}{c}$$

Remember how time dilates: $\Delta t' = \gamma \Delta t \rightarrow$

$$\frac{2L'\gamma^2}{c} = \Delta t' = \gamma \Delta t = \gamma \frac{2L}{c} \Rightarrow L' = \frac{L}{\gamma}$$

Since $\gamma > 1 \rightarrow$ Moving objects appear contracted (length contraction)

Remember: nothing physically contracts!

Summary so far

Assume Special Relativity postulates hold:
- The laws of physics are the same for all reference frames
- The speed of light is the same (c) in all reference frames

Consequences:
- Time dilation : clocks in motion run slower $\Delta t' = \gamma \Delta t$
- Length contraction : moving objects appear contracted $L' = \frac{L}{\gamma}$

REALLY??? Can we check this experimentally???

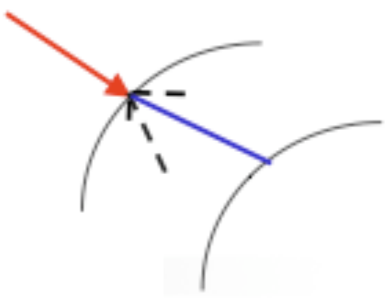
Application: Cosmic Ray Muons μ

Cosmic ray muons:

- Cosmic rays are energetic particles (mainly protons) coming from somewhere in the Universe
- When they hit the atmosphere they will produce showers of particles
- μ of particular interest because very penetrating and have long lifetime (2.2 μs)

Question: Can muons produced in the upper atmosphere reach the ground?

- Input:
  - Muon's velocity = 99.99% of velocity of light c
  - Atmosphere ~ 20 Km thick



Non relativistic approach:

$\Delta l = 0.9999c\Delta t = 0.6$ Km < 20 Km: NO, they cannot reach the ground

Relativistic approach

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \sim 71$$

Approach 1: our perspective

$\tau_\mu = 2.2\ \mu s$ in muon's reference frame

In our reference frame: $\tau' = \tau\gamma = 71 \times 2.2\ \mu s = 156\ \mu s$

→Now muon can travel: $\Delta l = 42$ Km: OK!

Approach 2: muons' perspective

The $\Delta l' = 20$ Km of atmosphere appear contracted

$\Delta l = \Delta l'/\gamma = 20\text{Km}/71 \sim 0.3$ Km that can easily be traveled in $\tau = 2.2\ \mu s$: OK!

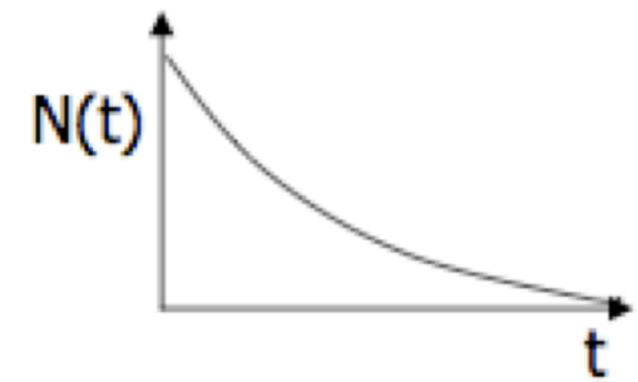
Relativity: same physics in all reference frames!

More on Cosmic Ray Muons

The rate of cosmic muons detected at sea level and on the top of Mount Everest is different. By how much?

Hypotheses:

- Muons are produced in the upper atmosphere: ~ 20 Km
- $\beta = 0.9999 \rightarrow \gamma = 1/\mathrm{sqrt}(1-v^2/c^2) \sim 71$
- Mount Everest ~ 8 Km
- Muons decay exponentially $N(t) = N_0 \exp(-t/\tau)$



Choose 1 Reference Frame and stay with it. →our RF

- $\tau'_\mu = 156\ \mu s$ in our R.F.
- At sea level:
  - $L = 20\mathrm{Km} \rightarrow T = 66\ \mu s \rightarrow N_{sea} = N_0 \exp(-66/156) = 0.65 N_0$
- On Mount Everest:
  - $L = 12\mathrm{Km} \rightarrow T = 40\ \mu s \rightarrow N_{Everest} = N_0 \exp(-40/156) = 0.77 N_0$

→At sea level expect ~ 15% less cosmic μ than on Mount Everest: OK!

How do lengths perpendicular to v transform?

Thought experiment

- Train moving towards a tunnel with velocity v = 0.9c
- Height of train in train's RF: $h_{train} = 3.5$ m
- Height of tunnel in tunnel's RF: $h' = 4.0$ m

If we have Lorentz contractions: $L' = L/\gamma$

$\gamma = 1/\text{sqrt}(1-0.9^2)=2.29$

In tunnel's reference frame: the train moves with $\beta = 0.9$

→$h'_{train} = h_{train}/\gamma = 3.5/2.29 = 1.5\text{m}$ →no problem: it will fit!

In train's reference frame: tunnel moves with velocity $\beta=0.9$

→$h_{tunnel} = h'_{tunnel}/\gamma = 4/2.29 = 1.7\text{m} < h_{train}$→they will smash!

→Different observers come to different conclusions →against relativity principle! →Lorentz contraction cannot happen when L perpendicular to v

WOW!!

powerful theory use of destructive measurements!

Lorentz transformation

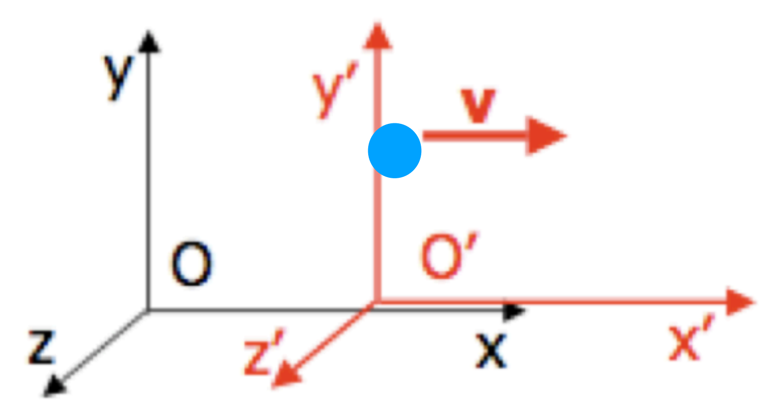
Time dilation and Length contraction are consequences of the Lorentz transformation

Consider 2 inertial reference frames: O and O′

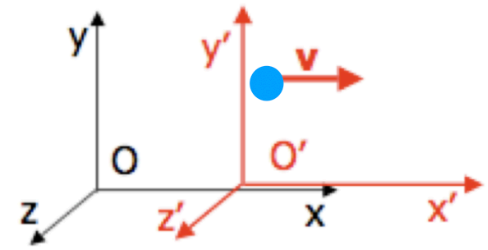
O′ is moving wrt O with velocity v || x-axis where

(x,y,z,t) are the coordinates in the $O_{reference}$ frame

(x′,y′,z′,t′) are the coordinates in the $O'_{reference}$ frame



Our path to the Lorentz transformation begins with the application of Einstein's postulates to a very familiar law of physics.



The figure show the essentials.

We assume that at time zero the object was at the point where the origins coincided.

With no net external force, it must travel at constant velocity.

Its position must be a linear function of time in both frames.

$$x' = u't' \qquad x = ut \tag{1}$$

Avoiding unnecessary assumptions, we allow that a given event in the object's life may occur at different locations and times in the two frames.

Suppose that the object beeps regularly.

We seek a relationship that would give us x' and t' of each such event in O' from its x and t in frame O.

At this point, our only guidelines are that constant velocity in one frame - position being a linear function of time - should imply constant velocity in the other.

As it turns out. this implies a fairly simple relationship: (x',t') must be related to (x,t) by a linear transformation

$$x' = Ax + Bt \qquad t' = Cx + Dt \tag{2}$$

where A,B,C, and D are constants, i.e., it is easy to show from these equations that if x/t is constant, the x'/t' is constant.

The main point is that for a free object to move at constant velocity in all frames, we require only that the transformation is linear.

It is not required that time be absolute - that t'= t - as in the Galilean transformations!

To determine the constants A, B, C, and D, we need to consider only three special cases of motion.

1. If the object were fixed at the origin of frame O', it would move at speed v relative to frame O, so equations (1) would become x'=0 and x=vt.

Inserting these into the first of equations (2), we obtain

$$B = -Av \tag{3}$$

2. If the object were fixed at the from O origin, it would have a velocity of -v according to someone in frame O', so equations (1) would be x'=-vt' and x=0.

Inserting these into (2) gives

$$D = -\frac{B}{v} \qquad \text{or} \qquad D = A \tag{4}$$

The algebra to derive these equations is simple.

Thus far, we have merely set up two frames with relative motion.

We have not invoked anything particularly “relativistic”.

With our third special case comes the crucial element.

3. If the object were a beam of light, Einstein’s second postulate demands that its speed be c in both frames.

Equations (1) become

$$x' = ct' \qquad x = ct \tag{5}$$

Inserting these into (2) and using (3) and (4) gives

$$C = -\frac{v}{c^2}A \tag{6}$$

With B, C, and D now in terms of A, we may write equations (2) as

$$x' = A(x - vt) \qquad t' = A\left(-\frac{v}{c^2}x + t\right) \tag{7}$$

We can now deduce A by a fairly simple argument if we first solve equations (7) for x and t.

$$x = \frac{1}{A\left(1 - \frac{v^2}{c^2}\right)}(x' + vt') \qquad t = \frac{1}{A\left(1 - \frac{v^2}{c^2}\right)}\left(+\frac{v}{c^2}x' + t'\right) \tag{8}$$

The only difference between the frames is that O' moves at +v relative to frame O, while O moves at -v relative to O'.

It follows that (7) and (8) have to be identical except for the sign of v, and this simply requires that their leading coefficients be equal.

$$\frac{1}{A\left(1 - \frac{v^2}{c^2}\right)} = A \quad \text{or} \quad A = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma_v \tag{9}$$

This factor is ubiquitous in special relativity.

It is often a measure of the departure of relativistic expectations from classical ones.

It increases continuously from essentially 1 at small speeds toward infinity as v approaches c.

Finally, inserting (9) into (7) and (8), we have the Lorentz transformation equations (dropping the subscript v from gamma).

$$x' = \gamma(x - vt) \qquad t' = \gamma\left(-\frac{v}{c^2}x + t\right) \tag{10}$$

$$x = \gamma(x' + vt') \qquad t = \gamma\left(+\frac{v}{c^2}x' + t'\right) \tag{11}$$

Note that equations (10) and (11) are actually equivalent but simply solved for different quantities.

We include both merely for convenience.

It is fairly easy argued that spatial coordinates perpendicular to the direction of relative motion are the same in both frames.

Thus

$$y' = y \qquad z' = z$$

Transformation of velocity

Consequence of Lorentz transformations
Observer in motion O' shoots a bullet with velocity u'_x || +x-axis
What is the velocity of the bullet u_x measured by O?

$$u_x = \frac{dx}{dt} = \frac{d\left(\gamma\left(x'+vt'\right)\right)}{d\left(\gamma\left(t'+\frac{v}{c^2}x'\right)\right)} = \frac{dx'+vdt'}{dt'+\frac{v}{c^2}dx'} = \frac{dx'/dt'+v}{1+\frac{v}{c^2}dx'/dt'} = \frac{u'_x+v}{1+\frac{u'_x v}{c^2}}$$

Conclusion:

$$u_x = \frac{u'_x+v}{1+\frac{u'_x v}{c^2}} \quad and \quad u'_x = \frac{u_x - v}{1-\frac{u_x v}{c^2}}$$

—> c' = c

Velocity not || to v

How do we deal with velocity not || to the relative motion of the 2 R.F.?
Observer in motion O' shoots a bullet with velocity u'_y perpendicular to v
What is the velocity of the bullet u_y measured by O?

$$u_y = \frac{dy}{dt} = \frac{dy'}{d\left(\gamma\left(t'+\frac{v}{c^2}x'\right)\right)} = \frac{dy'}{\gamma\left(dt'+\frac{v}{c^2}dx'\right)} = \frac{dy'/dt'}{\gamma\left(1+\frac{v}{c^2}dx'/dt'\right)} = \frac{u'_y}{\gamma\left(1+\frac{u'_x v}{c^2}\right)} \Rightarrow \frac{u'_y}{\gamma} \quad if \;\; u'_x = 0$$

When the velocity of the bullet has both x and y components:

$$u_y = \frac{u'_y}{\gamma\left(1+\frac{u'_x v}{c^2}\right)} \quad and \quad u'_y = \frac{u_y}{\gamma\left(1-\frac{u_x v}{c^2}\right)}$$

Momentum and Energy: definition

For a particle of mass m moving with velocity u

Classical definitions:

Momentum: $\vec{p} = m\vec{u}$

Kinetic energy: $E_{kin} = \frac{1}{2}mu^2$

Relativistic definition

Momentum: $\vec{p} = \gamma_u m\vec{u}$

Energy: $E = \gamma_u mc^2$

where γ_u is the relativistic γ factor: $\gamma_u = 1/\text{sqrt}(1-u^2/c^2)$

For low velocities, the new formulae reproduce old ones (Taylor!)

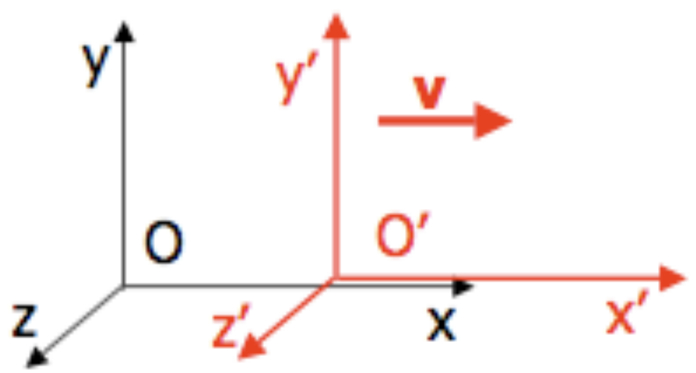
$$\vec{p} = \gamma_u m\vec{u} \sim \left(1 + \frac{1}{2}\frac{u^2}{c^2} - .....\right)m\vec{u} \sim m\vec{u}$$

$$E = \gamma_u mc^2 \sim \left(1 + \frac{1}{2}\frac{u^2}{c^2} - .....\right)mc^2 \sim mc^2 + \frac{1}{2}mu^2 = mc^2 + E_{kin}$$

Transformation of p and E

Consider 2 inertial reference frames: O and O'

O' is moving wrt O with velocity v || x-axis



How do momentum and energy Lorentz transform?

One can demonstrate that

$$E' = \gamma_v(E - \beta_v c p_x)$$

$$p'_x = \gamma_v(p_x - \beta_v E / c)$$

$$p'_y = p_y$$

$$p'_z = p_z$$

Let us see this in more detail:

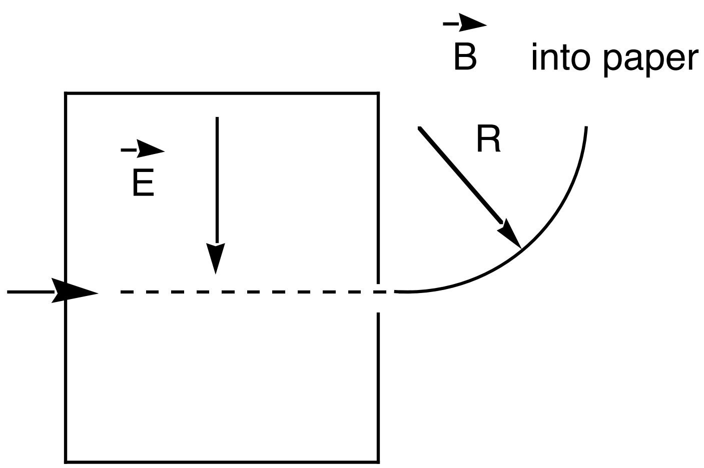
What happens to momentum and energy in SR? At first, use experiment to point the way. Later derive these results from first principles using interval and linear algebra.

The following result has been confirmed by experiment:

The force felt by a charged particle in electric and magnetic fields is given by the Lorentz force law

$$\vec{F} = q\left(\vec{E} + \frac{\vec{v}}{c} \times \vec{B}\right)$$

where $\vec{v}$ is the particle velocity. Consider the experimental setup below:



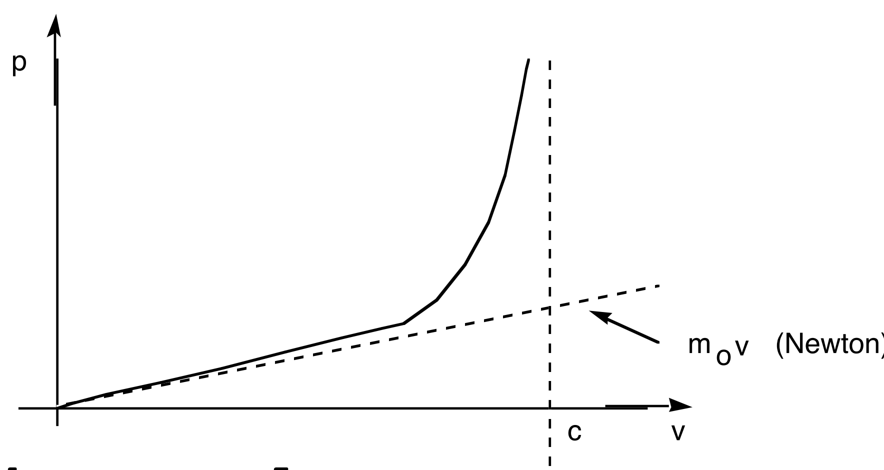
In box region, electric and magnetic fields adjusted so force = 0 for particle moving along dotted line with definite velocity. The electric force always points downward and magnetic force is always perpendicular to the velocity direction(upward in box for particle moving along dotted line). This means that particles with a particula velocity, namely,

$$q(-E + \frac{v}{c}B) = 0 \rightarrow \frac{v}{c} = \frac{E}{B}$$

pass undeflected though the box. The box is called a velocity selector. Outside the box there is no electric field, so the particle moves on a circular path (Force always perpendicular to the velocity) with a radius of

$$R = \frac{pc}{qB}$$

So measuring the radius corresponds to measuring the relativistic momentum. Thus, in the same experiment we can measure **both** the velocity and momentum **independently** and thus determine the relationship between them. A plot of experimental results looks like



This corresponds to the result

$$p = \gamma m_0 v \quad \text{where} \quad \gamma = \frac{1}{\sqrt{1-\beta^2}}$$

instead of the Newtonian assumption that $p = m_0 v$ where m_0 is the rest mass(only valid mass for a particle since we measure mass when a body is at rest). Any measurement of mass when a particle is moving is really a measurement of momentum and it would be incorrect to assume a different mass value for a moving object. There is no such thing as the "**relativistic mass**".

Therefore the relativistically correct form of the momentum is

$$\vec{p} = m_0\gamma(v)\vec{v}$$

Now what about relativistic energy? What is the relativistically correct form of the energy of a particle?

One way to generalize the concept of energy is to use the Newtonian definition of kinetic energy in conjunction with the relativistically correct definition of momentum. We proceed as follows: the formal definition of kinetic energy is given as

$$\Delta K = K - K_0 = work\ done\ by\ force = \int_{\vec{r}_0}^{\vec{r}} \vec{F}\cdot d\vec{r} = \int_{\vec{r}_0}^{\vec{r}} \frac{d\vec{p}}{dt}\cdot d\vec{r}$$

$$\vec{p} = m_0\gamma(v)\vec{v} \quad where \quad \gamma(v) = (1-\beta^2)^{-1/2}, \beta = \frac{v}{c}$$

Therefore

$$K - K_0 = \int_{\vec{r}_0}^{\vec{r}} \frac{d\left(m_0\gamma(v)\vec{v}\right)}{dt}\cdot\vec{v}dt = m_0\int_0^v \vec{v}\cdot d\left(\gamma(v)\vec{v}\right)$$

Since the kinetic energy = 0 when velocity = 0, we finally have

$$K = m_0\int_0^v \vec{v}\cdot d\left(\gamma(v)\vec{v}\right)$$

Now since

$$d(\gamma v^2) = d(\gamma\vec{v}\cdot\vec{v}) = \vec{v}\cdot d\left(\gamma\vec{v}\right) + \gamma\vec{v}\cdot d\vec{v}$$

we can write

$$K = m_0 \int_0^v (d(\gamma v^2) - \gamma \vec{v} \cdot d\vec{v}) = m_0 \int_0^v d(\gamma v^2) - m_0 \int_0^v \gamma \vec{v} \cdot d\vec{v} = m_0 \int_0^v d(\gamma v^2) - \frac{1}{2} m_0 \int_0^v \gamma d(v^2)$$

$$= m_0 \gamma v^2 - \frac{1}{2} m_0 c^2 \int_0^{v^2/c^2} \frac{du}{\sqrt{1-u}} = m_0 \gamma v^2 + m_0 c^2 \left(\frac{1}{\gamma} - 1 \right)$$

$$= m_0 c^2 \left(\gamma \beta^2 + \frac{1}{\gamma} \right) - m_0 c^2 = m_0 c^2 (\gamma - 1)$$

Check that this makes sense. Look at low velocity limit of this expression? Using

$$\gamma = (1 - \beta^2)^{-1/2} \rightarrow 1 + \frac{1}{2}\beta^2 = 1 + \frac{1}{2}\frac{v^2}{c^2}$$

we have

$$K = m_0 c^2 (\gamma - 1) \rightarrow m_0 c^2 \frac{1}{2}\frac{v^2}{c^2} = \frac{1}{2} m_0 v^2$$

as expected. If we rearrange this result we have

$$\gamma m_0 c^2 = K + m_0 c^2 = Energy(motion) + Energy(rest) = Total\ Energy = E$$

It is only the total energy that is conserved! We thus obtain Einstein's famous relation

$$E_{rest} = m_0 c^2$$

What is the connection to momentum? Some algebra gives the following results for relativistic objects

$$\frac{pc}{E} = \frac{m_0 \gamma v c}{\gamma m_0 c^2} = \frac{v}{c} = \beta \quad and \quad \left(\frac{E}{c}\right)^2 - p^2 = (m_0 c)^2$$

Are there any new predictions we can make from these results? The two relations above make the following interesting prediction:

$$v = c \rightarrow \beta = 1 \rightarrow E = pc \Rightarrow \left(\frac{E}{c}\right)^2 - p^2 = 0 = (m_0 c)^2$$

or the only objects that can travel at the speed of light must have a rest mass = 0! However, even though they have a zero rest mass, they still possess energy and momentum defying the classical equations! Such a particle has been observed .... it is the photon or the particle of light.

Now, in general, when the velocity is changing both its magnitude and direction we have

$$\vec{F} = \frac{d\vec{p}}{dt} = \frac{d\left(m_0 \gamma(v) \vec{v}\right)}{dt} = m_0 \frac{d}{dt}\left(\frac{\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}}\right)$$

One-Dimensional Motion

$$\vec{F} = m_0 \frac{d}{dt}\left(\frac{v}{\sqrt{1-\frac{v^2}{c^2}}} \right) = m_0 \frac{dv}{dt}\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \right) + m_0 v \frac{-\frac{1}{2}\left(-2\frac{v}{c^2}\right)\frac{dv}{dt}}{\left(1-\frac{v^2}{c^2}\right)^{3/2}}$$

$$= m_0 \frac{dv}{dt} \frac{1}{\left(1-\frac{v^2}{c^2}\right)^{3/2}} = m_0 \gamma^3 \frac{dv}{dt}$$

Newton's law is modified by the factor γ^3 which has a dramatic effect as $v \to c$. Now suppose that we have a constant force F = constant. We can then integrate the equation as follows:

$$F dt = m_0 \gamma^3(v) dv \to Ft = m_0 \int_0^v \gamma^3(v) dv$$

Now

$$\frac{d}{dv}(\gamma v) = \gamma + v\frac{d\gamma}{dv}$$

$$\frac{d\gamma}{dv} = \frac{d}{dv}\left(1-\frac{v^2}{c^2}\right)^{-1/2} = \frac{\frac{v}{c^2}}{\left(1-\frac{v^2}{c^2}\right)^{3/2}} = \gamma^3 \frac{v}{c^2}$$

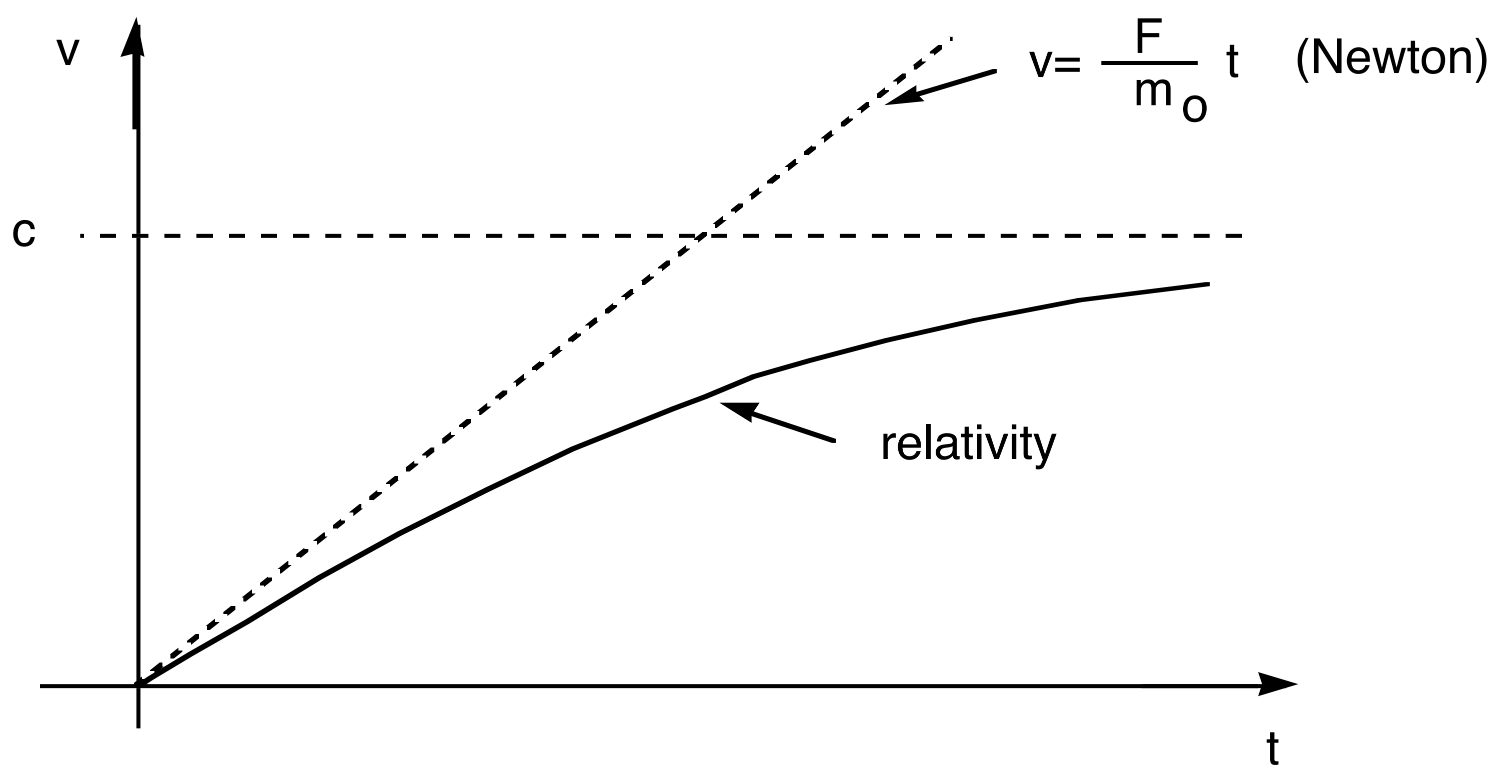
$$\frac{d}{dv}(\gamma v) = \gamma + \gamma^3 \frac{v^2}{c^2} = \gamma^3\left(\frac{1}{\gamma^2} + \frac{v^2}{c^2}\right) = \gamma^3$$

Therefore,

$$Ft = m_0 \int_0^v d(\gamma v) = m_0 \gamma v = m_0 \frac{v}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$F^2 t^2 = m_0^2 \frac{v^2}{1 - \frac{v^2}{c^2}} \rightarrow v^2 = \frac{\left(\frac{Ft}{m_0}\right)^2}{1 + \left(\frac{Ft}{m_0 c}\right)^2} \rightarrow v = \frac{dx}{dt} = \frac{F}{m_0 c} \frac{ct}{\sqrt{1 + \left(\frac{F}{m_0 c}\right)^2 t^2}}$$

A plot of v versus t is shown below.

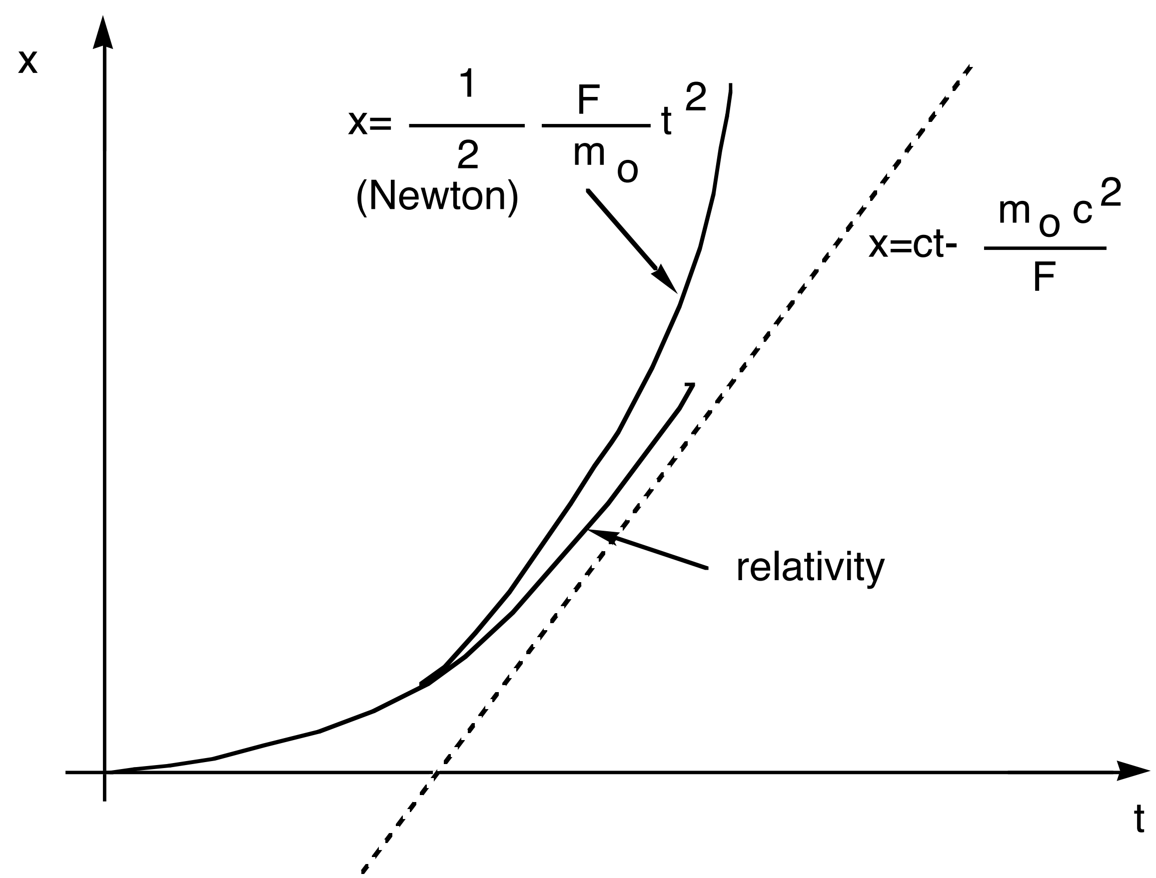


It is clear that no matter how long a constant force is applied we still have v < c. Continuing the integration

$$dx = \frac{F}{m_0 c} \frac{ct}{\sqrt{1+\left(\frac{F}{m_0 c}\right)^2 t^2}} dt \rightarrow x = \frac{F}{m_0} \int_0^t \frac{t}{\sqrt{1+\left(\frac{F}{m_0 c}\right)^2 t^2}} dt = \frac{F}{m_0} \left(\frac{m_0 c}{F}\right)^2 \times \int_0^{Ft/m_0 c} \frac{u}{\sqrt{1+u^2}} du$$

$$x = \frac{m_0 c^2}{F} \times \int_0^{Ft/m_0 c} d(\sqrt{1+u^2}) = \frac{m_0 c^2}{F} \left(\sqrt{1+\left(\frac{F}{m_0 c}\right)^2 t^2} - 1 \right)$$

A plot of x versus t is shown below.



Whenever, one does complex calculation you should check your results by calculating **limits** where the answer is known.

Letting t → 0 we get

$$\sqrt{1+\left(\frac{F}{m_0 c}\right)^2 t^2} \approx 1+\frac{1}{2}\left(\frac{F}{m_0 c}\right)^2 t^2$$

$$v \approx \frac{F}{m_0}t \quad , \quad a \approx \frac{F}{m_0} \quad , \quad \mathrm{x} \approx \frac{1}{2}\frac{F}{m_0}t^2 \quad \text{as expected}$$

Letting t → ∞ we get

$$v \to c \quad \text{and} \quad x \to ct \quad \text{as expected}$$

4-Vectors

What is an ordinary vector in 3-dimensional space?

A vector has many levels of complexity and is a very abstract mathematical object. A vector is a mathematical(geometrical) object that is representable by two numbers in two dimensions, three numbers in three dimensions, and so on. One characterization is to specify its magnitude or length and orientation or direction - imagine that it is a directed line segment. Only useful in 2D/3D - does not generalize

Consider a vector $\vec{A}$ representing some physical variable. Using cartesian unit vectors we can write

$$\vec{A} = \sum_{j=1}^{3} A_j \hat{e}_j$$

The components of the vector $A_i\ , i=1,2,3$ are its representation in a given coordinate system. We must choose a coordinate system in order to define the unit vectors. The coordinate system is not an essential part of the physics however. We can just as well use any other coordinate system to define unit vectors and the vector $\vec{A}$.

In particular, we consider another coordinate system with the same origin, but rotated from the first system. In another coordinate system we would write

$$\vec{A} = \sum_{j=1}^{3} A'_j \hat{e}'_j$$

Note that the vector $\vec{A}$ has not changed; only its representation (components) in the new system (new basis) has changed. We relate the two representations (components) as follows:

$$\sum_i A_i \hat{e}_i = \sum_i A'_i \hat{e}'_i$$

$$\hat{e}'_j \cdot \sum_i A_i \hat{e}_i = \hat{e}'_j \cdot \sum_i A'_i \hat{e}'_i = \sum_i A'_i \hat{e}'_i \cdot \hat{e}'_j = \sum_i A'_i \delta_{ij} = A'_j$$

$$A'_j = \sum_i A_i (\hat{e}'_j \cdot \hat{e}_i)$$

The coefficients $\left(\hat{e}'_j \cdot \hat{e}_i\right)$ are numbers that are determined by the specific rotation. They are independent of the vector $\vec{A}$. We now redefine a vector:

A vector in 3 dimensions is a set of 3 numbers $\{A_i\}$ (components) which transform under a rotation of the coordinate system according to

$$A'_j = \sum_i A_i(\hat{e}'_j \cdot \hat{e}_i) \qquad \text{definition in terms of transformations}$$

Any quantity which is unchanged by a coordinate transformation is called an **invariant** of the transformation. Since the principle of relativity requires that the results of physical theories (physical laws) be independent of the choice of coordinate system (must be inertial however), all physical laws must involve **only** invariants.

The dot product of two vectors is a scalar. Scalars are numbers that are independent of our choice of coordinate system. This gives us a method for **constructing** invariants. The dot product produces an invariant in the sense that

$$\vec{A}' \cdot \vec{B}' = \vec{A} \cdot \vec{B}$$

In particular, the norm or length-squared of a vector, $A^2 = \vec{A} \cdot \vec{A}$ is a scalar invariant. We now define a rotation

A rotation is any transformation which leaves

$$r^2 = \vec{r} \cdot \vec{r} = x^2 + y^2 + z^2$$

invariant

In Minkowski 4-dimensional spacetime we define vectors in a different manner. Both the ordinary space 3-dimensional and the Minkowski 4-dimensional vector definitions are special cases of a more general definition. The ordinary 3-dimensional definition corresponds to Euclidean geometry.

In Minkowski 4-dimensional spacetime we write the spacetime 4-vector in this way

$$\overleftarrow{s} = (ct, x, y, z)$$

and the scalar product of the vector with itself (its norm) as

$$\overleftarrow{s} \cdot \overleftarrow{s} = c^2t^2 - x^2 - y^2 - z^2 \quad \text{(note the minus signs)}$$

This is a scalar invariant under Lorentz transformations(it is the spacetime interval). In fact, any set of 4 numbers

$$\overleftarrow{A} = (A_0, A_1, A_2, A_3)$$

represents a Minkowski 4-vector if its norm defined by

$$\overleftarrow{A} \cdot \overleftarrow{A} = A_0^2 - A_1^2 - A_2^2 - A_3^2$$

is a scalar invariant. In addition, if a set of 4 numbers is a 4-vector then the components transform between frames via the Lorentz transformations as

$$A'_0 = \gamma(A_0 - \beta A_1)$$

$$A'_1 = \gamma(A_1 - \beta A_0)$$

$$A'_2 = A_2$$

$$A'_3 = A_3$$

for relative motion along the 1-axis.

It is in this sense that spatial and time variables are **not distinct entities** but are simply **different components** of the same vector and transform into each other under Lorentz transformations.

This corresponds to a non-Euclidean geometry.

Another 4-vector is $d\overleftarrow{s} = (cdt, dx, dy, dz)$ since it is the difference of two 4-vectors. Hence, its norm

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$$

is a Lorentz invariant.

A related quantity of great importance is

$$d\tau^2 = \frac{ds^2}{c^2}$$

(dividing an invariant by an invariant means that we still have an invariant). In particular,

$$d\tau^2 = dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2)$$

Consider a displacement $d\overleftarrow{s}$ between two events on the worldline of a moving particle. In the rest frame of the particle, $dx = dy = dz = 0$ and hence $d\tau = dt$ in the particle rest frame (the events are separated only by time). $d\tau$ is the time interval between the two events measured in the rest frame and is thus the **proper time.** It is a **Lorentz invariant.**

Time Dilation (the easy way)

Consider an observer at rest in x',y',z',t' system. In this system the proper time between two events is $d\tau = dt'$. In the x,y,z,t system moving with velocity v relative to the first frame, the time interval between the same two events is given by

$$d\tau^2 = dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2)$$

But $d\tau$ is an invariant or its value is the same in all frames. We therefore have

$$dt'^2 = dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2)$$

$$\left(\frac{dt'}{dt}\right)^2 = 1 - \frac{1}{c^2}\left(\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right)$$

$$= 1 - \frac{v^2}{c^2} = \frac{1}{\gamma^2}$$

Therefore, $dt = \gamma dt'$ which is the time dilation formula.

We did not need to introduce hypothetical experiments or discussions of simultaneity to obtain this result. That is an example of the power of using 4-vectors.

Other 4-Vectors

Using $d\overleftarrow{s} = (cdt, dx, dy, dz)$ and dividing by the Lorentz invariant $d\tau$ yields another 4-vector

$$\frac{d\overleftarrow{s}}{d\tau} = \left(c\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau}\right) = \overleftarrow{u} = 4 - vector\ velocity$$

Its norm is an invariant so it can be calculated in any frame. We pick the rest frame where

$$\overleftarrow{u} = (c,0,0,0) \rightarrow u^2 = c^2 = \text{invariant}$$

For a moving particle where the x,y,z,t system moves with velocity -v relative to the rest frame of the particle we have $dt = \gamma d\tau$ and thus

$$\overleftarrow{u} = \left(c\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau} \right) = (\gamma c, \gamma \vec{v})$$

Since the rest mass m_0 is a Lorentz invariant, $m_0\overleftarrow{u}$ is a 4-vector with dimensions of momentum. We define the 4-momentum as

$$\overleftarrow{\rho} = m_0\overleftarrow{u} = m_0\gamma(c, \vec{v}) = \left(\frac{E}{c}, \vec{p} \right)$$

We already saw that

$$\rho^2 = \left(\frac{E}{c} \right)^2 - p^2 = (m_0 c)^2 = \text{invariant}$$

Since the variables $E\ and\ \vec{p}$ are components of a 4-vector they must obey the Lorentz transformations

$$\frac{E'}{c} = \gamma\left(\frac{E}{c} - \beta p_x \right)$$

$$p'_x = \gamma\left(p_x - \beta\frac{E}{c} \right)$$ WOW!!!!

$$p'_y = p_y$$

$$p'_z = p_z$$

We will use these relations to prove that a magnetic field is observed in frames moving relative to fixed charged particles whereas only electric fields are observed in the rest frame of the charged particles. Magnetic fields are a consequence of special relativity!!

Finally we confirm our identification of the energy. We define the 4-vector Minkowski force as

$$\overleftarrow{\phi} = \frac{d\overleftarrow{\rho}}{d\tau} = \left(\frac{d\gamma m_0 c}{d\tau}, \frac{d\vec{p}}{d\tau} \right)$$

If dt is the time interval in the observer's frame corresponding to the interval of proper time $d\tau$ then $dt = \gamma d\tau$ and we get

$$\overleftarrow{\phi} = \gamma \left(\frac{d\gamma m_0 c}{dt}, \frac{d\vec{p}}{dt} \right) = \gamma \left(\frac{d\gamma m_0 c}{dt}, \vec{F} \right)$$

With this construction, the 4-momentum is conserved(constant) when the 4-force is zero. This corresponds to energy and momentum conservation. If the 4-force is zero in one frame then it is zero in all frames and hence if energy and momentum are conserved in one frame they are conserved in all frames. In Newtonian physics

$$\vec{F} \cdot \vec{v} = \frac{dE}{dt}$$

where E = total energy. Let us look at the corresponding quantity in 4-dimensions

$$\overleftarrow{\phi} \cdot \overleftarrow{u} = \gamma \left(\frac{d\gamma m_0 c}{dt}, \vec{F} \right) \cdot \gamma (c, \vec{v}) = \gamma^2 \left(\frac{d\gamma m_0 c^2}{dt} - \vec{F} \cdot \vec{v} \right)$$

Now the scalar product is an invariant and thus we can evaluate it in the rest frame of the particle. In this frame $\vec{F} \cdot \vec{v} = 0$ since $\vec{v} = 0$

We also have $\frac{d\gamma m_0 c^2}{dt} = \gamma^3 m_0 v \left(\frac{dv}{dt} \right) = 0$ since v=0. Therefore

$$\overleftarrow{\phi} \cdot \overleftarrow{u} = 0 = \gamma^2 \left(\frac{d\gamma m_0 c^2}{dt} - \vec{F} \cdot \vec{v} \right) \rightarrow \vec{F} \cdot \vec{v} = \frac{d\gamma m_0 c^2}{dt} \rightarrow E = \gamma m_0 c^2$$

as we indicated earlier. In this sense the momentum and energy variables are **not distinct entities** but are simply **different components** of the same vector and transform into each other under Lorentz transformations.

Transformation of the Force

We have
$$\vec{F} = \frac{d\vec{p}}{dt}$$

$$F_x = \frac{dp_x}{dt} \quad , \quad F_y = \frac{dp_y}{dt} \quad , \quad F_z = \frac{dp_z}{dt}$$

The Lorentz transformations for $\vec{p}$ and t are given by

$$\Delta p'_x = \gamma \Delta p_x - \beta\gamma \Delta E / c \quad , \quad \Delta p'_y = \Delta p_y \quad , \quad \Delta p'_z = \Delta p_z$$

$$\Delta t' = \gamma \Delta t - \beta\gamma \Delta x / c$$

Therefore,

$$F'_x = \frac{dp'_x}{dt'} = \lim_{\Delta t' \to 0} \frac{\Delta p'_x}{\Delta t'} = \lim_{\Delta t' \to 0} \frac{\gamma \Delta p_x - \beta\gamma \Delta E / c}{\gamma \Delta t - \beta\gamma \Delta x / c} = \lim_{\Delta t' \to 0} \frac{\Delta p_x}{\Delta t} = \frac{dp_x}{dt} = F_x$$

or the force component parallel to the relative frame motion is unchanged. It has the same value in both frames.

The transverse components behave differently.

$$F'_y = \frac{dp'_y}{dt'} = \lim_{\Delta t' \to 0} \frac{\Delta p'_y}{\Delta t'} = \lim_{\Delta t' \to 0} \frac{\Delta p_y}{\gamma \Delta t - \beta\gamma \Delta x / c} = \lim_{\Delta t' \to 0} \frac{\Delta p_y}{\gamma \Delta t} = \frac{1}{\gamma}\frac{dp_y}{dt} = \frac{1}{\gamma} F_y$$

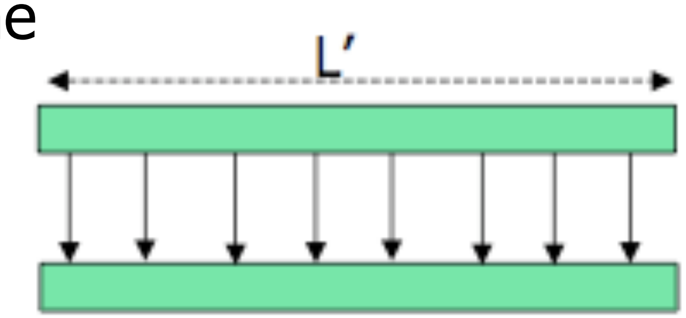
or a force component perpendicular to the relative frame motion observed in the moving frame is smaller by a factor $1/\gamma$ than the value determined by observers in the rest frame of the particle.

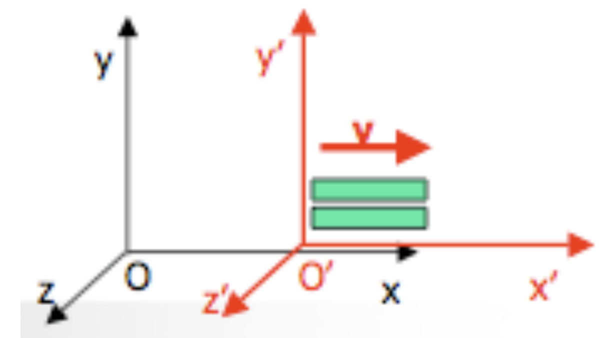
In this derivation we have used the fact that for small Δt both Δx and ΔE are proportional to $(\Delta t)^2$

Electric Fields in Motion

Consider a parallel plate capacitor

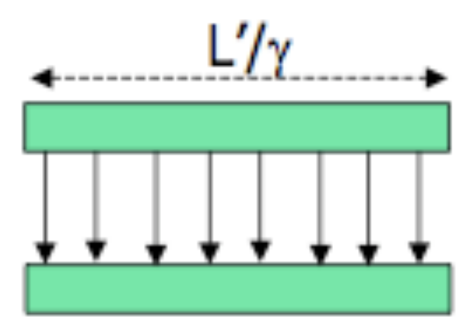
- Square plates of side L′ in the xz plane
- Charge Q′ distributed on the plates:
  - $\sigma' = Q'/L'^2$
- Electric field || y-axis:
  - $E' = 4\pi\sigma' = 4\pi Q'/A' = 4\pi Q'/L'^2$





capacitor at rest in primed frame

The capacitor is now boosted with velocity v in the x direction



How does E transform? Q=Q′ charge is Lorentz invariant

$E = 4\pi\sigma = 4\pi Q/A = 4\pi Q'/LL' = 4\pi Q'/(L'/\gamma)\, L' = \gamma 4\pi Q'/A' = \gamma E'$

Conclusion: $$E'_\perp = \frac{E_\perp}{\gamma}$$

also $\sigma = \gamma\sigma'$

Expected result: the field lines get more dense as L shrinks…

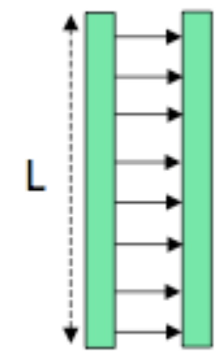
Now orient the capacitor with the plates in the yz plane:

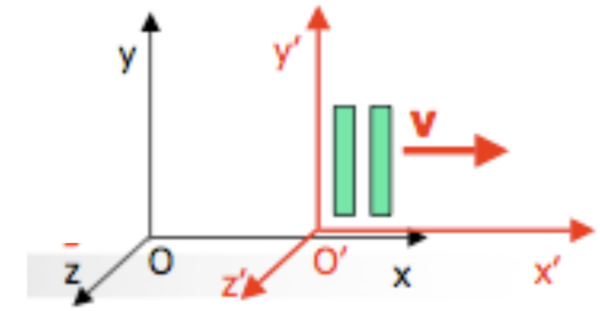
Charge Q′ distributed on the plates:

$\sigma' = Q'/L'^2$

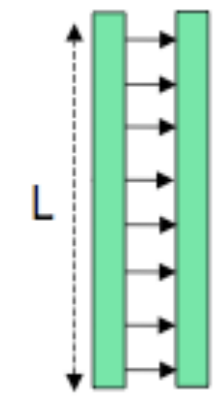
Electric field || x-axis:

$E' = 4\pi\sigma' = 4\pi Q'/L'^2$





Boost the capacitor again with velocity v in the x direction



How does E transform? Q=Q′ and A=A′

$E = 4\pi\sigma = 4\pi Q/A = 4\pi Q'/L_y L_z = 4\pi Q'/A' = E'$

Conclusion: $E'_{||} = E_{||}$ Same transformations as the FORCE!!!

Expected result: the field lines keep the same distance…

How does the capacitance change?

C proportional to A/d

In 1st case C decreases because A decreases and d is unchanged.
In 2nd case, C increases because A unchanged and d is decreases.

High Energy Particle Physics

A special reference frame is the center of mass or zero momentum system frame. It is very useful when discussing high energy particle reactions.

We consider a collision between two particles with rest masses m_1 and m_2. We assume that particle 1 is moving with velocity $\vec{u}$ in the laboratory system and that particle 2 is at rest in that system. We have the energy-momentum 4-vectors

$$\overleftarrow{p}_1 = \left(\frac{E_1}{c}, p_1, 0, 0\right) \quad and \quad \overleftarrow{p}_2 = \left(\frac{E_2}{c}, 0, 0, 0\right)$$

and the total energy-momentum

$$\overleftarrow{P} = \overleftarrow{p}_1 + \overleftarrow{p}_2 = \left(\frac{E_1 + E_2}{c}, p_1, 0, 0\right)$$

In a new frame moving along the x-axis with speed V we have

$$P'_1 = \Gamma\left(p_1 - \frac{V}{c}\frac{E_1 + E_2}{c}\right) \quad , \quad P'_2 = P'_3 = 0$$

where

$$\Gamma = \left(1 - \frac{V^2}{c^2}\right)^{-1/2}$$

In the center of mass system, $V=V_{CM}$ and $\vec{P} = 0$. This says that

$$P'_1 = \Gamma\left(p_1 - \frac{V_{CM}}{c}\frac{E_1+E_2}{c}\right) = 0 \rightarrow V_{CM} = \frac{p_1 c^2}{E_1+E_2} \qquad \text{NR:} \qquad V_{CM} = \frac{m_1 v_1}{m_1+m_2}$$

The energy available for physical processes such as the production of new particles or inelastic events is the total energy in the center of mass system, E′. In the center of mass system the total energy-momentum 4-vector is

$$\left(\frac{E'}{c},0,0,0\right)$$

We can find E′ by using the fact that the norm of the energy-momentum 4-vector is invariant

$$\left(\frac{E'}{c}\right)^2 = \left(\frac{E_1+E_2}{c}\right)^2 - p_1^2$$

or

$$E'^2 = E_1^2 + E_2^2 + 2E_1E_2 - p_1^2c^2 = E_1^2 + E_2^2 + 2E_1E_2 - \left(E_1^2 - m_1^2c^4\right)$$

$$= m_1^2c^4 + 2E_1E_2 + E_2^2$$

We have

$$E_1 = \gamma_1 m_1 c^2 \quad \text{and} \quad E_2 = m_1 c^2 \qquad , \qquad \gamma_1 = \left(1 - \frac{u^2}{c^2}\right)^{-1/2}$$

Therefore

$$E' = \left(\gamma_1 m_1 + m_2\right)c^2 = \text{ total energy in laboratory system}$$

and

$$E' = \left(m_1^2 + m_2^2 + 2\gamma_1 m_1 m_2\right)^{1/2} c^2$$

The fraction of energy available for physical processes is

$$\frac{E'}{E} = \frac{\left(m_1^2 + m_2^2 + 2\gamma_1 m_1 m_2\right)^{1/2}}{\gamma_1 m_1 + m_2}$$

For the special case $m_1 = m_2 = m$ we have

$$\frac{E'}{E} = \sqrt{\frac{2}{1+\gamma_1}}$$

At low velocity or low energy of the incident particle (the one that is moving), we have

$$\gamma_1 \approx 1 \rightarrow \frac{E'}{E} = 1 \rightarrow \text{ all energy available}$$

In this case, most of the energy is rest energy and kinetic energy is unimportant. In the high speed or high energy limit we have

$$\frac{E'}{E} = \sqrt{\frac{2}{1+\frac{E_1}{mc^2}}} \rightarrow \sqrt{\frac{2mc^2}{E_1}}$$

Thus, the useful fraction of energy decreases as $E_1^{-1/2}$. For example, in a 300 GeV accelerator (STOP and define eV)

Now

$$1\, GeV = 10^9\, eV = 1.6 \times 10^{-10}\, J$$

where

Definition of an Electron-Volt of Energy

Accelerate an electron (charge e) through a potential difference of 1 Volt over a distance L.

Electric field = E = 1/L
Force = F = eE = e/L
acceleration a = F/m = e/mL

$v^2 = 2aL = 2e/m$

$E_k = mv^2/2 = e = 1.6 \times 10^{-19}$ J = 1 eV

For example, in a 300 GeV accelerator
an accelerated proton ($mc^2 \approx 1\,GeV$) colliding with a hydrogen target(protons) has

$$\frac{E'}{E} = \rightarrow \sqrt{\frac{2}{300}} = 0.082$$

or only 25 GeV is available for reactions!!! We will show how to fix this up shortly.

Let us look at **production reactions** in another way. Suppose that we have two particles that interact with each other(one is at rest -- the target) and produce N final particles. The high energy available from the incident particle is converted into mass of newly created particles. We ask the question: What is the minimum energy needed by the incident particle in order to produce the final state of N particles?

In the final state we have

$$\left(\frac{\sum_{i=1}^{N} E_i}{c}, \sum_{i=1}^{N} \vec{p}_i \right) \quad where \quad E_i^2 = p_i^2 + m_i^2 c^4 \quad , \quad i = 1,2,3....,N$$

Now, the norm of the energy-momentum 4-vector is invariant in time and across different frames. Therefore

norm in laboratory before = norm in center of mass after

This gives

$$\left(\frac{E_{inc}}{c} + m_{\text{target}} c \right)^2 - p_1^2 = \left(\frac{\sum_{i=1}^{N} E_{i,CM}}{c} \right)^2 - \left(\sum_{i=1}^{N} \vec{p}_{i,CM} \right)^2$$

By definition, however, $\sum_{i=1}^{N} \vec{p}_{i,CM} = 0$. After some algebra we have

$$E_{inc} = \frac{\left(\sum_{i=1}^{N} E_{i,CM}\right)^2 - \left(m_{\text{inc}}c^2\right)^2 - \left(m_{\text{target}}c^2\right)^2}{2m_{\text{target}}c^2}$$

This is a minimum when $\sum_{i=1}^{N} E_{i,CM}$ is a minimum or when

$$\sum_{i=1}^{N} E_{i,CM} = \sum_{i=1}^{N} m_i c^2$$

or all the particles are at rest in the center of mass system after the collision (what are they doing in the laboratory system?).

Therefore the minimum energy needed by the incident particle (this is called the **threshold energy**) is

$$E_{inc,threshold} = \frac{\left(\sum_{i=1}^{N} m_i c^2\right)^2 - \left(m_{\text{inc}}c^2\right)^2 - \left(m_{\text{target}}c^2\right)^2}{2m_{\text{target}}c^2}$$

For example, consider the reaction

$$p + p \rightarrow p + p + \pi + \pi + \pi$$

where a proton is incident on another proton producing two protons and three pi mesons. The threshold energy is

$$E_{p,threshold} = \frac{\left(2m_p + 3m_\pi\right)^2 - 2m_p^2}{2m_p} = \left(m_p + 6m_\pi + \frac{9}{2}\frac{m_\pi^2}{m_p}\right)c^2$$

Clearly, a very non-intuitive answer (classically = $E_{p,threshold} = \left(m_p + 3m_\pi\right)c^2$!)

Now let us consider the difference between a particle accelerator where one particle is accelerated and collides with a second particle at rest (as above=laboratory system) and two particle accelerators where each particle is accelerated in the same way (colliding beams=center of mass system). We have

Single Accelerator

$$\left(\frac{E_{total,lab}}{c}, \vec{p}_{total,lab}\right) = \left(\frac{E_1 + m_2 c^2}{c}, \vec{p}_1\right) \quad , \quad E_1^2 = p_1^2 c^2 + m_1^2 c^4$$

Colliding Beams

$$\left(\frac{E_{total,CM}}{c}, \vec{p}_{total,CM}\right) = \left(\frac{2E}{c}, 0\right) \quad , \quad E = \text{energy of each particle}$$

In the first case the accelerator must produce energy E_1 and in the second case each accelerator must produce energy E.

The two accelerators are equivalent (same energy available for physical processes) if

$$\left(\frac{E_1 + m_2 c^2}{c}, \vec{p}_1 \right)^2 = \left(\frac{2E}{c}, 0 \right)^2$$

Algebra gives the result

$$E = \frac{1}{2}\sqrt{m_1^2 c^4 + m_2^2 c^4 + 2m_2 c^2 E_1}$$

If we consider the case of very high energy accelerators where

$$E_1 >> m_i c^2$$

we have

$$E = \frac{1}{2}\sqrt{2m_2 c^2 E_1}$$

Suppose we want to build a single 10 TeV accelerator(1 TeV = 10^3 GeV) so that $E_1 = 10^4\, GeV$. This is very difficult to design and requires the development of significant new equipment($$$$$$).

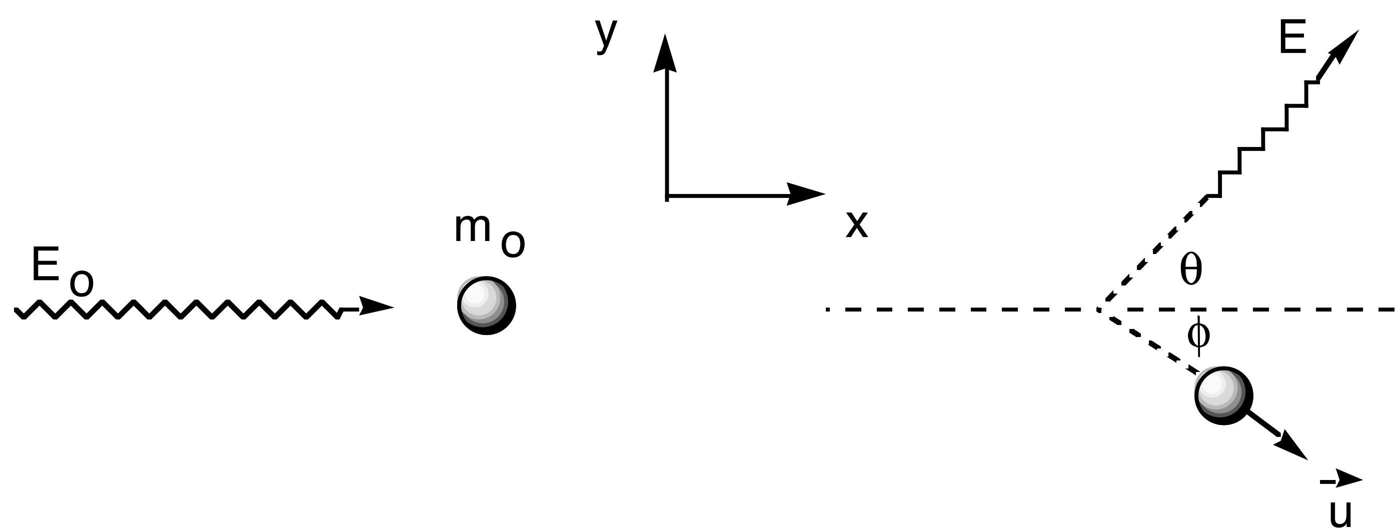
If instead we build two smaller accelerators and use them in the colliding beams configuration, then we get the same available energy with

$$E = \frac{1}{2}\sqrt{2E_1} = \sqrt{5000} = 71\, GeV$$

which we already know how to build. In fact, if we use an old single accelerator of this size that already exists, we then only have to build one small new accelerator ($$).

High Energy Collisions

Let us look at the collision processes at high energy using conservation of energy and momentum. We consider a collision in which the incident particle has zero rest mass (photon) and the target particle is at rest. If the target particle is an electron, then this is the so-called **Compton Effect.** The process looks like



The photon momentum is E_0 / c . After the collision the photon is scattered through an angle θ with energy E and the electron recoils at an angle ϕ with velocity $\vec{u}$. The final electron energy is

$$E_e = \gamma(u) m_0 c^2 = \frac{m_0 c^2}{\sqrt{1 - \frac{u^2}{c^2}}}$$

Conservation of energy gives $E_0 + m_0 c^2 = E + E_e$. Conservation of momentum gives (x and y directions)

$$\frac{E_0}{c} = \frac{E}{c}\cos\theta + p\cos\phi$$

$$0 = \frac{E}{c}\sin\theta - p\sin\phi$$

where

$$\vec{p} = \gamma m_0 \vec{u} \quad or \quad E_e^2 = p^2c^2 + m_0^2c^4$$

We want to eliminate reference to the electron and find the new photon energy(that is what is detected in the experiment).

$$\frac{E_0}{c} = \frac{E}{c}\cos\theta + p\cos\phi \rightarrow p\cos\phi = \frac{E_0}{c} - \frac{E}{c}\cos\theta \rightarrow p^2\cos^2\phi = \left(\frac{E_0}{c} - \frac{E}{c}\cos\theta\right)^2$$

$$0 = \frac{E}{c}\sin\theta - p\sin\phi \rightarrow p\sin\phi = \frac{E}{c}\sin\theta \rightarrow p^2\sin^2\phi = \left(\frac{E}{c}\sin\theta\right)^2$$

Adding these equations we get

$$p^2c^2 = E_e^2 - m_0^2c^4 = E_0^2 - 2E_0E\cos\theta + E^2$$

Using the energy conservation equation we have (after algebra)

$$E = \frac{E_0}{1 + \left(\frac{E_0}{m_0c^2}\right)(1 - \cos\theta)}$$

The first thing to note is that E > 0. This means that a free electron cannot absorb a photon completely; there will always be a scattered photon of some energy. If we convert to wavelengths using

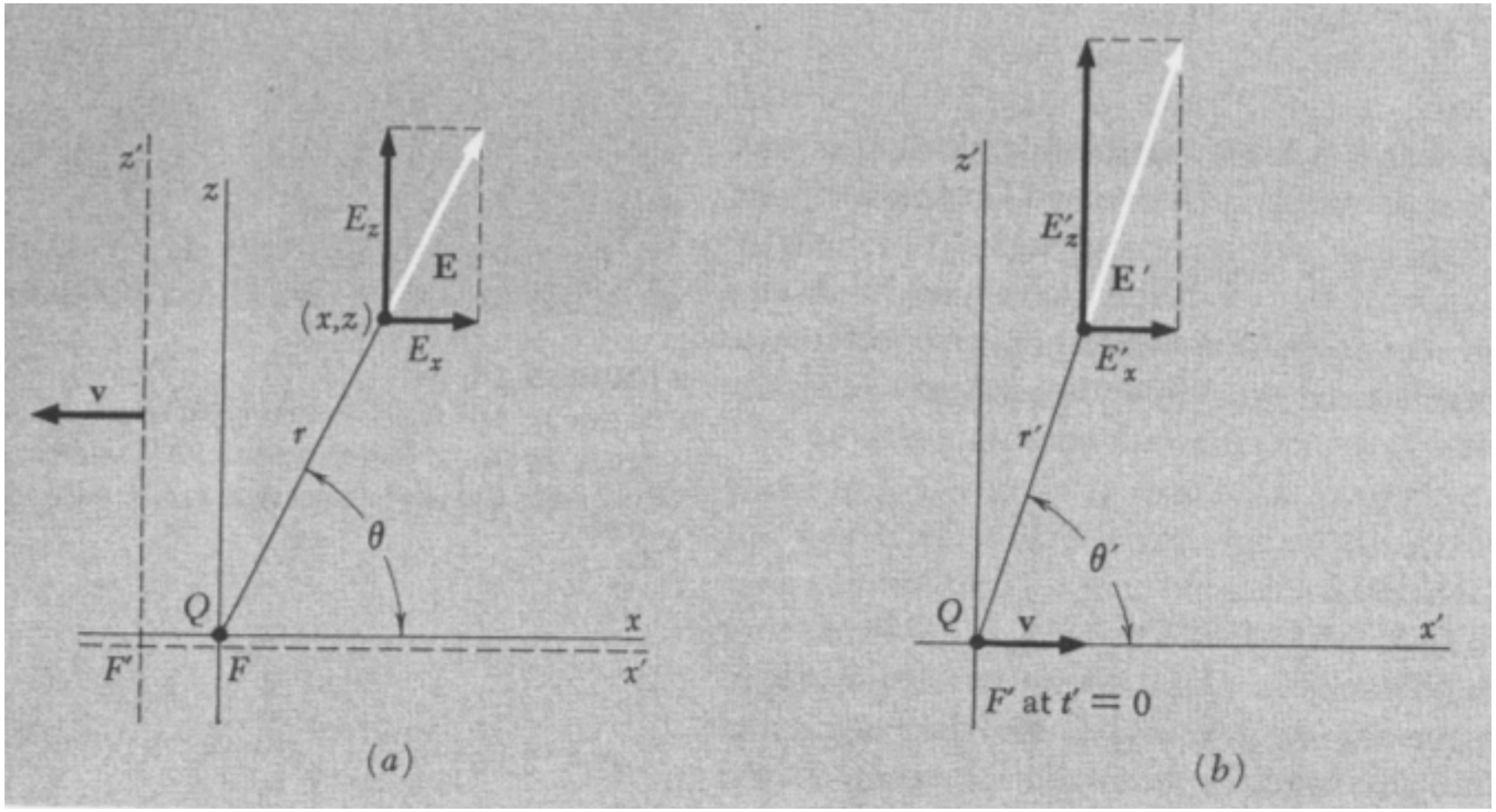
$$E = h\nu = h\frac{c}{\lambda}$$

we get

$$\lambda - \lambda_0 = \frac{h}{m_0 c}(1 - \cos\theta)$$

The shift in wavelength at a given angle is independent of the incident photon energy.

Digression.... Field of a Point Charge Moving with Constant Velocity



Electric field of point charge in rest frame of charge

Electric field of point charge in frame of charge moving with v

Purcell shows that:

$$E_x = \frac{Qx}{(x^2 + z^2)^{3/2}} \quad , \quad E_z = \frac{Qz}{(x^2 + z^2)^{3/2}}$$

$$E_x' = E_x = \frac{\gamma Q x'}{((\gamma x')^2 + z'^2)^{3/2}} \quad , \quad E_z' = \gamma E_z = \frac{\gamma Q z'}{((\gamma x')^2 + z'^2)^{3/2}}$$

$$\frac{E_z'}{E_x'} = \frac{z'}{x'} \Rightarrow$$ direction of E′ makes the same angle with the x′ axis as the radius vector r′

Hence E′ points radially outward along line drawn radially outward from INSTANTANEOUS position of Q (as in (b) above).

Think about that statement!! If Q pass origin of primed system at 12 Noon (primed time) observer anywhere in primed system reports E-field in local vicinity was pointing (at 12 Noon primed time) exactly radially from the origin.

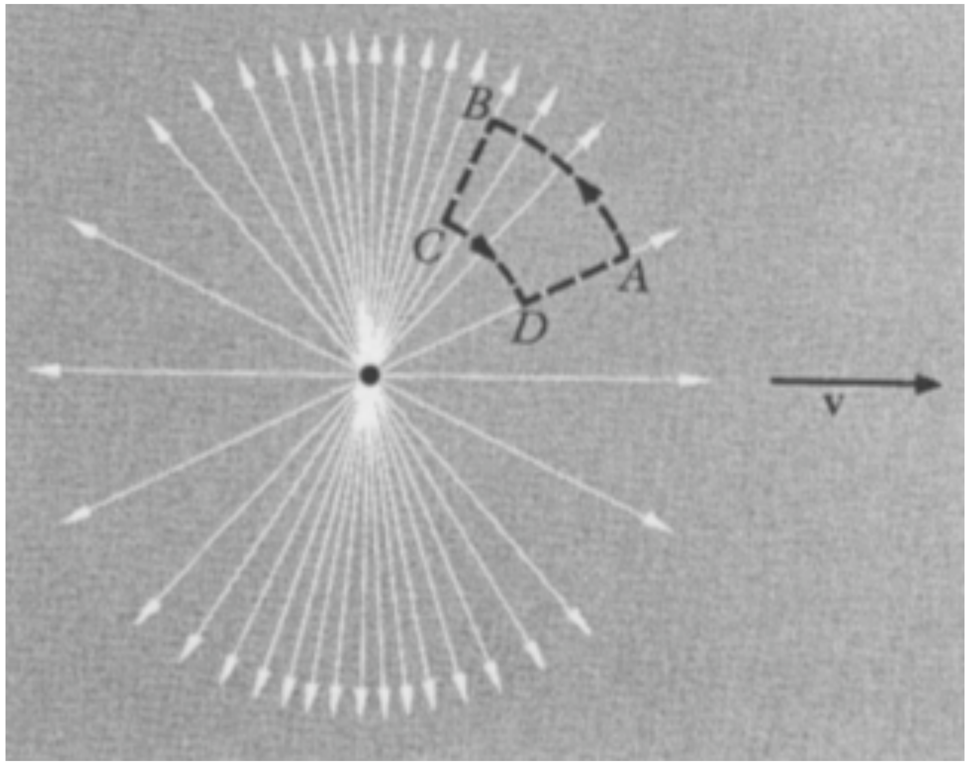
This is not, however, instantaneous transmission of information. The particle has been moving with constant velocity forever such that it was programmed to pass the origin at 12 Noon. Information has been available for long time. PAST HISTORY of particle determined observed field(causality is preserved).

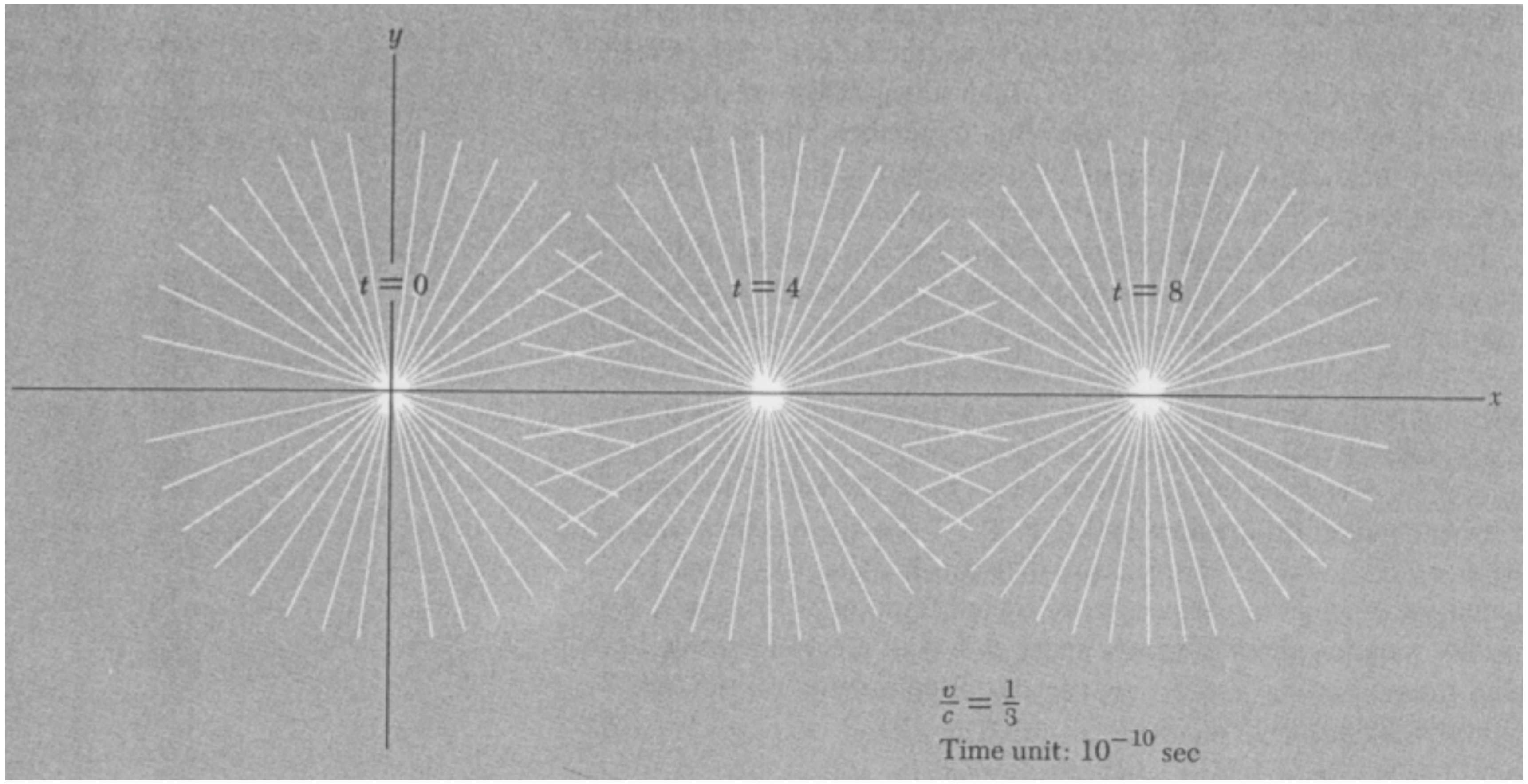
As will see, strange stuff happens when the programmed motion is changed.

If v is large, the field lines bunch up as shown:

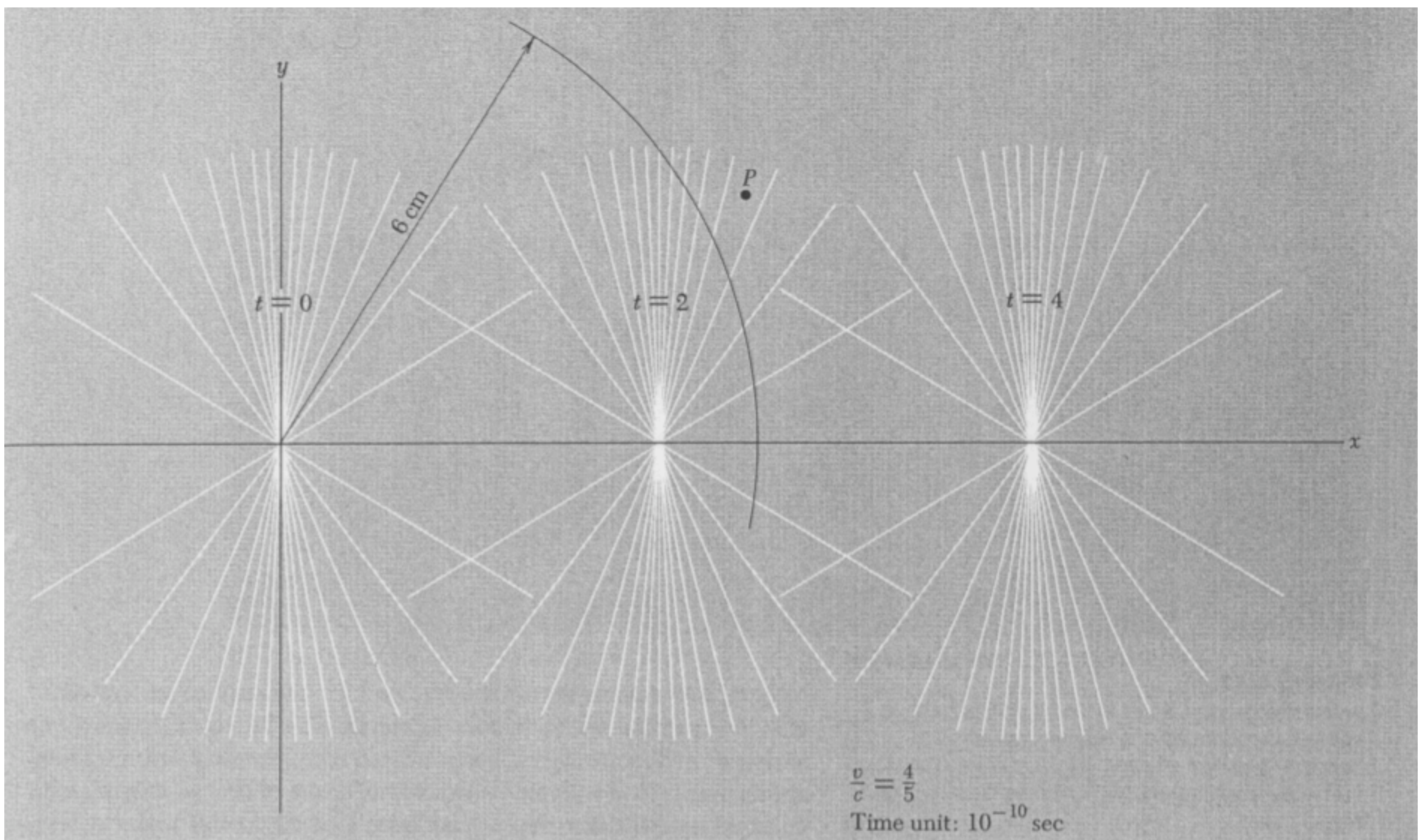
Not spherically symmetric! Exists preferred direction. Cannot be produced by any stationary charge distribution! Line integral of E around closed path is not equal to zero (look at ABCD path). E-field on BC is stronger than on AD.

Figure below: E-field v/c=1/3; not much bunching.
2nd figure - E-field v/c=0.8; strong bunching.





y
x
t = 0
t = 4
t = 8
v/c = 1/3
Time unit: 10^-10 sec

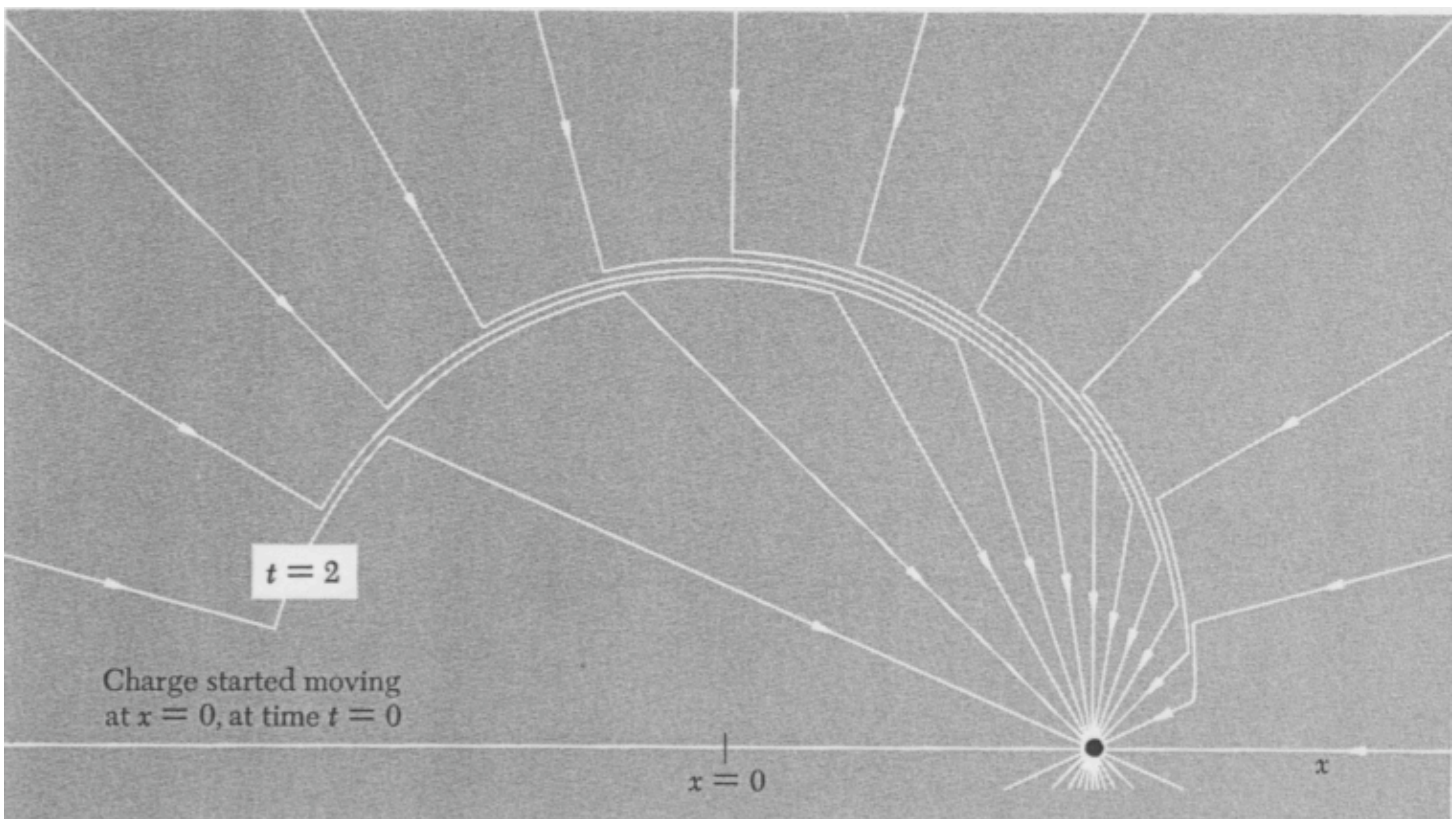


y
x
6 cm
P
t = 0
t = 2
t = 4
v/c = 4/5
Time unit: 10^-10 sec

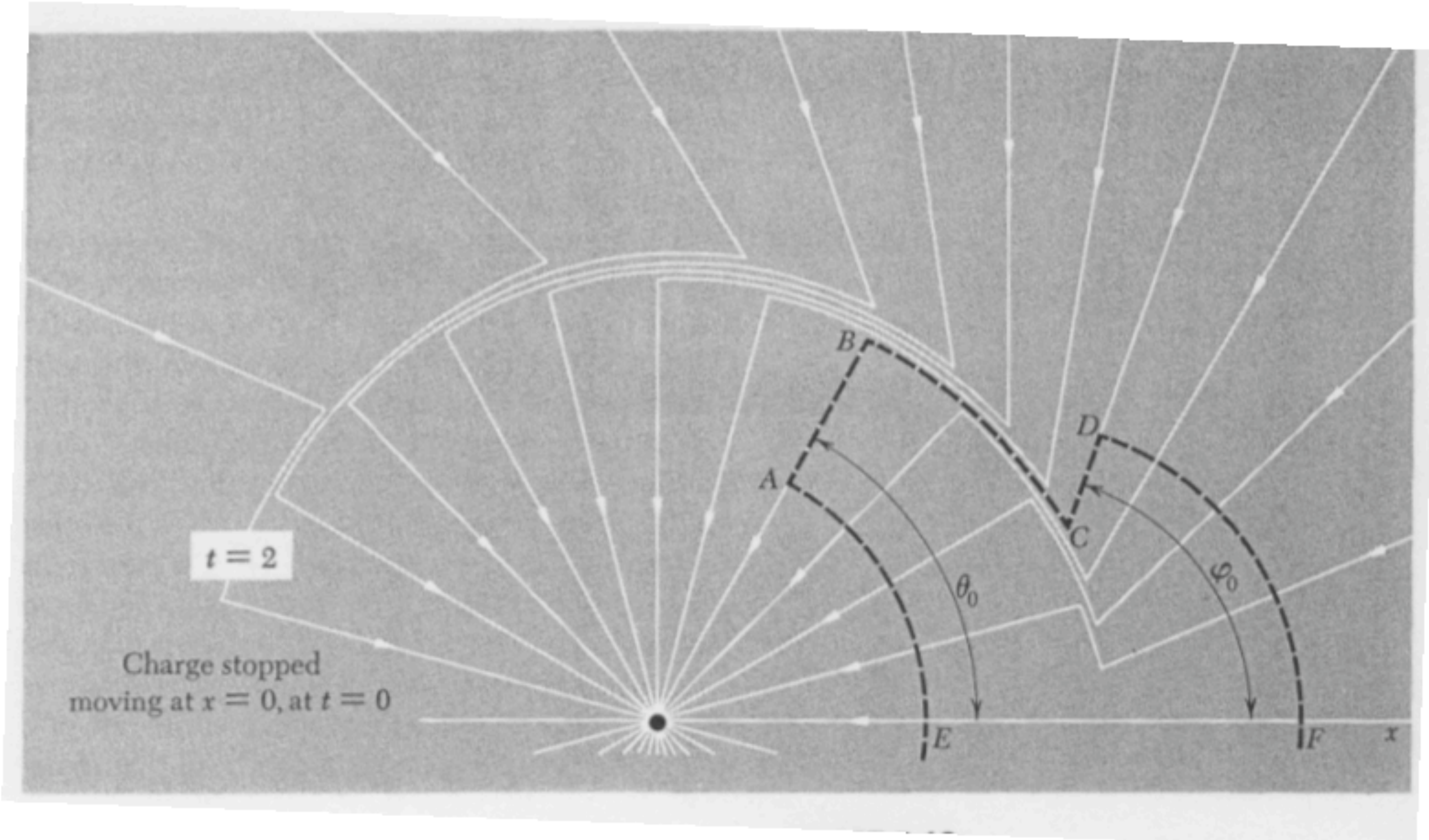
Field of Charge That Starts

Suppose charge has been at rest at origin waiting for clock to read t=0. At t =0, large acceleration causes charge to suddenly move with velocity v and then charge continues with v from then on. Last figure is not correct representation of field. We must take into account finite velocity of light or finite velocity of transmission of information. After a time t, only point within a sphere of radius r = ct can "know" charge has started to move! Outside of that sphere, observer must still think charge is at the origin!

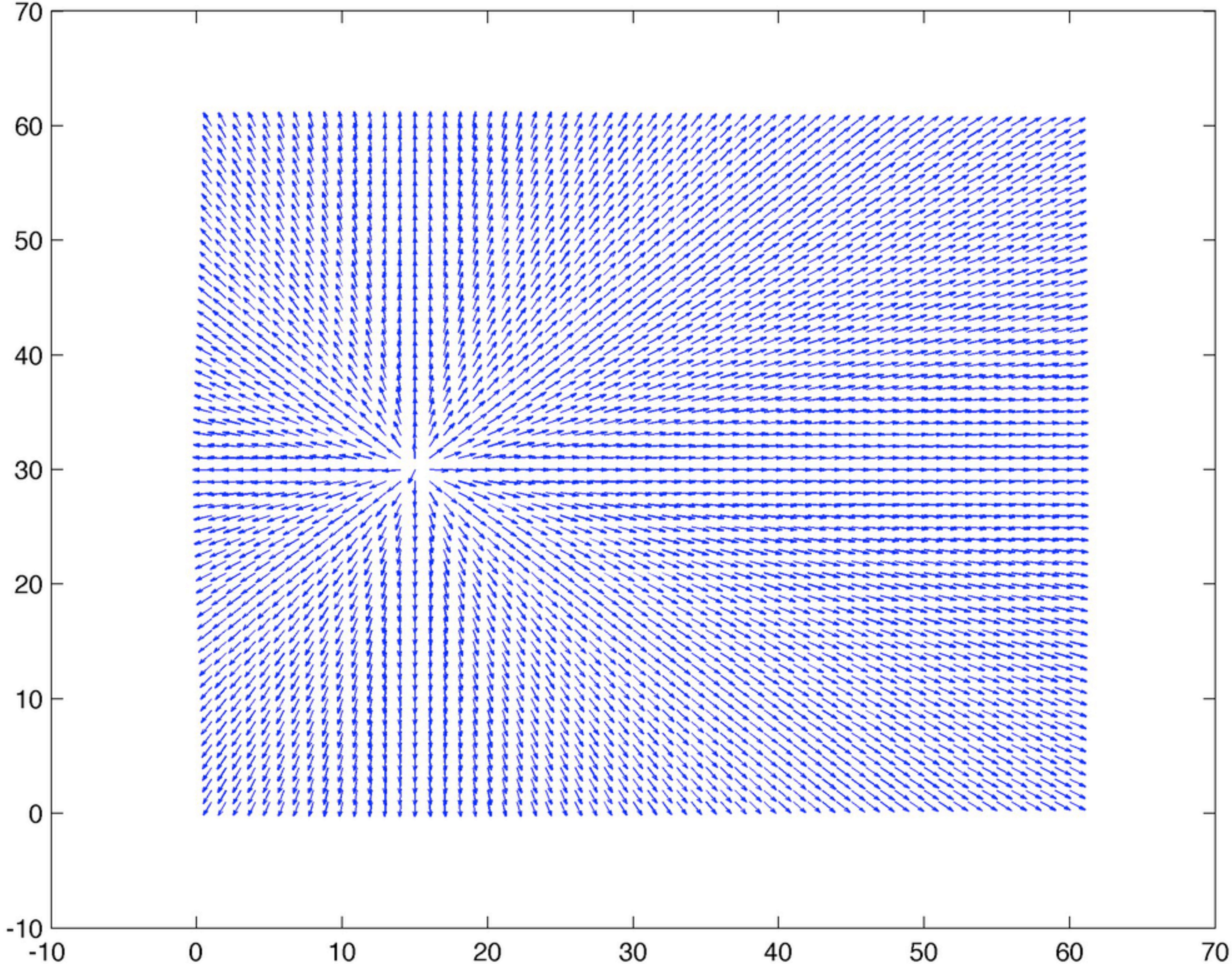
In side the sphere, PAST HISTORY (before t=0) does not matter; the field must reflect the fact that the charge started moving at t=0.



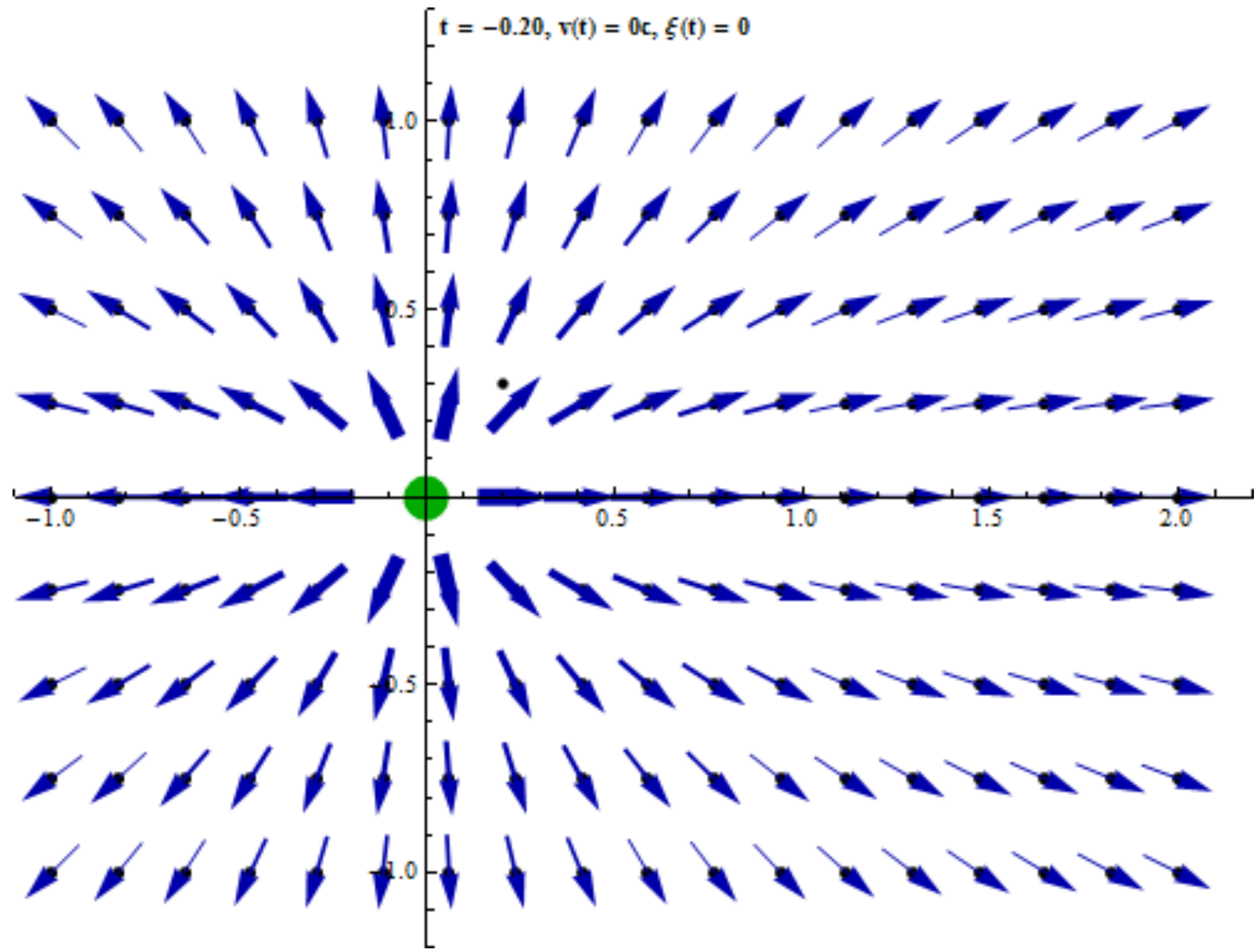
Field of Charge That Starts and then Stops



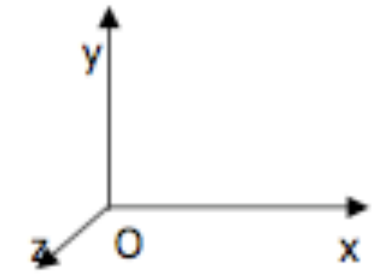
Details of how to connect field lines inside sphere and outside spher discussed in Purcell.



http://www.its.caltech.edu/~phys1/java/phys1/MovingCharge/MovingCharge.html


t = −0.20, v(t) = 0c, ξ(t) = 0
1.0
0.5
−0.5
−1.0
−1.0
−0.5
0.5
1.0
1.5
2.0

Please pay attention: this is difficult!



Force by current on moving charge

Description in the Lab Frame

Electrically neutral wire carrying a current

Positive charge density $\lambda_+ = \lambda_+^{REST} = \lambda_0$.

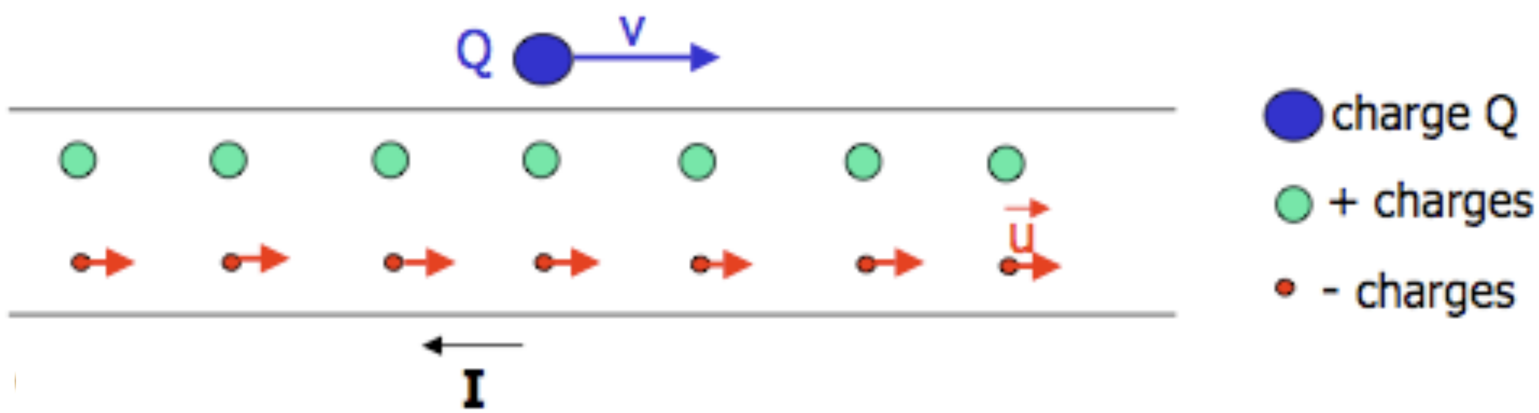
Note: these charges are at rest in the Lab

Negative charge density $\lambda_- = \lambda_-^{MOT} = -\lambda_0$ (remember wire is neutral)

Note: these charges are moving with velocity u

→ $-\lambda_0$ is **not** the density of the electrons in their reference frame O

A charge Q outside the wire moves to right with velocity v



λ of negative charges in their Reference Frame(RF)

We said: $\lambda_- = \lambda_-^{MOT} = -\lambda_0$ (in Lab Frame)

What is charge density in their RF? λ_-^{REST}

First attempt:

Charge density $\lambda_0 = Q/L$ where L = length of the wire in Lab frame

In lab frame: $\lambda_-^{MOT} = Q/L = -\lambda_0$

In O (at rest with -charges), length of wire appears contracted: $L' = L/\gamma$

→λ_-^{REST} = Q/L′= Qγ/L = -γλ_0 → λ_-^{REST} > λ_-^{MOT} **WRONG!**

Why? There is no such thing as the wire. Just the length of + and – charges which happen to be the same in the Lab reference frame but not elsewhere.

Second attempt:
The electrons will "think": our length in our own RF is L'_- . In the reference frame of the lab, boosted wrt us by a velocity –u, this length will be contracted by a factor γ: L'_- = γL

→λ_-^{REST} = Q/L′ = Q/γL = -λ_0/γ → λ_-^{REST} < λ_-^{MOT}

Force by current on moving charge

What forces act on the charge Q? Lab frame:
- Wire is neutral: no electric field E
- Current will generate magnetic field B:
  - Current in the wire: I = dq/dt = $\lambda_0$dx/dt=λ_0u

Ampere's law: $B = \frac{2I}{cr}$ right-hand rule - wraps around wire!

Magnetic force acting on charge Q: $\vec{F} = q\left(\vec{E} + \frac{\vec{v}}{c} \times \vec{B}\right)$

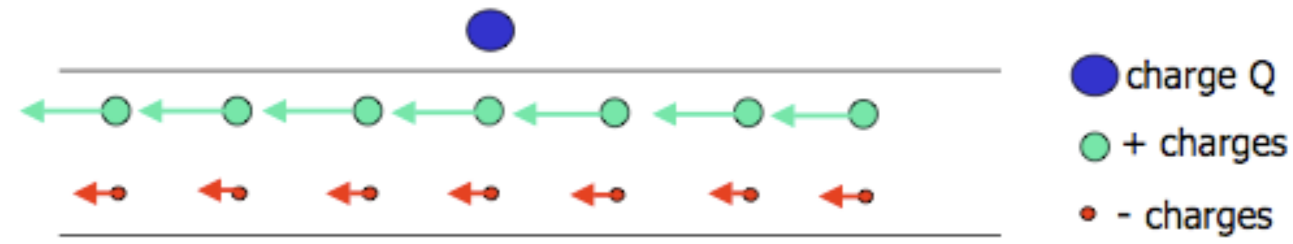
$F = Q\frac{v}{c}B = Q\frac{2\lambda_0 uv}{c^2 r}$ (magnitude)

Direction?
- Right hand rule: → repulsive force $\vec{v} \times \vec{B}$
- Note: I opposite to v electrons!

v → X → B

Force in charge’s rest frame?

Let’s now move to the rest frame of Q:



Velocities involved:

Charge Q: at rest by definition

Negative charges in the wire: velocity u’ = (u-v)relativistic sum $$u' = \frac{u-v}{1-uv/c^2}$$

Positive charges in the wire: velocity –v

Is there any force acting on Q?

There must be: Relativity Principle! Consistent descriptions in both frames.

Force is either there or not! No magnetic force however: the charge is at rest!

Charge densities in Q’s RF

Are we in trouble?

Let’s see what happens to the charges in the wire

Positive charges: moving in this RF with -v

Charge density in charge’s reference frame: $\lambda_+' = Q/L' = \gamma_v\lambda_0$

Negative charges:

$$\lambda'_- = \gamma_{u'}\lambda_-^{REST} = \gamma_{u'}\frac{-\lambda_0}{\gamma_u} = \gamma_u\gamma_v(1-\beta_u\beta_v)\frac{(-\lambda_0)}{\gamma_u} = -\gamma_v(1-\beta_u\beta_v)\lambda$$ as shown below

Goal: calculate $\gamma_{u'}$ Let's start calculating

$$\frac{1}{\gamma_{u'}^2} = 1-\beta_{u'}^2 = 1-\left(\frac{u'}{c}\right)^2 = 1-\left(\frac{\frac{u-v}{1-uv/c^2}}{c}\right)^2 = 1-\left(\frac{\beta_u-\beta_v}{1-\beta_u\beta_v}\right)^2 = \frac{1-2\beta_u\beta_v+\beta_u^2\beta_v^2-(\beta_u-\beta_v)^2}{(1-\beta_u\beta_v)^2}$$

$$= \frac{(1-\beta_v^2)(1-\beta_u^2)}{(1-\beta_u\beta_v)^2} = \frac{1}{\gamma_u^2\gamma_v^2(1-\beta_u\beta_v)^2} \Rightarrow \gamma_{u'} = \gamma_u\gamma_v(1-\beta_u\beta_v)$$

Force in charge's rest frame

Net charge density of wire in Q's reference frame(add + and - contributions):

$$\lambda'_{NET} = \lambda'_{+} + \lambda'_{-} = \gamma_v \lambda_0 - \gamma_v (1 - \beta_u \beta_v)\lambda_0 = \gamma_v \beta_u \beta_v \lambda_0 = \gamma_v \frac{uv}{c^2}\lambda_0$$

In this RF there is a net + charge on the wire! ➙ generates Electric field!

$$E' = \frac{2\lambda'_{NET}}{r} = \gamma_v \frac{2uv}{rc^2}\lambda_0$$

Electric field ➙ force F' will act on the charge Q

$$F' = QE' = \gamma_v \frac{2Quv}{rc^2}\lambda_0 \quad (repulsive)$$

Is there a Magnetic field as well?

Yes, but it does not exert any force on Q because Q is at rest

Comparison of forces in the 2 RFs

In lab frame: we see a repulsive magnetic force acting on charge Q:

$$F_{Lab} = Q\frac{v}{c}B = Q\frac{2\lambda_0 uv}{rc^2}$$

In Q's rest frame(charge frame) wire is not neutral ➙ Repulsive electric force acting on charge Q:

$$F_{Q_{RF}} = QE' = \gamma_v Q\frac{2\lambda_0 uv}{rc^2}$$

no magnetic force on Q because Q is at rest

Are results consistent? Remember the rules for transforming forces!

Yes! We have seen that forces in direction perpendicular to v transform as $F_{Lab} = \frac{F_{Q_{RF}}}{\gamma_v}$ and this is exactly what we find!

Thoughts on this problem

Is the comparison fair?

You might object on the following grounds:

In one RF we have a magnetic force, in the other an electric force

Are we comparing apples and oranges?

No, on the contrary!

This results proves that Electricity and Magnetism are intimately connected!

Electric and Magnetic forces are not separate phenomena!!!!

The electric force acting on a charge in that charge's rest frame is exactly what is needed to explain the magnetic force in a frame in which that charge is moving!

Physics is consistent!

Principle of relativity demands that the 2 observers will come to the same conclusions.

The details of the calculation (Electric? Magnetic?) are different in the different RF, but ultimately irrelevant.

This is the essence of SR! It also tells us that electric and magnetic fields are really the same thing. "Electricity" and "magnetism" are not separate phenomena: they are different specific manifestations of a single thing: "electromagnetism"

The Origins of Magnetism

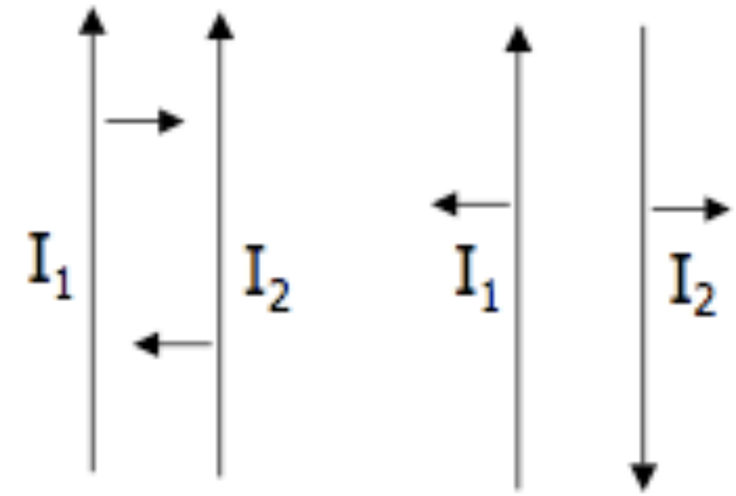
- Ancient Greeks noticed that pieces of mineral magnetite (an oxide of iron) had very special properties:
  - Could attract a piece of iron, but no effect on Au, Ag, Cu, etc
  - Can attract or repel piece of magnetite depending on relative orientation
- By the 12th century people could build a magnetic compass
  - A small magnetic needle is suspended so it can pivot around vertical axis
  - The needle will always come to rest with one end pointing North
  - By definition we call that end "North" and the other "South"

- Experiments proved that:
  - Like poles repel, unlike poles attract
  - North and South cannot be separated in a magnet
  - Magnetic forces can be very strong!

The big step forward

- In 1820 Oersted realized that current flowing in wire made the needle of compass swing
  - The direction depends on the direction of the current
- BIG discovery: proves that Electricity and Magnetism are related!
- Soon after, Ampere's experiment with parallel wires carrying current

If currents are parallel, wires attract
If anti-parallel, wires repel
No force on a stationary charge nearby …
Note: wires are overall neutral!



Magnetic force

From more refined observations:

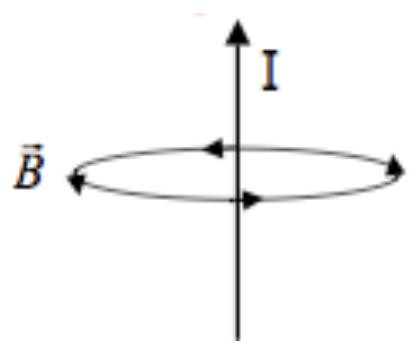
Magnitude: $\left|\vec{F}\right| \propto I_1 I_2 \Rightarrow F \propto v$ of charges in motion

Direction: $\vec{F} \perp \vec{v}$

Interpretation

Some field (magnetic field B) is created by the charges in motion

Magnetic force is proportional to cross product $\vec{v} \times \vec{B}$



$$\vec{F}_{magnetic} = q\,\frac{\vec{v}}{c} \times \vec{B}$$

Direction of B: B curls around current (right hand rule)
Iron fillings can be used to visualize B field lines

Note: this is an empirical law so far

Magnetic Field B of a magnet bar

What kind of B does a magnet bar produce?

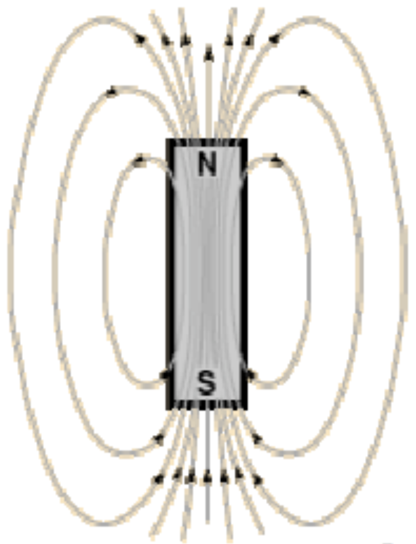
Strength and direction

How to visualize it?

Direction: map B field moving compass around the bar

Direction and strength:

sprinkle iron fillings on a magnet bar



Lorentz force

When a charged particle moves in electric (E) and magnetic (B) fields it feels a force ($F_{Lorentz}$):

$$\vec{F}_{Lorentz} = q\left(\vec{E} + \frac{\vec{v}}{c} \times \vec{B}\right)$$

This formula defines the magnetic field B

Units of B in cgs:

[B] = [F]/[q] = dyne/esu = Gauss (G)

Note: [B] = [E]

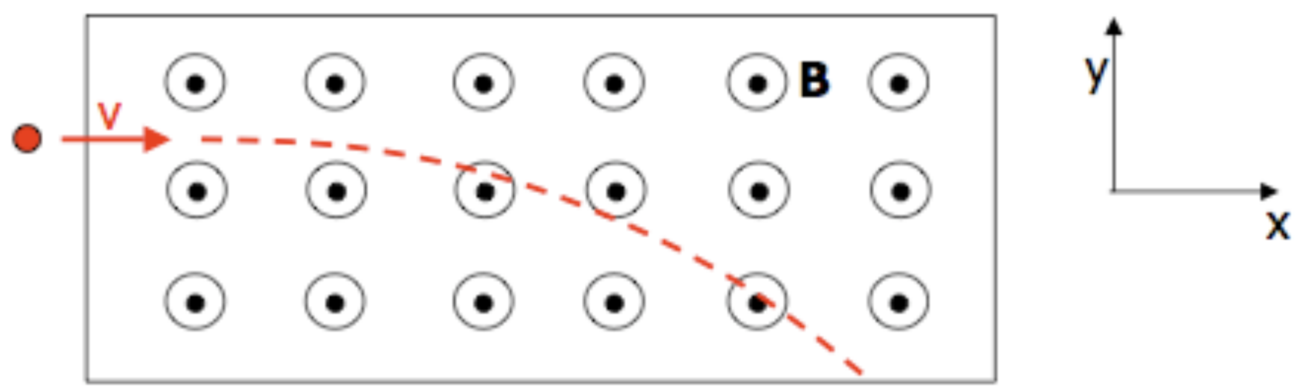
Units of B in SI:

[B] = [F]/[q v] = N s /(m C) = Tesla (T)

Conversion: 1 T = 10^4 G

Trajectory in magnetic fields

A particle of charge q and mass m moves with velocity v || +x-axis in a magnetic field B || +z-axis (out of the page):



What is the trajectory of q in the magnetic field?

$$\vec{F} = q\frac{\vec{v}}{c} \times \vec{B}$$

v , B and F (or a) are always perpendicular→circular motion!

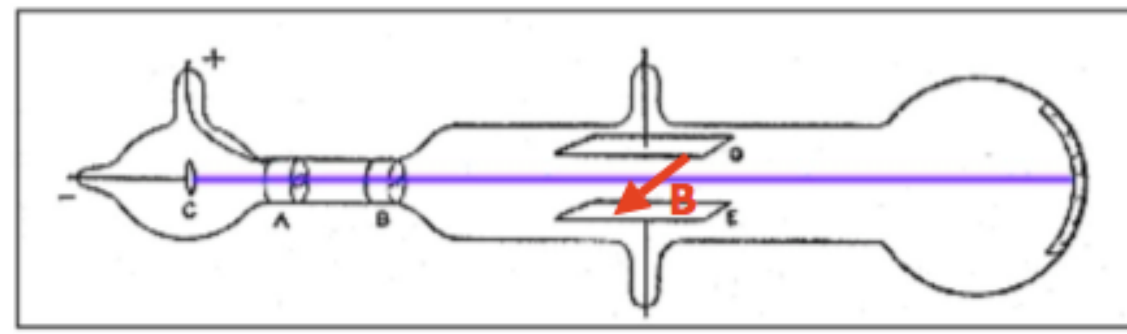
$$F_{Lorentz} = \frac{qvB}{c} = ma_{centripetal} = m\frac{v^2}{R} \Rightarrow R = \frac{mvc}{qB}$$

Deflection of electron beam by B

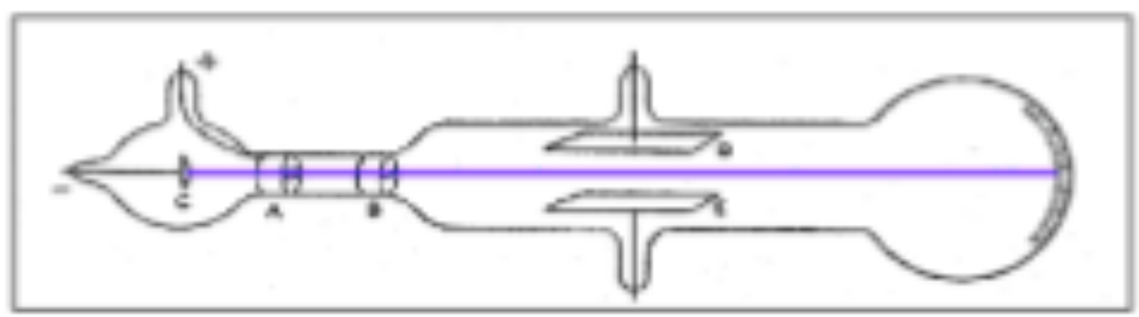
An electron beam is produced by a cathode in a vacuum tube
- Velocity of electrons: v_e

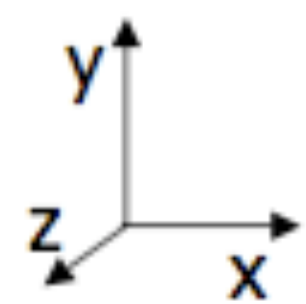
Magnetic field B perpendicular to v_e is produced by current in a wire or by permanent magnet

What do we expect to happen?
- Electrons curve according to Lorentz force



J.J. Thompson's experiment





Discovery of electrons and measurement of e/m_e in 1897

The idea:
- A beam of "cathode rays" crosses a region with E and B present

Choosing v_e || x-axis, B || z-axis, E || y-axis→$F_{Lorentz}$ || $F_{Electric}$ || y

E and B can be adjusted so $F_{magnetic} = -F_{electric}$ that e will go straight

Electric field alone causes a shift: $F = qE = ma \Rightarrow \Delta y = \frac{1}{2}at^2 = \frac{1}{2}\frac{qE}{m}t^2$

$$t = \frac{L}{v} \Rightarrow \Delta y = \frac{1}{2}\frac{qE}{m}\left(\frac{L}{v}\right)^2 = \frac{qEL^2}{2mv^2}$$

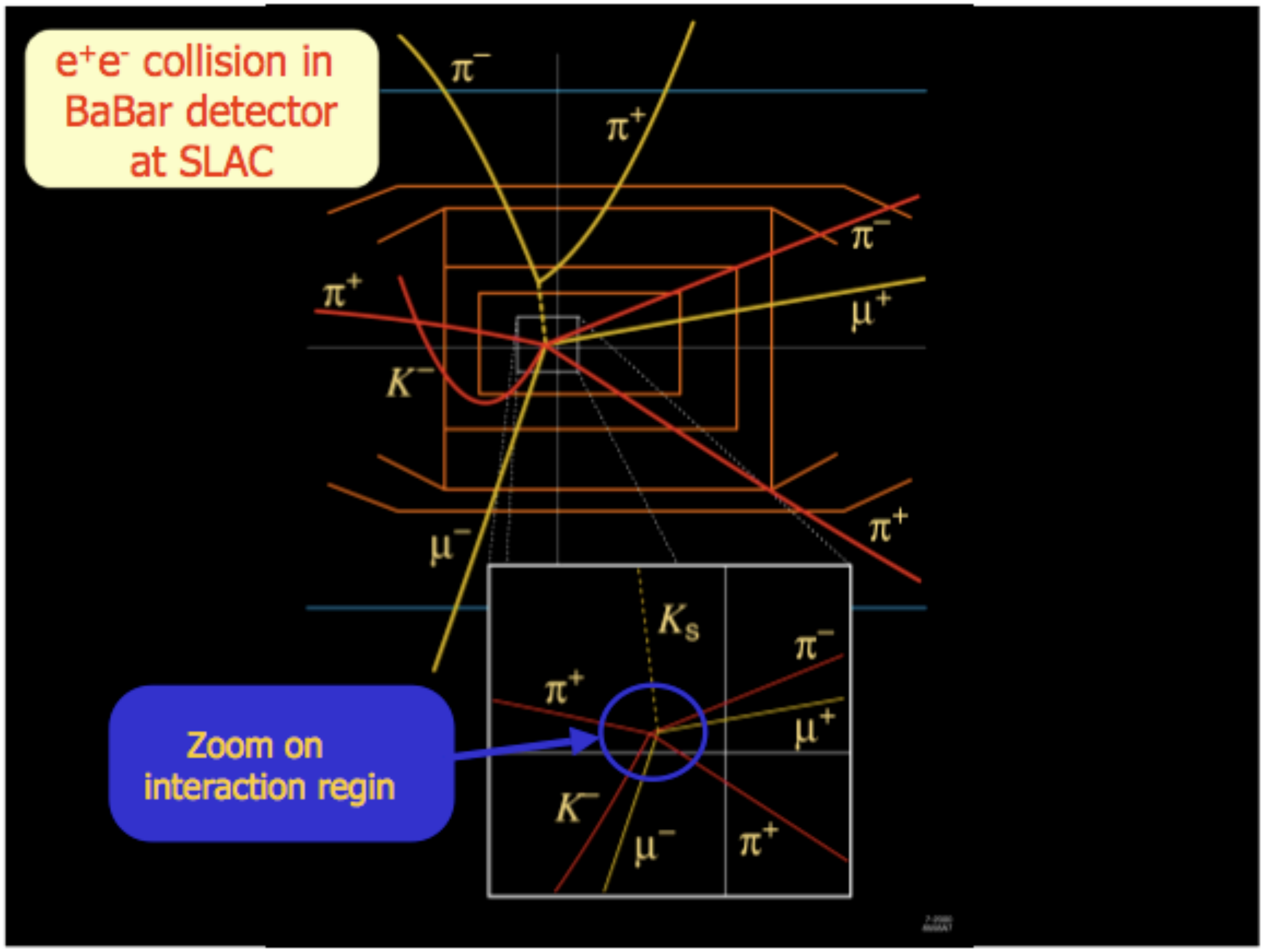
Now turn on B and set it to cancel the shift due to E: $F_B = F_E \Rightarrow \frac{qvB}{c} = qE \Rightarrow v = c\frac{E}{B}$

Substituting this in the previous equation gives:

$$\frac{e}{m_e} = \frac{q}{m} = \frac{2v^2\Delta y}{EL^2} = \frac{2c^2E\Delta y}{B^2L^2}$$

Application in modern physics

- Tracking detectors in modern particle physics
- The problem
  - High energy collisions between elementary particles (such as e+ e-) produce many particles (protons, electrons, pions, muons,…)
  - How can we “see” these particles?
    - Build detectors that can “visualize” the trajectory of charged particles using the fact that particles ionize the material they cross
  - How can I measure the properties of these particles?
    - E.g.: measure momentum, energy, mass, etc.
    - Immerse the detector in a very strong magnetic field B ~ 2 T
    - Charged particles will curve according to $R = \frac{mvc}{qB}$
      - Direction measures the charge
    - Radius of curvature measures momentum p = mv



Magnetic force and work

Moving a charge in an electric field E requires work: $W_{12} = -q\int_1^2 \vec{E}\cdot d\vec{s}$

How much work does it take to move a charge in a magnetic field?

$$dW = \vec{F}\cdot d\vec{s} = \vec{F}\cdot\vec{v}dt = \frac{q}{c}\left(\vec{v}\times\vec{B}\right)\cdot\vec{v}dt = 0$$

→No work is needed to move a particle in a magnetic field because v and F are always perpendicular!

Force on a current

A magnetic field will exert a force on a current

Since a current is just a stream of moving charges!

Current I flowing in a wire can be seen a density of charges λ moving with velocity v:

$$I=\lambda v$$

The force dF exerted on the infinitesimal wire dl is:

$$d\vec{F} = (\lambda dl)\frac{\vec{v}}{c} \times \vec{B}$$

Rewrite this in terms of the current:

$$d\vec{F} = \left(\frac{\lambda v}{c}\right) d\vec{l} \times \vec{B} = \left(\frac{I}{c}\right) d\vec{l} \times \vec{B}$$

Total force F:

$$\vec{F} = \left(\frac{I}{c}\right) \int_{wire} d\vec{l} \times \vec{B}$$

For a long straight wire in a constant magnetic field:

$$\vec{F} = \left(\frac{I}{c}\right) L\hat{n} \times \vec{B} \qquad \hat{n} = direction\ of\ current$$

Ampere's law

In electrostatics, the electric field E and its sources (charges) are related by Gauss's law:

$$\int_{surface} \vec{E} \cdot dA = 4\pi Q_{enclosed}$$

Why useful? When symmetry applies, E can be easily computed

Similarly, in magnetism the magnetic field B and its sources (currents) are related by Ampere's law:

$$\oint_C \vec{B} \cdot d\vec{s} = \frac{4\pi}{c} I_{enclosed}$$

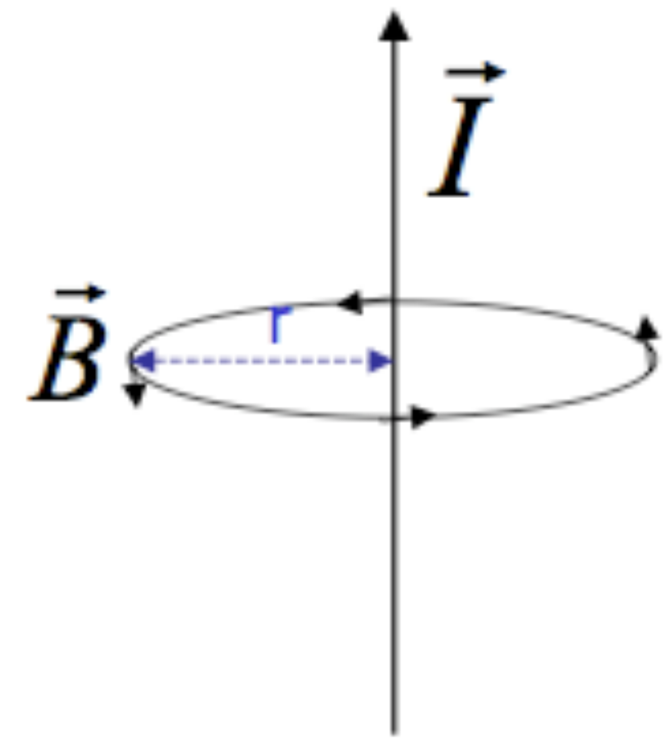
Why useful? When symmetry applies, B can be easily computed Note: This is a line integral!

Use wire result from earlier to derive Ampere's law. B created by current in a wire

Long, straight wire in which flows a current I.
Magnetic field B created by I

$$\vec{B} = \frac{2I}{cr}\hat{\varphi}$$

$$\int_{path} \vec{B} \cdot d\vec{s} = \int_{path} \frac{2I}{cr}\hat{\varphi} \cdot r d\varphi \hat{\varphi} = \frac{2I}{c}\int_0^{2\pi} d\varphi = \frac{2I}{c} 2\pi = \frac{4\pi I}{c} = \frac{4\pi}{c} I_{enclosed}$$



Application of Ampere's law: B created by current in a wire

$$\oint_C \vec{B} \cdot d\vec{s} = B(r) 2\pi r = \frac{4\pi}{c} I_{enclosed} \Rightarrow \vec{B} = \frac{2I}{cr}\hat{\varphi}$$

Direction: using right hand rule

Note: $B_{wire} \sim 1/r$. Does this look familiar?

Remember E created by a line of charge:

Coincidence? Not at all …

$$E(r) = \frac{2\lambda}{r}$$

Force between 2 wires

Force on wire 1 due to magnetic field B created by wire 2:

$$\vec{F}_1 = \frac{I_1}{c} L\hat{n} \times \vec{B}_2$$

Magnetic field created by wire 2: $\vec{B}_2 = \frac{2I_2}{cr}\hat{\varphi}_2$

Total force F: $F = \frac{2I_1 I_2}{c^2 r} L$

Usually we quote the force/unit length:

$$\frac{F}{L} = \frac{2I_1 I_2}{c^2 r}$$

Direction? $\vec{F} \propto \vec{I}_1 \times \hat{\varphi}_2$

Using right hand rule:

I$_1$ and I$_2$ parallel: attractive
I$_1$ and I$_2$ anti-parallel: repulsive

$\Uparrow \quad \bullet \quad \Uparrow$; $I_1 \quad I_2$; $\hat{\varphi}_2$

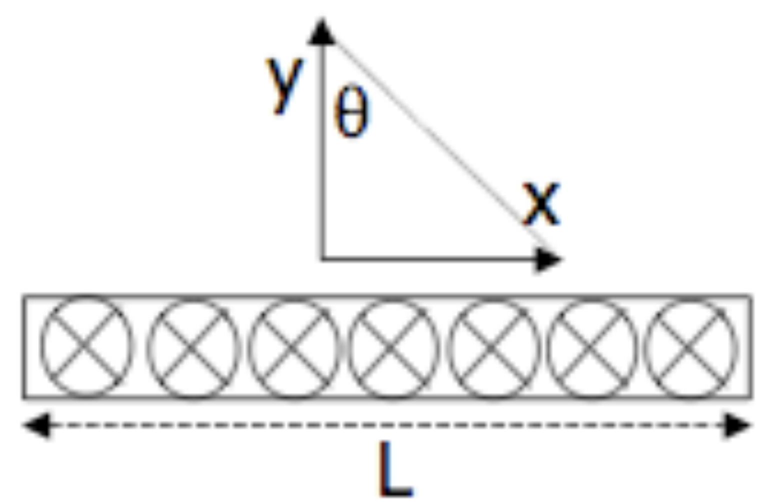
$\vec{F}_1 \propto \vec{I}_1 \times \hat{\varphi}_2 \Rightarrow attractive$

$\Uparrow \quad \times \quad \Downarrow$; $I_1 \quad I_2$; $\hat{\varphi}_2$

$\vec{F}_1 \propto \vec{I}_1 \times \hat{\varphi}_2 \Leftarrow repulsive$

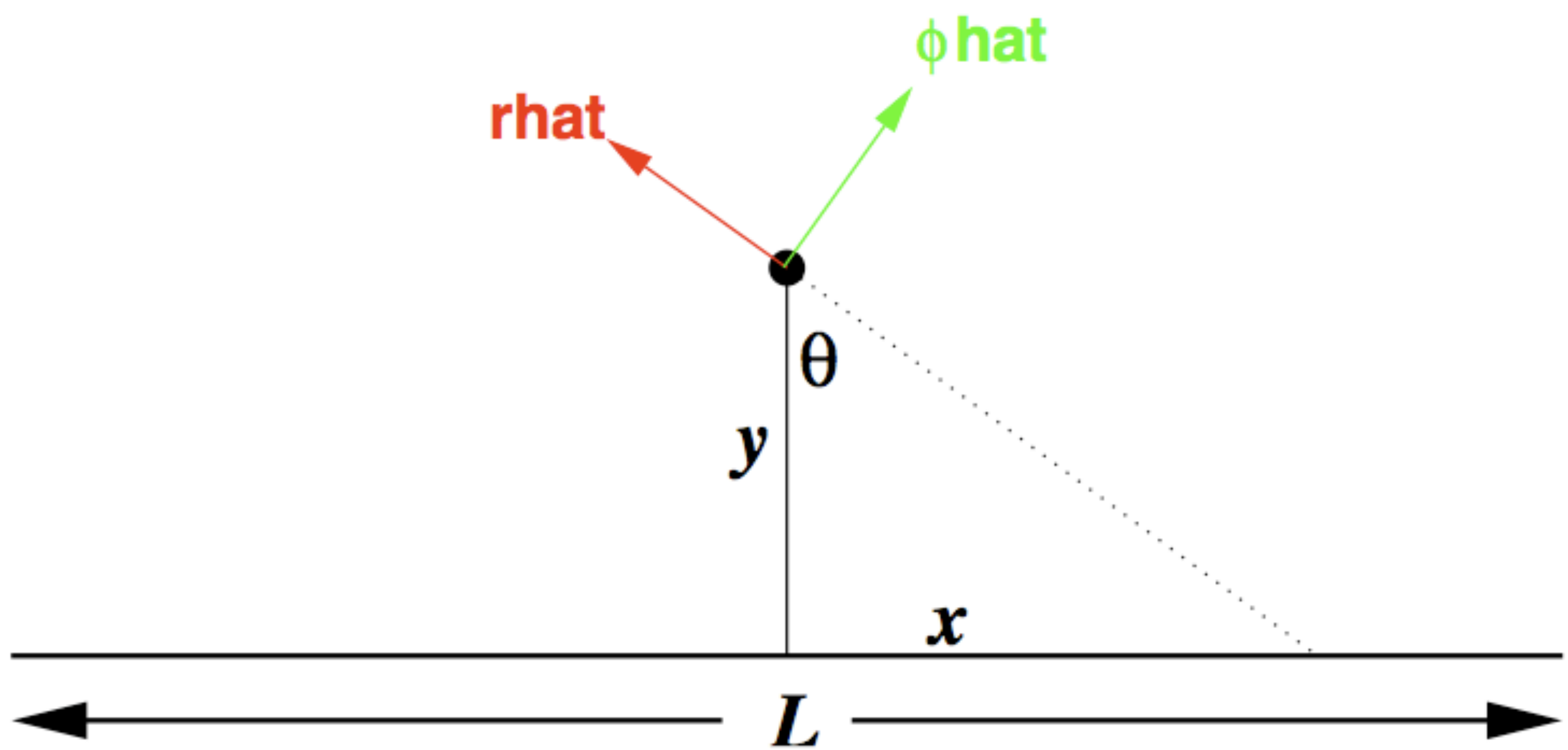
Another application of Ampere's law: **B created by sheet of current**

Calculate the magnetic field B created by current flowing in a sheet of conductor
Current || -z axis (into the page)
Width of sheet of conductor: L
Current in a metal sheet ~ N parallel wires



Solution:
B from a wire is known: $\vec{B} = \frac{2I}{cr}\hat{\varphi}$

Here is the sheet. In this figure, I've squished all the wires down into a nice, thin sheet. Consider the contribution to the magnetic field arising from a "chunk" of current to the right of the midpoint:



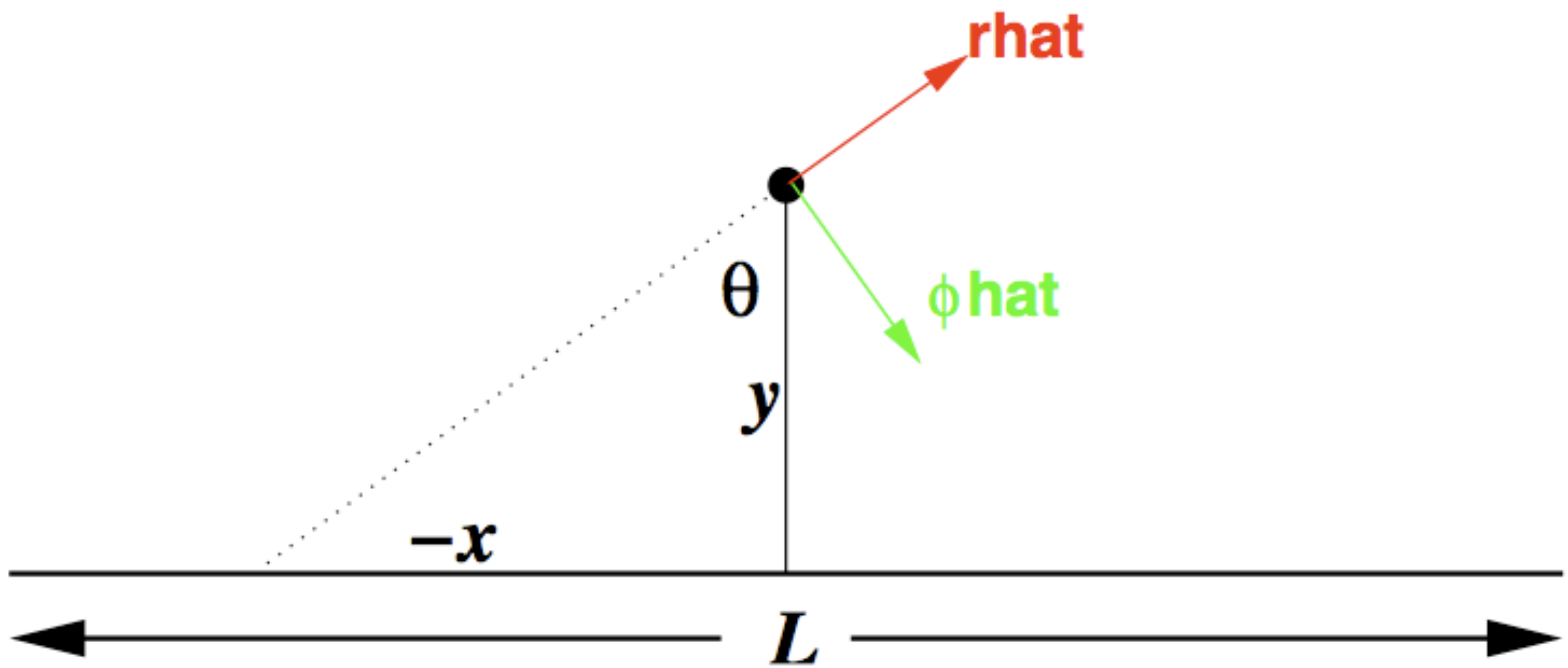
The radial direction from that chunk of current points along the red $\hat{r}$ vector that I've drawn. The magnetic field points along $\hat{\phi}$, which is rotated 90° clockwise (right hand rule) from $\hat{r}$. It is drawn in green. The magnetic field arising from this chunk of current is

$$d\vec{B}_{\text{right}} = \frac{2K}{c}\frac{dx}{r}\hat{\phi}_{\text{right}} \,,$$

K = current/length

where $r = \sqrt{x^2 + y^2}$. (Don't forget that the current in this drawing is flowing into the page.)

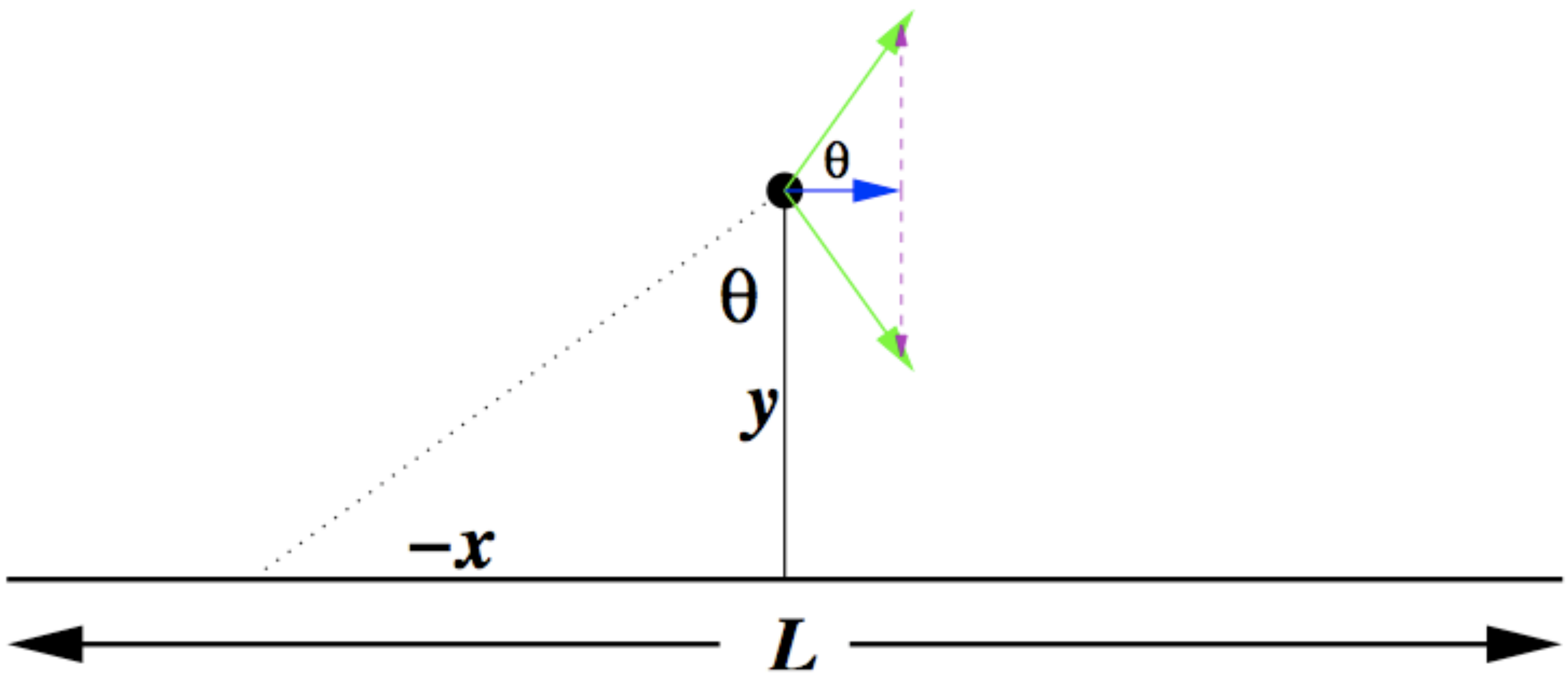
Now consider the field from a chunk to the left of the midpoint:



I've again drawn the radial direction from the chunk of current $\hat{r}$ in red; the magnetic field points along $\hat{\phi}$, rotated 90° clockwise from $\hat{r}$. The magnetic field arising from this chunk of current is

$$d\vec{B}_{\rm left} = \frac{2K}{c}\frac{dx}{r}\hat{\phi}_{\rm left} \; .$$

Now superpose the two contributions: The vertical components (drawn purplish) of the



fields clearly cancel each other out. The horizontal components by contrast reinforce one another. In other words, we can ignore the vertical component of the field and just take the horizontal one.

How do we get the horizontal component? It will clearly be something like $dB_x = |d\vec{B}|[\text{some trig function}](\theta)$; the question is which trig function to use. To get the correct trig function, we need to study the triangles in the figure very carefully — the "little" triangle that is built out of the components of $d\vec{B}$ is congruent with the big one that is built out of x and y. Careful drawing[1] shows that θ belongs where I've drawn it in the figure above.

This shows us that the trig function we need is cosine; this in turn tells us that to get the magnetic field we want

$$dB_x = \frac{2K}{c}\frac{dx}{r}\cos\theta\ .$$

Substituting for r and $\cos\theta$ and setting up the integral, we have

$$\begin{aligned}\vec{B} &= \frac{2K}{c}\hat{x}\int_{-L/2}^{L/2}\frac{dx}{\sqrt{x^2+y^2}}\frac{y}{\sqrt{x^2+y^2}} \\ &= \frac{2Ky}{c}\hat{x}\int_{-L/2}^{L/2}\frac{dx}{x^2+y^2}\ .\end{aligned}$$

The rest of the derivation follows from here.

Summary: The key confusing bit is that the horizontal component of the field was given by the cosine of θ (with θ defined as in the drawing). Our usual intuition told us that it should be the sine of θ. However, the intuition we've developed so far is based on the idea that the field direction points along the hypotenuse of the triangle (i.e., along $\hat{r}$). The magnetic field actually points at right angles to the hypotenuse. This extra 90° changes the sine which we might have expected into the cosine that we actually get.

Summary of the summary: magnetic fields can be a pain in the butt.

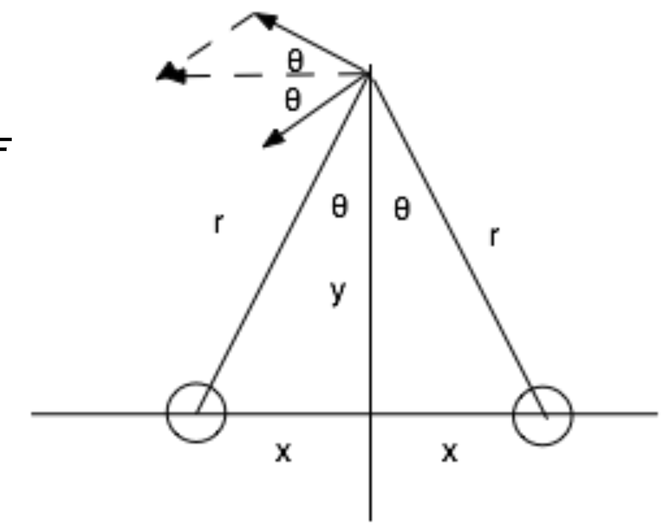
So re-iterating, we just apply superposition…

Direction: symmetry says that for y > 0: B || +x; for y < 0: B || -x

vertical components cancel in pairs

Magnitude: integrate dB = B field from each infinitesimal wire

$$d\vec{B} = \frac{2}{c}\left(\frac{I}{L}dx\right)\frac{1}{\sqrt{x^2+y^2}}\cos\theta\hat{x} \Rightarrow \vec{B} = \frac{2}{c}\hat{x}\frac{I}{L}\int_{-L/2}^{L/2}\frac{dx}{\sqrt{x^2+y^2}}\frac{y}{\sqrt{x^2+y^2}}$$



$$\vec{B} = \frac{2}{c}\hat{x}\frac{I}{L}y\int_{-L/2}^{L/2}\frac{dx}{x^2+y^2}$$ parallel to sheet!

Messy integration; take limit L→∞

$$\int_{-\infty}^{\infty}\frac{dx}{x^2+y^2} = \frac{\pi}{|y|} \Rightarrow \vec{B} = \frac{2\pi K y}{c|y|}\hat{x} \quad , \quad K = \text{current per unit length}$$

$$\vec{B} = \begin{cases} +\dfrac{2\pi K}{c}\hat{x} & y>0 \quad \text{(above the plane)} \\ -\dfrac{2\pi K}{c}\hat{x} & y<0 \quad \text{(below the plane)} \end{cases}$$

Note: magnitude of B does not depend on y. As for E of sheet of charges.

What is the change of B (parallel to sheet) across the sheet of current? $\Delta B = \dfrac{4\pi K}{c}$

Does it ring a bell? Yes, ΔE across a plane of charge! $\Delta E = 4\pi\sigma$

Another similarity between electric/magnetic fields .. must be more than a pure coincidence

Ampere's law in SI units

In SI Ampere's law takes the form: $\oint_C \vec{B}\cdot d\vec{s} = \mu_0 I_{enclosed}$

where $\mu_0 = 4 \times 10^{-7}$ N/A^2 is the magnetic permeability of free space

Be careful not to mix cgs and SI formulae!

To convert cgs ➙ SI: multiply by $\mu_0 c/(4\pi)$

Examples:

Magnetic field created by a wire: $\left(\vec{B} = \frac{2I}{cr}\hat{\varphi}\right)_{cgs} \Rightarrow \left(\vec{B} = \frac{\mu_0 I}{2\pi r}\hat{\varphi}\right)_{SI}$

Force between 2 wires:

Note: factor 1/c missing in $F_{Lorentz}$ in SI

$$\left(\frac{F}{L} = \frac{2I_1 I_2}{c^2 r}\right)_{cgs} \Rightarrow \left(\frac{F}{L} = \frac{\mu_0 I_1 I_2}{2\pi r}\right)_{SI}$$

Divergence of B

Consider the B produced by a wire of current: $\vec{B} = \frac{2I}{cr}\hat{\varphi}$

Calculate its divergence in Cartesian coordinates:

Was this a smart choice if variables???

Only if you REALLY like to spend time on math…

Using cylindrical coordinates:

$$\nabla\cdot\vec{B} = \frac{1}{r}\frac{d}{dr}(rB_r) + \frac{1}{r}\frac{dB_\varphi}{d\varphi} + \frac{dB_z}{dz} = 0$$

This is a **general property** of the magnetic field: $\nabla\cdot\vec{B} = 0$

Similar equation for E: $\nabla \cdot \vec{E} = 4\pi\rho$

The divergence of E is related to the density of electric charges

The divergence of B must be related to the density of magnetic charges

→ Magnetic monopoles don't exist

(There may be magnetic monopoles leftover from the Early Universe, but never observed experimentally so far except at Stanford!)

Thoughts on B

What exactly is a magnetic field B?

Why does it have so much in common with electric field E?

Why should there be a field that acts only on moving charges?

Answer: Special Relativity as we saw earlier.

Relativity: the physics must be the same in all reference frames

A charge at rest for observer 1 appears in motion to observer 2 that moves with a certain velocity w.r.t. observer 1:

Observer 1 will measure an electric field

Observer 2 will measure a magnetic field

Calculating attractive or repulsive force acting on a test charge in the 2 reference frames will lead to the same conclusions.

Change in B at a Current Sheet

We already discussed what happens to the B-field at a sheet of current using direct integration. We found that

$$\vec{B} = \begin{cases} +\dfrac{2\pi K}{c}\hat{x} & y > 0 \quad \text{(above the plane)} \\ -\dfrac{2\pi K}{c}\hat{x} & y < 0 \quad \text{(below the plane)} \end{cases}$$

K = current/length

or that the change of B (parallel to sheet) across the sheet of current is $\Delta B = \dfrac{4\pi K}{c}$

We can also derive this result using Ampere's law: $\oint_C \vec{B} \cdot d\vec{s} = \dfrac{4\pi}{c} I_{enclosed}$

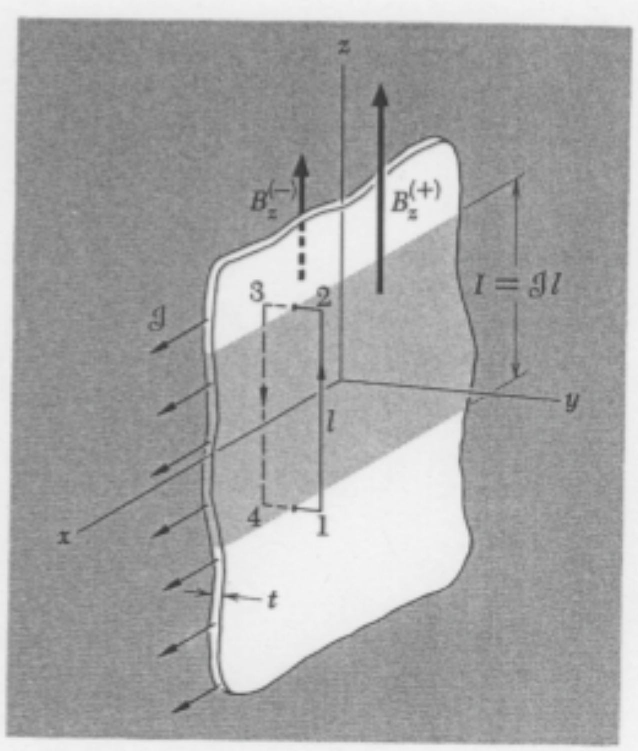


FIGURE 6.21
At a sheet of surface current there must be a change in the parallel component of **B** from one side to the other.

Sheet located in xz plane. ∞ extent. Thickness t (not important). J= current density --> every cm of height in z-direction includes a ribbon of current Jt = surface current density or sheet current density = K. K determines change in B-field across sheet. Also assume another field possibly present in z-direction from another source(why fields as shown).

Consider line integral of B around rectangle shown. Since no B-field perpendicular to sheet, we have line integral =

$$\ell(B_z^+ - B_z^-)$$

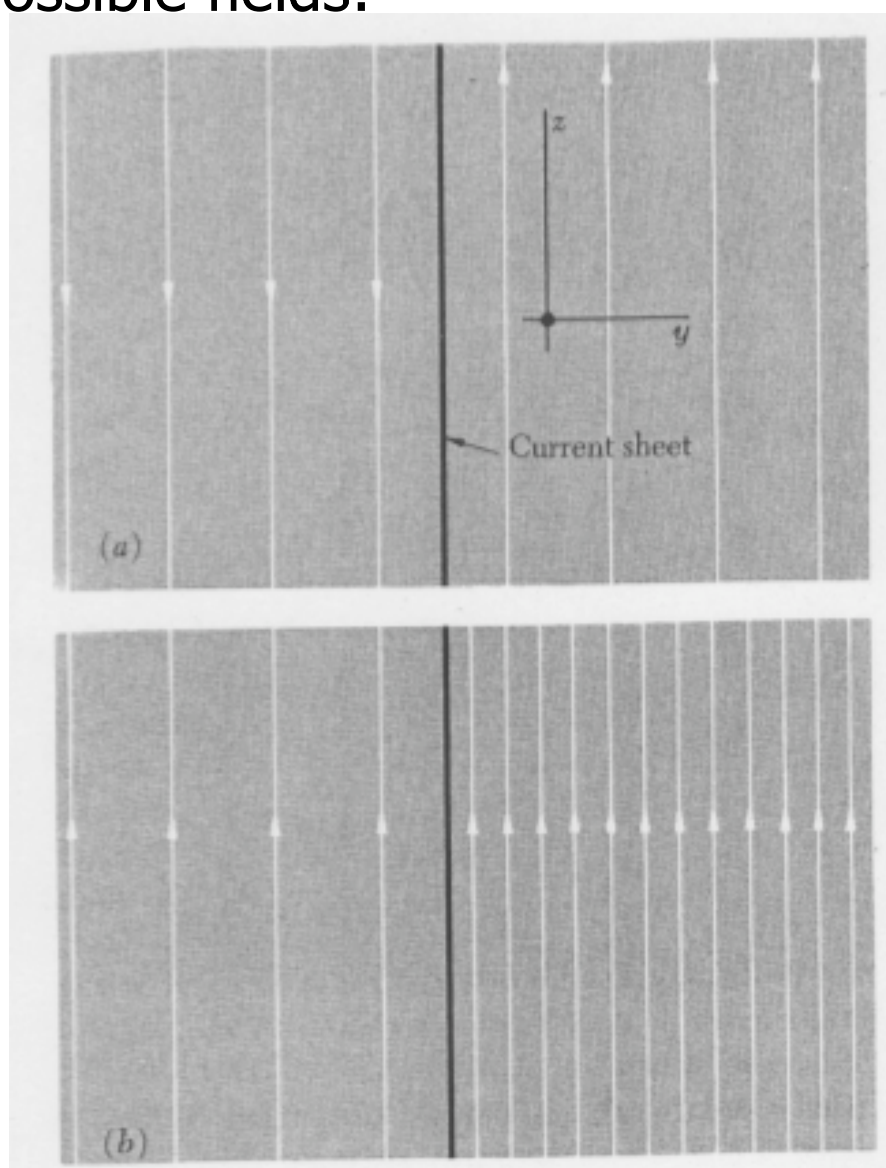
Current enclosed by rectangle is Kl. Thus,

$$\ell(B_z^+ - B_z^-) = \frac{4\pi K\ell}{c} \Rightarrow B_z^+ - B_z^- = \frac{4\pi K}{c}$$

Current sheet of density K --> jump in B-field parallel to surface and perpendicular to K.
If sheet = only current source then

$$B_z^+ = -B_z^- = \frac{2\pi K}{c}$$

Some possible fields:



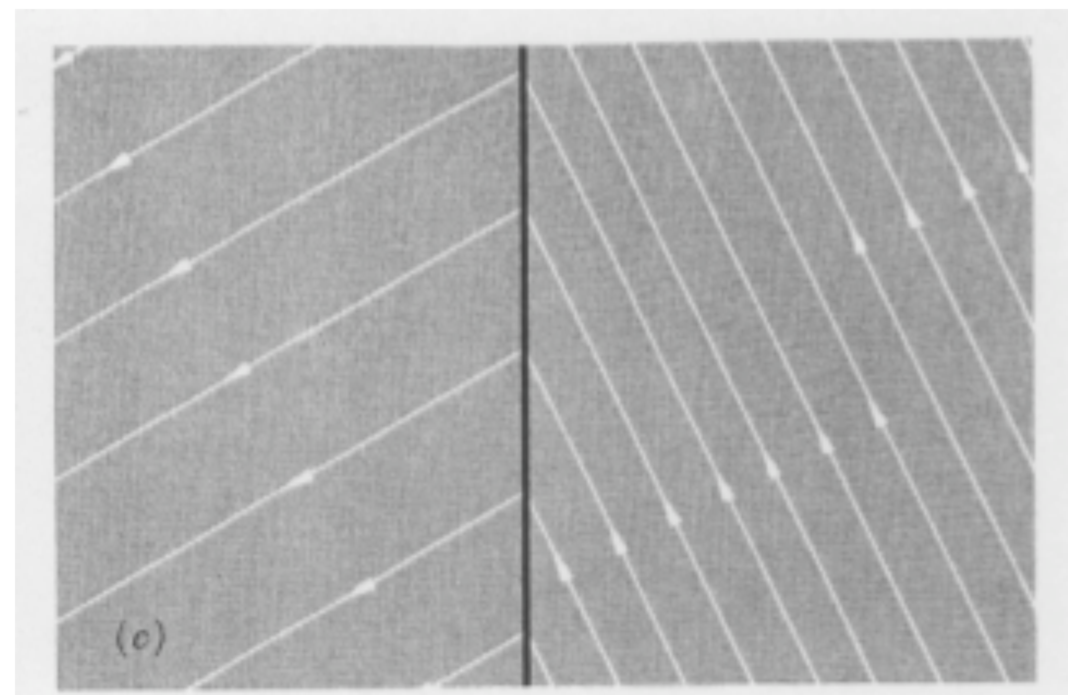


FIGURE 6.22
Some possible forms of the total magnetic field near a current sheet. Current flows in the $\hat{\mathbf{x}}$ direction (out of the page). (*a*) The field of the sheet alone. (*b*) Superposed on a uniform field in the *z* direction (this is like the situation in Fig. 6.21). (*c*) Superposed on a uniform field in another direction. In every case the component B_z changes by $4\pi\mathcal{K}/c$, on passing through the sheet, with no change in B_y.

Also if we have 2 sheets with currents in opposite direction ,then we have as shown. The field between the sheets is

$$\frac{4\pi K}{c}$$

and there is zero field outside!

Purcell calculates the Force/cm^2 on the sheets as

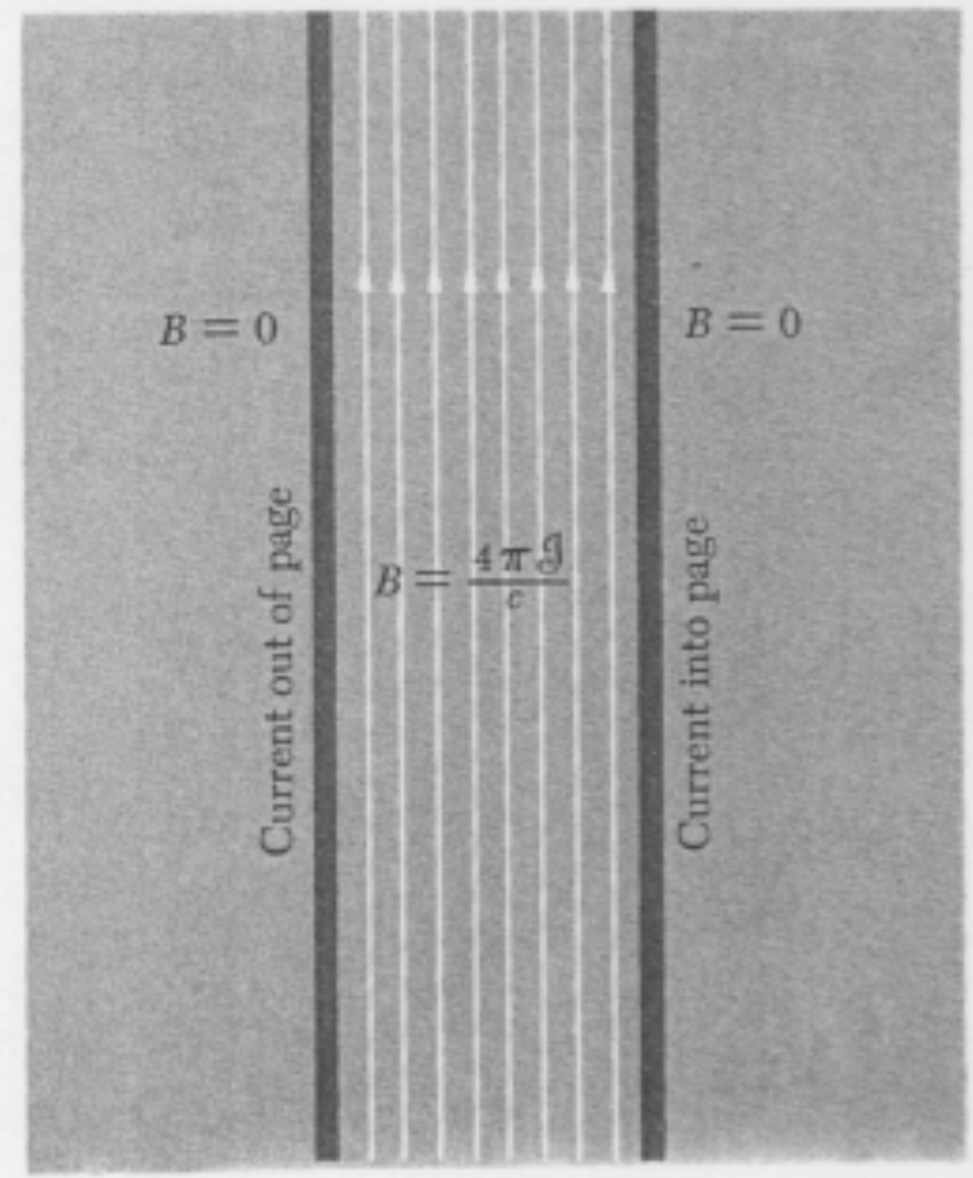
$$Force\,/\,cm^2 = \frac{1}{8\pi}\left(\left(B_z^+\right)^2 - \left(B_z^-\right)^2\right)$$

Force is perpendicular to surface and proportional to area --> like stress caused by hydrostatic pressure.

Force is repulsive for opposite currents(same as wires).

This just reflects fact that energy density is large inside the sheets and that produces a large pressure effect!

FIGURE 6.23
The magnetic field between plane-parallel current sheets.



How the fields transform:

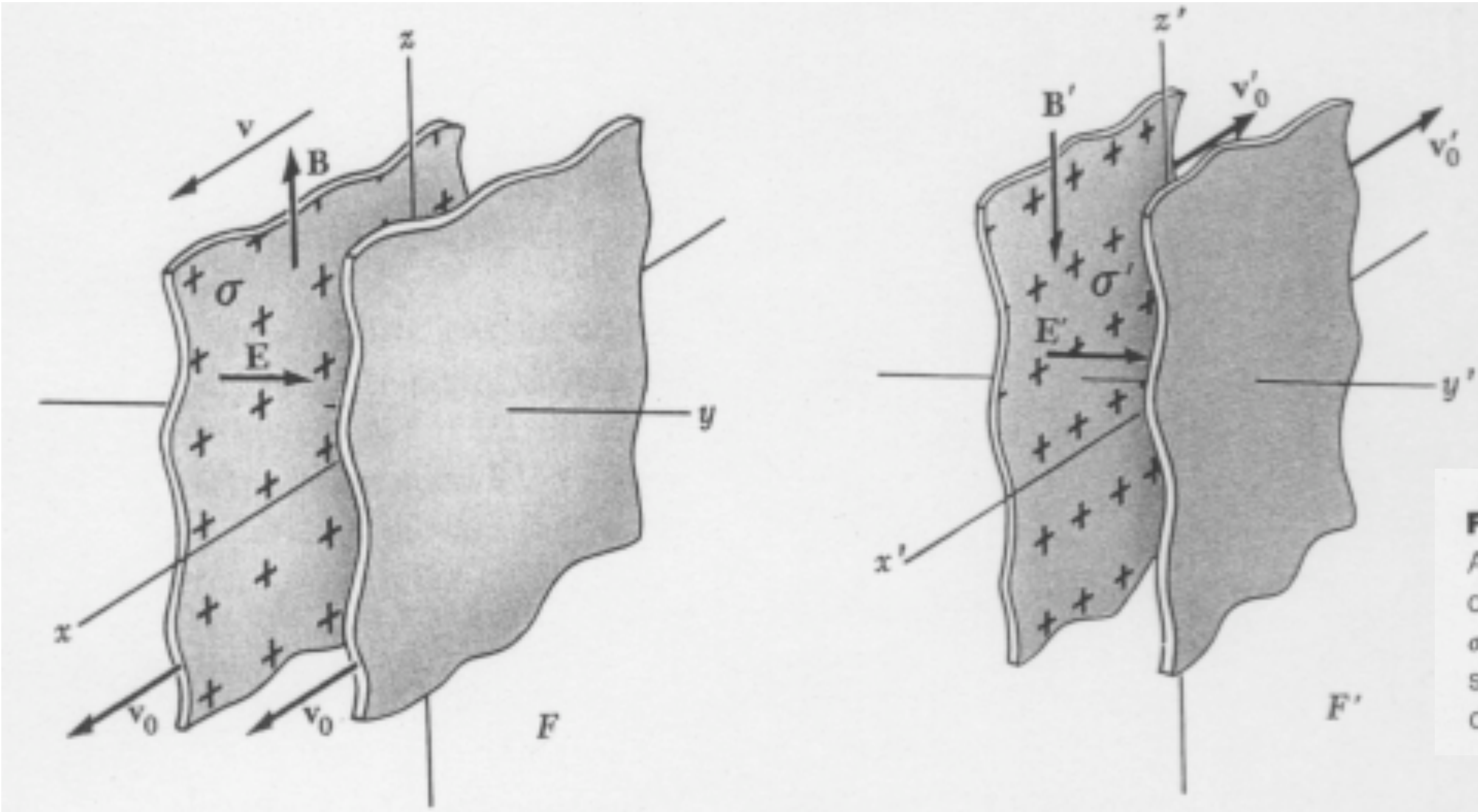


FIGURE 6.25
As observed in the frame *F* on the left, the surface charge density is σ and the surface current density is σv_0. Frame *F′* on the right moves in the *x* direction with speed *v* as seen from *F*. In *F′* the surface charge density is σ' and the current density is $\sigma' v_0'$.

Sheet of charge moving parallel to itself = surface current. I have σ and speed v, K=σv. Imagine two sheets of surface charge as shown(parallel to xz-plane). In frame F(xyz) surface charge = ±σ (amount of charge/area as measured in F ≠ density in rest frame of charges - smaller by 1/γ. In F uniform E-field -m +y-direction and = 4πσ (Gauss law). In F both sheets moving in +x-direction with speed v_0 --> pair of current sheets $K_x = \sigma v_0$. Thus B-field between sheets is

$$B_z = \frac{4\pi K}{c} = \frac{4\pi\sigma v_0}{c}$$

Frame F' moves (as seen by F) with speed v in positive x-direction. What fields will observer in F' measure?

In F' the x' velocity of charge-bearing sheet is v_0' =(using velocity addition formula)

$$v_0' = \frac{v_0 - v}{1 - \frac{v_0 v}{c^2}} = c\,\frac{\beta_0 - \beta}{1 - \beta_0\beta}$$

Density in rest frame of charges is $\sigma / \gamma_0 = \sigma\sqrt{1 - v_0^2 / c^2}$

Density in frame F' of charges is $\sigma' = \sigma \frac{\gamma_0'}{\gamma_0}$, $\gamma_0' = 1/\sqrt{1 - v_0'^2 / c^2}$

So that(same algebra as before) $\sigma' = \sigma\gamma(1 - \beta_0\beta)$

Surface current density in F' is charge density x charge velocity

$$K' = \sigma' v_0' = \sigma\gamma(1 - \beta_0\beta)c\,\frac{\beta_0 - \beta}{1 - \beta_0\beta} = \sigma\gamma\left(v_0 - v\right)$$

Since we now know how sources appear in F', Principle of Relativity tells us fields in F'.
Use identical formulas as in F but with primed values!!!!

$$E_y' = 4\pi\sigma' = \gamma\left[4\pi\sigma - \frac{4\pi\sigma}{c}\frac{v}{c}\right] = \gamma\left[E_y - \beta B_z\right]$$

$$B_z' = \frac{4\pi}{c}K' = \gamma\left[\frac{4\pi\sigma v_0}{c} - 4\pi\sigma\frac{v}{c}\right] = \gamma\left[B_z - \beta E_y\right]$$

If sheets oriented parallel to xy-plane instead then connect E_z' and E_z and B_y' and B_y. Get same form as above some difference in signs(due to B-field directions). Purcell also shows that fields in direction of motion do NOT change! Thus, full(not just for special example used to derive them) general transformations are given by

$$E_x' = E_x \quad , \quad E_y' = \gamma\left[E_y - \beta B_z\right] \quad , \quad E_z' = \gamma\left[E_z + \beta B_y\right]$$

$$B_x' = B_x \quad , \quad B_y' = \gamma\left[B_y + \beta E_z\right] \quad , \quad B_z' = \gamma\left[B_z - \beta E_y\right]$$

Astonishing fact: complete symmetry between E and B. If interchange E and B we would get SAME equations! Previously, though B is v/c effect (smaller than E) due to relativistic changes in E fields. Magnetic phenomena in macroscopic world(where we exist) are certainly different from electric phenomena. Macro world is NOT symmetrical wrt to E and B fields. With no sources, however, we find E and B connected in highly symmetrical way.

E and B fields appear to be different aspect or components of the same physical entity - the electromagnetic field. The same EM field viewed in different frames has different values of E and B components (similar to rotated vectors or Lorentz-transformed 4-vectors). EM field is NOT a vector however. It is a new mathematical object called a TENSOR (2nd rank) (Physics 50)

$$F_{\mu\nu} = \begin{bmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{bmatrix}$$

Another way to express transformations:

$$\vec{E} = \vec{E}_{||} + \vec{E}_{\perp} \quad , \quad \vec{E}' = \vec{E}_{||}' + \vec{E}_{\perp}'$$

$$\vec{B} = \vec{B}_{||} + \vec{B}_{\perp} \quad , \quad \vec{B}' = \vec{B}_{||}' + \vec{B}_{\perp}'$$

$$\vec{E}_{||}' = \vec{E}_{||} \quad , \quad \vec{E}_{\perp}' = \gamma\left[\vec{E}_{\perp} + \vec{\beta} \times \vec{B}_{\perp}\right]$$

$$\vec{B}_{||}' = \vec{B}_{||} \quad , \quad \vec{B}_{\perp}' = \gamma\left[\vec{B}_{\perp} - \vec{\beta} \times \vec{E}_{\perp}\right]$$

Remarkably simple relation between E and B fields in important class of cases. Suppose frame exists(F) where B-field = 0 in some region. Then in any other frame F′ which moves with velocity βc wrt to F, we have

$$\vec{E}_{||}' = \vec{E}_{||} \quad , \quad \vec{E}_{\perp}' = \gamma \vec{E}_{\perp} \quad , \quad \vec{B}_{||}' = 0 \quad , \quad \vec{B}_{\perp}' = -\gamma \vec{\beta} \times \vec{E}_{\perp}$$

But since $\vec{\beta} \times \vec{E}_{||} = 0$ in any case for $E_{||}$ parallel to β by definition, we have

$$\vec{B}_{\perp}' = \vec{B}_{\perp}' + \vec{B}_{||}' = \vec{B}' = -\vec{\beta} \times \left(\gamma \vec{E}_{\perp} + \vec{E}_{||}\right) = -\vec{\beta} \times \left(\vec{E}_{\perp}' + \vec{E}_{||}'\right)$$

$$\vec{B}' = -\vec{\beta} \times \vec{E}'$$

Holds true in every frame if B=0 in one frame. Similarly, if E=0 in one frame then

$$\vec{E}' = \vec{\beta} \times \vec{B}'$$

Very useful relations because only involve fields in a single frame; See Purcell.

What we learned about magnetism so far…

Magnetic Field B

Experiments: currents in wires generate forces on charges in motion

Force exerted on charge q with velocity v: $\vec{F} = q\frac{\vec{v}}{c} \times \vec{B}$

Explanation: there must exist a magnetic field B

Special Relativity: B is just E seen from another reference frame…

Ampere's Law: $\oint_C \vec{B} \cdot d\vec{s} = \frac{4\pi}{c} I_{enclosed}$

Application: B generated by current in a wire: $\vec{B} = \frac{2I}{cr}\hat{\varphi}$

Ampere's law in differential form

Apply Stoke's theorem to Ampere's law:

$$\oint_C \vec{B} \cdot d\vec{s} = \frac{4\pi}{c} I_{enclosed} \Rightarrow \int_S \left(\nabla \times \vec{B}\right) \cdot d\vec{S} = \frac{4\pi}{c} \int_S \vec{J} \cdot d\vec{S}$$

$$\int_S \left(\nabla \times \vec{B} - \frac{4\pi}{c}\vec{J}\right) \cdot d\vec{S} = 0 \Rightarrow \nabla \times \vec{B} = \frac{4\pi}{c}\vec{J}$$

➙Ampere's law in differential form: $\nabla \times \vec{B} = \frac{4\pi}{c}\vec{J}$

Toward Maxwell's equations

Let's collect all the equations in differential form that we found so far:

$\nabla \cdot \vec{E} = 4\pi\rho$ Relates E and charge density (ρ) -Gauss

$\nabla \cdot \vec{B} = 0$ No magnetic monopoles!

$\nabla \times \vec{E} = 0$ E is a conservative field

$\nabla \times \vec{B} = \frac{4\pi}{c}\vec{J}$ Relates B and its sources (J) -Ampere

Not complete Maxwell's equations yet, but we are getting closer…

Vector potential A

Definition of potential for electric field:

ϕ(P) = work needed to move a unit charge from reference to P

Relationship between ϕ and E: $\vec{E} = -\nabla\phi$

Bonus:

$$If\ \vec{E} = -\nabla\phi \Rightarrow \nabla \times \vec{E} = 0\ because\ \nabla \times \left(\nabla\phi\right)\ \forall\ \phi$$

Reminder: $\hat{e}_i \times \hat{e}_j = \varepsilon_{ijk}\hat{e}_k$

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if ijk = even permutation of 123} \\ -1 & \text{if ijk = odd permutation of 123} \\ 0 & \text{otherwise} \end{cases}$$

$$\varepsilon_{123} = \varepsilon_{312} = \varepsilon_{231} = 1 \quad , \quad \varepsilon_{213} = \varepsilon_{132} = \varepsilon_{321} = -1$$

$$\varepsilon_{113} = \varepsilon_{112} = \varepsilon_{121} = \varepsilon_{131} = \varepsilon_{311} = \varepsilon_{211} = \varepsilon_{111} = \varepsilon_{221} = \varepsilon_{223} = \varepsilon_{212} = 0$$

$$\varepsilon_{232} = \varepsilon_{322} = \varepsilon_{122} = \varepsilon_{222} = \varepsilon_{331} = \varepsilon_{332} = \varepsilon_{313} = \varepsilon_{323} = \varepsilon_{133} = \varepsilon_{233} = \varepsilon_{333} = 0$$

$$\vec{A}\times\vec{B}=\varepsilon_{ijk}A_iB_j\hat{e}_k=\left(A_2B_3-A_3B_2\right)\hat{e}_1+\left(A_3B_1-A_1B_3\right)\hat{e}_2+\left(A_1B_2-A_2B_1\right)\hat{e}_3$$

$$curl(\vec{f})=\nabla\times\vec{f}=\varepsilon_{ijk}\nabla_i f_j\hat{e}_k=\left(\nabla_2 f_3-\nabla_3 f_2\right)\hat{e}_1+\left(\nabla_3 f_1-\nabla_1 f_3\right)\hat{e}_2+\left(\nabla_1 f_2-\nabla_2 f_1\right)\hat{e}_3$$

$$=\left(\frac{\partial f_z}{\partial y}-\frac{\partial f_y}{\partial z}\right)\hat{x}+\left(\frac{\partial f_x}{\partial z}-\frac{\partial f_z}{\partial x}\right)\hat{y}+\left(\frac{\partial f_y}{\partial x}-\frac{\partial f_x}{\partial y}\right)\hat{z}$$

$$\nabla\times\nabla\phi=curl(grad(\phi))=\left(\frac{\partial}{\partial y}\frac{\partial\phi}{\partial z}-\frac{\partial}{\partial z}\frac{\partial\phi}{\partial y}\right)\hat{x}+\left(\frac{\partial}{\partial z}\frac{\partial\phi}{\partial x}-\frac{\partial}{\partial x}\frac{\partial\phi}{\partial z}\right)\hat{y}+\left(\frac{\partial}{\partial x}\frac{\partial\phi}{\partial y}-\frac{\partial}{\partial y}\frac{\partial\phi}{\partial x}\right)\hat{z}$$

$$Since\quad\frac{\partial^2\phi}{\partial y\partial z}=\frac{\partial^2\phi}{\partial z\partial y}\quad in\ physical\ systems,we\ have$$

$$\nabla\times\nabla\phi=0$$

Alternatively,

$$\nabla\times\nabla\phi=\varepsilon_{ijk}\frac{\partial}{\partial x_j}\frac{\partial}{\partial x_k}\phi=\varepsilon_{ijk}\frac{\partial^2\phi}{\partial x_j\partial x_k}=\frac{1}{2}\left(\varepsilon_{ijk}\frac{\partial^2\phi}{\partial x_j\partial x_k}+\varepsilon_{ikj}\frac{\partial^2\phi}{\partial x_k\partial x_j}\right)$$

$$=\frac{1}{2}\left(\varepsilon_{ijk}\frac{\partial^2\phi}{\partial x_j\partial x_k}-\varepsilon_{ijk}\frac{\partial^2\phi}{\partial x_k\partial x_j}\right)=\frac{\varepsilon_{ijk}}{2}\left(\frac{\partial^2\phi}{\partial x_j\partial x_k}-\frac{\partial^2\phi}{\partial x_k\partial x_j}\right)=0$$

Symmetric x antisymmetric clever argument - very useful!!!!

Another useful derivation. Will do again in Physics 50 so useful to try to understand now.

$$\nabla\times\left(\nabla\times\vec{B}\right)=\nabla\times\left(\varepsilon_{ijk}\nabla_j B_k\hat{e}_i\right)=\left(\nabla\times\hat{e}_i\right)\left(\varepsilon_{ijk}\nabla_j B_k\right)=\left(\varepsilon_{lmn}\nabla_m\left(\hat{e}_i\right)_n\hat{e}_l\right)\left(\varepsilon_{ijk}\nabla_j B_k\right)$$

$$=\left(\varepsilon_{lmn}\nabla_m\delta_{in}\hat{e}_l\right)\left(\varepsilon_{ijk}\nabla_j B_k\right)=\left(\varepsilon_{lmi}\varepsilon_{ijk}\nabla_m\nabla_j B_k\hat{e}_l\right)=\varepsilon_{ilm}\varepsilon_{ijk}\nabla_m\nabla_j B_k\hat{e}_l$$

$$=\left(\delta_{lj}\delta_{mk}-\delta_{lk}\delta_{mj}\right)\nabla_m\nabla_j B_k\hat{e}_l=\nabla_k\nabla_j B_k\hat{e}_j-\nabla_j\nabla_j B_k\hat{e}_k=\nabla\left(\nabla\cdot\vec{B}\right)-\nabla^2\vec{B}$$

Can we introduce something similar to φ for B?

Goal: enforce $div\vec{B}=\nabla\cdot\vec{B}=0$

Since $\nabla\cdot\nabla\times\vec{f}=0$ for any $\vec{f}$

$$\nabla\bullet\nabla\times\vec{f}=\left(\nabla_m\hat{e}_m\right)\bullet\left(\varepsilon_{ijk}\nabla_i f_j\hat{e}_k\right)=\varepsilon_{ijk}\nabla_m\nabla_i f_j\hat{e}_m\bullet\hat{e}_k=\varepsilon_{ijk}\nabla_m\nabla_i f_j\delta_{mk}=\varepsilon_{ijk}\nabla_k\nabla_i f_j=0$$

we define

$$\vec{B}=\nabla\times\vec{A}$$

A is called “vector potential” in analogy with φ

A, however, is not connected to work or energy (but to angular momentum)

Non-Uniqueness

Electrostatics: given a charge distribution and boundary conditions the potential ϕ is uniquely identified (= 0 at ∞)

Magnetism: does it work the same for A? No, there are infinite number of A's corresponding to a single B

Example: $\vec{B} = -B_0\hat{z}$ Find A that creates this B field. Q: what current creates this B?

Requirements:

$$\begin{Bmatrix} B_x = \dfrac{\partial A_z}{\partial y} - \dfrac{\partial A_y}{\partial z} = 0 \\ B_y = \dfrac{\partial A_x}{\partial z} - \dfrac{\partial A_z}{\partial x} = 0 \\ B_z = \dfrac{\partial A_x}{\partial y} - \dfrac{\partial A_y}{\partial x} = 0 \end{Bmatrix} \Rightarrow$$

Possible solutions:

$$\vec{A} = -yB_0\hat{x} \quad , \quad \vec{A} = xB_0\hat{y}$$

$$\vec{A} = \frac{B_0}{2}\left(-y\hat{x} + x\hat{y}\right) \quad , \quad \vec{A} = \text{infinite others!}$$

➜ we are given one "coupon" or choice of A to simplify equations whenever needed.

Poisson's equation for A

Electrostatics:

$$\begin{Bmatrix} \vec{E} = -\nabla\phi \\ \nabla \cdot \vec{E} = 4\pi\rho \end{Bmatrix} \Rightarrow \nabla^2\phi = -4\pi\rho$$

Poisson's equation

Magnetism:

$$\begin{Bmatrix} \vec{B} = \nabla \times \vec{A} \\ \nabla \times \vec{B} = \dfrac{4\pi}{c}\vec{J} \end{Bmatrix} \Rightarrow \nabla \times \nabla \times \vec{A} = \frac{4\pi}{c}\vec{J} \Rightarrow \nabla\left(\nabla \cdot \vec{A}\right) - \nabla^2\vec{A} = \frac{4\pi}{c}\vec{J}$$

We used the identity:

$$\nabla \times \nabla \times \vec{A} = \nabla\left(\nabla \cdot \vec{A}\right) - \nabla^2\vec{A}$$

Use your coupon now!

Choosing $\nabla \cdot \vec{A} = 0 \quad \Rightarrow \nabla^2 \vec{A} = -\frac{4\pi}{c}\vec{J}$ Note: Laplacian is scalar operator ➞ A || J !

Solving Poisson's equation for A

How do you solve $\nabla^2 \vec{A} = -\frac{4\pi}{c}\vec{J}$ Goofy equation

Think of it in cartesian coordinates(components):

$$\nabla^2 A_x = -\frac{4\pi}{c}J_x \quad , \quad \nabla^2 A_y = -\frac{4\pi}{c}J_y \quad , \quad \nabla^2 A_z = -\frac{4\pi}{c}J_z$$

Each component of A related to component of J via Laplacian operator

Suppose $\nabla \cdot \vec{A} \neq 0$. *Then choose* $\vec{A}'$ *s.t.*

$\vec{A}' = \vec{A} + \nabla f \;\; s.t. \;\; \nabla \cdot \vec{A}' = 0 \;\; i.e.,$

$\nabla \cdot \vec{A}' = \nabla \cdot \vec{A} + \nabla^2 f = 0 \rightarrow \nabla^2 f = -\nabla \cdot \vec{A}$

This does not change $\vec{B}$ *i.e.,*

$\vec{B}' = \nabla \times \vec{A}' = \nabla \times \vec{A} + \nabla \times \nabla f = \nabla \times \vec{A} = \vec{B}$

So can always choose $\nabla \cdot \vec{A} = 0$

Remember Poisson's equation $\nabla^2 \phi = -4\pi\rho$ and its general solution for φ $\phi = \int_V \frac{\rho}{r} dV$

Same as our new equation if replace

$$\phi \rightarrow \vec{A} \quad and \quad \rho \rightarrow \frac{\vec{J}}{c} \quad \Rightarrow \vec{A} = \frac{1}{c}\int_V \frac{\vec{J}}{r} dV$$

For current flowing in a wire: $\vec{J} = I d\vec{l}$ ➜ $\vec{A} = \frac{I}{c}\int_{wire} \frac{d\vec{l}}{r}$

Biot-Savart Law

Find $\vec{B}$ produced from current knowing that $\vec{A} = \frac{I}{c}\int_{wire} \frac{d\vec{l}}{r}$

$$\vec{B} = \nabla \times \vec{A} = \nabla \times \frac{I}{c}\int_{wire} \frac{d\vec{l}}{r} = \frac{I}{c}\int_{wire} \nabla \times \frac{d\vec{l}}{r}$$

Using the fact that $\nabla \times (a\vec{b}) = a\left(\nabla \times \vec{b}\right) + (\nabla a) \times \vec{b}$ we can write:

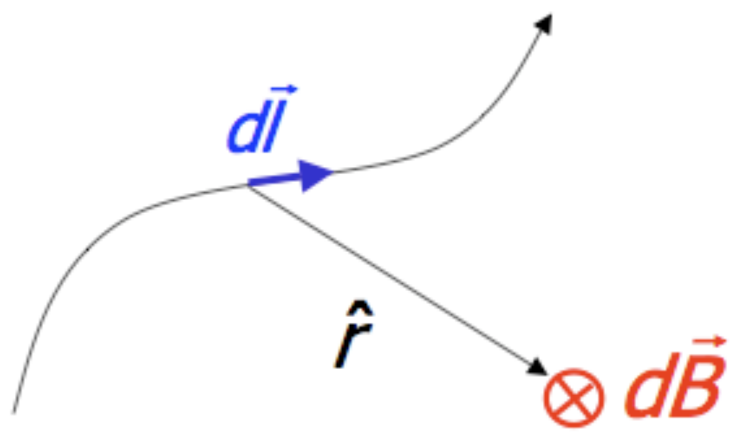
$$\vec{B} = \frac{I}{c}\left[\int_{wire} \frac{1}{r}\nabla \times d\vec{l} + \int_{wire} \left(\nabla \frac{1}{r}\right) \times d\vec{l}\right]$$

Since $\nabla \times d\vec{l} = 0 \left(d\vec{l} \text{ is a constant!} \right)$ *and* $\nabla \frac{1}{r} = -\frac{\hat{r}}{r^2}$

$$\Rightarrow \vec{B} = \frac{I}{c} \int_{wire} d\vec{l} \times \frac{\hat{r}}{r^2}$$

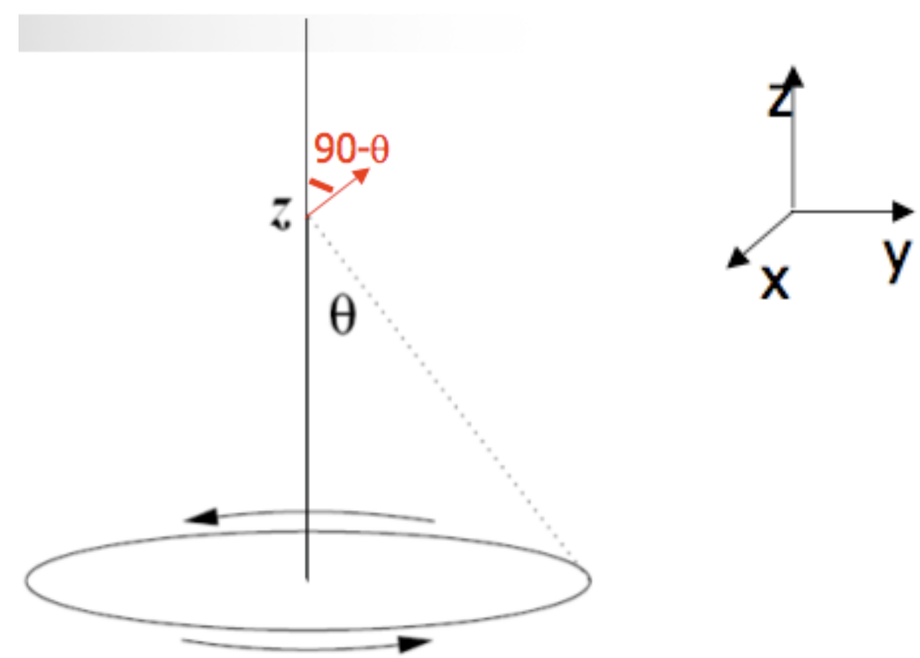
Biot-Savart Law: illustration

Biot-Savart: $d\vec{B} = \frac{I}{c} d\vec{l} \times \frac{\hat{r}}{r^2}$



dB is perpendicular to current and to radial direction
E.g.: if you have dl || x and r || y → B || z

Application of Biot-Savart: B from loop of current

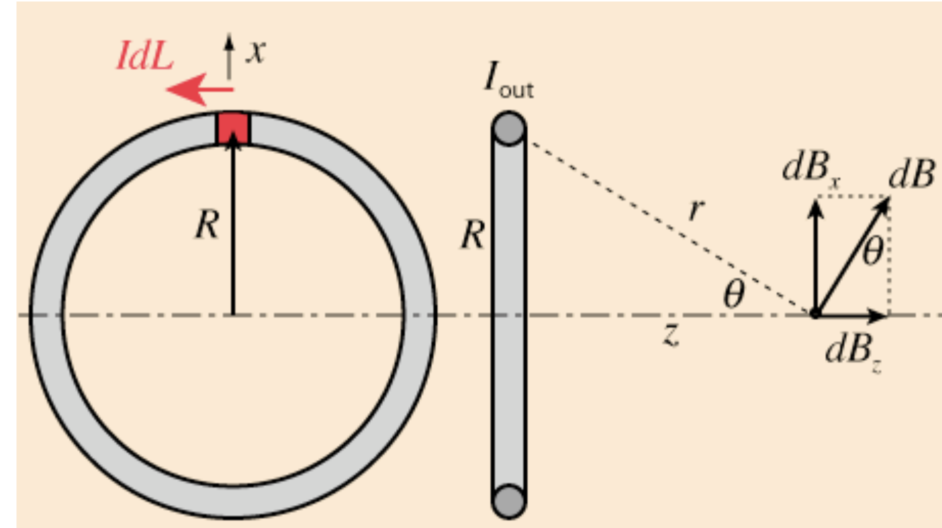


- Calculate B created by a loop of current
  - Radius: R
  - Distance from center of the loop: z
- Solution on axis
  - Apply Biot-Savart
  - Determine direction of dB
  - Symmetry→only component || z survives

$$B = \int_{wire} (d\vec{B})_z = \int_{wire} \frac{I}{cr^2} \left| d\vec{l} \times \hat{r} \right| \sin\theta$$

$$\left| d\vec{l} \times \hat{r} \right| = \left| d\vec{l} \right| = R d\varphi \quad , \quad \sin\theta = R/r \quad , \quad r = \sqrt{R^2 + h^2}$$

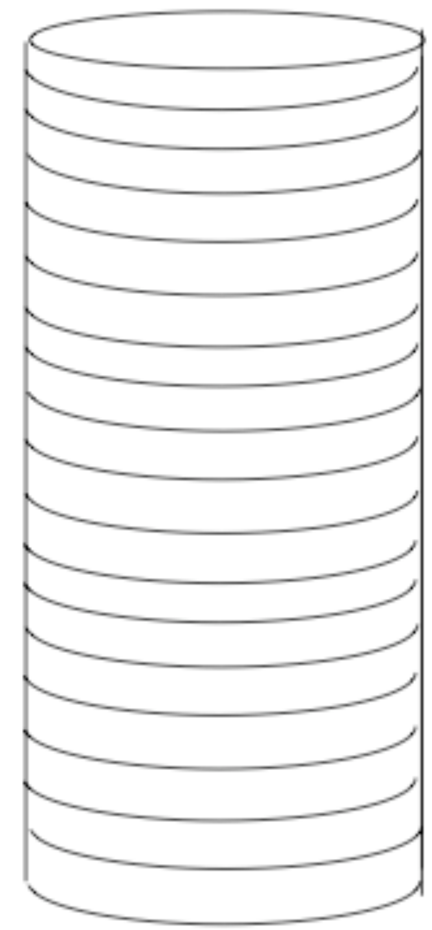
$$\vec{B} = \frac{I}{cr^2} R \sin\theta \int_0^{2\pi} d\varphi \hat{z} = \frac{2\pi I R^2}{c\left(R^2 + z^2\right)^{3/2}} \hat{z} \Rightarrow \vec{B}_{loop\ center} = \frac{2\pi I}{cR} \hat{z}$$



Application of Biot-Savart: B from solenoid

What if we stack N rings over a length L?
Use result of single loop + superposition:

Single ring: $$d\vec{B} = \hat{z} \frac{2\pi R^2}{c\left(R^2 + z^2\right)^{3/2}} dI$$



Integrate on all rings (in the middle of the solenoid)

$$\vec{B} = \hat{z} \int_{-L/2}^{L/2} \frac{2\pi R^2}{c\left(R^2 + z^2\right)^{3/2}} nI dz = \hat{z} \frac{2\pi nI}{c} \int_{-L/2}^{L/2} \frac{R^2}{\left(R^2 + z^2\right)^{3/2}} dz = \hat{z} \frac{2\pi nI}{c} \frac{2L}{\sqrt{L^2 + 4R^2}}$$

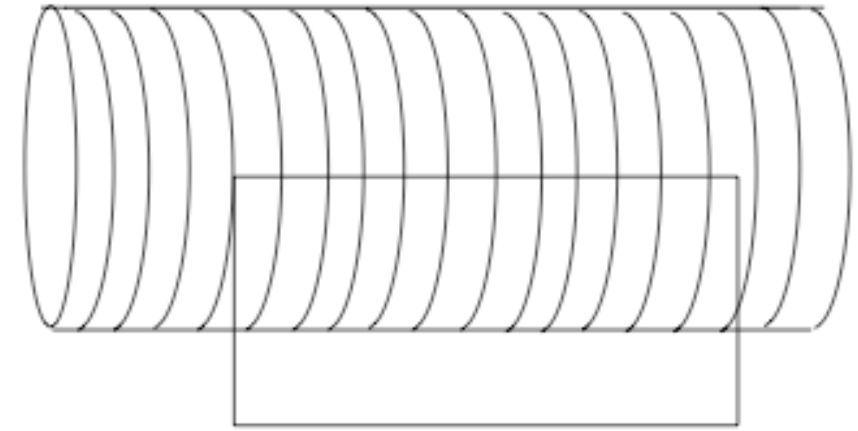
with n = N/L

For L >> R: $$\vec{B} = \hat{z} \frac{4\pi nI}{c}$$ Note: so far result valid on axis only…

Solenoid and Ampere's law: This value turns out to describe the magnetic field everywhere in the interior of the solenoid for L >> R. One can also prove that B outside the solenoid is ≈0 for L >> R. It suddenly jumps to $4\pi nI / c$ inside (like crossing sheet of current!).

Ampere can be used to simply prove that B does not depend on r:

$$\oint_{\text{rectangle}} \vec{B}\cdot d\vec{l} = \frac{4\pi}{c} I_{enclosed}$$



Since $\vec{B}$ is || z and present only inside the solenoid:

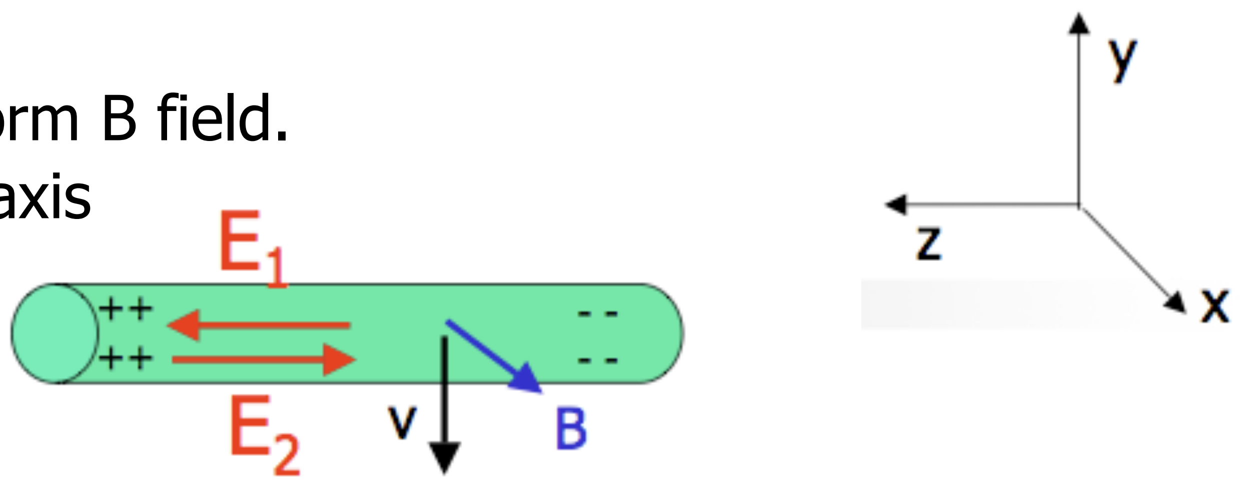
$$B(r) = \frac{4\pi}{c}\frac{N}{L} I = \frac{4\pi}{c} nI$$ no dependence on r

Moving rod in uniform B

Let's move a conducting rod in a uniform B field.

Charges move with velocity v || -y axis

B || x axis



What happens?

Lorentz force: $\vec{F}_{Lorentz} = q\frac{\vec{v}}{c}\times\vec{B}$

Separation of charges can be explained by an effective electric field E_1 || z

$\vec{F}_{Lorentz} = q\vec{E}_1$ Gives same separation as force due to B.

Separation of charges creates opposite electric field E_2 that exactly compensates E_1 and equilibrium is established:

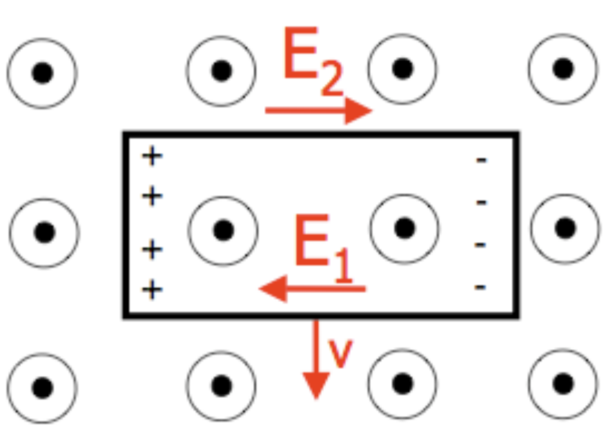
$$\vec{E}_2 = -\frac{\vec{v}}{c}\times\vec{B}$$

Moving a loop in uniform B(out of paper)

Now move a rectangular loop of wire in uniform B

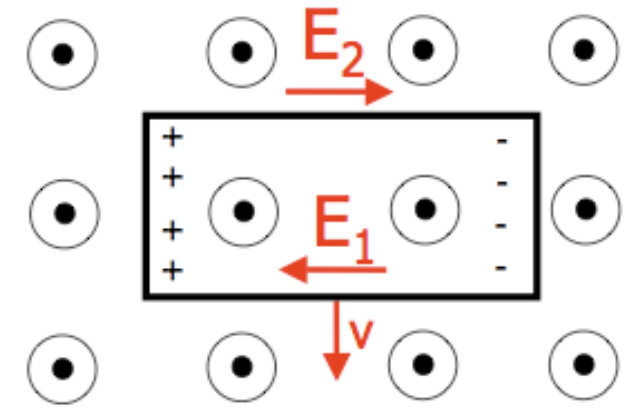
Same velocity || -y

Same B|| x



What happens?

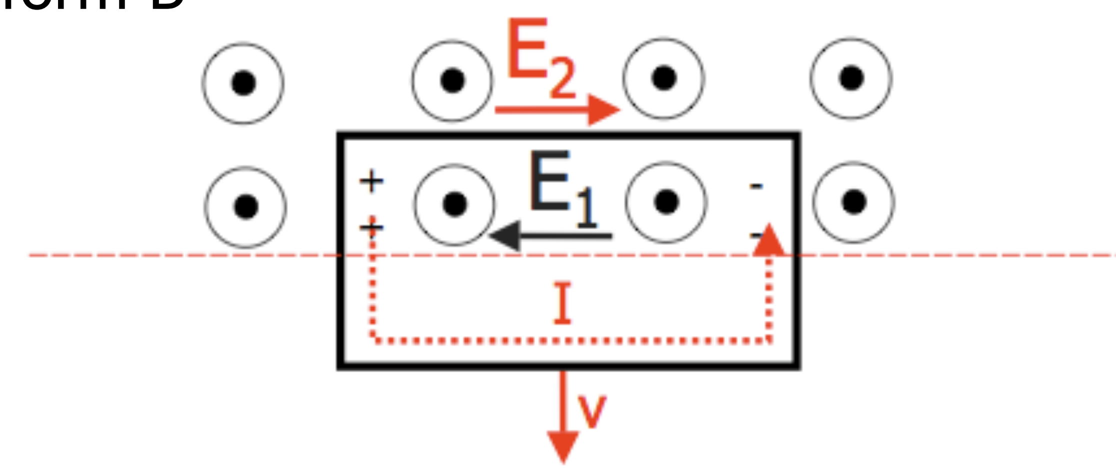
- Lorentz force ➙ E_1
- E_1 ➙ separation of charges on the wire
- Separation of charges creates opposite electric field $E_2 = -E_1$:



What if B is non-uniform?

Now move the rectangular loop of wire in non-uniform B

- Velocity v
- $B = B_0$ above ---
- $B = 0$ below ---



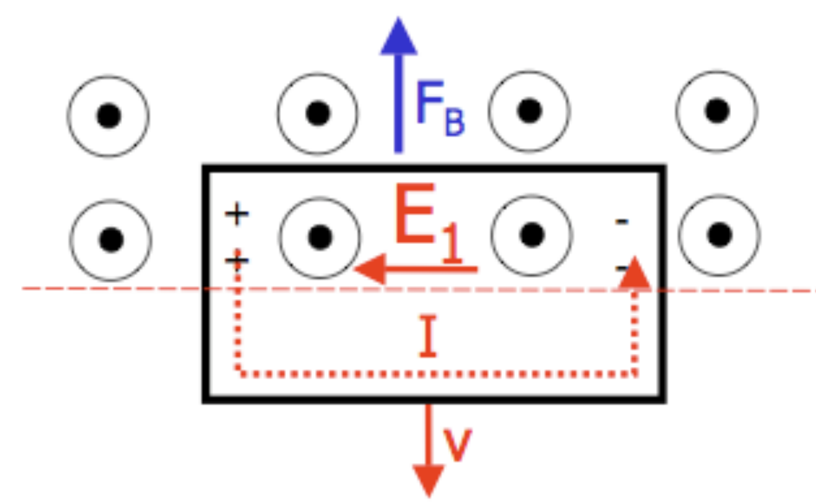
What happens?

- Lorentz force ➙ E_1
- E_1 ➙ separation of charges on the wire
- Separation of charges causes charges to flow in the loop
  (no opposing force in the bottom part!): current I!
- This phenomenon is called **electromagnetic induction**

Comments on induction:

Please notice the following:

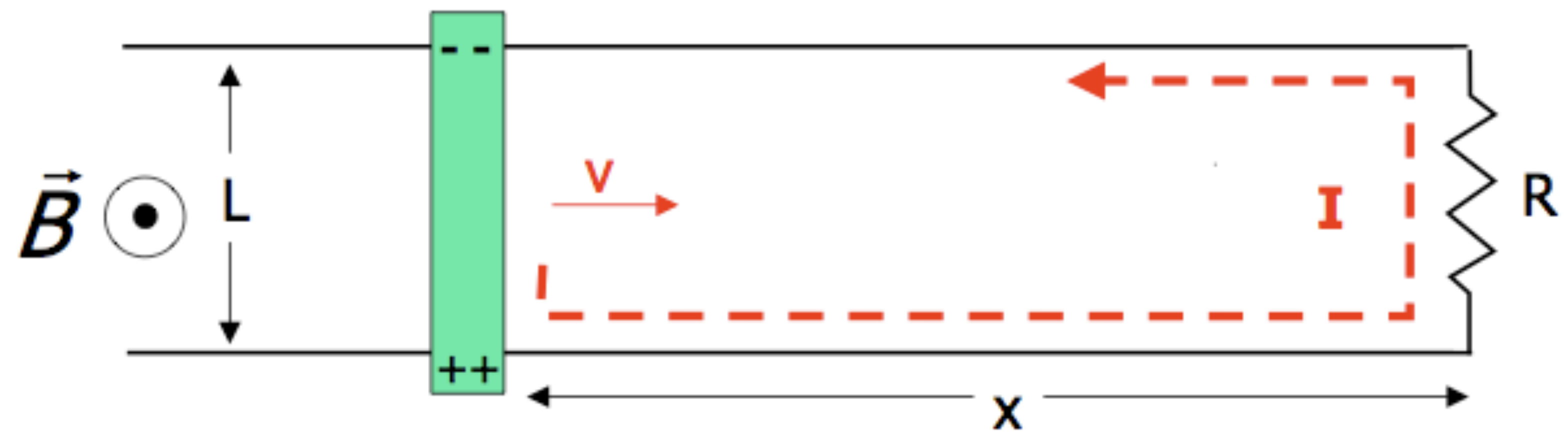
End of electrostatics! $\oint_{loop} \vec{E} \cdot d\vec{l} \neq 0 \quad or \quad \nabla \times \vec{E} \neq 0$



The current flowing in top leg of the loop will feel a force F_B (from B) pointing up, opposite to v, trying to slow down the motion that causes current to flow ➙ Lenz's law

Induced emf

Consider a sliding conducting bar on rails closed on a resistor R in a region of constant magnetic field B



Charge separation in the bar will induce current → e.m.f.

$$emf = \frac{1}{q} W(- \rightarrow +) = \frac{1}{q} \int_{-}^{+} \vec{F} \cdot d\vec{s} = \frac{1}{c} \int_{-}^{+} \left(\vec{v} \times \vec{B} \right) \cdot d\vec{s} = \frac{vBL}{c}$$

Current flowing in the loop: $I = \frac{emf}{R} = \frac{vBL}{cR}$

Faraday's law

EMF in the loop: $emf = \frac{vBL}{c} = \frac{BL}{c} \frac{dx}{dt}$

Magnetic flux in the rectangle is defined as: $\Phi_B = BLx$ Magnetic flux = B x Area

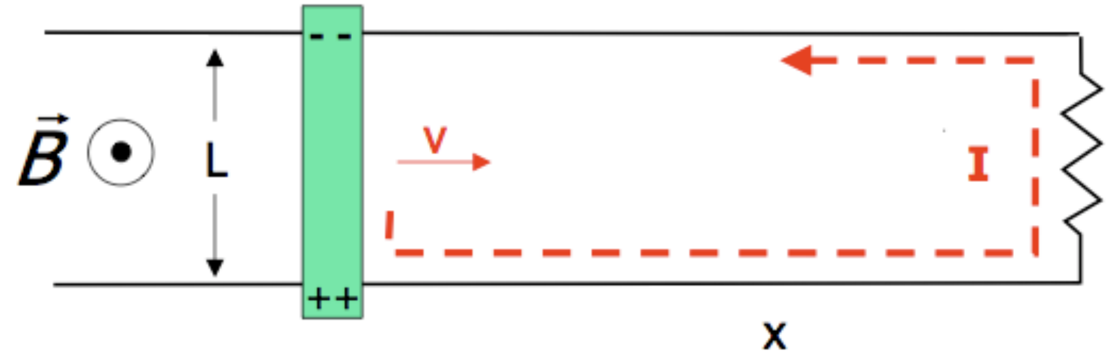
Combine the two keeping in mind that given the direction of v, the flux decreases with time:

Faraday's law: $emf = -\frac{1}{c} \frac{\partial \Phi_B}{\partial t}$

The minus sign is important: Lenz's law

It indicates that the direction of the current is such to oppose the changes in flux of B: ~ "electromagnetic inertia"

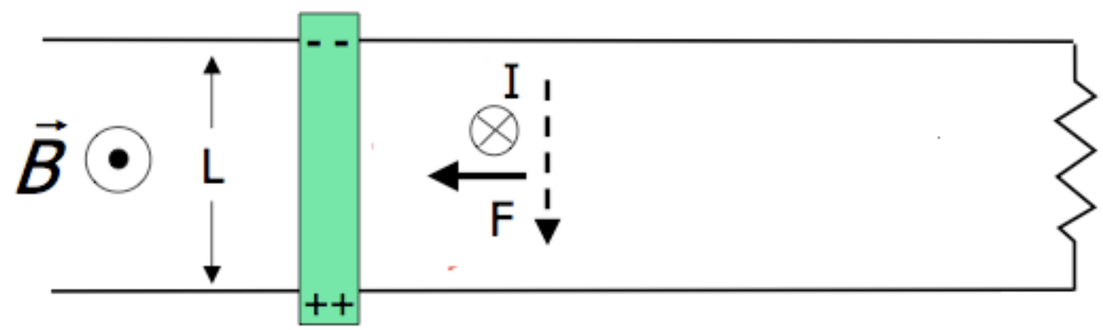
Thoughts on Lenz's law(1)



$$e.m.f. = -\frac{1}{c}\frac{\partial \Phi_B}{\partial t}$$

The current generated in wire opposes changes in flux of B
- v is L ➙ R:
  - Flux of B decreases over time ➙ e.m.f. is created with direction that compensates this change: counterclockwise
- v is R ➙ L:
  - Flux of B increases over time ➙ e.m.f. is created with direction that compensates this change: clockwise

Thoughts on Lenz's law (2)



$$e.m.f. = -\frac{1}{c}\frac{\partial \Phi_B}{\partial t}$$

When current flows in magnetic field it feels a force: $\vec{F} \propto \vec{I} \times \vec{B}$

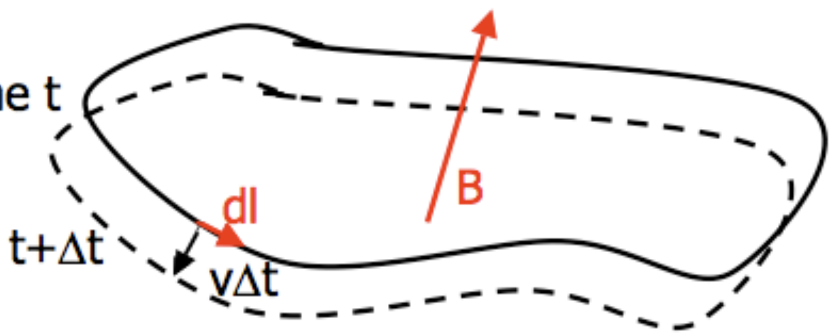
Lenz: the force will be will try to slow down the bar
- If I clockwise:
  - It creates a B pointing into the board $\vec{I} \times \vec{B}$ points to the left
- If I counterclockwise:
  - It creates a B pointing out of the board $\vec{I} \times \vec{B}$ points to the right

Note: the – sign in Lenz's law is what allows conservation of energy

General proof of Faraday's law

Consider a loop of arbitrary shape moving with velocity v through a static magnetic field B

At time t, the flux through the loop is: $\Phi_B = \int_S \vec{B} \cdot d\vec{a}$



How does it change when t ➙ t+Δt?

$$\Delta\Phi_B = \Phi_B(t+\Delta t) - \Phi_B(t) = \Phi_{ribbon} = \int_{ribbon} \vec{B} \cdot d\vec{a}$$

On the ribbon, the area element is:

$$d\vec{a} = (\vec{v}\Delta t) \times d\vec{l}$$

Surface S spans loop at time t; imagine moving loop down; at t+Δt new surface spanning loop = old(S) + ribbon

This means that:

$$\Delta\Phi_B = \int_{ribbon} \vec{B} \cdot d\vec{a} = \int_{ribbon} \vec{B} \cdot (\vec{v}\Delta t) \times d\vec{l} = \int_{ribbon} \Delta t \vec{B} \cdot \left(\vec{v} \times d\vec{l}\right)$$

Using the identity $\vec{a} \cdot (\vec{b} \times \vec{c}) = (\vec{a} \times \vec{b}) \cdot \vec{c})$ we obtain:

$$\Delta\Phi_B = \int_{loop} \Delta t \vec{B} \cdot \left(\vec{v} \times d\vec{l}\right) = \Delta t \int_{loop} \left(\vec{B} \times \vec{v}\right) \cdot d\vec{l} = -\Delta t \int_{loop} \left(\vec{v} \times \vec{B}\right) \cdot d\vec{l}$$

$$For\ \Delta t \to 0: \quad \frac{\partial \Phi_B}{\partial t} = -c \int_{loop} \left(\frac{\vec{v}}{c} \times \vec{B}\right) \cdot d\vec{l}$$

Since $\frac{\vec{v}}{c} \times \vec{B}$ is the magnetic force for a unit charge

➙ its line integral on the loop is the work necessary to move a unit charge around the wire: = the e.m.f!

$$\rightarrow emf = -\frac{1}{c}\frac{\partial \Phi_B}{\partial t}$$

Work from B??? Since F_B perpendicular to v , how can we have any work?

Faraday's law: $emf = -\frac{1}{c}\frac{\partial \Phi_B}{\partial t}$

This means that $\vec{v}/c \times \vec{B}$ integrated over the loop is the work that we have to do to move a unit charge around the loop

But last time we proved that B cannot do work

Are these 2 statements inconsistent???

No, the work done to move the charges is not done by B

It's done by whoever is moving the loop in B!!

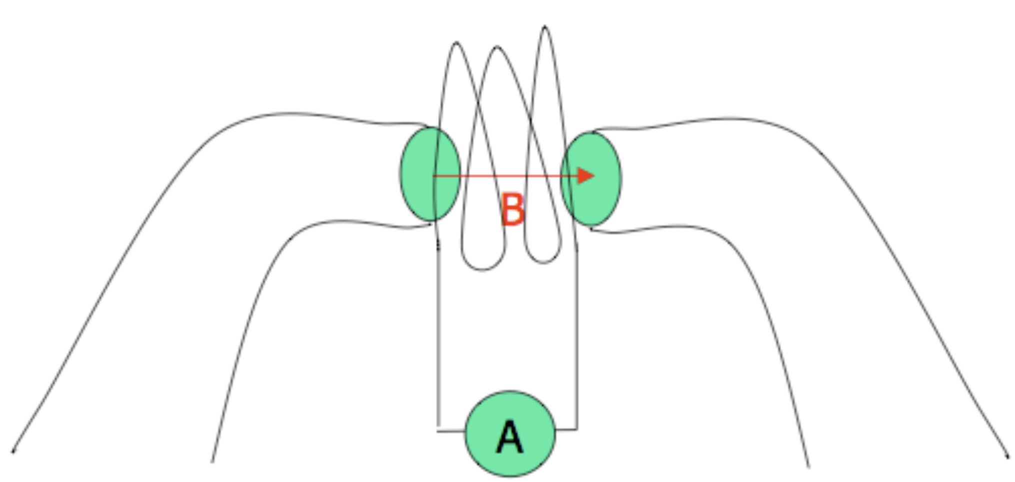
Verification of Faraday's law

Faraday's law: $emf = -\frac{1}{c}\frac{\partial \Phi_B}{\partial t}$

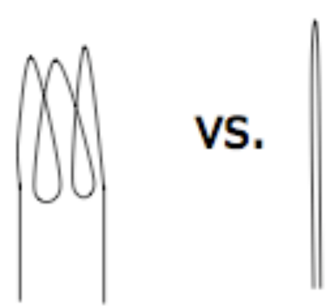
What does it mean?

emf is produced when the flux of B changes over time

area of the loop cannot be null!



-move loop in B
-current flows in wire
-if we used instead a wire with 0 area: no I



"Relativity"

What if loop is static and B changes?

Relativity tells me that we should get the same result

Same problem from another reference frame!!

Does this make sense?

Charges do not move in the other reference frame

What causes the force? The induced electric field.

Since e.m.f. is the work necessary to move a unit charge around the loop:

$$emf = \oint_C \vec{E} \cdot d\vec{l}$$

Example: magnet bar moving in the loop

Application of Lenz's law

Disk falling in a magnetic field B

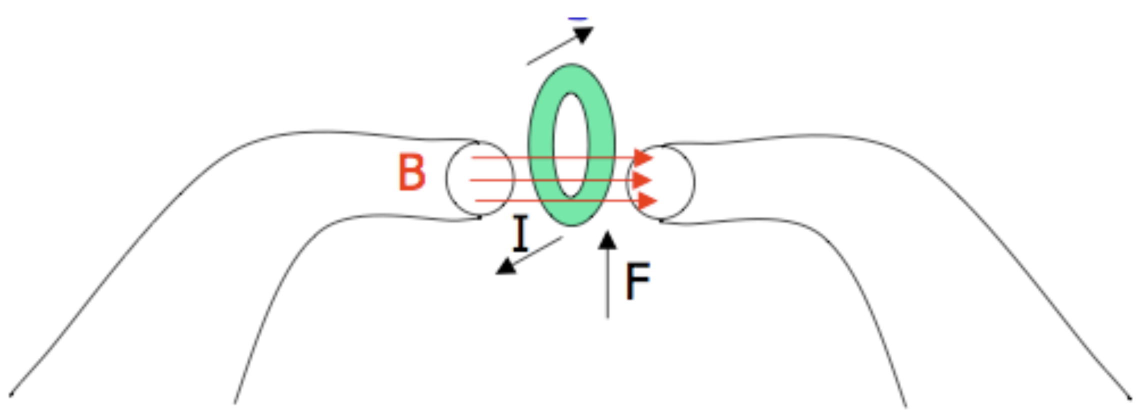
Create B with electromagnets

What happen if we drop a loop of conductor?

With and without B

What if we drop a full disk

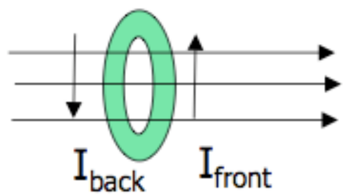
What if we drop a loop with a cut?



Explanation

Falling Loop:

B perpendicular to loop is limited in space ➙ flux changes during fall induced ➙loop will levitate (Eddie currents)



Falling Disk : Will slow down

Falling open ring : Will not levitate because loop is broken and current cannot flow

Faraday's law in differential form

Faraday's law in integral form: $emf = -\frac{1}{c}\frac{\partial \Phi_B}{\partial t}$

Left term (apply Stokes): $emf = \oint_C \vec{E} \cdot d\vec{l} = \int_S \left(\nabla \times \vec{E}\right) \cdot d\vec{a}$

Right term:

$$-\frac{1}{c}\frac{\partial \Phi_B}{\partial t} = -\frac{1}{c}\frac{\partial}{\partial t}\int_S \vec{B} \cdot d\vec{a} \Rightarrow \int_S \left(\nabla \times \vec{E} + \frac{1}{c}\frac{\partial \vec{B}}{\partial t}\right) \cdot d\vec{a} = 0$$

Since this is valid for any surface:

- curl E is no longer zero: bye bye electrostatics!
- Explicit link between E and B, as expected from relativity!

$$\nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}$$

Another step toward Maxwell's equations…

All the equations in differential form that we found so far:

$\nabla \cdot \vec{E} = 4\pi\rho$ ⇐ Relates E and charge density (ρ) - Gauss

$\nabla \cdot \vec{B} = 0$ ⇐ Magnetic field lines are closed

$\nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}$ ⇐ Change in B creates E - Faraday

$\nabla \times \vec{B} = \frac{4\pi}{c}\vec{J}$ ⇐ Relates B and its sources (J) - Ampere

Another step toward Maxwell's equations: one last missing ingredient…

Can you guess what?

- Hint #1: Symmetry can guide you…
- Hint #2: Take the divergence of Faraday's law and see what happens…
- BUT first Inductance, EMFs, Lenz and Faraday....... and then answer question!

Faraday's law in differential form

Faraday's law in integral form: $emf = -\frac{1}{c}\frac{\partial \Phi_B}{\partial t}$

Left term (apply Stokes): $emf = \oint_C \vec{E} \cdot d\vec{l} = \int_S \left(\nabla \times \vec{E}\right) \cdot d\vec{a}$

Right term:

$$-\frac{1}{c}\frac{\partial \Phi_B}{\partial t} = -\frac{1}{c}\frac{\partial}{\partial t}\int_S \vec{B} \cdot d\vec{a} \Rightarrow \int_S \left(\nabla \times \vec{E} + \frac{1}{c}\frac{\partial \vec{B}}{\partial t}\right) \cdot d\vec{a} = 0$$

Since this is valid for any surface:

$$\nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}$$

- curl E is no longer zero: bye bye electrostatics!
- Explicit link between E and B, as expected from relativity!

Another step toward Maxwell's equations…

All the equations in differential form that we found so far:

$\nabla \cdot \vec{E} = 4\pi\rho$ ⇐ Relates E and charge density (ρ) - Gauss

$\nabla \cdot \vec{B} = 0$ ⇐ Magnetic field lines are closed

$\nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}$ ⇐ Change in B creates E - Faraday

$\nabla \times \vec{B} = \frac{4\pi}{c}\vec{J}$ ⇐ Relates B and its sources (J) - Ampere

Another step toward Maxwell's equations: one last missing ingredient…

Can you guess what?

- Hint #1: Symmetry can guide you…
- Hint #2: Take the divergence of Faraday's law and see what happens…
- BUT first Inductance, EMFs, Lenz and Faraday....... and then answer question!

Electromagnetic Inductance

Cu pendulum in B field

A copper pendulum is oscillating

Application of Lenz's law

Turn on the magnetic field for the following 3 different situations:

Pendulum #1

B crosses area with cuts

No effect

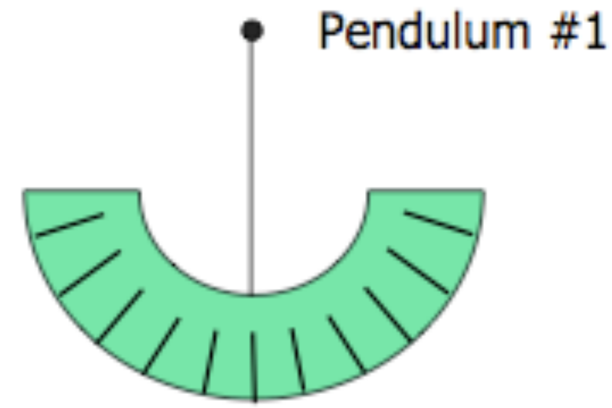
B crosses area above cuts

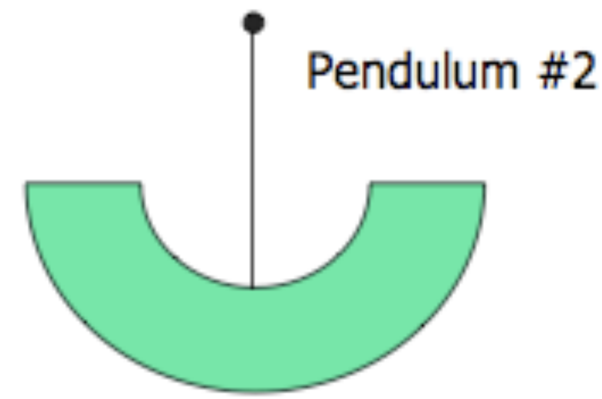
Stops slowly: Lenz's law

Pendulum #2

No cuts in Cu

Stops abruptly: Lenz's law





Three ways of creating emf

Faraday's law can be used to build generators:

$$emf = -\frac{1}{c}\frac{\partial}{\partial t}\int_S \vec{B}\cdot d\vec{S}$$

3 ways of creating emf:

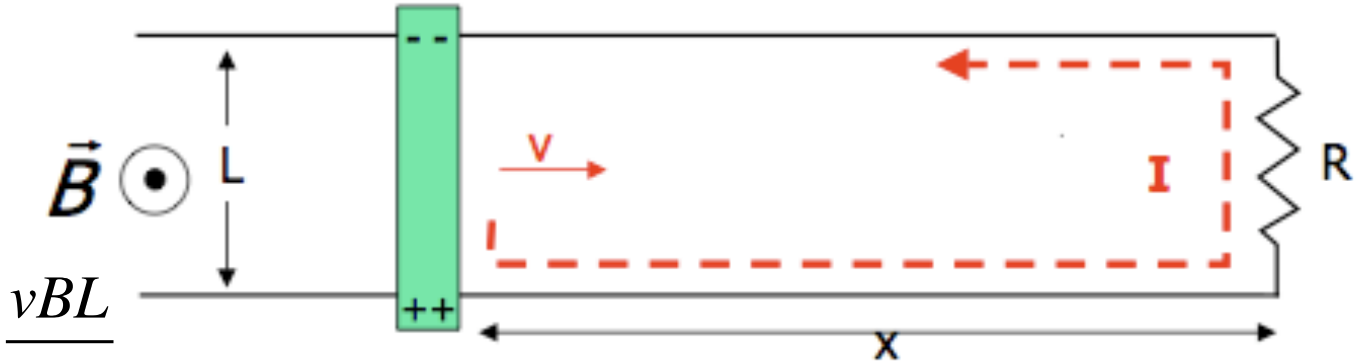
Vary the area: S=S(t)

Vary the angle between B and dS

Vary magnitude of B: B=B(t)

Changing the area

Sliding rod on rails:

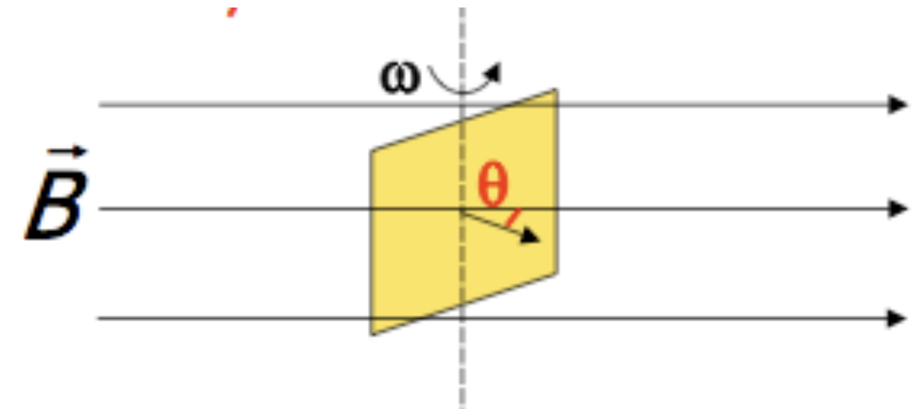


As derived earlier: $emf = \frac{vBL}{c}$

Because of Lenz's law, direction of current is counterclockwise to oppose the change of flux of B . . .

Changing angle between B and S

Constant B and loop rotating around its axis with angular velocity ω



If S is the area of the loop:

$$|emf| = \frac{1}{c}\frac{\partial}{\partial t}\left(BS\cos\omega t\right) = \frac{\omega}{c}BS\sin\omega t$$

This is an easy way to build an AC power generator

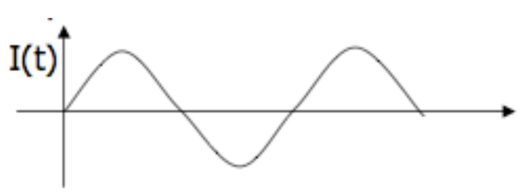
DC vs. AC current

DC current

Electrons flow all in the same direction at the same rate

AC current

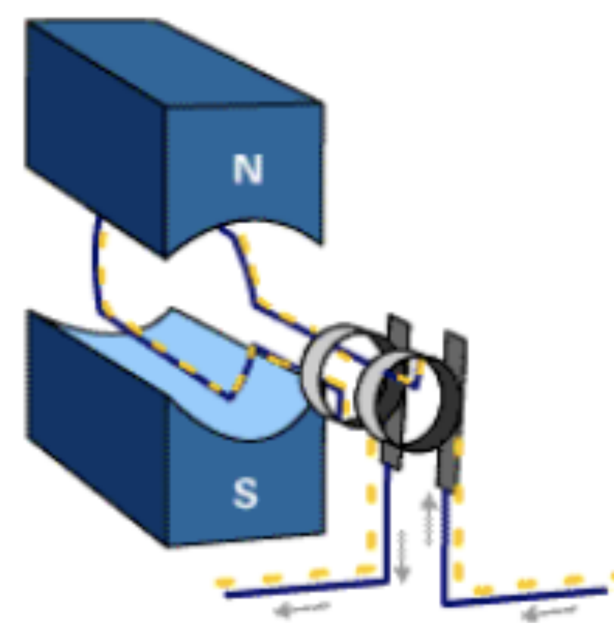
The flow of electron varies with time in amplitude and direction:



DC/AC generator

Uses DC to power electromagnet and induce AC on rotating loop

Why AC? Easier to step up and down for efficient transportation



Changing magnitude of B

Suppose you have a way to vary over time the magnitude of B: B = B(t)

Flux of B: $$\Phi_B = \int_S \vec{B}(t)\cdot d\vec{a} = B(t)S\cos\theta$$

Generated emf : $$|emf| = \frac{1}{c}\frac{\partial \Phi_B}{\partial t} = \frac{1}{c}S\frac{\partial B(t)}{\partial t}$$

How to create B = B(t)?

Loop of wire: $B \propto I$

If I = I(t) → B = B(t)

→AC in a solenoid will do the trick!

Induced emf – qualitative

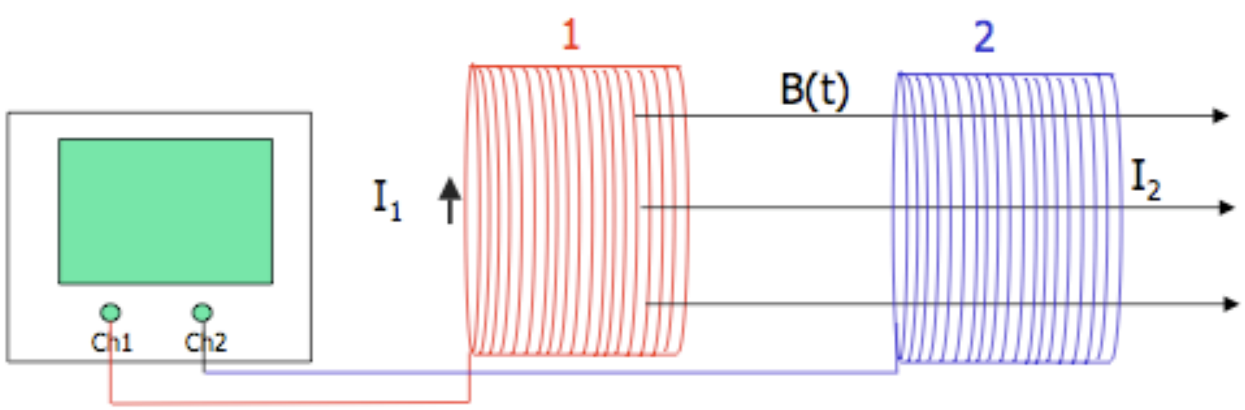
AC current circulating in a loop (solenoid 1) will induce current in a nearby loop (2)

Explanation:

I(1) in a loop creates B pointing out of the loop

Faraday: $$|emf(2)| = \frac{1}{c}\frac{\partial \Phi_2(B_1)}{\partial t}$$

What if we rotate 2 by 90 deg?



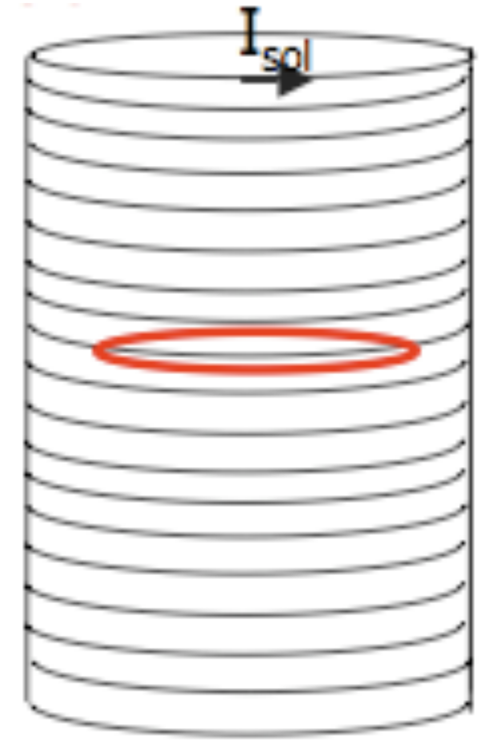
Induced emf – quantitative

Consider a loop of wire with radius r inside a long solenoid

Solenoid:

N = # of loops, l = total length ➙ n = N/l

$I_{sol} = I_{sol}(t)$



What is the emf generated in the loop?

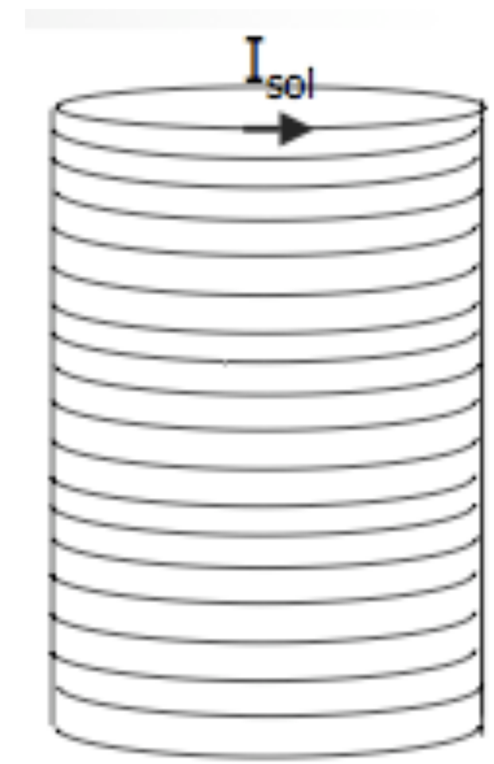
Find B inside solenoid: $$B_{sol} = \frac{4\pi n I_{sol}(t)}{c}$$

emf generated in loop:

$$|emf| = \frac{1}{c}\frac{\partial \Phi_B}{\partial t} = \frac{1}{c}\left(\pi r^2\right)\frac{\partial B(t)}{\partial t} = \frac{4\pi^2 n r^2}{c^2}\frac{\partial I_{sol}(t)}{\partial t}$$

➙The emf will depend on the geometry of the setup and on the rate of change of the I over time

Induced emf on solenoid itself



What if the “loop” is the solenoid itself?

Will any e.m.f. be created?

Remember Faraday’s law: $$emf = -\frac{1}{c}\frac{\partial}{\partial t}\int_S \vec{B}\cdot d\vec{S}$$

B inside solenoid: $$B_{sol} = \frac{4\pi n I_{sol}(t)}{c}$$

Flux of B through each loop: $$\Phi_B^{1\,loop} = BS_{1\,loop} = \frac{4\pi n I_{sol}(t)}{c}\pi R^2$$

Flux of B through N loops: $$\Phi_B^{Tot} = N\Phi_B^{1\,loop} = \frac{4\pi^2 R^2 N^2}{cl} I_{sol}(t)$$

Induced e.m.f. on solenoid: $$|emf| = \frac{4\pi^2 R^2 N^2}{c^2 l}\frac{\partial I_{sol}(t)}{\partial t}$$

Self Inductance L

Self-induced emf in the solenoid: $$|emf| = \frac{4\pi^2 R^2 N^2}{c^2 l}\frac{\partial I_{sol}(t)}{\partial t} \Rightarrow |emf| = L\frac{\partial I_{sol}(t)}{\partial t}$$

Let's examine this in detail:

- emf depends on change over time of current: dI/dt
- A bunch of constants depending on geometry called **self inductance L**

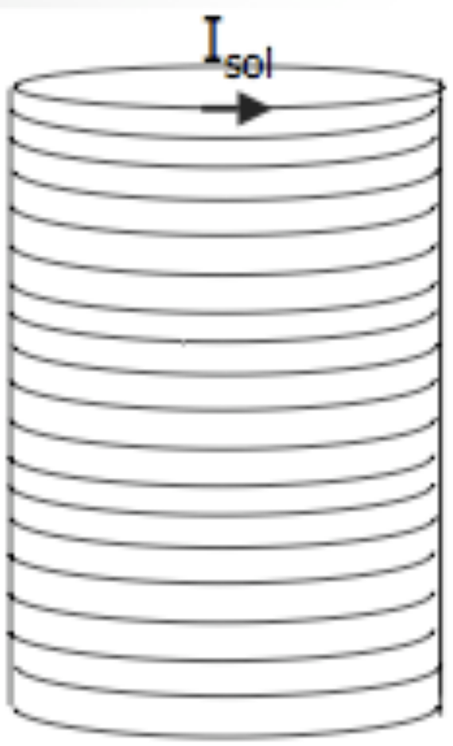
For a solenoid: $$L_{sol} = \frac{4\pi^2 R^2 N^2}{c^2 l}$$

Units:

cgs: $$[L] = \frac{[emf]}{[current]/[time]} = \frac{esu/cm}{(esu/\sec)/\sec} = \frac{\sec^2}{cm}$$

SI: $$[L] = \frac{[emf]}{[current]/[time]} = \frac{Volt}{Amp/\sec} = Henry(H)$$

Back emf



Magnitude of induced e.m.f. on solenoid: $$|emf| = \frac{4\pi^2 R^2 N^2}{c^2 l}\frac{\partial I_{sol}(t)}{\partial t}$$

How about the direction? And the effect?

Use Lenz's law to predict direction of induced current

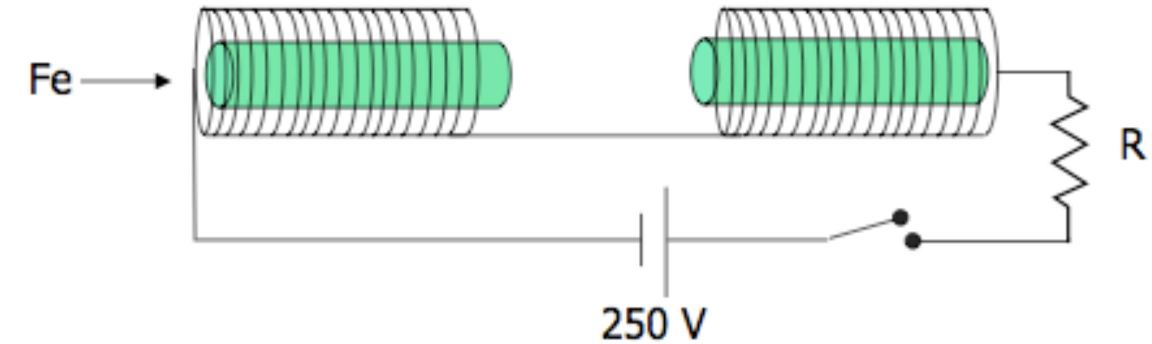
- If I_{sol} increases ➙ B increases ➙ flux increases
- I_{loop} will fight change ➙ opposite direction to I_{sol}
- If I_{sol} decreases ➙ B decreases ➙ flux decreases
- I_{loop} will fight change ➙ same direction as I_{sol}

Conclusion:

The inductance always opposes the change in the current

The e.m.f. created is called back emf as it acts back on the circuit trying to oppose changes

Example of back emf



Close switch: I flows (~80 A)
Open switch: big spark due to back emf

Energy stored in inductors

Consider an inductor L in which we start a current I flowing.
As soon as the current starts flowing, a back-emf tries to oppose this current.
Power needed to deal with the back-emf:

$$P = I \times emf = IL\frac{\partial I(t)}{\partial t}$$

Calculate work to increase the current from 0 ➙ I when t: 0 ➙ t

$$W = \int_{t=0}^{t} Pdt = \int_{t=0}^{t} LI\frac{\partial I}{\partial t}dt = L\int_{t=0}^{t} IdI = \frac{1}{2}LI^2$$

Energy stored in the inductor: $W = \frac{1}{2}LI^2$

How is energy stored in inductors?

We created a magnetic field where there was none: work necessary to create the magnetic field is the energy stored in the B itself
Same as energy stored in electric field of a capacitor
Not surprising: special relativity!

Energy density of magnetic field (solenoid example)
Energy stored in solenoid: $U_L = LI^2/2$
Self inductance of a solenoid: $L = 4\pi^2R^2N^2 / lc^2$
B created by solenoid: $B = 4\pi NI/lc$

$$\Rightarrow U_L = \frac{1}{2}LI^2 = \frac{1}{2}\frac{4\pi^2 R^2 N^2}{c^2 l}I^2 = \frac{1}{8\pi}\left(\pi R^2 l\right)\left(\frac{4\pi N}{cl}I\right)^2 = (volume)\frac{B^2}{8\pi}$$

Energy density of B: $U_B = \frac{B^2}{8\pi}$

Similar to energy density of the electric field: $U_E = \frac{E^2}{8\pi}$

How do we calculate L ?

Just some examples…

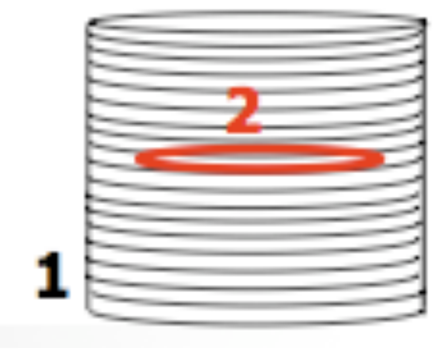
Strategy 1:

L is the proportionality constant between induced emf and variation over time of current: $|emf| = L\frac{\partial I(t)}{\partial t}$

Strategy 2:

Exploit the fact that energy stored in the magnetic field is theenergy stored in the inductor: $Energy\ stored\ in\ B = \int_V \frac{B^2}{8\pi}dV = \frac{1}{2}LI^2$

Mutual inductance



Back to the loop inside the solenoid

Label solenoid with 1 and loop with 2

e.m.f. induced on loop (ε_2) depends on dI/dt and a constant M_{21}

$$\varepsilon_2 = M_{21}\frac{\partial I_1}{\partial t}$$

where M_{21} is the coefficient of mutual inductance

For this particular configuration we already calculated that $M_{21} = \frac{4\pi^2 r^2 N}{c^2 l}$

Now do the opposite: run a current $I_2(t)$ in the loop and calculate emf induced on solenoid (ε_1): $\varepsilon_1 = M_{12}\frac{\partial I_2}{\partial t}$

How to calculate M_{12} ???

No need to calculate it! Reciprocity theorem ➙ $M_{12} = M_{21}$

Reciprocity theorem

Consider 2 loops of wire:

Current I runs through loop 1. What is Φ through loop 2 due to 1?

$$\Phi_{21}^{B} = \int_{S_2} \vec{B}_1 \cdot d\vec{a}_2$$

Now rewrite this result in terms of vector potential and use Stokes:

$$\Phi_{21} = \int_{S_2} \vec{B}_1 \cdot d\vec{a}_2 = \int_{S_2} \left(\nabla \times \vec{A}_1\right) \cdot d\vec{a}_2 = \int_{C_2} \vec{A}_1 \cdot d\vec{l}_2$$

Since $\vec{A}_1 = \frac{I}{c} \oint_{C_1} \frac{d\vec{l}_1}{r}$ we obtain $\Phi_{21} = \frac{I}{c} \int_{C_2} \int_{C_1} \frac{d\vec{l}_1}{r} \cdot d\vec{l}_2 = \Phi_{12}$

Same fluxes ➙ if currents are the same: $M_{21} = M_{12}$

Transformers

Devices to step up (or down) AC currents

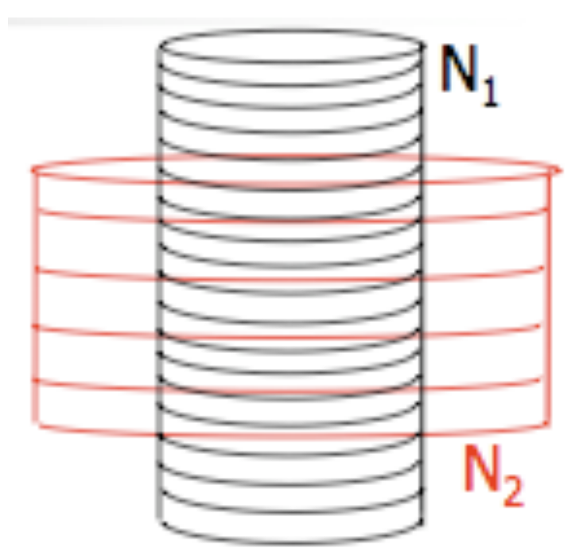
Practical application of mutual inductance

Simplest implementation:

Primary solenoid (black): N_1 turns

Secondary solenoid (red): N_2 turns

I(t) in the primary will induce a varying Φ_B through itself:



$$\varepsilon_1 = \frac{N_1}{c} \frac{d\Phi_B}{dt}$$

where Φ_B = magnetic flux through single turn

Flux is the same in second solenoid → induced emf is: $\varepsilon_2 = \frac{N_2}{c}\frac{d\Phi_B}{dt}$

Comparing: $\varepsilon_2 = \varepsilon_1 \frac{N_2}{N_1}$

Depending on number of turns we can
increase voltage ($N_2 > N_1$)
reduce the voltage ($N_2 < N_1$)

Called "stepping up" or "stepping down".

Long distance power transmission uses AC currents (not DC).

Voltage stepped up 480000 volts, sent down lines, stepped down to house in stages.

Transmission efficiency is improved by increasing the voltage, which reduces the current in the conductors while keeping the power transmitted nearly equal to the power input.

The reduced current flowing through the conductor reduces the losses in the conductor. Losses are proportional to I^2. Thus cutting current in half reduces losses by a factor of 4.

Maxwell's equations so far

$\nabla \cdot \vec{E} = 4\pi\rho$ ⇐ Relates E and charge density (ρ) - Gauss

$\nabla \cdot \vec{B} = 0$ ⇐ Magnetic field lines are closed

$\nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}$ ⇐ Change in B creates E - Faraday

$\nabla \times \vec{B} = \frac{4\pi}{c}\vec{J}$ ⇐ Relates B and its sources (J) - Ampere

Is this set of equations completely consistent?
Not quite…

To see inconsistency, take the divergence of Ampere's law

$$\nabla \cdot \left(\nabla \times \vec{B}\right) = \frac{4\pi}{c}\nabla \cdot \vec{J}$$

$$\nabla \cdot \vec{J} = -\frac{4\pi}{c}\frac{\partial \rho}{\partial t} \qquad \text{(continuity equation)}$$

$$\left\{\begin{array}{l} \oint_S \vec{J} \cdot d\vec{A} = \oint_V \nabla \cdot \vec{J} dV \\ -\frac{\partial Q_{inside}}{\partial t} = -\frac{\partial}{\partial t}\int_V \rho dV \end{array}\right\} \Rightarrow \int_V \left(\nabla \cdot \vec{J} + \frac{\partial \rho}{\partial t}\right) dV = 0 \Rightarrow \nabla \cdot \vec{J} + \frac{\partial \rho}{\partial t} = 0$$

$$\nabla \cdot \left(\nabla \times \vec{B}\right) = 0 \qquad \text{(ALWAYS)}$$

$$\Rightarrow \frac{\partial \rho}{\partial t} = 0$$

Ampere's law works only when $\partial \rho / \partial t = 0$ which works in most cases but not always:
➙Ampere's law is incomplete! It is inconsistent with the continuity equation in general!

Fixing the inconsistency

Since $\nabla \cdot \left(\nabla \times \vec{V} \right) = 0$ we can add some term to the right hand side such that its divergence will be identically 0 (won't change LHS of equation)

Generalized Ampere's law: $\nabla \times \vec{B} = \frac{4\pi}{c} \vec{J} + \vec{F}$ Assumption

What is F? We know that divergence of RHS must be = 0, i.e.,

$$\nabla \cdot \left(\nabla \times \vec{B} \right) = 0 = \nabla \cdot \left(\frac{4\pi}{c} \vec{J} + \vec{F} \right) \Rightarrow \nabla \cdot \left(c\vec{F} \right) = -4\pi \nabla \cdot \vec{J} = 4\pi \frac{\partial \rho}{\partial t}$$

$$\Rightarrow \nabla \cdot \left(c\vec{F} \right) = 4\pi \frac{\partial \rho}{\partial t}$$ Similar to Gauss's law!

Take time derivative of Gauss's law:

$$\frac{\partial}{\partial t} \left(\nabla \cdot \vec{E} \right) = 4\pi \frac{\partial \rho}{\partial t} \Rightarrow \frac{\partial}{\partial t} \left(\nabla \cdot \vec{E} \right) = \nabla \cdot \left(\frac{\partial \vec{E}}{\partial t} \right) = \nabla \cdot \left(c\vec{F} \right)$$

since time and space derivatives commute

$$\Rightarrow \vec{F} = \frac{1}{c} \frac{\partial \vec{E}}{\partial t}$$

Displacement currents

Generalized Ampere's equation $\nabla \times \vec{B} = \frac{4\pi}{c} \vec{J} + \frac{1}{c} \frac{\partial \vec{E}}{\partial t}$

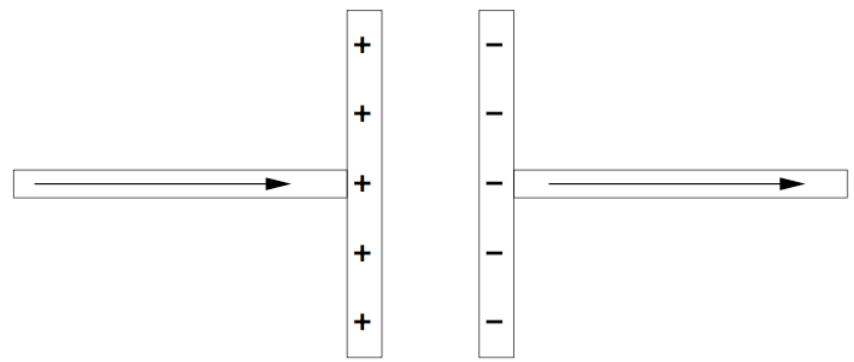
This can also be written as: $\nabla \times \vec{B} = \frac{4\pi}{c} \left(\vec{J} + \vec{J}_d \right)$

With J_d = **displacement** current (density): $\vec{J}_d = \frac{1}{4\pi} \frac{\partial \vec{E}}{\partial t}$

What is the J_d? (first proposed by Maxwell(not based on continuity but on symmetry!))
Not a real current:
does not describe charges flowing through some region
But it acts like a real current:
whenever have changing E field, can treat its effect as if due to as a real current J_d

What is a displacement current?

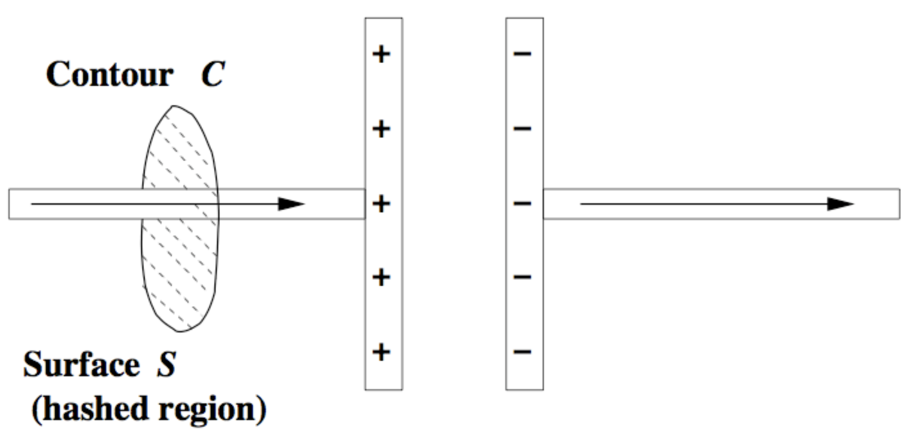
Consider a current flowing in a circuit and charging a capacitor C as shown below:

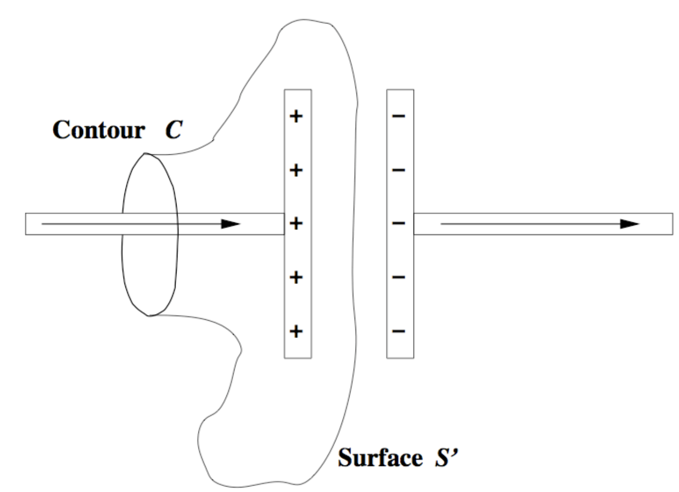


Standard integral Ampere's law:

$$\oint_C \vec{B} \cdot d\vec{l} = \frac{4\pi}{c} I_{enclosed} = \frac{4\pi}{c} \int_S \vec{J} \cdot d\vec{a}$$

Let's choose the path C and the surface S as in the drawing below(left): It all makes sense!





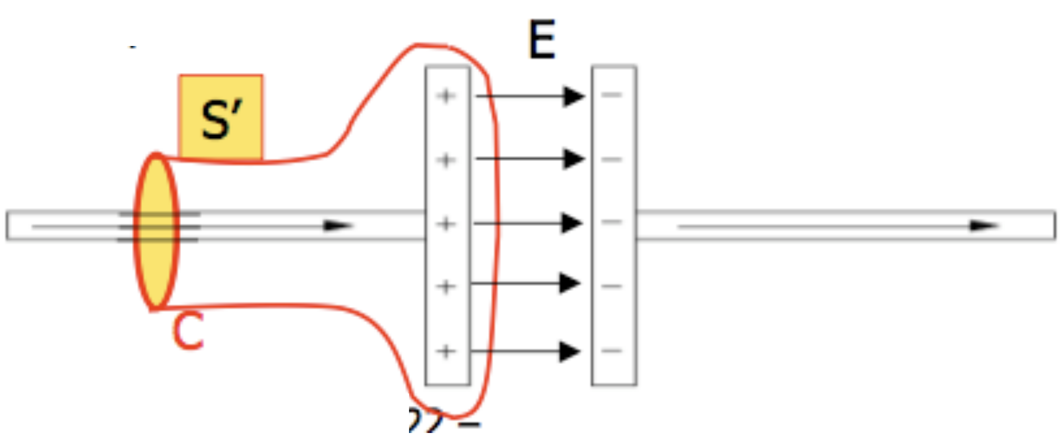
Now choose the same path C but the surface S' (right figure)
(OK by Stokes…) No standard current J through the surface (no charge crosses C!)
But there is a flux of displacement current J_d through the plates!

We can use the generalized Ampere's Law:

$$\oint_C \vec{B} \cdot d\vec{l} = \frac{4\pi}{c}\left(I_{enclosed} + I_d\right)$$

$$with \quad I_d = \int_{S'} \vec{J}_d \cdot d\vec{a} = \frac{1}{4\pi}\int_{S'} \frac{\partial \vec{E}}{\partial t} \cdot d\vec{a} = \frac{1}{4\pi}\frac{\partial}{\partial t}\int_{S'} \vec{E} \cdot d\vec{a} = \frac{1}{4\pi}\frac{\partial \Phi_E}{\partial t}$$

The displacement current is related to the change over time of the flux of the electric field. In the example above, the electric field E is the one produced in between the plates of the capacitor C



The electric field E:

Points in the same direction as the current (+x)

At a given instant in time: $\vec{E} = \frac{4\pi Q}{A}\hat{x}$

The flux of E will then be:

$$\Phi_E = 4\pi Q \quad \text{(yes, Gauss's law!)}$$

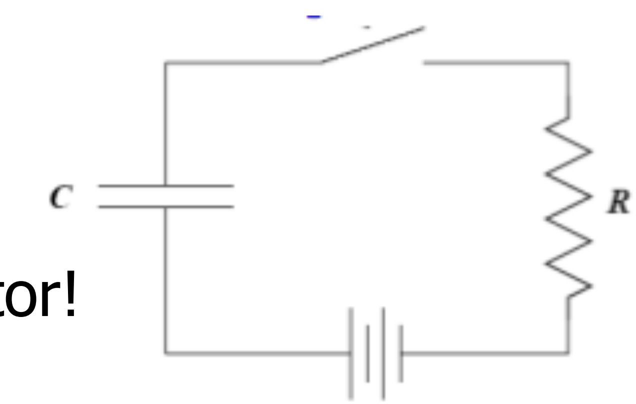
The rate of the change of this flux is: $\frac{\partial \Phi_E}{\partial t} = 4\pi \frac{\partial Q}{\partial t} = 4\pi I$

Where I is the current that is charging the capacitor

Comparing this with earlier results: $I_d = \int_S \vec{J} \cdot d\vec{a} = \int_{S'} \vec{J}_d \cdot d\vec{a} = I$

→ generalized Ampere’s Law is **valid** no matter what surface we use

The importance of displacement currents



When we examined the following circuit:

I said the same current I was flowing in each circuit element.

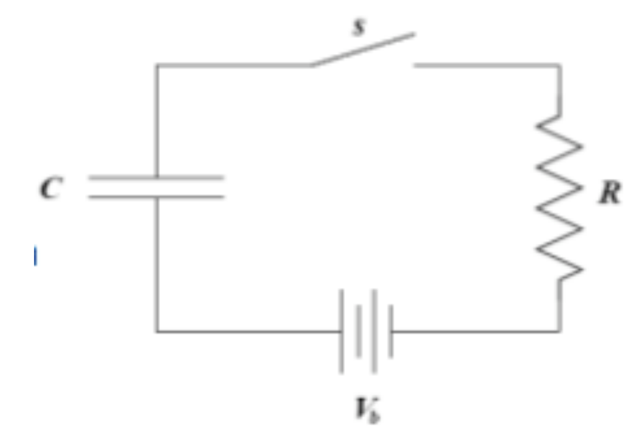
BUT NO ONE YELLED AT ME!!!!!!

How is it possible? No current flows through the plates of a capacitor!

Displacement currents fix this inconsistency!

Displacement current “continues” the “real” current across the capacitors ensuring the validity of Kirchoff’s laws.

Displacement current: application



Consider the following RC circuit:

As C charges up, I_d flows

I_d induces B inside the plates

Assuming cylindrical plates of radius a

Calculate B inside the plates

Find E(t): $E(t) = 4\pi\sigma = \dfrac{4\pi Q(t)}{\pi a^2}$

Displacement current density: $J_d = \dfrac{1}{4\pi}\dfrac{\partial E(t)}{\partial t} = \dfrac{1}{\pi a^2}\dfrac{\partial Q(t)}{\partial t} = \dfrac{I(t)}{\pi a^2}$

Remember that $I(t) = \dfrac{V_b}{R} e^{-t/RC}$

Magnetic field inside the plate (Ampere's law): $\oint_C \vec{B} \cdot d\vec{l} = \dfrac{4\pi}{c} \int_C \vec{J}_d \cdot d\vec{a} \Rightarrow B(r) = \dfrac{2rV_b}{ca^2 R} e^{-t/RC}$

Direction: RH Rule - Thumb in direction of J_d and fingers curl like B.

Maxwell's equations(complete!): differential form

$$\left.\begin{cases}\nabla\cdot\vec{E}=4\pi\rho \\ \nabla\cdot\vec{B}=0 \\ \nabla\times\vec{E}=-\dfrac{1}{c}\dfrac{\partial\vec{B}}{\partial t} \\ \nabla\times\vec{B}=\dfrac{4\pi}{c}\vec{J}+\dfrac{1}{c}\dfrac{\partial\vec{E}}{\partial t}\end{cases}\right\}cgs \qquad \left.\begin{cases}\nabla\cdot\vec{E}=\rho/\varepsilon_0 \\ \nabla\cdot\vec{B}=0 \\ \nabla\times\vec{E}=-\dfrac{\partial\vec{B}}{\partial t} \\ \nabla\times\vec{B}=\mu_0\vec{J}+\mu_0\varepsilon_0\dfrac{\partial\vec{E}}{\partial t}\end{cases}\right\}SI$$

The last equations are the generalized Ampere's law

Note: when Maxwell introduced the term dE/dt in the generalized Ampere's law, his arguments were based purely on symmetry

Yes, he was a theorist! ☺

Maxwell's equations: integral form

$$\Phi_E=\int_S\vec{E}\cdot d\vec{a}=4\pi Q_{enclosed}$$ (Gauss's law)

$$\Phi_B=\int_S\vec{B}\cdot d\vec{a}=0$$ (Magnetic field line are closed)

$$emf=\oint_C\vec{E}\cdot d\vec{l}=-\frac{1}{c}\frac{\partial\Phi_B}{\partial t}$$ (Faraday's law)

$$\oint_C\vec{B}\cdot d\vec{l}=\frac{4\pi}{c}\left(\vec{I}+\vec{I}_d\right)$$ (Generalized Ampere's law)

where the currents $\vec{I}$ *and* $\vec{I}_d$ are defined as $\vec{I}=\int_S\vec{J}\cdot d\vec{a}$ *and* $\vec{I}_d=\frac{1}{4\pi}\frac{\partial\Phi_E(S)}{\partial t}$

$$\Phi_E = \int_S \vec{E} \cdot d\vec{a} = 4\pi Q_{enclosed}$$

Electric flux through closed surface S = 4π time charge enclosed by S

$$\Phi_B = \int_S \vec{B} \cdot d\vec{a} = 0$$

Magnetic flux through closed surface S = 0

$$emf = \oint_C \vec{E} \cdot d\vec{l} = -\frac{1}{c}\frac{\partial \Phi_B}{\partial t}$$

EMF induced around closed contour C = -1/c times rate of change of magnetic flux through surface S bounded by C

$$\oint_C \vec{B} \cdot d\vec{l} = \frac{4\pi}{c}\left(\vec{I} + \vec{I}_d\right)$$

Line integral of magnetic field around closed contour C = 4π/c times the sum of the total current -- real + displacement -- pass through that contour.

3 good reasons to remember Maxwell's equations

They compactly and beautifully summarize all the E&M we learned so far!

You will see them on T shirts for the rest of your life around the world:
Better to get familiar with them ASAP!

On the first day of P112 next year you will need to remember them:
save your honor (and mine)

Maxwell equations in vacuum

What happens when we write Maxwell's equations in vacuum?
Vacuum: no sources, ρ=0 and J=0

$$\begin{Bmatrix} \nabla \cdot \vec{E} = 4\pi\rho \\ \nabla \cdot \vec{B} = 0 \\ \nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} \\ \nabla \times \vec{B} = \frac{4\pi}{c}\vec{J} + \frac{1}{c}\frac{\partial \vec{E}}{\partial t} \end{Bmatrix} \Rightarrow \begin{Bmatrix} \nabla \cdot \vec{E} = 0 \\ \nabla \cdot \vec{B} = 0 \\ \nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} \\ \nabla \times \vec{B} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t} \end{Bmatrix}$$

Except for a – sign, these equations are exquisitely symmetric(Maxwell saw this)!

Consequence: an electric field E varying in time will create a magnetic field B;
a B field varying in time creates an E field:

E and B are intimately related!

Maxwell equations in vacuum: solution

$$\begin{cases} \nabla \cdot \vec{E} = 0 & (1) \\ \nabla \cdot \vec{B} = 0 & (2) \\ \nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} & (3) \\ \nabla \times \vec{B} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t} & (4) \end{cases}$$

How to solve these equations?
- Uncouple them!
  - Separate E and B in equations

How?
- Take the curl of equations (3) and (4)
- Use other equations((1) and (4)) as needed

Start from (3):

$$\nabla \times \left(\nabla \times \vec{E}\right) = -\frac{1}{c}\nabla \times \frac{\partial \vec{B}}{\partial t}$$

$$\nabla\left(\nabla \cdot \vec{E}\right) - \nabla^2 \vec{E} = -\frac{1}{c}\frac{\partial \nabla \times \vec{B}}{\partial t} = -\frac{1}{c^2}\frac{\partial}{\partial t}\frac{\partial \vec{E}}{\partial t}$$

$$\nabla^2 \vec{E} = \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2}$$

Now repeat the procedure starting from (4)

$$\nabla \times \left(\nabla \times \vec{B}\right) = \frac{1}{c}\nabla \times \frac{\partial \vec{E}}{\partial t}$$

$$\nabla\left(\nabla \cdot \vec{B}\right) - \nabla^2 \vec{B} = -\frac{1}{c}\frac{\partial \nabla \times \vec{E}}{\partial t} = -\frac{1}{c^2}\frac{\partial}{\partial t}\frac{\partial \vec{B}}{\partial t}$$

$$\nabla^2 \vec{B} = \frac{1}{c^2}\frac{\partial^2 \vec{B}}{\partial t^2}$$

These results are special cases of a known equation: the **wave equation**

Wave Equations: For $\nabla^2 f = \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2}$ Solution = any function f that has well-behaved derivatives and arguments of function have specified form $f = f(x \pm vt)$

Note: we are restricting ourselves to the 1D case; extension to 3D later

Solution of wave equation: proof

Prove that $f = f(x \pm vt)$ is a solution of the wave equation

Just calculate time and space derivatives.

Keep in mind that $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$

Define: u = x ± vt

$$\frac{\partial f(x \pm vt)}{\partial t} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial t} = \pm v\frac{\partial f}{\partial u} \Rightarrow \frac{\partial^2 f(x \pm vt)}{\partial t^2} = v^2\frac{\partial^2 f}{\partial u^2}$$

$$\frac{\partial f(x \pm vt)}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} = \frac{\partial f}{\partial u} \Rightarrow \frac{\partial^2 f(x \pm vt)}{\partial x^2} = \frac{\partial^2 f}{\partial u^2}$$

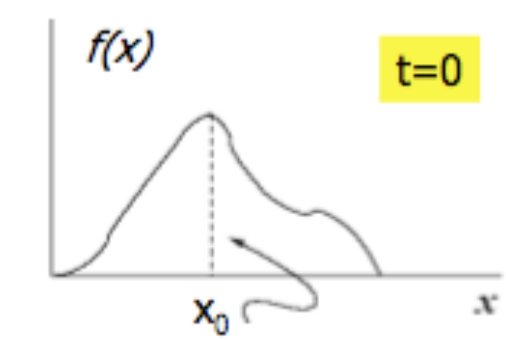
Plug the above results into the equation $\frac{\partial^2 f}{\partial u^2} = \frac{1}{v^2}v^2\frac{\partial^2 f}{\partial u^2} \Rightarrow identity$

As we wanted to prove!

Wave equation solution

What is a function such as f = f(x+vt)?

Assume v=1 cm/s

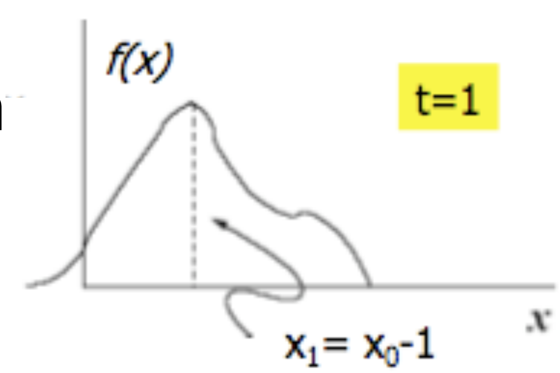


At time t=0: Position of the max: x_0

At time t=1s: Peak still occurs when the argument of f is x_0

But since the time is not 0, function will be shifted in x by vt=1cm

Position of the max: $x_1 = x_0 - 1$. Thus,



f = f(x+vt) represents a wave traveling in -x direction with velocity v

EM waves ⇔ Wave equation: $$\nabla^2 f = \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2}$$

Solution: f = f(x±vt)

Any function of argument x±vt

These solution represent waves traveling with velocity v

x-vt represents a wave traveling in the +x direction

x+vt represents a wave traveling in the –x direction

Maxwell's equation: $$\nabla^2 \vec{f} = \frac{1}{c^2}\frac{\partial^2 \vec{f}}{\partial t^2}$$

Same equation! Only difference: v=c

Solution: EM waves traveling with speed of light

➙ light IS an EM wave!!!

EM waves in SI

This same result looks much more interesting in SI.

Maxwell's equations in SI: $$\nabla^2 \vec{f} = \mu_0 \varepsilon_0 \frac{\partial^2 \vec{f}}{\partial t^2}$$

where ε_0 is the permittivity of free space and μ_0 is the permeability of free space

Maxwell's equations tell us what the velocity of an EM wave is: $$v = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$$

ε_0 and μ_0 can be measured➙we can predict velocity of EM waves:

$$\mu_0 = 4\pi \times 10^{-7} C^2 N^{-1} m^2 \quad , \quad \varepsilon_0 = 8.85418 \times 10^{-12} N \sec^2 C^{-2}$$

➙v=2.998x10^8 m/s^2 which is the speed of light!

Maxwell was the first to realize that E&M equations were leading to a wave equation that was propagating at the speed of light: light is an EM wave!

Plane waves: a particularly important and useful class of waves

Definition of plane waves in the most general form:

$$\vec{E} = \vec{E}_0 \sin\left(\vec{k}\cdot\vec{r} - \omega t\right) = \vec{E}_0 \sin\left(k_x x + k_y y + k_z z - \omega t\right)$$

$$\vec{B} = \vec{B}_0 \sin\left(\vec{k}\cdot\vec{r} - \omega t\right) = \vec{B}_0 \sin\left(k_x x + k_y y + k_z z - \omega t\right)$$

$$with \quad \vec{k} = wavevector; \ \left|\vec{k}\right| = wavenumber; \ \hat{k} = propagation\ direction$$

Example: k || x ➙ $\vec{E} = \vec{E}_0 \sin\left(k_x x - \omega t\right)$

Fourier Theorem(Physics 50- fseries.m):

Any periodic function can be expressed as a linear combination of sin and cos functions

➙ sin and cos are the building blocks of all waves!

This is what makes plane waves so important! We can build up all waves as superpositions of plane waves.

Plane waves vs f(x-ct)

Goal: Prove that plane waves satisfy wave equation

We proved that f(x±ct) satisfies the wave equation

How to connect x±ct to the the argument of plane waves $\vec{k}\cdot\vec{r} \pm \omega t$?

$$\sin\left(\vec{k}\cdot\vec{r} - \omega t\right) = \sin\left[k\left(\hat{k}\cdot\vec{r} - \frac{\omega}{k}t\right)\right] = f\left(\hat{k}\cdot\vec{r} - \frac{\omega}{k}t\right)$$

When $\hat{k} \parallel \hat{x} : f\left(x - \frac{\omega}{k}t\right) \Rightarrow \omega = ck$

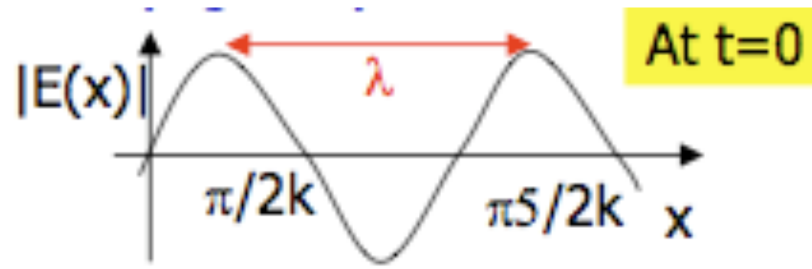
More on k and ω

Choose a system of coordinates so that our wave vector k is oriented || to x axis:
➙plane wave solution for E is $\vec{E} = \vec{E}_0 \sin(k_x x - \omega t)$

Let's consider only the spatial variation of the wave (e.g. t=0):

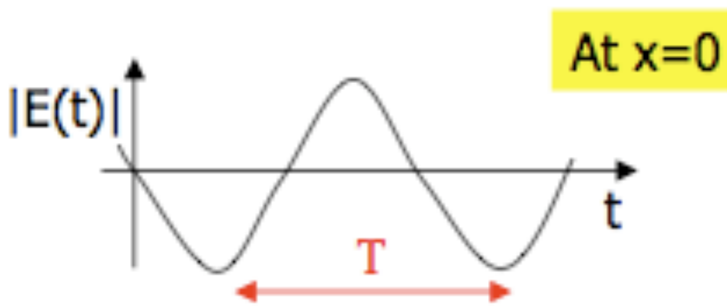
$$\vec{E} = \vec{E}_0 \sin(k_x x)$$

λ = wavelength = 2π/k



Let's now consider the time variation of the wave (e.g. x=0):

$$\vec{E} = -\vec{E}_0 \sin(\omega t)$$



Relationships between variables:

$$\omega = \frac{2\pi}{T} = 2\pi\nu \quad , \quad \omega = ck \quad , \quad \lambda\nu = c$$

Do plane wave satisfy Maxwell's equations?

EM waves are a consequence of Maxwell's equations in the sense that we used the 4 Maxwell's Equations to derive the wave equations for E and B:

$$\begin{Bmatrix} \nabla^2 \vec{E} = \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} \\ \nabla^2 \vec{B} = \frac{1}{c^2}\frac{\partial^2 \vec{B}}{\partial t^2} \end{Bmatrix} \Rightarrow \begin{Bmatrix} \vec{E} = \vec{E}_0 \sin(\vec{k}\cdot\vec{r} - \omega t) \\ \vec{B} = \vec{B}_0 \sin(\vec{k}\cdot\vec{r} - \omega t) \end{Bmatrix}$$

Does the solution of the EM wave equation satisfy all Maxwell's Equations?
Not necessarily! Let's start with Gauss's law: $\nabla \cdot \vec{E} = 0$

$$\nabla \cdot \vec{E} = \nabla \cdot \left(\vec{E}_0 \sin\left(k_x x + k_y y + k_z z - \omega t \right) \right)$$

$$= \left(E_{0x} k_x + E_{0y} k_y + E_{0z} k_z \right) \cos\left(k_x x + k_y y + k_z z - \omega t \right)$$

$$= \vec{k} \cdot \vec{E}_0 \cos\left(\vec{k} \cdot \vec{r} - \omega t \right)$$

$$\Rightarrow \nabla \cdot \vec{E} = 0 \quad when \quad \vec{k} \cdot \vec{E}_0 = 0 \Rightarrow \vec{E} \perp \vec{k} = \text{wave's direction of propagation}$$

More constraints on plane waves

Constraints following from $\nabla \cdot \vec{B} = 0$

$$\nabla \cdot \vec{B} = \vec{k} \cdot \vec{B}_0 \cos\left(\vec{k} \cdot \vec{r} - \omega t \right) = 0 \Rightarrow \vec{k} \cdot \vec{B}_0 = 0$$

$$\Rightarrow \vec{B} \perp \vec{k} = \text{wave's direction of propagation}$$

Divergence equations imply radiation's E and B fields are orthogonal to direction of propagation of plane wave - lie in a plane

Constraints following from curl equations $\nabla \times \vec{E} = -\frac{1}{c} \frac{\partial \vec{B}}{\partial t}$

Time derivative does not change direction of B → $\vec{E} \perp \vec{B}$

Same conclusion follows from: $\nabla \times \vec{B} = \frac{1}{c} \frac{\partial \vec{E}}{\partial t}$

Conclusion: $\vec{k} \perp \vec{E} \perp \vec{B} \perp \vec{k}$

Let's now calculate:

$$\frac{1}{c} \frac{\partial \vec{E}}{\partial t} = \frac{1}{c} \frac{\partial \left(\vec{E}_0 \cos\left(\vec{k} \cdot \vec{r} - \omega t \right) \right)}{\partial t} = \frac{\omega}{c} \vec{E}_0 \sin\left(\vec{k} \cdot \vec{r} - \omega t \right) = k \vec{E}_0 \sin\left(\vec{k} \cdot \vec{r} - \omega t \right)$$

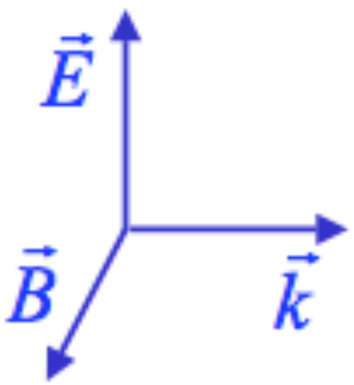
$$\nabla \times \vec{B} = \nabla \times \vec{B}_0 \cos\left(k_x x + k_y y + k_z z - \omega t \right)$$

Using $\nabla \times (\vec{v}s) = (\nabla \times \vec{v})s + (\nabla s) \times \vec{v}$ and the fact that B_0 = constant

$$\nabla \times \vec{B} = \nabla \times \vec{B}_0 \cos\left(k_x x + k_y y + k_z z - \omega t\right)$$
$$= \left(\nabla \times \vec{B}_0\right)\cos\left(\vec{k}\cdot\vec{r} - \omega t\right) + \left(\nabla \cos\left(k_x x + k_y y + k_z z - \omega t\right)\right) \times \vec{B}_0$$
$$= -\left(k_x \hat{x} + k_y \hat{y} + k_z \hat{z}\right)\sin\left(k_x x + k_y y + k_z z - \omega t\right) \times \vec{B}_0$$
$$= -\left(\vec{k} \times \vec{B}_0\right)\sin\left(\vec{k}\cdot\vec{r} - \omega t\right)$$

Therefore,

$\Rightarrow \vec{E}_0 = -\hat{k} \times \vec{B}_0 \Rightarrow \vec{k}, \vec{E}$ *and* $\vec{B}$ are a right-handed coordinate system



Important consequences:

In cgs, E and B have the same magnitude

$$\left|\vec{E}_0\right| = \left|\hat{k} \times \vec{B}_0\right| = \left|\vec{B}_0\right| \Rightarrow E_0 = B_0$$

$\vec{E}_0 \times \vec{B}_0$ is parallel to the direction of propagation of the wave

$$\vec{E}_0 = -\hat{k} \times \vec{B}_0 \Rightarrow \vec{E}_0 \times \vec{B}_0 = -\hat{k} \times \vec{B}_0 \times \vec{B}_0 = -\vec{B}_0 \underbrace{\left(\hat{k}\cdot\vec{B}_0\right)}_{=0} + \underbrace{\left(\vec{B}_0 \cdot \vec{B}_0\right)}_{=\left|\vec{B}_0\right|^2 = \left|\vec{E}_0\right|^2}\hat{k}$$

$$\Rightarrow \vec{E}_0 \times \vec{B}_0 = \left|\vec{E}_0\right|^2 \hat{k}$$ parallel to k

This says that E x B has an important physical meaning as we will soon see.

Summary: radiation so far

source free Maxwell equations:

$$\vec{\nabla}\cdot\vec{E} = 0 \qquad \vec{\nabla}\times\vec{E} = -\frac{1}{c}\frac{\partial\vec{B}}{\partial t}$$

$$\vec{\nabla}\cdot\vec{B} = 0 \qquad \vec{\nabla}\times\vec{B} = \frac{1}{c}\frac{\partial\vec{E}}{\partial t}\ .$$

can be rewritten as wave equation for $\vec{E}$ and $\vec{B}$

$$\frac{\partial^2\vec{E}}{\partial t^2} - c^2\nabla^2\vec{E} = 0$$

$$\frac{\partial^2\vec{B}}{\partial t^2} - c^2\nabla^2\vec{B} = 0\ .$$

A particularly instructive solution to the wave equations are the *plane wave* forms:

$$\vec{E}(\vec{r},t) = \vec{E}_0\sin(\vec{k}\cdot\vec{r} - \omega t)$$

$$\vec{B}(\vec{r},t) = \vec{B}_0\sin(\vec{k}\cdot\vec{r} - \omega t)\ .$$

This solution represents an electromagnetic wave propagating in the $\hat{k} = \vec{k}/k$ direction (where $k = \sqrt{\vec{k}\cdot\vec{k}} = \sqrt{k_x^2 + k_y^2 + k_z^2}$).

By considering how the wave behaves at some fixed time, we learned that k is simply related to the wavelength λ: $k = 2\pi/\lambda$.

The requirement that this solution satisfy the wave equation tells us that

$$\omega = ck\ .$$

From the definition $\omega = 2\pi\nu$ (angular frequency is 2π radians times "regular" frequency), we then obtain

$$\lambda\nu = c\ .$$

Finally, requiring that the plane wave solution satisfy all of Maxwell's equations leads to some important constraints on the vector amplitudes $\vec{E}_0$ and $\vec{B}_0$. These constraints are:

- The amplitudes are orthogonal to the propagation direction: $\hat{k} \cdot \vec{E}_0 = 0$, $\hat{k} \cdot \vec{B}_0 = 0$.
- The amplitudes are orthogonal to each other $\vec{E}_0 \cdot \vec{B}_0 = 0$.
- The amplitudes have the same magnitude: $|\vec{E}_0| = |\vec{B}_0|$.
- The propagation direction is parallel to $\vec{E} \times \vec{B}$.

These are important and rather constraining conditions. Nonetheless, they leave us with a great deal of freedom in the amplitudes. This freedom is described in terms of the radiation's *polarization state.*

Polarization of EM waves

Did we use all of our freedom in choosing the waves?

No, we can still choose a property called the "polarization state"

Linear polarization:

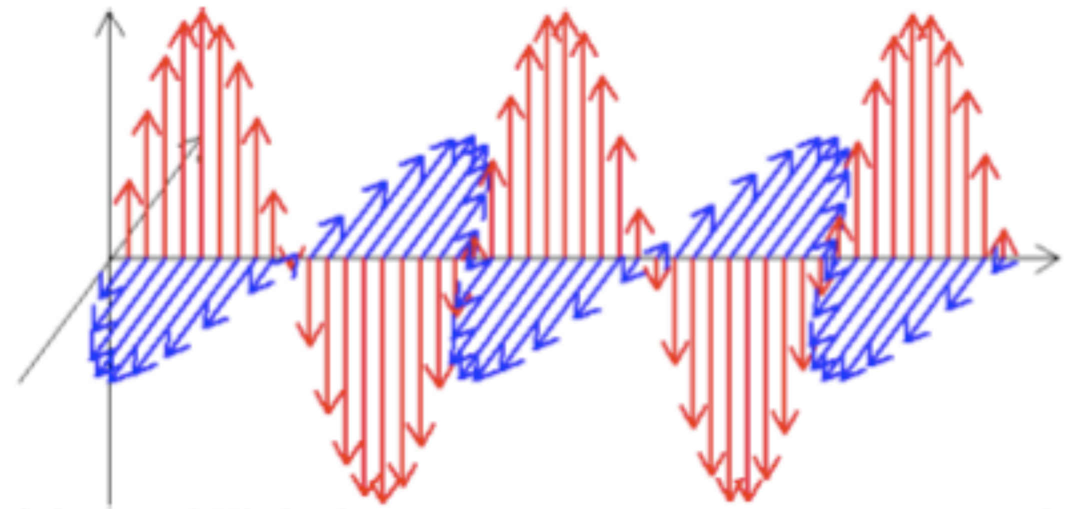
Consider a plane wave propagating in the x direction
Choose the coordinate system so that at t=0 E || y and B || z

If the directions of E_0 and B_0 are constant in time, the wave is "linearly polarized"

$$\vec{E} = E_0 \cos(kx - \omega t)\hat{y}$$

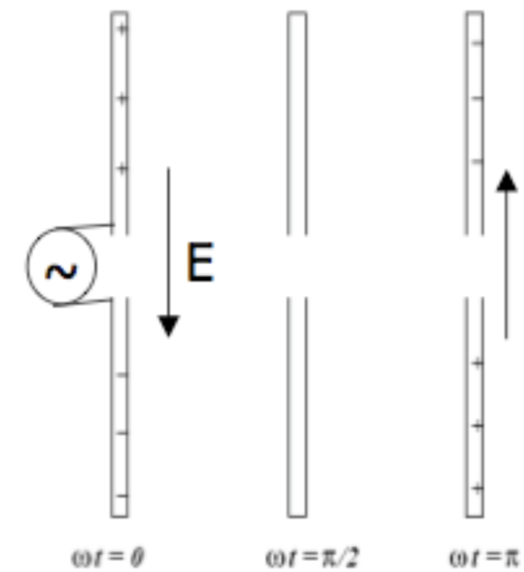
$$\vec{B} = B_0 \cos(kx - \omega t)\hat{z}$$

Note: direction of polarization
= direction of electric field
(a definition)

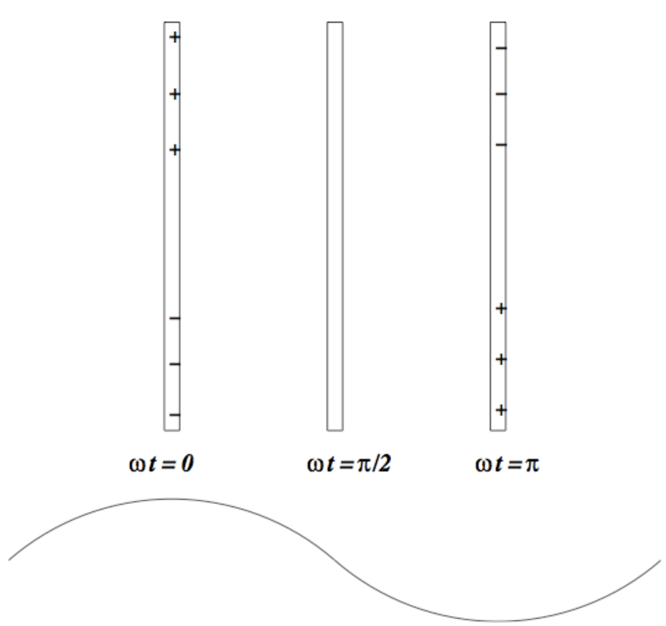


Linear Polarization of EM waves(all E or B vectors || to single fixed line)

How to produce linearly polarized waves?
Oscillating charge distribution in a conductor
= Broadcasting antenna
How to produce such a charge?
Long conductor driven by oscillating current
Field generated (in phase with current) shown below



t --->



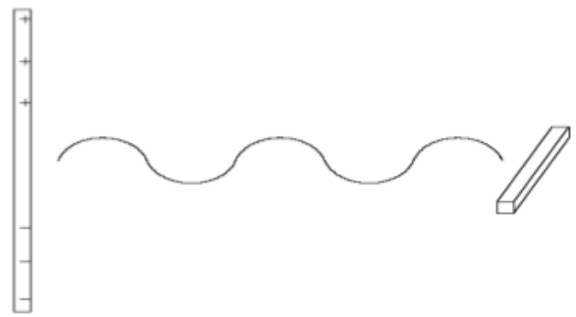
Examples:

(1) $\vec{E} = E_0 \hat{x} \sin(ky - \omega t) \Rightarrow \vec{B} = -E_0 \hat{z} \sin(ky - \omega t)$

(2) $\vec{E} = \left(E_0 \hat{x} + 2E_0 \hat{y}\right) \sin(kz - \omega t)$

How do we receive or measure the signal?

Receiving antenna (RCL circuit)~ E



Oscillating field in signal oscillates charges in receiver antenna --> current --> measure or use as in radio

For linear polarization:
When receiving antenna is || to broadcasting one, good reception
When receiver is perpendicular to broadcasting antenna: no reception because there is not enough room for charges to oscillate

Polaroids

- Sheet of plastic embedded with organic molecules extended in one direction
  - They can carry current in that particular direction: behave like antennas!
- When linearly polarized light hits the polaroid:
  - If E is aligned with orientation of molecules:
    - Charges move→current is generated→plastic heats: light stopped
- If E is perpendicular to orientation of molecules (“preferred direction”):
  - Charges will not be able to move in that direction: light goes through
- Conclusion:
  - Polaroids are transparent to light polarized || to their preferred direction and opaque to light polarized in the direction perpendicular to their preferred direction

Polaroids and polarization direction

- What happens when the light is polarized in a direction in between the preferred direction and its perpendicular?
- Example: light polarized along x axis; polaroid oriented at angle θ

$$\vec{E} = E_0 \cos(kz - \omega t)\hat{x}$$

$$\hat{p} = \hat{x}\cos\theta + \hat{y}\sin\theta$$

Light will go through partially

Since E has a component || to preferred direction of polaroid

E coming out is "overlap" between incoming E and polaroid's orientation

$$\left|\vec{E}_{out}\right| = \vec{E} \cdot \hat{p} = E_0 \cos\theta \cos(kz - \omega t)$$

$$\vec{E}_{out} = \left|\vec{E}_{out}\right| \hat{p} \quad \text{(parallel to polaroid's orientation)}$$

Conclusion:

Polaroids reduce the amplitude of linearly polarized light by cosθ(angle between E and polaroid's orientation) and rotate the orientation of E by θ

Polarization of random light

Light from a bulb, sunlight, etc is not polarized

Superposition of many plane waves, each with its own polarization

$$\vec{E}_{random} = \sum_i E_0 \left(\hat{x}\cos\theta_i + \hat{y}\sin\theta_i\right)\cos(kz - \omega t)$$

When light passes through a polaroid becomes linearly polarized

If polaroid is oriented || x axis:

$$\vec{E}_{out} = \sum_i E_0 \left(\hat{x}\cos\theta_i\right)\cos(kz - \omega t) = E_0 \hat{x}\cos(kz - \omega t)\sum_i \cos\theta_i$$

Conclusion:

Polaroids can be used to produce linearly polarized light

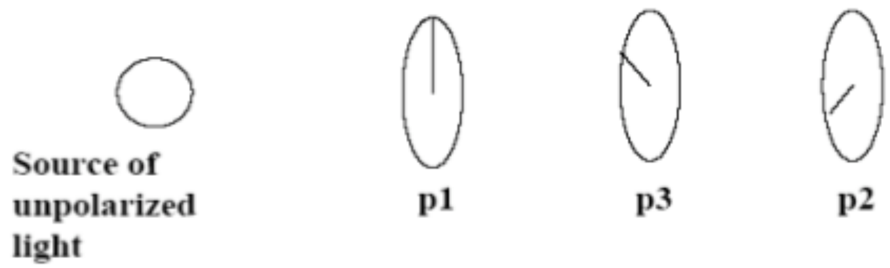
The intensity of the light will be reduced

Demo: 3 vs 2 polaroids

2 polaroids with orthogonal preferred direction will block light

First polaroid (P1) polarizes light in the direction x (for example)

Second polaroid (P2) oriented in the y direction, but E is now just || x



Now place a third polaroid P3 in between P1 and P2 (at 45°)

P1 will polarize light || x

P3 will select only component || to its preferred direction and rotate direction of polarization by 45°. $E_0' = E_0 \cos 45^o$

P2 will select component y direction that now is not 0 anymore.

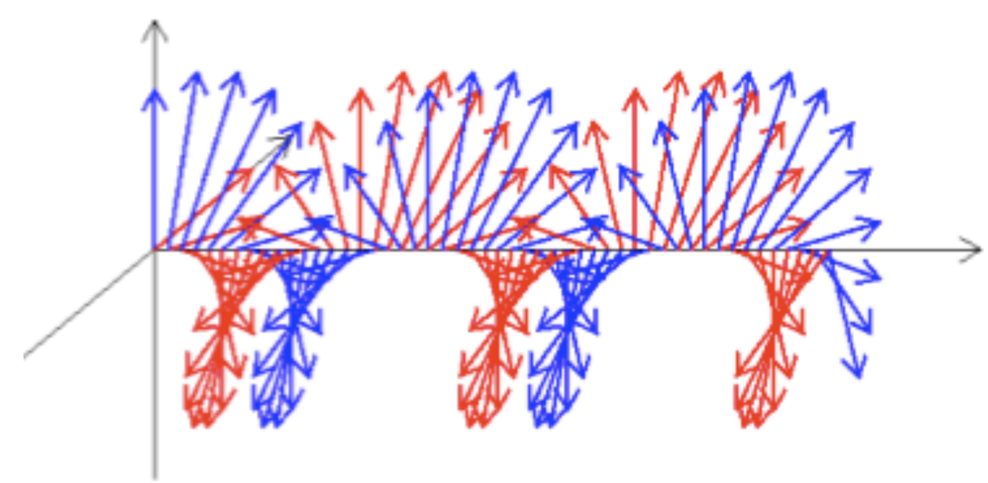
Intensity further reduced, but not 0! $E_0' = E_0(\cos 45^o)^2 = E_0/2$

Circular polarization

Consider a wave with the following form:

$$\vec{E} = E_0 \hat{x} \sin(kz - \omega t) + E_0 \hat{y} \cos(kz - \omega t)$$

$$\vec{B} = B_0 \hat{x} \sin(kz - \omega t) + B_0 \hat{y} \cos(kz - \omega t)$$



What is it?

Easier to understand if we look at z=0

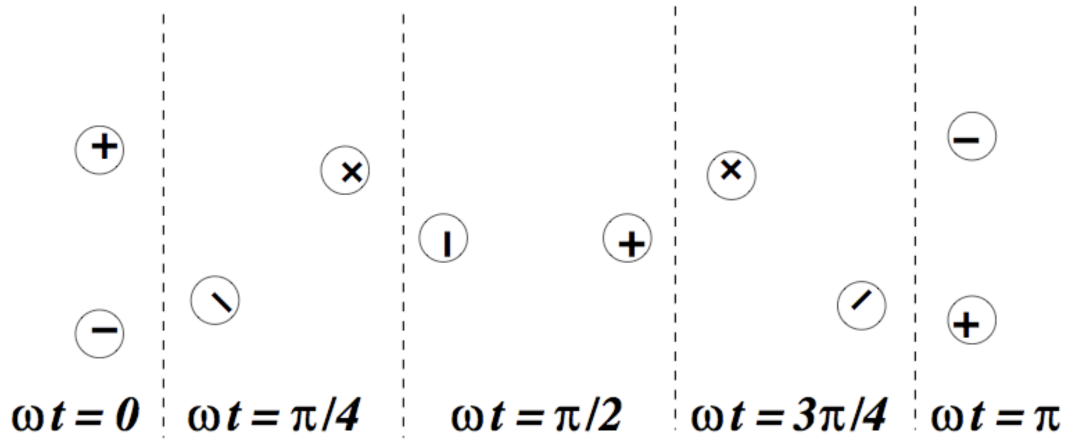
$$\vec{E} = -E_0 \hat{x} \sin(\omega t) + E_0 \hat{y} \cos(\omega t)$$

$$\vec{B} = -B_0 \hat{x} \sin(\omega t) + B_0 \hat{y} \cos(\omega t)$$

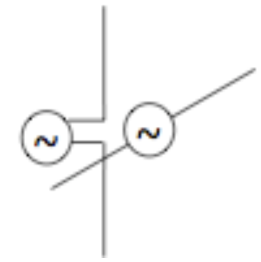
Electric and magnetic fields rotate at frequency ω

Circular polarization because E and B vector tips describe circles over time

How to produce it? Rotating dipole



produces circularly polarized light propagating out of board



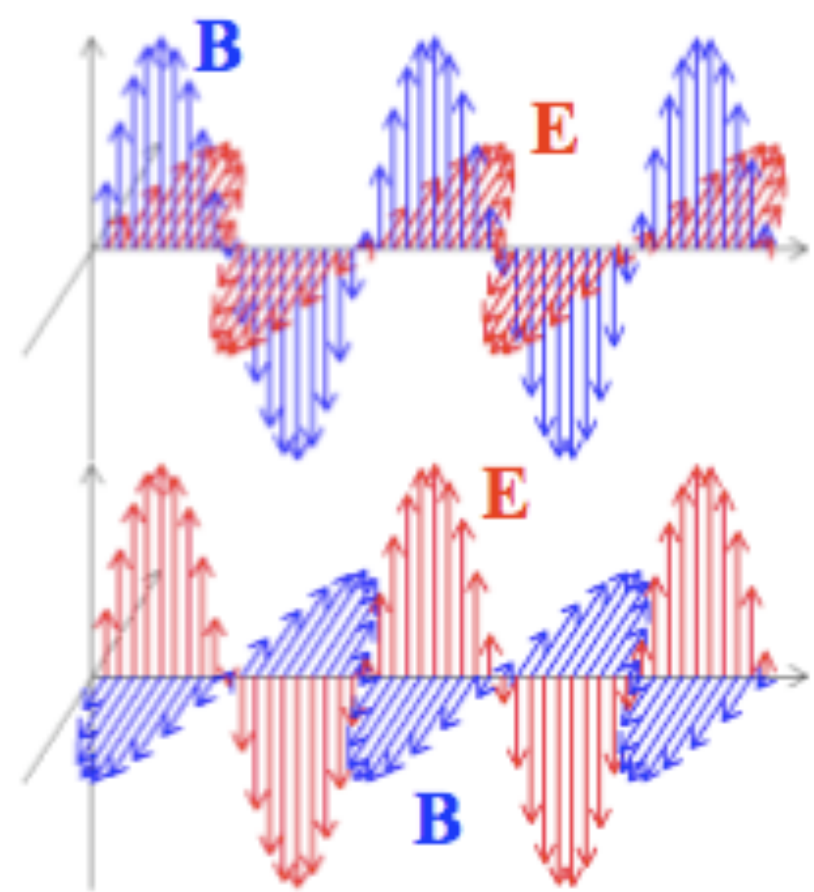
or 2 antennas at 90° driven by currents out of phase by 90°

Note: circular polarization does exist naturally in nature; in fact nature likes to produce it!

Elliptical Polarization

For a given k-vector, there are 2 independent solutions for the plane waves, e.g. 2 possible directions of the E-vector

$$\vec{E}_1 = E_0 \hat{x} \sin(kz - \omega t + \phi_1)$$
$$\vec{E}_2 = E_0 \hat{y} \cos(kz - \omega t + \phi_2)$$



All other solutions are just linear combinations of these

$\phi_1 = \phi_2$: linear polarization

$\phi_1 = \phi_2 + 90^o$: circular polarization

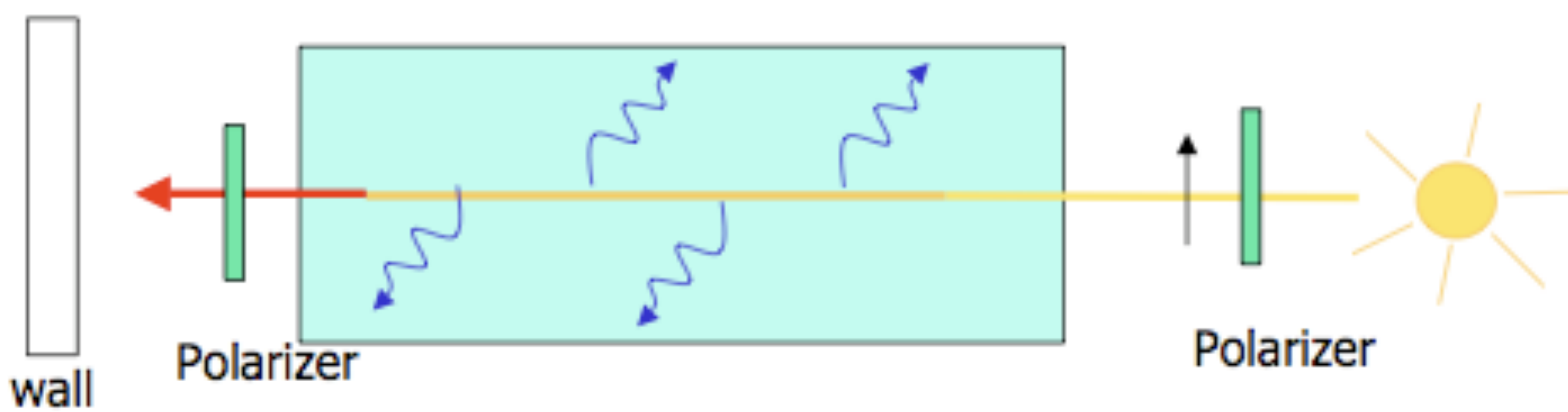
All the rest: elliptical polarization

For example, many organic molecules have a "chirality", meaning that they twist in a particular way. This turns out to mean that they interact with clockwise circular polarization *differently* than they interact with counterclockwise circular polarization!

If we were to shoot linearly polarized light into a medium containing such molecules, it would interact with the clockwise component of the linear wave differently than it would interact with the counterclockwise component. In other words, the medium would separate the linear polarization into its two circular components. We might expect any number of funky things to happen in this case.

Sugar solution experiment

Light goes through a Polaroid, an optically active sugar solution, a second Polaroid



First polarizer creates linearly polarized wave:
- Overlap of right-handed and left-handed circularly polarized waves.
  - RH and LH waves propagate at different speeds in the solution,
    - causing linear polarization direction to rotate as light penetrates in the solution.

Since the effect depends on wavelength λ, different colors are rotated differently.
The second polarizer selects polarization direction at exit, selecting different λ(colors)

Polaroids

- Sheet of plastic embedded with organic molecules extended in one direction
  - They can carry current in that particular direction: behave like antennas!
- When linearly polarized light hits the polaroid:
  - If E is aligned with orientation of molecules:
    - Charges move→current is generated→plastic heats: light stopped
- If E is perpendicular to orientation of molecules (“preferred direction”):
  - Charges will not be able to move in that direction: light goes through
- Conclusion:
  - Polaroids are transparent to light polarized || to their preferred direction and opaque to light polarized in the direction perpendicular to their preferred direction

Polaroids and polarization direction

- What happens when the light is polarized in a direction in between the preferred direction and its perpendicular?
- Example: light polarized along x axis; polaroid oriented at angle θ

$$\vec{E} = E_0 \cos(kz - \omega t)\hat{x}$$

$$\hat{p} = \hat{x}\cos\theta + \hat{y}\sin\theta$$

Light will go through partially

Since E has a component || to preferred direction of polaroid

E coming out is "overlap" between incoming E and polaroid's orientation

$$\left|\vec{E}_{out}\right| = \vec{E}\cdot\hat{p} = E_0\cos\theta\cos(kz-\omega t)$$

$$\vec{E}_{out} = \left|\vec{E}_{out}\right|\hat{p} \quad \text{(parallel to polaroid's orientation)}$$

Conclusion:

Polaroids reduce the amplitude of linearly polarized light by cosθ(angle between E and polaroid's orientation) and rotate the orientation of E by θ

Polarization of random light

Light from a bulb, sunlight, etc is not polarized

Superposition of many plane waves, each with its own polarization

$$\vec{E}_{random} = \sum_i E_0\left(\hat{x}\cos\theta_i + \hat{y}\sin\theta_i\right)\cos(kz-\omega t)$$

When light passes through a polaroid becomes linearly polarized

If polaroid is oriented || x axis:

$$\vec{E}_{out} = \sum_i E_0\left(\hat{x}\cos\theta_i\right)\cos(kz-\omega t) = E_0\hat{x}\cos(kz-\omega t)\sum_i\cos\theta_i$$

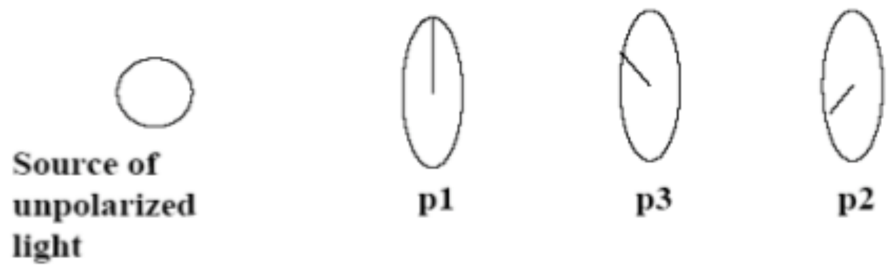
Conclusion:

Polaroids can be used to produce linearly polarized light

The intensity of the light will be reduced

Demo: 3 vs 2 polaroids

- 2 polaroids with orthogonal preferred direction will block light
  - First polaroid (P1) polarizes light in the direction x (for example)
  - Second polaroid (P2) oriented in the y direction, but E is now just || x



- Now place a third polaroid P3 in between P1 and P2 (at 45^o)
  - P1 will polarize light || x
  - P3 will select only component || to its preferred direction and rotate direction of polarization by 45^o. $E_0'=E_0\cos45^o$
  - P2 will select component y direction that now is not 0 anymore.
  - Intensity further reduced, but not 0! $E_0'=E_0(\cos45^o)^2= E_0/2$

Circular polarization

- Consider a wave with the following form:

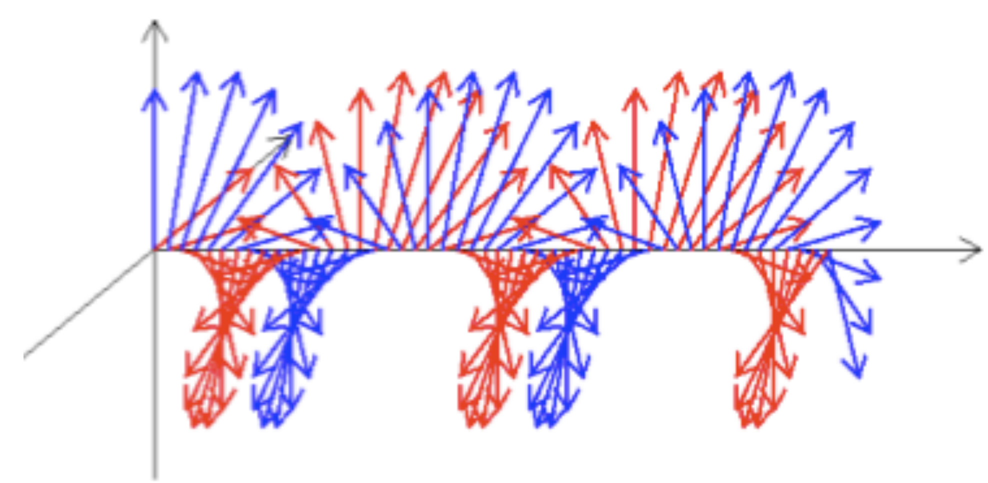
$$\vec{E} = E_0\hat{x}\sin(kz-\omega t)+E_0\hat{y}\cos(kz-\omega t)$$

$$\vec{B} = B_0\hat{x}\sin(kz-\omega t)+B_0\hat{y}\cos(kz-\omega t)$$

- What is it?
  - Easier to understand if we look at z=0

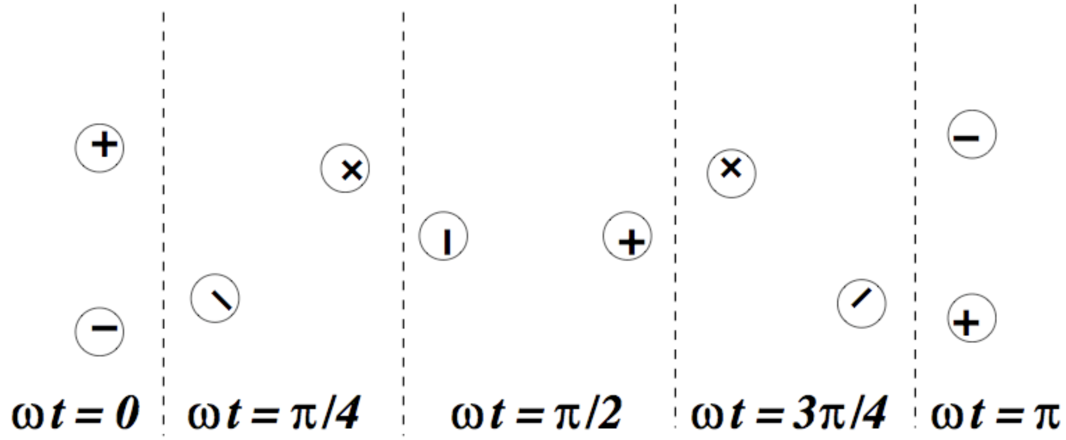
$$\vec{E} = -E_0\hat{x}\sin(\omega t)+E_0\hat{y}\cos(\omega t)$$

$$\vec{B} = -B_0\hat{x}\sin(\omega t)+B_0\hat{y}\cos(\omega t)$$

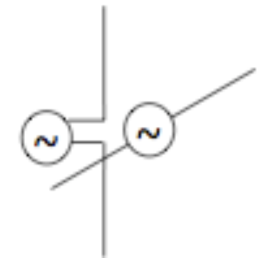


- Electric and magnetic fields rotate at frequency ω
  - Circular polarization because E and B vector tips describe circles over time

How to produce it? Rotating dipole



produces circularly polarized light propagating out of board

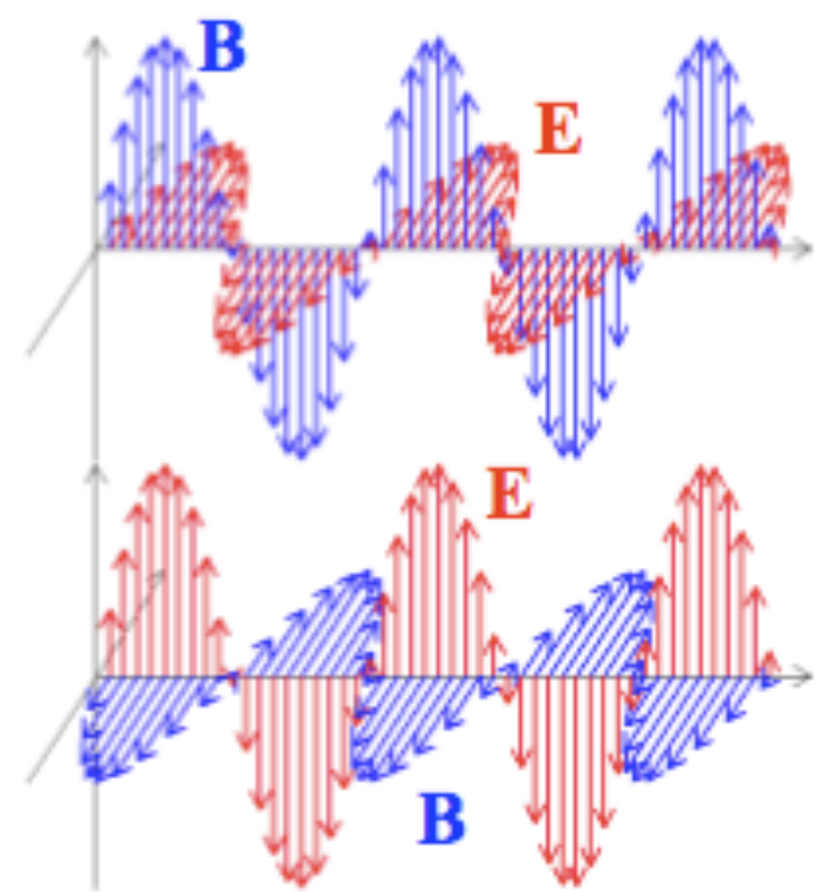


or 2 antennas at 90^o driven by currents out of phase by 90^o

Note: circular polarization does exist naturally in nature; in fact nature likes to produce it!

Elliptical Polarization

For a given k-vector, there are 2 independent solutions for the plane waves, e.g. 2 possible directions of the E-vector



$$\vec{E}_1 = E_0 \hat{x} \sin(kz - \omega t + \phi_1)$$
$$\vec{E}_2 = E_0 \hat{y} \cos(kz - \omega t + \phi_2)$$

All other solutions are just linear combinations of these

- $\phi_1 = \phi_2$: linear polarization
- $\phi_1 = \phi_2 + 90^o$: circular polarization
- All the rest: elliptical polarization

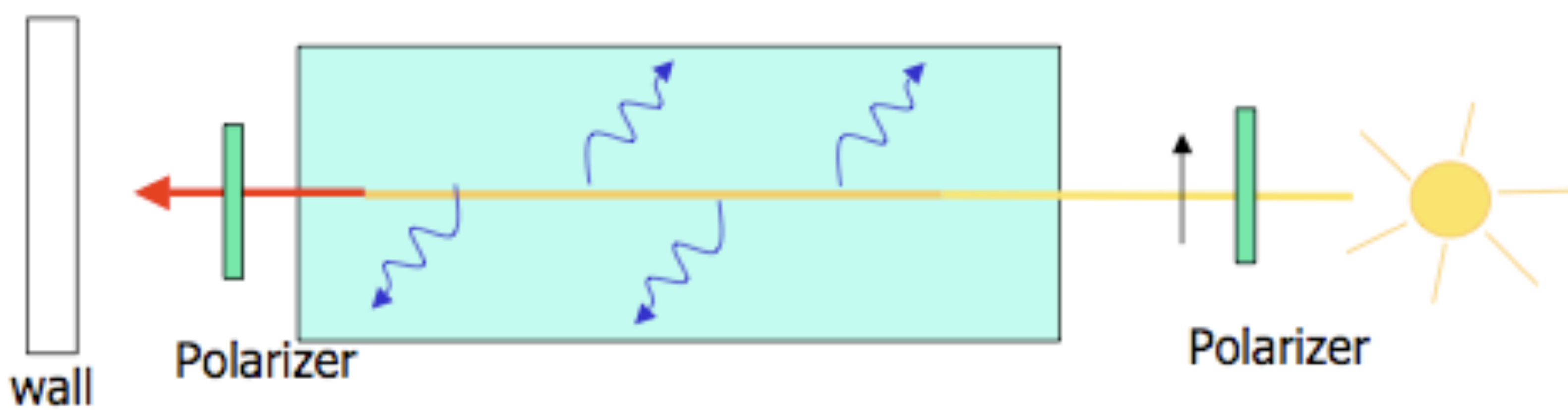
For example, many organic molecules have a "chirality", meaning that they twist in a particular way. This turns out to mean that they interact with clockwise circular polarization *differently* than they interact with counterclockwise circular polarization!

If we were to shoot linearly polarized light into a medium containing such molecules, it would interact with the clockwise component of the linear wave differently than it would interact with the counterclockwise component. In other words, the medium would separate the linear polarization into its two circular components. We might expect any number of funky things to happen in this case.

Think different bases for space!!!

Sugar solution experiment

Light goes through a Polaroid, an optically active sugar solution, a second Polaroid



First polarizer creates linearly polarized wave:

Overlap of right-handed and left-handed circularly polarized waves.

RH and LH waves propagate at different speeds in the solution,

causing linear polarization direction to rotate as light passes through the solution.

Since the effect depends on wavelength λ, different colors are rotated differently.

The second polarizer selects polarization direction at exit, selecting different λ(colors)

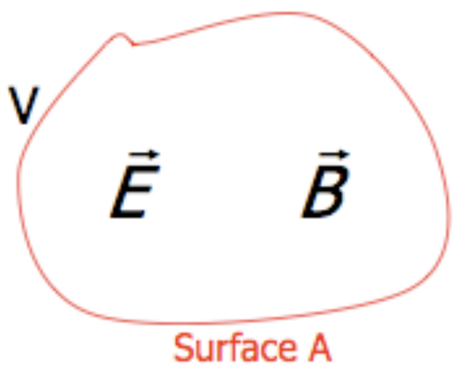
EM Energy

EM radiation carries energy

Obvious if you think about the fact that is the light from the sun that keeps us warm…

How does this energy propagate?

Consider a volume V of surface area A containing E and B fields



$$\text{Energy density}: u = \frac{energy}{volume} = \frac{1}{8\pi}\left(\vec{E}\cdot\vec{E} + \vec{B}\cdot\vec{B}\right)$$

$$\text{Total energy}: U = \int_V u\,dV = \frac{1}{8\pi}\int_V \left(\vec{E}\cdot\vec{E} + \vec{B}\cdot\vec{B}\right)dV$$

The Poynting vector

How does the energy U change over time?

$$\frac{\partial U}{\partial t} = \frac{1}{8\pi}\frac{\partial}{\partial t}\int_V \left(\vec{E}\cdot\vec{E} + \vec{B}\cdot\vec{B}\right)dV = \frac{1}{4\pi}\int_V \left(\frac{\partial \vec{E}}{\partial t}\cdot\vec{E} + \frac{\partial \vec{B}}{\partial t}\cdot\vec{B}\right)dV$$

Remembering that in vacuum we have $\nabla\times\vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}$, $\nabla\times\vec{B} = +\frac{1}{c}\frac{\partial \vec{E}}{\partial t}$

$$\frac{\partial U}{\partial t} = \frac{c}{4\pi}\int_V \left(\left(\nabla\times\vec{B}\right)\cdot\vec{E} - \left(\nabla\times\vec{E}\right)\cdot\vec{B}\right)dV$$

Remembering that $\nabla\cdot\left(\vec{E}\times\vec{B}\right) = -\vec{E}\cdot\left(\nabla\times\vec{B}\right) + \vec{B}\cdot\left(\nabla\times\vec{E}\right)$ Proof on next slide.

$$\Rightarrow \frac{\partial U}{\partial t} = -\frac{c}{4\pi}\int_V \nabla\cdot\left(\vec{E}\times\vec{B}\right)dV = -\int_V \nabla\cdot\vec{S}\,dV$$

where we defined the **Poynting vector** S as $\vec{S} = \frac{c}{4\pi}\vec{E}\times\vec{B}$

$$\nabla\cdot\left(\vec{E}\times\vec{B}\right)=\left(\nabla_m\hat{e}_m\right)\cdot\left(\varepsilon_{ijk}E_iB_j\hat{e}_k\right)=\varepsilon_{ijk}\nabla_m\left(E_iB_j\right)\hat{e}_m\cdot\hat{e}_k=\varepsilon_{ijk}\nabla_m\left(E_iB_j\right)\delta_{mk}$$

$$=\varepsilon_{ijk}\nabla_k\left(E_iB_j\right)=\varepsilon_{ijk}\nabla_k\left(E_i\right)B_j+\varepsilon_{ijk}E_i\nabla_k\left(B_j\right)$$

$$=B_j\left(\varepsilon_{kij}\nabla_k\left(E_i\right)\right)+E_i\left(-\varepsilon_{kji}\nabla_k\left(B_j\right)\right)$$

$$=-\vec{E}\cdot\left(\nabla\times\vec{B}\right)+\vec{B}\cdot\left(\nabla\times\vec{E}\right)$$

Interpretation of the Poynting vector

Given: $$\frac{\partial U}{\partial t} = -\int_V \nabla \cdot \vec{S} dV \overset{Stokes}{\rightarrow} \frac{\partial U}{\partial t} = -\int_A \vec{S} \cdot d\vec{A} = -\Phi_{\vec{S}}(A)$$

➙The rate of change of EM energy in the volume V is given by the flux of the Poynting vector S through the surface A

Minus sign: dA points outward ➙ U increases when S is opposite to dA

Interpretation of Poynting vector:

$$\vec{S} = \frac{c}{4\pi} \vec{E} \times \vec{B}$$ represents the flux of energy through A/unit time

It points in the direction of the EM energy flow $\vec{E}_0 \times \vec{B}_0 = \left|\vec{E}_0\right|^2 \hat{k}$

The flux of S through a surface gives the power through A

$$Power\ through\ A : \int_A \vec{S} \cdot d\vec{A}$$

Poynting vector: dimensional analysis

What are the units of the Poynting vector?

$$\left[\vec{S}\right] = \left[\frac{c}{4\pi} \vec{E} \times \vec{B}\right] = [c][B][E] \overset{cgs}{=} [c][E]^2$$

$$= \frac{length}{time} \frac{energy}{volume} = \frac{energy}{time \times area} = \frac{power}{area}$$

In cgs: [S]=erg $s^{-1}cm^{-2}$

as expected if the flux of S is the power through area A

Note: Magnitude of S is know as Intensity I

Intense source of radiations emit a lot of power per unit area

Application: plane waves

Consider a linearly polarized plane wave: $$\begin{cases}\vec{E}=\vec{E}_0\sin(kz-\omega t)\hat{x}\\ \vec{B}=\vec{B}_0\sin(kz-\omega t)\hat{y}\end{cases}$$

Poynting vector associated with it:

$$\vec{S}=\frac{c}{4\pi}\vec{E}\times\vec{B}=\frac{c}{4\pi}E_0^2\sin^2(kz-\omega t)\hat{z}$$

This can be compared to the energy density of the wave:

$$u=\frac{1}{8\pi}\left(\vec{E}\cdot\vec{E}+\vec{B}\cdot\vec{B}\right)=\frac{1}{4\pi}E_0^2\sin^2(kz-\omega t)$$

$$\Rightarrow\vec{S}=uc\hat{z}=uc\hat{k}$$

This is similar to $\vec{J}=\rho\vec{v}$

➙another way to show that S tells us about the flow of energy!

Usually the oscillation is very fast (e.g.: visible~10^{14}Hz)

➙ all that matters is the average energy density <S> and intensity <I>:

$$\left\langle\vec{S}\right\rangle=c\hat{k}\frac{E_0^2}{8\pi}\quad,\quad\left\langle I\right\rangle=c\frac{E_0^2}{8\pi}$$

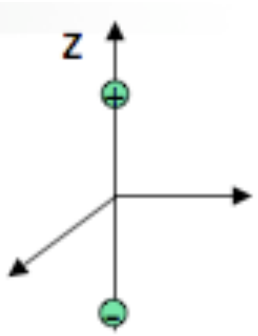
Application: Dipole radiation(advanced stuff (P112) but trust me for the moment)

Radiation emitted by a dipole oriented along the z-axis in spherical coordinates:

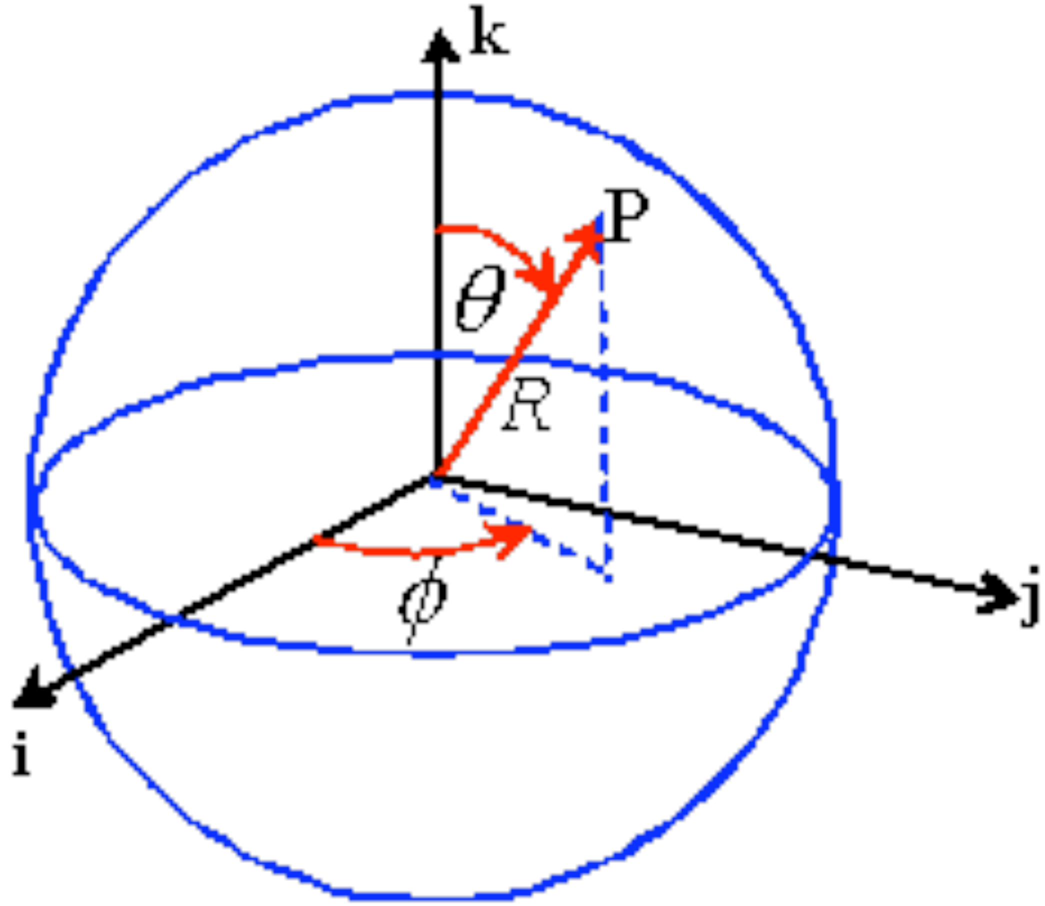
$$\begin{cases}\vec{E}=\dfrac{\omega^2 p}{c^2}\sin\theta\dfrac{\sin(kr-\omega t)}{r}\hat{\theta}\\ \vec{B}=\dfrac{\omega^2 p}{c^2}\sin\theta\dfrac{\sin(kr-\omega t)}{r}\hat{\phi}\end{cases}$$

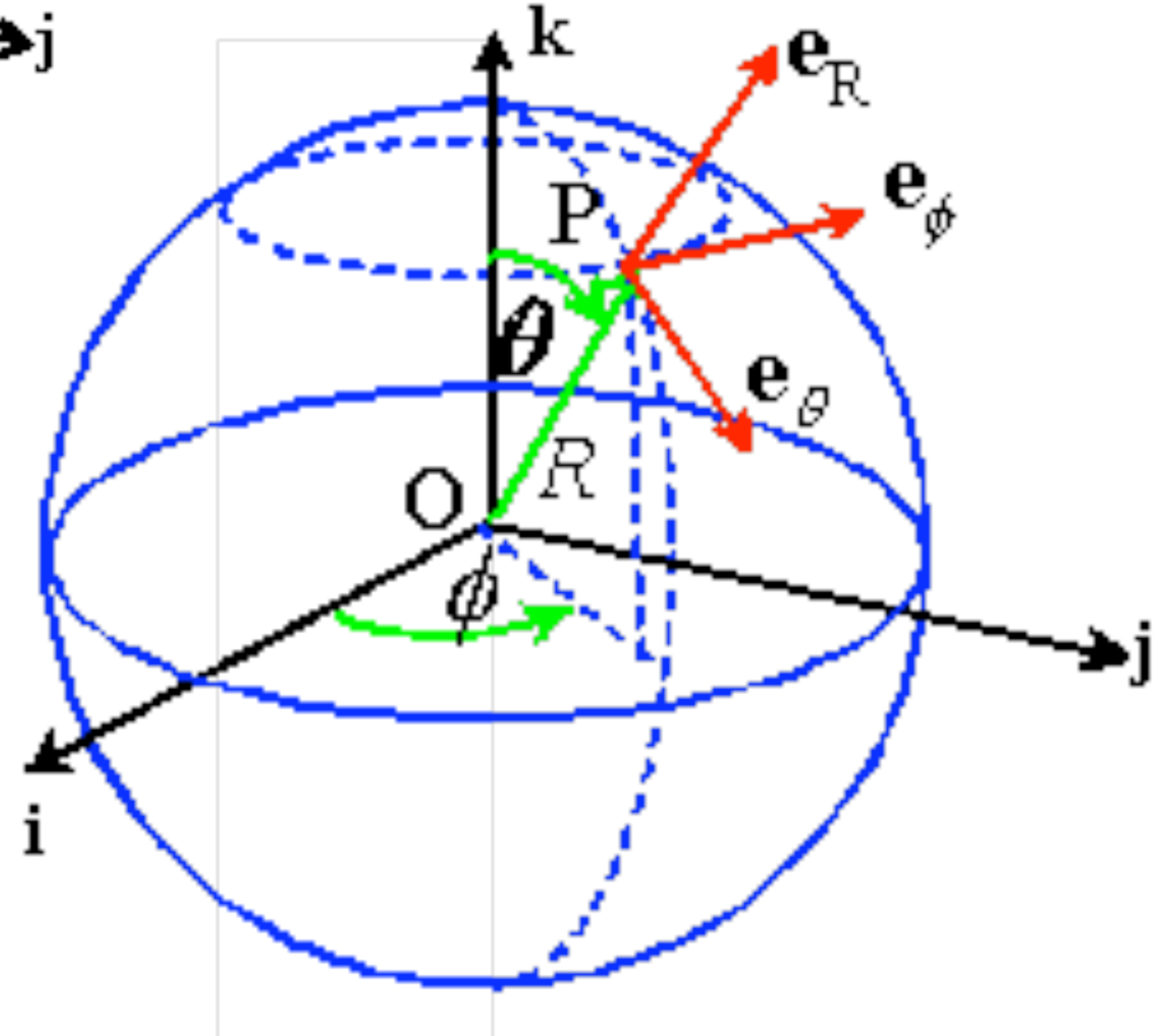
Note: solution valid for r >> λ = 2π/k

p = qd = dipole vector(see next slide)



This radiation propagates radially, with some angular dependence


k
P
θ
R
φ
j
i

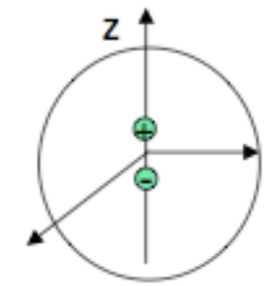

k
e_R
e_φ
P
θ
e_θ
R
O
φ
j
i

Poynting vector: $$\vec{S} = \frac{c}{4\pi}\vec{E}\times\vec{B} = \frac{1}{4\pi c^3}\omega^4 p^2 \sin^2\theta \frac{\sin^2(kr-\omega t)}{r^2}\hat{r}$$

$$\Rightarrow \langle \vec{S} \rangle = \frac{\omega^4 p^2}{8\pi r^2 c^3}\sin^2\theta\hat{r}$$

Note: Poynting vector (and I) falls as $1/r^2$: this should be intuitive. Why?

Draw a sphere of radius R around the dipole centered in origin: R >> d



Compute power radiated through the sphere:

$$\left\langle \frac{\partial U}{\partial t} \right\rangle = \int_A \langle \vec{S} \rangle \cdot d\vec{A} = \int_A \frac{\omega^4 p^2}{8\pi r^2 c^3}\sin^2\theta\hat{r}\cdot d\vec{A}$$

Since $d\vec{A} = R^2 \sin\theta d\theta d\phi \hat{r}$

$$\left\langle \frac{\partial U}{\partial t} \right\rangle = \frac{\omega^4 p^2}{8\pi R^2 c^3} R^2 \int_0^{2\pi} d\phi \int_0^{\pi} \sin^3\theta d\theta$$

Since

$$\int_0^{\pi} \sin^3\theta d\theta = \frac{4}{3} \Rightarrow \left\langle \frac{\partial U}{\partial t} \right\rangle = \frac{\omega^4 p^2}{3c^3}$$ (so-called Larmor formula)

Note: power through sphere does not depend on R

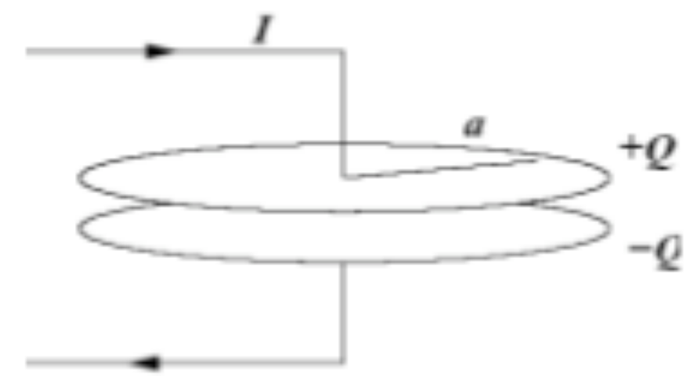
Why? S falls as $1/r^2$, area increases as r^2

→Power through S (flux through S) is constant: Energy is conserved

Application: capacitor

The Poynting vector applies to ANY situation in which both E and B appear, not just when we have radiation

Example: charging capacitor



Between plates: $\vec{E} = -\frac{4\pi Q}{A}\hat{z} = -\frac{4Q}{a^2}\hat{z}$ Points down

From generalized Ampere law: $\vec{B}(r) = \frac{2Ir}{ca^2}\hat{\phi}$ Curls around

Calculate Poynting vector:

$$\vec{S} = \frac{c}{4\pi}\vec{E}\times\vec{B} = -\frac{c}{4\pi}\frac{4Q}{a^2}\frac{2Ir}{ca^2}\hat{z}\times\hat{\phi} = \frac{2IQr}{\pi a^4}(-\hat{r})$$

$\hat{z}\uparrow$

$\hat{\phi}\swarrow \searrow\hat{r}$

Note: what is important here is the direction of S:

S points into the center of the capacitor as it should: the plates are charging up!

Momentum carried by EM wave

Since EM waves carry energy it's not surprising that they carry momentum as well

In relativity, E and p are related by $E^2 = p^2c^2 + m^2c^4$

For EM radiation, m=0: $E^2 = p^2c^2 \Rightarrow p = \frac{E}{c}$

Remember that $\left[\vec{S}\right] = \frac{energy}{time\times area} \Rightarrow \frac{1}{c}\left[\vec{S}\right] = \frac{energy/c}{time\times area} = \frac{momentum}{time\times area}$

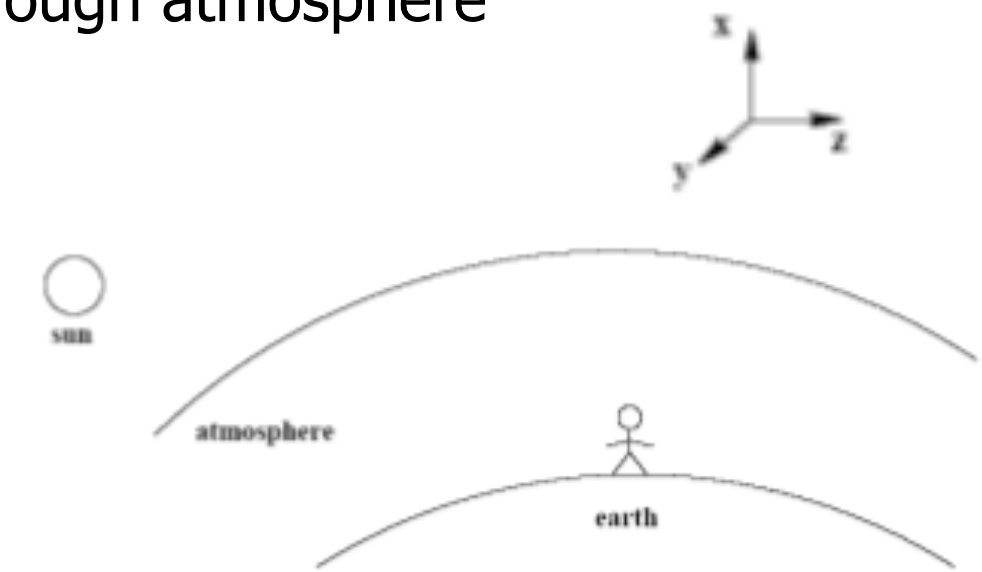
Dimensional analysis will also tell us that: $\frac{1}{c}\left[\vec{S}\right] = \frac{momentum}{time\times area} = \frac{force}{area} = pressure$

→ Radiation pressure must exist!

Scattering of light

When we send light into a medium, the light is scattered in many directions

Example: light from Sun (unpolarized) passing through atmosphere
Propagation of light || z(initially)
We look up in x direction
What kind of light do we see?



Since light propagates || z: no polarization || z
We measure the light (with our eyes!) along the x direction: no polarization || x
→The light we see must be linearly polarized along the y direction
This is actually not really true because the light scatters multiple times, but it suggests the general tendency
→scattered light would be more intense in direction perpendicular to polarization direction
What if the put a giant Polaroid in front of the Sun?
Rotating the Polaroid would allow us to change intensity of the light:
Max intensity when polarization direction is || y axis (orthogonal to "up")
Dark when polarization direction is || x axis (parallel to "up")

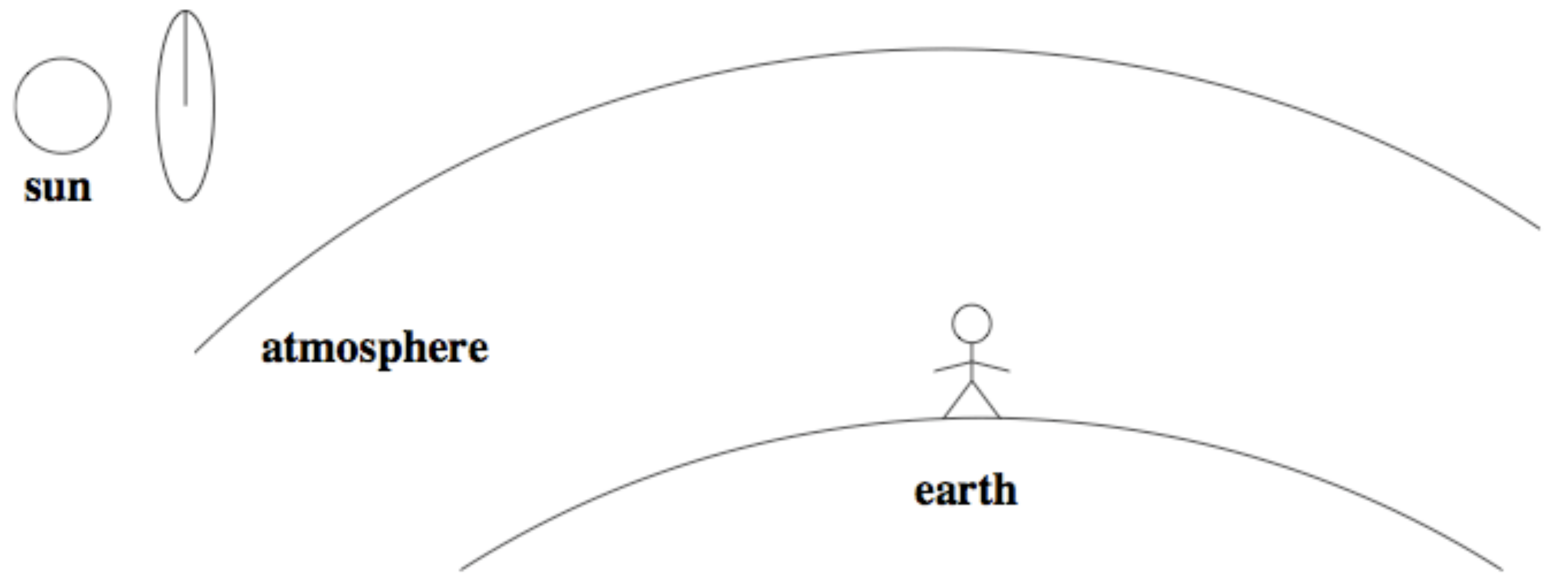


Figure 1: Dark above my head.

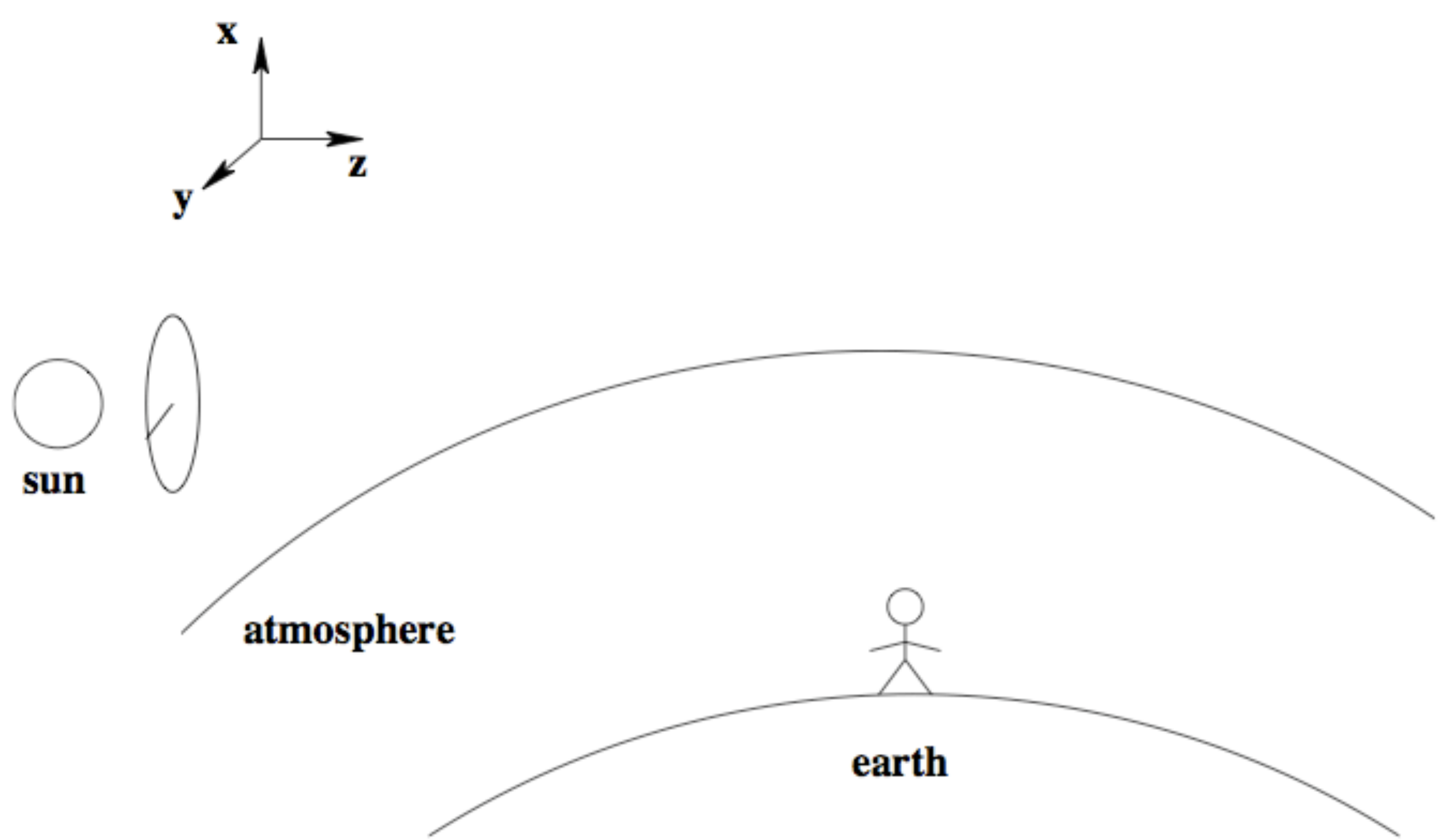


Figure 2: Bright above my head.

How is light scattered?

Light hits a molecule; the E-field shakes the molecule's charges with input frequency ω

➙dipole radiation(molecules re-radiate in all directions)

Are all frequencies scattered in the same way?

Electric fields of scattered radiation depend on acceleration of(dipole) charges

$E_{scattered} \propto \partial^2 d / \partial t^2 \propto \omega^2$ if dipole moment of the shaken molecule goes as d~cosωt

Intensity of scattered radiation: $I \propto E^2_{scattered} \propto \omega^4 \propto \lambda^{-4}$

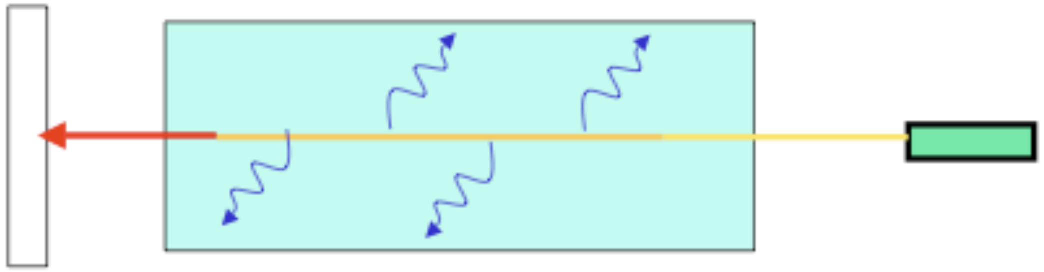
Since $\lambda_{red} \sim 2\lambda_{blue}$ ➙blue is scattered 16 times more than red

The sky is blue during the day when we look away from sun and light we see is light that has scattered out of initial propagation direction. All wavelengths in sunlight but smaller wavelength have higher scattering probability. Also why sunset is red. At sunset we are looking at the sun along initial propagation direction and looking through large thickness of atmosphere. Most short wavelength stuff scatters out and only long wavelength stuff remains.

Sunset experiment

Solution of distilled water and salt (NaCl)

Unpolarized light is shining through it to the wall



Add H_2SO_4 to make $Na_2S_2O_3 5H_2O$ (Na thiosulfate): big molecule!

Lights starts scattering: fog; light on the wall becomes red first and then dark as all the light is scattered toward the audience (as in sunset)

What happened?

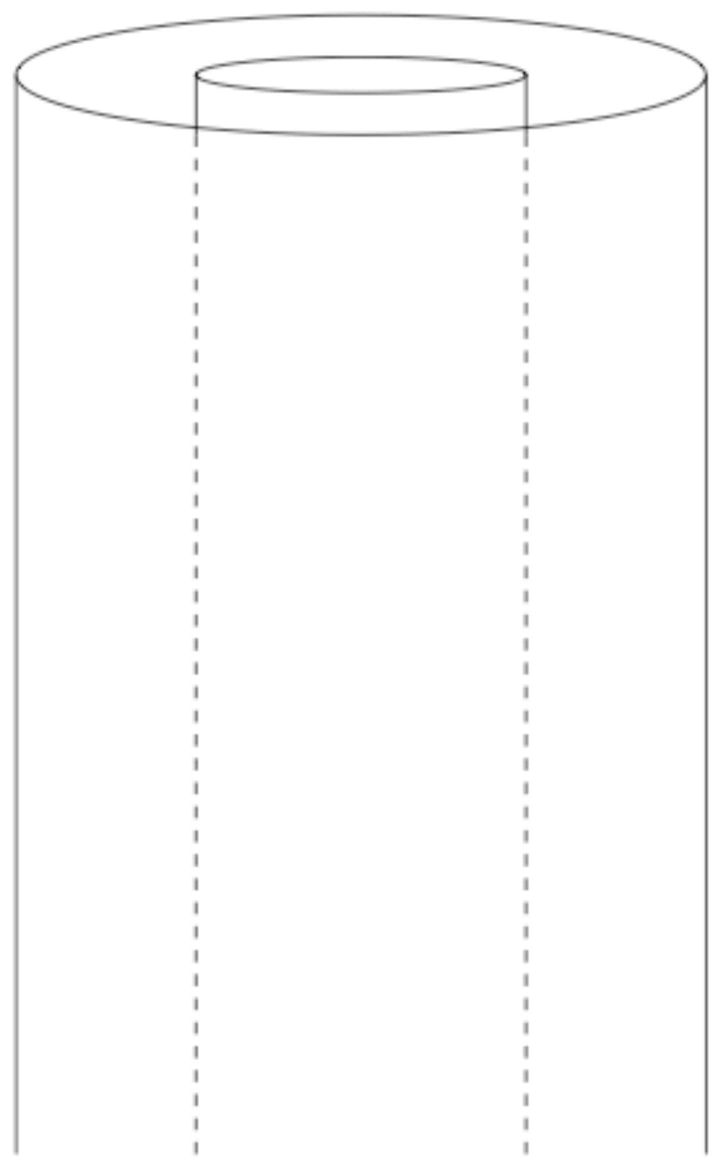
Chemical reaction creates bigger and bigger molecules that scatter more and more light. Blue light is scattered first. Red makes it for a while but eventually scatters too.

Note: light is polarized!

Transmission lines

Suppose we want to send a signal from one device to another — our goal is to get electromagnetic energy from point A to point B. What's the best way to do this? If we just use wires, we're going to have problems — the wires will act like antennas and we'll lose a lot of energy to radiation. Instead, we would like to set up some kind of structure that *shields* the signal, so that no field leaks outside. This setup is called a transmission line.

The simplest example of a transmission line is a coaxial cable: a pair of conducting tubes nested in one another. Equal and opposite currents flow on the inner and the outer tubes,



Inner radius: ***a***

Outer radius: ***b***

so that the net current "seen" outside the cylinder is zero. Note that this is *not* a simple DC current — it is a highly oscillatory AC current. This means that there must be some charge distribution as well. Equal and opposite charges "live" on the inner and outer tubes of the coax.

Suppose that the inductance per unit length of this cable is L', and that its capacitance per unit length is C'. Consider a short length Δx of this cable. This little segment has a total inductance of $L'\Delta x$. If the current flowing down the cable is I and has the rate of change $\partial I/\partial t$, then the voltage drop due to inductance over this short length is

$$\Delta V = V(x + \Delta x) - V(x) = -(L'\Delta x)\frac{\partial I}{\partial t} \ .$$

Divide both sides by Δx and take the limit:

$$\frac{\partial V}{\partial x} = -L'\frac{\partial I}{\partial t} \ .$$

Consider now the charge on this little segment: since it is at potential $V(x)$, and its capacitance is $C'\Delta x$, the amount of charge on this segment is

$$\Delta Q = (C'\Delta x)V(x) \ .$$

Divide and take the limit:

$$\frac{\partial Q}{\partial x} = C'V(x) \ .$$

Let's take another derivative of this:

$$\begin{aligned}
\frac{\partial^2 Q}{\partial x^2} &= C'\frac{\partial V}{\partial x} \\
&= -C'L'\frac{\partial I}{\partial t} \\
&= C'L'\frac{\partial^2 Q}{\partial t^2} \ .
\end{aligned}$$

In going from the first to the second line, we used our previous result relating $\partial V/\partial x$ and $\partial I/\partial t$; in going from the second to the third line, we used $I = -\partial Q/\partial t$ (discharging capacitor current).

Rearranging this, we see yet another wave equation!

$$\frac{\partial^2 Q}{\partial t^2} - \frac{1}{C'L'}\frac{\partial^2 Q}{\partial x^2} = 0 \,.$$

With a little bit of tweaking, you should be able to convince yourself that this equation holds replacing Q with V and I as well. The speed of propagation of this wave is

$$v = \frac{1}{\sqrt{L'C'}} \,.$$

Now, for the coaxial cable introduced above, you should be able to show very easily that

$$\begin{aligned} C' &= \frac{1}{2\ln(b/a)} \\ L' &= \frac{2\ln(b/a)}{c^2} \,. \end{aligned}$$

This means that

$$v = \frac{1}{\sqrt{L'C'}} = c \,.$$

The wave propagates down the transmission line at the speed of light! Not surprising, at least in retrospect.

One other very important quantity determines the characteristic of a transmission line: its impedance. Following our intuition from our study of AC circuits, we define the impedance as the ratio of the voltage to the current:

$$Z = \frac{V}{I} = \frac{V}{\Delta Q/\Delta t} \,.$$

As we already discussed, for a small segment of the cable $\Delta Q = (C'\Delta x)V$, so

$$\frac{\Delta Q}{\Delta t} = C'V\frac{\Delta x}{\Delta t} = \frac{C'V}{\sqrt{L'C'}} = V\sqrt{\frac{C'}{L'}}$$

where we have used the fact that $\Delta x/\Delta t$ is just the speed of propagation down the line.

Putting everything together, we find

$$Z = \sqrt{\frac{L'}{C'}} \,.$$

For the coaxial cable, this yields

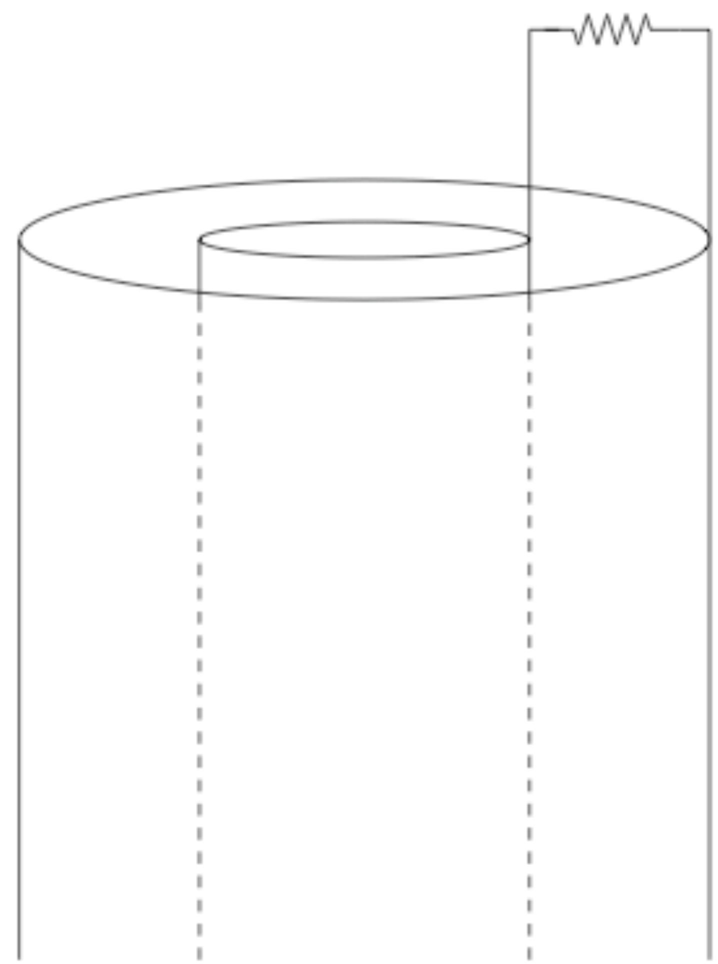
$$Z = \frac{2\ln(b/a)}{c} \,.$$

Using the conversion factor 1 Ohm $= 1.1 \times 10^{-12}$ sec/cm, this means

$$Z = 60\,\text{Ohms}\,\ln(b/a)\ .$$

Many coaxial cables are actually filled with a dielectric material (which we have not discussed in detail); this has the effect of increasing the capacitance, which reduces the line's characteristic impedance. 50 Ohms is a value that is commonly encountered (e.g., in cable TV systems).

One of the most important consequences of the impedance of a transmission line is in determining how it must be *terminated*. Roughly speaking, a terminated cable has the form



The resistor might actually be some electronic device, like a television or a computer; then again, it might just be a resistor. The point is that the signal which is flowing down the cable is then fed into this device or resistor.

If the resistance of this "load" precisely matches the impedance of the cable, then we have *matched* the impedances. Because V/I does not change in going from the cable to the resistor, it is as though the cable were of infinite length! There is no way for the signal to "know" whether it is in the resistor or in the cable.

Suppose the resistance were ∞ Ohms — an open circuit. Then, the currents would essentially bounce off the "resistor" and head back where they came from. The signal would just reverse course — you'd get a horrible reflection back up the line. Suppose the resistance were 0 Ohms — a short circuit. Then, the currents would just flow from the inner tube to the outer tube, and vice versa — the signal would reverse course *and* switch sign on its amplitude!

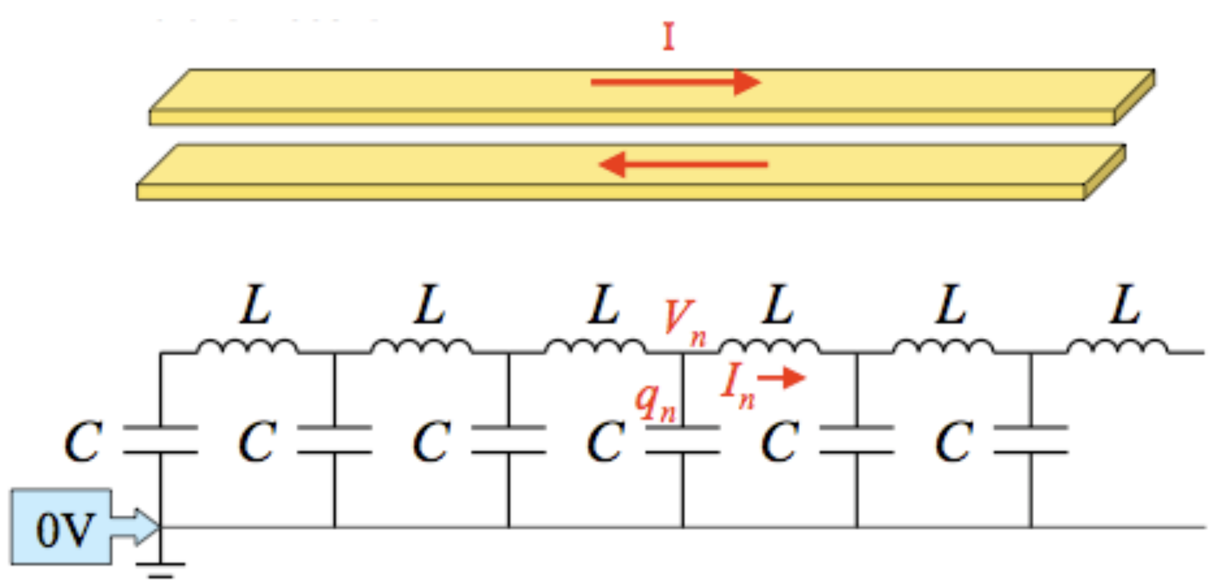
In general, if the impedance of the cable doesn't match the impedance of the load, you get reflections of this sort. This is one reason why cheap cable TV equipment can give you a horrible picture — there are these extra signals bouncing around, which eventually arrive at the TV late, and possible with a 180° phase shift. This causes ghost images to appear, and generally confuses your equipment.

Transmission lines(again)

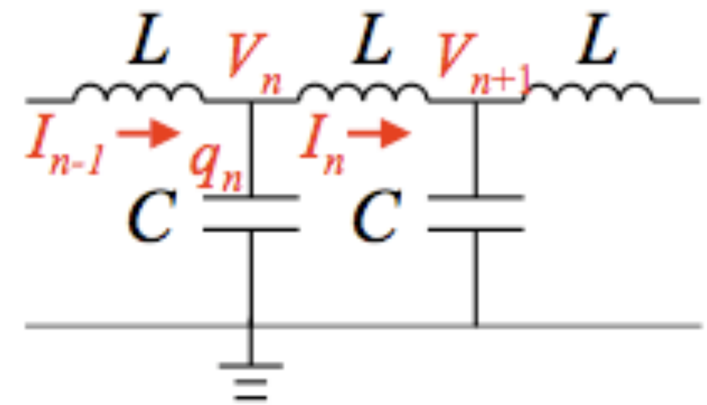
- Transmission line = a pair of conductors used to transmit a signal
  - Goal: send electric signals from A to B connected by the transmission line
  - Examples: coaxial cable (nested cylindrical conductors), parallel wires
- Questions:
  - Can we have the output signal be exactly the same as input signal?
    - Not smaller or distorted
- How fast can the signal propagate?
  - Can the transmission be instantaneous? No! It has to be <= c

Model of transmission line --- Simplest example of transmission line

Parallel ribbons



Charge, Voltage and Current

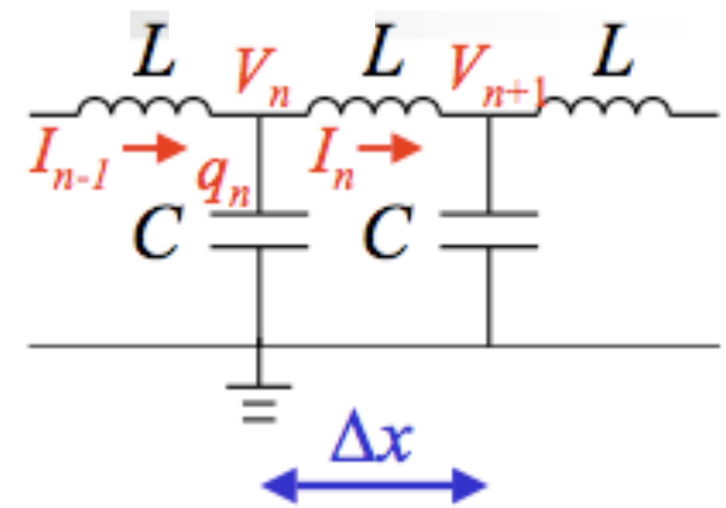


Voltage across C $\quad V_n = \dfrac{q_n}{C}$

Voltage across L $\quad V_{n+1} - V_n = -L\dfrac{dI_n}{dt}$

Charge conservation $\quad I_{n-1} - I_n = \dfrac{dq_n}{dt}$

Going Continuous



Assume we have L's and C's at every Δx

Replace indices with position

$$V_n \rightarrow V(x)\,,\, I_n \rightarrow I(x)\,,\, q_n \rightarrow q(x)$$

The equations become:

$$V_{n+1} - V_n = -L\frac{dI_n}{dt} \Rightarrow -L\frac{\partial I(x)}{\partial t} = V(x+\Delta x) - V(x) = \frac{\partial V(x)}{\partial x}\Delta x$$

$$I_{n-1} - I_n = \frac{dq_n}{dt} = C\frac{dV_n}{dt} \Rightarrow C\frac{\partial V(x)}{\partial t} = I(x-\Delta x) - I(x) = -\frac{\partial I(x)}{\partial x}\Delta x$$

Wave Equation

$$\Rightarrow \left\{ \begin{aligned} -L\frac{\partial I(x,t)}{\partial t} &= \frac{\partial V(x,t)}{\partial x}\Delta x \\ C\frac{\partial V(x,t)}{\partial t} &- \frac{\partial I(x,t)}{\partial x}\Delta x \end{aligned} \right\} \quad Define: \left\{ \begin{aligned} \frac{L}{\Delta x} &= L' \\ \frac{C}{\Delta x} &= C' \end{aligned} \right.$$

Take space derivative or first equation and plug in the second when needed:

$$-L'\frac{\partial}{\partial x}\frac{\partial I(x,t)}{\partial t} = \frac{\partial^2 V(x,t)}{\partial x^2} = -L'\frac{\partial}{\partial t}\frac{\partial I(x,t)}{\partial x}$$

$$\frac{\partial^2 V(x,t)}{\partial x^2} = L'\frac{\partial}{\partial t}C'\frac{\partial V(x,t)}{\partial t} = L'C'\frac{\partial^2 V(x,t)}{\partial t^2}$$

$$\frac{\partial^2 V(x,t)}{\partial x^2} = L'C'\frac{\partial^2 V(x,t)}{\partial t^2} \Rightarrow wave\ equation\ with\ v = \frac{1}{\sqrt{L'C'}}$$

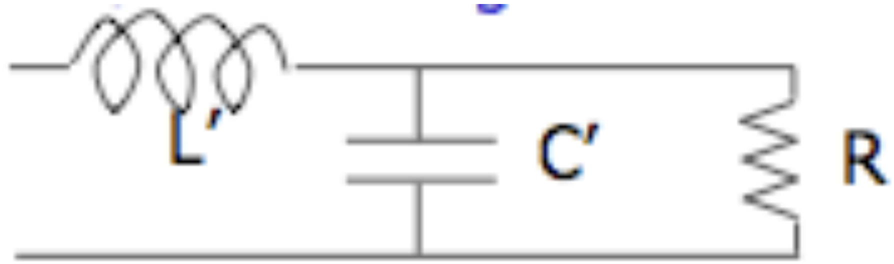
Wave Propagation Velocity

$$v = \frac{1}{\sqrt{L'C'}}$$

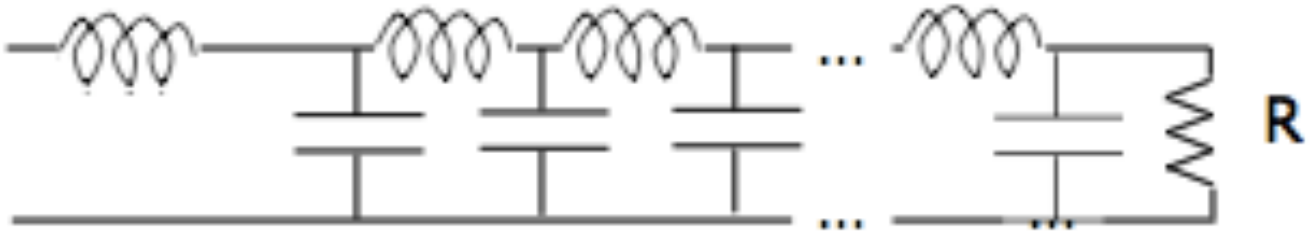
- Can velocity be as big as we like?
  - It cannot be true if we consider special relativity
    - v can never be made faster than c
- This tells us something about L′ and C′…
- Let’s calculate what v is for parallel wires

- Transmission line = a pair of (twisted) cables used to transmit a signal
  - Current flows in one direction on one cable and comes back on the other cable
- If terminated correctly, Z is purely real: $Z \sim R_{termination}$
- Find R when capacitance per unit length = C′ and inductance per unit length = L′

In theory:



In practice infinite sum of infinitesimal elements C and L:



Calculate Z of the last piece and impose that it's purely real.

$$Z_{eq} = i\omega L + \left(\frac{1}{R} + i\omega C\right)^{-1} = i\omega L + \frac{R}{1 + i\omega RC} = \frac{i\omega L - \omega^2 RLC + R}{1 + i\omega RC} \overset{impose}{=} R$$

$$\Rightarrow i\omega L - \omega^2 RLC + R = R + i\omega R^2 C$$

Ignoring term with LC (small) : $\Rightarrow R = \sqrt{\frac{L}{C}} = \sqrt{\frac{L'}{C'}}$

What happens when transmission line is terminated correctly?

- Z is purely real: Z~$R_{termination}$ ➙ Z is a constant of the cable:
  - Z does not depend on how long the cable is!
- If $R \neq \sqrt{L'/C'}$:
  - ➙Z will depend on how long the cable is and on the frequency of the signal
  - ➙Distortions of the signal!

How an EM wave looks in a different frame:

Plane EM wave in vacuum. Direction of travel in frame F = $\hat{n}$. E and B fields measured somewhere at some time in F by observer in F. What will be measured by an observer in frame F' passing particular point at that time?

F' relative motion with speed v in x-direction wrt F. Earlier derivations (Chapter 6) gave

$$E_x' = E_x \quad , \quad E_y' = \gamma\left[E_y - \beta B_z\right] \quad , \quad E_z' = \gamma\left[E_z + \beta B_y\right]$$

$$B_x' = B_x \quad , \quad B_y' = \gamma\left[B_y + \beta E_z\right] \quad , \quad B_z' = \gamma\left[B_z - \beta E_y\right]$$

Key to answer is the way two particular scalar quantities transform. Purcell does algebra to show that we have two invariants under Lorentz transformations:

$$\vec{E} \cdot \vec{B} = \vec{E}' \cdot \vec{B}' \quad , \quad E^2 - B^2 = E'^2 - B'^2$$

Invariance of these two quantities=important general property of EM field(not just EM wave). For EM wave fields implications are simple and direct. Know plane wave has E-field perpendicular to B-field and B=E in magnitude. In this case each of the invariants = 0.

If an invariant is =0 in any frame, then it =0 in all frames. Thus, any Lorentz transformation will leave the E-field and B-field perpendicular and equal in magnitude.

A light wave looks like a light wave in all reference frames!!!!

This should not surprise anyone - that is what we assumed to derive SR!

Einstein wondered at age 16 what one would observe if one could “catch up” with a light wave.

Let $E_y = E_0$, $E_x=E_z=0$, $B_z=E_0$, $B_x=B_y=0$. This is an EM wave traveling in the x-direction. This is confirmed by the the fact that the Poynting vector (direction of momentum flow) points in the x-direction:

$$\vec{E} \times \vec{B} = E_0^2 \hat{x}$$

Using identity $\gamma^2(1-\beta^2)=1$ we find

$$E_y' = E_0\sqrt{\frac{1-\beta}{1+\beta}} \quad , \quad B_z' = E_0\sqrt{\frac{1-\beta}{1+\beta}}$$

Note reappearance of k-factor I introduced in my derivation of SR!!

As observed in F′, the amplitude of the wave is reduced. The wave velocity is still c in F′ as it is in F. The EM wave has no rest frame. In the limit $\beta=1$, the amplitudes of the E and B fields observed in F′ are reduced to zero.

The EM wave has vanished!!!!!

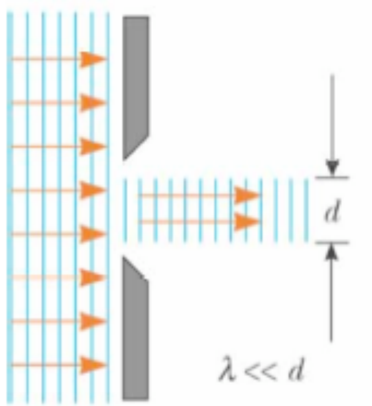
That finishes our study of E & M at this level. More to come in Physics 112 Seminar.

Assumptions of geometric optics

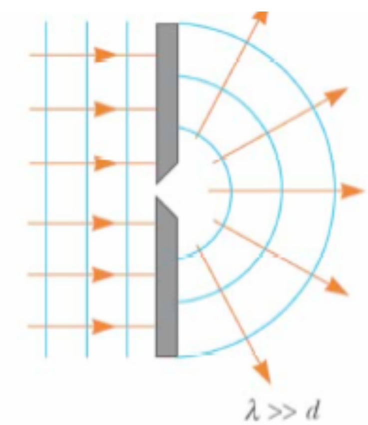
- The description of the propagation of light with rays is applicable if the wavelength of the light λ is much less that the size of optical apertures d encountered by the light.

- If λ becomes comparable or smaller than d then the wave character of light comes into play and the incident wave diffracts at the aperture.

$\lambda \ll d$

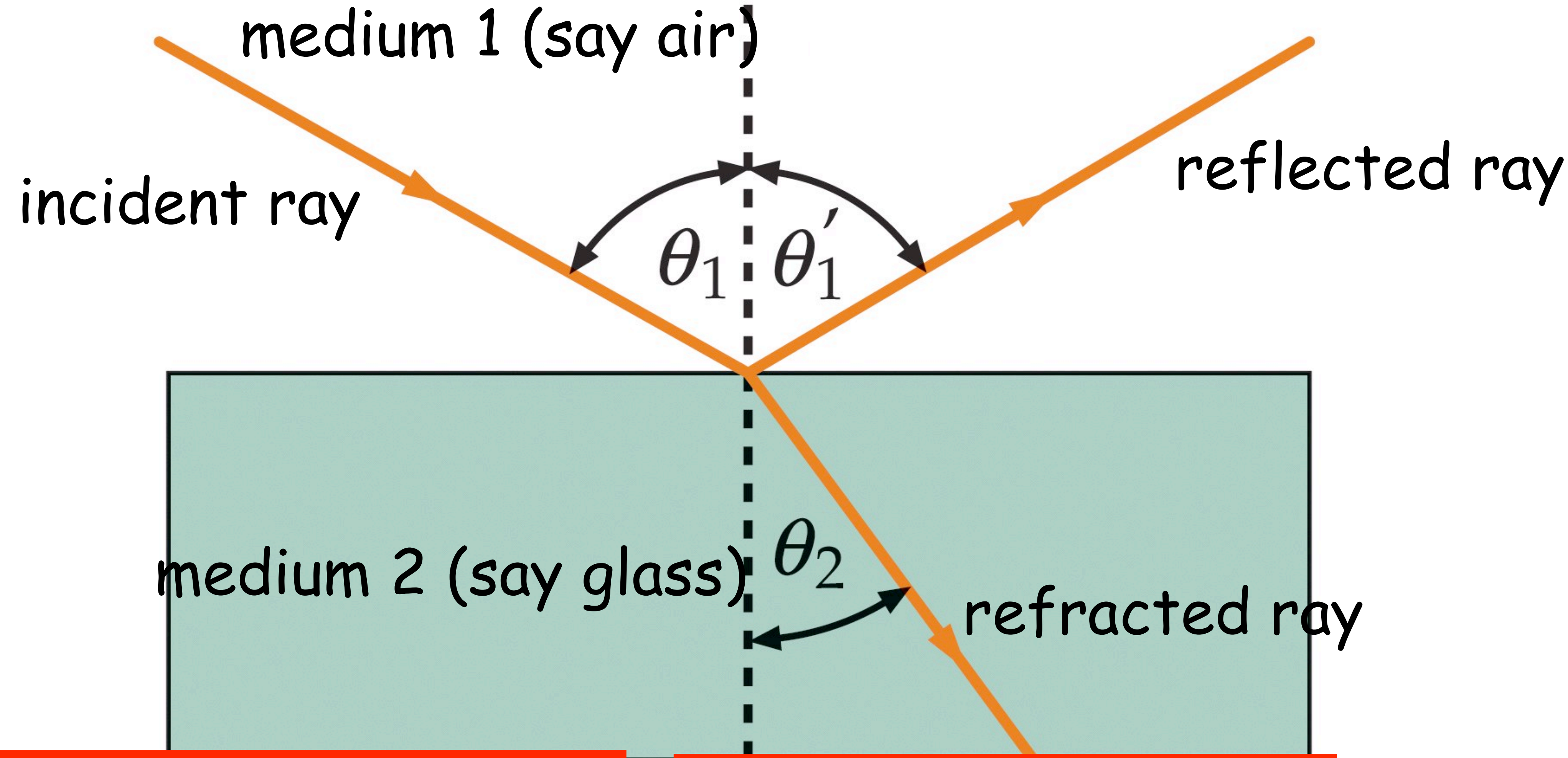


- For now we assume $\lambda \ll d$ and rays in a given optical medium travel in straight lines.



$\lambda > d$

reflection and refraction

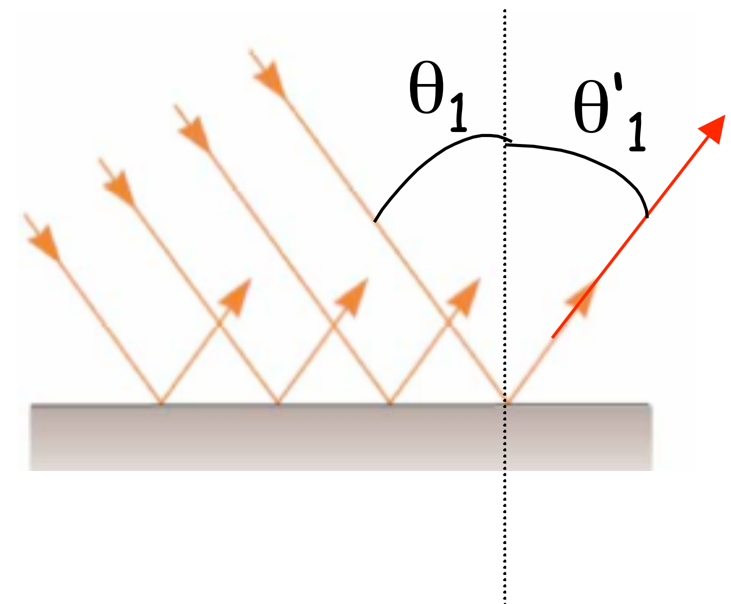


How are the angles related?

How are the intensities related?

Law of Reflection

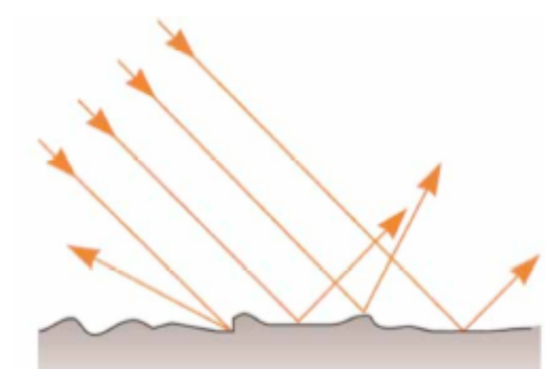
- Spectral reflection occurs from surfaces that are "flat" meaning the size of imperfections on the surface are small compared to the wavelength of the light.



Law of reflection

$\theta_1 = \theta'_1$

- Diffuse reflection occurs from "rough" surfaces.



Velocity of light in a transparent medium

- Refraction occurs when light passes through a boundary between two media having different light propagation velocities.
- In a vacuum the velocity of light is $c = 3.00 \times 10^8$ m/s.
- In a transparent medium such as glass light is absorbed and re-emitted by atoms/molecules of the material causing a slowing of its propagation velocity.

material	velocity of light/c
water	0.75
typical glass	0.64
diamond	0.41

Definition of the index of refraction

- Define n = index of refraction of a material to be $n = c/v$
- For a vacuum n is identically 1 and for most gases under standard conditions n is nearly one (n_{air} = 1.0003).
- For transparent liquids and solids the index of refraction varies between about 1.3 and 2.4

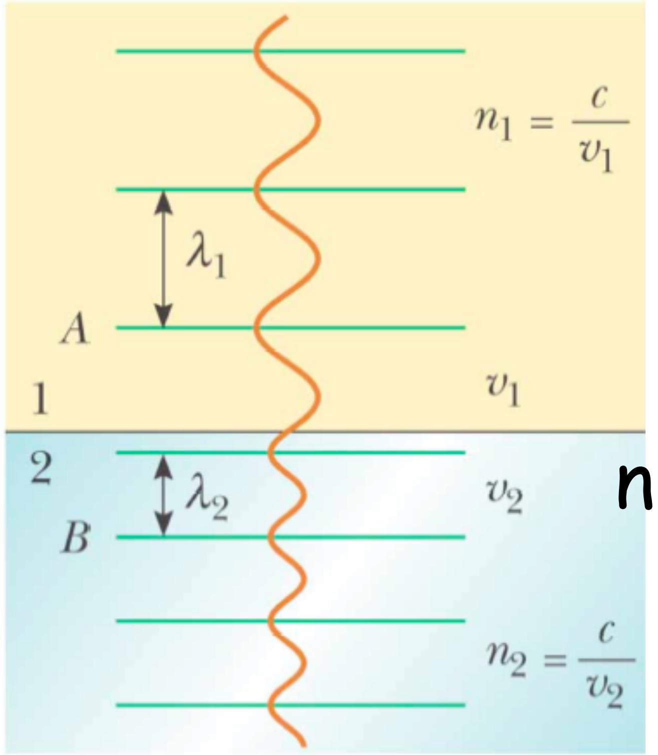
material	index of refraction
water	1.333
typical glass	~ 1.5
diamond	2.419

Light propagation in transparent media

- As light passes into a material with index of refraction n:
  - The velocity changes:

    $v_{material} = c_{vacuum}/n$
  - The frequency does not change
  - Therefore the wavelength does change:

    $\lambda_{material} = v_{material}/f = c_{vacuum}/fn = \lambda_{vacuum}/n$

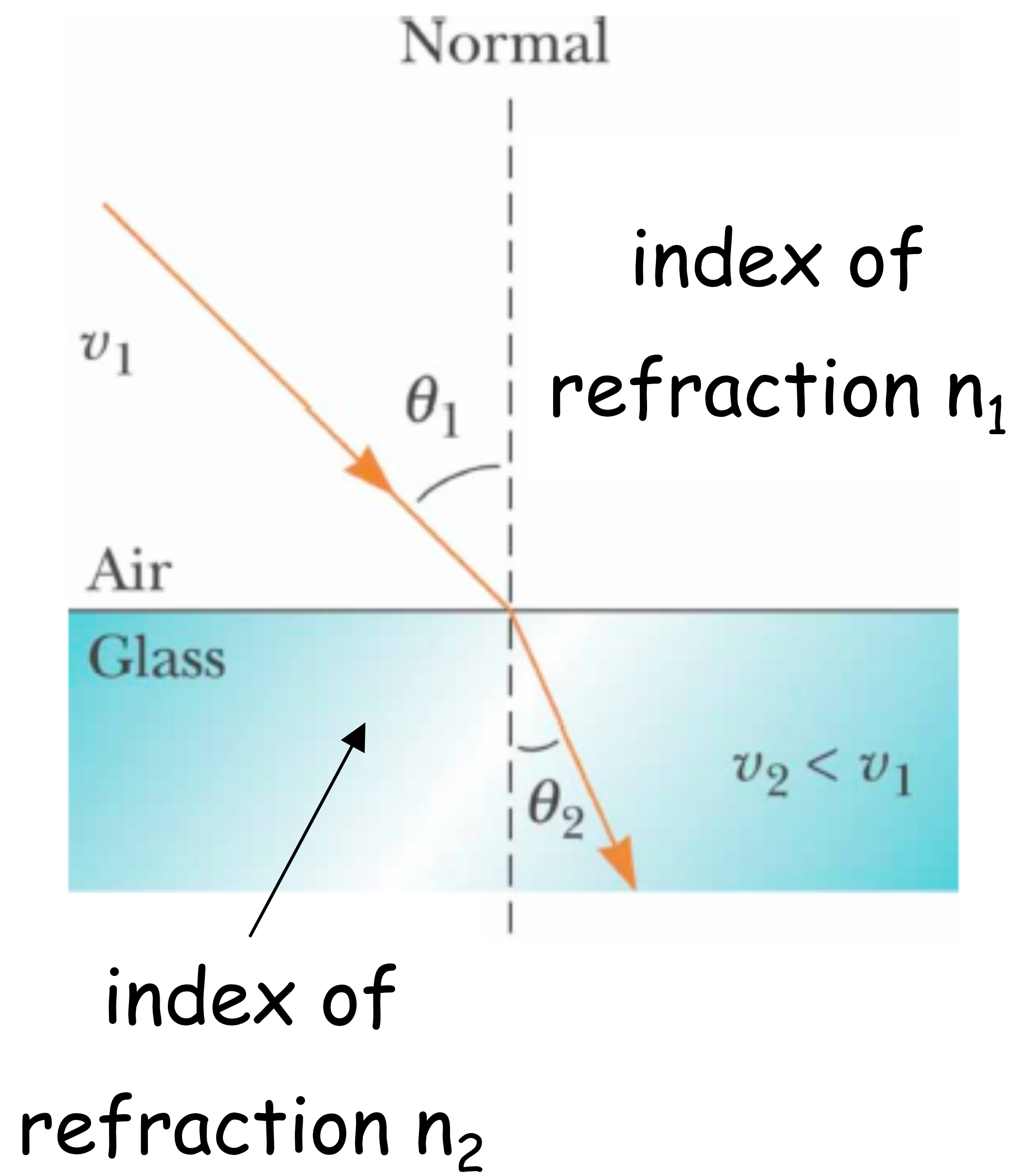


$n_2 > n_1$

Law of refraction = Snell's Law

- The relationship between the angle of the incident and reflected ray is given by a law first deduced from experiment in ~ 1600.

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

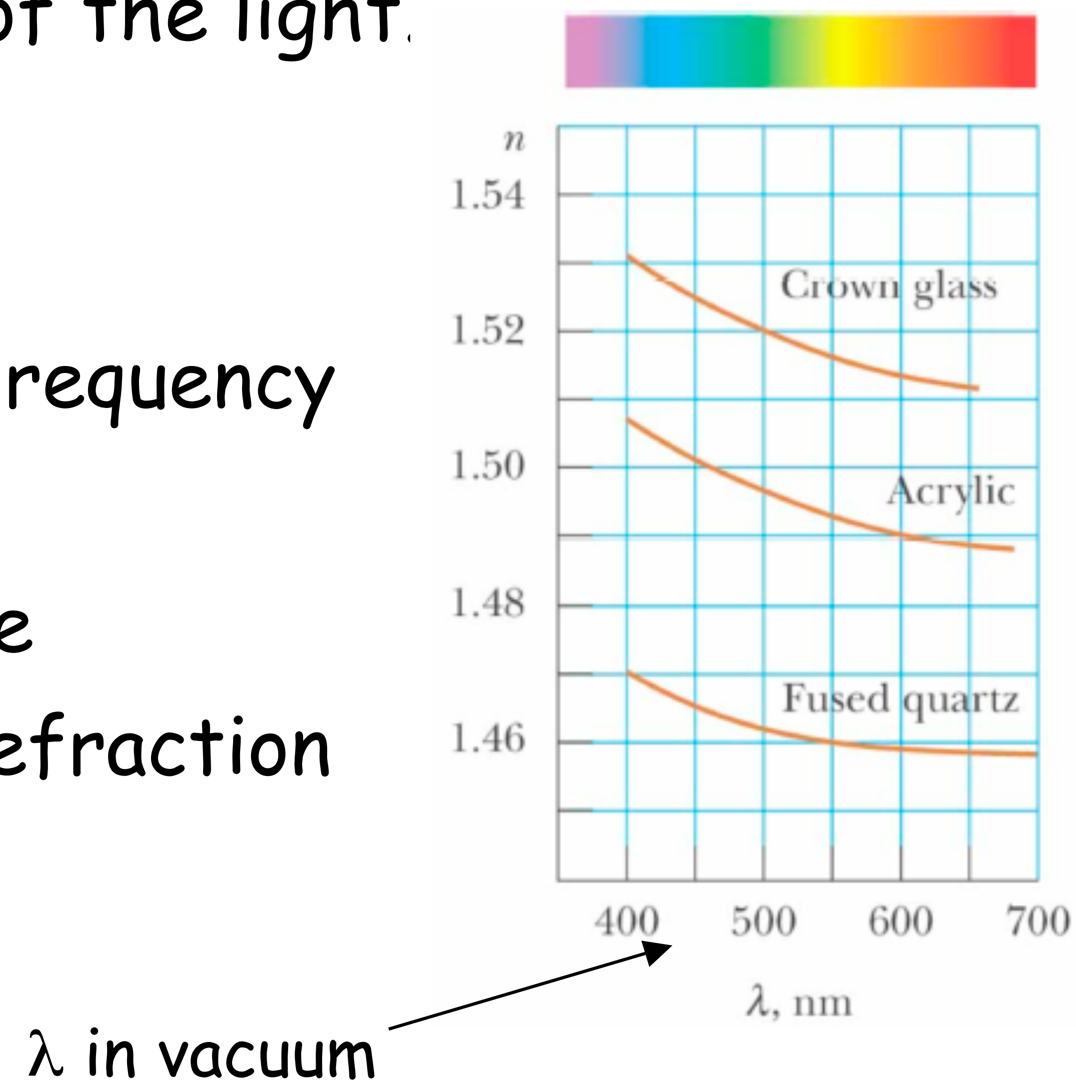


Properties of the index of refraction

- The velocity of light in a medium (not a vacuum) depends slightly on the frequency of the light.

- Therefore the index of fraction n depends on frequency

- As a consequence the angle of bending of light from refraction will depend on frequency.



Dispersion of light

- White light entering a medium (say glass) from air ($n \sim 1.0$, incident angle θ_1) will be separated into colors since $n_g = n_g(f)$.

$$\sin\theta_2(f) = [\sin\theta_1] / n_g(f)$$

dispersed colors in glasss

incident white light

Overview

- There are obviously a tremendous range of applications of the laws of geometric optics. We will proceed as follows:

  1. Reflection and refraction from flat surfaces

  2. Other properties of light propagation (polarization, etc.)

  3. Reflection from a curved surface (mirrors)

  4. Refraction from curved surfaces (lenses)

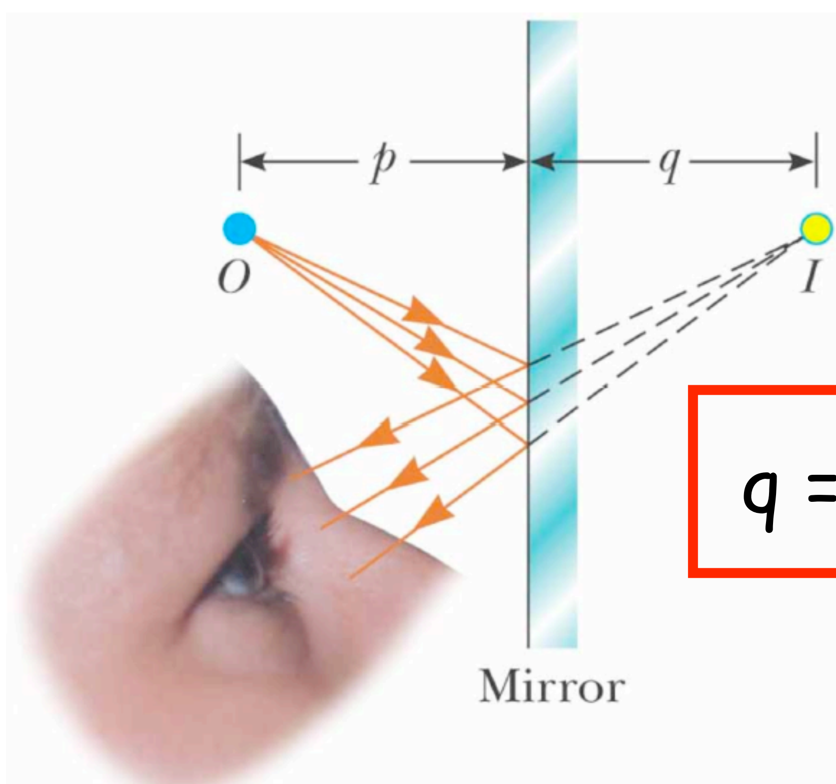
  5. Combinations of mirrors and lenses to form optical instruments:

     - the eye
     - magnifiers, microscopes
     - telescopes

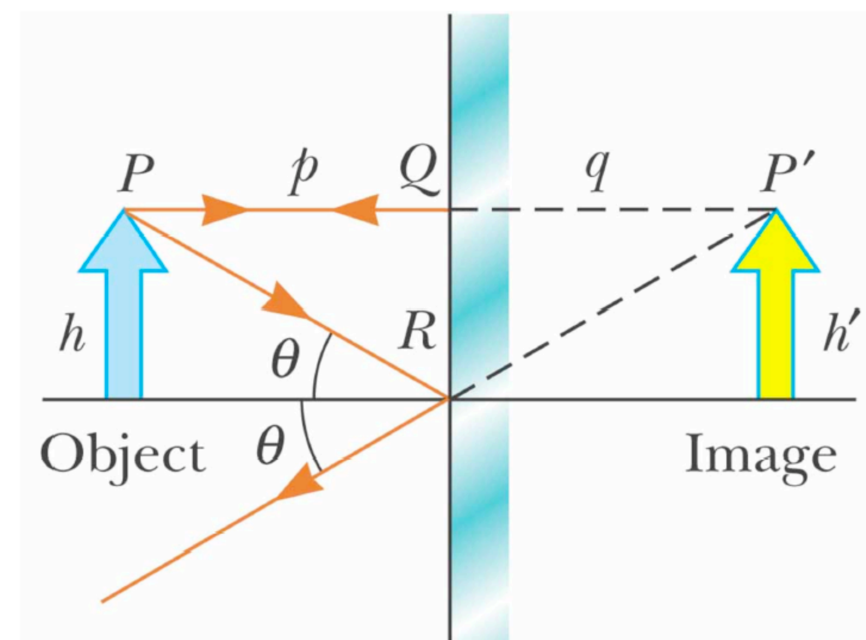
Reflection from flat surfaces

- For a single flat reflecting surface (plane mirror)
  - The reflected light forms an image behind the mirror.
  - The image distance = the object distance
  - The size of the image = that of the object

$h' = h$

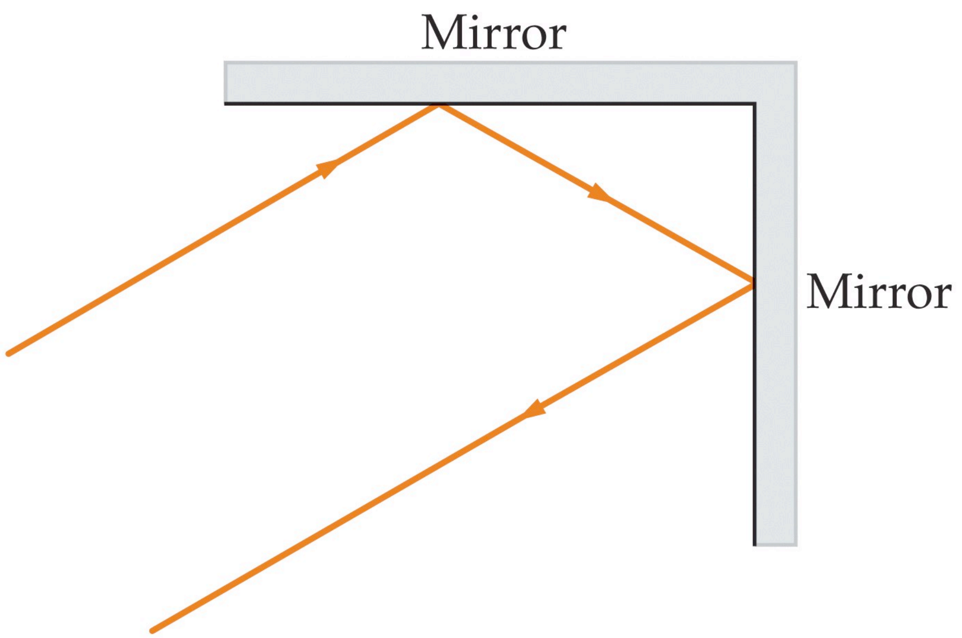


$q = -p$



Retro reflection from a corner cube

- If two flat mirrors are placed at right angles, any incident light will be reflected back along the incident direction.



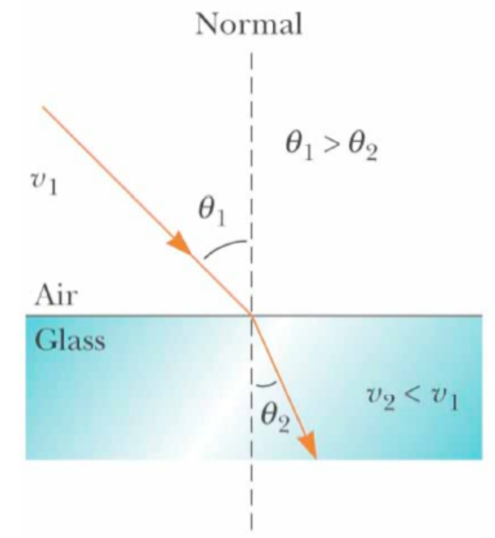
- Using a reflecting corner cube the light from any direction is reflected back upon itself.

- Used in highway to mark road divisions from reflected headlights and as part of car/bicycle rear red reflectors.

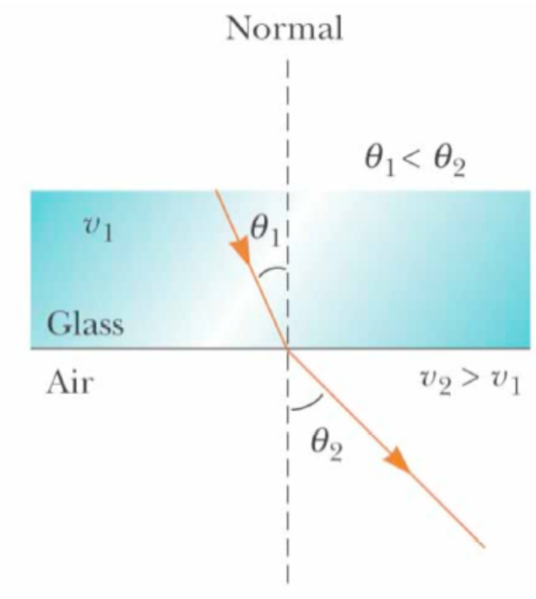
Refraction through a flat surface

- Snell's Law applies at each refracting surface.

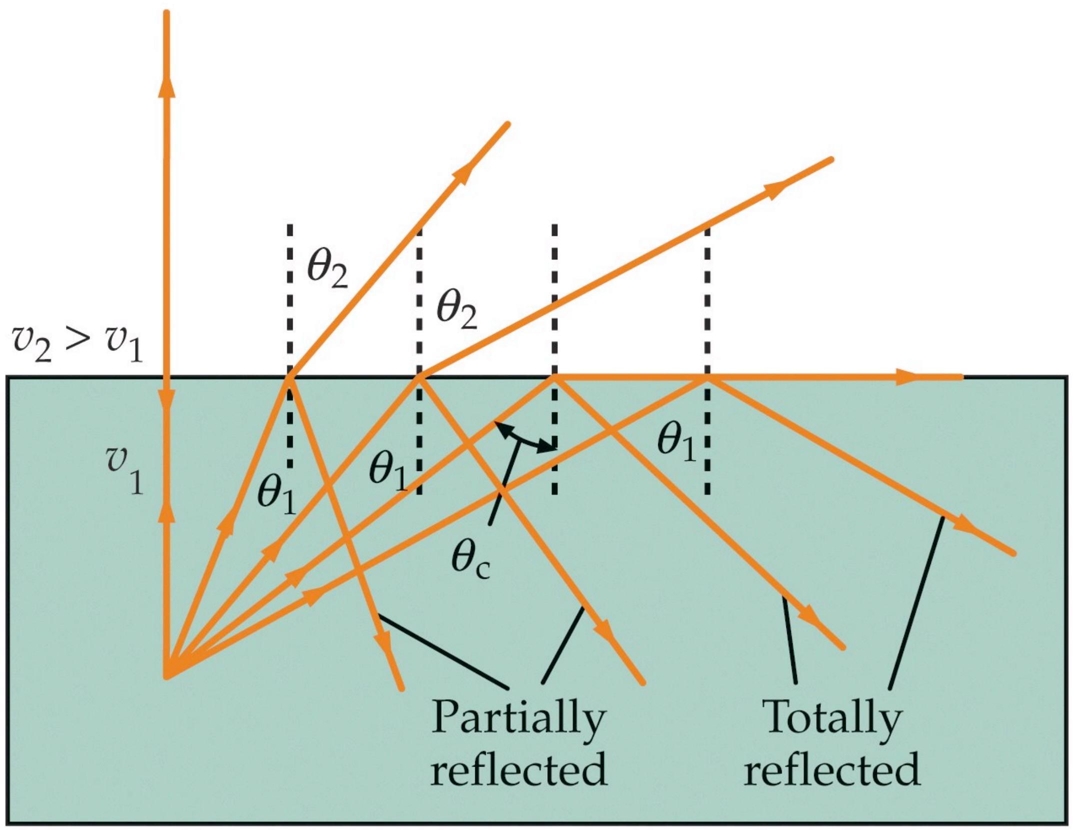
$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

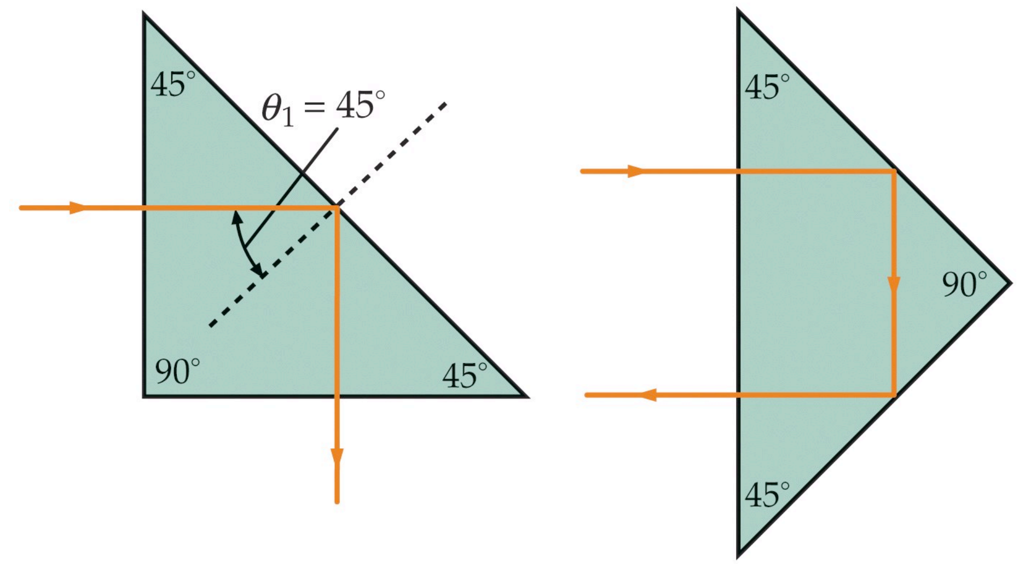


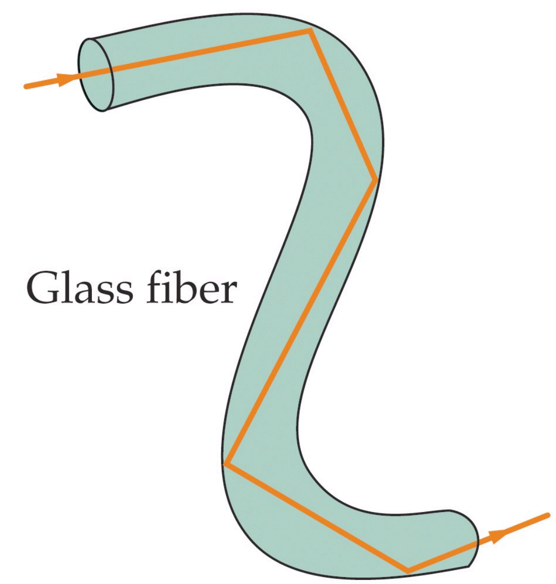
- In the case of light traveling from say glass (n_g) to air ($n = 1$), the refracted light intensity will be zero when $n_g \sin\theta_c = \sin 90^o$
  For $\theta > \theta_c$ there is total internal reflection and the surface acts as a perfect mirror



Total internal reflection

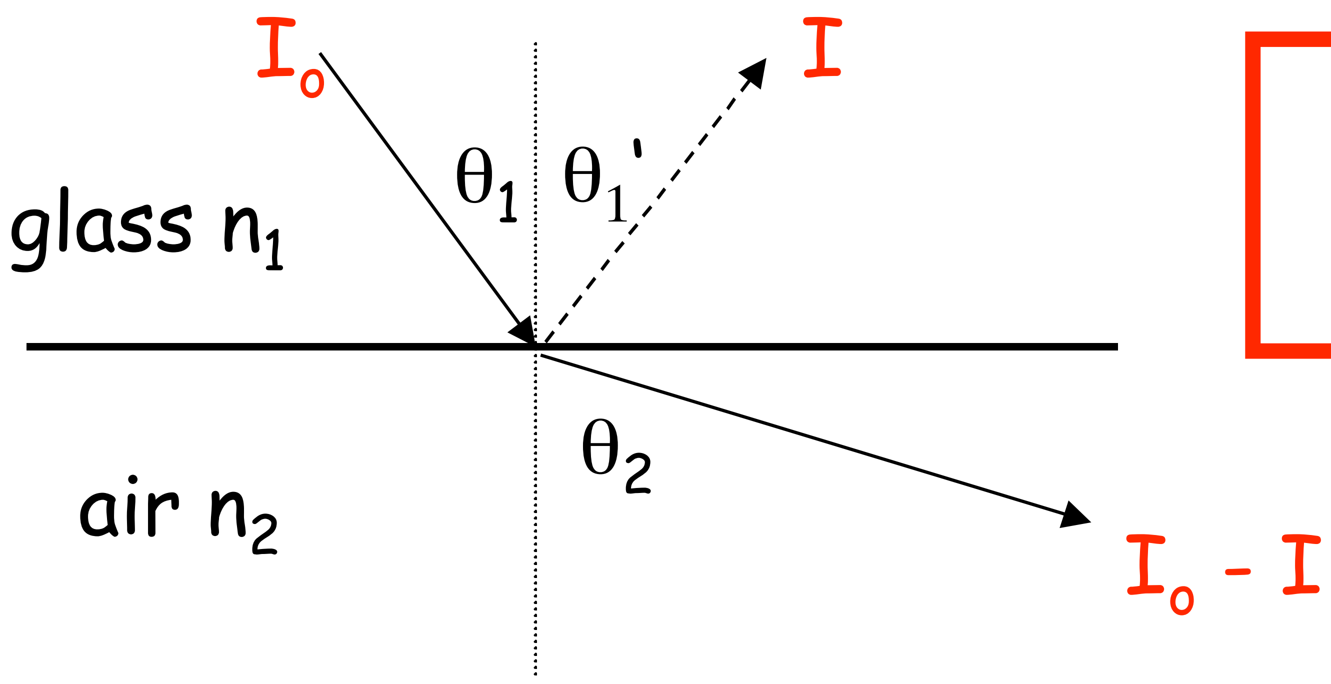






Intensity sharing at an optical boundary

- When light enters an optical surface the intensity is shared between the reflected and refracted beams. The intensity sharing depends both on the angle of incidence and the polarization of the light (more below).
- Consider light emerging from glass to air.

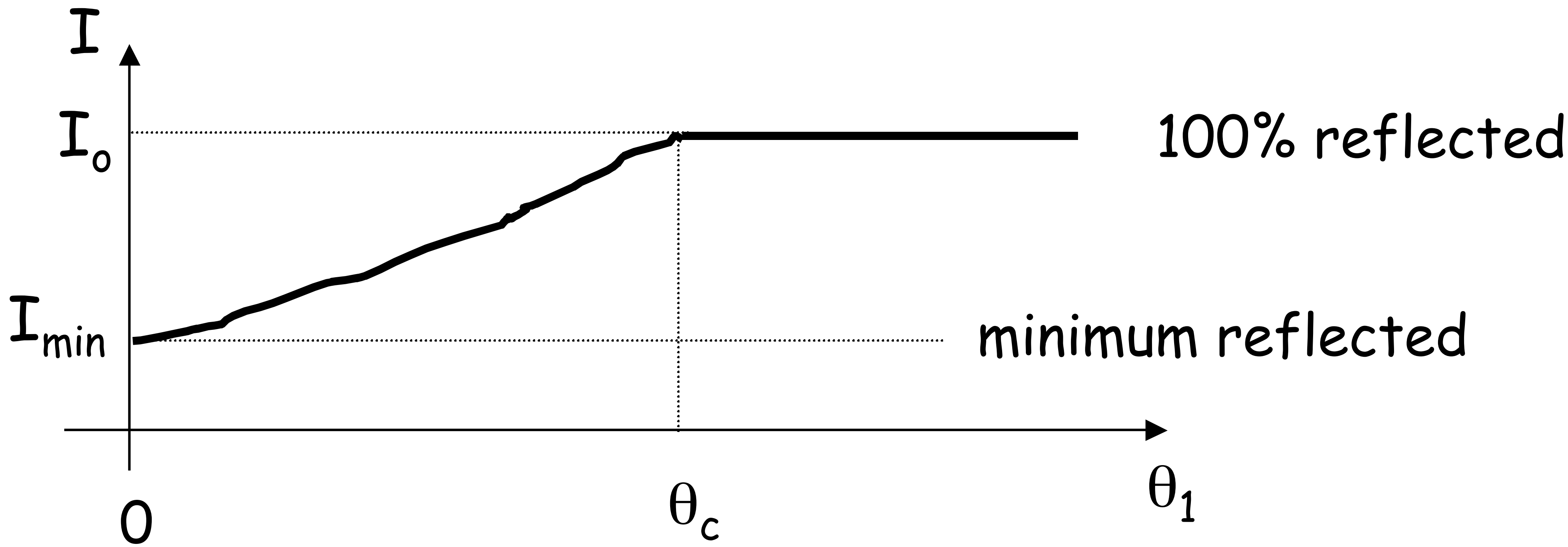


$$\theta_1 = \theta'_1$$

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

Intensity sharing at an optical boundary

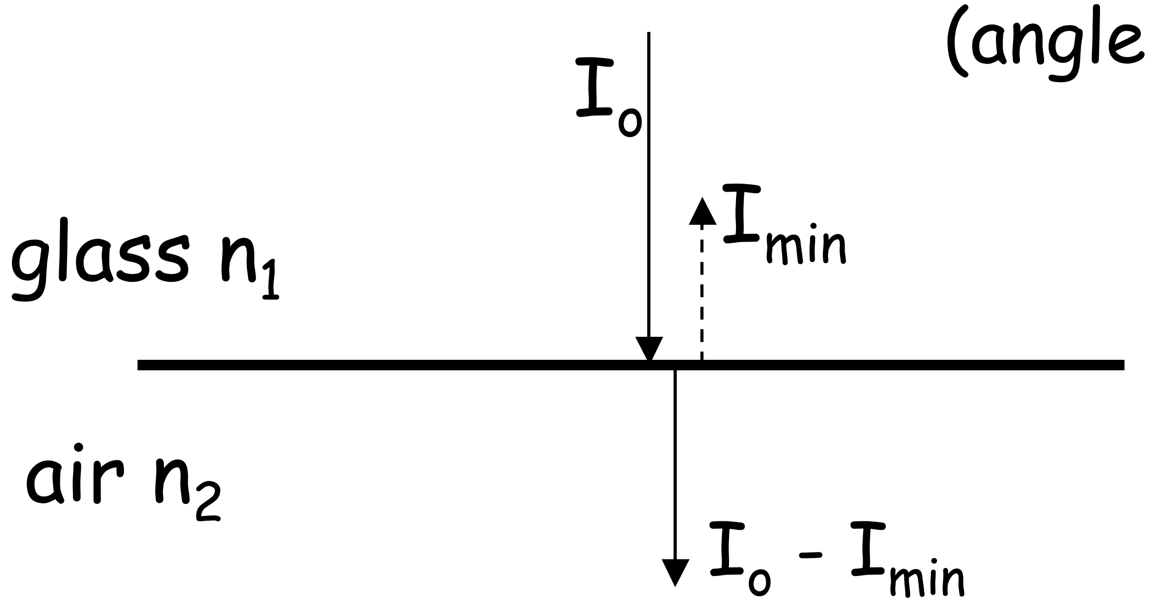
- For the case $n_1 > n_2$ (light from glass into air e.g.) the intensity of the light in the reflected beam (I) varies with the angle of incidence (θ_1) as follows (for unpolarized incident light).



- As discussed previously the angle for total internal reflection is given by: $\theta_c = \sin^{-1}(n_2/n_1)$

Intensity sharing at an optical boundary

- The reflected light has minimum intensity (maximum transmitted) when the light strikes perpendicular to the surface (angle of incidence = θ_1 = 0).

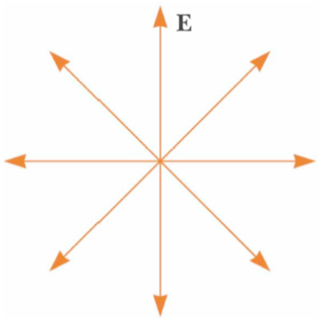


- At normal incidence the intensity of the reflected beam is:

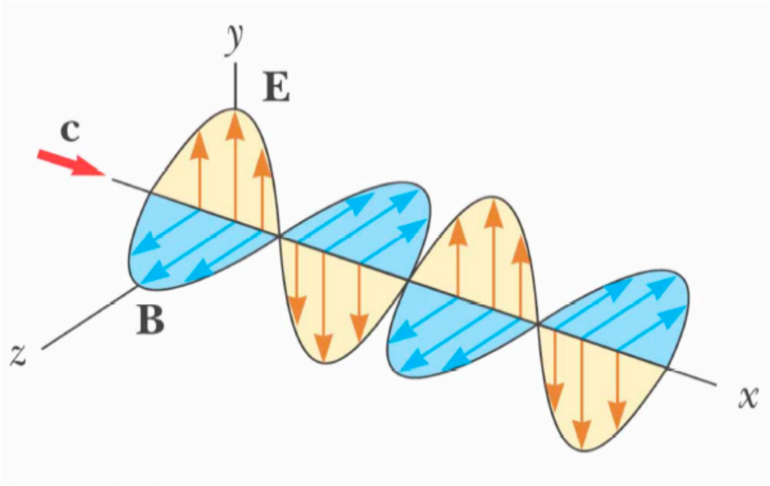
$$I_{min} = [(n_1 - n_2)]/(n_1 + n_2)]^2 \, I_o$$

Polarization of light

- In unpolarized light there is no preferred direction of E field oscillation in the plane perpendicular to the light propagation.



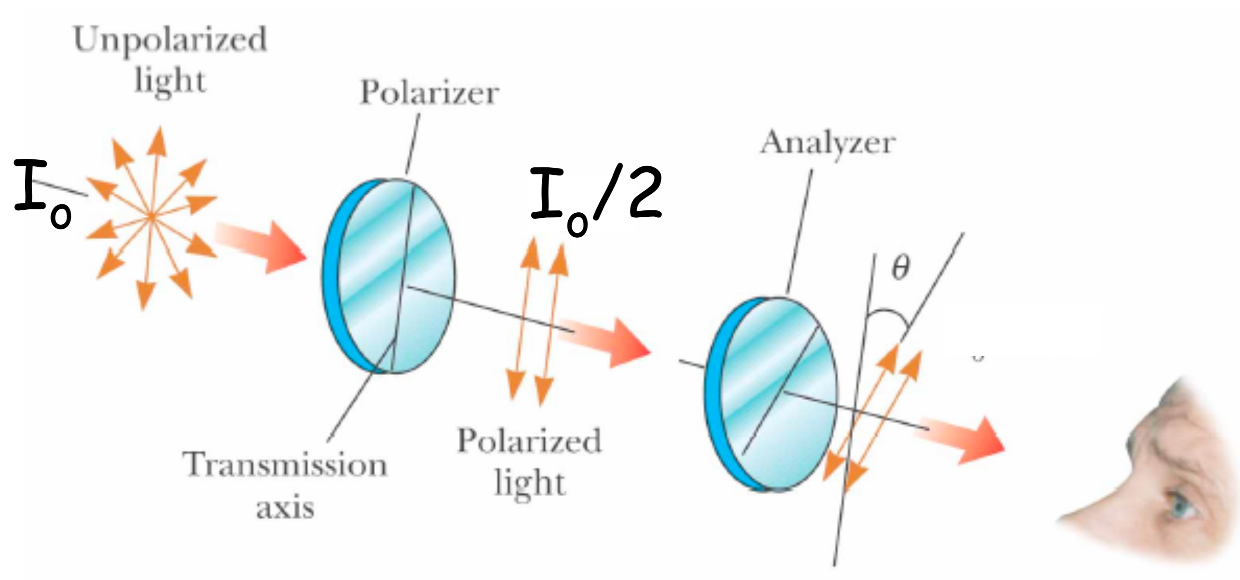
- Linearly polarized light is light prepared so that the electric field vector oscillation is aligned in some direction, say along the y axis.

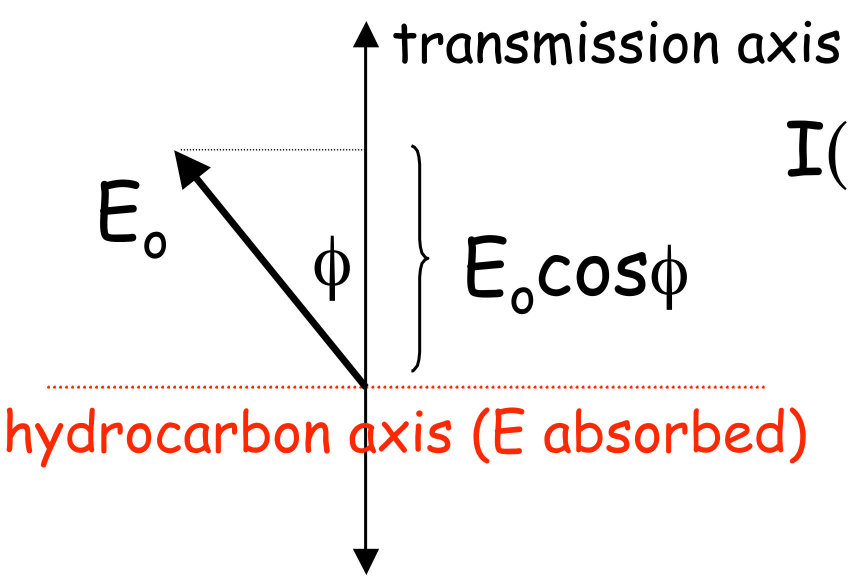


Polarization of light

- Linearly polarized light can be produced in a variety of ways.
- The most direct is to use a material that selectively absorbs light with E field oscillations along a specific axis.
- The light emerging from this material is linearly polarized perpendicular to the absorption axis.
- Materials (polaroid film) produced with long hydrocarbon chains aligned along an axis, can be treated to make the material conducting along this axis.
- Electric field oscillations along the conducting axis are absorbed, and the perpendicular oscillations transmitted.

Linear polarization using polaroid sheets





$$I_o = E_o^2/(2c\,\mu_o)$$

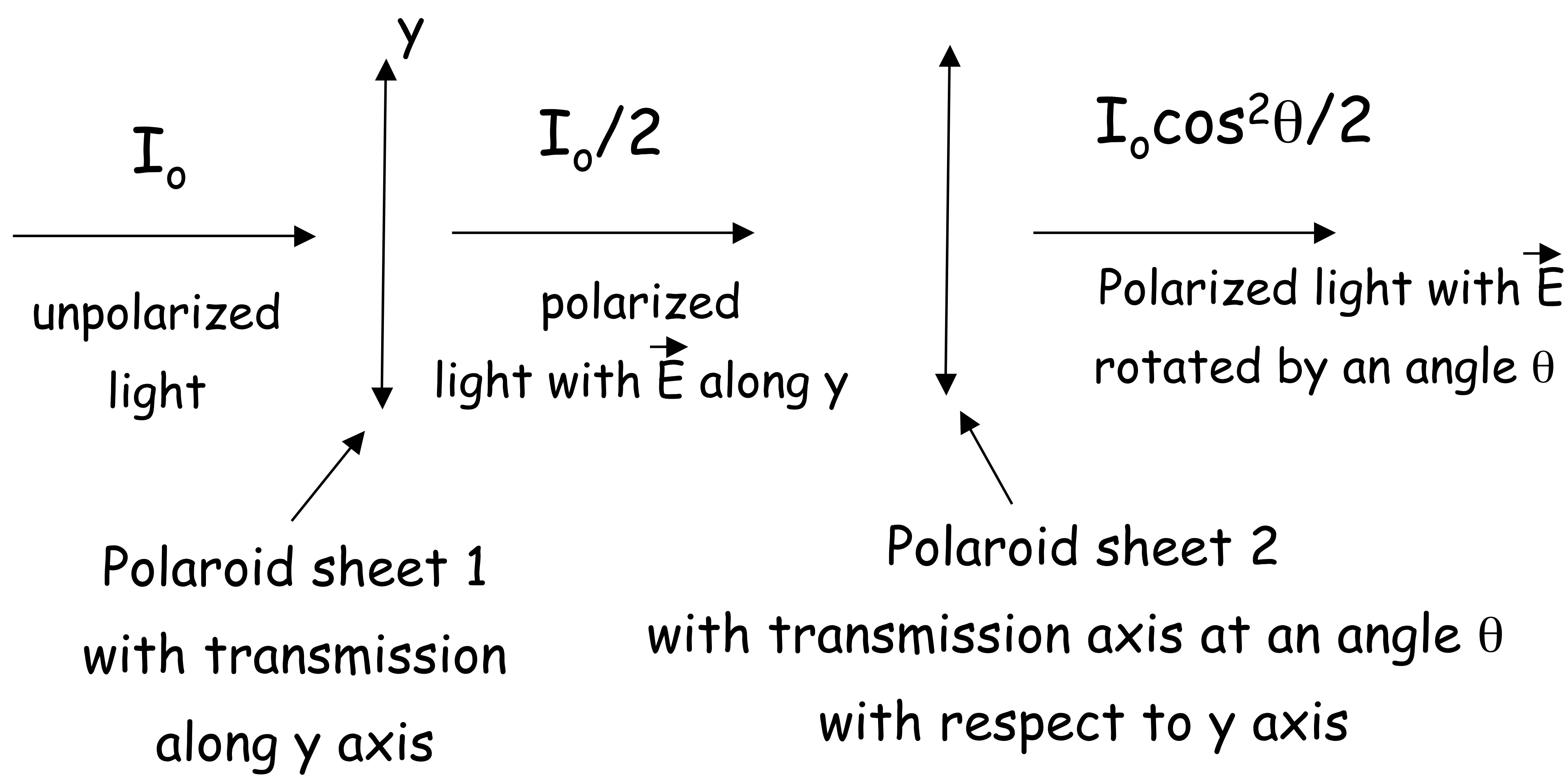
$$I(\phi) = (E_o\cos\phi)^2/(2c\,\mu_o) = I_o\cos^2\phi$$

averaging over ϕ: $\langle\cos^2\phi\rangle = 1/2$

$$I_{transmitted} = I_o/2$$

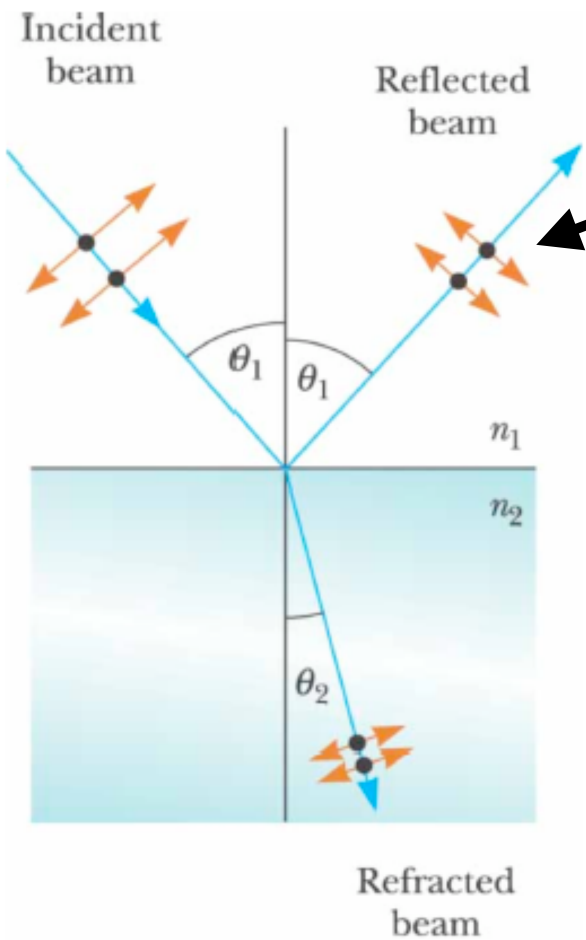
Linear polarization using polaroid sheets

- Using two polaroid sheets, the intensity of light can be continuously varied by rotating the second sheet with respect to the first.



Polarization by reflection

- Another way to produce polarized light is by reflection from a boundary between two optical media (say air and glass).

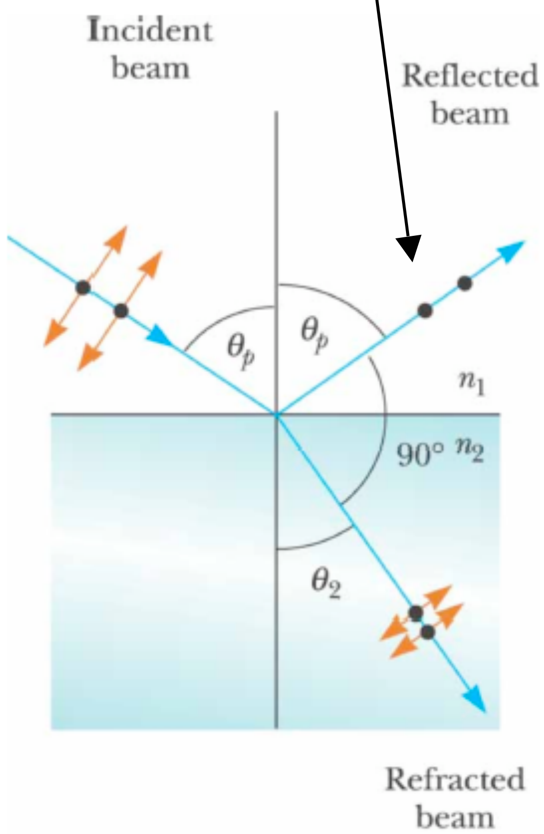


E field components parallel and perpendicular (•) to the plane of incidence

The E field vectors parallel to the surface boundary (•) are preferentially reflected

Polarization by reflection

- At a certain angle, the reflected light beam is 100% linearly polarized parallel to the surface.
- This occurs when the angle between the reflected and refracted rays is 90^o as shown in the figure.



$$n_1 \sin\theta_p = n_2 \sin\theta_2 = n_2 \sin(90^o - \theta_p)$$

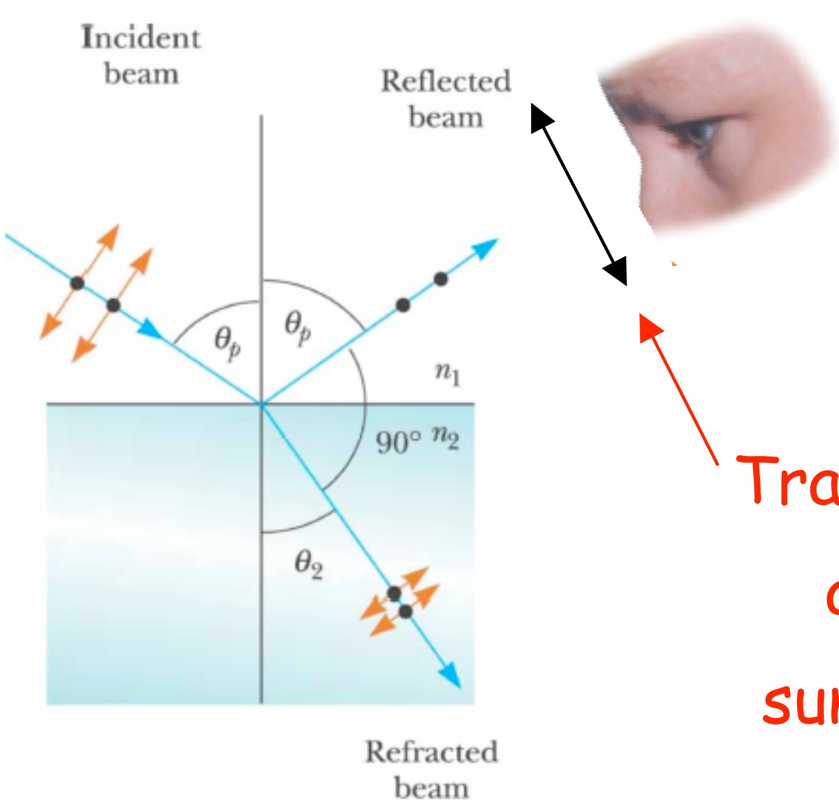
$$n_1 \sin\theta_p = n_2 \cos\theta_p$$

$$\Rightarrow \quad \tan\theta_p = n_2 / n_1$$

The polarizing angle θ_p is referred to as Brewster's angle.

Polarization by reflection

- The fact that light reflected at an angle near Brewster's angle is ~100% polarized can be used to reduce glare with polaroid sun glasses.
- Orientate the transmission axis of the polaroid to be vertical, thus blocking the horizontally polarized reflection



Transmission axis of sun glasses

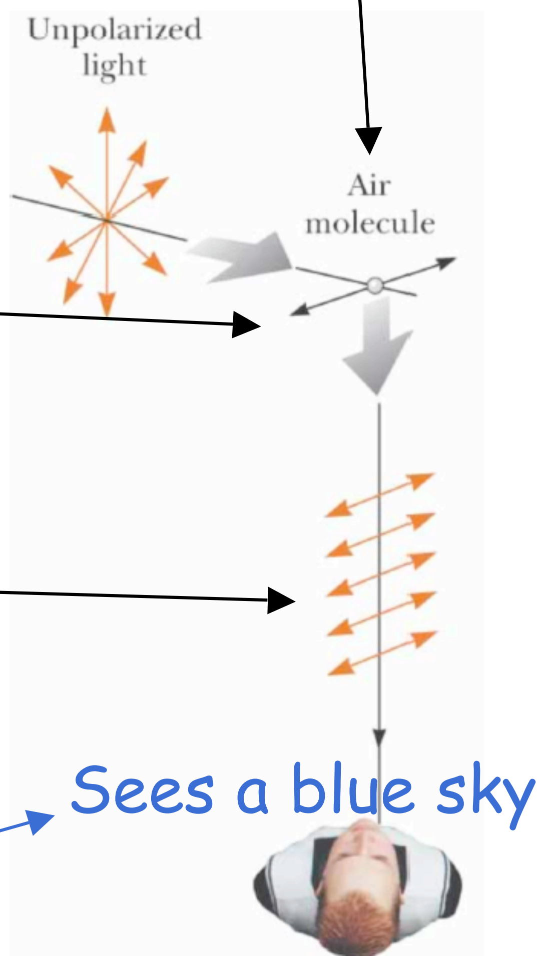
$$\tan \theta_p = n_2 / n_1$$

air -> glass $\theta_p = 56°$

air -> water $\theta_p = 53°$

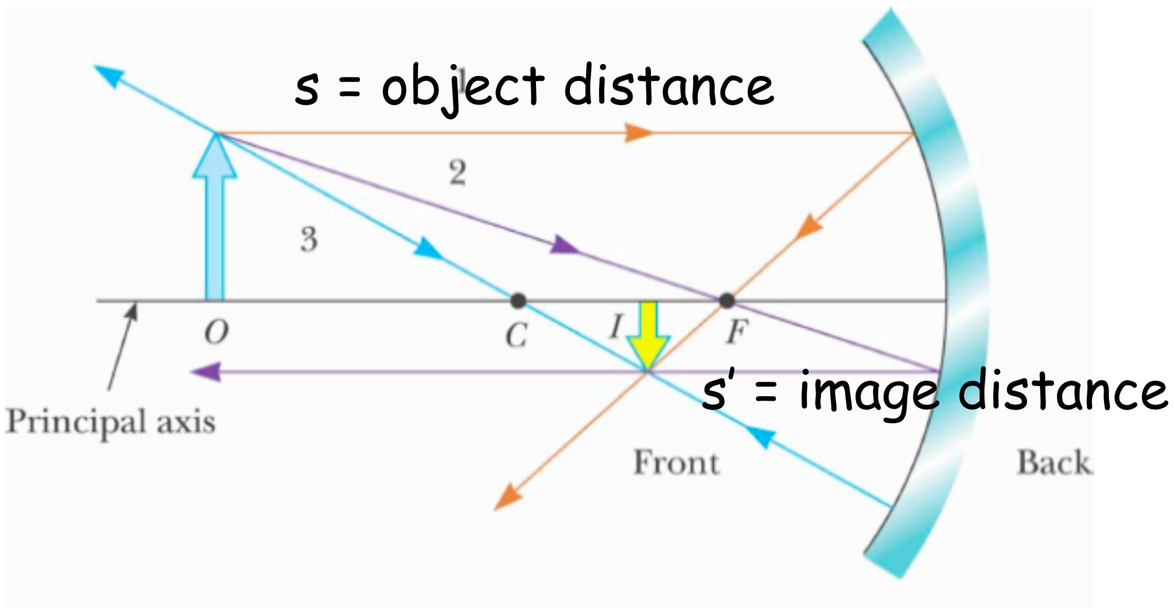
Polarization by scattering

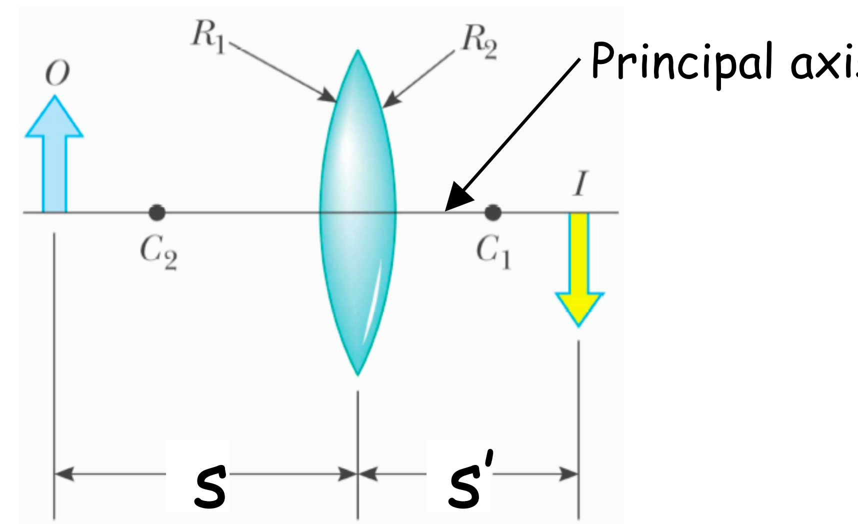
- A third way to create polarized light is by scattering off molecules in a gas or liquid.
- The horizontal component of the E field oscillations excite a horizontal oscillation of the charges in the molecules.
- This radiates (as in a dipole antenna) preferentially perpendicular to the oscillations => vertically downward.
- This light has two characteristics:
  1. It is linearly polarized as shown
  2. The intensity ~ ω^4 ~ $1/\lambda^4$, therefore blue (shorter λ) is scattered more than red light.



Preview

- We will obtain the equations that describe the images formed by spherical mirrors and thin lenses.





- This is done under the assumption that all rays make a small angle with respect to the principal axis of the mirror/lens:

θ < ~ 0.1 radian => sinθ ~ tanθ ~ θ (paraxial rays)

Preview and notation

- The location and character of images can then be calculated form the laws of reflection and refraction.
- We will use the following notation:

  s = object distance $\quad$ s' = image distance

  y = the object height $\quad$ y' = the image height

  f = the focal length of the mirror or lens

  R = the radius of curvature of the spherical surface
- Then for a mirror or a thin lens, for paraxial rays:

  $1/s + 1/s' = 1/f$ and M = the magnification = $y'/y = -s'/s$

  The sign convention is important as we will discuss.

Preview and terminology

- The image formed by an object has two important characteristics.

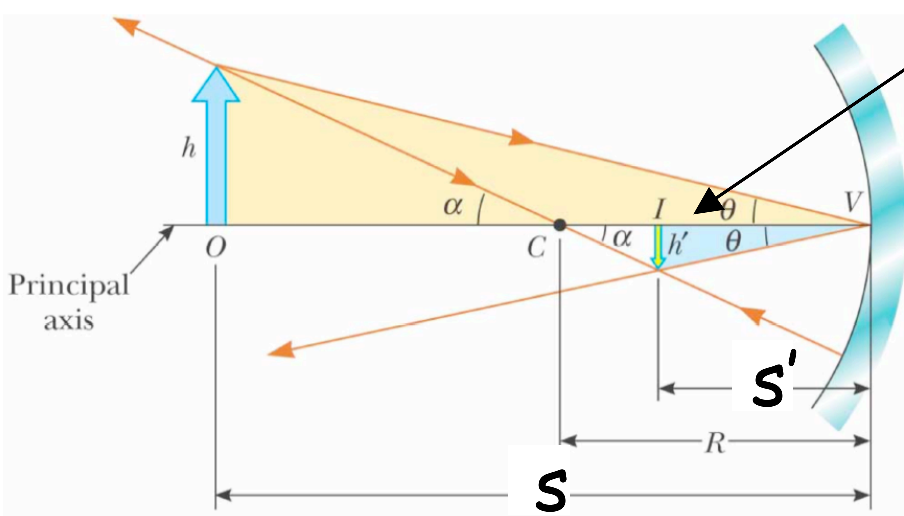
  1. The transverse size of the image relative to the object

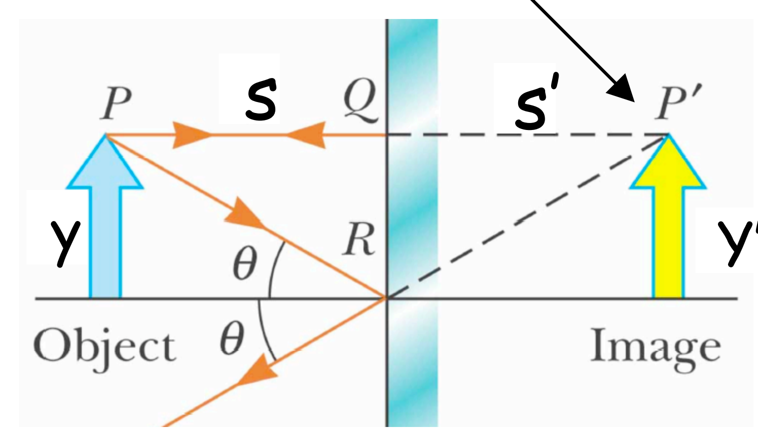
     M = the magnification = y′/y (upright or inverted)

  2. The image formed can be real or virtual

     light rays pass through a real image

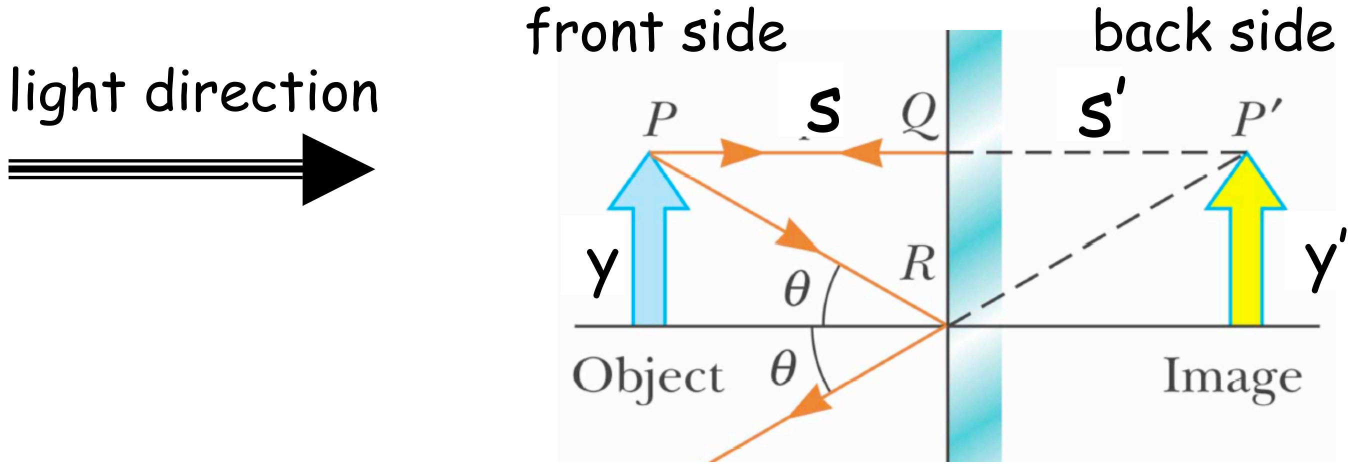
     extrapolation of light rays pass through a virtual image





A plane mirror

- A plane (flat) mirror forms an image on the back side of the mirror.



Sign convention:

s and s' are

+ on the front side

- on the back side

- For an object on the front side the image is always virtual and upright:

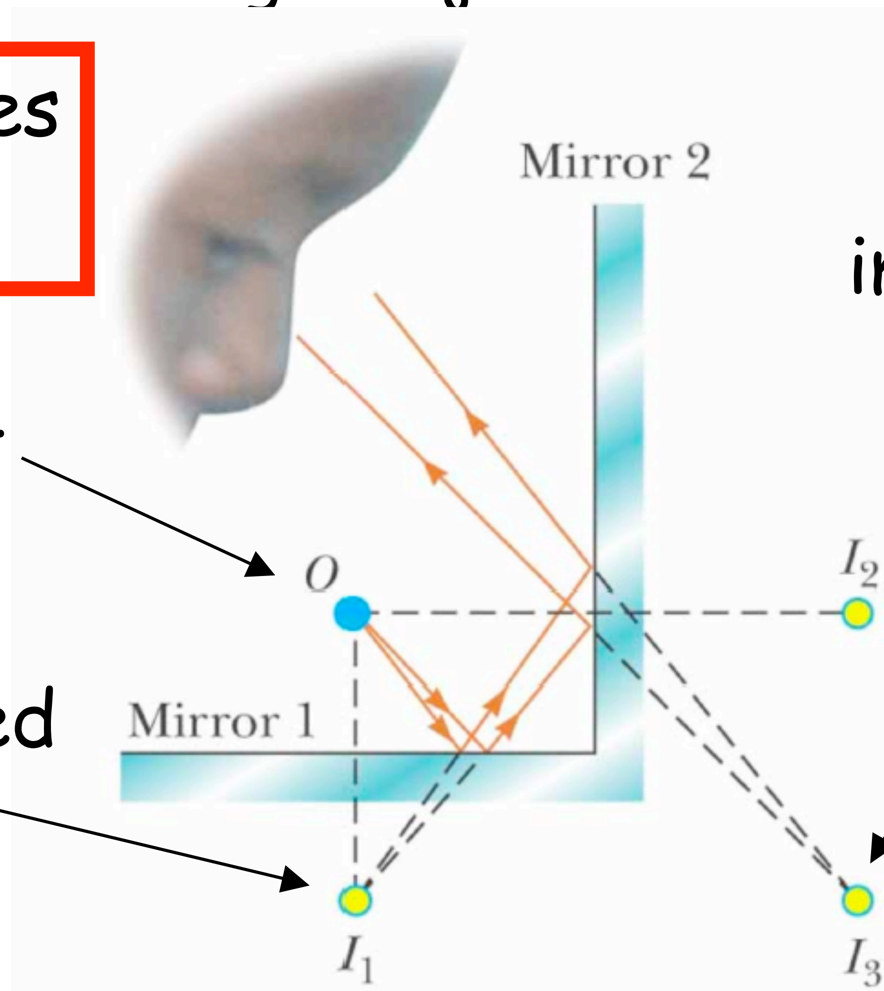
$$s' = -s$$

$$M = y'/y = +1$$

Plane mirrors

- Multiple images of an object are formed by two plane mirrors.
- Consider two at right angles as shown. There will be three virtual images of a single object.

observer sees 3 images



object

image I_1 formed in mirror 1

image I_2 formed in mirror 2

image I_3 of I_2 in mirror 1 and I_1 in mirror 2

Spherical mirrors: the basics

- This analysis will use only " $\theta_1 = \theta_1'$ " and will be valid for paraxial rays forming the image.
- The mirror is described by a single parameter:

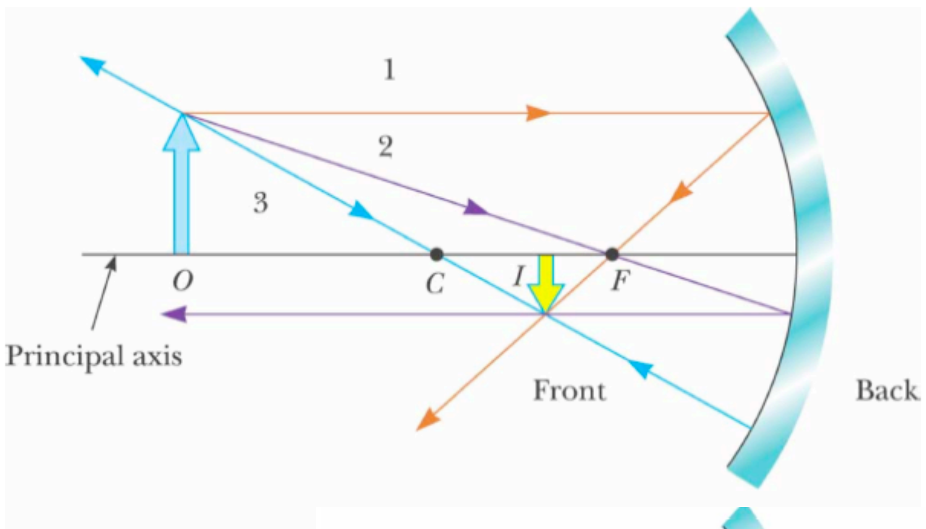
  R = the radius of curvature of the mirror surface
- The sign convention is the following:

  s, s' and R are + (-) on the front (back) side of the mirror

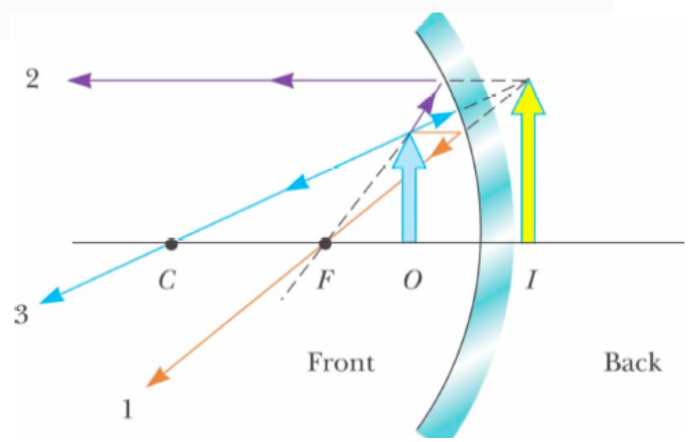
  y and y' are + (-) if upright (inverted)
- The problem is the following:

  Given R (the mirror) and s and y (the object description) what are s' and y' (the image description)?
- Note that the image is real (virtual) for q positive (negative)

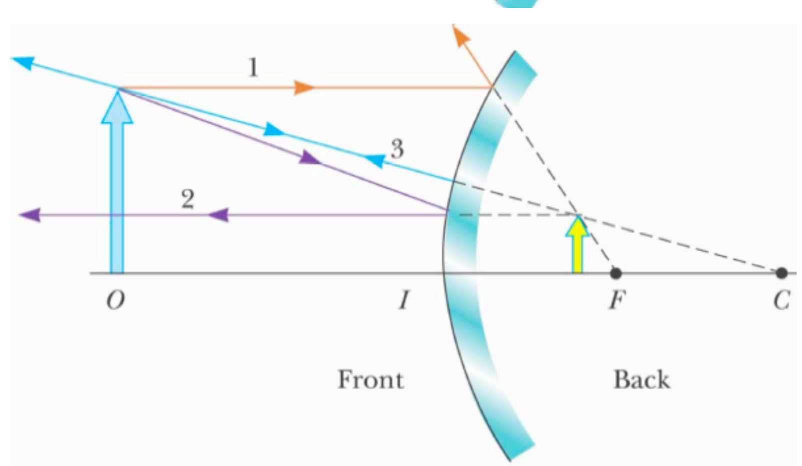
Spherical mirrors: examples



Concave mirror:
real, inverted image



Concave mirror:
virtual, upright image

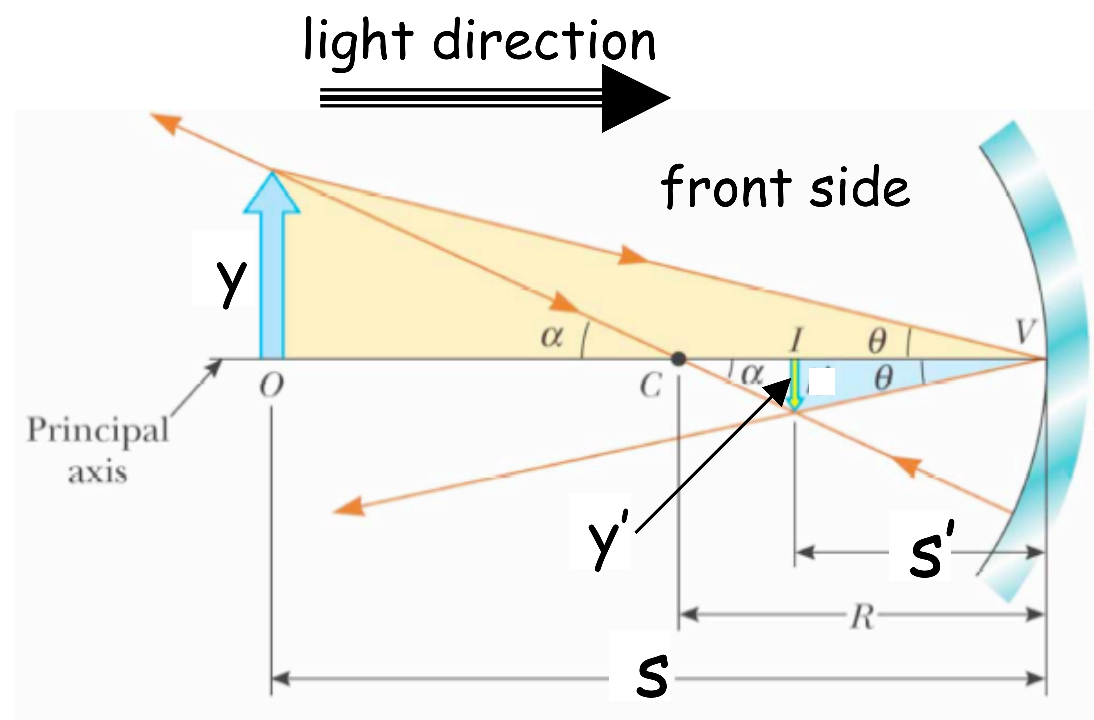


Convex mirror:
virtual, upright image

Spherical mirrors: the details

$\tan\theta = y/s = -y'/s'$

Magnification $M = y'/y = -s'/s$



$\tan\alpha = y/(s-R) = -y'/(R-s')$

$-y'/y = (R-s')/(s-R)$

Substitute $y'/y = -s'/s$

$\Rightarrow s'/s = (R-s)/(s-R)$

Solve to obtain:

$1/s + 1/s' = 2/R$

Valid for any paraxial ray

In this example:

s and s' are +

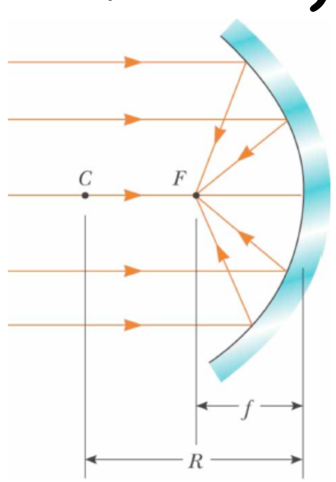
y is + and y' is -

Spherical mirrors: the details

- The derivation was done for the special case of a concave mirror with a real object and image.
- It is not difficult to show that the result is completely general:

  $M = y'/y = -s'/s$ and $1/s + 1/s' = 2/R$

  for any object and image combination for concave and convex mirrors (using the sign convention defined above).
- Note also that if the object is at infinity (parallel rays hitting the mirror) then the image is formed at R/2



Define $R/2 = f$ = focal length of the mirror

Then $1/s + 1/s' = 1/f$

Summary: the mirror equation for paraxial rays

- For spherical mirrors with radius R:

$$1/s + 1/s' = 1/f$$

$$M = \text{transverse magnification} = y'/y = -s'/s$$

where s (s') = object (image) distance

y (y') = object (image) height

f = focal length = R/2

- The sign convention is:

  s, s', R are + (-) on the front (back) side of the mirror

  y and y' are + (-) if upright (inverted)

Ray diagrams for mirrors

- It is often useful to use some simple rays to geometrically reconstruct the location and character of an image.

  Ray 1: a ray entering parallel to the axis will be reflected to pass through the focal point of the mirror

  Ray 2: a ray passing through the focal point will be reflected parallel to the principal axis

  Ray 3: a ray passing through the center of curvature will be reflected back upon itself.

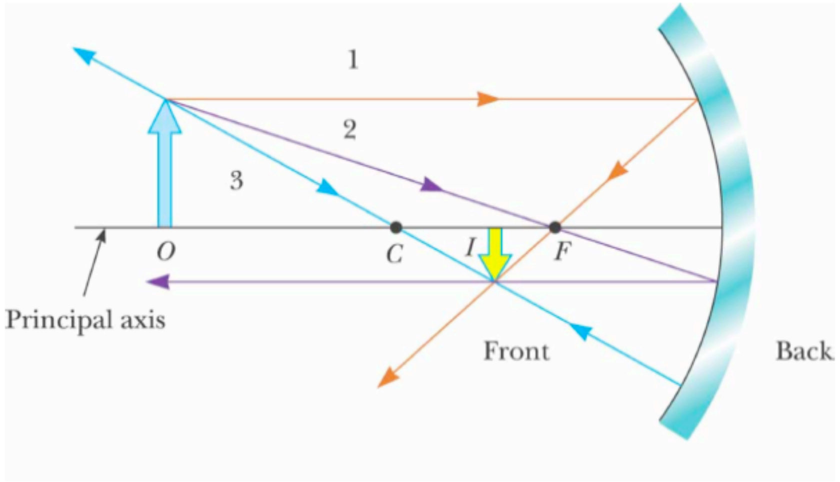
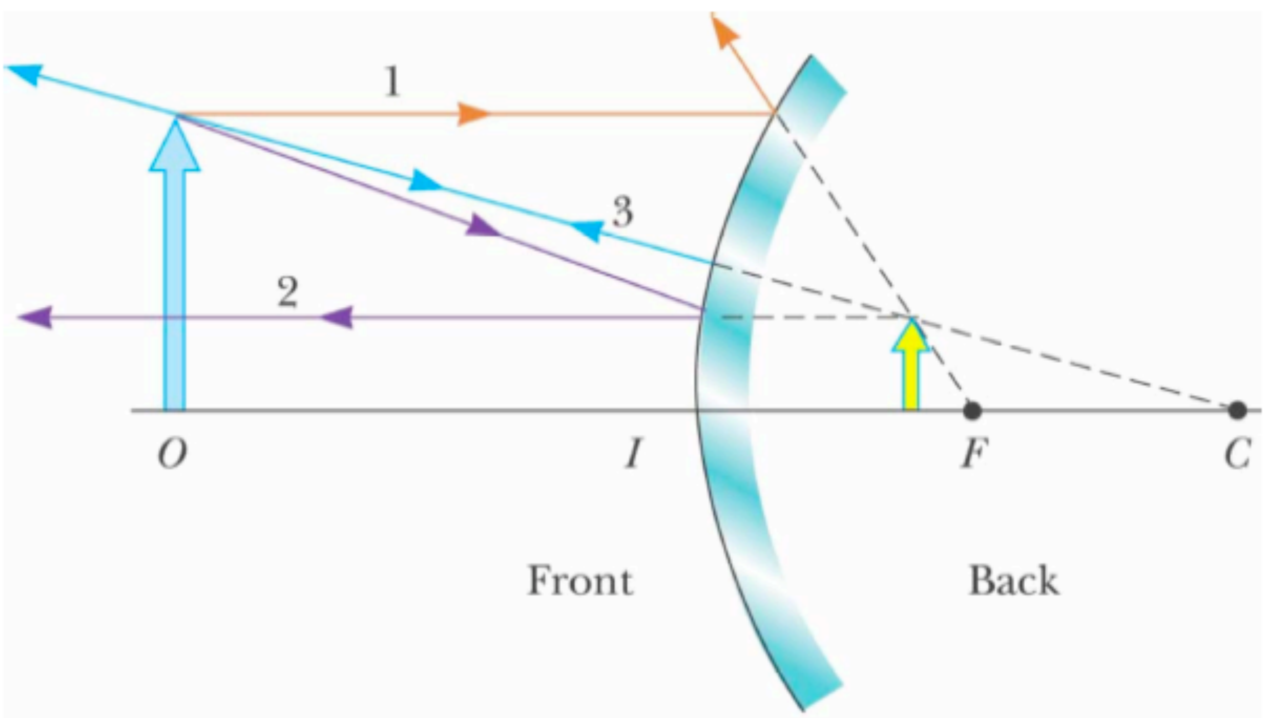


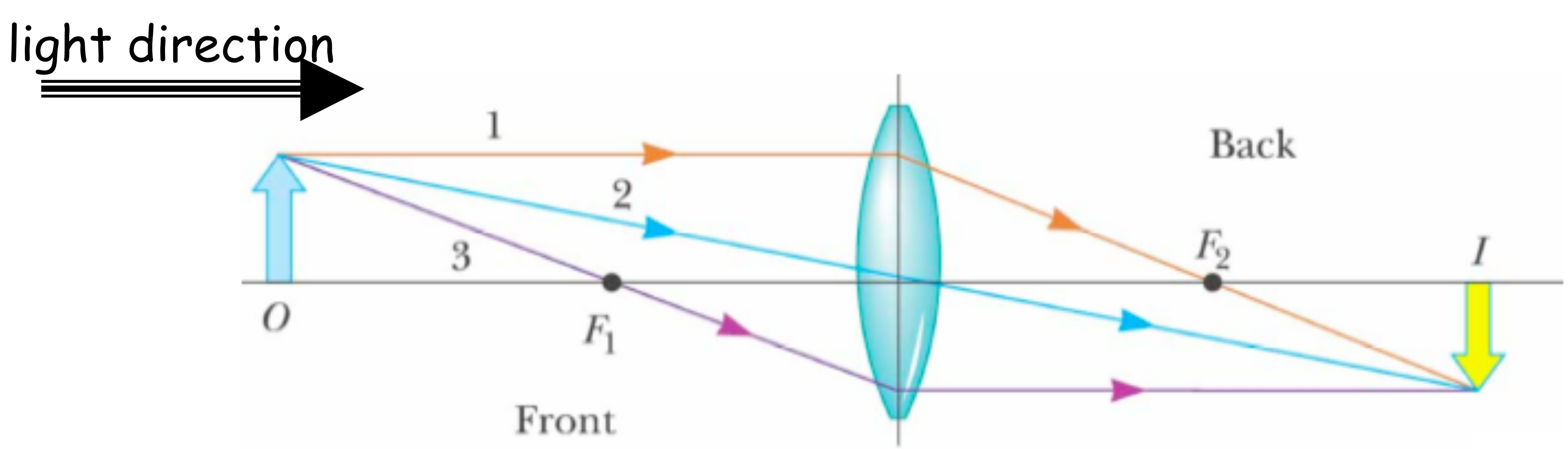
Image reconstruction using simple rays



Convex mirror:
virtual, upright
reduced, image

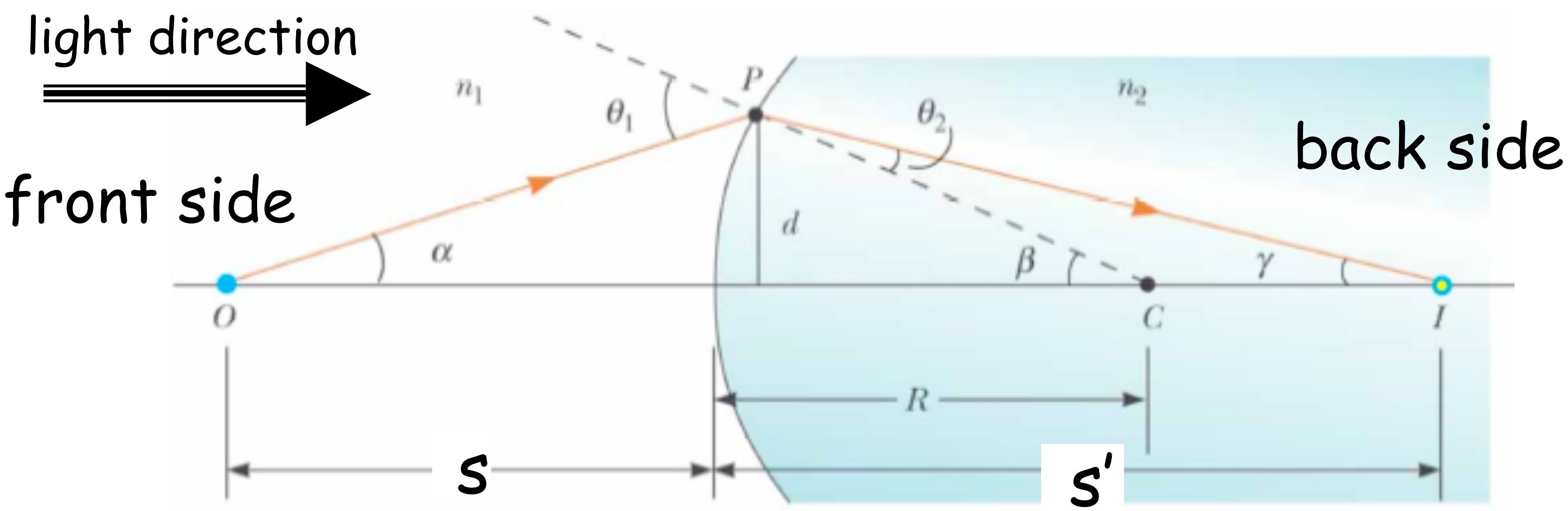
Thin lenses: the basics

- The terminology and conventions applied to mirrors will be used for images formed by lenses.
- In this case the light from an object passes through two refracting surfaces to form the image.
- We assume the rays are paraxial so that we can use $\sin\theta \sim \theta$ and therefore Snell's law is $n_1 \theta_1 \sim n_2 \theta_2$ at each surface.



Refraction from one spherical surface

- First the sign convention.



- s is + (-) on the front (back) side
- y and y' are + (-) if upright (inverted)

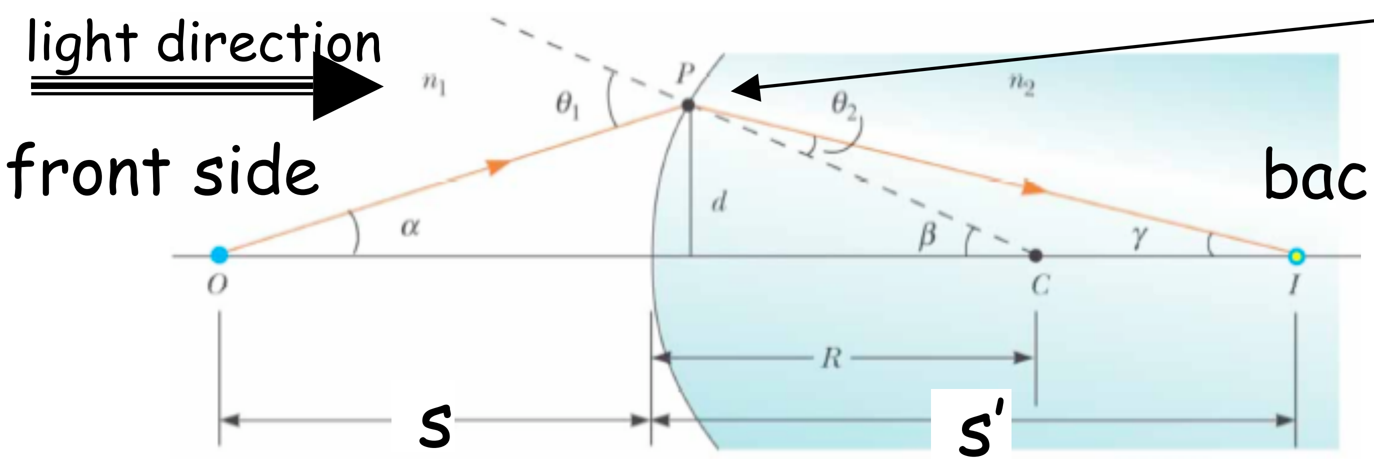
} same as for mirrors

- s' is + (-) on the back (front)side of the surface
- R is + (-) on the back (front) side of the surface

} opposite to mirrors

Refraction from one spherical surface

- Now calculate the image distance s′



$$n_1 \theta_1 = n_2 \theta_2$$

$$\theta_1 = \alpha + \beta$$

$$\theta_2 = \beta - \gamma$$

- Substitute for θ_1 and θ_2:

$$n_1 (\alpha + \beta) = n_2 (\beta - \gamma)$$

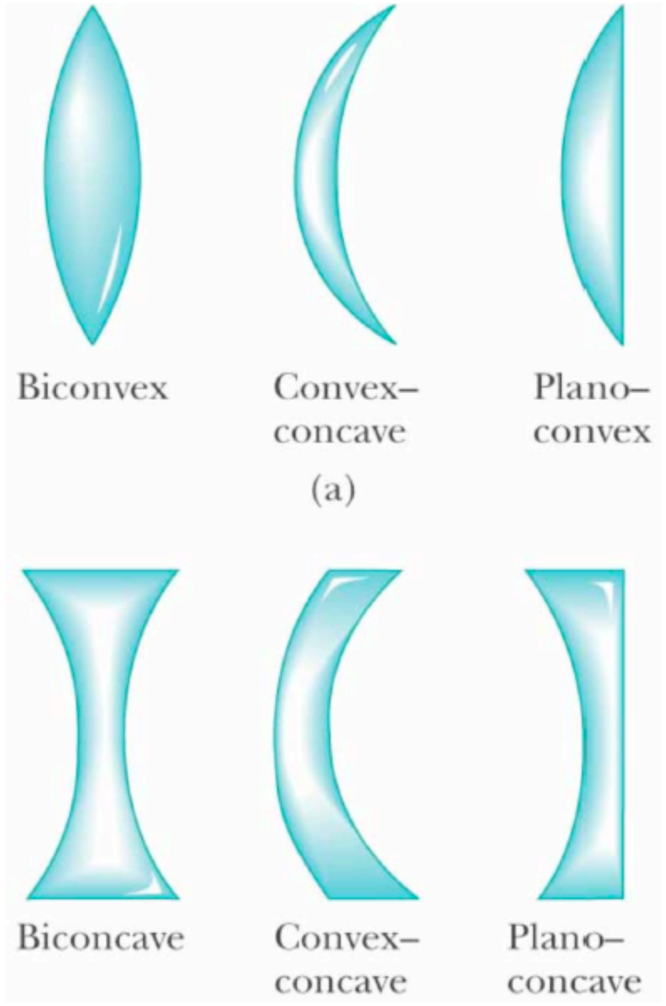
$$n_1 \alpha + n_2 \gamma = (n_2 - n_1) \beta$$

- For paraxial rays: $\tan\alpha = \alpha = d/s$; $\tan\beta = \beta = d/R$; $\tan\gamma = \gamma = d/s'$

$$n_1/s + n_2/s' = (n_2 - n_1)/R$$

A thin lens: two refracting surfaces

- A lens forms an image as the result of refraction from two surfaces.
- There are six different basic lenses as shown



- For paraxial rays and thin lenses (see derivation below):

$$1/s + 1/s' = 1/f$$

$$M = - s'/s$$

A thin lens: two refracting surfaces

Surface 1

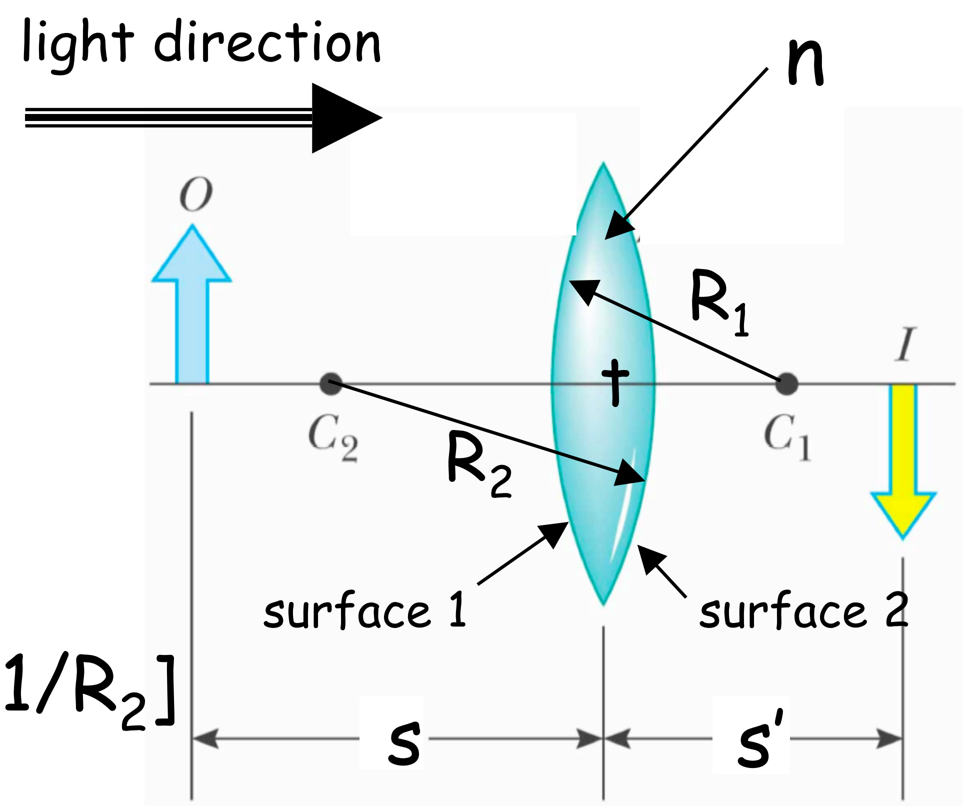
$$1/s + n/s_1' = (n-1)/R_1$$

Surface 2

$$n/s_2 + 1/s' = (1-n)/R_2$$

Add above 2 equations

$$1/s + 1/s' + [n/s_1' + n/s_2] = (n-1)[1/R_1 - 1/R_2]$$



Also $s_2 = t - s_1'$; for a thin lens $t \sim 0$ and $s_2 = - s_1'$

$$1/s + 1/s' = (n-1)[1/R_1 - 1/R_2]$$

The thin lens equation

Images formed by a thin lens

- As for a mirror we define f as the point where parallel rays are focused. f can be + or - depending on the shape of the lens.

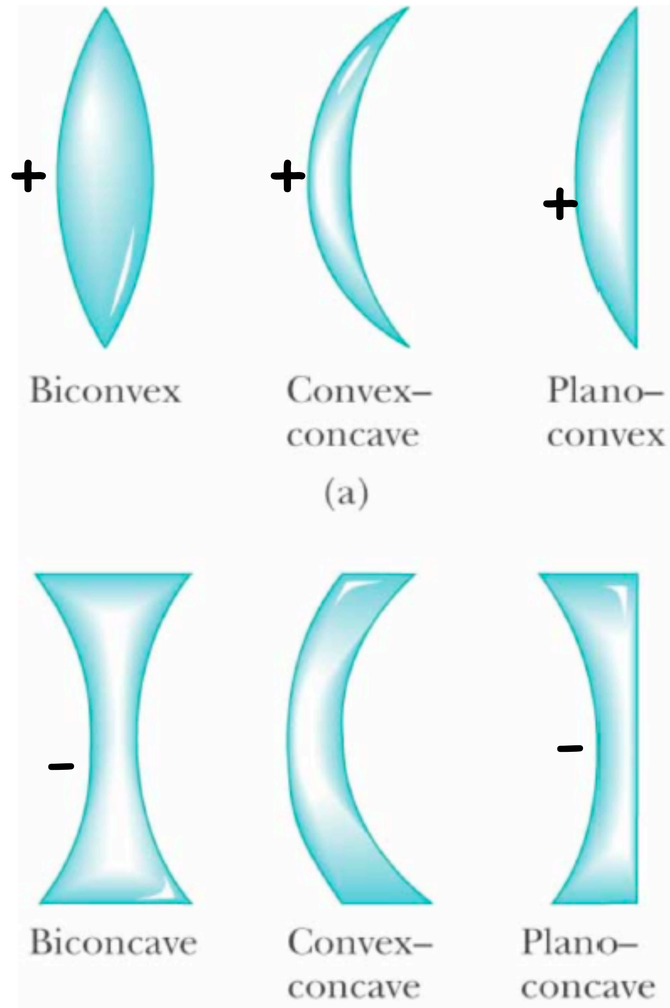
$$1/f = (n-1)[1/R_1 - 1/R_2]$$

Positive f => converging lens

Negative f => diverging lens

- The thin lens equation becomes

$$1/s + 1/s' = 1/f$$



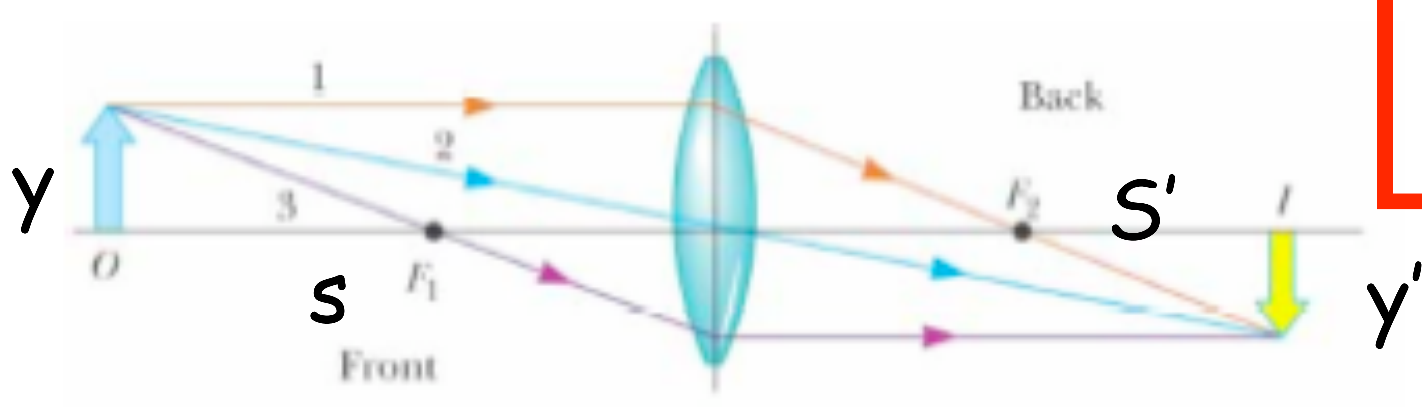
Ray diagrams for lenses

- As for mirrors a few simple light rays can be used to reconstruct the location and character of an image.

  Ray 1: a ray entering parallel to the axis will be refracted to pass through the focal point of the lens.

  Ray 2: a ray passing through the center of the lenses will be undeflected.

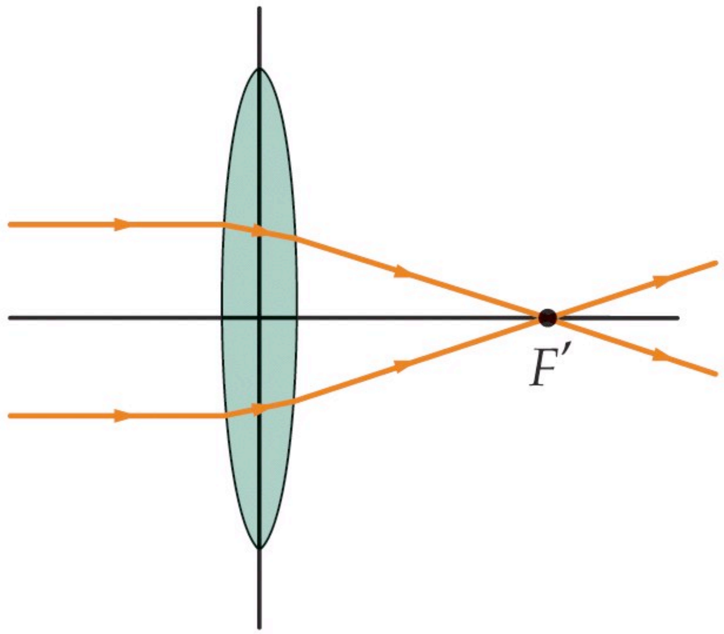
  Ray 3: a ray passing through the focal point will be refracted parallel to the axis.

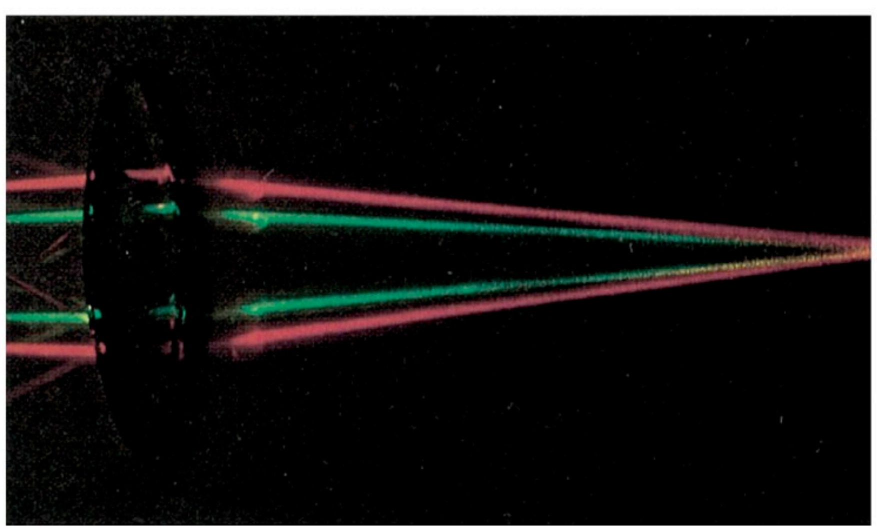


$$M = y'/y = -s'/s$$

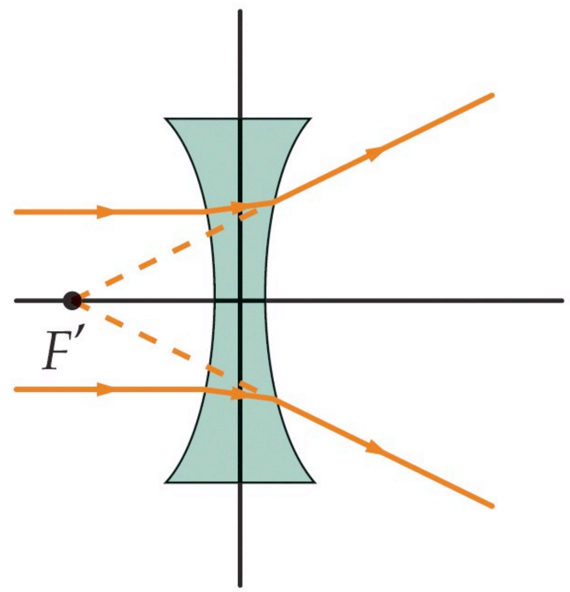
Converging and Diverging Lenses

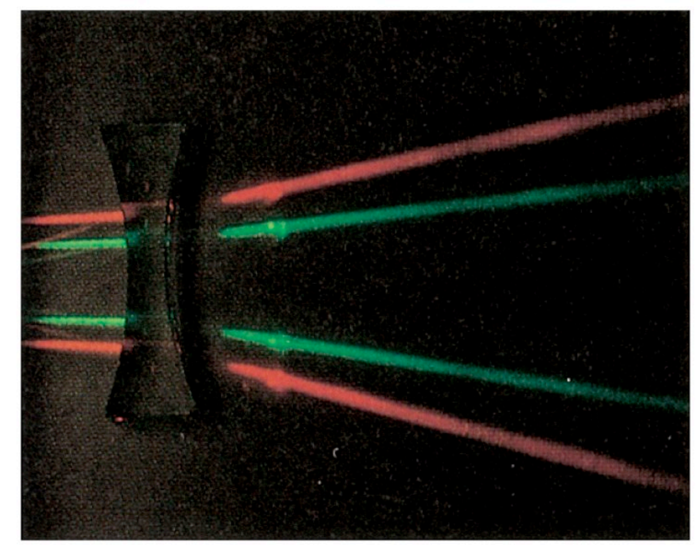
Converging or
Positive lens (f +)





Diverging or
Negative lens (f -)





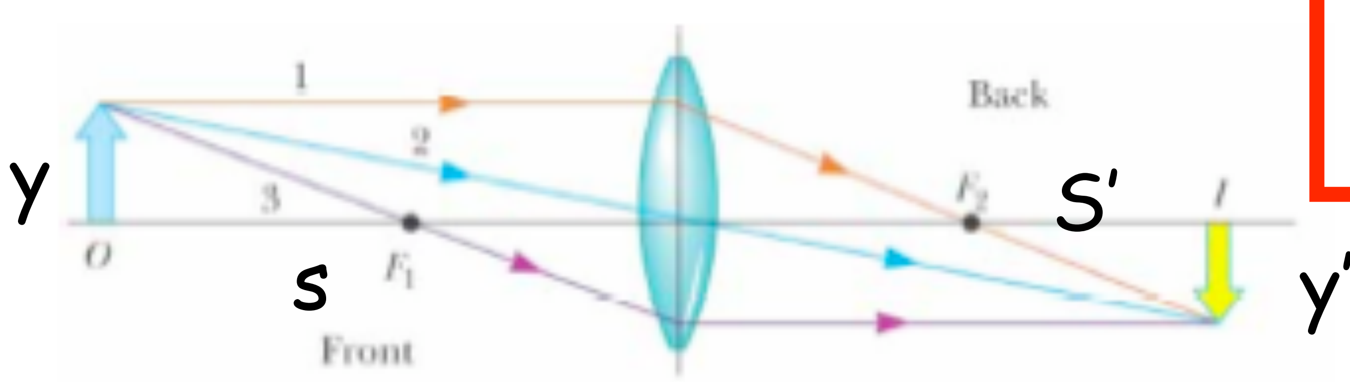
Ray diagrams for lenses

- As for mirrors a few simple light rays can be used to reconstruct the location and character of an image.

  Ray 1: a ray entering parallel to the axis will be refracted to pass through the focal point of the lens.

  Ray 2: a ray passing through the center of the lenses will be undeflected.

  Ray 3: a ray passing through the focal point will be refracted parallel to the axis.



$$M = y'/y = -s'/s$$

Images formed by multiple optical elements

- Optical instruments generally use multiple lenses and/or mirrors to improve the quality of the image.
- The determination of the location and character of the image formed involves repeated application of the basic mirror and thin lens equations.
- We will start with some simple examples, and then move on to specific optical instruments (microscopes, telescopes, etc.)

Images formed by multiple optical elements

- The approach is the following:

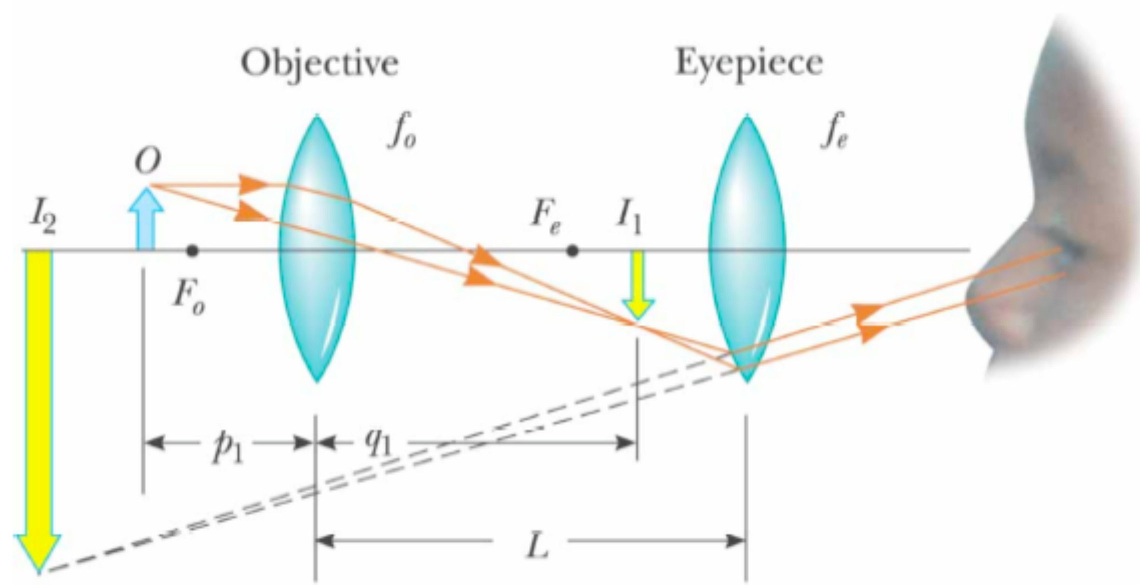
  1. Starting with the object, calculate the location and magnification of the image formed by the first lens/mirror.

  2. Take this image as the object for the second lens/mirror. Redefine the sign convention relative to this second lens/mirror.

  3. Calculate the image formed by the second lens/mirror.

  4. Repeat for any additional optical elements.

Example: compound microscope



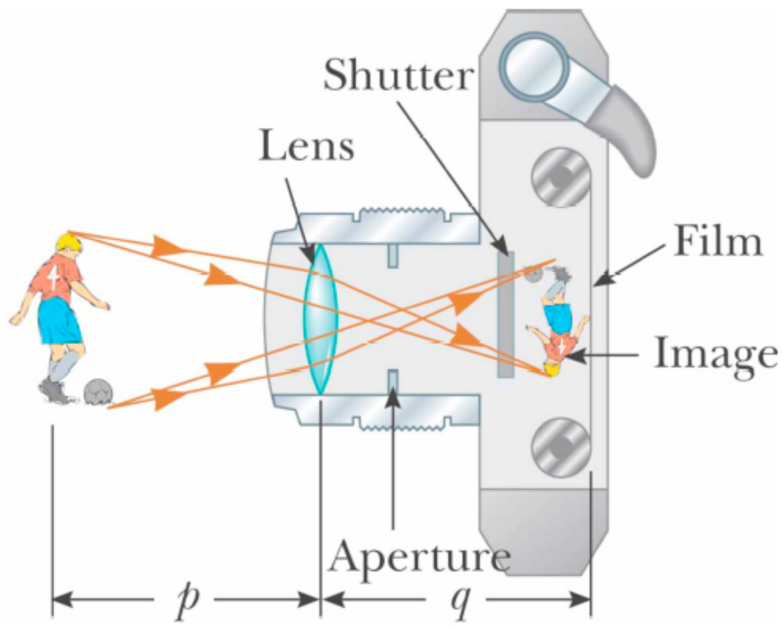
Why optical instruments?

- The purpose of optical instruments is to improve the viewing of objects beyond that possible with your unaided eye.
- The image is enhanced in 3 ways:

1. Make the image sharper: Often the lens of the unaided eye forms a blurred image on the retina. Glasses, contacts or a lens replacement can fix this.
2. Make the object viewed "bigger" so that finer details can be observed.
3. Make the object brighter. The eye has an aperture of about 5mm, and therefore can collect a limited amount of light.

The Camera

- A camera forms the image of an object on film or a charged-coupled device (CCD) to permanently record the image.
- A camera has the following basic parts:
  - a converging lens (focal length f) to form a real image of an object
  - a device to move the position of the lens relative to the film so that the real image is formed on the film (focusing the camera)
  - a variable aperture (D) to control the light intensity
  - a shutter that rapidly opens and closes to "freeze" an object in motion



The Camera

- When taking a picture with a camera you adjust (either manually or using internal electronic processing logic):
  - The lens - film separation: depends on object distance
  - The shutter speed say 1/30, 1/60, … 1/500s: depends on the motion of the object
  - The aperture D: depends on the intensity of the light source
  - The "f-number" is defined to be f/D: the intensity of the light on the film is proportional to $1/(\text{f-number})^2$

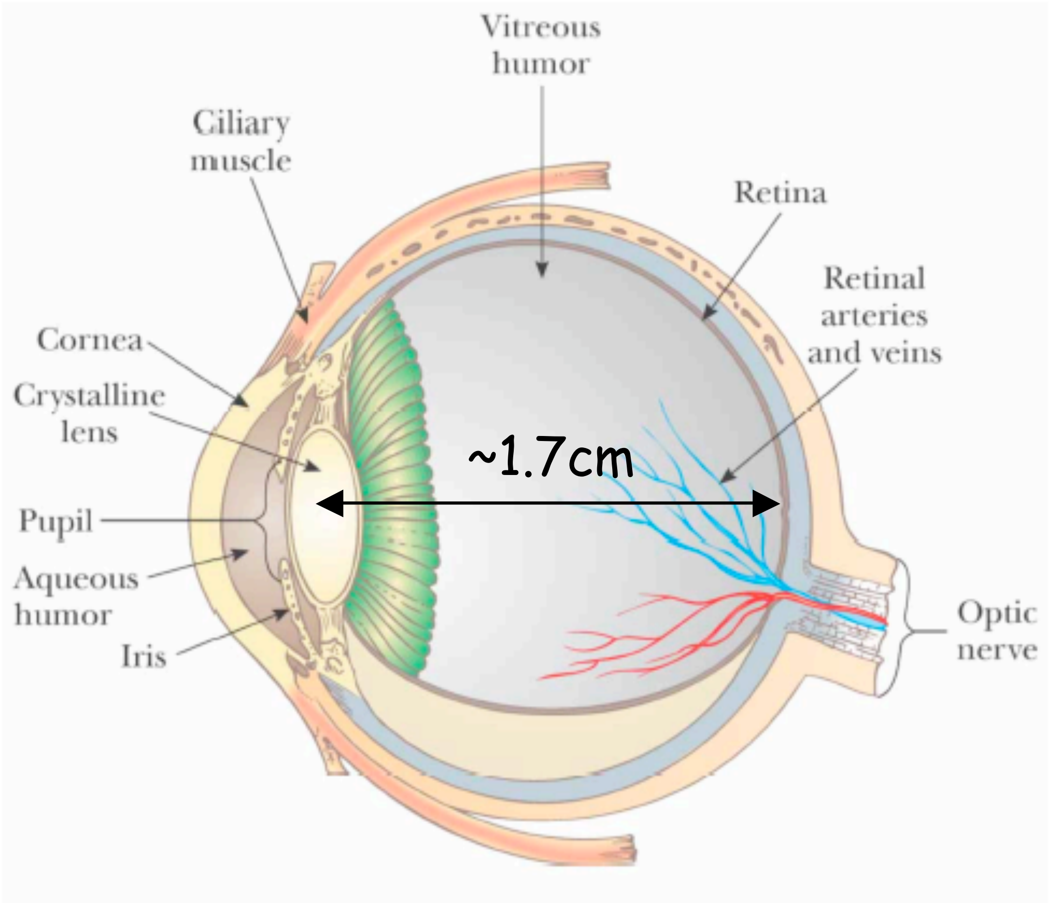
    Standard f-numbers for cameras are 2.8, 4, 5.6, 8, 11, 16.

The operation of the amazing human eye

- Our eyes are the biological optical transducers that take light signals and transform them to electrical impulses that are processed and interpreted by our brain.

- There is some analogy to a camera with a digital CCD read out.

- But the eye + brain's ability to detect and analyze light signals is truly amazing.



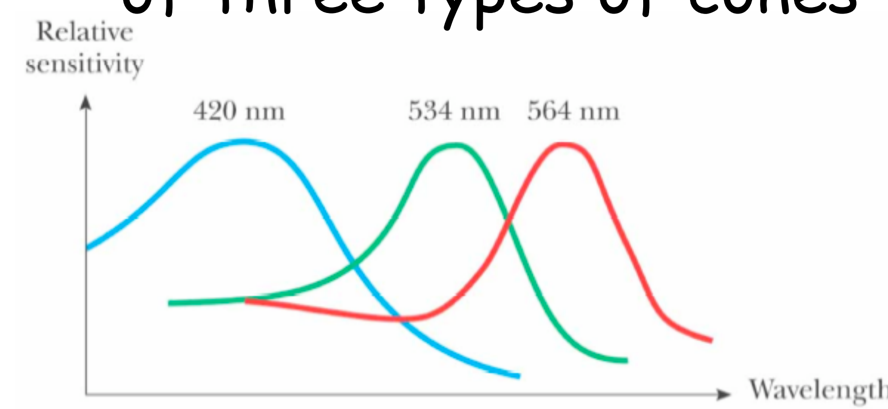
The eye

- The intensity of light entering the eye is controlled by a variable aperture called the iris. Intensities from bright sun to few photons can be detected.
- Ciliary muscles are used to modify the shape the lens and change its focus length to form an image on the back readout surface.
- This process is called accommodation. An eye can focus objects at distances from a far point (usually infinity) to a near point as close as about 25 cm for young eyes.
- The retina is the back surface of the eye. It is composed of read out sensors (rods and cones) that transform the light signals to electrical pulses that are sent via the optic nerve to the brain.

The eye

- The cones do the high resolution readout and are located near the center of the of retina. Three different types of cones respond to different wavelength ranges of the light spectrum. This is the information we use to construct the concept of color.

Spectral response of three types of cones



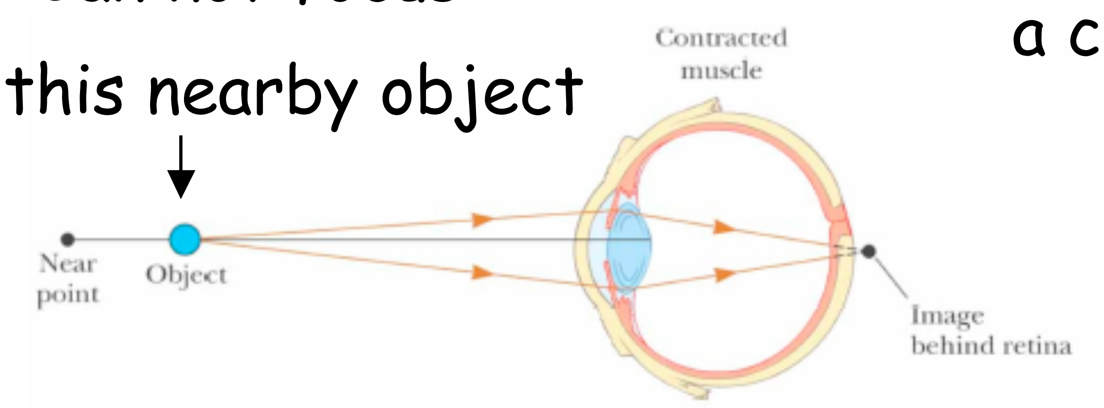
- The rods located on the periphery of the retina, are less dense and do not differentiate colors.

- There are on the order of a 100 million cones/rods used for readout.

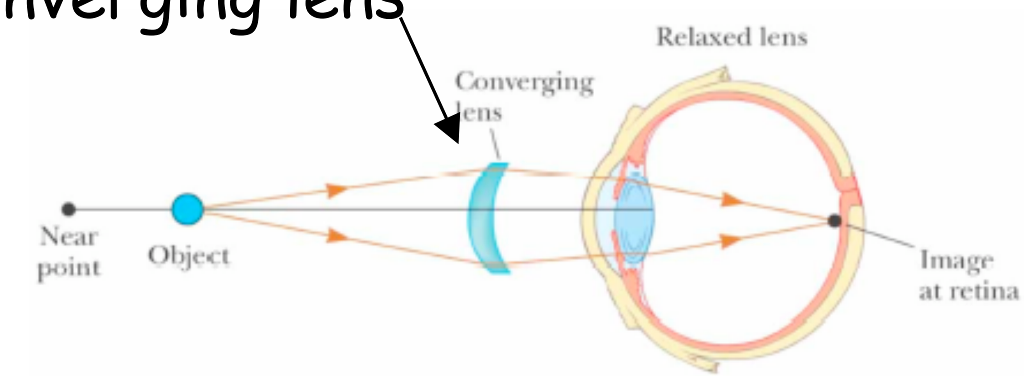
Imperfections of the eye

- A common imperfection, especially as people age, is a reduction of the flexibility of the ciliary -lens system. This limits the accommodation of the eye to focus on distant and close objects.
- Farsighted person: can focus on distant objects but not those nearby. Their near point is beyond the "normal" 25 cm.
- This can be corrected by placing a converging lens in front of the eye.

Can not focus on this nearby object



Correct with a converging lens



Improvement of the eye's performance

- Let's assume you have normal eyes, or they are corrected to be normal using glasses:

  far point is at infinity

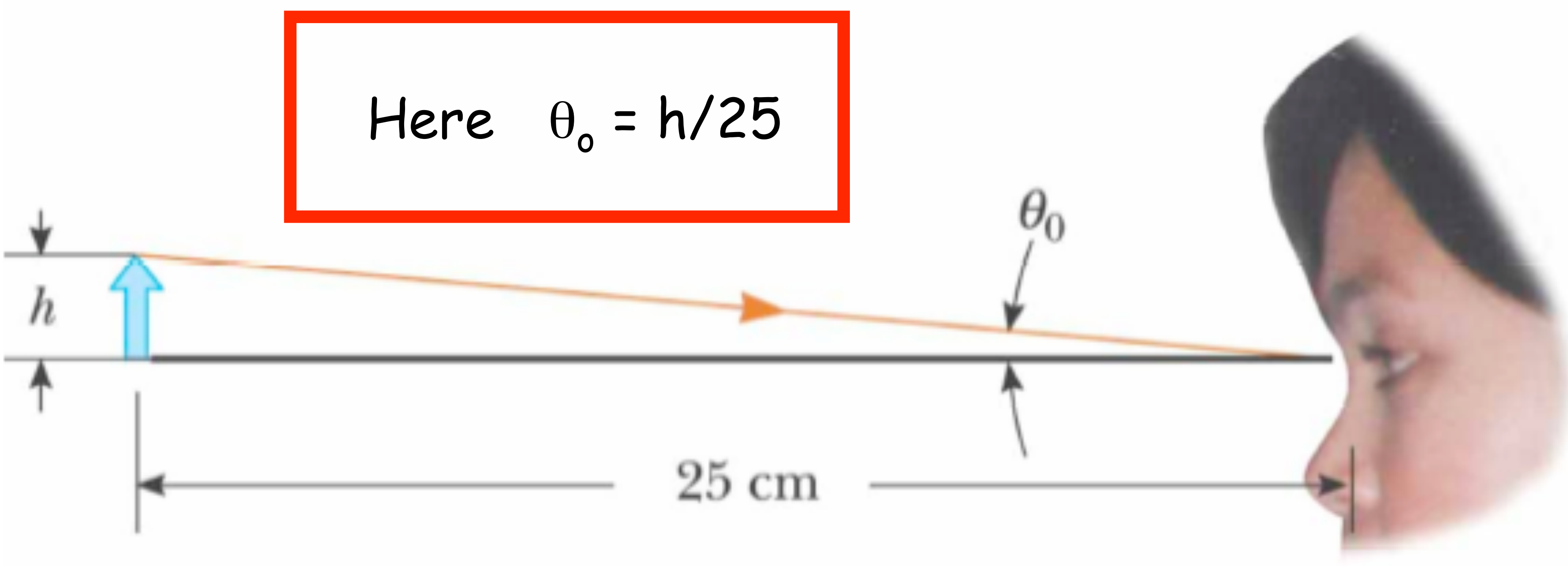
  near point is at ~ 25 cm

- As discussed this "normal" eye is a marvelous instrument to convert light signals to electrical signals for you brain to process and you to "see".
- However it has limitations in recording details of very small objects (say a bacteria) and very distant objects (say craters on the moon).
- Microscopes allow you to see very small objects.
- Telescopes allow you to see very distant objects

The simple magnifier

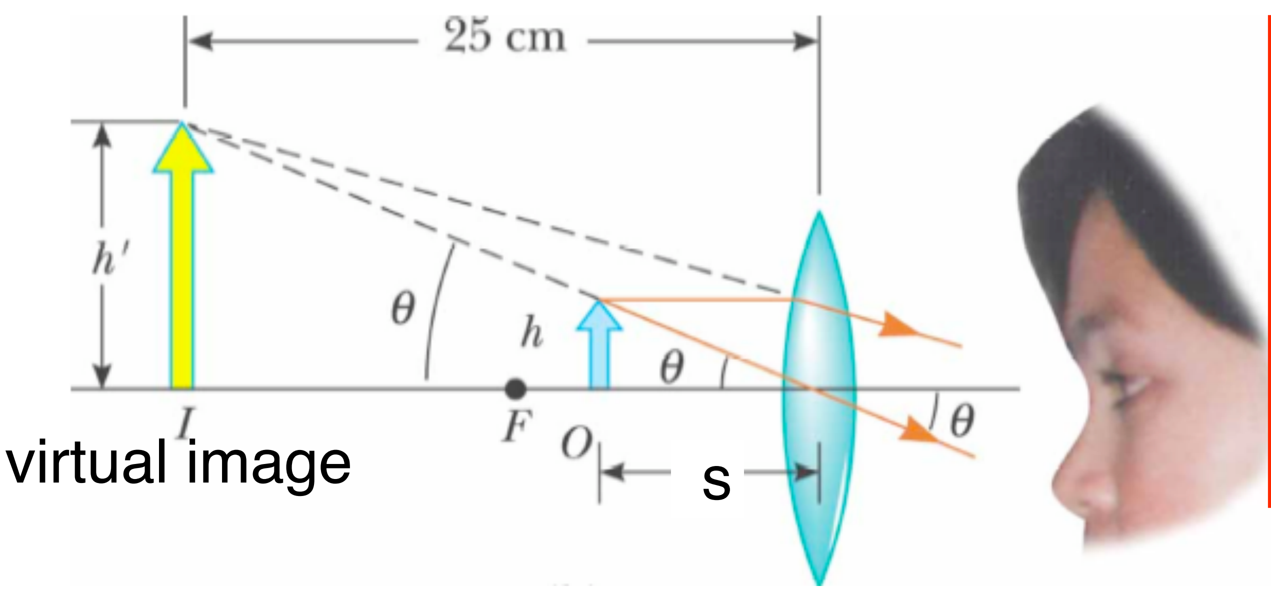
- The simplest enhancement to viewing small, close objects is a magnifying glass -- a short focal length converging lens.
- Consider viewing a small object with your unaided eye, with the object brought as close to the eye as possible = the near point of the eye of 25 cm.

Here $\theta_o = h/25$



The simple magnifier

- Now view the object through a converging lens (the magnifier) with the object placed within the focal point of the lens. The observer then views a virtual image of the object.
- Adjust the object distance so that the virtual image is now at the near point of the eye as shown below ($s' = -25$cm)
- From the figure: $1/s + 1/-25 = 1/f$ or $s = 25f/(25+f)$
- Therefore $\theta = h/s = h(25+f)/25f$ and using $m = \theta/\theta_o$ with $\theta_o = h/25$



m = angular magnification

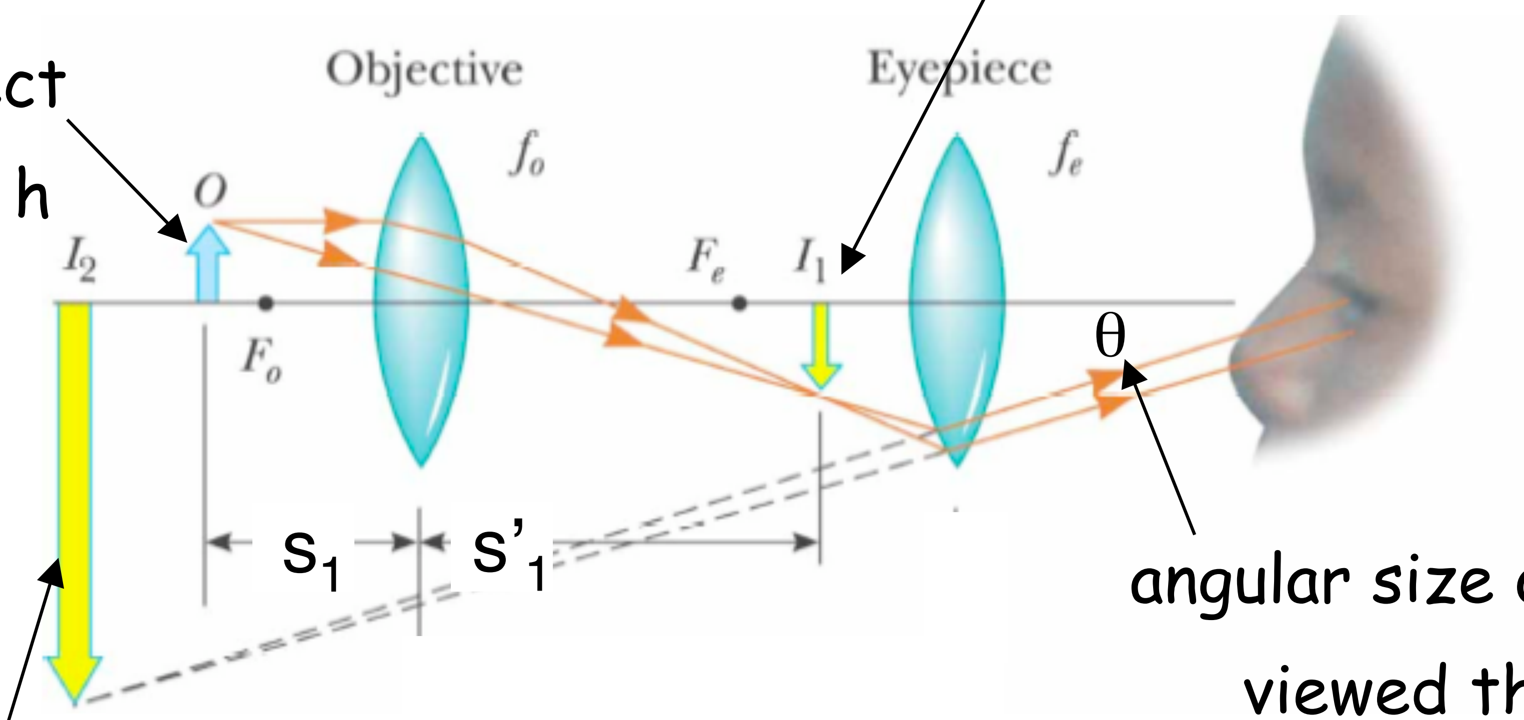
= 1 + 25/f(cm)

(for the image at the eye's near point)

Operation of a compound microscope

- The objective lens forms a real image of the object.

small object of height h

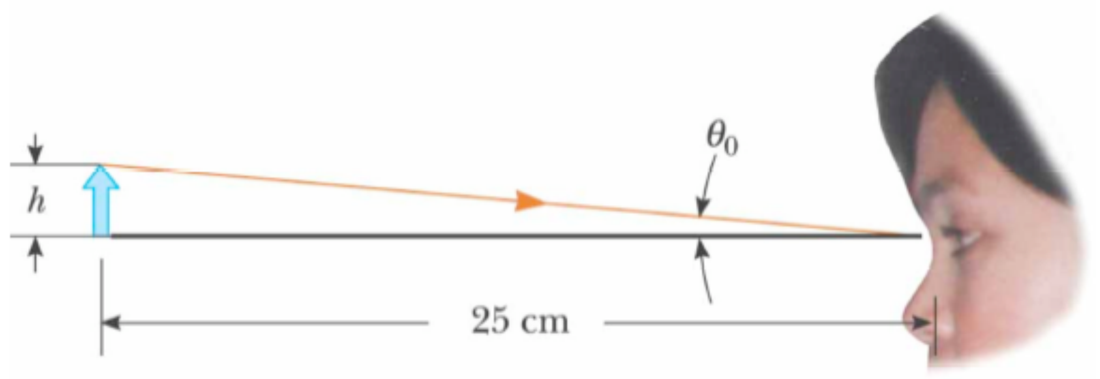


angular size of image when viewed through the compound microscope

- The eyepiece views this real image and provides a virtual image to the eye.

Magnification of a compound microscope

- We want to calculate angular magnification = m = θ / θ_o where as defined above θ_o = angular size of the object when viewed with the unaided eye and θ = the angular size when viewed through the microscope.



- Step 1: find θ_o

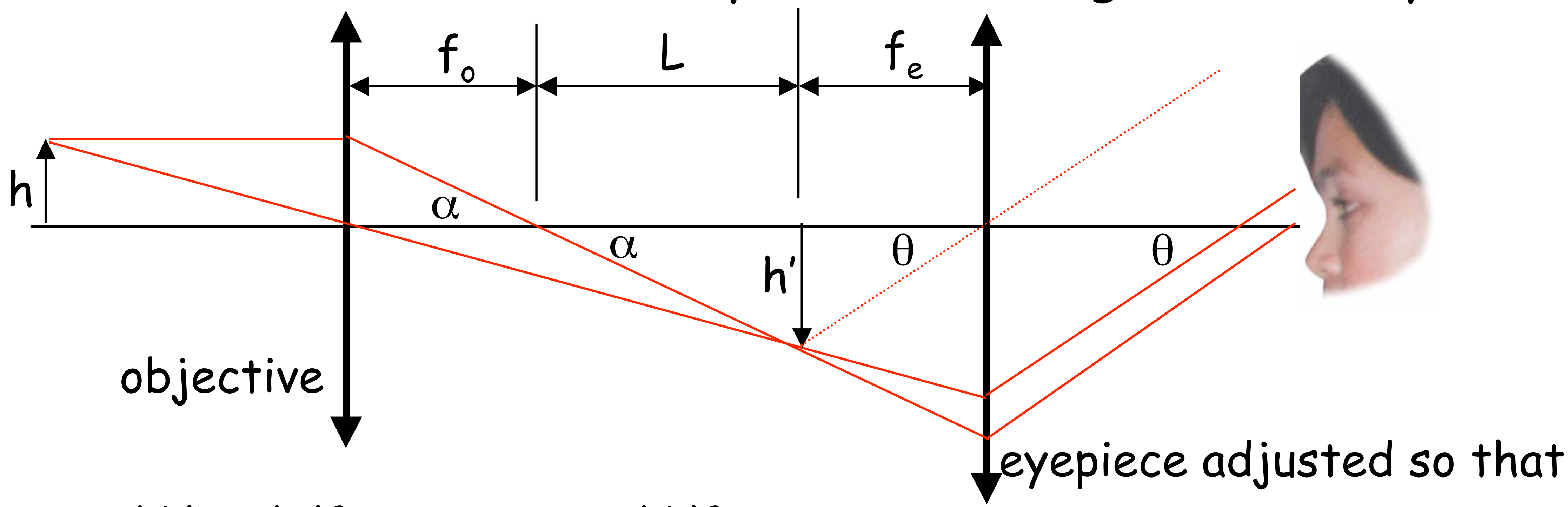
  Using an unaided eye

  with the object place at the near point of the eye:

  θ_o = h/25 cm (assuming a near point of 25 cm)

Magnification of a compound microscope

- Step 2: find the angular size of the final image = θ . Do this for the case of a relaxed eye => final image at infinity.



1. $\alpha = h'/L = h/f_o$ and $h' = L\,h/f_o$

2. $\theta = h'/f_e$ and $h' = \theta\,f_e$

3. Therefore $L\,h/f_o = \theta\,f_e$ and $\theta = L\,h/(f_o f_e)$

Magnification of a compound microscope

- From Step 1: $\theta_o = h/25\text{ cm}$
- From Step 2: $\theta = L\,h/(f_o f_e)$
- And finally angular magnification $m = \theta/\theta_o$ is:

$$m = -L\,25/(f_o f_e)$$

Compound microscope adjusted for relaxed eye viewing by an eye with a near point at 25 cm

insert - sign since image inverted

Matrix Operations - The idea of a matrix arises in this way. Suppose we have a pair of linear equations

$$U = Ax + By$$

$$V = Cx + Dy$$

where A,B,C and D are known constants and x and y are variables. We can write this pair of equations as:

$$\begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

where [..] is called a **matrix.**

$$\begin{bmatrix} U \\ V \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} x \\ y \end{bmatrix}$$

are 2x1 matrices or **column matrices** or **column vectors.**

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

is a 2x2 matrix. We can have n x n matrices, n x 1 column vectors and 1 x n **row vectors.** We can then write

$$C_2 = \begin{bmatrix} U \\ V \end{bmatrix} \qquad S = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \qquad C_1 = \begin{bmatrix} x \\ y \end{bmatrix}$$

so that

$$C_2 = SC_1$$

The equations define the operation of **matrix multiplication.** We also have

$$L = PU + QV$$

$$M = RU + TV$$

or

$$\begin{bmatrix} L \\ M \end{bmatrix} = \begin{bmatrix} P & Q \\ R & T \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix}$$

or

$$C_3 = KC_2$$

These give the relations

$$L = P(Ax + By) + Q(Cx + Dy)$$

$$M + R(Ax + By) + T(Cx + Dy)$$

or

$$\begin{bmatrix} L \\ M \end{bmatrix} = \begin{bmatrix} PA + QC & PB + QD \\ RA + TC & RB + TD \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

or

$$C_3 = FC_1 = (KS)C_1$$

where

$$F = KS = \begin{bmatrix} PA + QC & PB + QD \\ RA + TC & RB + TD \end{bmatrix}$$

Thus, we obtain the rule for matrix multiplication

$$\begin{bmatrix} P & Q \\ R & T \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} PA+QC & PB+QD \\ RA+TC & RB+TD \end{bmatrix}$$

Generalizing to arbitrary size we have

$$(C_2)_i = \sum_k S_{ik}(C_1)_k \qquad (C_3)_i = \sum_k K_{ik}(C_2)_k$$

$$(C_3)_i = \sum_j \left(\sum_k K_{ik} S_{kj} \right)(C_1)_j = \sum_j F_{ij}(C_1)_j$$

$$F_{ij} = \sum_k K_{ik} S_{kj}$$

Matrix Addition/Subtraction

$$F_{ij} = K_{ij} \pm S_{ij}$$

Special Matrices

Null matrix: $A_{ij} = 0 \text{ , } all \ \ i, j$

Unit or identity matrix I : $A_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

Diagonal matrix: $A_{ij} = 0 \quad i \neq j$

Matrix Transpose: $\left(A^T\right)_{ij} = A_{ij} \qquad \left(AB\right)^T = B^T A^T$

Determinants (interested in 2x2 matrices only)

$$P = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad , \quad \det(P) = \begin{vmatrix} A & B \\ C & D \end{vmatrix} = AD - BC$$

$$\det(AB) = \det(A)\det(B)$$

Matrix Inversion (interested in 2x2 matrices only)

$$P = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad , \quad P^{-1} = \begin{bmatrix} D & -B \\ -C & A \end{bmatrix}$$

where

$$PP^{-1} = \det(P)I$$

Note that det(P) = 1 (unimodular) for matrices of interest to us.

Matrix Diagonalization - Suppose we have the relation

$$M = F\Lambda F^{-1}$$

where Λ is a diagonal matrix. This implies that

$$M^N = F\Lambda^N F^{-1}$$

Thus, once the diagonalizing transformation F has been found, the N^{th} power of the original matrix is found by taking the diagonal matrix to the N^{th} power, which is easy since all we have to do is replace each diagonal element by its N^{th} power.

The diagonal elements of the matrix Λ are the eigenvalues of M and the columns of the diagonalizing matrix are its eigenvectors. This is shown below:

$$C_r = column\ vector\ with\ 1\ in\ r^{th}\ place$$

$$F_r = FC_r = r^{th} column\ of\ F$$

$$MF_r = (F\Lambda F^{-1})(FC_r) = F\Lambda C_r = \lambda_r F$$

$$\lambda_r = r^{th}\ diagonal\ element\ of\ \Lambda$$

Eigenvalues and Eigenvectors for 2x2 Unimodular Matrices - In general, eigenvalues are determined using the characteristic equation

$$\det(\lambda I - M) = 0$$

For a 2x2 unimodular matrix, we have

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

$$\det(M) = 1 = AD - BC$$

$$\det(\lambda I - M) = 0 = \det\begin{bmatrix} \lambda - A & -B \\ -C & \lambda - D \end{bmatrix}$$

$$= (\lambda - A)(\lambda - D) - BC$$

$$= (\lambda - A)(\lambda - D) + (1 - AD)$$

or

$$\lambda^2 - (A+D)\lambda + 1 = 0 = \lambda^2 - (TR(M))\lambda + \det(M)$$

so that we have two eigenvalues satisfying

$$\lambda_1 + \lambda_2 = A + D = Tr(M) \quad , \quad \lambda_1\lambda_2 = 1 = \det(M)$$

and we find

$$\lambda_{1,2} = \frac{1}{2}\left[(A+D) \pm \sqrt{(A+D)^2 - 4}\,\right]$$

If Tr(M)=A+D lies between 2 and -2 in value, then the two eigenvalues can be written conveniently in terms of an angle θ, chosen so that it lies between 0 and π and such that A+D=2cosθ. We then find

$$\lambda_1 = \cos\theta + i\sin\theta = \exp(i\theta)$$

$$\lambda_2 = \cos\theta - i\sin\theta = \exp(-i\theta)$$

If Tr(M)=A+D is greater than 2 or less than -2, we can choose a positive quantity t such that A+D=2cosh(t) (or -2cosh(-t) if A+D is negative). The eigenvalues then take the form

$$\lambda_1 = \exp(t) \;\; (or \;\; -\exp(t) \;\; if \;\; (A+D) \;\; is \;\; negative)$$

$$\lambda_2 = \exp(-t) \;\; (or \;\; -\exp(-t) \;\; if \;\; (A+D) \;\; is \;\; negative)$$

We determine the diagonalizing matrix F as follows.

$$M = F\Lambda F^{-1} \Rightarrow MF = F\Lambda$$

$$MF = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = F\Lambda$$

where AD-BC=1 since det(M)=1. On multiplying we get

$$\begin{bmatrix} AF_{11}+BF_{21} & AF_{12}+BF_{22} \\ CF_{11}+DF_{21} & CF_{12}+DF_{22} \end{bmatrix} = \begin{bmatrix} F_{11}\lambda_1 & F_{12}\lambda_2 \\ F_{21}\lambda_1 & F_{22}\lambda_2 \end{bmatrix}$$

which gives

$$\frac{F_{11}}{F_{21}} = \frac{\lambda_1 - D}{C} = \frac{B}{\lambda_1 - A} \quad , \quad \frac{F_{12}}{F_{22}} = \frac{\lambda_2 - D}{C} = \frac{B}{\lambda_2 - A}$$

where the eigenvectors are

$$F_1 = \begin{bmatrix} F_{11} \\ F_{21} \end{bmatrix} \quad , \quad F_2 = \begin{bmatrix} F_{12} \\ F_{22} \end{bmatrix}$$

One possible solution is

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} = \begin{bmatrix} \lambda_1 - D & \lambda_2 - D \\ C & C \end{bmatrix}$$

Matrix methods in Paraxial(Gaussian) Optics

Let us discuss how matrices can be used to describe the geometric formation of images by a centered lens system - a succession of spherical refracting surfaces all centered on the same optical axis.

The results are valid within two approximations. First, the basic assumption of all geometric optics is that the wavelength of light is small and the propagation of light can be explained in terms of individual rays instead of wavefronts. As can be shown by means of Huygens' construction, if light waves are allowed to travel without encountering any obstacles, they are propagated along a direction which is normal to the wavefronts. The concept of a geometric ray is a idealization of this wavenormal. Each ray obeys Fermat's principle of least time - if we consider the neighborhood of any short section of the raypath, the path which the ray chooses between a given entry point and a given exit point is that which minimizes the time taken.

The second approximation is that we only consider paraxial rays - those that remain close to the axis and almost parallel to it so that we can use the first-order approximations for sines or tangents of any angles.

Ray-Transfer Matrices - Now consider the propagation of a paraxial ray through a system of centered lenses. We use a Cartesian coordinate system where the z-axis (horizontal) represents the optical axis. The y-axis is orthogonal to the optical axis. All rays will lie in the yz-plane and close to the z-axis.

The trajectory of a ray as it passes through the various refracting surfaces of the system will consist of a series of straight lines. We specify a ray by the height y of a point on the ray and the angle (anticlockwise from z-axis is positive direction) the ray makes with the z-axis as shown in figure 1.

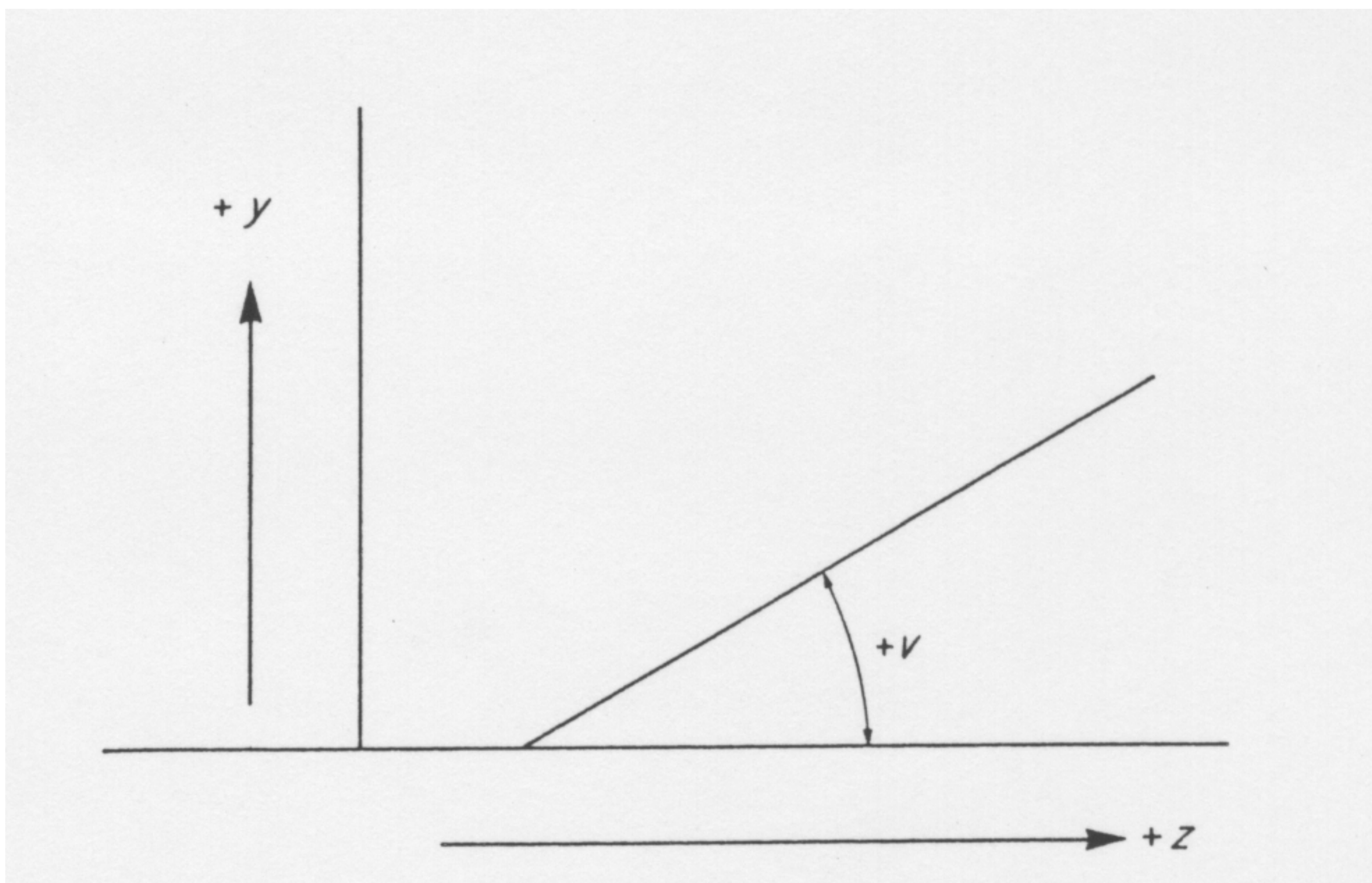


Figure 1

Later calculations are made more convenient by replacing the ray angle v with the corresponding **optical direction-cosine** V = nv (or strictly, V = nsin(v)) where n is the refractive index of the medium in which the ray is traveling. The optical direction-cosine has the property that, by Snell's law, it will remain unchanged as it crosses a plane boundary between two different media. It will also make all matrices involved unimodular(unitary with determinant = 1), which will be very useful.

As a ray passes through a refracting lens system, there are only two basic types of process that we need to consider in order to determine its progress:

(a) A **translation** during which the ray simply moves in a straight line to the next refracting surface. We need to know the distance translated (thickness of material) and the corresponding refractive index n.

(b) **Refraction** at the boundary surface between two regions of different refractive index. To determine how much bending the ray undergoes, we need to know the radius of curvature of the refracting surface and the two values of refractive index.

We now investigate the effect that each of these two basic elements has on the y-value and the V-value of a ray passing between two reference planes(one on either side of the element), namely RP_1 and RP_2, orthogonal to the optical axis(z-axis).

The ray first passes through RP_1 with value y_1 and V_1, then through the optical element and then through RP_2, with values y_2 and V_2. We want equations expressing y_2 and V_2 in terms of y_1 and V_1 and the properties of the optical element. We will find that for both types of elements (translation and refraction) the equations are linear and thus can be written as

$$\begin{bmatrix} y_2 \\ V_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} y_1 \\ V_1 \end{bmatrix}$$

where the determinant of the transformation matrix(ray-transfer matrix) is equal to 1, i.e., AD-BC=1. Alternatively, if we want to trace a ray backwards, the matrix equation can be inverted to give

$$\begin{bmatrix} y_1 \\ V_1 \end{bmatrix} = \begin{bmatrix} D & -B \\ -C & A \end{bmatrix} \begin{bmatrix} y_2 \\ V_2 \end{bmatrix}$$

The ray-transfer matrix representing the optical element is the matrix product of all the ray-transfer matrices of which it is composed.

The Translation Matrix

The figures 2(a) and 2(b) below show tow examples of a ray which travels a distance t to the right between the reference planes.

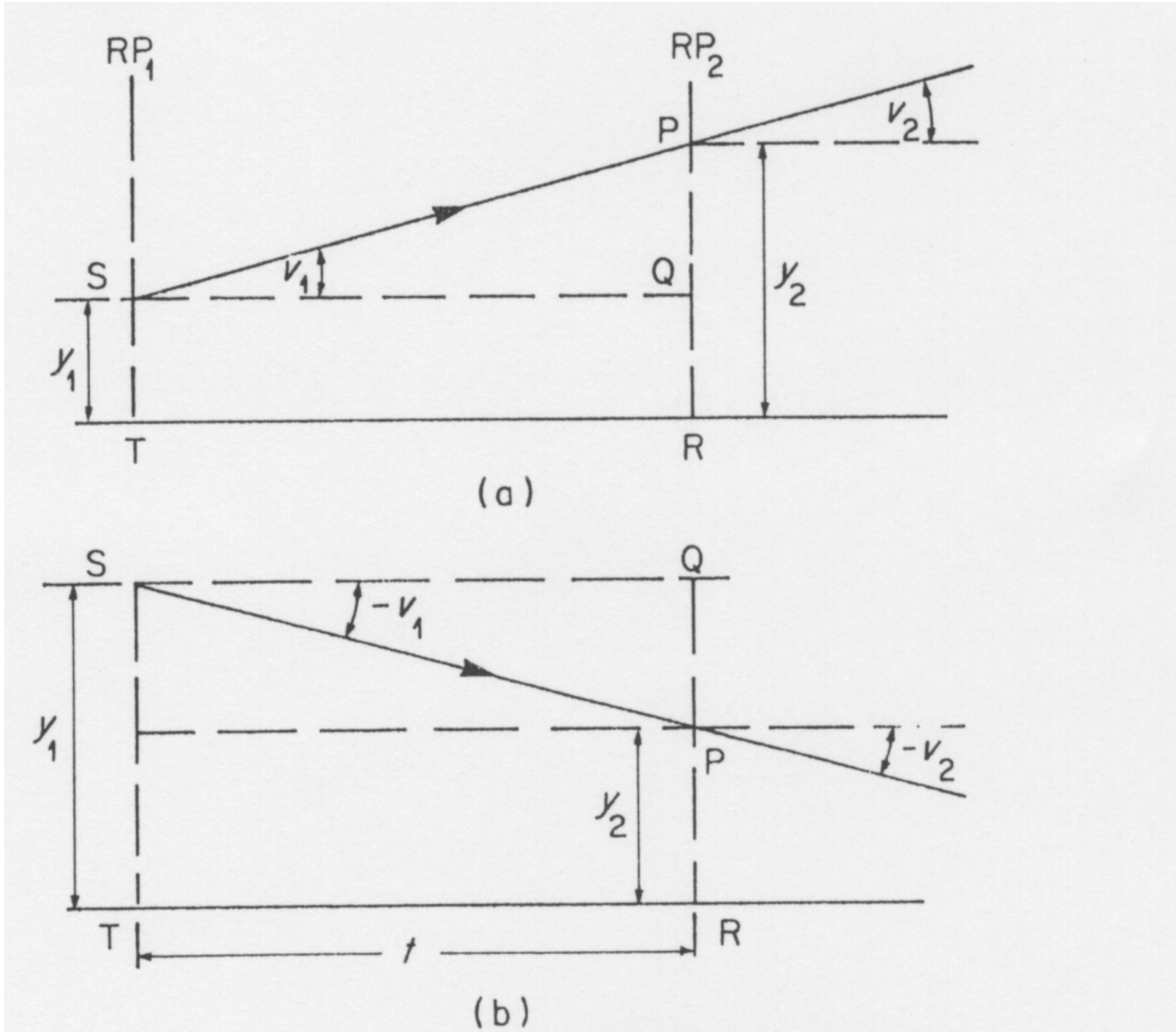


Figure 2

Clearly the angle of the ray will remain unchanged during the translation, but not the distance y from the optical axis. Figure 2(a) illustrates the case where both y- and V-values remain positive and Figure 2(b) illustrates the case where the V-value is negative.

In both figures the angles v are greatly exaggerated(generally less than 0.1 rad or 6°)

Referring to Figure 2(a)	Referring to Figure 2(b)
$y_2 = RP = RQ + QP$	$y_2 = RP = RQ - QP$
$= TS + SQ\tan(PSQ)$	$= TS - SQ\tan(PSQ)$
$= y_1 + t\tan(v_1)$	$= y_1 - t\tan(-v_1)$
$= y_1 + tv_1$	$= y_1 + tv_1$

If n is the refractive index of the medium between RP_1 and RP_2, we can write

$$y_2 = y_1 + tv_1 = y_1 + \left(\frac{t}{n}\right)nv_1 = (1)y_1 + TV_1$$

Clearly

$$T = \frac{t}{n} = \text{reduced thickness}$$

It is clear from the diagrams that $v_1 = v_2$ so that we can write

$$V_2 = nv_2 = nv_1 = (0)y_1 + (1)V_1$$

The pair of equations can be written using matrices as

$$\begin{bmatrix} y_2 \\ V_2 \end{bmatrix} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ V_1 \end{bmatrix}$$

Thus, the matrix representing a translation to the right through a reduced distance T is

$$\Im = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$$

Clearly $\det(\Im) = 1$.

Compound Layers and Plane-Parallel Plates

If we imagine that we divide the distance t into two segments t_1 and t_2, both with index of refraction n, we can write two successive translation matrices

$$\mathfrak{I}_1 = \begin{bmatrix} 1 & T_1 \\ 0 & 1 \end{bmatrix} \quad , \quad \mathfrak{I}_2 = \begin{bmatrix} 1 & T_2 \\ 0 & 1 \end{bmatrix}$$

$$T_1 = \frac{t_1}{n} \quad , \quad T_2 = \frac{t_2}{n}$$

and we note that

$$\mathfrak{I}_1\mathfrak{I}_2 = \mathfrak{I}_2\mathfrak{I}_1 = \begin{bmatrix} 1 & T_1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & T_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & T_1 + T_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} = \mathfrak{I}$$

According to our rules above we should write $\mathfrak{I} = \mathfrak{I}_1\mathfrak{R}\mathfrak{I}_2$ where $\mathfrak{R}$ is the refraction matrix at the boundary between the two segments. Clearly, $\mathfrak{R} = I$ = identity matrix in this case(same medium on both sides of boundary). We will prove this shortly. A similar situation applies if we divide into many segments. We have

$$\mathfrak{I} = \mathfrak{I}_1\mathfrak{R}\mathfrak{I}_2\mathfrak{R}\mathfrak{I}_3\ldots\ldots\ldots\ldots\mathfrak{R}\mathfrak{I}_N = \mathfrak{I}_1\mathfrak{I}_2\mathfrak{I}_3\ldots\ldots\ldots\ldots\mathfrak{I}_N = \begin{bmatrix} 1 & \sum_{i=1}^{N} T_i \\ 0 & 1 \end{bmatrix}$$

Clearly, the order of the segments in this case is irrelevant.

When we say that a plane-parallel glass plate of refractive index n and thickness t has a reduced-thickness t/n we are using very appropriate language. If we are looking at an object on the far side of the plate the light actually takes longer to reach us then if the plate were absent, but the object certainly looks closer. Replacing a layer of air by the same thickness of glass makes the world seem closer to the observer by a distance

$$\frac{t}{1} - \frac{t}{n} = \frac{t(n-1)}{n}$$

which is about 1/3 the thickness of the plate. For an object submerged in water, the factor (n-1)/n is only about 1/4, but the depth t may be very large - even a bear fishing with his paw needs to know about reduced thickness!

The Refraction Matrix

We now consider the action of a curved surface separating two regions of refractive index n_1 and n_2. The radius of curvature of the surface will be taken to be positive if the center of curvature lies to the right of the surface. This situation is shown in Figure 3 below, which shows a surface of positive curvature with the refractive index n_2 on the right of the surface greater than that on the left (n_1). The ray also shows that we have positive y- and V-values on both sides of the surface.

The angles are again greatly exaggerated. As a consequence, RP_1 appears well separated from RP_2. But for paraxial rays, the separation between these two planes will be $r(1-\cos\alpha)$ and is very small since α (like v_1 and v_2) is very small. We therefore have $y_1 = y_2$.

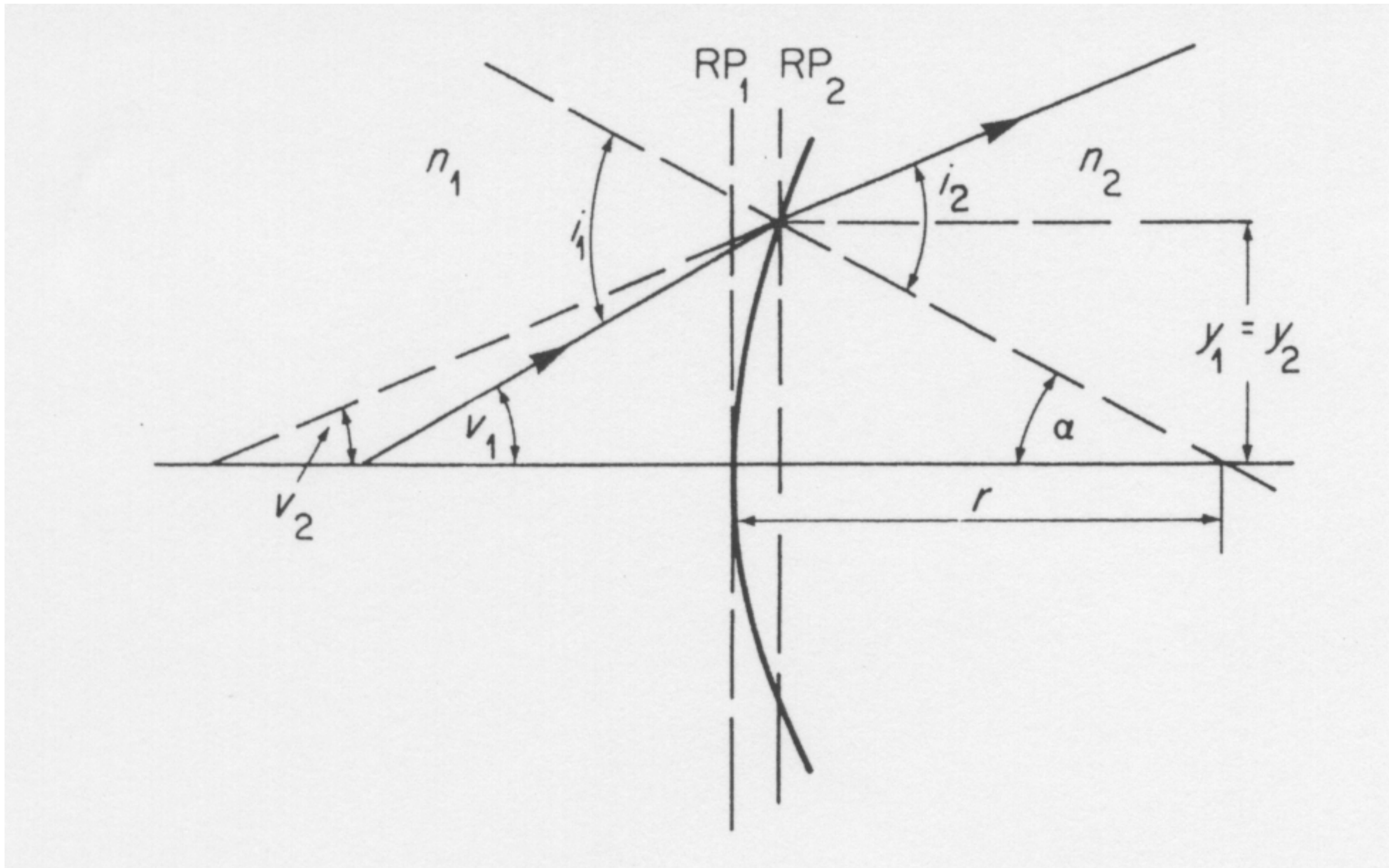


Figure 3

Applying Snell's law in the diagram, we have

$$n_1 \sin i_1 = n_2 \sin i_2 \rightarrow n_1 i_1 = n_2 i_2$$

in the paraxial approximation. But simple geometry says

$$i_1 = v_1 + \alpha = v_1 + \frac{y_1}{r} \quad and \quad i_2 = v_2 + \alpha = v_2 + \frac{y_2}{r}$$

Hence,

$$n_1\left(v_1 + \frac{y_1}{r}\right) = n_2\left(v_2 + \frac{y_1}{r}\right)$$

or

$$V_1 + \frac{n_1 y_1}{r} = V_2 + \frac{n_2 y_1}{r}$$

Thus, rearranging the equations in matrix form, we get

$$\begin{bmatrix} y_2 \\ V_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\dfrac{n_2 - n_1}{r} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ V_1 \end{bmatrix}$$

The quantity $(n_2-n_1)/r$ is the refractive power of the surface. We thus have the refraction matrix

$$\Re = \begin{bmatrix} 1 & 0 \\ -\dfrac{n_2 - n_1}{r} & 1 \end{bmatrix}$$

Even though this was derived for a special case (all quantities positive), it turns out to be general and works in all other cases. Two special cases are:

$n_2 = n_1$ which says $\Re = I$

and $r \to \infty$ (a plane surface) which also says that $\Re = I$

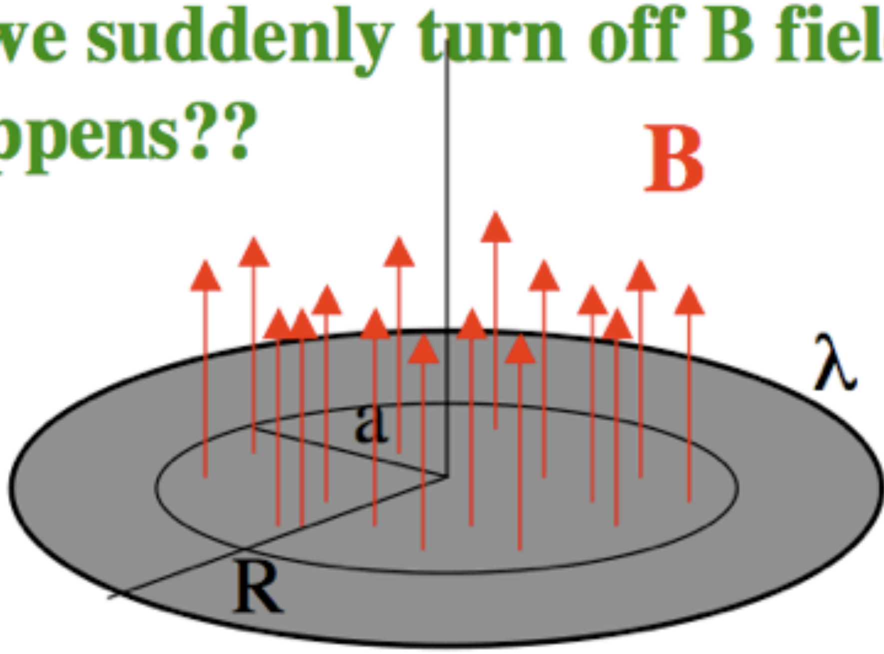
We used this result earlier when discussing plane-parallel plates. With these simple matrices we can do all of geometrical optics (Physics 50).

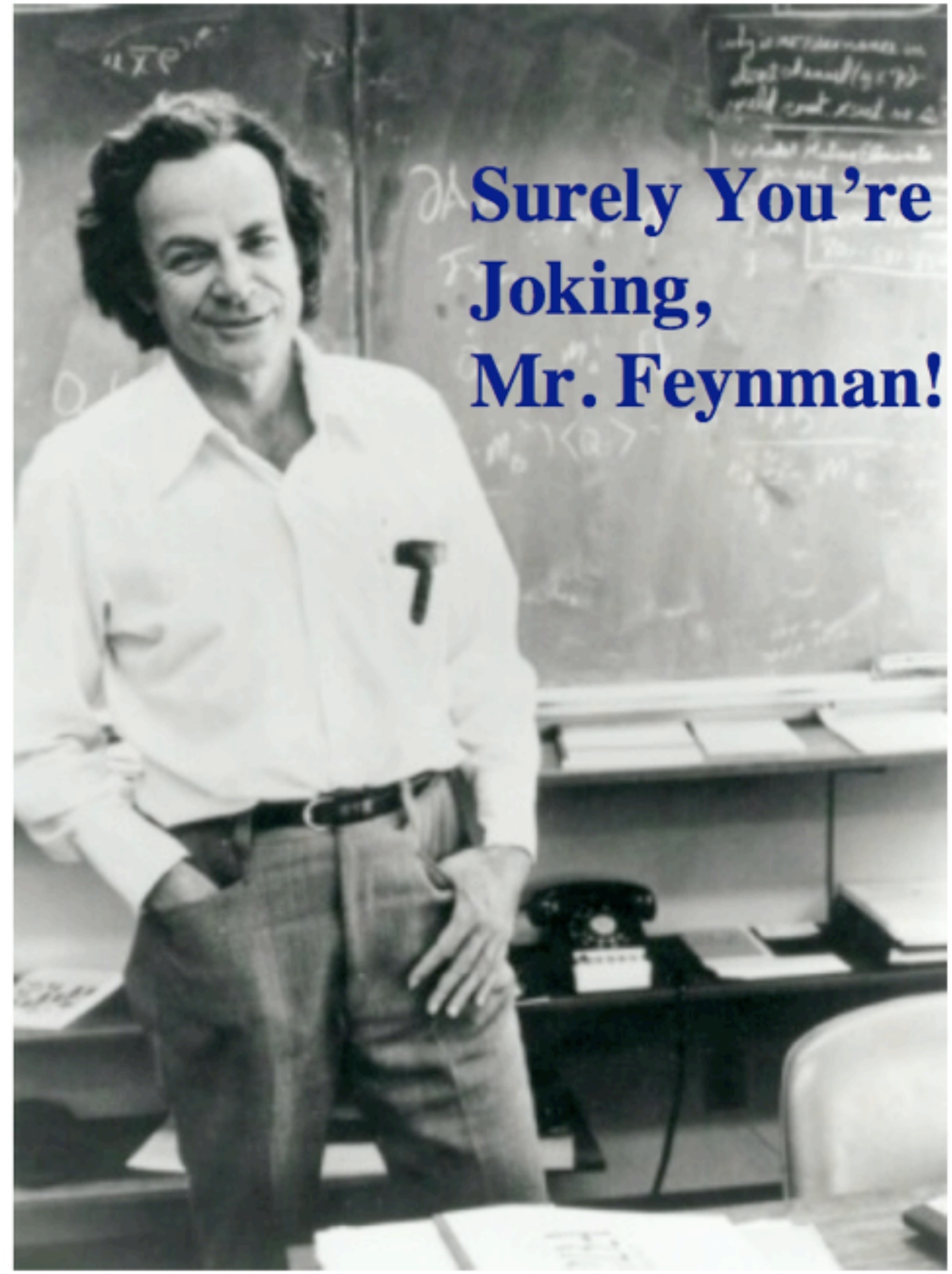
Feynman's Paradox

Consider a line charge λ on a wheel of radius R. The line charge is glued on to the non-conducting wheel, which is hung so it can rotate freely.

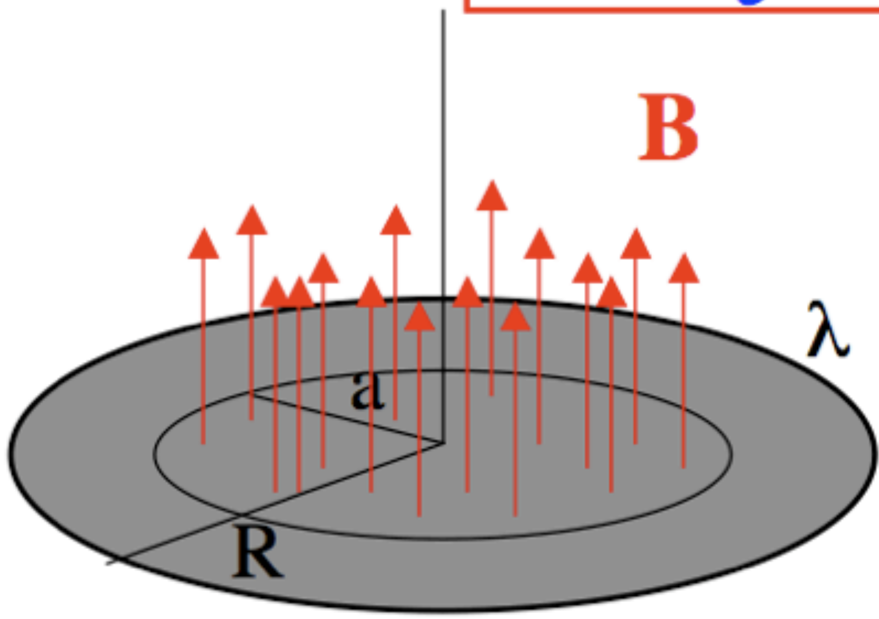
A uniform magnetic field B is in the central region of the wheel r<a.

If we suddenly turn off B field, what happens??





Feynman's Paradox



What happens when B field is turned off??

Before B field turned off: Flux $\Phi_B = B\pi a^2$

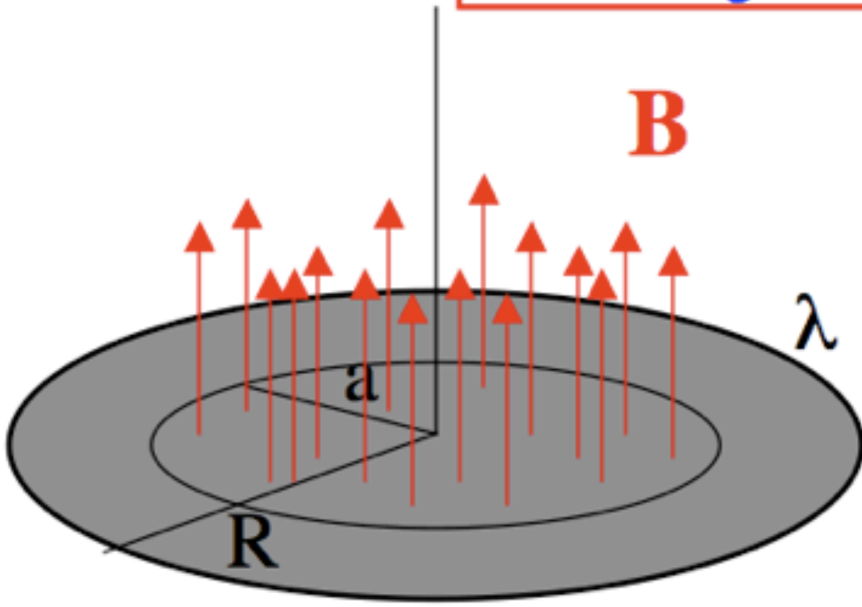
After B field turned off: Flux $\Phi_B = B\pi a^2$ goes to zero

An induced current will oppose the change in flux (Faraday's Law)

The wheel will rotate to make charges move ⇒ current

What is the angular momentum?

Feynman's Paradox



First calculate torque, then angular momentum

Torque

$$\vec{\tau} = \oint \vec{R} \times d\vec{F} = \oint \vec{R} \times \lambda\, dl\, \vec{E} = \lambda R \oint E\, dl\, \hat{k} = \lambda R\left(-\frac{d\Phi_B}{dt}\right)\hat{k} = -\hat{k}\,\lambda R \frac{d}{dt}\int \vec{B}\cdot d\vec{a}$$

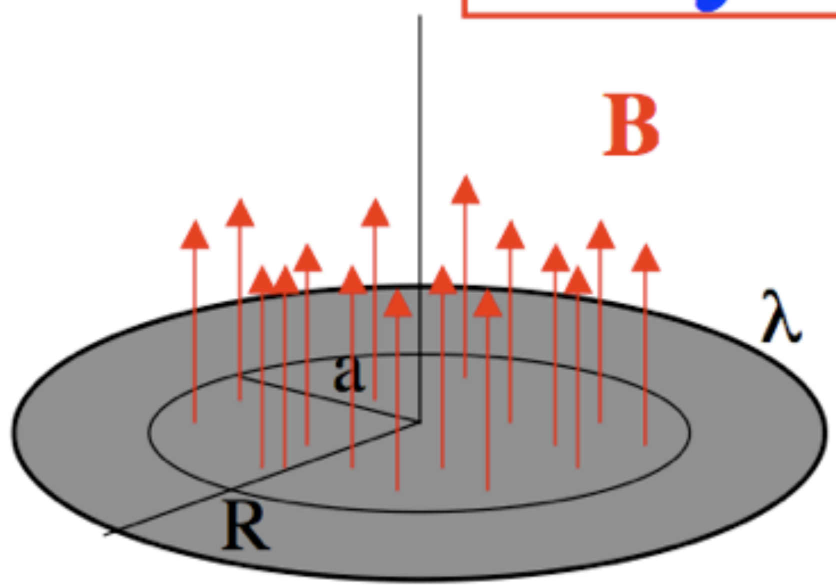
$$= -\lambda R \hat{k}\frac{d}{dt}\int \vec{B}\cdot d\vec{a} = -\lambda R \pi a^2 \frac{\partial \vec{B}}{\partial t}$$

Angular Momentum $\vec{L} = \int \vec{\tau}\, dt$ (like $\vec{p} = \int \vec{F}\, dt$)

$$= \int -\lambda R \pi a^2 \frac{\partial \vec{B}}{\partial t}\, dt = -\lambda R \pi a^2\, \hat{k} \int_{B_0}^{0} dB$$

$$\vec{L} = \lambda R \pi a^2\, B_0\, \hat{k} \qquad \vec{L} = I\omega$$

Feynman's Paradox



Does it really start spinning?

What about conservation of angular momentum?????

$$\vec{L}_{before} = ? = 0?$$

$$\vec{L}_{after} = ? \neq 0??$$

Is angular momentum not conserved??
Impossible!

Which do we flush down the toilet???
Conservation of angular momentum, or Faraday's Law???
Two laws of physics, but both apparently cannot be true here…
Either this wheel starts rotating, and angular momentum is not conserved, or it doesn't, but then Faraday's Law/Maxwell's Eqns are untrue! ???

HELP!

Poynting Vector

Poynting vector: $\vec{S} = \frac{1}{\mu_0} \vec{E} \times \vec{B}$

It 'points' ("poynts") in direction of energy flow.

It is the energy flux density… the energy transported by EM fields..

The momentum associated with EM field is:

$\vec{p}_{EM} = \mu_0 \varepsilon_0 \vec{S} = \varepsilon_0 \vec{E} \times \vec{B}$ where $\vec{p}_{EM}$ is momentum density

As long as E and B are non-zero, there is momentum in electromagnetic field.

Phew! Angular momentum is conserved, amount lost by field is gained by rotating disk!!! we keep both conservation of angular momentum and Maxwell's Eqns!

Radiation Pressure

Livermore Labs: recreate neutron star atmospheres using big-power lasers to compress and heat matter. (weapons too?)

Laser Cooling:

Laser Cooling cooling can reduce the temperatures of atoms to a few billionths of a degree above zero Kelvin.. Enable study of trapped atoms.

Eddington Luminosity: (or Eddington limit) is the largest luminosity that can pass thru a layer of gas in hydrostatic equilibrium - assuming spherical symmetry. Set the outward radiation pressure equal to the gravitational force inward, then can use the mass-luminosity relationship to set a limit on the maximum mass of a star. The luminosity of a star can't exceed the Eddington limit- if it did, the outer gas layers would blow away

Radiation Pressure: Solar Sails

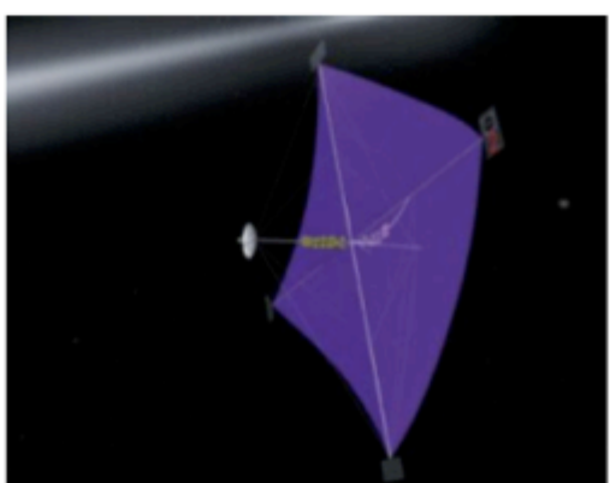
Sunlight (electromagnetic radiation) ⇒ momentum density

On earth: Intensity is ~1.3 kW/m^2 -- small (=.13 W/cm^2)

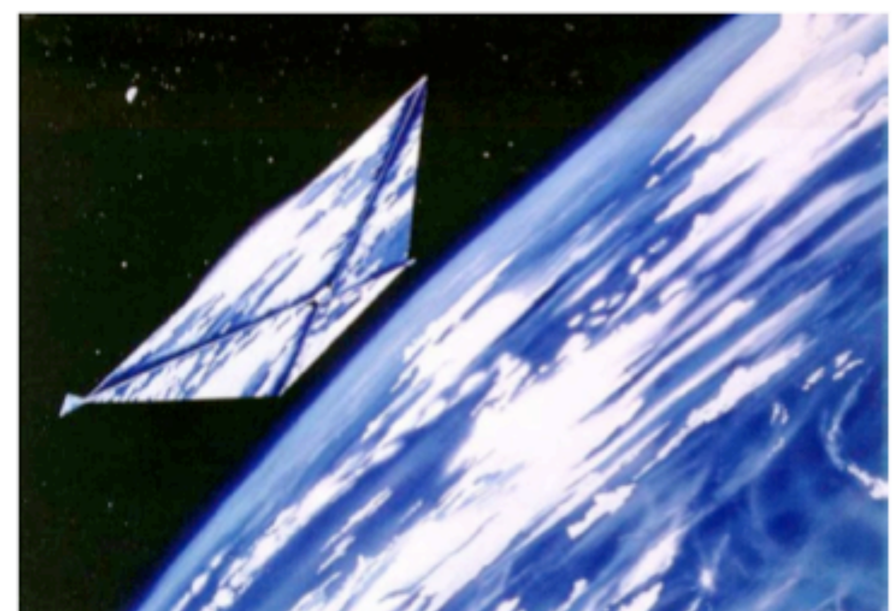
Spacecraft with solar sails use radiation pressure (no rockets) for propulsion. Sails of such a spacecraft are usually made out of a large reflecting panel.
The size of each panel is maximized to allow the largest possible flux of incident photons, and hence largest possible total momentum transfer from the incident radiation.
Because the surface is reflective, the momentum transferred by the photons is twice their momentum. Pressure ~ 9×10^{-6} N/m^2 = .1μg/cm^2

$$\langle\vec{S}\rangle = \frac{1}{2}c\varepsilon_0 E_0^{\,2} = I \text{ (Intensity)}$$

Radiation Pressure $P = \frac{I}{c}$ (absorbed) or $P = \frac{2I}{c}$ (reflected)



NASA solar sail



Radiation Pressure

Radiometer: Vacuum bulb, 4 vanes, one white/one black side

What happens when light hits a black (or white) side of vane?

Ideal: Photons are absorbed on the black side of each vane, but are reflected on the white side, transferring more momentum to the white side, producing rotation in the direction of the black side with the white side trailing. YES! RADIATION PRESSURE!

Real life: Photons are absorbed by the black side, but mostly reflected by the white side, thus heating the black side more. The weaker the vacuum, the more pronounced the effect. Air molecules coming into contact with the vanes will obtain more heat energy from the black sides, colliding with greater momentum transfer from that side, thus producing rotation away from the black side.

in "real life" it doesn't turn in the direction you expect from radiation pressure arguments!

Radiometer

1- radiometer is glass bulb with partial vacuum

2- inside bulb, on low friction spindle, is rotor with 4 vertical metal vanes around axis

3- vanes polished white on one side, black on other

4- expose to sunlight, artificial light, or infrared radiation (even hand nearby enough)

5- vanes turn with no apparent motive power; light pressure far too small to move vanes

6- cooling radiometer causes rotation in opposite direction

7- with radiant energy source is heat engine which operate on temperature differences

8- black side of vane hotter than other side

Motion without external radiation

1- heat radiometer in absence of light source, turns with black sides trailing

2- cool glass quickly in absence of light source, turns with white sides are trailing

3- turns backwards because black sides give off more heat;cool quicker than white sides