Electricity and magnetism notes[1]

The impact of applications of electric and magnetic phenomena is apparent (eg lightbulbs, radio, telephones, computer circuits, etc.). In this course we will study the basic theory behind these impressive applications, which may all be summarized concisely in *Maxwell's equations* :

$$\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_o} \ , \quad \nabla \times \vec{E} = -\partial_t \vec{B} \ ,$$

$$\nabla \cdot \vec{B} = 0 \ , \quad \nabla \times \vec{B} = \mu_o \left(\vec{J} + \varepsilon_o \partial_t \vec{E} \right) \ .$$

We will devote ourselves both to a *physical* and *mathematical* understanding of what the quantities in these equations represent, by considering the experimental effects they encapsulate and applying some theory of partial differential equations to analyze in detail some particular situations and applications.

[1]Connor Jackman, connor.jackman@cimat.mx, for a 2022 course at CIMAT.

I. Statics

§1: calculus and linear algebra

We begin with a brief review of some relevant concepts from vector calculus and linear algebra (one may wish to jump straight to the next section and refer back to this section for reference).

It will be important during the course to write expressions in various coordinate systems. To emphasize this, we will base our review in a 3-dimensional vector space V with an inner product, $V \times V \to \mathbb{R}, (u, v) \mapsto u \cdot v$. Of course upon the *choice* of an orthonormal basis for V we may identify $V \cong \mathbb{R}^3 \ni (x, y, z)$ and for concreteness the reader may think of $\mathbb{R}^3$ with its standard dot product.

The inner product on V is equivalent to the Euclidean notions of length and angle. To pass from the geometric concepts to the algebraic inner product, one defines $u \cdot v$ as an (oriented) length of orthogonal projection of u onto the line spanned (and oriented) by v:

$$(*) \quad u \cdot v = |u||v| \cos\theta$$

where $\theta \in [0, \pi]$ is the angle between u and v and $|u|, |v|$ are the lengths of u, v. One may then establish that this operation is bilinear and symmetric, giving a (positive definite) inner product. Conversely, given an inner product, lengths of vectors and angles between vectors are defined using $(*)$.

Length, area, volume: The inner product on V determines distances and lengths of curves. The *distance* between two points $p, q \in V$ is defined as:

$$dist(p, q) := |p - q| = \sqrt{(p - q) \cdot (p - q)}.$$

The *length* of a curve, $\mathcal{C} \subset V$ from p to q is defined as:

$$length(\mathcal{C}) := \lim_{\varepsilon \to 0} \sum \Delta s_i.$$

taken over partitions $p = p_1 < p_2 < ... < p_n = q$ of the curve with $\Delta s_i := |p_{i+1} - p_i|$ and $\varepsilon := \max \Delta s_i$. When $\mathcal{C}$ is parametrized by $[t_o, t_1] \ni t \mapsto \gamma(t) \in \mathcal{C}$, one has:

$$length(\mathcal{C}) = \int_{t_o}^{t_1} |\dot\gamma(t)| \; dt$$

where $\dot\gamma(t) := \lim_{\varepsilon \to 0} \frac{\gamma(t+\varepsilon) - \gamma(t)}{\varepsilon} = \frac{d\gamma}{dt}(t) \in V$ is the *velocity* of the parametrized curve $\gamma(t)$. One writes $ds = |\dot\gamma(t)| \; dt$ called the *arc-length element* along the curve and lengths of curves are written as $\int_{\mathcal{C}} ds$.

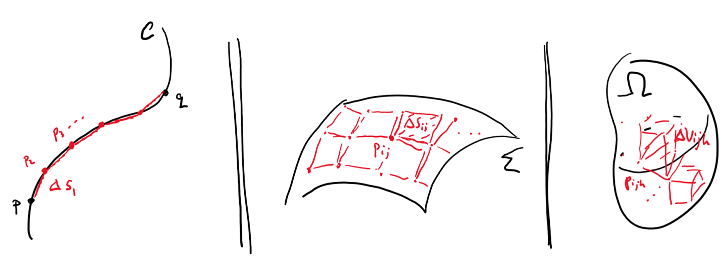

Figure 1. Lengths, areas and volumes are defined via approximation by sums of segment lengths, parallelogram areas and parallelpiped volumes. The curves surfaces and regions –rectifiable sets– we consider in this course will all be parametrized by smooth functions having well defined lengths, areas or volumes (the limits are well-defined) given by certain integrals.

Similarly, we obtain a notion of area of surfaces. First, the area of a parallelogram, $u \wedge v$, spanned by two vectors u, v is:

$$Area(u \wedge v) := |u \wedge v| = |u||v| \sin\theta = \sqrt{|u|^2|v|^2 - (u \cdot v)^2}.$$

The *area* of a surface, $\Sigma \subset V$, may be defined by:

$$Area(\Sigma) := \lim_{\varepsilon \to 0} \sum \Delta S_{i,j}$$

taken over partitions, $p_{i,j}$, ('grids') of the surface with $\Delta S_{i,j}$ the area of the parallelgram with vertex $p_{i,j}$ and $\varepsilon := \max \Delta S_{i,j}$. When Σ is parametrized by $[a_o, a_1] \times [b_o, b_1] \ni (a,b) \mapsto \varphi(a,b) \in \Sigma$, we have:

$$Area(\Sigma) = \int_{b_o}^{b_1} \int_{a_o}^{a_1} |\varphi_a \wedge \varphi_b| \; dadb$$

where $\varphi_a(a,b) := \lim_{\varepsilon \to 0} \frac{\varphi(a+\varepsilon,b) - \varphi(a,b)}{\varepsilon} = \frac{\partial \varphi}{\partial a}(a,b)$ and $\varphi_b(a,b) := \lim_{\varepsilon \to 0} \frac{\varphi(a,b+\varepsilon) - \varphi(a,b)}{\varepsilon} = \frac{\partial \varphi}{\partial b}(a,b)$ are the partial derivatives of $\varphi(a,b)$. One writes $dS = dA = |\varphi_a \wedge \varphi_b| \; dadb$ for the *area element* of the surface and surface area integrals are written $\iint_\Sigma \; dA$ or $\int_\Sigma \; dA$.

Volumes of regions are also determined. The volume of a parallelpiped, $u \wedge v \wedge w$, spanned by the vectors u, v, w is:

$$Vol(u \wedge v \wedge w) := |u \wedge v \wedge w| = |v \wedge w||u|| \cos \varphi| = |u \cdot (v \times w)|$$

where φ is the angle between u and the normal to the parallelogram $v \wedge w$. Here $v \times w \in V$ is the *cross product* of v and w, geometrically defined as a vector normal [1] to v, w with length $|v \wedge w|$. The *volume* of a region $\Omega \subset V$ is then:

$$Vol(\Omega) := \lim_{\varepsilon \to 0} \sum \Delta V_{i,j,k}$$

over partitions, $p_{i,j,k}$, ('grids') of Ω and with $\Delta V_{i,j,k}$ the volume of the paralleliped with vertex $p_{i,j,k}$ of the grid and $\varepsilon = \max \Delta V_{i,j,k}$. When Ω is parametrized by $\varphi(a,b,c)$ one has:

$$Vol(\Omega) = \int \int \int |\varphi_a \wedge \varphi_b \wedge \varphi_c| \; dadbdc$$

Where $\varphi_a, \varphi_b, \varphi_c$ are the partial derivatives of $\varphi(a,b,c)$. One writes $dV = |\varphi_a \wedge \varphi_b \wedge \varphi_c| \; dadbdc$ for the *volume element* in the region Ω, and volume integrals are written $\iiint_\Omega \; dV$ or $\int_\Omega \; dV$.

Mass, density: Curves, surfaces or regions are thought of as a continuous representations of physical objects. In reality, at a sufficiently small microscopic level, any physical object consists of collections of atomic particles. In the large –as is our common experience in interacting with physical objects– this microscopic detail is ignored and an object is idealized as a *continuum* or continuous subset of space.

Properties of these physical objects idealized as continua may be represented by 'densities' or functions on the continua. Let us consider mass and density of a continuum.

Consider some matter distributed in space. Then average densities of this matter over given regions, $\Omega \subset V$ (compact), may be given by:

$$\langle \rho \rangle_\Omega := \frac{mass(\Omega)}{Vol(\Omega)}$$

where $mass(\Omega)$ is the measure of the total mass of matter contained in Ω. In practice these measurements may be made over smaller and smaller regions $\Omega \ni p$ surrounding a given point, p, until there is no noticeable fluctuation in $\langle \rho \rangle_\Omega$ or the region Ω has become so small as to be indistinguishable from the point p. Then the density at p is defined as $\langle \rho \rangle_\Omega$. This 'ideal' density function, $\rho : V \to \mathbb{R}$, is then:

$$\rho(p) := \lim_{Vol(\Omega) \to 0} \frac{mass(\Omega)}{Vol(\Omega)}, \quad \Omega \ni p.$$

The density function or *volume density* associated to the matter is characterized by:

$$mass(\Omega) = \int_\Omega \rho \; dV$$

[1] Chosen by *right hand rule*: $v \times w$ points along the thumb of ones right hand when the index finger is directed along v and middle finger along w.

ie the total mass of the matter in some region Ω is given by averaging the density function over the region.

Likewise one has *surface densities* and *linear densities* as mass per unit area or unit length of a continua modeled as a surface or curve. They are given by $\sigma : \Sigma \to \mathbb{R}, \lambda : \mathcal{C} \to \mathbb{R}$ through:

$$\sigma(p) = \lim_{Area(D)\to 0} \frac{mass(D)}{Area(D)}, \quad \lambda(p) = \lim_{length(I)\to 0} \frac{mass(I)}{length(I)}$$

where $p \in D \subset \Sigma$ or $p \in I \subset \mathcal{C}$. For such subsets, one then has:

$$mass(D) = \int_D \sigma \; dA, \quad mass(I) = \int_I \lambda \; ds.$$

We may write for example $\rho \; dV = dm = d\mu$ (or $\sigma \; dA = d\mu$ or $\lambda \; ds = d\mu$) for the *mass element* of the matter.

Differentials, chain rule, change of variable: A function $f : V \to W$ between vector spaces has, at each $p \in V$, its linearization or *differential*:

$$d_p f : V \to W, \quad v \mapsto \frac{d}{dt}|_{t=0} f(c(t))$$

where $c(t)$ is a curve in V with $c(0) = p, \dot{c}(0) = v$. Essentially all functions we consider will be differentiable[1] so that $d_p f$ is always a well-defined linear map.

The *chain rule* gives the behaviour of differentials under composition. For $U \xrightarrow{f} V \xrightarrow{g} W$,

$$d_p(g \circ f) = \left(d_{f(p)} g\right) \left(d_p f\right).$$

As for integrals, the *change of variables* formula allows one to express an integral in different coordinates (parametrizations). The single variable case –substitution– reads:

$$\int_a^b f(x) \; dx = \int_{a'}^{b'} f(\varphi(y))\varphi'(y) \; dy$$

for $\varphi : [a', b'] \to [a, b], \; y \mapsto \varphi(y) = x$ bijective (and differentiable). The multivariable case reads:

$$\int_D f(x) \; d^n x = \int_{D'} f(\varphi(y)) |\det d_y \varphi| \; d^n y$$

for $\varphi : D' \subset \mathbb{R}^n \to D \subset \mathbb{R}^n, \; y \mapsto \varphi(y) = x$. Here $d^n x = dx_1 ... dx_n$ (and $d^n y = dy_1 ... dy_n$) are the standard volume elements on $\mathbb{R}^n$ and $\det d_y \varphi$ is the *Jacobian* of the transformation φ.

Vector fields, flows: A smooth choice of a vector based at each point is called a *vector field*, ie a (differentiable) map, $X : V \to V, p \mapsto X_p$.

A vector field has a corresponding first order ode, $\dot{p} = X_p$, and by uniqueness and existence theorem there is for each $p \in V$ a unique solution curve $p(t)$ with $p(0) = p$ of this ode (defined at least for $t \in (-\delta, \delta)$ some $\delta(p) > 0$). Unless otherwise mentioned, we will consider vector fields which are *complete*, meaning the solutions are defined for all time. Such a vector field may be visualized by its *flow*, for each $t \in \mathbb{R}$ we have $\varphi_t : V \to V, p \mapsto p(t) =: \varphi_t(p)$. These 'time t flow maps' are differentiable (smooth dependence of solutions on initial conditions), and satisfy $\varphi_0 = id, \varphi_{t+s} = \varphi_t \circ \varphi_s$ for any $t, s \in \mathbb{R}$ (in particular they are invertible).

Although the flow of a given vector field can seldom be found explicitely, its mere existence is a useful conceptual device. Namely one can imagine a fluid continuum flowing and deforming through space as a function of time. The 'streamlines' of this fluid are the curves traced by individual points of the fluid parametrized by time, and their velocities are the values of the vector field. Many properties and physical concepts related to vector fields, eg flux, are motivated by thinking in this way.

[1] That is, by definition, that there exists some linear map $d_p f$ s.t. $\lim_{|v|\to 0} \frac{f(p+tv)-f(p)-d_p f(v)}{|v|} = 0$.

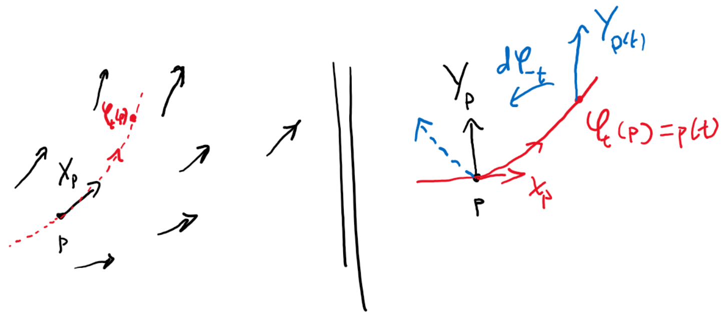

Figure 2. Vector fields may be visualized by their flows. The Lie derivative wrt X measures changes of objects under its flow.

Next, let us mention a fundamental type of derivative or 'operator' determined by a vector field, the *Lie derivative*. Roughly the Lie derivative (wrt the vector field X) of an object measures how the object changes under the flow of X.

The Lie derivatives of a wide class of objects may be defined (namely general 'tensor fields'), here let us give the construction for two special cases. Given a function, $f : V \to \mathbb{R}$, the Lie derivative of f wrt the vector field X is the function:

$$Xf = \mathscr{L}_X f : V \to V, \quad (Xf)(p) := \frac{d}{dt}|_{t=0} f(\varphi_t(p))$$

where φ_t is the flow of X. Note that $(Xf)(p) = d_p f(X_p)$. Similarly given another vector field, $Y : V \to V$, its Lie derivative wrt X is the vector field:

$$\mathscr{L}_X Y = [X, Y] : V \to V, \quad (\mathscr{L}_X Y)_p := \frac{d}{dt}|_{t=0} \left(d_{p(t)}\varphi_{-t}\right)\left(Y_{p(t)}\right)$$

where $p(t) = \varphi_t(p)$. Note that as well $(\mathscr{L}_X Y)_p = \frac{d^2}{dtds}|_{t=s=0} \varphi_{-t} \circ \psi_s \circ \varphi_t(p)$, where ψ_s is the flow of Y, so that in this sense Lie derivative measures how the flows of X and Y commute.

Via the visualization of a vector field by its flow we may also imagine measuring changes in integrals or averages over regions as the regions are deformed under the flow. The divergence of a vector field can be motivated in this way as measuring rate of volume change under the flow. Namely, for a region Ω, consider the regions $\Omega_t := \varphi_t(\Omega)$. Using the change of variables formula one may establish:

$$\frac{d}{dt}|_{t=0} Vol(\Omega_t) = \int_\Omega div(X)\ dV$$

where $div(X) : V \to \mathbb{R}$ is the function defined by $p \mapsto tr(d_p X)$.

Line, surface integrals: In addition to density integrals, one meets other integrals over curves and surfaces that are orientation dependent.

We will first describe line integrals in terms of flow. Given a vector field X and oriented curve, $\mathcal{C}$, think of X as representing a flowing fluid, $p \mapsto \varphi_t(p)$, and imagine at a given instant a small tube[1] is placed around the curve $\mathcal{C}$. The fluid will now be restricted in this tube to move along it, and we ask what is the average velocity of the fluid in this tube (counted positive if in the direction $\mathcal{C}$ is oriented and negative if opposite). Letting T be the unit tangent vector to $\mathcal{C}$ directed along its orientation, this *line integral* is then:

$$\int_{\mathcal{C}} X \cdot T\ ds.$$

When a vector field F represents a force field, line integrals of F along a curve have the important interpretation as giving the *work* done by the forces in moving an object along the path $\mathcal{C}$.

[1]The ends of this imagined tube should have some kind of valves which allow fluid out of but not into the tube.

Surface or 'flux' integrals may also be interpreted nicely in terms of flow. Let X be a vector field with flow φ_t and Σ a surface oriented with unit normal ν. We would like to measure the rate of fluid passing through Σ (positive in the direction of the chosen normal and negative if opposite). Letting Ω_t be the region $\{\varphi_s(\Sigma) : 0 \leq s \leq t\}$ this flux of X across Σ is then the rate of change of the oriented volume: $\frac{d}{dt}|_{t=0}Vol_o(\Omega_t)$, which one computes is given by:

$$\int_\Sigma X \cdot \nu \; dA.$$

Differential operators: The line, surface and (oriented) density integrals have relations generalizing the fundamental theorem of calculus in terms of certain operators ('grad, curl, div').

The *gradient* of a function $f : V \to \mathbb{R}$ is a vector field $\nabla f = grad(f)$ on V defined at $p \in V$ through:

$$\nabla_p f \cdot v = d_p f(v)$$

for every $v \in V$. This vector thus determines rates of change or directional derivatives of the function f. Note that for curves $\mathcal{C} \ni p$, containing p and with fixed (unit) tangent direction $\hat{v}$ at p, we have:

$$\nabla_p f \cdot \hat{v} := \lim_{length(\mathcal{C}) \to 0} \frac{f(p_1) - f(p_o)}{length(\mathcal{C})}$$

where p_o, p_1 are the endpoints of $\mathcal{C}$. The integral theorem for gradients and line integrals is:

$$\int_{\mathcal{C}} \nabla f \cdot T \; ds = f(p_1) - f(p_o).$$

Similarly the *curl* of a vector field $X : V \to V$ is a vector field $\nabla \times X = curl(X)$ defined at $p \in V$ through:

$$(\nabla \times X)_p \cdot \hat{n} := \lim_{Area(\Sigma) \to 0} \frac{\oint_{\partial\Sigma} X \cdot T \; ds}{Area(\Sigma)}$$

where $\Sigma \ni p$ is an oriented surface containing p with unit normal $\hat{n}$ at p. The boundary line integral in the numerator is also called the *circulation* of X around the closed curve $\partial\Sigma$. It measures rate of rotation of the flow of X along this closed curve.

The integral theorem –*Stoke's theorem*– for curls and surface integrals is:

$$\int_\Sigma (\nabla \times X) \cdot \nu \; dA = \int_{\partial\Sigma} X \cdot T \; ds.$$

Also the *divergence* of a vector field, X, is a function $\nabla \cdot X = div(X)$ defined at $p \in V$ through:

$$(\nabla \cdot X)(p) := \lim_{Vol(\Omega) \to 0} \frac{\oint_{\partial\Omega} X \cdot \nu \; dA}{Vol(\Omega)}$$

where $\Omega \ni p$ is a region containing p. The flux integral in the numerator measures the rate of volume change of the flow of X through $\partial\Omega$, and so the divergence may be interpreted as a density giving the rate of volume change under the flow of X (equivalent to our definition with Lie derivatives above).

The integral theorem –*Gauss' theorem*– for divergence and density integrals is:

$$\int_\Omega \nabla \cdot X \; dV = \int_{\partial\Omega} X \cdot \nu \; dA.$$

In the course, we will also meet some other fundamental differential operators two of which we mention now. The *Laplacian* of a function, is another function

$$\Delta f := div(grad(f)) = \nabla \cdot (\nabla f).$$

The *Hessian* of a function is related to its second differential, defined at each $p \in V$ as the (symmetric) bilinear form:

$$d_p^2 f(u, v) := \frac{d^2}{dtds}|_{t=s=0} f(p + tu + sv).$$

Then the Hessian is the symmetric linear map: $H(f)_p : V \to V$ defined through:

$$d_p^2 f(u, v) = (H(f)_p u) \cdot v.$$

One has for instance that $tr(H(f)_p) = \Delta f(p)$.

Tensors, forms: The basic notions of vector calculus are also expressed in the language of differential forms. This efficient notation naturally generalizes to manifolds of arbitrary dimensions (in this class we will consider the case $n = 4$ of 'space-time'). Let us begin with a quick recap of the relevant linear algebra.

The *dual space* to a vector space V is the set V^* of all linear maps (functionals) from $V \to \mathbb{R}$. The space V^* is itself a vector space, with 'pointwise' addition and scalar multiplication: given $\alpha, \beta \in V^*$ and $\lambda \in \mathbb{R}$ then $\alpha + \lambda\beta : V \to \mathbb{R}, v \mapsto \alpha(v) + \lambda\beta(v)$. We denote the *natural pairing* of evaluation between vector space and dual by $\alpha(v) =: (\alpha, v)$ for $\alpha \in V^*, v \in V$.

Observe there is a canonical (no choice of basis needed) isomorphism of $V^{**} \cong V$, by $v \mapsto i_v,\ \ i_v(\alpha) = (\alpha, v)$. Without further structure on V there is no canonical isomorphism between V and V^*. However given a basis $v_1, ..., v_n$ of V, there is a corresponding *dual basis*, denoted $v^1, ..., v^n$, for V^* by $(v^k, v_j) := \delta_j^k$.

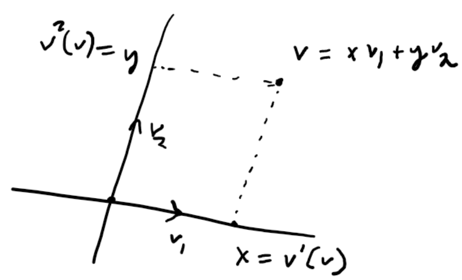

Figure 3. The dual basis to a given basis of a vector space are the projections onto the coordinate axes.

A non-degenerate scalar product, $\langle \cdot, \cdot \rangle$, on V determines *musical isomorphisms* between V and V^* by:

$$V \xrightarrow{\flat} V^*, \quad v^\flat(w) = \langle v, w \rangle$$

with inverse denoted $V^* \xrightarrow{\sharp} V, \quad \alpha \mapsto \alpha^\sharp$. Conversely a symmetric isomoprhism $L : V \to V^*$ determines an inner product $v \cdot w := (Lv, w)$.

Now we consider bilinear forms in the language of *tensors*. A *bilinear form* on V is a map:

$$\beta : V \times V \to \mathbb{R}$$

s.t. for any fixed $v \in V$ the maps $V \to \mathbb{R}$ by $w \mapsto \beta(v, w), w \mapsto \beta(w, v)$ are linear, ie elements of V^*. We call the form *symmetric* (resp. *skew-symmetric*) if $\beta(v, w) = \beta(w, v)$ (resp. $\beta(v, w) = -\beta(w, v)$) for all $v, w \in V$.

The set of bilinear forms on V is itself a vector space, with pointwise addition and scalar multiplication. This vector space is denoted:

$$V^* \otimes V^*$$

and also called the *tensor product*[1] of V^* with V^*.

Given a basis, $v_1, ..., v_n$, for V a bilinear form has a matrix representation. For $x = x^1 v_1 + ... + x^n v_n, y = y^1 v_1 + ... + y^n v_n$, we have by bilinearity the expansion:

$$\beta(x, y) = \sum \beta_{ij} x^i y^j, \quad \beta_{ij} := \beta(v_i, v_j)$$

[1] Tensor products may be also defined via a universal property.

or in matrix form:

$$\beta(x,y) = \begin{pmatrix} x^1 & \dots & x^n \end{pmatrix} \begin{pmatrix} \beta_{11} & \dots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \dots & \beta_{nn} \end{pmatrix} \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix}.$$

For $v^1, ..., v^n$ the corresponding dual basis to V^*, we define:

$$v^i \otimes v^j : V \times V \to \mathbb{R}, \quad (x, y) \mapsto v^i(x)v^j(y) = x^i y^j.$$

These bilinear forms then form a basis for $V^* \otimes V^*$, and the matrix or coordinate expansion of a general bilinear form is given by:

$$\beta = \sum \beta_{ij} v^i \otimes v^j.$$

Note that when β is symmetric, we have:

$$\beta = \sum \beta_{ij} v^i v^j$$

where $v^i v^j := \frac{v^i \otimes v^j + v^j \otimes v^i}{2} = v^j v^i$, and when β is skew symmetric that:

$$\beta = \sum_{i<j} \beta_{ij} v^i \wedge v^j$$

where $v^i \wedge v^j := v^i \otimes v^j - v^j \otimes v^i = -v^j \wedge v^i$. The vector space of symmetric bilinear forms on V is denoted $Sym^2(V^*) \subset V^* \otimes V^*$ and skew-symmetric bilinear forms on V by $\bigwedge^2 V^* \subset V^* \otimes V^*$.

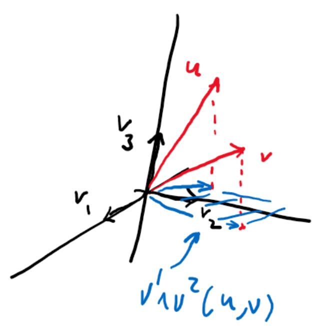

Figure 4. Anti-symmetric bilinear forms return (oriented) areas of projections onto coordinate planes. This geometric interpretation still holds for anti-symmetric k-multi-linear forms.

Similarly, one has k-multilinear forms, $\otimes^k V^* := \overbrace{V^* \otimes ... \otimes V^*}^{k-times}$ or tensor products $V^* \otimes W^*$ between vector spaces V and W as the vector space of bilinear forms $V \times W \to \mathbb{R}$.

A non-degenerate scalar product on V determines scalar products on its tensor products via the musical isomorphisms. Namely:

$$\otimes^k V \xrightarrow{\flat} (\otimes^k V)^* = \otimes^k V^*$$

sending a k-multilinear map B on V^* to the k-multilinear map $B^\flat$ on V by $B^\flat(v_1, ..., v_k) := B(v_1^\sharp, ..., v_k^\sharp)$. The inverse of $\flat$ is again denoted by $\sharp$. Then one has (non-degenerate) scalar products:

$$\langle A, B \rangle := (B^\flat, A), \quad \langle \alpha, \beta \rangle := (\alpha, \beta^\sharp)$$

for $A, B \in \otimes^k V, \alpha, \beta \in \otimes^k V^*$. Note that by restriction, one obtains scalar products on the symmetric or skew-symmetric k-multilinear forms on V (denoted $Sym^k(V^*), \bigwedge^k V^*$).

Observe that when V has dimension n then $\bigwedge^k V = \bigwedge^k V^* = 0$ for $k > n$ while $\bigwedge^n V, \bigwedge^n V^*$ are 1-dimensional (determined by the scalar value of such an n-form on a given basis). When V has a non-degenerate scalar product, $\langle \cdot, \cdot \rangle$ with signature σ, an (oriented) volume form on V may be defined to be an element $\omega_{vol} \in \bigwedge^n V^*$ such that $\langle \omega_{vol}, \omega_{vol} \rangle = (-1)^\sigma$. Given an oriented volume form, ω_{vol}, there is only one other choice, namely $-\omega_{vol}$. An ordered orthonormal basis $e_1, ..., e_n$, of V, determines:

$$\omega_{vol} := e^1 \wedge .... \wedge e^n.$$

The choice of an orientation of V, in addition to the scalar products, determines the *Hodge-$*$* operation on skew-symmetric forms. For $dim(V) = n$, this operator is given by

$$\bigwedge{}^k V^* \xrightarrow{*} \bigwedge{}^{n-k} V^*, \quad \beta \mapsto *\beta$$

through $\alpha \wedge (*\beta) = \langle \alpha, \beta \rangle \, \omega_{vol}$ for any $\alpha \in \bigwedge^k V^*$.

We have used above the notation $u \wedge v$ for oriented parallelograms in the 3-dimensional V with an inner product. Let us explain now this notation. An element of $u \wedge v \in \bigwedge^2 V$ represents a skew-symmetric bilinear map $V^* \times V^* \to \mathbb{R}$. Since V^* is 3-dimensional, unless $u \wedge v \equiv 0$, there is a 1-dimensional kernel

$$\alpha \in V^*, \ \text{ s.t. } (u \wedge v)(\alpha, \beta) = (\alpha, u)(\beta, v) - (\alpha, v)(\beta, u) = 0, \ \forall \beta \in V^*.$$

Then for such a non-zero $\alpha \in \ker(u \wedge v)$, the plane $\ker \alpha \subset V$ is that spanned by u and v and $|u \wedge v|$ is the (oriented) area of the parallelogram spanned by u, v. Thus oriented areas in 3-dimensional Euclidean space are represented by *bivectors* (elements of $\bigwedge^2 V$)[1].

Now we come to *differential forms.* A 1-form on V is –similarly to a vector field– a smooth choice of a dual vector in V^* at each point of V, ie a (differentiable) map $\omega : V \to V^*, p \mapsto \omega_p$. For example the differential of a function, $p \mapsto d_p f$. These 1-forms are objects which may be integrated along (oriented) curves. Namely for C parametrized by $\gamma(t)$, we take:

$$\int_C \omega := \int_{t_o}^{t_1} \omega(\dot{\gamma}) \, dt$$

Similarly a 2-form on V is a smooth choice of skew-symmetric bilinear forms on V at each point of V, ie $\omega : V \to \bigwedge^2 V^*$. These 2-forms are objects which may be integrated along oriented surfaces. For an oriented surface, Σ, parametrized by $\varphi(a, b)$, with $(a, b) \in D \subset \mathbb{R}^2$, we take:

$$\int_\Sigma \omega := \int_D \omega(\varphi_a, \varphi_b) \, dadb.$$

A 3-form on V is a differentiable map $\omega : V \to \bigwedge^3 V^*$ and may be integrated over oriented regions. Note that $\omega = \rho \omega_{vol}$ for some $\rho : V \to \mathbb{R}$. When the region Ω is parametrized by $\varphi(a, b, c)$, with $(a, b, c) \in U \subset \mathbb{R}^3$, we take

$$\int_\Omega \omega := \int_U \omega(\varphi_a, \varphi_b, \varphi_c) \, dadbdc.$$

It should be clear that this type of integration generalizes easily to arbitrary dimensions.

We will now give an explicit coordinate expression to see that integrals of 2-forms are certain surface integrals. First note that –as we have used implicitly at times above– a basis of V 'slides' to give a basis at every tangent space to V at a point $p \in V$. Let e_1, e_2, e_3 be an orthonormal basis and denote the coordinates

$$\mathbb{R}^3 \ni (x, y, z) \longleftrightarrow xe_1 + ye_2 + ze_3 \in V.$$

The corresponding dual basis is denoted dx, dy, dz, eg $dx(x, y, z) = x$. A 2-form is then given by:

$$\omega = \omega_1 \, dy \wedge dz + \omega_2 \, dz \wedge dx + \omega_3 \, dx \wedge dy$$

for some functions $\omega_j : \mathbb{R}^3 \to \mathbb{R}$.

[1] In dimension 3, a (non-zero) bivector $B : V^* \times V^* \to \mathbb{R}$ has a 1-dimensional kernel, $\ker(B) = \{\alpha : B(\alpha, \cdot) \equiv 0\}$ determining the plane $\ker \alpha \subset V$ and the parallelograms in this plane with area $|B|$.

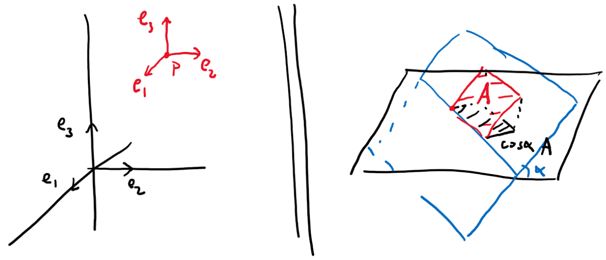

Figure 5. A basis for a vector space may be translated to a given point to form a basis for the vectors based at that point. An area element in a given plane projects to a multiple of the area element in the coordinate planes.

The area elements along a surface may be thought of as functions on parallelograms tangent to the surface at a given point: $dA(u \wedge v) = |u \wedge v|$. When the surface is oriented by unit normal ν, then when evaluated on oriented parallelograms we have:

$$\cos\alpha_1 \; dA = dy \wedge dx, \quad \cos\alpha_2 \; dA = dz \wedge dx, \quad \cos\alpha_3 \; dA = dx \wedge dy$$

where α_j's are the angles between ν and the basis vectors e_j. Since ν is unit, we have:

$$\int_\Sigma \omega = \int_\Sigma (\omega_1 \cos\alpha_1 + \omega_2 \cos\alpha_2 + \omega_3 \cos\alpha_3) \; dA = \int_\Sigma \vec{\omega} \cdot \nu \; dA$$

for the vector field $\vec{\omega} := \omega_1 e_1 + \omega_2 e_2 + \omega_3 e_3$. Hence integrals of 2-forms are the same as surface (flux) integrals. Observe that we may interpret these surface integrals as weighted sums of oriented areas of projections to the coordinate planes. This generalizes to allow geometric interpretations of integrals of k-forms as weighted sums of projected oriented k-volumes. Explicitely, for a surface parametrized by $\varphi(a,b) = (x(a,b), y(a,b), z(a,b))$:

$$\int \vec{\omega} \cdot (\varphi_a \times \varphi_b) \; dadb = \int \omega_1 \det \begin{pmatrix} y_a & y_b \\ z_a & z_b \end{pmatrix} + \omega_2 \det \begin{pmatrix} x_b & x_a \\ z_b & z_a \end{pmatrix} + \omega_3 \det \begin{pmatrix} x_a & x_b \\ y_a & y_b \end{pmatrix} \; dadb$$

The integral theorems we have met above may be summarized via the *exterior derivative* operation on differential forms. Let ω be a differential k-form on V and define a differential $k+1$ form, $d\omega$, on V by:

$$d\omega_p(v_1, ..., v_{k+1}) := \lim_{\varepsilon \to 0} \frac{\int_{\partial P_\varepsilon} \omega}{\varepsilon^{k+1}}$$

where P_ε is a parallelogram with sides $\varepsilon v_1, ..., \varepsilon v_k$ based at $p \in V$. One then has the *generalized Stoke's theorem*:

$$\int_R d\omega = \int_{\partial R} \omega$$

where ω is a k-form and R a $k+1$-dimensional oriented region with boundary ∂R.

The Lie derivatives we have met above of vector fields and functions may also be defined for differential k-forms. For a vector field X the Lie derivative of a k-form ω is a k-form $\mathscr{L}_X \omega$ which measures the change of ω under the flow of X by:

$$\mathscr{L}_X \omega_p(v_1, ..., v_k) := \frac{d}{dt}|_{t=0} \omega_{\varphi_t(p)}(d_p\varphi_t v_1, ..., d_p\varphi_t v_k).$$

One has the important transport equation, $\frac{d}{dt}|_{t=0} \int_{\varphi_t(R)} \omega = \int_R \mathscr{L}_X \omega$.

A scalar product (and orientation) of V induces a great deal of structure on the differential forms of V. First there is a Hodge-$*$ operator taking k forms to $n-k$ forms, defined pointwise, $\alpha \wedge (*\beta) = \langle \alpha, \beta \rangle \omega_{vol}$.

This operator determines a *co-derivative* to the exterior derivative, denotes δ, taking $n-k$ forms to $n-k-1$ forms by:

$$\delta := (-1)^k * \circ\ d\ \circ *^{-1}, \quad \delta\omega := (-1)^k * d(*^{-1}\omega).$$

The reason for the sign conventions is that –when the integrals have sense – one has an inner product $\langle\langle\cdot,\cdot\rangle\rangle$ on differential forms:

$$\langle\langle\alpha,\beta\rangle\rangle := \int \alpha \wedge *\beta$$

for which:

$$\langle\langle d\alpha,\beta\rangle\rangle = \langle\langle\alpha,\delta\beta\rangle\rangle.$$

The important Laplacian operator may be extended from functions to differential forms by:

$$\Delta := d\circ\delta + \delta\circ d$$

A fundamental decomposition theorem, the *Hodge decomposition*, describes the vector space of differential k-forms on an oriented scalar product space V, $\Omega^k(V)$ by:

$$\Omega^k(V) = im(d) \oplus im(\delta) \oplus \mathscr{H}^k(V)$$

where $d : \Omega^{k-1}(V) \to \Omega^k(V), \delta : \Omega^{k+1}(V) \to \Omega^k(V)$ are the exterior derivative and co-derivative, while $\mathscr{H}^k(V) = \{\omega : \Delta\omega = 0\}$ are the *harmonic* differential k-forms. Moreover, the sum above is orthogonal with respect to the scalar product $\langle\langle\cdot,\cdot\rangle\rangle$. In this course we will see some applications of this theorem.

Coordinates: To derive explicit formulas, one chooses coordinates to parametrize regions of space. We present here the formulas above in some standard coordinate systems.

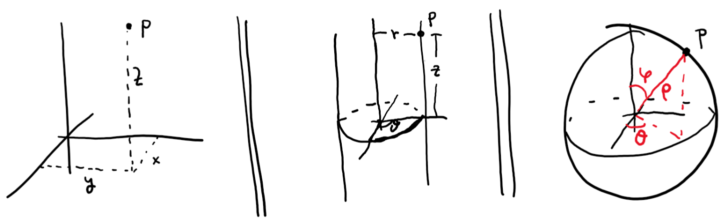

Figure 6. Cartesian, cylindrical, or spherical coordinates parametrize points in space.

Cartesian coordinates, (x,y,z), are determined by choosing an orthonormal basis, e_1, e_2, e_3 for V:

$$\mathbb{R}^3 \ni (x,y,z) \longleftrightarrow xe_1 + ye_2 + ze_3 \in V.$$

Given a parametrized curve, $\gamma(t) = x(t)e_1 + y(t)e_2 + z(t)e_3$, we then have:

$$|\dot\gamma|^2 = \dot x^2 + \dot y^2 + \dot z^2$$

and the arc-length of curves is given by:

$$\int_C ds = \int \sqrt{\dot x^2 + \dot y^2 + \dot z^2}\ dt$$

where $ds = \sqrt{ds^2}$ and for dx, dy, dz the dual basis to e_1, e_2, e_3:

$$ds^2 = dx^2 + dy^2 + dz^2.$$

The oriented volume element is:

$$dV_o = dx \wedge dy \wedge dz = \omega_{vol}$$

and $dV = |dV_o|$. A parametrized region, $\varphi(u,v,w) = (x(u,v,w), y(u,v,w), z(u,v,w))$ then has volume:

$$\int_\Omega dV = \int |\det(\varphi_u, \varphi_v, \varphi_w)|\ dudvdw$$

and oriented volume:

$$\int_\Omega dV_o = \int \det(\varphi_u, \varphi_v, \varphi_w)\ dudvdw.$$

A surface with unit normal $\nu = \nu^1 e_1 + \nu^2 e_2 + \nu^3 e_3$ has oriented area element:

$$dA_o = i_\nu \omega_{vol} = \nu^1\ dy \wedge dz + \nu^2\ dz \wedge dx + \nu^3\ dx \wedge dy$$

and area element $dA = |dA_o|$. A parametrized surface $(x(u,v), y(u,v), z(u,v))$ then has area:

$$\int_\Sigma dA = \int \sqrt{(y_u z_v - y_v z_u)^2 + (z_u x_v - z_v x_u)^2 + (x_u y_v - x_v y_u)^2}\ dudv.$$

The differential of a function, $f(x,y,z)$, at the point p is given in this basis by

$$d_p f(v^1 e_1 + v^2 e_2 + v^3 e_3) = (f_x v^1 + f_y v^2 + f_z v^3)|_p$$

so in the dual basis or matrix form by:

$$(f_x\ dx + f_y\ dy + f_z\ dz)|_p, \qquad \text{or} \qquad \begin{pmatrix} f_x & f_y & f_z \end{pmatrix}|_p.$$

For a vector field, $X = X^1 e_1 + X^2 e_2 + X^3 e_3$, line integrals are given by:

$$\int_C X \cdot T\ ds = \int X(\gamma(t)) \cdot \dot{\gamma}(t)\ dt = \int_C X^1\ dx + X^2\ dy + X^3\ dz$$

They are also written $\int_C X \cdot d\vec{s}$. Surface integrals are given by:

$$\int_\Sigma X \cdot \nu\ dA = \int X(\varphi(u,v)) \cdot (\varphi_u \times \varphi_v)\ dudv = \int_\Sigma X^1\ dy \wedge dz + X^2\ dz \wedge dx + X^3\ dx \wedge dy.$$

They are also written $\int_\Sigma X \cdot d\vec{S}$. Oriented density integrals may be written $\int_\Omega \rho\ dx \wedge dy \wedge dz$.
The differential operators are given in these coordinates by:

$$\nabla f = f_x\ e_1 + f_y\ e_2 + f_z\ e_3, \quad \nabla \cdot X = X^1_x + X^2_y + X^3_z$$

$$\nabla \times X = (X^3_y - X^2_z)\ e_1 + (X^1_z - X^3_x)\ e_2 + (X^2_x - X^1_y)\ e_3.$$

Note that the musical isomorphism and Hodge-$*$ operators are given by:

$$(e_1)^\flat = dx, \quad (e_2)^\flat = dy, \quad (e_3)^\flat = dz$$

$$*dx = dy \wedge dz, \quad *\,dy = dz \wedge dx, \quad *\,dz = dx \wedge dy$$

And $*\omega_{vol} = 1$ (here $* \circ * = id$, so for instance $*(dy \wedge dz) = dx$). They are defined similarly on $\bigwedge^k V$. We have:

$$*(u \wedge v) = u \times v = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ u^1 & u^2 & u^3 \\ v^1 & v^2 & v^3 \end{pmatrix}$$

allowing one to remember the formulas for the operators by placing $\nabla = \partial_x\ e_1 + \partial_y\ e_2 + \partial_z\ e_3$.
The musical isomorphisms and Hodge-$*$ allow us to associate to a function f or vector field X the differential forms:

$$*f = f\ \omega_{vol}, \quad X^\flat =: \omega^1_X, \quad *\,(X^\flat) =: \omega^2_X$$

Note that $\omega_X^2 = i_X\omega_{vol}$. Then for instance:

$$df = \omega^1_{\nabla f}, \quad d\omega^1_X = \omega^2_{\nabla\times X}, \quad d\omega^2_X = (\nabla\cdot X)\omega_{vol}$$

and one may write our integral theorems as:

$$\int_{\mathcal{C}} df = f|_{\partial\mathcal{C}}, \quad \int_\Sigma d\omega^1_X = \int_{\partial\Sigma}\omega^1_X, \quad \int_\Omega d\omega^2_X = \int_{\partial\Omega}\omega^2_X.$$

The Laplacian, $\Delta f = *d*df$, is given by:

$$\Delta f = f_{xx} + f_{yy} + f_{zz}.$$

Cylindrical coordinates, (r,θ,z), relate to Cartesian coordinates by:

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = z.$$

Points $p \in V$ are given by:

$$p(r,\theta,z) = r\ e_r + z\ e_3$$

where $e_r := \cos\theta e_1 + \sin\theta e_2$. Set $p_\theta := r\partial_\theta(e_r)$ and note that $p_r = e_r, p_\theta, p_z = e_3$ form an oriented *orthogonal* basis at each point (when $r \neq 0$). For $dr, d\theta, dz$ the dual basis:

$$ds^2 = dr^2 + r^2 d\theta^2 + dz^2$$

$$dV_o = r\ dr\wedge d\theta\wedge dz$$

$$dA_o = r\left(\nu^r\ d\theta\wedge dz + \nu^\theta\ dz\wedge dr + \nu^z\ dr\wedge d\theta\right)$$

when the unit normal is written $\nu = \nu^r p_r + \nu^\theta p_\theta + \nu^z p_z$. Also:

$$\int_\Sigma dA = \int\sqrt{r^2(\theta_u z_v - \theta_v z_u)^2 + (z_u r_v - z_v r_u)^2 + r^2(r_u\theta_v - r_u\theta_v)^2}\ dudv$$

$$\int_{\mathcal{C}} X\cdot T\ ds = \int_{\mathcal{C}} X^r\ dr + r^2 X^\theta\ d\theta + X^z\ dz$$

$$\int_\Sigma X\cdot\nu\ dA = \int_\Sigma r\left(X^r\ d\theta\wedge dz + X^\theta\ dz\wedge dr + X^z\ dr\wedge d\theta\right)$$

where $X = X^r p_r + X^\theta p_\theta + X^z p_z$. The differential operators are:

$$\nabla f = f_r\ p_r + \frac{f_\theta}{r^2}\ p_\theta + f_z\ p_z$$

$$\nabla\times X = \frac{1}{r}\det\begin{pmatrix} p_r & p_\theta & p_z \\ \partial_r & \partial_\theta & \partial_z \\ X^r & r^2X^\theta & X^z\end{pmatrix}$$

$$\nabla\cdot X = \frac{1}{r}\partial_r(rX^r) + \partial_\theta(X^\theta) + \partial_z(X^z)$$

$$\Delta f = \frac{1}{r}\partial_r(rf_r) + \frac{1}{r^2}f_{\theta\theta} + f_{zz}$$

Spherical coordinates, (ρ,φ,θ), relate to Cartesian coordinates by:

$$x = \rho\sin\varphi\cos\theta, \quad y = \rho\sin\varphi\sin\theta, \quad z = \rho\cos\varphi.$$

Points $p \in V$ are given by:

$$p(\rho,\varphi,\theta) = \rho\ e_\rho$$

where $e_\rho := \sin\varphi\cos\theta e_1 + \sin\varphi\sin\theta e_2 + \cos\varphi e_3$. We set: $p_\varphi := \rho\partial_\varphi(e_\rho)$, $p_\theta := \rho\partial_\theta(e_\rho)$ and note that $p_\rho = e_\rho, p_\varphi, p_\theta$ form an oriented *orthogonal* basis at each point (when $\varphi \not\equiv_\pi 0$). For $d\rho, d\varphi, d\theta$ the dual basis:

$$ds^2 = d\rho^2 + \rho^2\ d\varphi^2 + \rho^2\sin^2\varphi\ d\theta^2,$$

$$dV_o = \rho^2\sin\varphi\ d\rho\wedge d\varphi\wedge d\theta,$$

$$dA_o = \rho^2\sin\varphi\left(\nu^\rho\ d\varphi\wedge d\theta + \nu^\varphi\ d\theta\wedge d\rho + \nu^\theta\ d\rho\wedge d\varphi\right)$$

when the unit normal is written $\nu = \nu^\rho p_\rho + \nu^\varphi p_\varphi + \nu^\theta p_\theta$. Also:

$$\int_\Sigma dA = \int \rho\ \sqrt{\rho^2\sin^2\varphi(\varphi_u\theta_v - \varphi_v\theta_u)^2 + \sin^2\varphi(\theta_u\rho_v - \theta_v\rho_u)^2 + (\rho_u\varphi_v - \rho_v\varphi_u)^2}\ dudv$$

$$\int_C X\cdot T\ ds = \int_C X^\rho\ d\rho + \rho^2 X^\varphi\ d\varphi + \rho^2\sin^2\varphi X^\theta\ d\theta$$

$$\int_\Sigma X\cdot\nu\ dA = \int_\Sigma \rho^2\sin\varphi\left(X^\rho\ d\varphi\wedge d\theta + X^\varphi\ d\theta\wedge d\rho + X^\theta\ d\rho\wedge d\varphi\right)$$

where $X = X^\rho p_\rho + X^\varphi p_\varphi + X^\theta p_\theta$. The differential operators are:

$$\nabla f = f_\rho\ p_\rho + \frac{f_\varphi}{\rho^2}\ p_\varphi + \frac{f_\theta}{\rho^2\sin^2\varphi}\ p_\theta$$

$$\nabla\times X = \frac{1}{\rho^2\sin\varphi}\det\begin{pmatrix} p_\rho & p_\varphi & p_\theta \\ \partial_\rho & \partial_\varphi & \partial_\theta \\ X^\rho & \rho^2 X^\varphi & \rho^2\sin^2\varphi X^\theta \end{pmatrix}$$

$$\nabla\cdot X = \frac{1}{\rho^2}\partial_\rho(\rho^2 X^\rho) + \frac{1}{\sin\varphi}\partial_\varphi(\sin\varphi X^\varphi) + \partial_\theta(X^\theta)$$

$$\Delta f = \frac{1}{\rho^2}\partial_\rho(\rho^2 f_\rho) + \frac{1}{\rho^2\sin\varphi}\partial_\varphi(\sin\varphi f_\varphi) + \frac{1}{\rho^2\sin^2\varphi} f_{\theta\theta}.$$

The coordinate systems above are examples of what are called orthogonal coordinates. In general a system of coordinates, $p(u,v,w)\in V$ are called *orthogonal coordinates* when its coordinate curves are orthogonal, ie p_u, p_v, p_w are an oriented orthogonal basis at each point.

Set $a = |p_u|, b = |p_v|, c = |p_w|$ and let du, dv, dw be the dual basis to p_u, p_v, p_w. Then:

$$ds^2 = a^2\ du^2 + b^2\ dv^2 + c^2\ dw^2$$

$$dV_o = abc\ du\wedge dv\wedge dw$$

$$\int_C X\cdot T\ ds = \int_C X^u\ a^2 du + X^v\ b^2 dv + X^w\ c^2 dw$$

$$\int_\Sigma X\cdot\nu\ dA = \int_\Sigma abc\,(X^u\ dv\wedge dw + X^v\ dw\wedge du + X^w\ du\wedge dv)$$

$$\nabla f = \frac{f_u}{a^2}\ p_u + \frac{f_v}{b^2}\ p_v + \frac{f_w}{c^2}\ p_w$$

$$(abc)\ \nabla\times X = \det\begin{pmatrix} p_u & p_v & p_w \\ \partial_u & \partial_v & \partial_w \\ a^2 X^u & b^2 X^v & c^2 X^w \end{pmatrix}$$

$$(abc)\ \nabla\cdot X = \partial_u(abcX^u) + \partial_v(abcX^v) + \partial_w(abcX^w)$$

$$(abc)\ \Delta f = \partial_u(\frac{bc}{a} f_u) + \partial_v(\frac{ac}{b} f_v) + \partial_w(\frac{ab}{c} f_w)$$

for $X = X^u\ p_u + X^v\ p_v + X^w\ p_w$. We remark that the formulas are often written in the *orthonormal basis*:

$$e_u := p_u/a, \quad e_v := p_v/b, \quad e_w := p_w/c.$$

So eg, $X = x^u e_u + x^v e_v + x^w e_w$ (and $x^u = aX^u, x^v = bX^v, x^w = cX^w$).

There is a lot of information contained in this (dense) section (we imagine some –but not necessarily all– of it may be 'review'). One can also see sections 1-3 of Feynman volume II or ch. 1 of Griffiths. Essentially any (multi-variable) calculus book will also go into more detail on these concepts. See for example Shey's *Div, grad, curl and all that* or Marsden and Weinstein (vol. III, available here) or these online notes. For some references with differential forms, one can see Spivak's *Calculus on manifolds*, Fleming's *Functions of several variables*, part II of Arnold's *Mathematical methods of classical mechanics*, or Jänich's *Vector analysis*.

Exercises:

1. Let $\gamma(t), t \in [t_o, t_1]$ parametrize a curve. For $f : [\tau_o, \tau_1] \to [t_o, t_1], \tau \mapsto t = f(\tau)$ with $f'(\tau) > 0$ and parametrization $\tau \mapsto \gamma(f(\tau)) =: \Gamma(\tau)$ of the curve, show that:
$$\int_{\tau_o}^{\tau_1} |\frac{d\Gamma}{d\tau}| \ d\tau = \int_{t_o}^{t_1} |\frac{d\gamma}{dt}| \ dt.$$

2. (a) For a parametrized curve given in cylindrical coordinates, $r(t), \theta(t), z(t)$, show the norm of its velocity squared is:
$$\dot{r}^2 + r^2 \ \dot{\theta}^2 + \dot{z}^2.$$
(b) For a parametrized curve given in spherical coordinates, $\rho(t), \varphi(t), \theta(t)$, show the norm of its velocity squared is:
$$\dot{\rho}^2 + \rho^2 \ \dot{\varphi}^2 + \rho^2 \sin^2 \varphi \ \dot{\theta}^2.$$

3. From the geometric definition, $u \cdot v := |u||v| \cos \theta$, of dot product show that the dot product is bilinear and symmetric.

4. From the geometric definition, $|u \times v| = |u||v| \sin \theta$ and directed along the normal to u, v by right hand rule, of cross product show that cross product is bilinear and anti-symmetric.

5. Show that $u \times (v \times w) = (u \cdot w) \ v - (u \cdot v) \ w$.

6. Let $f : V \to W$ be a differentiable function between vector spaces. For $v_1, ..., v_n$ and $w_1, ..., w_m$ bases of V and W we have
$$\mathbb{R}^n \to \mathbb{R}^m, x = (x_1, ..., x_n) \mapsto (f_1(x), ..., f_m(x))$$
by $v = x_1 v_1 + ... + x_n v_n \mapsto f(v) = f_1(x) w_1 + ... + f_m(x) w_m$. For $p \in V$ show that, in these bases, the linear map $d_p f : V \to W$ is represented by the $m \times n$ matrix:
$$\begin{pmatrix} \partial_{x_1} f_1 & \partial_{x_2} f_1 & ... & \partial_{x_n} f_1 \\ \partial_{x_1} f_2 & \partial_{x_2} f_2 & ... & \partial_{x_n} f_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} f_m & \partial_{x_2} f_m & ... & \partial_{x_n} f_m \end{pmatrix} |_p.$$

7. Determine the surface area, by explicitly evaluating the integrals, of:
(a) A circular cylinder of radius r and height h,
(b) A cone of height h and radius r,
(c) A spherical lune on a sphere of radius r with opening angle θ.

8. Consider a surface given as a graph, $z = f(x, y)$ for $(x, y) \in D \subset \mathbb{R}^2$. Show that its surface area is given by:
$$\int_D \sqrt{1 + f_x^2 + f_y^2} \ dxdy.$$

9. Determine the flux of the vector field $X(x, y, z) = (0, 0, z)$ through
(a) An upper hemisphere of a sphere of radius r: $\{x^2 + y^2 + z^2 = r^2, \quad z > 0\}$.
(b) A vertical cylinder of height h and radius r: $\{x^2 + y^2 = r^2, 0 \leq z \leq h\}$.

10. Let $A_t : V \to V, t \in \mathbb{R}$ be a (smooth[1]) curve of invertible linear transformations of the vector space V. Set $\dot{A}_t = \frac{d}{dt} A_t$, ie $\dot{A}_t : V \to V$ is the linear transformation $v \mapsto \frac{d}{d\varepsilon}|_{\varepsilon=0} A_{t+\varepsilon} v$. Show that
$$\frac{d}{d\varepsilon}|_{\varepsilon=0} \det A_{t+\varepsilon} = \det(A_t) tr(A_t^{-1} \dot{A}_t) = \det(A_t) tr(\dot{A}_t A_t^{-1}).$$

[1] Meaning that when expressed in a basis as a matrix, the entries of the matrices are smooth functions of t.

11. From the limit definition, $\nabla \cdot X(p) = \lim_{vol(\Omega)\to 0} \frac{\int_{\partial\Omega} X\cdot\nu\ dA}{vol(\Omega)}$ with $p \in \Omega$, show that:

    (a) In Cartesian coordinates, $X = Pe_1 + Qe_2 + Re_3$, we have : $\nabla \cdot X = P_x + Q_y + R_z$.

    (b) In cylindrical coordinates, $X = Pp_r + Qp_\theta + Rp_z$, we have: $\nabla \cdot X = \frac{1}{r}\partial_r(rP) + \partial_\theta(Q) + \partial_z(R)$

12. For f a function and X a vector field, show that:

    (a) $\nabla \times (\nabla f) = 0$,

    (b) $\nabla \cdot (\nabla \times X) = 0$.

13. For f, g functions, X a vector field and Ω a (compact) region with smooth boundary $\partial\Omega$, verify:

    (a) $\nabla \cdot (fX) = \nabla f \cdot X + f\nabla \cdot X$, deduce that $\int_\Omega \nabla f \cdot X + f\nabla \cdot X\ dV = \int_{\partial\Omega} fX \cdot d\vec{S}$.

    (b) $\int_\Omega f\Delta g - g\Delta f\ dV = \int_{\partial\Omega}(f\nabla g - g\nabla f)\cdot d\vec{S}$ (called *Green's identity*).

14. Consider simple curve ∂D in the plane bounding the planar region D. Let X be a vector field on the plane. Show that:

$$\int_D \nabla \cdot X\ dA = \int_{\partial D} X \cdot n\ ds$$

    where n is the outward unit normal to ∂D (called *Green's theorem*).

15. For a vector field X, set $\Delta X = \nabla(\nabla \cdot X) - \nabla \times (\nabla \times X)$ for the *vector Laplacian* of X. In Cartesian coordinates, with $X = (P, Q, R)$ show that $\Delta X = (\Delta P, \Delta Q, \Delta R)$.

16. The *solid angle*, Ω, of an oriented surface Σ measured from a point p_o is the oriented area of its projection onto the unit sphere centered at p_o. One has:

$$\Omega = \int_\Sigma \frac{\cos\alpha}{r^2}\ dA$$

    where r is the distance from p_o to $p \in \Sigma$ and α is the angle between the unit normal to the surface and the ray from p_o to $p \in \Sigma$. Determine the solid angle of a sphere of radius r (oriented with outward unit normal) measured from

    (a) a point inside the sphere,

    (b) a point outside the sphere.

§2 reference frames and fields

In this section, we will review the basis of classical mechanics. During the course, the fundamental ideas of electricity and magnetism will be developed in this framework. By the end of the course, we will see how our study of electricity and magnetism motivates the modification of classical mechanics to special relativity.

Properties of physical objects are determined by making measurements. The fundamental *position and time* measurements at which a property of an object is measured are *relative concepts*, meaning they are described in reference to other objects.
Such descriptions require the choice of a coordinate system or *reference frame*, given by:

- at each instant of time, an origin in space,
- at each instant of time, coordinate axes based at the chosen origin,
- choice of an origin of time,
- choice of length and time units (eg meters, seconds).

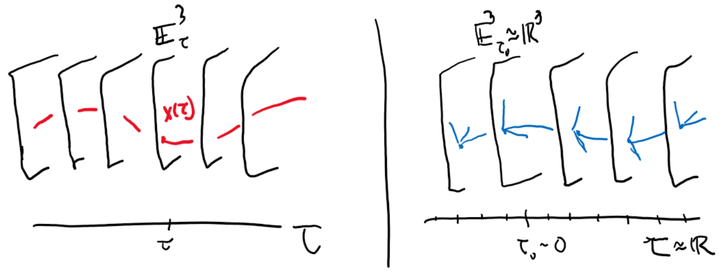

Figure 7. The classical *space-time* consists of the collection of all positions, $\mathbb{E}^3_\tau$, at a given instant $\tau \in \mathcal{T}$ where $\mathcal{T}$ is *the* time-line. We will write $\mathcal{M} := \bigsqcup_{\tau \in \mathcal{T}} \mathbb{E}^3_\tau$ for spacetime. The choices for a reference frame yield identifications $\mathcal{T} \cong \mathbb{R}, \mathbb{E}^3_\tau \cong \mathbb{R}^3$ and $\mathcal{M} \cong \mathbb{R}^3 \times \mathbb{R}$.

Once a reference frame is chosen, space-time is identified with $\mathbb{R}^3 \times \mathbb{R} \ni (x, t)$. *Any* other choice of reference frame is then related to this one by a transformation of the form:

$$(A(t)x + b(t), ct + d) = (\tilde{x}, \tilde{t}) \in \mathbb{R}^3 \times \mathbb{R}$$

where $A(t) \in \mathrm{O}_3, b(t) \in \mathbb{R}^3, c \in \mathbb{R}_{>0}, d \in \mathbb{R}$.
The main objective in physics is to describe and predict the properties of an object over the course of time. In particular, the *motion* of an object consists in describing its position in some reference frame as a function of time. Analysis of an objects' motion is based first on the fundamental idea of *free* or *natural motion*. This is the motion the object *would* take in the absence of *any* external influences on it. Any deviations in an objects motion from its free motion is said to be caused by *forces* having been applied to the object.
As position and time measurements are relative concepts, so too is motion: the observed motion of an object depends on ones chosen reference frame. Defining free motions thus goes hand in hand with defining particular reference frames in which free motions have a standard description. A reference frame is called *inertial* when the points defining its origin and axes are free from any external influences. The dynamic or predictive portion of classical mechanics is summarized in the following 'axioms' or 'laws of motion', essentially defining free motions and forces.

Newton's 1st law: An inertial reference frame exists. In an inertial reference frame, the free motion of a particle is of uniform velocity along a straight line.

Newton's 2nd law: The change in velocity (acceleration) of a particle in an inertial reference frame subject to a force is directly proportional to the force. This constant of proportionality is the *mass* of the object.

Newton's 3rd law: Whenever object A exerts a force $\vec{f}$ on object B then object B exerts the force $-\vec{f}$ on object A.

When the forces influencing a system are known, classical mechanics reduces describing the resulting motion to the study of second order differential equations. Formulas for the forces due to certain influences are determined by experiments or observations. For example some of the most impressive predictions of classical mechanics have come in astronomy from the:

Universal law of gravitation: the force due to gravity between two particles is proportional to the product of their masses over the distance squared between them.

This 'inverse square' law of gravitation is motivated by the observations of the motions of the planets, summarized at that time most precisely in 'Kepler's laws'. The proportionality constant is labelled G and called the *universal gravitational constant.* In standard units, $G \approx 6.67 \times 10^{-11}\ m^3/(kg \cdot s^2)$.

Physical objects have more properties than their positions at given times. In general any quality which may be measured at each point of space and instant of time is called a *field.* For example, one may measure temperatures at various times and locations, $T : \mathcal{M} \to \mathbb{R}$, given in a reference frame by a function, $\mathbb{R}^3 \times \mathbb{R} \to \mathbb{R},\ \ (x,t) \mapsto T(x,t)$ representing this temperature (scalar) field. Similarly, a distribution of matter determines its (mass) density field, $\rho : \mathcal{M} \to \mathbb{R}$.

To finish we will describe the potential and field theories for classical gravitation, noting that we will apply essentially exactly the same considerations in our study of electrostatics. Also, let us remark that unless otherwise mentioned if we are using a reference system we will always assume it is inertial.

Consider an isolated particle of mass m, currently located in our inertial reference frame at the fixed position $x \in \mathbb{R}^3$. *If* we were to place a particle ('test particle') of mass m' at the position x' then this particle would experience a gravitational force:

$$\vec{f}(x') = Gmm' \frac{x - x'}{|x - x'|^3}.$$

This force depends on the mass m' of the test particle. A measurable property of space at each position x' produced by the presence of the point mass m at x and independent of the test particle is thus:

$$\vec{G}(x') := Gm \frac{x - x'}{|x - x'|^3}$$

called the *gravitational field* produced by the presence of the point mass m currently at the position x. It is a field in the above sense, since to measure $\vec{G}$ at a given time and location we measure the acceleration –in an inertial frame– that a test particle placed at this location receives due to the presence of the particle of mass m at its location at this time.

The gravitational potential field is closely related to work, which in turn is closely related to energy. Consider a field of forces in space at a given instant, $\vec{f}(x)$ is the force at $x \in \mathbb{R}^3$. The *work* done by the forces when one moves a point mass along an oriented curve $\mathcal{C}$ is the line integral:

$$W := \int_{\mathcal{C}} \vec{f} \cdot T\ ds = \int_{\mathcal{C}} \vec{f} \cdot d\vec{s}.$$

Now, the work –done by the gravitational forces produced by the particle of mass m located at x– to move a test particle of mass m' from some fixed reference position, x_o, to the position x' is:

$$W = \int_{\mathcal{C}_{x_o,x'}} \vec{f} \cdot d\vec{s} = -Gmm' \int_0^1 \frac{\gamma(\tau) \cdot \gamma'(\tau)}{|\gamma(\tau)|^3}\ d\tau = Gmm' \left(\frac{1}{|x - x'|} - \frac{1}{|x - x_o|} \right)$$

where $\mathcal{C}_{x_o,x'}$ is a curve from x_o to x' parametrized as $[0,1] \ni \tau \mapsto \gamma(\tau) + x \in \mathbb{R}^3$. In particular the work is independent of the path from x_o to x', and we have: $\nabla W = \vec{f}$. Moreover, the fixed reference position x_o has no effect on ∇W. It is customary to take $|x_o| \to \infty$ and

$$U := \frac{Gmm'}{|x - x'|}$$

as the *force function*, with $\nabla U = \vec{f}$ giving the force field experienced by a test particle of mass m' due to the particle of mass m. Moreover, if the gravitational force field has done work U to bring the mass m' to x' then in bringing it there 'we' or whatever is moving it has invested work $-U$ to get to this configuration. This 'stored work' in the position x' is called the *potential energy*:

$$V := -U = -\frac{Gmm'}{|x - x'|}.$$

The above functions, W, U, V, all depend on the mass m' of the test particle, we obtain the test particle independent (scalar) field over space:

$$\varphi(x') := -\frac{Gm}{|x - x'|}$$

called the *gravitational potential* produced by the particle of mass m at position x, satisfying:

$$\nabla\varphi = -\vec{G}.$$

Note that the equation above does not determine φ uniquely, rather upto addition of an arbitrary constant. However φ is determined uniquely by $\nabla\varphi = -\vec{G}$ and the *additional* condition that $\varphi(x') \to 0$ as $|x'| \to \infty$. Moreover, note that φ and $\vec{G}$ are undefined at x.

These considerations of the fields produced by a single point mass yield more general expressions using the:

Principle of superposition: The net result of applying two forces $\vec{f}_1, \vec{f}_2$ to a particle is that of applying their vector sum: $\vec{f}_1 + \vec{f}_2$.

By induction, this principle determines the resultant net force when any number of forces act on a given particle, and consequently the gravitational fields and gravitational potentials of any finite system of point masses are determined by summation. In the case of a (compact) continuum[1], Ω, with mass density ρ,

$$(*) \quad \vec{G}(x') := G \int_{x\in\Omega} \frac{x - x'}{|x - x'|^3}\ \rho(x)\ dV, \quad \varphi(x') := -G \int_{x\in\Omega} \frac{\rho(x)\ dV}{|x - x'|}.$$

The gravitational field, gravitational potential and mass density are related by the:

Classical gravitational field equations: The gravitational field and gravitational potential, $\vec{G}, \varphi$, defined by $(*)$ produced by a mass distribution with density ρ satisfy:[2]

$$\nabla\varphi = -\vec{G}, \quad \Delta\varphi = 4\pi G\rho.$$

proof: First, we consider the fields produced by a point mass m, located at position x in an inertial frame. Then for a sphere of radius r centered at the point mass, we have $\int_{S^2_r} \vec{G} \cdot \nu\ dA = -4\pi Gm$. Note that φ is only a function of the radial distance to the point mass, and so as well $\Delta\varphi$. By the divergence theorem, $\int_{\varepsilon\le\rho\le r} \Delta\varphi\ dV = \int_{S^2_\varepsilon} \vec{G} \cdot d\vec{S} - \int_{S^2_r} \vec{G} \cdot d\vec{S} = 0$ for any $0 < \varepsilon < r$, and so $\Delta\varphi \equiv 0$ away from x. Now for any smooth function f vanishing outside a compact set, we have by Green's identity that:

$$\int_{\varepsilon\le\rho\le r} \varphi\Delta f\ dV = \int_{S^2_r - S^2_\varepsilon} \varphi\partial_\nu f - f\partial_\nu\varphi\ dA = \int_{S^2_\varepsilon} f\partial_\nu\varphi - \varphi\partial_\nu f\ dA$$

[1] Likewise, for a surface density σ or linear density λ, one has eg $\varphi(x') := -G \int_{x\in\Sigma} \frac{\sigma(x)\ dA}{|x-x'|}$ or $\varphi(x') := -G \int_{x\in\mathcal{C}} \frac{\lambda(x)\ ds}{|x-x'|}$.

[2] This second equation is called the *Poisson equation*.

when $r >> 0$ is sufficiently large so that $f \equiv 0$ around S^2_r. Sending $\varepsilon \to 0$ and using our explicit expressions for φ due to a point mass, we obtain:

$$\int_{\mathbb{R}^3} \varphi \Delta f \; dV = 4\pi G m \; f(x).$$

In this way, $\Delta\varphi$, may be thought of as a *generalized function* [1] satisfying $\Delta\varphi = 4\pi G m \; \delta_x$ for δ_x the *dirac-delta function* at x. On the other hand, the mass density ρ of a point mass is $\rho = m\delta_x$, so that for a point mass we have $\Delta\varphi = 4\pi G\rho$. Now, we obtain the general result by superposition and differentiation under the integral sign: $\Delta\varphi(x') = 4\pi G \int \rho(x)\delta(x - x') \; dV = 4\pi G\rho(x')$. $\square$

Finally, we remark that given the gravitational fields, the *dynamics* is determined through $\vec{f} = m'\vec{G}$ being the gravitational force on a particle of mass m'. For two bodies for instance, one has the ode's:

$$m_1\ddot{x}_1 = m_1\vec{G}_2(x_1), \quad m_2\ddot{x}_2 = m_2\vec{G}_1(x_2)$$

where $\vec{G}_j$ are the gravitational fields produced by the point mass m_j at x_j. Since its introduction, this aspect of classical gravity –*instantaneous action at a distance*– has met with objections (including from Newton himself). Namely, the appearance of a point mass *instantly* exerts a gravitational influence over *all* points of space, which is not intuitive. These objections were soon forgotten or at least overlooked, as one may still obtain impressive and highly accurate predictions with the classical theory.

[1] See R. Strichartz, *A guide to distribution theory and Fourier transforms.* World Scientific Publishing Company, 2003.

Exercises:

1. The orthogonal group (or rotation group) –denoted by O_3– are linear maps $A : \mathbb{R}^3 \to \mathbb{R}^3$ which preserve the dot product: $A\vec{u} \cdot A\vec{v} = \vec{u} \cdot \vec{v}, \forall \vec{u}, \vec{v} \in \mathbb{R}^3$.

   (a) Show that $A \in O_3 \iff A^T A = I$. Deduce that $\det A = \pm 1$.

   Orientation preserving elements of O_3 (that is those with determinant 1) are denoted SO_3.

   (b) Let $\mathbb{R} \ni t \mapsto A(t) \in SO_3$ be a smooth curve of rotations with $A(0) = I$. Show that $\Omega := \frac{d}{dt}|_{t=0} A(t)$ is a skew-symmetric linear map ($\Omega^T = -\Omega$).

   The set of skew-symmetric linear maps are denoted by $\mathfrak{so}_3$.

   (c) For $\vec{\omega} \in \mathbb{R}^3$ consider $\Omega_{\vec{\omega}} : \mathbb{R}^3 \to \mathbb{R}^3, \vec{v} \mapsto \vec{\omega} \times \vec{v}$. Show that $\Omega_{\vec{\omega}} \in \mathfrak{so}_3$ and $\mathbb{R}^3 \to \mathfrak{so}_3, \vec{\omega} \mapsto \Omega_{\vec{\omega}}$ is a vector space isomorphism.

   We call $\vec{\omega}$ an *infinitesimal rotation axis* and $\Omega_{\vec{\omega}}$ an infinitesimal rotation.

2. Let the length and time units be fixed. For $(x, t) \in \mathbb{R}^3 \times \mathbb{R}$ coordinates from an inertial reference frame, show that the coordinates $(\tilde{x}, \tilde{t}) \in \mathbb{R}^3 \times \mathbb{R}$ are also those of an inertial frame iff one has:[1]

   $$(Ax + b + tv, t + d) = (\tilde{x}, \tilde{t})$$

   for some fixed $A \in O_3, b, v \in \mathbb{R}^3$ and $d \in \mathbb{R}$.

3. A fundamental principle in classical mechanics is the *Galilean principle of relativity*. Stating that '*the equations of motion for a closed system*[2] *are the same in all inertial reference frames*'. Consider a closed system consisting of two particles with equations of motion $m_j \ddot{x}_j = f_j(x_1, x_2, \dot{x}_1, \dot{x}_2, t)$ in some inertial reference frame $(x, t) \in \mathbb{R}^3 \times \mathbb{R}$. [3]

   (a) Show that the forces do not depend on time, ie we may write $m_j \ddot{x}_j = f_j(x_1, x_2, \dot{x}_1, \dot{x}_2)$.

   (b) Show that the forces depend only on the mutual distances and velocities, ie we may write $m_j \ddot{x}_j = f_j(x_1 - x_2, \dot{x}_1 - \dot{x}_2)$.

   (c) Show that the forces are 'rotation equivariant', ie we have $f_j(A(x_1 - x_2), A(\dot{x}_1 - \dot{x}_2)) = A f_j(x_1 - x_2, \dot{x}_1 - \dot{x}_2)$ for any $A \in O_3$.

4. Let $A(t) \in SO_3$ be rotation by angle ωt around the $\hat{k}$ (z)-axis. Show that

   $$\frac{d}{dt}|_{t=s} A(t)A(s)^{-1} = \frac{d}{dt}|_{t=s} A(s)^{-1}A(t) = \Omega_{\vec{\omega}}$$

   where $\vec{\omega} = \omega \hat{k}$.

5. For $A(t)$ as in the previous problem, let $A(t)\vec{y} = \vec{x}$ be a uniformly rotating coordinate system. Show that:

   $$\ddot{\vec{y}} = -2\vec{\omega} \times \dot{\vec{y}} - \vec{\omega} \times (\vec{\omega} \times \vec{y}) + A^{-1}\ddot{\vec{x}}.$$

6. Calculate the gravitational field, $\vec{G}$, produced by an (infinite) plane with constant surface density σ.

7. Calculate the gravitational field, $\vec{G}$, produced by a homogeneous ($\sigma \equiv cst.$) spherical shell (a sphere) of radius r.

8. Calculate the gravitational field, $\vec{G}$, produced by a homogeneous ($\rho \equiv cst.$) solid ball of radius r.

[1] Such transformations of $\mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3 \times \mathbb{R}$ form what is called the *Galilean group*.

[2] A *closed system* is one for which the particles of the system are subject to no external forces, ie only due to forces from the other particles of the system.

[3] More precisely, the Galilean principle asserts that for any other inertial reference frame $(\tilde{x}, \tilde{t})$ the equations of motion have the same form of $m_j \frac{d^2\tilde{x}_j}{dt^2} = f_j(\tilde{x}_1, \tilde{x}_2, \frac{d\tilde{x}_1}{dt}, \frac{d\tilde{x}_2}{dt}, \tilde{t})$ as those in the (x, t) inertial frame.

9. Consider the gravitational field $\vec{G}$ produced by a point of mass m. Show that for any closed surface, Σ, enclosing the point mass and oriented with outward unit normal, one has:

$$\int_\Sigma \vec{G}\cdot d\vec{S} = -4\pi G m.$$

10. Let a parametrized plane curve, $\gamma(t)$, be given in polar coordinates by $r(t), \theta(t)$. Show that

$$\frac{1}{2}\int_{t_1}^{t_2} r(t)^2\dot{\theta}(t)\ dt$$

is the sectorial area between $\gamma(t_1)$ and $\gamma(t_2)$. Deduce that the sectorial area as a function of t increases at a constant rate if and only if $r^2\dot{\theta} = cst$.

11. In this exercise we outline how Newton's inverse square law may be deduced from Kepler's laws[1]. Suppose a force field $\vec{f}$ in space has the property that all of its trajectories (solutions to $\ddot{q} = \vec{f}(q)$) satisfy Kepler's laws with the 'sun' at the origin.

(a) Consider a trajectory $q(t)$ (a conic section by assumption). Show that $q(t)\times\dot{q}(t) = cst$. By differentiating, show that $q(t)\times\vec{f}(q(t)) = 0$ and deduce that $\vec{f}(q(t))$ is proportional to $q(t)$. Since the trajectory $q(t)$ was arbitrary, we may write:

$$\vec{f}(q) = -\frac{f(q)}{|q|}q$$

for some $f:\mathbb{R}^3\to\mathbb{R}$.

(b) Consider polar coordinates, r, θ, on the plane containing a trajectory $q(t)$. Show that:

$$\ddot{r} = \frac{C^2}{r^3} - f(q), \quad r^2\dot{\theta} = C$$

for some constant C.

(c) Using that a conic section in a plane with focus at the origin may be given in polar coordinates by

$$r = \frac{p}{1+e\cos\theta}$$

for constants e, p, show that along the trajectory $q(t)$ we have:[2]

$$\vec{f}(q) = -\frac{C^2}{p}\frac{q}{|q|^3}$$

with C the constant from part (b).

(d) For an elliptic trajectory, apply Kepler's 3rd law to deduce that $\vec{f}(q) = -k\frac{q}{|q|^3}$ with k a constant.

[1] These state: 1) the orbit of an object around the sun traces out a conic section having one focus at the sun. 2) As the object moves along its orbit the sectorial area as a function of time increases at a constant rate. 3) For an elliptic orbit, the ratio of the period squared to the major axis cubed is constant. A lovely article on Kepler's laws is: A. Givental, *Kepler's laws and conic sections.* Arnold Mathematical Journal 2.1 (2016): 139-148.

[2] Suggestion: rewrite the equations of motion from part (b) in terms of $\rho := 1/r$ and reparametrized wrt θ using $d\theta = \frac{C}{r^2}dt$.

§3 electrostatics

We will begin with the following situation: to examine the effects due to a collection of electrically charged objects located at *fixed* positions in a given inertial frame. Similarly to mass, every object may be assigned a total charge (unlike mass, charge is *signed*). The basis[1] for this electrostatics is:

Coulomb's law: The force between two *static* particles with charges q_1, q_2 is proportional to the product of the charges over the distance squared between them,

$$\vec{f}_{12} = -\vec{f}_{21} = \frac{q_1 q_2}{4\pi\varepsilon_o}\frac{x_2 - x_1}{|x_2 - x_1|^3}.$$

Here $\vec{f}_{ij}$ is the force on q_j due to q_i when the charges are located at the positions x_j and the constant of proportionality is given by $\frac{1}{4\pi\varepsilon_o} = 9 \times 10^9\ N \cdot m^2/C^2$ in standard units.

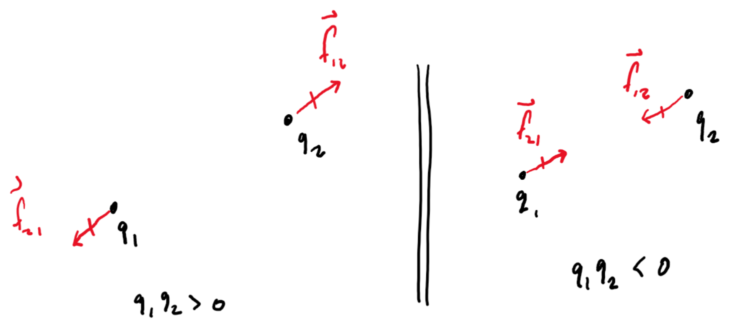

Figure 8. Coulomb's law gives the form for the force between two fixed charged particles.

We will be concerned with determining the effects a given static charge distribution produces, that is to say the *electric field* the charges generate. The electric field due to a configuration of charges is a vector field, $\vec{E}$, whose value at a given position, x', is measured by placing a test charge q' at the position and taking:

$$\vec{E}(x') := \vec{f}/q'$$

where $\vec{f}$ is the force on the test charge due to the charge distribution. That the electric field is well-defined (independent of the test charge q') follows from Coulomb's law and the principle of superposition.

More precisely, one may first consider the electric field generated by a point charge q fixed at the position $x \in \mathbb{R}^3$, which by Coulomb's law is:

$$\vec{E}(x') = \frac{q}{4\pi\varepsilon_o}\frac{x' - x}{|x' - x|^3}.$$

By superposition, a charge distribution with charge density ρ generates the electric field:

$$\vec{E}(x') = \frac{1}{4\pi\varepsilon_o}\int_{x\in\mathbb{R}^3}\frac{\rho(x)(x' - x)}{|x' - x|^3}\ dV,$$

so that electrostatics at its core is reduced to the evaluation of (typically complicated) integrals.

There are however interesting techniques that are relevant for the analysis of electrostatic problems. We will get a lot of mileage out of the superposition principle by observing some properties of the electric field, $\vec{E}$, generated by a point charge q. First note that $\nabla \cdot \vec{E} = 0$ away from the point charge so that for any

[1]This law was derived from experiments by various scientists and published in its final form by Coulomb (1785) by use of a torsion balance (the same device would be used by Cavendish (1798) to determine the first reliable measurement of the gravitational constant G). See this lecture of Feynman (§7-6) for a description of this device and remarkable measurement.

region Ω containing the point charge in its interior one finds: $\int_{\partial\Omega} \vec{E} \cdot d\vec{S} = \frac{q}{\varepsilon_o}$. Now, by superposition,

Gauss' law: for a static configuration of charges with charge density ρ, generating the electric field $\vec{E}$,

$$\int_{\partial\Omega} \vec{E} \cdot d\vec{S} = \frac{1}{\varepsilon_o} \int_{\Omega} \rho \, dV = \frac{Q_{int}}{\varepsilon_o}.$$

Where Q_{int} is the total charge contained in the region Ω.

Next observe that the electric field generated by a point charge is a gradient field: for any curve C from x_o to x_1 not passing through the point charge, one computes: $\int_C \vec{E} \cdot d\vec{s} = \frac{q}{4\pi\varepsilon_o}\left(\frac{1}{r_o} - \frac{1}{r_1}\right)$ where r_j are the distances from the endpoints x_j of the curve to the point charge. It is customary to take the 'base-point' $|x_o| \to \infty$ and: $\varphi := \frac{q}{4\pi\varepsilon_o r}$ with r the distance to the point charge, as the *electric potential* which has: $-\nabla\varphi = \vec{E}$. By superposition, we obtain:

Electric potential: a *bounded* static configuration of charges with charge density ρ, generates a gradient electric field $\vec{E}$,

$$\vec{E} = -\nabla\varphi.$$

Where $\varphi(x') := \int_{x\in\mathbb{R}^3} \frac{\rho(x)}{4\pi\varepsilon_o|x'-x|} \, dV$ is called the *electric potential*.

Note that when the charge distribution is not bounded there is typically still a potential function, well defined upto a constant by taking some basepoint and considering the work, $\int_C \vec{E} \cdot d\vec{s}$, done by the forces due to the charge distribution to move a unit charge along a curve C beginning at the basepoint. Also we comment that our formula $\vec{E} = -\nabla\varphi$ is meant to hold only where the electric field is actually defined (eg for a point charge it has no sense at the location of the point charge).

In sufficiently symmetric charge configurations, the Gauss law coupled with a symmetry argument often yields an efficient way to determine the electric fields. As well it may be simpler in some cases to determine the electric potential –consisting of one integral– rather than the electric field directly (consisting of three integrals, one for each component).

Before considering some more intricate electrostatic situations, we summarize our results so far in the following table:

Maxwell's equations in electrostatics	
Integral form	Differential form
$\int_{\partial\Omega} \vec{E} \cdot d\vec{S} = \frac{1}{\varepsilon_o} \int_{\Omega} \rho \, dV$	$\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_o}$
$\oint_C \vec{E} \cdot d\vec{s} = 0$	$\nabla \times \vec{E} = 0$
Electrostatic potential theory	
$\vec{E} = -\nabla\varphi$	$\Delta\varphi = -\frac{\rho}{\varepsilon_o}$

Table 1. Electrostatics may be summarized as a special case of Maxwell's equations. Note that the vector $\vec{J}$ in Maxwell's equations represents a 'current density' and is zero in electrostatics (charges are not moving), $\nabla \cdot \vec{B} = 0, \nabla \times \vec{B} = 0$.

The electric potential also presents the following rephrasing of electrostatics. One may consider a given charge distribution ρ and seek a solution to the Poisson equation, $\Delta\varphi = -\frac{\rho}{\varepsilon_o}$, which 'vanishes at infnity'. Then we have $\vec{E} = -\nabla\varphi$. By our discussion above, the solution to this Poisson equation –upto a constant– is given by $\varphi(x') = \int \frac{\rho(x) \, dV}{|x'-x|}$. Electrostatics is thus reduced to understanding solutions of the Poisson equation. Note that wherever there is no charge present, $\rho(x') = 0$, the electric potential is *harmonic*: $\Delta\varphi(x') = 0$.

Conductors: We now consider electrostatic situations involving conducting materials: materials in which charges may move freely. As opposed to the situation where the charge density is given and we seek the electric field, here one usually does not know the entire charge distribution, but rather some properties of the electric field are known.

To explain these properties, we first consider a conducting material, a region Ω, with zero net charge in a static state. This does *not* mean that there are no charges in Ω, rather the charges constituting the conducting material are in 'perfect balance'. At the microscopic level, the atoms constituting the material contain the same number of electrons and protons or, as a continuum one may think of a zero charge density $\rho_+ + \rho_-$ over Ω with $\rho_\pm$ the densities of positive and negative charges having $\rho_+(x) = -\rho_-(x),\ x \in \Omega$.

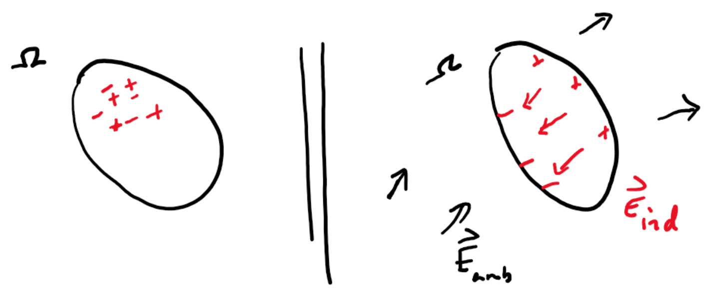

Figure 9. An isolated conductor, Ω, with no net charge will consist of an equal balance of positive and negative charges throughout the conductor. When in the presence of an ambient electric field, $\vec{E}_{amb}$, the charges inside Ω will rearrange and produce their own electric field $\vec{E}_{ind}$ giving a net electric field $\vec{E} = \vec{E}_{amb} + \vec{E}_{ind}$. Unless this rearrangement has led to conditions with $\vec{E} = 0$ inside Ω and $\vec{E} \perp \partial\Omega$ the charges inside the conductor will continue to rearrange themselves –ie not be in a static situation.

Now, if the conducting material is in the presence of an electric field then, as the charges constituting the conductor are free to move, they will in general no longer remain in their state of perfect balance but move to a new equilibrium as a result of the ambient electric field. That is at a local level the charges in the conducting material will be rearranged, no longer having a perfect balance of positive and negative charges, and producing its own electric field. Now unless this rearrangement of charges in Ω has led to a situation in which the total electric field is zero inside Ω and perpendicular to $\partial\Omega$, the charges of Ω would still move. So we arrive at the following conditions for a static situation involving a conducting material Ω:

- the electric field inside Ω is zero,
- the electric field is perpendicular to the boundary of Ω.

Let us remark that these properties yield how conductors are used for 'electric shielding', namely inside an empty conductor the electric field due to outside charges does not penetrate. From our reasoning above, it is not clear that a conductor placed in an ambient electric field will in fact redistribute over enough time to produce a net electric field with the above conditions. However, it is possible to give some mathematical justification for the experimental fact that the above two conditions are essentially realized after a very short rearrangement 'transient' time. Regardless, we see that placing a conductor into an electric field with the above equilibrium conditions will remain a static situation. To make use of this information, it is useful to reformulate in terms of electric potential: let φ be an electric potential for the electric field generated by a configuration of (fixed) charges with density ρ and a conductor Ω (with $\rho|_\Omega = 0$). Then:

- $\varphi|_\Omega = cst.$

Moreover, the electric field, $\vec{E} = -\nabla\varphi$, is the result of the charge distribution ρ and a charge distribution σ over $\partial\Omega$ (outward unit normal ν) with:

$$\partial_\nu \varphi = \nabla\varphi \cdot \nu = -\frac{\sigma}{\varepsilon_o}.$$

A related question with conductors is to consider a solitary conductor, Ω, having some net charge, Q. If the charges constituting the conductor are in a static situation, how will these excess charges be distributed

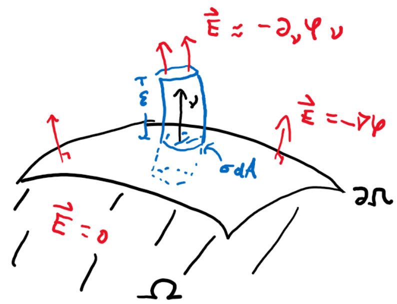

Figure 10. In a static situation, the electric field is zero inside the conductor and normal to the boundary. Applying Gauss' law to a small tube of height ε around $\partial\Omega$ one obtains when $\varepsilon \to 0$ that $\int_D \frac{\sigma}{\varepsilon_o}\, dA = \int_D \vec{E} \cdot \nu\, dA$ for any $D \subset \partial\Omega$ so that $\frac{\sigma}{\varepsilon_o}\nu = \vec{E}$ over $\partial\Omega$.

over the conductor? As before, the assumption that we are static leads us to require that the distribution of charges in Ω leads to a situation producing a vanishing electric field inside Ω and an exterior electric field that along the boundary is perpendicular to $\partial\Omega$. Thus this situation is encapsulated as well by requiring $\varphi|_\Omega = cst.$ and the charges are then distributed over the boundary with density $\sigma = -\varepsilon_o \partial_\nu \varphi$.

We now show the answers to these questions are unique (also existence can be proved, but is more involved). Namely there is exactly one charge distribution σ over the surface of a conductor leading to a static situation. The main tool for uniqueness is the maximum principle[1] for harmonic functions, giving:

Dirichlet boundary conditions: Given a compact region Ω and $f : \partial\Omega \to \mathbb{R}$, there is exactly one function $u : \Omega \to \mathbb{R}$ satisfying:

$$\Delta u = 0, \quad u|_{\partial\Omega} = f.$$

proof: We will not prove existence [2]. For uniqueness, suppose u_1, u_2 are solutions and consider $v := u_1 - u_2$. Then v is harmonic in Ω and constant (zero) on $\partial\Omega$. By the maximum principle, $v = 0$ is constant. $\square$

Static conductors: Given a (compact) conductor Ω, ambient charge distribution ρ, and constant V, there is exactly one electric potential, φ, satisfying:

$$\Delta\varphi = -\frac{\rho}{\varepsilon_o} \text{ on } \Omega^c,$$

$$\varphi|_\Omega = V = cst., \quad \varphi(x) = O\left(\frac{1}{|x|}\right),\ |x| \to \infty.$$

proof: Again, we only prove uniqueness. Suppose φ_1, φ_2 are two solutions and set $u := \varphi_1 - \varphi_2$. The problem may be divided into two parts: an interior and exterior problem. Inside Ω, we have $\Delta u = 0$ and $u|_{\partial\Omega} = 0$ so that $\varphi_1 = \varphi_2$ inside Ω. Outside Ω, we have $\Delta u = 0$ and $u|_{\partial\Omega} = 0$ however Ω^c is not compact so we may not apply the maximum principle as before. Instead let $x_o \in \Omega^o$ be an interior point of Ω. Take $x_o = 0$ as our origin and consider the spherical inversion:

$$x \mapsto \frac{x}{|x|^2} = x'.$$

This map (also called *Kelvin transform*) sends Ω^c to a bounded region, Ω' with x_o the 'image of infinity'. In fact, the function:

$$u'(x') := \frac{u(x'/|x'|^2)}{|x'|}$$

[1]This states that a harmonic function, $\Delta u = 0$, defined on a compact set Ω attains its maximal and minimal values on $\partial\Omega$.

[2]See Arnold's lectures on pde's (in particular lectures 7,8, 12) for some of the ideas, or for example Evans' partial differential equations 6.3.2 for a proof.

is harmonic on $\Omega'\backslash\{x_o\}$. The condition that $\varphi_j(x) = O(\frac{1}{|x|})$ as $|x| \to \infty$ means that u' is bounded around x_o, so by the removable singularities theorem for harmonic functions extends to be harmonic over Ω'. Now $\Delta u' = 0$ and $u'|_{\partial\Omega'} = 0$ over Ω' so that $u' = 0$ and hence $u(x) = \frac{u'(x/|x|^2)}{|x|} = 0$ on Ω^c. $\square$

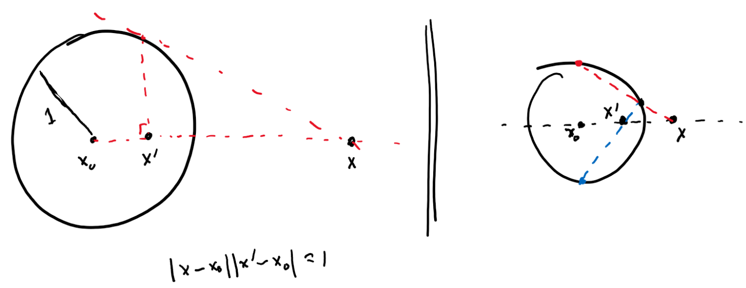

Figure 11. Spherical inversion about a point $x_o \in \mathbb{R}^3$ is a map from $\iota : \mathbb{R}^3\backslash x_o \to \mathbb{R}^3\backslash x_o$ (with $\iota^2 = id$). It is a *conformal map* (angle preserving) with conformal factor $\lambda(x)^4 = \frac{1}{|x-x_o|^4}$. In particular since $\Delta\lambda = 0$ is harmonic on $\mathbb{R}^3\backslash x_o$, inversion may be used to send harmonic functions u to harmonic functions $u' := \lambda\ u \circ \iota$.

As we will see with capacitance, the constant value V of $\varphi|_\Omega$ may be related to the total charge on the conductor, namely through: $\int_{\partial\Omega} \partial_\nu \varphi\ dA = -Q/\varepsilon_o$ with Q the net charge of the conductor. It is also common to consider problems with *grounded conductors*, meaning one takes the solution with $\varphi|_\Omega = 0$, the idea being that the conductor is connected to a 'reservoir' of charges at zero potential (eg the earth).

We remark that our description here is of an 'ideal' conductor: *all* charges constituting the conductor are free to move and the conductor has an unlimited supply of charges (so that it may always redistribute to induce an electric field canceling *any* ambient field in its interior). For a more realistic 'physical' conductor this may not always be the case. Namely, in a physical conductor it is only a certain finite supply of electrons that may move freely, while the protons and their 'close' electrons are more rigidly fixed by the atomic structure of the material. However, provided the ambient electric field is not too strong a sufficiently large supply of free electrons (eg the conductor is grounded) is sufficient to produce the equilibrium conditions we have stated above. We will consider a more realistic description of some effects when the free charges are not sufficient to overcome the ambient field (or there are simply no completely free charges) when we study dielectrics (insulators).

Capacitors: First consider an isolated conductor Ω with no net charge. There is no electric field, and the electric potential is constant throughout space (zero with our condition at infinity). Upon charging the conductor, an electric field will be produced with an electric potential having a new constant value over the conductor. Thus charging a conductor is related to changing the (constant) value of electric potential over the conductor. In fact this relation is linear:

Capacitance of a conductor: If the net charge on a conductor is changed by δQ then the electric potential changes by δV through:

$$\delta Q = C\ \delta V$$

where C is a constant (depending on the conductor) called the *capacitance* of the conductor. It is measured in [1] *Farads* (F = C/V).

proof: Consider the electric potential, φ_1, in Ω^c resulting from a net charge Q_1 on the conductor:

$$\Delta\varphi_1 = 0, \quad \varphi_1|_\Omega = V_1, \quad \varphi_1(x) = O\left(\frac{1}{|x|}\right)$$

[1] After M. Faraday.

for some constant V_1 and:

$$Q_1 = -\varepsilon_o \int_{\partial\Omega} \partial_\nu \varphi_1 \ dA.$$

Likewise, if we give the conductor a net charge Q_2, then the resulting electric potential φ_2 in Ω^c satisfies as well $\Delta\varphi_2 = 0$ and $\varphi_2(x) = O(\frac{1}{|x|})$ with

$$\varphi_2|_\Omega = V_2, \quad Q_2 = -\varepsilon_o \int_{\partial\Omega} \partial_\nu \varphi_2 \ dA$$

for some constant V_2. In this case we may scale solutions. By uniqueness: $V_1\varphi_2 = V_2\varphi_1$. Integrating over the boundary gives:

$$\frac{Q_1}{V_1} = \frac{Q_2}{V_2} = C$$

for some constant C, yielding the linear relation between the increments: $\delta Q := Q_2 - Q_1, \ \ \delta V := V_2 - V_1$. $\square$

Similarly, for a conductor in the presence of ambient electric charges, we first consider the grounded solution:

$$\Delta\varphi_{gr} = -\frac{\rho}{\varepsilon_o}, \quad \varphi_{gr}|_\Omega = 0, \quad \varphi_{gr}(x) = O\left(\frac{1}{|x|}\right).$$

In general, this grounded solution will have some associated net charge on the conductor:

$$Q_{gr} := -\varepsilon_o \int_{\partial\Omega} \partial_\nu \varphi_{gr} \ dA.$$

By linearity, one obtains:

Charge of a conductor in an ambient field: Let the conductor Ω with net charge Q be in the presence of ambient charges, with density ρ. Then the resulting static electric field is generated by the potential φ with $\Delta\varphi = -\frac{\rho}{\varepsilon_o}, \ \ \varphi(x) = O(\frac{1}{|x|})$ on Ω^c and $\varphi|_\Omega = V$ where [1]:

$$Q = Q_{gr} + CV.$$

proof: By uniqueness and linearity, the general solution may be written as $\varphi = \varphi_{gr} + \varphi_V$ where

$$\Delta\varphi_V = 0, \quad \varphi_V|_\Omega = V, \quad \varphi_V(x) = O\left(\frac{1}{|x|}\right)$$

with a constant $V = \varphi|_\Omega$. The total charge on the conductor is then $Q = -\varepsilon_o \int_{\partial\Omega} \partial_\nu \varphi \ dA = Q_{gr} + CV$. $\square$

A *capactitor* is a collection of conductors (at fixed locations). As conductors may store charge, so may a capacitor (we will consider capacitors consisting of *two* conductors). The capacitance of a capacitor consisting of the conductors Ω_1, Ω_2 is defined as the constant of proportionality:

$$Q = C \ V$$

between the potential difference, $V_+ - V_- =: V$, of the two capacitors when charged to $Q > 0$ and $-Q < 0$.

It requires doing work, $W(V)$, to create a potential difference V between two conductors, so that capacitors store energy. Note that the work to move some charge from one conductor to the other depends on the present potential difference between the conductors, ie the present amount of charge on the conductors. To compute this energy (the work required to charge the capacitors to given charges Q and $-Q$), one may consider moving small amounts of charges q_j from one conductor to the other. Then for $\sum q_j = Q$, the total work will be approximately $\sum_1^n V(Q_j)q_j$ where $Q_j := \sum_1^{j-1} q_k$. Letting $\max\{q_j\} \to 0$ we get:

$$W = \int_0^Q V(q) \ dq = \int_0^Q \frac{q}{C} \ dq = \frac{Q^2}{2C} = \frac{1}{2}CV^2$$

[1] That is, the linear relation $\delta Q = C\delta V$ continues to hold in the presence of ambient charges (since $V_{gr} = 0$).

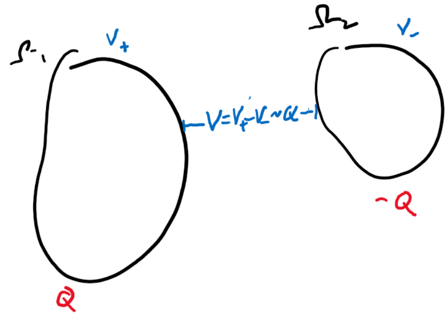

Figure 12. When two conductors are given net charges Q and $-Q$ the potential difference between them is proportional to Q.

for the work stored in a capacitor (energy) with charges $Q, -Q$ on the two conductors.

Dielectrics: We now consider materials at the other extreme from conductors, namely *insulators* or *dielectrics*, in which the constituent charges of the material are not free to move but may only be displaced 'slightly'. A dielectric in the presence of an ambient electric field will thus produce its own induced field as a result of this small displacement of charges in the material.

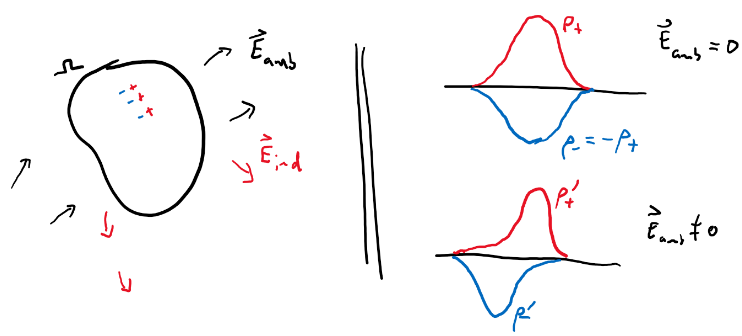

Figure 13. In the presence of an ambient electric field, the charges in a dielectric undergo a small displacement producing their own electric field induced by (in response to) the ambient field.

Since the charge displacements are small, a good approximation to the principal effects of a dielectric may be given by its *polarization*: approximating the displaced charge distribution by dipoles.

To derive some formulas, one may proceed by considering the displacements in response to an ambient electric field of the (positive) charges, ρ_+, in the dielectric given by some transformation ψ_ε (the flow of some vector field X). The negative charges, $\rho_- = -\rho_+$, are then displaced by $\psi_{-\varepsilon}$. We then compute the:

Dipole expansion: Let $\vec{P} := 2\varepsilon\rho_+ X$ be the *dipole density* of a dielectric Ω in response to an ambient electric field. Then *upto order* ε, the displaced charges of Ω generate the potential:

$$4\pi\varepsilon_o\varphi(x) = \int_{y\in\Omega} \vec{P}(y)\cdot\nabla\left(\frac{1}{|x-y|}\right)\,dV = \int_{y\in\Omega}\frac{\rho_b(y)}{|x-y|}\,dV + \int_{y\in\partial\Omega}\frac{\sigma_b(y)}{|x-y|}\,dA$$

due to the densities $\rho_b = -\nabla\cdot\vec{P}$, $\sigma_b = \vec{P}\cdot\nu$, of the displaced charges.

proof: Let ρ_+', ρ_-' be the distributions of the displaced positive and negative charges of the dielectric under the ambient field. By charge conservation, for any region R, we have:

$$\int_{\psi_\varepsilon(R)} \rho_+'\,dV = \int_R \rho_+\,dV$$

where the transformation ψ_ε represents the displacement of the positive charges. As this holds for any region R, one obtains by change of variable:

$$\rho'_+(\psi_\varepsilon(y)) \det d_y\psi_\varepsilon = \rho_+(y)$$

for $y \in \Omega$. Likewise, $\rho'_-(\psi_{-\varepsilon}(y)) \det d_y\psi_{-\varepsilon} = \rho_-(y)$, for the negative charges [1]. Now, for the potential:

$$4\pi\varepsilon_o\varphi(x) = \int_{u\in\psi_\varepsilon(\Omega)} \frac{\rho'_+(u)\ dV}{|x-u|} + \int_{u\in\psi_{-\varepsilon}(\Omega)} \frac{\rho'_-(u)\ dV}{|x-u|}$$

$$= \int_{y\in\Omega} \rho_+(y)\left(\frac{1}{|x-\psi_\varepsilon(y)|} - \frac{1}{|x-\psi_{-\varepsilon}(y)|}\right)\ dV$$

using change of variable and that the undisturbed dielectric is neutral: $\rho_+(y) = -\rho_-(y)$. The difference of inverse distances may be expanded in ε the same way as with a dipole and using $\psi_{\pm\varepsilon}(y) = y \pm \varepsilon X + O(\varepsilon^2)$ to obtain $4\pi\varepsilon_o\varphi(x) = \int_{y\in\Omega} \vec{P}(y)\cdot\nabla(\frac{1}{|x-y|})\ dV + O(\varepsilon^2)$ where the gradient is wrt y and $\vec{P} := 2\varepsilon\rho_+ X$. Integration by parts then yields the second expression for φ with $\rho_b = -\nabla\cdot\vec{P}$ and $\sigma_b = \vec{P}\cdot\nu$. Note as well that these density expressions may be obtained directly as well by expansion of the equations of charge conservation in ε to get, for example, $\rho' = \rho'_+ + \rho'_- = -\nabla\cdot\vec{P} + O(\varepsilon^2)$. □

The response of a dielectric to an ambient electric field depends on the structural properties of the material. Typically one considers the ambient field, $\vec{E}_o$, produced by ambient charges (also called free charges), ρ_o, to be known while the complications of the material structure are contained in $\vec{P}$ and the resulting density, ρ_b, of the displaced charges in the dielectric (called *bound charges*). The main goal is usually to describe the total electric field, $\vec{E}$, produced by the ambient electric field and the dielectric. For this one may consider the following 'shift' to write Maxwell's equations without reference to the bound charges:

$$\nabla\cdot\varepsilon_o\vec{E} = \rho_o + \rho_b = \rho_o - \nabla\cdot\vec{P}$$

$$\Rightarrow \nabla\cdot\vec{D} = \rho_o$$

where $\vec{D} := \varepsilon_o\vec{E} + \vec{P}$ is called the *dielectric displacement*.

For this shift to be useful, one requires some information on $\vec{P}$. The final polarization of the material is a response to the ambient field and the induced field created by the dielectric, ie to the total field, so that the polarization, $\vec{P}(\vec{E})$, is some function [2] of $\vec{E}$ with $\vec{P}(0) = 0$. Expanding in $\vec{E}$, one has:

$$\vec{P} = \varepsilon_o\boldsymbol{\chi}\vec{E} + O(|\vec{E}|^2)$$

where $\boldsymbol{\chi}$ is the *susceptibility* tensor of the material (a matrix with position dependent entries). Ignoring higher order terms, we have 'linear dielectrics', a good model of materials when the electric fields involved are not too strong. A material may be *isotropic*: $\boldsymbol{\chi} = \chi Id$ for some function χ or *homogeneous*: $\boldsymbol{\chi} = cst$. In the linear case:

$$\vec{D} = \varepsilon_o(Id + \boldsymbol{\chi})\vec{E} = \boldsymbol{\varepsilon}\vec{E}$$

where $\boldsymbol{\varepsilon} = \varepsilon_o(Id + \boldsymbol{\chi})$ is called the *permittivity* tensor of the material (in the homogeneous and isotropic case, $\boldsymbol{\varepsilon} = \varepsilon Id$ with $\varepsilon = \varepsilon_o(1+\chi)$ a constant).

The simplest case involving dielectrics is of an isotropic and homogeneous dielectric filling *all* of space. Then $\vec{D} = \varepsilon\vec{E}$ with ε constant and for ambient charges ρ_o, we have: $\nabla\cdot\vec{E} = \frac{\rho_o}{\varepsilon}$, $\nabla\times\vec{E} = 0$, so:

$$\vec{E} = \frac{\varepsilon}{\varepsilon_o}\vec{E}_o$$

[1] The condition of charge conservation may be written as well in differential form: $\nabla\cdot(\rho_\pm X) = 0$ where $X(y) = \frac{d}{d\varepsilon}|_{\varepsilon=0}\psi_\varepsilon(y)$.

[2] There are some materials which may be naturally polarized, meaning $\vec{P}(0) \neq 0$. As well, certain materials may retain a given induced polarization, so even after an ambient electric field is removed they remain polarized.

and the total electric field is a rescaling of $\vec{E}_o$ (the ambient electric field).

More generally, one may consider dielectric materials filling different regions of space with possible jumps in the permittivity. For instance we may have an isotropic and homogeneous dieletric in the region Ω so that $\vec{D} = \varepsilon\vec{E}$ in Ω and $\vec{D} = \varepsilon_o\vec{E}$ in Ω^c. Similarly to conductors, the situation may be formulated in terms of potentials with certain boundary conditions.

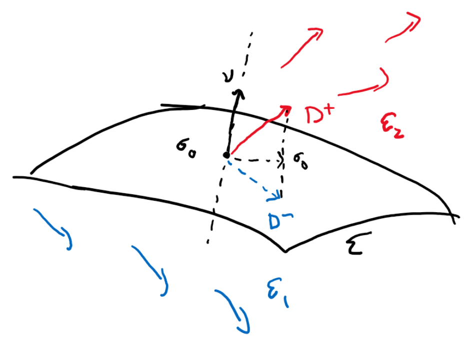

Figure 14. If the ambient charges are distributed over a surface with density σ_o, then the dielectric displacement has a jump in its normal component proportional to the charge density: $(D_+ - D_-)\cdot\nu = \sigma_o$. When the surface is taken as an interface between two dielectric regions, with permittivities $\varepsilon_1, \varepsilon_2$, the condition may be written in terms of the potential: $\varepsilon_1\partial_\nu\varphi_- - \varepsilon_2\partial_\nu\varphi_+ = \sigma_o$.

The uniqueness for dielectric fields is based on:

Neumann boundary conditions: Given a compact region Ω and function $f : \partial\Omega \to \mathbb{R}$ with zero average over $\partial\Omega$, then upto addition of constants there is a unique harmonic function, $u : \Omega \to \mathbb{R}$ satisfying:

$$\Delta u = 0, \quad \partial_\nu u|_{\partial\Omega} = f.$$

proof: Note it is necessary that f have zero average. If there is a solution, u, then: $\int_{\partial\Omega} f\, dA = \int_{\partial\Omega} \partial_\nu u\, dA = \int_\Omega \Delta u\, dV = 0$. Consider the difference, $v := u_1 - u_2$, of two solutions. Then $\int_\Omega |\nabla v|^2\, dV = \int_{\partial\Omega} v\partial_\nu v\, dA = 0$ so that $\nabla v \equiv 0$ in Ω and so v is constant. $\square$

Dielectric potentials: Consider a compact homogeneous and isotropic dielectric, Ω, with permittivity ε in the presence of ambient charges ρ_o. Then the electric field is generated by the unique potential φ, continuous over $\partial\Omega$ and solving:

$$\Delta\varphi = \begin{cases} -\frac{\rho_o}{\varepsilon} & \text{in } \Omega \\ -\frac{\rho_o}{\varepsilon_o} & \text{in } \Omega^c \end{cases},$$

$$\varepsilon\partial_\nu\varphi_- - \varepsilon_o\partial_\nu\varphi_+ = \sigma_o, \text{ on } \partial\Omega, \quad \varphi(x) = O\left(\frac{1}{|x|}\right), \quad |x| \to \infty.$$

We summarize these equations of electrostatics with materials in the following table:

Conductors	
$\Delta\varphi = -\frac{\rho}{\varepsilon_o}$	$\sigma = -\varepsilon_o \partial_\nu \varphi$
$\varphi\|_{\partial\Omega} \equiv V = cst.$	$\varphi(x) = O(\frac{1}{\|x\|})$.
Capacitance	
$\delta Q = C\ \delta V$	$W = \frac{CV^2}{2}$.
Dielectrics	
$\vec{P} \cdot \nu = \sigma_b, \quad \nabla \cdot \vec{P} = \rho_b$	$\vec{D} = \varepsilon_o \vec{E} + \vec{P}, \quad \nabla \cdot \vec{D} = \rho_o$
Linear Dielectrics	
$\vec{P} = \varepsilon_o \boldsymbol{\chi} \vec{E}$	$\vec{D} = \boldsymbol{\varepsilon} \vec{E}$
$\boldsymbol{\varepsilon} = \varepsilon_o(Id + \boldsymbol{\chi})$	
Homogeneous: $\boldsymbol{\chi} \equiv cst.$	Isotropic: $\boldsymbol{\chi} = \chi\ Id$
$\varepsilon = \varepsilon_o(1 + \chi)$	

Table 2. Equations of electrostatics involving materials.

EXERCISES:

1. Determine the electric field generated by a uniformly charged line, with charge density $\lambda = cst.$
2. Let X be a vector field on $\mathbb{R}^3$ with $\nabla \cdot X = 0$. For $p \in \mathbb{R}^3$, set
$$Y(p) := \int_0^1 tX(tp) \times p\ dt.$$
Show that $\nabla \times Y = X$.
3. For $x_o, x \in \mathbb{R}^3$ and $r := |x - x_o|$, show that $\nabla \frac{1}{r} = -\nabla^o \frac{1}{r}$ where ∇ is gradient with respect to x and ∇^o gradient with respect to x_o.
4. Let X be a vector field on $\mathbb{R}^3$ and $\Omega \subset \mathbb{R}^3$ a compact region with smooth boundary. Show that $\int_\Omega \nabla \times X\ dV = \int_{\partial\Omega} \nu \times X\ dA$.
5. Let $\vec{B}$ be a vector field on $\mathbb{R}^3$ with $\nabla \cdot \vec{B} = 0$ and $\nabla \times \vec{B} = 0$. Suppose that $\vec{B}$ satisfies the 'conditions at infinity' of $\lim_{r\to\infty} r\vec{B} = 0$. Show that $\vec{B} = 0$ (here feel free to make use of the formulas in the notes on the Helmholtz decomposition).
6. Let $\Omega \subset \mathbb{R}^3$ be a compact region with smooth boundary $\partial\Omega$.

   (a) Let f be a smooth function on Ω. Show that $\nabla \cdot (f\nabla f) = |\nabla f|^2 + f\Delta f$.

   (b) Suppose u is a harmonic function, $\Delta u = 0$, on Ω, with $u|_{\partial\Omega} = 0$. Show that $u = 0$ (consider the flux of $u\nabla u$ through $\partial\Omega$).
7. Consider inversion over the unit circle in the plane, $\mathbb{R}^2 \backslash \mathbf{0} \ni \mathbf{x} \mapsto \frac{\mathbf{x}}{|\mathbf{x}|^2} = \mathbf{x}' \in \mathbb{R}^2 \backslash \mathbf{0}$.

   (a) For two points $\mathbf{x}_1, \mathbf{x}_2$ (not along the same line through the origin) show that the triangles $\Delta(\mathbf{0}, \mathbf{x_1}, \mathbf{x_2})$ and $\Delta(0, \mathbf{x_2'}, \mathbf{x_1'})$ are similar.

   (b) Given two points, $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ show that $\mathbf{c} \in \mathbb{R}^2$ lies on the circle with diameter $\overline{\mathbf{ab}}$ iff the triangle $\Delta(\mathbf{a}, \mathbf{b}, \mathbf{c})$ has a right angle at $\mathbf{c}$.

   (c) Show that inversion sends lines (not passing through $\mathbf{0}$) to circles (passing through $\mathbf{0}$).

(d) Show that inversion preserves angles [1].

8. For inversion over the unit circle in the plane, $\mathbf{x} \mapsto \frac{\mathbf{x}}{|\mathbf{x}|^2} = \mathbf{x}'$

   (a) For $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2 \backslash \mathbf{0}$ show that $|\mathbf{x_1}'\mathbf{x_2}'| = \frac{|\mathbf{x}_1\mathbf{x}_2|}{|\mathbf{x}_1||\mathbf{x}_2|}$.

   (b) Let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ be ordered points on some circle (vertices of a quadrilateral $\square(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})$ inscribed in some circle). Show *Ptolemy's theorem*:

$$|\mathbf{ab}||\mathbf{cd}| + |\mathbf{ad}||\mathbf{cb}| = |\mathbf{ac}||\mathbf{bd}|.$$

9. Let $\iota : \mathbb{R}^3 \backslash 0 \to \mathbb{R}^3 \backslash 0, \ \ \mathbf{x} \mapsto \frac{\mathbf{x}}{|\mathbf{x}|^2} = \mathbf{x}'$ be spherical inversion about a unit sphere centered at the origin.

   (a) Show that inversion is a conformal map (preserves angles).

   (b) For $\vec{u}, \vec{v}$ two vectors at $\mathbf{x}$ show that $d\iota_{\mathbf{x}}\vec{u} \cdot d\iota_{\mathbf{x}}\vec{v} = \frac{1}{|\mathbf{x}|^4} \ \vec{u} \cdot \vec{v}$.

10. Let $f : \mathbb{R}^3 \to \mathbb{R}^3, \ \ \mathbf{x} \mapsto \mathbf{x}' = f(\mathbf{x})$ be a conformal mapping with $df_{\mathbf{x}}\vec{u} \cdot df_{\mathbf{x}}\vec{v} = \lambda(\mathbf{x})^4 \ \vec{u} \cdot \vec{v}$ and $\Delta\lambda = 0$.

    (a) For $u' : \mathbb{R}^3 \to \mathbb{R}$ harmonic show that[2] $\lambda\Delta(u' \circ f) = -2\nabla\lambda \cdot \nabla(u' \circ f)$.

    (b) Set $u := \lambda \ u' \circ f$. Show that u is harmonic.

11. Show *Liouville's theorem*: if $u : \mathbb{R}^3 \to \mathbb{R}$ is harmonic and bounded then u is constant.

12. Let $x_o \in \mathbb{R}^3$ and $\varphi(x) = \frac{1}{|x - x_o|}$. Show that $\varphi'(x) := \frac{\varphi(x/|x|^2)}{|x|}$ may be written as:

$$\varphi'(x) = \frac{1}{|x_o||x - x_o'|}$$

    where $x_o' := x_o/|x_o|^2$.

13. Determine the static electric field generated by a (grounded) conducting cylinder, $\Omega = \{x^2 + y^2 \leq 1\}$, and uniformly charged line with charge density λ parallel and at distance r_o from the cylinders axis.

14. Consider a capacitor consisting of two conductors, Ω_1, Ω_2 with capacitance C.

    (a) Consider the conductors are both given charge Q with resulting potential values V_1, V_2 over Ω_1, Ω_2. Show:

$$Q = C_o V$$

    for some constant C_o and $V = V_1 - V_2$ the potential difference between the conductors.

    (b) Suppose the capacitors are given charges Q_1, Q_2 with resulting potential values V_1, V_2 over the conductors. Set $V := V_1 - V_2$, $Q_{diff} := Q_1 - Q_2$, and $Q_{tot} := Q_1 + Q_2$. Show:

$$2V = \frac{Q_{diff}}{C} + \frac{Q_{tot}}{C_o}.$$

15. Consider a capacitor consisting of two concentric conducting cylinders of radii $a < b$ and height h. For $h >> b - a$, and approximating the fields of charged cylinders by those of infinite cylinders show the capacitance is:

$$C = \frac{2\pi\varepsilon_o h}{\log(b/a)}.$$

16. Consider a capacitor consisting of two conductors Ω_1, Ω_2 having capacitance C. Show that:

$$\frac{Q^2}{C} = \int_{\mathbb{R}^3} \varepsilon_o |\vec{E}|^2 \ dV$$

    where $\vec{E}$ is the electric field generated by the conductors when charged to Q and $-Q$.

[1] Suggestion: consider the images of two intersecting lines under inversion (two circles intersecting at $\mathbf{0}$) and consider the angle between the tangents to these circles at $\mathbf{0}$.

[2] See our formula for Laplacian in orthogonal coordinates at the top of pg. 15. Here this reads: $\lambda^6\Delta u' = \nabla \cdot (\lambda^2\nabla(u' \circ f))$.

17. Consider a charge distribution ρ_1, with corresponding potential (vanishing at infinity) φ_1. For any other other charge distribution ρ_2 with potential φ_2 vanishing at infinity show *Green's reciprocity theorem*:

$$\int_{\mathbb{R}^3} \rho_1 \varphi_2 \, dV = \int_{\mathbb{R}^3} \rho_2 \varphi_1 \, dV.$$

18. Consider an interface of two dielectrics: Ω_1, Ω_2 are (homogeneous and isotropic) dielectrics with permittivities $\varepsilon_1, \varepsilon_2$ sharing a common boundary (interface) $\partial\Omega_1 = \partial\Omega_2$.

    Let $\vec{E}$ be the electric field generated by these dielectrics in the presence of some ambient charges, ρ_o (with $\rho_o = 0$ on the interface). Show that the electric field satisfies the following *refraction law*:

$$\varepsilon_2 \cot \alpha_2 = \varepsilon_1 \cot \alpha_1$$

    where α_j is the angle between $\vec{E}$ and the normal to the interface on the side Ω_j.

19. Consider a collection of ambient charges with density ρ_o and potential φ_o. Show that if a homogeneous and isotropic dielectric, Ω, with permittivity ε is placed with boundary along the level sets of φ_o then the resulting electric field is given by:

$$\vec{E} = \begin{cases} \vec{E}_o & \text{in } \Omega^c \\ \frac{\varepsilon_o}{\varepsilon} \vec{E}_o & \text{in } \Omega \end{cases}$$

    where $\vec{E}_o = -\nabla \varphi_o$ is the electric field produced by the ambient charges.

20. Determine the electric field when a point charge q is placed on the interface, $\{z = 0\}$, between two dielectric half spaces, $\Omega_1 = \{z < 0\}$, $\Omega_2 = \{z > 0\}$, having permittivities $\varepsilon_1, \varepsilon_2$.

21. Consider a solid ball of radius R, which is uniformly polarized: the dipole density $\vec{P} = cst.$ over the ball. Determine the electric field generated by this dipole distribution on the ball.

§4 magnetostatics

We know examine the following situation: to determine the effects produced by a collection of *steady* currents [1] in a given inertial frame.

The connection of magnetism to electricity was first given by Oersted, who noticed that a current produces a magnetic field. Conversely, a magnetic field exerts a force on moving charges:

Lorentz (magnetic) force: The force, $\vec{f}_m$, of a magnetic field $\vec{B}$ on a test particle of charge q moving with velocity $\vec{v}$ is:

$$\vec{f}_m = q\,\vec{v} \times \vec{B}.$$

Thus a magnetic field may be measured by 'throwing' charges in various directions. The basis for this magnetostatics (analogous to Coulomb's law for electrostatics) is the:

Biot-Savart law: The magnetic field produced by a steady current $\vec{I}$ along the curve C is given by:

$$\vec{B}(x) = \frac{\mu_o}{4\pi} \int_{y \in C} \frac{\vec{I}(y) \times (x-y)}{|x-y|^3}\, ds$$

where $\frac{\mu_o}{4\pi} = 1.3 \times 10^{-6} N/A^2$.

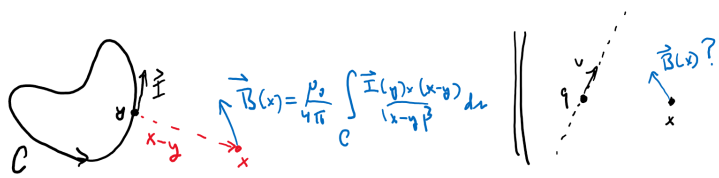

Figure 15. The Biot-Savart law gives the magnetic field produced by a wire carrying a steady current. It is reasonable to wonder why we have not begun –as with electrostatics– with the case of a point charge moving with velocity v and stated a simpler law more analogous to Coulomb's law for point charges. One complication is that a solitary moving point charge does not produce a *steady* current, so we fall outside our present realm of statics. *Approximately* however a point charge q moving with velocity v produces the magnetic field $\vec{B}(x) \approx \frac{\mu_o}{4\pi} qv \times \frac{x-y(t)}{|x-y(t)|^3}$ where $y(t)$ is the position of the point charge at time t.

By superposition, a steady current density[2] $\vec{J}$ produces a magnetic field:

$$\vec{B}(x) = \frac{\mu_o}{4\pi} \int_{y \in \mathbb{R}^3} \frac{\vec{J}(y) \times (x-y)}{|x-y|^3}\, dV.$$

Thus, when the steady currents are known, magnetostatics at its core is reduced to the evaluation of (typically complicated) integrals. As with electrostatics, there are analogous properties of the magnetic field that underly its study and are especially useful in symmetric situations:

Gauss' law for magnetostatics: the magnetic field, $\vec{B}$, produced by a (bounded) current density $\vec{J}$ satisfies:

$$\nabla \cdot \vec{B} = 0.$$

Ampere's law: the magnetic field, $\vec{B}$, produced by a (bounded) current density $\vec{J}$ satisfies:

$$\nabla \times \vec{B} = \mu_o \vec{J}.$$

proof: Let R be any compact region with smooth boundary ∂R. We will compute that $\int_{\partial R} \vec{B} \cdot d\vec{S} = 0$, so that $\int_R \nabla \cdot \vec{B}\, dV = 0$ for any R and hence $\nabla \cdot \vec{B} = 0$.

[1] Also called *direct currents*, ie the currents (directions and rates of moving charges) do not depend on time (but may depend on position).

[2] Likewise, a surface density $\vec{K}$, of current over the surface Σ produces $\vec{B}(x) = \frac{\mu_o}{4\pi} \int_{y \in \Sigma} \frac{\vec{K}(y) \times (x-y)}{|x-y|^3}\, dA$.

Since the currents are bounded, we have $\vec{J} = 0$ outside some ball, B_r, of radius r. Then:

$$\int_{\partial R} \vec{B} \cdot d\vec{S} = \frac{\mu_o}{4\pi} \int_{x_o \in \partial R} \left(\int_{y \in B_r} \nabla^o\left(\frac{1}{|x_o - y|}\right) \times \vec{J}(y) \, dV \right) \cdot \nu \, dA.$$

Since $\frac{x_o - y}{|x_o - y|^3} = -\nabla^o \frac{1}{|x_o - y|}$, where ∇^o indicates gradient wrt x_o. From the product rule $\nabla \times (fX) = \nabla f \times X + f \nabla \times X$, and changing the order of integration, the integral is proportional to:

$$\int_{y \in B_r} \left(\int_{x_o \in \partial R} \nabla^o\left(\frac{1}{|x_o - y|}\right) \times \vec{J}(y) \cdot \nu \, dA \right) dV = \int_{y \in B_r} \left(\int_{x_o \in \partial R} \nabla^o \times \frac{\vec{J}(y)}{|x_o - y|} \cdot \nu \, dA \right) dV = 0$$

since the divergence of a curl is zero. Note that this same computation gives $\vec{B} = \nabla \times \vec{A}$ where

$$\vec{A}(x) := \frac{\mu_o}{4\pi} \int_{y \in \mathbb{R}^3} \frac{\vec{J}(y)}{|x - y|} \, dV$$

is the magnetic vector potential. Next, for Ampere's law, we will make use of *charge conservation*. Namely, for a given region R with smooth boundary ∂R, then the rate at which charge is *leaving* R is $\int_{\partial R} \vec{J} \cdot \nu \, dA = \int_R \nabla \cdot \vec{J} \, dV$. On the other hand, the rate the total charge inside R is changing is $\int_R \partial_t \rho \, dV$ so that:

$$\int_R \partial_t \rho \, dV = -\int_R \nabla \cdot \vec{J} \, dV.$$

Since R was arbitrary, we have the *continuity equation*:

$$\nabla \cdot \vec{J} = -\partial_t \rho.$$

In the static situation $\partial_t \rho = 0$ so that $\nabla \cdot \vec{J} = 0$. Now we return to Ampere's law:

$$\nabla^o \times \vec{B}(x_o) = \frac{\mu_o}{4\pi} \nabla^o \times \int_{y \in B_r} \nabla^o\left(\frac{1}{|x_o - y|}\right) \times \vec{J}(y) \, dV = \frac{\mu_o}{4\pi} \int_{y \in B_r} \nabla^o \times \left(\nabla^o \times \frac{\vec{J}(y)}{|x_o - y|}\right) dV.$$

Now we use the definition of the vector Laplacian, $\Delta X = \nabla(\nabla \cdot X) - \nabla \times (\nabla \times X)$, as well as that $\Delta^o\left(\frac{1}{|x_o - y|}\right) = -4\pi \delta_{x_o}$, so that:

$$\nabla^o \times \vec{B}(x_o) = \mu_o \vec{J}(x_o) + \frac{\mu_o}{4\pi} \nabla^o \int_{y \in B_r} \nabla^o \cdot \left(\frac{\vec{J}(y)}{|x_o - y|} \right) dV.$$

For the remaining integral term, note that $\nabla \cdot \left(\frac{\vec{J}(y)}{|x_o - y|}\right) = -\nabla^o \cdot \left(\frac{\vec{J}(y)}{|x_o - y|}\right)$ since $\nabla \cdot \vec{J} = 0$, and where ∇ denotes divergence wrt y. Thus, using the divergence theorem,

$$\int_{y \in B_r} \nabla^o \cdot \left(\frac{\vec{J}(y)}{|x_o - y|} \right) dV = -\int_{y \in B_r} \nabla \cdot \left(\frac{\vec{J}(y)}{|x_o - y|} \right) dV = -\int_{y \in \partial B_r} \frac{\vec{J}(y) \cdot \nu}{|x_o - y|} \, dA = 0$$

since the currents are bounded (so $\vec{J} = 0$ on ∂B_r). In general one obtains the same results if the currents decay sufficiently rapidly at infinity so that the last integral goes to zero as $r \to \infty$. $\square$

We remark that the 'magnetic Gauss' law' is often not named, or referred to as the condition for *no magnetic monopoles*. As well, Ampere's law is the magnetic analogue to the Gauss' law in electrostatics. Namely the line integral of $\vec{B}$ along the boundary, $\partial \Sigma$, of an oriented surface Σ is proportional to the rate of charge (current), $I_\Sigma = \int_\Sigma \vec{J} \cdot d\vec{S}$, passing through the surface:

$$\int_{\partial \Sigma} \vec{B} \cdot d\vec{s} = \mu_o I_\Sigma.$$

That the magnetic field is divergence free is equivalent to it being the curl of some vector field (also called a *solenoidal vector field*). So we have a vector analogue of potential.

Magnetic potential: the magnetic field, $\vec{B}$, produced by a (bounded) current density $\vec{J}$ is a solenoidal vector field:

$$\vec{B} = \nabla \times \vec{A}$$

where $\vec{A}(x) = \frac{\mu_o}{4\pi} \int_{y \in \mathbb{R}^3} \frac{\vec{J}(y)}{|x-y|} \, dV$ is called the *magnetic vector potential*.

proof: In fact we have already essentially seen this in our verification of the magnetic Gauss' law, since for a general surface Σ, we have:

$$\int_\Sigma \vec{B} \cdot d\vec{S} = \int_{x_o \in \Sigma} \nabla^o \times \left(\frac{\mu_o}{4\pi} \int_{y \in B_r} \frac{\vec{J}(y)}{|x_o - y|} \, dV \right) \cdot \nu \, dA = \int_\Sigma \nabla \times \vec{A} \cdot d\vec{S}.$$

$\square$

Since the vector Laplacian in Cartesian coordinates is a componentwise Laplacian, we have by applying our observations on electric potentials componentwise that the magnetic potential satisfies the 'vector Poisson equation': $\Delta \vec{A} = -\mu_o \vec{J}$. The additional condition, that $\vec{A}$ vanish at infinity, determines $\vec{A}$ uniquely, as well as implies $\nabla \cdot \vec{A} = 0$ (from charge conservation: $\nabla \cdot \vec{J} = 0$).

The magnetic potential is useful for similar reasons to the electric potential. When symmetry considerations do not suffice, the integrals giving the components of the magnetic potential may often be more manageable than those giving the magnetic field directly.

We summarize our results so far in the following table:

Maxwell's equations in magnetostatics	
Integral form	Differential form
$\oint_\Sigma \vec{B} \cdot d\vec{S} = 0$	$\nabla \cdot \vec{B} = 0$
$\int_{\partial\Sigma} \vec{B} \cdot d\vec{s} = \mu_o \int_\Sigma \vec{J} \cdot d\vec{S}$	$\nabla \times \vec{B} = \mu_o \vec{J}$
Magnetostatic potential theory	
$\vec{B} = \nabla \times \vec{A}, \quad \nabla \cdot \vec{A} = 0$	$\Delta \vec{A} = -\mu_o \vec{J}$

Table 3. Magnetostatics may be summarized as a special case of Maxwell's equations.

We will now briefly consider the behaviour of materials in response to ambient magnetic fields. When a material is placed in an ambient magnetic field, $\vec{B}_o$, it may become *magnetized*, producing a magnetic field $\vec{B}_b$ of its own.

As with polarization of dielectrics, the principal effects of this magnetization may be described by a *magnetic dipole density*, or *magnetization*, $\vec{M}$. If the material occupies the region Ω, then this induced magnetic field is generated by the vector potential:

$$\vec{A}_b(x) = \frac{\mu_o}{4\pi} \int_{y \in \Omega} \frac{\vec{M}(y) \times (x-y)}{|x-y|^3} \, dV.$$

Integrating by parts, we have:

$$\vec{A}_b(x) = \frac{\mu_o}{4\pi} \left(\int_{y \in \Omega} \frac{\nabla \times \vec{M}(y)}{|x-y|} \, dV + \int_{y \in \partial\Omega} \frac{\vec{M}(y) \times \nu}{|x-y|} \, dA \right)$$

so that the induced magnetic field, $\vec{B}_b$, is that produced by the *bound current densities*:

$$\vec{J}_b = \nabla \times \vec{M}, \quad \text{in } \Omega$$

$$\vec{K}_b = \vec{M} \times \nu, \quad \text{in } \partial\Omega.$$

Let $\vec{B} = \vec{B}_o + \vec{B}_b$ be the total magnetic field produced by the ambient charges and the response of the magnetized material. Applying a similar 'shift' to the magnetostatic equations that we applied with dielectrics, we take [1] $\vec{H} := \frac{1}{\mu_o}\vec{B} - \vec{M}$ so that Ampere's law, $\nabla \times \vec{B} = \mu_o(\vec{J}_o + \vec{J}_b)$ may be written:

$$\nabla \times \vec{H} = \vec{J}_o$$

where $\vec{J}_o$ is the the ambient or 'free' current density producing the ambient magnetic field $\vec{B}_o$.

In general, these considerations are not very useful unless some assumption or information is known about $\vec{M}$. *Linear materials*, are studied under the assumption that

$$\boldsymbol{\mu}\vec{H} = \vec{B}$$

for some (in general position dependent) matrix $\boldsymbol{\mu}$ called the *permeability tensor* of the material. It follows that:

$$\vec{M} = \boldsymbol{\chi}_m\vec{H}$$

for $\boldsymbol{\mu} = \mu_o(Id + \boldsymbol{\chi}_m)$, and where $\boldsymbol{\chi}_m$ is called the *magnetic susceptibility tensor* of the material.

Linear materials are called *homogeneous* when $\boldsymbol{\mu}$ is a constant matrix, and *isotropic* when $\boldsymbol{\mu} = \mu Id$ for some function μ. Homogeneous and isotropic materials are then characterized by a constant *permeability* μ (and *magnetic susceptibility* χ_m) related through:

$$\mu = \mu_o(1 + \chi_m).$$

Time independent, or *static*, solutions to Maxwell's equations may be split into a system of (parallel) electrostatic and magnetostatic systems. We summarize the main results in the table below.

The material on statics corresponds to ch. 2-6 of Griffiths, where ch. 2-4 cover electrostatics and ch. 5,6 magnetostatics. In Feynman, see lectures 4-12 on electrostatics and lectures 13-15 on magnetostatics, as well as lectures 30-37 which go further on materials.

[1] This vector is given various names. Sometimes it is (confusingly) called the magnetic field, and then $\vec{B}$ is given a different name. One might also find it called the *magnetic displacement* or *magnetizing field*. We will just call it 'H'.

Electrostatics	Magnetostatics
$\vec{f}_{el} = q\vec{E}$	$\vec{f}_{mag} = q(v \times \vec{B})$
Lorentz force: $\vec{f} = q(\vec{E} + v \times \vec{B})$	
λ, σ, ρ	$\vec{I}, \vec{K}, \vec{J}$
Charge conservation: $\nabla \cdot \vec{J} = -\partial_t \rho = 0$	
$\vec{E}(x) = \frac{1}{4\pi\varepsilon_o} \int_{y\in\mathbb{R}^3} \frac{\rho(y)(x-y)}{\lvert x-y\rvert^3}\ dV$	$\vec{B}(x) = \frac{\mu_o}{4\pi} \int_{y\in\mathbb{R}^3} \frac{\vec{J}(y)\times(x-y)}{\lvert x-y\rvert^3}\ dV$
$\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_o}, \quad \nabla \times \vec{E} = 0$	$\nabla \cdot \vec{B} = 0, \quad \nabla \times \vec{B} = \mu_o \vec{J}$
$\varphi(x) = \frac{1}{4\pi\varepsilon_o} \int_{y\in\mathbb{R}^3} \frac{\rho(y)}{\lvert x-y\rvert}\ dV$	$\vec{A}(x) = \frac{\mu_o}{4\pi} \int_{y\in\mathbb{R}^3} \frac{\vec{J}(y)}{\lvert x-y\rvert}\ dV$
$\Delta\varphi = -\frac{\rho}{\varepsilon_o}, \quad \vec{E} = -\nabla\varphi$	$\Delta\vec{A} = -\mu_o\vec{J}, \quad \vec{B} = \nabla \times \vec{A}$
Materials	
Polarization: $\vec{P}$	Magnetization: $\vec{M}$
$\nabla \cdot \vec{P} = -\rho_b, \quad \vec{P} \cdot \nu = \sigma_b$	$\nabla \times \vec{M} = \vec{J}_b, \quad \vec{M} \times \nu = \vec{K}_b$
$\vec{D} = \varepsilon_o\vec{E} + \vec{P}, \quad \nabla \cdot \vec{D} = \rho_o$	$\vec{H} = \frac{1}{\mu_o}\vec{B} - \vec{M}, \quad \nabla \times \vec{H} = \vec{J}_o$
Linear materials	
$\vec{P} = \varepsilon_o \boldsymbol{\chi}\vec{E}, \quad \vec{D} = \boldsymbol{\varepsilon}\vec{E}$	$\vec{M} = \boldsymbol{\chi}_m\vec{H}, \quad \boldsymbol{\mu}\vec{H} = \vec{B}$
$\boldsymbol{\varepsilon} = \varepsilon_o(Id + \boldsymbol{\chi})$	$\boldsymbol{\mu} = \mu_o(Id + \boldsymbol{\chi}_m)$

Table 4. Maxwell's equations in the static (no time dependence) case split into electrostatics and magnetostatics.

Exercises:

1. Determine the magnetic field, $\vec{B}$, produced by a steady current $\vec{I} = \lambda\vec{v}$ (with $\lambda, \vec{v}$ constants and $I := \lambda|\vec{v}|$) flowing along an infinite line directed by $\vec{v}$.
2. Determine a magnetic vector potential, $\vec{B} = \nabla \times \vec{A}$, for the magnetic field of the previous problem.
3. Suppose $\nabla \times \vec{A} = \nabla \times \vec{A}'$ for two vector fields $\vec{A}, \vec{A}'$ on $\mathbb{R}^3$. Show that $\vec{A}' = \vec{A} + \nabla f$ for some function f.
4. For a function f and vector field X, show that: $\nabla \times (fX) = \nabla f \times X + f\nabla \times X$.
5. Consider two parallel (straight line) wires with steady constant currents I_1, I_2 and seperated by a distance d. Show that the force per unit length on the wires is given by *Ampere's force law*:
$$f = \frac{\mu_o}{2\pi}\frac{I_1 I_2}{d}.$$
6. Determine the magnetic field produced by a steady current flowing along an infinite plane with current surface density $\vec{K} = cst.$ (tangent to the plane).
7. Consider a compact region, Ω, with smooth boundary $\partial\Omega$ and magnetization (magnetic dipole density) $\vec{M}$. Integrating by parts (see exercise # 4 on pg. 33), show that the vector potential may be written:
$$\vec{A}(x) = \frac{\mu_o}{4\pi}\int_{y\in\Omega} \frac{\vec{M}(y) \times (x-y)}{|x-y|^3} = \frac{\mu_o}{4\pi}\left(\int_{y\in\Omega} \frac{\vec{J}_b(y)}{|x-y|}\ dV + \int_{y\in\partial\Omega} \frac{\vec{K}_b(y)}{|x-y|}\ dA\right)$$
where $\vec{J}_b = \nabla \times \vec{M}, \quad \vec{K}_b = \vec{M} \times \nu$ (and ν the exterior unit normal to $\partial\Omega$).

8. Let $\mathcal{C} = \partial\Sigma$ be a closed curve oriented with unit tangent T which is the boundary of some surface Σ. For a (smooth) function $f : \mathbb{R}^3 \to \mathbb{R}$ show that:

$$\oint_{\mathcal{C}} fT\ ds = \int_{\Sigma} \nu \times \nabla f\ dA$$

where ν is the unit normal to Σ.

9. Consider a simple closed *planar* curve, $\mathcal{C}$, along which flows a steady current $\vec{I} = IT$ for I constant and T the unit tangent to the curve. Let x_o be a fixed origin. Show that the magnetic potential may be expanded as:

$$\vec{A}(x) = \frac{\mu_o}{4\pi}\vec{\mu} \times \frac{x - x_o}{|x - x_o|^3} + O(\frac{1}{|x - x_o|^3})$$

as $|x - x_o| \to \infty$ and where $\vec{\mu} = IA\hat{n}$, for $\hat{n}$ the unit normal to the plane containing $\mathcal{C}$.

10. Determine the magnetic field produced by a uniformly magnetized solid ball, $\vec{M} = cst.$.

11. Consider a solid ball with uniform charge density ρ. Determine the magnetic field produced when the ball is rotated at uniform angular speed about a fixed axis passing through its center.

12. Consider a motion, $\mathbf{x}(t)$, of a charged test particle subject to the magnetic field $\vec{B}$.

(a) Show that the velocity $|\dot{\mathbf{x}}|$ of the test particle remains constant.

(b) If $\vec{B}$ is perpendicular to some plane, and the test particle begins in this plane with an initial velocity tangent to this plane, show that the test particle remains in the plane.

13. Consider a magnetic field $(0, 0, B(x, y))$, perpendicular to the xy-plane. Show that the curvature [1]of a test particles motion, $\mathbf{x}(t)$, in the xy-plane is given by:

$$\kappa(t) = \frac{q}{mv}B(\mathbf{x}(t))$$

where $v = |\dot{\mathbf{x}}|$ and q, m are the charge and mass of the test particle.

14. Show that the motions, $\mathbf{x}(t)$, of a charged test particle in a constant magnetic field $\vec{B}$ are circles or helices.

15. Consider a plane curve, parametrized by *arc-length*: $\mathbf{x}(s)$ with $|\frac{d\mathbf{x}}{ds}| = 1$.

(a) Show that $\frac{d^2\mathbf{x}}{ds^2}$ is perpendicular to $\frac{d\mathbf{x}}{ds}$.

(b) Write $\mathbf{x}(s) = (x(s), y(s))$ and $\frac{d\mathbf{x}}{ds}(s) = (\cos\theta(s), \sin\theta(s))$. Show that:

$$\frac{d\theta}{ds} = \frac{d^2y}{ds^2}\frac{dx}{ds} - \frac{d^2x}{ds^2}\frac{dy}{ds}.$$

16. Show that a plane curve is a circle of radius $r > 0$ iff it has curvature $k(s) = 1/r$

[1]The curvature of a plane curve $(x(t), y(t)) = \mathbf{x}(t)$ is given by $\kappa(t) = \frac{\ddot{y}\dot{x} - \ddot{x}\dot{y}}{|\dot{\mathbf{x}}|^3}$.

§5 examples

We collect here various examples of static (time-independent) electromagnetic fields.

EXAMPLES OF ELECTROSTATIC FIELDS:

- A *uniformly charged plane*, with charge density $\sigma = cst.$ produces an electric field of strength:

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

directed along the outer normals to the plane (and zero on the plane).

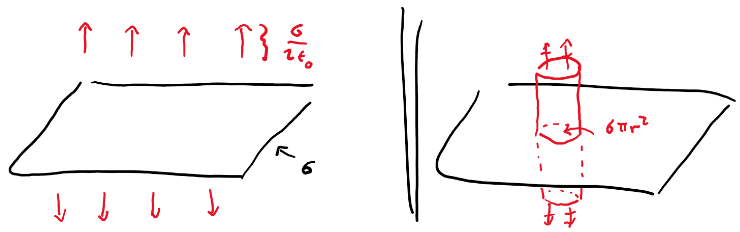

Figure 16. A uniformly charged plane (shown here with charge density $\sigma > 0$) generates an electric field normal to the plane with constant strength.

This may be derived by direct evaluation of the integral:

$$\vec{E} = \frac{\sigma}{4\pi\varepsilon_o}\int_0^\infty\int_0^{2\pi}\frac{(-r\cos\theta, -r\sin\theta, z)}{(r^2+z^2)^{3/2}}\, rd\theta dr.$$

Note that in handling these integrals one may argue by symmetry that only the vertical component is non-zero. Namely at a given point above the plane, the electric field remains the same under rotations about the axis from this point to the plane –since the plane and constant charge distribution remains unchanged. Hence the electric field must be directed along this vertical axis. This general principle – *a symmetry in the charge distribution leads to a symmetry in the resulting electric field* – is very useful when it can be exploited.

Alternately, the result here can be derived more simply by applying the Gauss' law: consider a vertical cylinder, Ω, centered on the plane then:

$$2E\pi r^2 = \int_{\partial\Omega}\vec{E}\cdot d\vec{S} = \frac{\pi r^2\sigma}{\varepsilon_o} \Rightarrow E = \frac{\sigma}{2\varepsilon_o}.$$

This electric field is a potential field. We may take as a base-point a point on the plane (in Cartesian coordinates as the $z = 0$ plane) and have:

$$\varphi = -\frac{\sigma}{2\varepsilon_o}|z|.$$

Note that the usual formula, with a base-point 'at infinity' for electric potential, $\varphi(x') = \int \frac{\sigma\, dA}{4\pi\varepsilon_o|x-x'|}$, gives a diverging integral here as the charge distribution is infinite.

- A *uniformly charged sphere* of radius R with charge density $\sigma = cst.$ produces a radial electric field with strength:

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

in the exterior of the sphere (here r is the distance to the center of the sphere and $Q = 4\pi R^2\sigma$ the total charge of the sphere) and a vanishing electric field in the interior of the sphere.

Figure 17. A uniformly charged sphere with total charge Q generates a radial electric field, with strength $\frac{Q}{4\pi\varepsilon_o r^2}$ outside the sphere and vanishing inside the sphere.

Note that the exterior field generated by the sphere is exactly the same as the field generated by a point charge Q at the center of the sphere.

One may derive this result by explicit evaluation of the integrals:

$$\vec{E} = \frac{\sigma}{4\pi\varepsilon_o}\int_0^{\pi}\int_0^{2\pi}\frac{(-R\sin\varphi\cos\theta, -R\sin\varphi\sin\theta, r - R\cos\varphi)}{(r^2 - 2rR\cos\varphi + R^2)^{3/2}}\, R^2\sin\varphi d\theta d\varphi.$$

Note that symmetry considerations allow one to only consider the integral giving the radial component of $\vec{E}$ (here the last component). Another option is to determine the electric potential, by evaluating the integral:

$$\varphi = \frac{\sigma}{4\pi\varepsilon_o}\int_0^{\pi}\int_0^{2\pi}\frac{R^2\sin\varphi}{(r^2 - 2rR\cos\varphi + R^2)^{1/2}}\, d\theta d\varphi$$

and yields $\varphi(r) = \frac{Q}{4\pi\varepsilon_o r}$ outside the sphere and $\varphi \equiv \frac{Q}{4\pi\varepsilon_o R} = cst.$ inside the sphere. These are fine exercises to practice integrating techniques, but there are more efficient ways to obtain the result.

A quite efficient method is using the Gauss' law (and the symmetry argument that the field is radial). Then for a sphere of radius r centered with the charged sphere:

$$E(r)4\pi r^2 = \int_{S_r^2}\vec{E}\cdot d\vec{S} = \begin{cases}\frac{Q}{\varepsilon_o} & r > R\\ 0 & r < R\end{cases}.$$

Another method to efficiently find the electric potential is to apply some considerations of harmonic functions. By symmetry, we may consider that $\varphi(r)$ depends only on the distance to the center of the sphere, and by Poisson's equation satisfies $\Delta\varphi = 0$ in the interior and exterior of the sphere. Now the only harmonic functions depending only on r are of the form:

$$a + b/r$$

for some constants a and b. Thus in the interior of the sphere, φ is constant ($1/r$ is not defined over the whole interior of the sphere) and its value may be obtained at the center of the sphere:

$$\varphi_{int} = \frac{1}{4\pi\varepsilon_o}\int_{S_R^2}\frac{\sigma dA}{R} = \frac{Q}{4\pi\varepsilon_o R}.$$

On the exterior of the sphere, the condition that $\varphi(r) \to 0$ as $r \to \infty$ gives $\varphi = b/r$ for some constant b. By continuity with φ_{int}, we have

$$\varphi_{ext} = \frac{Q}{4\pi\varepsilon_o r}.$$

Yet another approach to determine the electric field is by using solid angle and spherical inversion (see figure below).

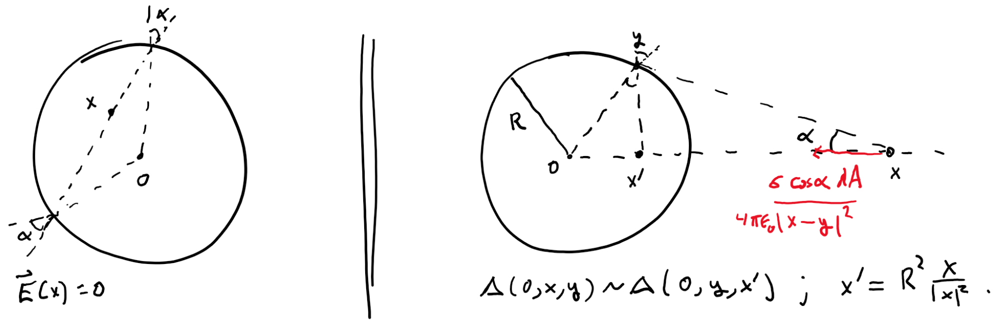

Figure 18. The exterior electric field generated by a uniformly charged sphere may also be computed using a spherical inversion and solid angle. To find the exterior field, one argues by symmetry that one only needs to compute the radial component, given at distance $r = |x|$ by: $E(r) = \frac{\sigma}{4\pi\varepsilon_o}\int_{y\in S^2_R} \frac{\cos\alpha dA}{|x-y|^2} = \frac{\sigma}{4\pi\varepsilon_o}\int_{y\in S^2_R} \frac{|x'-y|^2 d\Omega_{x'}}{|x-y|^2} = \frac{\sigma}{4\pi\varepsilon_o}\int_{y\in S^2_R} \frac{R^2 d\Omega_{x'}}{r^2} = \frac{Q}{4\pi\varepsilon_o r^2}$, where $d\Omega_{x'}$ is the element of solid angle from $x' = R^2\frac{x}{|x|^2}$ a spherical inversion of x (over the charged sphere). Similarly one may find that the interior field vanishes by considering $\frac{4\pi\varepsilon_o}{\sigma}\vec{E}_{int}(x) = \int_{y\in S^2_R} \frac{y-x}{|y-x|^3}\, dA = \int_{y\in S^2_R} \frac{y-x}{|y-x|}\frac{d\Omega_x}{\cos\alpha} = \int_{y'\in S^2_R} \frac{y'-x}{|y'-x|}\frac{d\Omega_x}{\cos\alpha} = -\frac{4\pi\varepsilon_o}{\sigma}\vec{E}_{int}(x)$ where y, y' are the two points on the sphere along the chords through x.

We also remark that similarly one obtains the field produced by a *uniformly charged ball* of radius R with charge density ρ as:

$$E(r) = \begin{cases} \frac{Q}{4\pi\varepsilon_o r^2} & r > R \\ \frac{Qr}{4\pi\varepsilon_o R^3} & r < R \end{cases}$$

with E the radial component of the electric field and $Q = \frac{4}{3}\pi R^3\rho$ the total charge of the ball.

- In general one calls the potentials produced by charge distributions over surfaces *single layer potentials*. For $\sigma : \Sigma \to \mathbb{R}$ the charge density over the surface Σ this single layer potential is then:

$$\varphi(x) = \frac{1}{4\pi\varepsilon_o}\int_{y\in\Sigma} \frac{\sigma(y)\, dA}{|x-y|}.$$

In general, this integral may diverge. Let us state some analytic results in a case when it is defined [1].

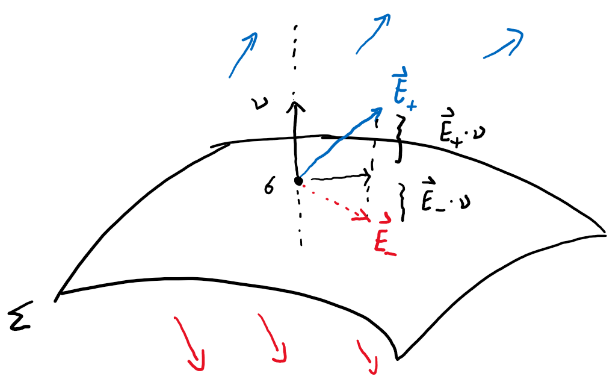

Figure 19. The electric field generated by a single layer potential has a jump in its normal direction proportional to the charge density, $\vec{E}_+ - \vec{E}_- = \frac{\sigma}{\varepsilon_o}\nu$, as one crosses the surface.

For Σ compact and $\sigma : \Sigma \to \mathbb{R}$ smooth, one may show that φ is defined and continuous over $\mathbb{R}^3$ and

[1] See for example Arnold's lectures on pde's section 9.3.

smooth on $\mathbb{R}^3\backslash\Sigma$. Moreover in this case, one has $\nabla\varphi = -\vec{E}$ over $\mathbb{R}^3\backslash\Sigma$ where:

$$\vec{E}(x) = \frac{1}{4\pi\varepsilon_o}\int_{y\in\Sigma}\frac{\sigma(y)(x-y)}{|x-y|^3}\,dA$$

is a smooth vector field over $\mathbb{R}^3\backslash\Sigma$. As φ is in general only continuous over Σ one finds a 'jump' in the direction of $\vec{E}$ over Σ. Namely, if Σ is oriented with unit normal ν, then:

$$\partial_\nu\varphi_- - \partial_\nu\varphi_+ = (\vec{E}_+ - \vec{E}_-)\cdot\nu = \frac{\sigma}{\varepsilon_o}$$

holds over Σ, where $\vec{E}_\pm$ are the directions of $\vec{E}$ from the outer ($+\nu$ side) and inner ($-\nu$ side) sides of Σ. In this case as well the tangential components are equal (in fact smooth on Σ):

$$(\vec{E}_+ - \vec{E}_-)\cdot v = 0$$

holds over Σ for any v tangent to Σ.

- We now find some static fields involving conductors using the *method of images*.

  Consider a conductor Ω consisting of a half space (region below a plane) and a point charge q located at position x_o above the plane (in Ω^c).

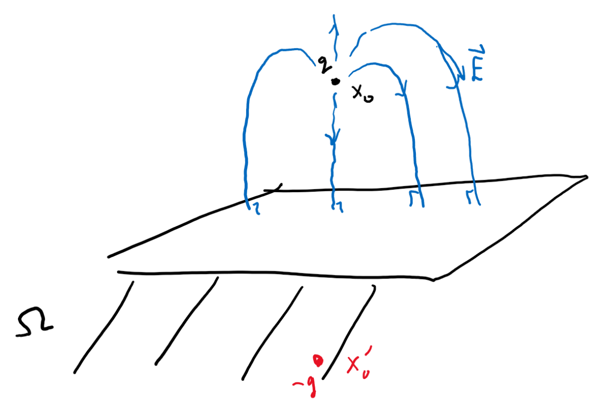

Figure 20. The static equilibrium field generated by a conducting half space and point charge q at x_o may be obtained by superimposing the field generated by q with the field generated by an imaginary charge $-q$ placed at the reflection x_o' of x_o over the plane.

By placing an imaginary 'image charge' $q' = -q$ at the reflection x_o' of x_o over the plane, we produce an electric field in Ω^c perpendicular to $\partial\Omega$, given by the potential:

$$\varphi_o(x) = \frac{q}{4\pi\varepsilon_o}\left(\frac{1}{|x-x_o|} - \frac{1}{|x-x_o'|}\right), \quad x\in\Omega^c$$

with $\varphi_o|_{\partial\Omega}\equiv 0$. Thus the electric field produced by the potential $\varphi_o(x),\ \ x\in\Omega^c,\ \ \varphi_o(x) := 0,\ \ x\in\Omega$ is the static field produced by a (grounded) conducting half space and point charge q at x_o.

Note that letting z be the distance of a point in Ω^c to $\partial\Omega$, we have solutions:

$$\varphi(x) = \varphi_o(x) - \frac{\sigma}{\varepsilon_o}z + V$$

for σ, V constants where $\varphi\equiv V$ in Ω.

For a general distribution of charges with density ρ and potential φ_{amb} in the upper half space Ω^c one may take image charges:

$$\rho'(x') := -\rho(x)$$

in Ω where x' is the reflection of x over the plane $\partial\Omega$ to determine the electrostatic fields:

$$\varphi(x) = \varphi_{amb}(x) + \varphi'_{amb}(x) - \frac{\sigma}{\varepsilon_o} z + V, \quad x \in \Omega^c$$

and $\varphi|_\Omega \equiv V$, where φ'_{amb} is the potential generated by the charges ρ'.

In general the method of images consists in trying to place imaginary charges *inside* [1] Ω, with density ρ' and potential φ' in a way so that the sum, $\varphi + \varphi'$, with the potential φ produced by the exterior ambient charges in Ω^c (with density ρ) is constant over $\partial\Omega$.

The method may also be understood 'in reverse'. Namely, one may take a given distribution of charges ρ with potential φ and place the boundary of a conductor $\partial\Omega$ along an equipotential surface: $\varphi \equiv V = cst$. Then when one side of the conductor is filled, the potential φ on Ω^c and V on Ω gives the equilibrium field produced by this conductor and the ambient charges in Ω^c.

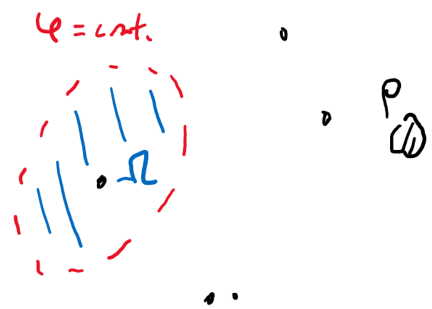

Figure 21. Given a charge distribution ρ with potential φ placing a conductor with boundary along a level set of φ gives an equilibrium configuration (with the same potential φ in Ω^c and the charges of ρ interior to Ω the imaginary 'image' charges).

Now we consider a solid conducting ball Ω, of radius R and a point charge q at position x_o at distance $r_o > R$ from the center of the ball.

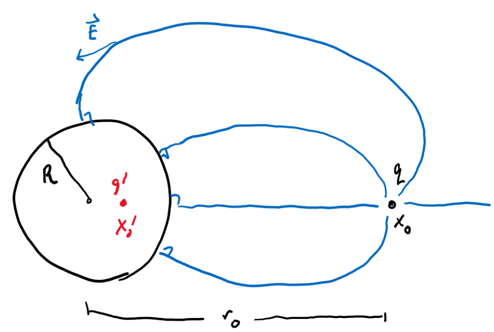

Figure 22. The equilibrium field generated by a (grounded) conducting ball of radius R and point charge q at x_o at distance r_o from the center of the ball may be obtained by superimposing the field generated by q and that generated by an imaginary charge $q' = -\frac{R}{r_o} q$ at the spherical inversion x'_o of x_o (along the ray from the balls center to x_o and at distance R^2/r_o from the center).

Let $\varphi(x) = \frac{q}{4\pi\varepsilon_o} \frac{1}{|x - x_o|}$ be the potential produced by q. Then $\Delta\varphi = 0$ on $\mathbb{R}^3 \backslash \{x_o\}$ and

$$\varphi'(x) := \frac{R}{|x|}\, \varphi(R^2 x/|x|^2)$$

[1] So that $\Delta\varphi' = 0$ over Ω^c.

is harmonic on $\mathbb{R}^3\backslash\{x_o'\}$ where $x_o' = R^2 \frac{x_o}{|x_o|^2}$ is the inversion of x_o over the sphere $\partial\Omega$ of radius R. Moreover:

$$\varphi'|_{\partial\Omega} = \varphi|_{\partial\Omega}$$

so that the potential:

$$\varphi_o(x) := \varphi(x) - \varphi'(x), \quad x \in \Omega^c$$

and $\varphi_o(x) \equiv 0$ in Ω, corresponds to the static electric field produced by the conducting grounded solid ball, Ω, of radius R and point charge q at x_o. Explicitly, one may work out:

$$\varphi_o(x) = \frac{1}{4\pi\varepsilon_o}\left(\frac{q}{|x-x_o|} + \frac{q'}{|x-x_o'|}\right), \quad x \in \Omega^c$$

where $q' := -\frac{R}{r_o}q$. The general solution, corresponding to a net charge Q on Ω is determined by adding to the above result a potential from uniformly charging the sphere to charge $Q - q'$. As we have seen above, exterior to Ω this potential is that of a point charge, so:

$$\varphi_Q(x) = \varphi_o(x) + \frac{1}{4\pi\varepsilon_o}\frac{Q-q'}{|x|}, \quad x \in \Omega^c$$

generates the static field with total charge Q on Ω (and $\varphi_Q|_\Omega = \frac{Q-q'}{4\pi\varepsilon_o R}$).

- A pair of opposite electric point charges, q, $-q$, seperated by a distance ℓ is called a *dipole*. In considering a conducting half space, we have given the expression for the potential generated by a dipole:

$$\varphi(x) = \frac{q}{4\pi\varepsilon_o}\left(\frac{1}{\sqrt{r^2 - r\ell\cos\theta + \ell^2/4}} - \frac{1}{\sqrt{r^2 + r\ell\cos\theta + \ell^2/4}}\right).$$

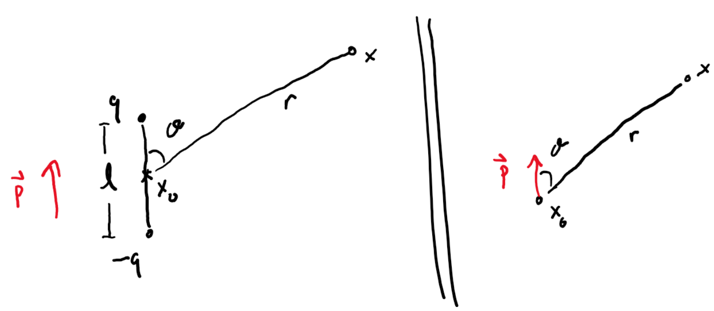

Figure 23. The electric field generated by a pair of opposite point charges at a fixed distance depends on the distance to the midpoint of the charges and an angle to the point. At distances far relative to the seperation between the point charges, one may approximate the dipole by a vector $\vec{p}$ at the point x_o.

At distances far from the dipole relative to the distance between the charges, $\frac{\ell}{r} << 1$, we expand:

$$\varphi(x) = \frac{p}{4\pi\varepsilon_o r^2}\left(\cos\theta + O(\frac{\ell}{r})\right)$$

where $p := \ell q$ is called the *dipole moment*.

Neglecting the higher order terms, we obtain the potential generated by a 'perfect dipole' or 'ideal dipole' (we may at times just call this a dipole when there is no risk of confusion) :

$$\varphi(x) := \frac{1}{4\pi\varepsilon_o}\frac{\vec{p}\cdot(x-x_o)}{|x-x_o|^3} = \frac{1}{4\pi\varepsilon_o}\vec{p}\cdot\nabla^o\left(\frac{1}{|x-x_o|}\right)$$

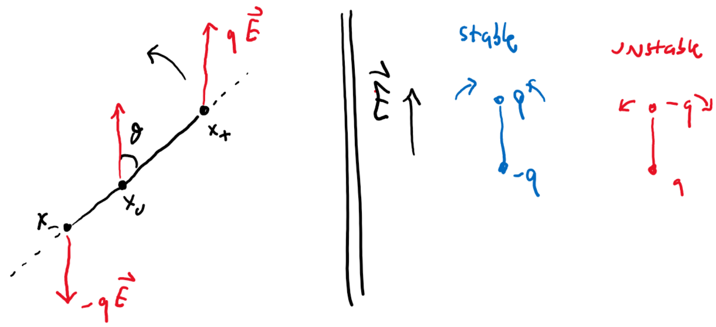

Figure 24. The dynamics of a dipole subject to an ambient electric field $\vec{E}$ tends to align the dipole with the electric field. The total force on the dipole is $q(\vec{E}(x_+) - \vec{E}(x_-))$ and the total torque about x_o is $(x_+ - x_0) \times q\vec{E}(x_+) - (x_- - x_o) \times q\vec{E}(x_-)$.

where x_o represents the location of the dipole (the location of the two 'infinitesimally close' charges) and the vector $\vec{p}$ directs the dipole (a vector from $-q$ to q).

One may consider the motion (dynamics) of a 'test' dipole under the influence of an ambient electric field, $\vec{E}$. First consider planar motions in a constant ambient electric field of strength E. Then:

$$\ddot{x}_o = 0, \quad \ddot{\theta} = -\frac{qME}{m_+ m_- \ell} \sin\theta$$

where $M = m_+ + m_-$ is the total mass, $Mx_o = m_+ x_+ + m_- x_-$ is the center of mass of the two charges and θ is the angle between $\vec{E}$ and the vector from $-q$ to q. In the general case one has:

$$M\ddot{x}_o = q\left(\vec{E}(x_+) - \vec{E}(x_-)\right), \quad v \times \ddot{v} = q\ v \times \left(\frac{1}{m_+}\vec{E}(x_+) + \frac{1}{m_-}\vec{E}(x_-)\right)$$

where x_o is the center of mass of the dipole and $v = x_+ - x_-$ the vector (of fixed length) from $-q$ to q. Letting $qv = \varepsilon\vec{p}$ and reparametrizing by $\tau := \sqrt{\varepsilon}t$ (with $\frac{df}{d\tau} = f'$) one has the expansion:

$$Mx_o'' = \varepsilon^2\ d_{x_o}\vec{E}(\vec{p}) + O(\varepsilon^3), \quad \vec{p} \times \vec{p}\,'' = \mu\ \vec{p} \times \vec{E}(x_o) + O(\varepsilon)$$

where $\mu := \frac{Mq^2}{m_+ m_-}$. The potential energy of the dipole in the electric field $\vec{E} = -\nabla\varphi$ is:

$$W = q\left(\varphi(x_o + m_- v/M) - \varphi(x_o - m_+ v/M)\right) = -\vec{p} \cdot \vec{E}(x_o) + O(\varepsilon^2)$$

with $qv = \varepsilon\vec{p}$. We remark as well that, for $(x_o, \vec{p}\,) \in \mathbb{R}^3 \times S^2$, these equations of motion may be written in Lagrangian (or as well Hamiltonian) form:

$$L = \frac{M|x_o'|^2}{2} + \frac{\varepsilon^2|\vec{p}\,'|^2}{2\mu} + \varepsilon^2\ \vec{p} \cdot \vec{E}(x_o) + O(\varepsilon^3).$$

- We consider some examples involving capacitance.

  First, a conducting solid ball of radius R, when given charge Q generates an exterior field ($r \geq R$) with potential, $\varphi = \frac{Q}{4\pi\varepsilon_o r}$, so the potential value over the ball is given by: $V = \varphi|_{r=R} = \frac{Q}{4\pi\varepsilon_o R}$. The capacitance is thus:

$$C = 4\pi\varepsilon_o R.$$

  Next, we determine the capacitance of a capacitor consisting of two concentric spherical conductors of radii $a < b$. When charged to Q and $-Q$, the outer conducting sphere generates zero electric field in its interior, while the inner (negatively charged) conducting sphere generates the radial electric field with potential $\varphi = -\frac{Q}{4\pi\varepsilon_o r}$ on the annular region between the spheres. The potential difference is then:

$$V = V_+ - V_- = V_b - V_a = \frac{Q}{4\pi\varepsilon_o}\left(\frac{1}{a} - \frac{1}{b}\right)$$

so that the capacitance is:

$$C = \frac{4\pi\varepsilon_o}{1/a - 1/b}.$$

An important and useful approximate situation is the *parallel plate capacitor*. Consider two congruent planar 'plates' or 'disks' of areas A and seperated by a relatively small distance $d << A$.

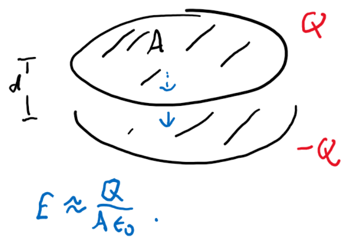

Figure 25. The electric field between two nearby charged plates is, away from the edges, well approximated by the uniform field generated between two uniformly charged planes (with charge densities $\pm\sigma = \pm Q/A$).

When the two plates are charged to Q and $-Q$ they approximately generate a uniform electric field in the region between the plates (away from the edges of the disk, the situation is nearly that of a uniformly charged infinite plane). The strength of this uniform field is then: $\frac{\sigma}{\varepsilon_o} = \frac{Q}{A\varepsilon_o}$, with potential difference: $V = \frac{Qd}{A\varepsilon_o}$, and capacitance:

$$C = \frac{A\varepsilon_o}{d}.$$

- We consider some examples involving dielectrics (that are homogeneous and isotropic).

  First, consider a solid ball of radius R dielectric with permittivity ε and a point charge q at its center. By symmetry, $\vec{D}$ is radial, and from $\nabla \cdot \vec{D} = q\delta_o$, we have:

$$\vec{D} = \frac{q}{4\pi}\frac{x}{|x|^3}$$

  where x is the position from the center of the ball. Hence:

$$\vec{E} = \frac{q}{4\pi}\begin{cases} \frac{x}{\varepsilon|x|^3} & |x| < R \\ \frac{x}{\varepsilon_o|x|^3} & |x| > R \end{cases}.$$

  As a slightly more general situation, we may consider a dielectric Ω with permittivity ε in the presence of ambient charges ρ_o. Let us describe the bound charge density. We compute:

$$-\rho_b = \nabla \cdot \vec{P} = \chi\varepsilon_o\nabla \cdot \vec{E} = \chi(\rho_o + \rho_b)$$

  over Ω, so that:

$$\rho_b = -\frac{\chi}{1+\chi}\rho_o = -(1 - \frac{\varepsilon_o}{\varepsilon})\rho_o.$$

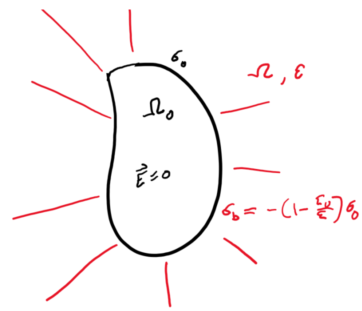

Figure 26. A charged conductor, Ω_o, with surface charge distribution σ_o, may be surrounded by a dielectric. The response of the dielectric accumulates bound charges σ_b over their interface (common boundary).

over Ω. Note that when no ambient charges are contained in Ω (that is $\rho_o|_\Omega \equiv 0$), then $\rho_b \equiv 0$. In this case, like a conductor, the dielectric produces a field only depending upon the accumulation of bound charges, σ_b, on its boundary.

In a similar vein, we may consider a dielectric, Ω, with permittivity ε, bordering a conductor Ω_o. Suppose that the conductor has been given a net charge, producing the field due to charges σ_o on its boundary. We view these as the ambient charges and introduce the dielectric Ω. The total field, $\vec{E}$, produced will vanish on the interior of the conductor and be given by $\frac{1}{\varepsilon}\vec{D}$ on the interior of the dielectric. From the boundary conditions:

$$\sigma_o = \vec{D}_+ \cdot \nu, \quad \sigma_o + \sigma_b = \varepsilon_o \vec{E}_+ \cdot \nu$$

we obtain:

$$\sigma_b = -\frac{\chi}{1+\chi}\sigma_o = -(1-\frac{\varepsilon_o}{\varepsilon})\sigma_o$$

for the distribution of bound surface charges due to the dielectric. Note that $\sigma_o + \sigma_b = \frac{\varepsilon_o}{\varepsilon}\sigma_o$, so by uniqueness of exterior fields to a conductor, we have (when the dielectric fills all of Ω_o^c) that:

$$\vec{E} = \frac{\varepsilon_o}{\varepsilon}\vec{E}_o$$

where $\vec{E}_o$ is the exterior electric field generated by the charged conductor in vacuum.

Lastly, let us consider two dielectric half spaces: $\Omega_1 = \{z > 0\}, \Omega_2 = \{z < 0\}$ with permittivities $\varepsilon_1, \varepsilon_2$. Suppose a point charge q is placed in Ω_1 at position x_o. We may apply a similar method of images to guess that the resulting electric field has a potential given piecewise by:

$$\varphi_1 = \frac{1}{4\pi\varepsilon_1}\left(\frac{q}{|x-x_o|} + \frac{q'}{|x-x_o'|}\right), \quad \text{in } \Omega_1$$

$$\varphi_2 = \frac{1}{4\pi\varepsilon_1}\left(\frac{q}{|x-x_o|} + \frac{q''}{|x-x_o''|}\right), \quad \text{in } \Omega_2$$

where $x_o' \in \Omega_2$ and $x_o'' \in \Omega_1$ lie along the perpendicular from x_o to the interface $z = 0$ plane.

These satisfy $\Delta\varphi_1 = -q\delta_{x_o}/\varepsilon_1$ in Ω_1 and $\Delta\varphi_2 = 0$ in Ω_2, so it remains to determine q', q'', x_o', x_o'' in order to satisfy the boundary conditions.

By continuity of φ:

$$\varphi_1|_{z=0} = \varphi_2|_{z=0}$$

one takes $q' = q''$ and x_o' the reflection of x_o'' over the interface.

By $D_+ \cdot \nu = D_- \cdot \nu$, we have the condition $\varepsilon_1 \partial_z \varphi_1|_{z=0} = \varepsilon_1 \partial_\nu \varphi_1 = \varepsilon_2 \partial_\nu \varphi_2 = \varepsilon_2 \partial_z \varphi_2|_{z=0}$, so take:

$$x_o'' = x_o, \quad q'' = q' = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2} q.$$

By uniqueness, the piecewise potential with these choices of x'_o, x''_o, q', q'' is the potential generating the resulting electric field.

Examples of magnetostatic fields:

- We consider the magnetic field generated by a uniform current, I, along a conducting wire, say running along the z-axis. Explicitly, one may use Biot-Savart to consider the integral:

$$\frac{\mu_o I}{4\pi}\int_{-\infty}^{\infty}\frac{(-y,x,0))}{(x^2+y^2+(z-z')^2)^{3/2}}\;dz'$$

  to determine the magnetic field at (x,y,z).

  Alternately, one may proceed using the integral theorems. Observe first that the magnetic field at a given point is perpendicular to the plane containing the wire and this point (from Biot-Savart) and by symmetry depends only on the distance from the point to the wire. Considering the line integral of the magnetic field along a circle of radius r around the wire we have by Ampere's law that:

$$2\pi r B(r) = \mu_o I \Rightarrow B(r) = \frac{\mu_o I}{2\pi r}$$

  where $B(r)$ is the strength of the magnetic field (see figure).

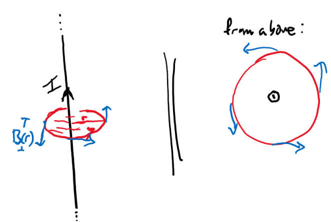

Figure 27. A steady and constant current flowing along an (infinite) straight line produces a rotational magnetic field around the wires axis whose strength varies inversely to the distance from the wire.

- We consider the magnetic field generated by a uniform current flowing along an infinite conducting plane. Let $\vec{K} = cst.$ be the surface current density. If we take the plane as the xy-plane and $\vec{K} = K\hat{i}$ along the x-axis, then explicitly one may evaluate the integrals:

$$\frac{\mu_o K}{4\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\hat{i}\times(x-x',y-y',z)}{((x-x')^2+(y-y')^2+z^2)^{3/2}}\;dx'dy'$$

  to determine the magnetic field at the point (x,y,z).

  Alternately, one may proceed using the integral theorems and some symmetry arguments. By translational symmetry parallel to the plane, $\vec{B}(z)$, will only depend on the height, z, above the plane. Moreover it is perpendicular to $\vec{K}$ from Biot-Savart. As well, under a reflection over the plane, the cross products change signs so that: $\vec{B}(-z) = -\vec{B}(z)$.

  We will first argue that $\vec{B}$ is parallel to the plane, ie has no vertical component. Since $\nabla\cdot\vec{B} = 0$, we consider an arbitrary cube, with two sides parallel to $\vec{K}$. Then the flux of $\vec{B}$ through this cube is zero. If the two 'horizontal' sides of the cube are at heights z_1, z_2 then this flux is:

$$(B_{ver}(z_1) - B_{ver}(z_2))\,A = 0$$

  where A is the area of the side of the cube and $B_{ver}(z)$ is the vertical component of $\vec{B}$ at height z. Thus B_{ver} is constant and by the reflectional symmetry, $B_{ver}(-z) = -B_{ver}(z) \Rightarrow B_{ver} \equiv 0$

To determine the magnitude, B, of $\vec{B}$ (parallel to the plane and perpendicular to $\vec{K}$), we integrate $\vec{B}$ along a rectangular loop perpendicular to $\vec{K}$ and with horizontal sides at height $z, -z$ to obtain by Ampere's law:

$$2\ell B(z) = \mu_o K \ell \Rightarrow B = \frac{\mu_o K}{2} = cst.$$

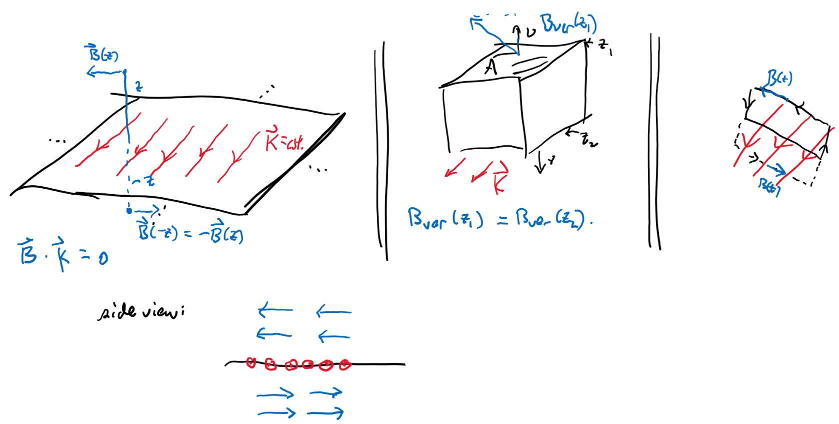

Figure 28. A steady and constant current flowing along an (infinite) plane produces a magnetic field of constant strength outside of the plane.

- Consider a surface current density $\vec{K}$ on a (infinite) cylinder of radius r_o, where the vector field $\vec{K}$ has constant norm, K, and is at a constant angle, α, from the horizontal longitude to the cylinder. The magnetic field generated by these currents is called a *solenoid* field. In practice, it provides a good approximation to the field generated by (tightly) wrapping a current carrying wire around a cylinder.

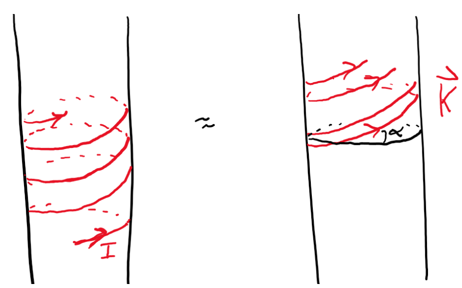

Figure 29. A *solenoid* consists of a tightly wound current carrying wire (approximately a constant surface current density $\vec{K}$) around a cylinder of radius r_o.

The magnetic field of the solenoid may be determined by the following symmetry considerations. We have symmetry by translations and rotations along the axis of the cylinder, so that the magnetic field strength, $B(r)$, depends only on the distance r to the cylinder's axis and the vector field $\vec{B}$ is invariant under these translations and rotations.

One may determine $\vec{B}$ as follows. First considering a cylinder, Σ_r, of radius r (and height h) with the same axis as the solenoid, we have from $\nabla \cdot \vec{B} = 0$, that:

$$0 = \int_{\Sigma_r} \vec{B} \cdot d\vec{S} = 2\pi r h B_{rad}(r) \Rightarrow B_{rad}(r) = 0$$

where B_{rad} is the radial component of $\vec{B}$ (note that the surface integrals of $\vec{B}$ over the top and bottom 'caps' of Σ_r cancel by the translational invariance of $\vec{B}$). Next, consider a circle, C_r, of radius r centered

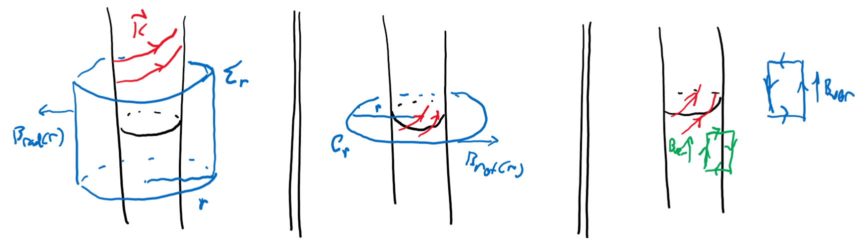

Figure 30. The components of the magnetic field generated by a solenoid may be determined by applying the integral theorems ($\nabla \cdot \vec{B} = 0$, and Ampere's law) to suitable curves and surfaces.

along the solenoids axis we have by Ampere's law :

$$0 = \oint_{C_r} \vec{B} \cdot d\vec{s} = 2\pi r B_{rot}(r) \Rightarrow B_{rot}(r) = 0, \quad r < r_o$$

$$\mu_o K \sin\alpha \; 2\pi r_o = \oint_{C_r} \vec{B} \cdot d\vec{s} = 2\pi r B_{rot}(r) \Rightarrow B_{rot}(r) = \frac{\mu_o K \sin\alpha \; r_o}{r}, \quad r > r_o$$

where r_o is the radius of the solenoid cylinder and B_{rot} is the rotational component component of $\vec{B}$. For the vertical component, B_{ver}, of $\vec{B}$ (directed along the axis of the solenoid), one may apply Ampere's law to rectangular loops parallel to the solenoids axis to obtain that B_{ver} has a constant value outside the solenoid and a (possibly different) constant value inside the solenoid. Outside the solenoid, we have $B_{ver} = 0$, since the magnetic field goes to zero as the distance to the currents goes to infinity. Finally, again by Ampere's law, inside the solenoid:

$$B_{ver} = \mu_o K \cos\alpha.$$

A solenoid is the basis for producing electromagnets, or examining magnetization of materials. In practice one winds a wire carrying current I around some finite cylinder to produce an approximately constant magnetic field inside the cylinder, with $B_{ver} = \mu_o n I \cos\alpha$ where n are the number of turns of the wire per unit length and α the 'pitch' of the winding. Often when the wire is very tightly (nearly horizontally) wound, α is small, and one further approximates by $\alpha = 0$. Note that with a relatively small current, I, one may still produce a strong magnetic field by making many turns (n large).

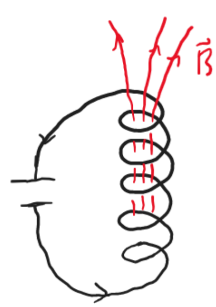

Figure 31. A current carrying wire wrapped around a cylinder in a helical shape produces an approximately constant solenoidal magnetic field in its interior and near its ends. It is the same type of magnetic field produced by a cylindrical magnet.

- A distribution of current on a surface produces what is called a single-layer magnetic potential. For $\vec{K}$ the surface density of current (tangent to the surface Σ), that is:

$$\vec{A}(x) = \frac{\mu_o}{4\pi} \int_{y \in \Sigma} \frac{\vec{K}(y)\; dA}{|x - y|}.$$

When Σ is compact and $\vec{K}$ is a smooth vector field on the surface, one may show that $\vec{A}$ is defined and continuous over $\mathbb{R}^3$ and smooth on $\mathbb{R}^3 \backslash \Sigma$. The magnetic field, $\vec{B} = \nabla \times \vec{A}$, is then defined and smooth over $\mathbb{R}^3 \backslash \Sigma$ and in general has a 'jump' in its direction when crossing Σ.

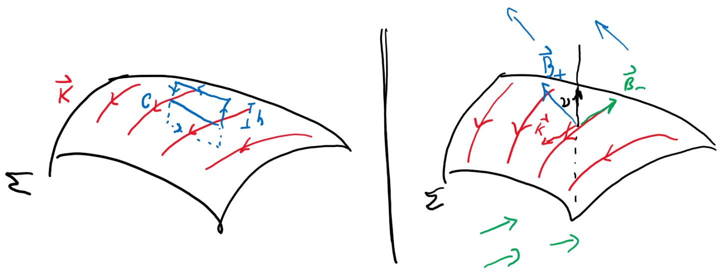

Figure 32. The magnetic field $\vec{B}$ produced by a surface density $\vec{K}$ on a surface Σ has a 'jump' over Σ in its tangential direction proportional to the current density. This 'jump' may be computed by considering a rectangle of small 'height' h over a curve $\mathcal{C}$ in the surface and applying Ampere's law to obtain, when $h \to 0$, that $\int_{\mathcal{C}} (\vec{B}_+ - \vec{B}_-) \cdot d\vec{s} = \mu_o \int_{\mathcal{C}} \vec{K} \cdot (\nu \times d\vec{s})$. Since $\mathcal{C}$ was arbitrary, one has $(\vec{B}_+ - \vec{B}_-) \cdot T = \mu_o (\vec{K} \times \nu) \cdot T$ for any T tangent to the surface. The same argument we used for surface densities of charges and with no magnetic monopoles, $\nabla \cdot \vec{B} = 0$, gives that $(\vec{B}_+ - \vec{B}_-) \cdot \nu = 0$.

Namely, if ν is a unit normal to the surface, then:

$$\vec{B}_+ - \vec{B}_- = \mu_o \vec{K} \times \nu$$

holds over Σ, where $\vec{B}_\pm$ are the direction of $\vec{B}$ from the outer ($+\nu$ side) and inner ($-\nu$ side) of Σ. Note that this relation contains $(\vec{B}_+ - \vec{B}_-) \cdot \nu = 0$, ie that the normal components are equal (in fact smooth) over Σ.

- We will introduce a *magnetic dipole* by computing a Taylor expansion of the field produced 'far' from a current loop. Let $\mathcal{C} = \partial\Sigma$ be a closed curve along which flows a constant steady current I. This current loop produces a magnetic potential:

$$\vec{A}(x) = \frac{\mu_o I}{4\pi} \oint_{y \in \mathcal{C}} \frac{d\vec{s}}{|x - y|}.$$

Fix an origin x_o, and set $\vec{x} := x - x_o, \vec{y} := y - y_o$. Then for $|\vec{y}|/|\vec{x}| << 1$, we have by Taylor expansion:

$$\vec{A}(x) = \frac{\mu_o I}{4\pi|\vec{x}|} \oint_{y \in \mathcal{C}} \left(1 + \frac{\vec{x} \cdot \vec{y}}{|\vec{x}|^2} + O(\frac{|\vec{y}|^2}{|\vec{x}|^2}) \right) \, d\vec{s}.$$

The first term integrates to zero, while for the second, one may take a surface Σ with $\partial\Sigma = \mathcal{C}$ and apply the integral theorem, $\int_{\partial\Sigma} f \; d\vec{s} = \int_\Sigma \nu \times \nabla f \; dA$, to obtain:

$$\vec{A}(x) = \frac{\mu_o}{4\pi} \frac{\vec{\mu} \times (x - x_o)}{|x - x_o|^3} + O\left(\frac{1}{|x - x_o|^3} \right)$$

where $\vec{\mu} := I \int_\Sigma \nu \; dA$ for ν unit normal to Σ with $\partial\Sigma = \mathcal{C}$. The vector $\vec{\mu}$ is called the *magnetic dipole moment* of the current loop $\mathcal{C}$.

When the loop $\mathcal{C}$ is a planar curve, one may take Σ as a surface in the plane so that:

$$\vec{\mu} = IAn$$

where A is the area enclosed by the current loop and n the unit normal to the plane containing the current loop.

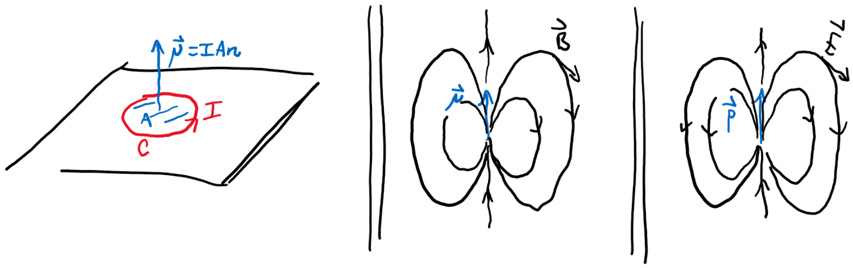

Figure 33. Taylor expansion of the magnetic field 'far' from a loop of current gives the field of an (ideal) magnetic dipole. It is the same type of field produced 'far' from a pair of opposite charges (an ideal electric dipole).

An 'ideal' or 'perfect' magnetic dipole corresponds to neglecting the higher order terms in this expansion. It is determined by a (fixed) dipole moment $\vec{\mu}$ and generates the magnetic potential:

$$\vec{A}(x) = \frac{\mu_o}{4\pi}\, \vec{\mu} \times \nabla^o \left(\frac{1}{|x - x_o|} \right).$$

One computes that the magnetic field generated by this (ideal) magnetic dipole is then:

$$\vec{B} = \nabla \times \vec{A} = -\nabla \varphi_m$$

where $\varphi_m(x) := \frac{\mu_o}{4\pi}\, \vec{\mu} \cdot \nabla^o \left(\frac{1}{|x-x_o|} \right)$. Comparing to the electric field generated by an (electric) dipole with dipole moment $\vec{p}$, we see this magnetic dipole field has exactly the same form with $\vec{\mu}$ replacing $\vec{p}$.

As with electric dipoles, a magnetic dipole in a constant ambient magnetic field will oscillate around allignment with the magnetic field, with $\vec{\mu}||\vec{B}$ being a stable equilibrium. This observation is the basis for designing electrical motors.

Far-field (multipole) expansions:

...

§6 Maxwell's equations

Now, we consider electric and magnetic fields in the general situation. Namely, we are in a fixed inertial frame in which there may be a given collection of charges moving in some way. What fields do such charges produce? The fields produced satisfy *Maxwell's equations*:

$$\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_o} \ , \quad \nabla \times \vec{E} = -\partial_t \vec{B} \ ,$$

$$\nabla \cdot \vec{B} = 0 \ , \quad \nabla \times \vec{B} = \mu_o \left(\vec{J} + \varepsilon_o \partial_t \vec{E} \right) \ .$$

Here $\rho(x,t)$ and $\vec{J}(x,t)$ are the charge and current densities (the 'sources') producing the fields $\vec{E}(x,t)$, $\vec{B}(x,t)$. Moreover, the effect such fields have on a test charge q moving with velocity $\vec{v}$ is via the *Lorentz force* :

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}).$$

We see, from our work in statics, there are two new terms to account for.

Experimental result	Expression (integral form)
Statics	
Coulomb's law	$\int_{\partial\Omega} \vec{E} \cdot d\vec{S} = \frac{Q_{int}}{\varepsilon_o}$, $\oint_C \vec{E} \cdot d\vec{s} = 0$
Oersted, Biot-Savart law	$\oint_\Sigma \vec{B} \cdot d\vec{S} = 0$, $\int_{\partial\Sigma} \vec{B} \cdot d\vec{s} = \mu_o I_\Sigma$
Dynamics	
induced currents (Faraday-Lenz law)	$\int_{\partial\Sigma} \vec{E} \cdot d\vec{s} = -\int_\Sigma \partial_t B \cdot d\vec{S}$
electromagnetic waves (Maxwell's correction)	$\int_{\partial\Sigma} \vec{B} \cdot d\vec{s} = \mu_o I_\Sigma + \mu_o \varepsilon_o \int_\Sigma \partial_t E \cdot d\vec{S}$

Table 5. The full Maxwell's equations have two new terms compared to their form in the static case. The (red) term involving time variations in the magnetic field expresses Faraday's experiments on induced currents, while the (blue) term involving variations in the electric field is due to Maxwell who included it on theoretical grounds (by the 'tip of his pen'). As we will see, this correction of Maxwell to Ampere's law predicts the existence of electromagnetic waves, and was later given experimental justification via the first detection of electromagnetic waves (by Hertz).

We will first consider the term involving time varying magnetic fields ($\partial_t \vec{B}$). The physical observation underlying this term (due to Faraday) is:

An electrical current is induced in a loop when

the flux of the magnetic field through the loop is varying in time.

Let us consider this observation in the following setting: a (rectangular) conducting loop moves over a half plane with constant magnetic field B perpendicular to (into) the half plane.

In this situation we can understand a current appearing in the conducting loop as it is moved without any new principles (than those we have developed in statics) since moving the current loops produces –by Lorentz– a force $f = qvB$ on the charges in the wire causing them to move and generate a current.

Alternately, let us consider that we fix the conducting loop and move the half plane. The same current is still observed appearing in the wire (this is intuitive since the wire shouldn't 'know' if it is moving over the half plane or the half plane is moving underneath it). However now our previous static principles do not suffice to explain the appearance of this current. The charges in the conducting loop are assumed to

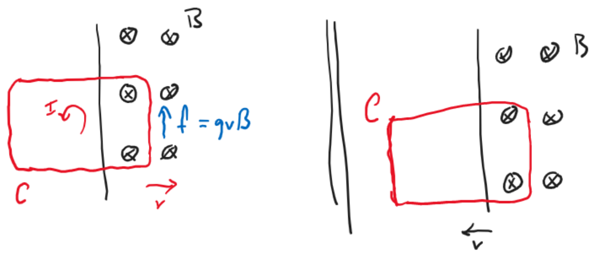

Figure 34. When a conducting loop moves over a magnetic half plane (into the page), a current is induced in the loop. Similarly, when a magnetic half plane is moved under a stationary conducting loop, a current is induced in the loop.

be initially at rest. A magnetic field produces no forces on stationary charges. What then is causing the current?

Something is thus missing from our static theory. Faraday's observations explain this situation by the flux of the magnetic field through the loop changing as we move the half plane. Moreover, by the Lorentz force law, the force that acts on stationary charges is an electric field. So, equivalently:

A time varying magnetic field produces an electric field.

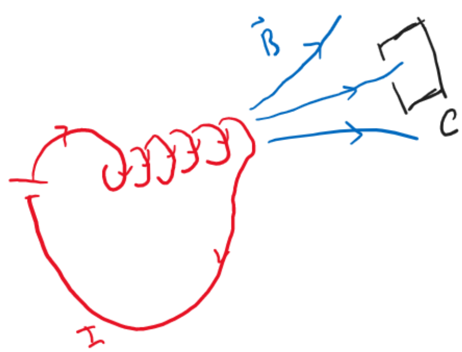

Figure 35. One may observe induced currents in a variety of situations. They may be the result of essentially two ways to change magnetic flux through $\mathcal{C}$. One may vary the magnetic field (eg changing the current I in the above figure), or one may move a conducting loop around in a given magnetic field (keep I and B fixed and move the loop $\mathcal{C}$).

The currents produced by this phenomenon are called *induced currents*, and as a summary of observation in various settings, are described in the following relation:

$$(*) \quad \mathcal{E} = -\frac{d\Phi}{dt}.$$

Here $\mathcal{E}$ is the *electromotive force* [1] on a conducting loop, and Φ is the flux of the magnetic field through the conducting loop. We may rearrange $(*)$ as the:

Faraday-Lenz law. Electromagnetic fields satisfy $(*)$ for any conducting loop iff they satisfy

$$\nabla \times \vec{E} = -\partial_t \vec{B}.$$

[1] See §11. This is really a work (per unit charge) defined by: $\mathcal{E} := \oint_{\mathcal{C}} \frac{\vec{f}_{em} \cdot d\vec{s}}{q}$ where $\vec{f}_{em}$ are the forces on a charge q (eg the chemical forces present in a battery moving charge from terminal to terminal).

proof: Let $\vec{E}(x,t), \vec{B}(x,t)$ be electromagnetic fields satisfying $(*)$ for any conducting loop (possibly deforming in time). Let $\mathcal{C}_t$ be the position of a moving conducting loop at time t. By Lorentz force, the emf on this current loop (at time t) is:

$$\mathcal{E} = \oint_{\mathcal{C}_t} (\vec{E} + \vec{v} \times \vec{B}) \cdot d\vec{s}$$

where $v(x_t) = \frac{d}{dt} x_t, \quad x_t \in \mathcal{C}_t$, are the velocities of the points on the conducting loop. On the other hand,

$$\Phi(t) = \int_{\Sigma_t} \vec{B} \cdot d\vec{S}$$

where $\partial\Sigma_t = \mathcal{C}_t$.

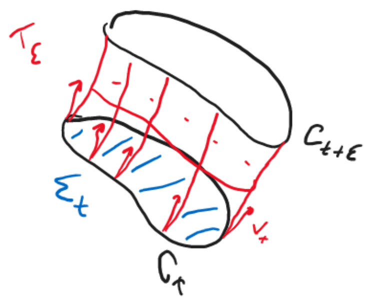

Figure 36. The time change in the flux of the magnetic field may be computed over a (moving) loop by considering the surface $\Sigma_{t+\varepsilon} = \Sigma_t + T_\varepsilon$ with $\partial\Sigma_{t+\varepsilon} = \mathcal{C}_{t+\varepsilon}$ and $\partial\Sigma_t = \mathcal{C}_t$, and where T_ε is the 'tube' formed by the positions over the conducting loop from time t to times $t+\varepsilon$.

To compute $\frac{d\Phi}{dt}$, set $\vec{B}_t(x) := \vec{B}(x,t)$, then:

$$\Phi(t+\varepsilon) - \Phi(t) = \int_{\Sigma_t} (\vec{B}_{t+\varepsilon} - \vec{B}_t) \cdot d\vec{S} + \int_{T_\varepsilon} \vec{B}_{t+\varepsilon} \cdot d\vec{S}$$

where the surface T_ε is that swept out by the moving loops $\mathcal{C}_s, \quad s \in [t, t+\varepsilon]$. Expanding in ε, we have:

$$\Phi(t+\varepsilon) - \Phi(t) = \varepsilon \left(\int_{\Sigma_t} \partial_t \vec{B}_t \cdot d\vec{S} - \oint_{\mathcal{C}_t} (\vec{v} \times \vec{B}_t) \cdot d\vec{s} \right) + O(\varepsilon^2),$$

or,

$$-\frac{d\Phi}{dt} = -\int_{\Sigma_t} \partial_t \vec{B} \cdot d\vec{S} + \oint_{\mathcal{C}_t} (\vec{v} \times \vec{B}) \cdot d\vec{s}.$$

Thus $(*)$ holds iff

$$\oint_{\mathcal{C}_t} \vec{E} \cdot d\vec{s} = -\int_{\Sigma_t} \partial_t \vec{B} \cdot d\vec{S}$$

for any curve, $\mathcal{C}_t = \partial\Sigma_t$ which is exactly the integral form of $\nabla \times \vec{E} = -\partial_t \vec{B}$. □

The term involving time variations in the electric field ($\partial_t \vec{E}$), was introduced in the 'opposite way' to the terms we have introduced to this point. Namely so far the equations have been derived to summarize experimental results. This new term was introduced –by Maxwell in 1864– before its effects had been observed in experiments. The effects of this term are difficult to observe experimentally essentially because creating a time varying electric field requires moving charges around, which in turn produce currents and the resulting current term, $\vec{J}$, is in general much larger than the term due to $\partial_t \vec{E}$.

A notable consequence of Maxwell's equations is the prediction of electromagnetic waves. These waves where measured by Hertz in 1887.

Maxwell's correction to Ampere's law. Electromagnetic fields satisfying Maxwell's equations, with

$$\nabla \times \vec{B} = \mu_o(\vec{J} + \varepsilon_o \partial_t \vec{E})$$

have conservation of charge $\nabla \cdot \vec{J} = -\partial_t \rho$.

proof: If we consider Ampere's law in the non-static case, then taking the divergence of both sides, and using charge conservation and Gauss law gives:

$$0 = \nabla \cdot (\nabla \times \vec{B}) = \mu_o(\nabla \cdot \vec{J}) = -\mu_o \partial_t \rho = -\mu_o \varepsilon_o \nabla \cdot \partial_t \vec{E},$$

which, if the charges are moving in time ($\partial_t \rho \neq 0$), is a contradiction. So if we are to have conservation of charge, then something is missing. Write $\nabla \times \vec{B} = \mu_o(\vec{J} + \vec{X})$, where $\vec{X}$ is the correction. Taking the divergence of this modified Ampere law gives the condition:

$$\varepsilon_o \nabla \cdot \partial_t \vec{E} = \nabla \cdot \vec{X}$$

so that $\vec{X} = \varepsilon_o \partial_t \vec{E} + \nabla \times \vec{Y}$ for some vector field $\vec{Y}$. Such corrections all satisfy conservation of charge, and in Maxwell's equations we take $\vec{Y} = 0$. $\square$

Note that Maxwell's correction to Ampere's law is not *equivalent* to conservation of charge. To justify it has this form (taking $\vec{Y} = 0$ in our previous computation) requires testing in physical experiments.

The basis for this experimental confirmation lies in the (non-homogeneous) *wave equation*, $c^2 \Delta f - f_{tt} = g$, which plays an analogous role that the Poisson equation for potentials played in statics. We study the wave equation in the next section.

Potentials. There is a choice of vector and scalar potential, $\vec{A}, \varphi$:

$$\vec{B} = \nabla \times \vec{A}, \quad \vec{E} + \partial_t \vec{A} = -\nabla \varphi$$

s.t. Maxwell's equations are satisfied iff:

$$\Delta \varphi - \frac{1}{c^2} \partial_t^2 \varphi = -\frac{\rho}{\varepsilon_o}, \quad \Delta \vec{A} - \frac{1}{c^2} \partial_t^2 \vec{A} = -\mu_o \vec{J}.$$

where $c = \frac{1}{\sqrt{\mu_o \varepsilon_o}} = 2.998 \times 10^8 \ m/s$ (coincides with the speed of light!).

proof: From $\nabla \cdot \vec{B} = 0$, we have a vector potential, $\vec{A}'$ with $\nabla \times \vec{A}' = \vec{B}$. All other options for vector potential are of the form $\vec{A}' + \nabla f$ for some function f. The Faraday-Lenz law reads:

$$\nabla \times (\vec{E} + \vec{A}'_t) = 0$$

so that $\vec{E} + \vec{A}'_t = -\nabla \varphi'$ for some function φ'. Gauss law and Maxwell's modification of Ampere's law are satisfied iff:

$$\Delta \varphi' + \nabla \cdot \vec{A}'_t = -\frac{\rho}{\varepsilon_o},$$

$$\Delta \vec{A}' - \varepsilon_o \mu_o \vec{A}'_{tt} = -\mu_o \vec{J} + \nabla \left(\nabla \cdot \vec{A}' + \mu_o \varepsilon_o \varphi'_t \right).$$

We would like to choose our potentials so that $\nabla \cdot \vec{A}' + \mu_o \varepsilon_o \varphi'_t = 0$.
Consider a potential $\vec{A} = \vec{A}' + \nabla f$, with scalar potential, $\varphi = \varphi' - f_t$. Taking f to be any solution to $\Delta f - \varepsilon_o \mu_o f_{tt} = -\mu_o \varepsilon_o \varphi'_t - \nabla \cdot \vec{A}'$ (here the right hand side is given, and we will see later that solutions exist), then we have:

$$\nabla \cdot \vec{A} = -\mu_o \varepsilon_o \varphi_t$$

so that the Gauss law and modified Ampere law take the form of wave equations. $\square$

Materials	
$\vec{D} = \varepsilon_o \vec{E} + \vec{P}$	$\vec{H} = \frac{1}{\mu_o}\vec{B} - \vec{M}$
$\nabla \cdot \vec{D} = \rho_o$	$\nabla \times \vec{H} = \vec{J}_o + \partial_t \vec{D}$
$\nabla \times \vec{E} = -\partial_t \vec{B}$	$\nabla \cdot \vec{B} = 0$

Table 6. A rearrangement of Maxwell's equations for materials.

Maxwell's equations can also be rearranged in a manner more convenient for studying materials (as in the table above).

It is reasonable to ask if Maxwell's equations are complete: are there further experiments or phenomena which will lead to more modifications of Maxwell's equations? It appears not. Since Maxwell published his equations in 1864, their consequences have been applied with great success in impressive applications and extensively tested (and verified) by experiment. Away from the atomic scale [1], no discrepancies or need for new terms has been found.

[1] The generalization to describe electromagnetic phenomena as well at the atomic scale is called *quantum electrodynamics*. See: R. Feynman, *QED: The Strange Theory of Light and Matter.* Princeton university press (1986).

Exercises:

1. Consider Maxwell's equations in *vacuum*: $\rho = 0, \vec{J} = 0$. Show that $\vec{E}, \vec{B}$ satisfy the wave equations:

$$\Delta \vec{E} - \frac{1}{c^2}\partial_t^2 \vec{E} = 0, \quad \Delta \vec{B} - \frac{1}{c^2}\partial_t^2 \vec{B} = 0$$

   with $c = \frac{1}{\sqrt{\varepsilon_o \mu_o}}$.

2. (a) Show from Coulomb's law that the units of ε_o are F/m, where m are meters, and F = C/V Farads (where V = Nm/C = J/C are Volts).

   (b) Show from the Biot-Savart law that the units of μ_o are $\frac{N}{A^2}$, where N are Newtons, and A = C/s Amperes (and s are seconds).

   (c) Show the units of $c = \frac{1}{\sqrt{\varepsilon_o \mu_o}}$ are m/s.

   (d) Using the values $\varepsilon_o = 0.885 \times 10^{-11}\ F/m$, $\mu_o = 1.257 \times 10^{-6}\ N/A^2$, check that $c = 2.998 \times 10^8\ m/s$.

3. Show there exists a choice[1] of vector and scalar potential $\vec{A}, \varphi$ with:

$$\vec{B} = \nabla \times \vec{A}, \quad \vec{E} + \partial_t \vec{A} = -\nabla \varphi$$

   such that $\nabla \cdot \vec{A} = 0$.

4. For the Coulomb gauge potentials of the previous problem, show that Gauss law and Maxwell's modification of Ampere's law are satisfied iff

$$\Delta \varphi = -\frac{\rho}{\varepsilon_o}, \quad \Delta \vec{A} - \frac{1}{c^2}\vec{A}_{tt} = -\mu_o \vec{J} + \frac{1}{c^2}\nabla \varphi_t$$

   (where $\varphi_t = \partial_t \varphi, A_{tt} = \partial_t^2 \vec{A}$).

5. Consider a one parameter family[2] of loops, $\mathcal{C}_t$. For each t, let X_t be a vector field on $\mathbb{R}^3$. Show that:

$$\frac{d}{d\varepsilon}|_{\varepsilon=0} \int_{T_\varepsilon} X_t \cdot d\vec{S} = \oint_{\mathcal{C}_t} (X_t \times v) \cdot d\vec{s}$$

   where T_ε is the 'tube' surface parametrized by $(\tau, s) \mapsto \gamma_\tau(s) =: \varphi(\tau, s)$, with $s \in [0,1], \tau \in [t, t+\varepsilon]$ and $v := \partial_\tau \varphi$ (the velocities of the points on the moving loops, see fig. 36).

6. Let $\vec{B}$ be a constant magnetic field, and $\mathcal{C}_r$ a circle of radius r. Let $\vec{\omega}$ be an axis of rotation perpendicular to $\vec{B}$.

   Determine the flux, $\Phi(t)$, of $\vec{B}$ through $\mathcal{C}_r$ as a function of time, when the circle, initially in the $\vec{\omega}, \vec{B}$ plane, is rotated uniformly (with constant angular speed $\omega = |\vec{\omega}|$) about the $\vec{\omega}$ axis through its center.

   For what configuration of the circle relative to $\vec{B}$ is $|\frac{d\Phi}{dt}|$ largest? Smallest?

---

[1] This is called the *Coulomb gauge*. The choice of potentials we made above is called the *Lorenz gauge*. In general choosing vector and scalar potentials is called choice of a gauge.

[2] For each $t \in \mathbb{R}$, $\mathcal{C}_t \subset \mathbb{R}^3$ is a loop. Moreover, the family is 'smooth' in the following sense: for $[0,1] \ni s \mapsto \gamma_t(s)$ a parametrization of $\mathcal{C}_t$, then the map $\mathbb{R} \times [0,1] \to \mathbb{R}^3, (t,s) \mapsto \gamma_t(s)$ is smooth.

§7 Waves

At the end of the last section, we saw the (non-homogeneous) *wave equation*:

$$\Delta u - \frac{1}{v^2} u_{tt} = -s,$$

where v is a constant and s a given function (a 'source'), plays the same fundamental role in electrodynamics that the Poisson equation played in electro and magneto statics.

In general, one interprets solutions of the wave equation as modeling vibrational or oscillating phenomena which propagate at the speed v. First, we consider the one dimensional and homogeneous ($s = 0$) case.

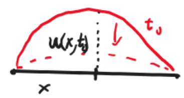

Figure 37. The 1d wave equation (J. d'Alembert) gives a model for small oscillations in a vibrating string or tightly stretched cord (eg a guitar string). The string's position at time t is represented by a graph $u(x,t)$ over its equilibrium position (x-axis).

One-dimensional wave equation: Any solution, $u(x,t)$, to the one-dimensional *homogeneous* wave equation, $v^2 u_{xx} = u_{tt}$, has the form:

$$u(x,t) = f(x - vt) + g(x + vt).$$

proof: Consider the change of variable $\xi = x - vt, \eta = x + vt$. Then

$$v^2 \partial_x^2 - \partial_t^2 = 4v^2 \partial_\xi \partial_\eta.$$

So a solution to the 1-d homogeneous wave equation in these new variables, $u(\xi,\eta)$, solves $\partial_\xi(\partial_\eta u) = 0$. Hence $\partial_\eta u = G(\eta)$ for some function G, and integrating once more gives $u = f(\xi) + g(\eta)$ for some functions f, g. We have the form above upon returning to the original variables, $x - vt = \xi, x + vt = \eta$. $\square$

The change of variables above is motivated by considering that the operator $v^2\partial_x^2 - \partial_t^2$ on C^2 functions is the same as successively applying the operators:

$$v\partial_x - \partial_t, \quad v\partial_x + \partial_t.$$

One may thus determine solutions to the wave equation by finding solutions to the order 1 pdes:

$$vu_x - u_t = 0, \quad vu_x + u_t = 0.$$

The key observation (a case of the *method of characteristics*), is that a solution, $u(x,t)$, to say $vu_x + u_t = 0$ remains constant over the lines $x = vt + x_o$. Thus its initial values, $u_o(x) = u(x,0)$, 'propagate' along the lines $x = vt + x_o$ to determine its values over the whole xt plane. Likewise one describes solutions to $vu_x - u_t = 0$, and then one obtains general solutions to the wave equation by superposition (summing two such solutions).

The general solution, may be related to initial conditions:

d'Alembert's formula: given initial *Cauchy data*, $u_o(x) = u(x,0), \quad \dot{u}_o(x) = \partial_t u(x,0)$ then

$$u(x,t) := \frac{u_o(x+vt) + u_o(x-vt)}{2} + \frac{1}{2v}\int_{x-vt}^{x+vt} \dot{u}_o(y)\, dy$$

is the unique solution to the wave equation satisfying these initial conditions.

proof: One writes $u(x,t) = f(x - vt) + g(x + vt)$ and solves $u_o(x) = f(x) + g(x), \quad \dot{u}_o(x) = v\,(g'(x) - f'(x))$ for f, g. $\square$

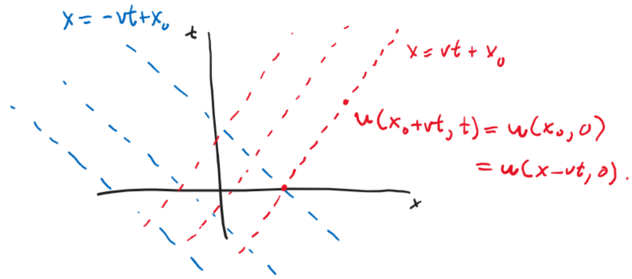

Figure 38. Solutions to the wave equation may be expressed as the sum of two types of left or right travelling waves: the initial values remain constant along the red (left travelling wave) or blue (right travelling wave) curves.

We next consider some *boundary conditions*:

Harmonics: Given $\ell > 0$, convergent sums of the form

$$u(x,t) = \sum_{n \geq 0} (a_n \cos \omega_n t + b_n \sin \omega_n t) \sin k_n x, \quad a_n, b_n \in \mathbb{R},$$

$k_n := \frac{n\pi}{\ell}, \omega_n := vk_n$ are periodic solutions of $v^2 u_{xx} = u_{tt}$ satisfying $u(0,t) = u(\ell,t) = 0$.

proof: One may obtain these expressions directly by *separation of variables*: seek a solution of the form $u(x,t) = X(x)T(t)$ leading to:

$$\frac{v^2 X''(x)}{X(x)} = \frac{T''(t)}{T(t)} = cst.$$

the bounded solutions correspond to a negative constant value, which we denote $-\omega^2$ (and $k := \omega/v$).
The boundary conditions, $X(0) = X(\ell) = 0$, are satisfied when $k = \frac{n\pi}{\ell}$, $n \in \mathbb{Z}$, and by linearity we may sum any number of such solutions (when the infinite sums have suitable convergence).
More conceptually, the procedure is analogous to how one would solve a linear system $\ddot{x} = Lx$. First determine the eigenvectors and eigenvalues $Lx_n = \lambda_n x_n$ of the matrix L. If we are interested in bounded solutions (physically these are in general the ones of interest), then we consider the negative eigenvalues, $\lambda_n = -\omega_n^2$, with corresponding solutions given by:

$$x(t) = (a_n \cos \omega_n t + b_n \sin \omega_n t) x_n.$$

For the wave equation, the linear operator, $v^2 \partial_x^2$ plays the role of the matrix L :

$$X(x) \mapsto v^2 X''(x)$$

and, with the boundary conditions $X(0) = X(\ell) = 0$, we find eigenvectors $X_n(x) = \sin k_n x$ for $k_n = \frac{n\pi}{\ell}$. $\square$

In particular, we have,

$$A \cos \omega(t + t_o) \sin k(x + x_o), \quad \omega = vk$$

where $A \in \mathbb{R}_{\geq 0}$ is called the *amplitude*, are particular solutions of the wave equation. These solutions go by many names: eg *harmonics* or *standing waves* or *fundamental modes*. They may be written as well in the notation of complex numbers as the real or imaginary parts of :

$$ae^{i(kx+\omega t)} + be^{i(kx-\omega t)}, \quad a, b \in \mathbb{C}$$

The *wavelength*, λ, and *period*, T, of a harmonic are:

$$\lambda = \frac{2\pi}{k}, \quad T = \frac{2\pi}{\omega}.$$

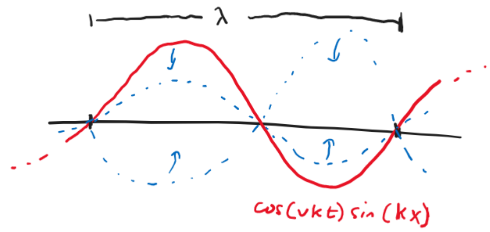

Figure 39. A standing wave or harmonic oscillates periodically.

The parameters k, ω are called the *wave number* and *angular frequency* respectively, while

$$f = \frac{1}{T}$$

is called the *frequency* (measured in *Hertz*, Hz $=$ sec^{-1}). Note that the relation $\omega = vk$ may be written:

$$f\lambda = v.$$

Fourier series: Any ℓ periodic, $u(x+\ell,t) = u(x,t)$, solution to $v^2 u_{xx} = u_{tt}$ may be expanded in harmonics. Moreover, for initial conditions $u(x,0) = u_o(x), \partial_t(u,0) = \dot{u}_o(x)$ with $u_o, \dot{u}_o$ of period ℓ, then

$$u(x,t) = a_0 + \sum_{n\in\mathbb{Z}\setminus 0} (a_n \cos\omega_n t + b_n \sin\omega_n t)e^{ik_n x}$$

where $k_n = \frac{2n\pi}{\ell}, \omega_n = vk_n$ and

$$a_n = \frac{1}{\ell}\int_{-\ell/2}^{\ell/2} u_o(x)e^{-ik_n x}\ dx, \quad b_n = \frac{1}{\omega_n \ell}\int_{-\ell/2}^{\ell/2} \dot{u}_o(x)e^{-ik_n x}\ dx$$

is the unique period ℓ solution to the wave equation satisfying these initial conditions.

proof: For each $n \in \mathbb{Z}$, the (complex valued) functions $X_n(x) := e^{ik_n x}$, $k_n = \frac{2\pi n}{\ell}$ are ℓ periodic functions of x and are eigenvectors of the operator $v^2\partial_x^2$ (with eigenvalue $-\omega_n^2$). Consider the inner product:

$$\langle f, g\rangle := \int_{-\ell/2}^{\ell/2} f(x)\bar{g}(x)\ dx$$

on $\mathbb{C}$ valued functions $f(x), g(x)$. Then:

$$\langle X_n, X_m\rangle = \begin{cases} \ell & n = m \\ 0 & n \neq m \end{cases}$$

so that the functions $X_n(x), n \in \mathbb{Z}$ are orthogonal. In fact, they are a basis for periodic functions [1] so that the initial conditions in this basis are:

$$u_o(x) = \sum a_n X_n, \quad \dot{u}_o(x) = \sum b_n X_n$$

where $a_n\ell = \langle u_o, X_n\rangle, b_n\ell = \langle \dot{u}_o, X_n\rangle$ and the solution with these initial conditions is given as above. Note that for a bounded periodic solution, it is necessary that $b_0 = \int_{-\ell/2}^{\ell/2} \dot{u}_o(x)\ dx = 0$ (otherwise one has an unbounded term $b_0 t$ translating the solution upwards). $\square$

[1] See eg, Arnold's lectures on pde's for some convergence properties.

The same 'linear algebra' idea we have used for periodic solutions may be applied to derive a general formula for solutions using the Fourier transform. To see the analogy, first we summarize the main formulas for Fourier series we have just used for period ℓ functions:

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{ik_n x}, \quad a_n = \frac{1}{\ell}\int_{-\ell/2}^{\ell/2} f(x)e^{-ik_n x},$$

with $k_n = \frac{2\pi n}{\ell}$. Note that $k_n - k_{n-1} = \frac{2\pi}{\ell}$.

For a non-periodic function, $f(x)$, one may apply these formulas to obtain periodic approximations to f. That is, restrict f to the interval $[-\ell/2, \ell/2]$ and extend periodically, obtaining a periodic function with a 'jump' at the endpoints.

Then, taking $\hat{f}(k_n) := \frac{\ell}{2\pi}a_n$, the formulas above may be written:

$$f(x) \approx \sum_{n=-\infty}^{\infty} \hat{f}(k_n)e^{ik_n x}(k_n - k_{n-1}), \quad \hat{f}(k_n) := \frac{1}{2\pi}\int_{-\ell/2}^{\ell/2} f(x)e^{-ik_n x}\ dx.$$

As we let $\ell \to \infty$, we obtain an approximation[1] to $f(x)$ over the whole real line:

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(k)e^{ikx}\ dk, \quad \hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ikx}\ dx.$$

The function $\hat{f}(k)$ is called the *Fourier transform* of $f(x)$.

Fourier transform: for initial conditions $u_o(x), \dot{u}_o(x)$ which vanish outside some compact interval, then

$$u(x,t) = \int_{-\infty}^{\infty} \cos vkt\ \hat{u}_o(k)e^{ikx}\ dk + \int_{-\infty}^{\infty} \frac{\sin vkt}{vk}\hat{\dot{u}}_o(k)e^{ikx}\ dk$$

where $\hat{u}_o, \hat{\dot{u}}_o$ are the Fourier transforms of $u_o, \dot{u}_o$, is the solution to the 1d wave equation with these initial conditions.

This last equation returns d'Alembert's formula in the 1d-case, and the same method may be applied to obtain a formula for general solutions in the n dimensional case. We now consider some special solutions in the spatial case.

Plane waves: Consider Maxwell's equations in vacuum: $\rho = 0, \vec{J} = 0$. Let us seek electric and magnetic fields which depend only on t and one Cartesian coordinate, say x. The divergence terms read:

$$\partial_x E_1 = 0, \quad \partial_x B_1 = 0$$

and the curl terms read:

$$(0, -\partial_x E_3, \partial_x E_2) = -\partial_t(B_1, B_2, B_3), \quad (0, -\partial_x B_3, \partial_x B_2) = \mu_o\varepsilon_o\partial_t(E_1, E_2, E_3).$$

Hence E_1, B_1 are constants. We take them to be zero. For the remaining components we have 1d-wave equations:

$$\partial_x^2 E_2 = -\partial_t\partial_x B_3 = \frac{1}{c^2}\partial_t^2 E_2$$

and similarly for E_3, B_2, B_3.

The solutions however are not entirely independent. Namely, if we write:

$$E_2 = f(x - ct) + g(x + ct)$$

[1] To justify equality in these formulas, one needs to give convergence conditions.

then from $\partial_t B_3 = -\partial_x E_2, c^2 \partial_x B_3 = -\partial_t E_2$, we have:

$$B_3 = \frac{f(x-ct) - g(x+ct)}{c}.$$

And a similar relation, for the E_3, B_2 components.

Note that the time dependent components are contained in the yz plane, perpendicular to the propagating direction x. As well the components come in dependent orthogonal pairs: E_2 with B_3 as well as E_3 with B_2. Such solutions are called *plane waves*: the wave propagates in a fixed direction (here the x-axis).

A special case is that of a *monochromatic plane wave*: when E_2 has a constant frequency (and so also constant wavelength). with $\omega = ck$,

$$E_2 = E_o \cos(\omega t - kx), B_3 = \frac{E_o}{c} \cos(\omega t - kx)$$

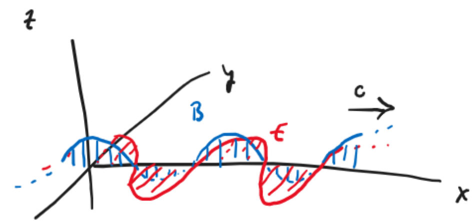

Figure 40. A monochromatic plane wave ($|B| = |E|/c$) travelling in a fixed direction (x-axis) with a fixed frequency.

Spherical waves: we seek solutions to the spatial wave equation that –in spherical coordinates– only depend on $r = |x|$.

The Laplacian in spherical coordinates on $u(r)$ is:

$$\Delta u = \frac{1}{r^2} \partial_r (r^2 \partial_r u) = u_{rr} + \frac{2}{r} u_r.$$

So the 3d-wave equation for $u(r,t)$ may be written:

$$v^2 \left(r u_{rr} + 2 u_r \right) = r u_{tt}.$$

Note that $r u_{rr} + 2u_r = (r u_r)_r + u_r = (r u_r + u)_r = (ru)_{rr}$, so that:

$$v^2 (ru)_{rr} = (ru)_{tt}$$

and ru satisfies the 1-d wave equation. Thus $u(r,t)$ has the form:

$$u(r,t) = \frac{f(r - vt) + g(r + vt)}{r}.$$

Now, we consider the spatial wave equation with a source term:

Retarded potentials: when the integral is defined, the function

$$u(x,t) = \frac{1}{4\pi} \int_{y \in \mathbb{R}^3} \frac{s(y, t - r/v)}{r} \, dV$$

for $r := |y - x|$ is a solution to the wave equation $\Delta u - \frac{1}{v^2} u_{tt} = -s$ with source $s(x,t)$.

proof: Consider an 'outgoing' spherical wave $u(r,t) = \frac{f(r-vt)}{r}$. This spherical wave satisfies the homogeneous wave equation for $r \neq 0$, but over $\mathbb{R}^3$ satisfies: $\Delta u - \frac{1}{v^2}u_{tt} = -4\pi f(-vt)\delta_o$ since $\Delta \frac{1}{r} = -4\pi\delta_o$. Setting $s(t) = 4\pi f(-vt)$, we have a solution

$$\frac{s(t-r/v)}{4\pi r}$$

to

$$\Delta u - \frac{1}{v^2}u_{tt} = -s(t)\delta_o.$$

We obtain solutions with general sources by superposition[1]. Write

$$s(x,t) = \int_{y\in\mathbb{R}^3} s(y,t)\delta(x-y)\ dV.$$

Each source $s(y,t)\delta(x-y)$, for *fixed* y has solution $\frac{s(y,t-r/v)}{4\pi r}$, with $r = |x-y|$ and the solution above is obtained by replacing the integrands by their respective solutions:

$$s(x,t) = \int_y s(y,t)\delta(x-y)\ dV = -\int_y \Box_x \frac{s(y,t-r/v)}{4\pi r}\ dV = -\Box_x \int_y \frac{s(y,t-r/v)}{4\pi r}\ dV$$

where $\Box := \Delta - \frac{1}{v^2}\partial_t^2$ is the wave operator or *d'Alembertian.* $\Box$

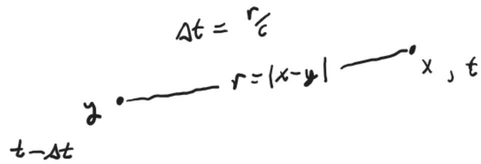

Figure 41. A solution, $u(x,t)$, to the wave equation depends on the sources, $s(y, t-r/c)$, at earlier 'retarded' times. The time delay, Δt, between the source at y and its effect on u is the time it takes to travel from y to x at speed c.

General formulas for solutions to the wave equation may also be derived using the Fourier transform (see exercises). The solution given above by the use of these retarded potentials is sufficient for 'solving' Maxwell's equations via the Lorenz gauge:

Lorenz gauge	
$\vec{B} = \nabla \times \vec{A}$	$\vec{E} = -\vec{A}_t - \nabla\varphi$
$\nabla \cdot \vec{A} = -\mu_o\varepsilon_o\varphi_t$	
$\vec{A}(x,t) = \frac{\mu_o}{4\pi}\int_{y\in\mathbb{R}^3} \frac{\vec{J}(y,t')}{r}\ dV$	$\varphi(x,t) = \frac{1}{4\pi\varepsilon_o}\int_{y\in\mathbb{R}^3} \frac{\rho(y,t')}{r}\ dV$

Table 7. Solving Maxwell's equations given charge and current densities $\rho(x,t), \vec{J}(x,t)$. The potentials of the Lorenz gauge are given by integrals involving the retarded potentials, with $r = |x-y|$, $t' = t - r/c$ and $c = \frac{1}{\sqrt{\mu_o\varepsilon_o}}$.

[1]Compare to the analogous method for writing solutions to the Poisson equation. We have $\Delta\frac{1}{4\pi r} = -\delta_o$, so that $\rho(x) = \int_y \rho(y)\delta(x-y)\ dV = -\int_y \rho(y)\Delta_x \frac{1}{4\pi|x-y|}\ dV = -\Delta_x \int_y \frac{\rho(y)\ dV}{4\pi|x-y|}$.

§8 Energy

We consider in this section energy associated to charge or current configurations.

First, we will consider some static cases. For a capacitor, consisting of conductors Ω_1, Ω_2 with capacitance C, we have seen above that the work required to charge the capacitors to charges $Q, -Q$ is:

$$W = \frac{1}{2}\frac{Q^2}{C}.$$

The energy of a configuration may be thought of as the 'stored work' in the configuration, ie how much input work is required to arrange the given configuration. Thus, the energy of the capacitors charged to $Q, -Q$ is $W = \frac{1}{2}\frac{Q^2}{C}$. In general:

Energy of static charge distribution: The amount of work required to arrange a static charge distribution ρ is:

$$W_{el} = \frac{1}{2}\int_{\mathbb{R}^3} \rho\varphi \; dV = \frac{\varepsilon_o}{2}\int_{\mathbb{R}^3} |\vec{E}|^2 \; dV$$

where $\vec{E} = -\nabla\varphi$ is the electric field generated by the charge distribution ρ.

proof: First, we consider the work required to assemble a collection of point charges $q_1, ..., q_N$ at given positions $x_1, ..., x_N$. There is no work required to bring the first charge q_1 to x_1, since there are no other charges present. Now with q_1 in place, we bring q_2 to x_2 which requires work:

$$\frac{q_1 q_2}{4\pi\varepsilon_o |x_1 - x_2|}$$

since q_2 must be moved to x_2 in the presence of the field due to q_1 at x_1. Successively placing the charges in the presence of the fields due to the previously placed charges, one obtains:

$$W = \frac{1}{2}\sum_{i \neq j} \frac{q_i q_j}{4\pi\varepsilon_o |x_i - x_j|}$$

for the total energy contained in the configuration of point charges. For a continuous distribution, ρ, then:

$$W = \frac{1}{2}\int_{(x,y)\in\mathbb{R}^3\times\mathbb{R}^3} \frac{\rho(x)\rho(y)}{4\pi\varepsilon_o |x-y|} dV_x dV_y = \frac{1}{2}\int_{x\in\mathbb{R}^3} \rho(x)\varphi(x) \; dV$$

since $\int_{y\in\mathbb{R}^3} \frac{\rho(y)}{4\pi\varepsilon_o|x-y|} \; dV = \varphi(x)$. To obtain the expression involving $\vec{E} = -\nabla\varphi$, we use the Gauss equation, $-\varepsilon_o\Delta\varphi = \rho$, and $\nabla\cdot(\varphi\nabla\varphi) = |\nabla\varphi|^2 + \varphi\Delta\varphi$ to obtain:

$$\frac{1}{2}\int_{B_r} \rho\varphi \; dV = \frac{\varepsilon_o}{2}\left(\int_{B_r} |\nabla\varphi|^2 \; dV - \int_{\partial B_r} \varphi\nabla\varphi\cdot d\vec{S}\right)$$

where B_r is a solid ball of radius r. Letting $r \to \infty$ and using that $\varphi = O(\frac{1}{r}), \nabla\varphi = O(\frac{1}{r^2})$ (bounded charges), the boundary term goes to zero so that $W = \frac{\varepsilon_o}{2}\int_{\mathbb{R}^3} |\vec{E}|^2 \; dV$. $\square$

The expression for the total energy as an integral involving $\vec{E}$, may be interpreted as the electric field storing energy. Namely one calls:

$$\mathscr{W}_{el} := \frac{\varepsilon_o}{2}|\vec{E}|^2$$

the (field) *energy density* for the static charge distribution.

To derive an analogous expression for the energy of a given current distribution, we will begin by considering the amount of work to produce a current I in a given conducting loop $\mathcal{C}$.

First, we have an analogue of capacitance:

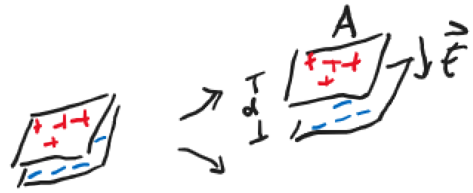

Figure 42. To create an electric field in a small region of volume, one can imagine separating two equally charged plates to produce an essentially uniform field between the plates. The act of separating the plates requires work, so one regards the electric field in the volume as containing stored work (energy).

(self) Inductance of a current loop: if a conducting loop $\mathcal{C}$ is supplied with a steady current I then the resulting magnetic flux, Φ, through the loop is proportional to I:

$$\Phi = LI$$

where L is a constant (depending on the loop $\mathcal{C}$) called the (self) *inductance* of the loop. It is measured in *Henries* (H = Wb/A, where Wb = T$\cdot$m^2 are *Webers*).

proof: By Biot-Savart, the magnetic field $\vec{B}(x)$ at a given point is proportional to the current I. Hence so too is its surface integral, $\int_\Sigma \vec{B} \cdot d\vec{S}$, over a surface, $\partial\Sigma = \mathcal{C}$, spanning the loop. $\square$

We may determine the work required to generate a current I in the loop $\mathcal{C}$, by the following considerations. To achieve the current I requires providing an increasing current $I(t)$ with, say, $I(0) = 0, I(1) = I$ in the loop. As the current grows to I the magnetic flux through the loop is time dependent, given by $\Phi(t) = LI(t)$. Now by Faraday's law,

$$L\dot{I} = \dot{\Phi} = -\mathcal{E}_{back},$$

the changing magnetic flux induces a *back emf*, $\mathcal{E}_{back}$, in the loop, driving current in the opposite direction to that which we are applying. Thus, to overcome this back emf, *we* must provide additional work $-\mathcal{E}_{back} = L\dot{I}$ per unit charge around the loop.

To determine the total work we provide against this back emf as the current is increased to I, we consider the *power* [1] required at a given time:

$$\dot{W} = \int_{\mathcal{C}} \vec{f} \cdot \vec{v}\ ds = \int_{\mathcal{C}} \frac{\vec{f}}{\lambda} \cdot \lambda\vec{v}\ ds = L\dot{I}I$$

since $\lambda\vec{v}\ ds = Id\vec{s}$ and $\int_{\mathcal{C}} \frac{\vec{f}}{\lambda} \cdot d\vec{s} = -\mathcal{E}_{back} = L\dot{I}$. Integrating, the total work W to drive a current I in the loop is then:

$$W = \frac{1}{2}LI^2.$$

In general:

Energy of steady current distribution: The amount of work required to arrange a steady current distribution $\vec{J}$ is:

$$W_{mag} = \frac{1}{2}\int_{\mathbb{R}^3} \vec{J} \cdot \vec{A}\ dV = \frac{1}{2\mu_o}\int_{\mathbb{R}^3} |\vec{B}|^2\ dV$$

where $\vec{B} = \nabla \times \vec{A}$ is the magnetic field generated by the current distribution $\vec{J}$.

[1] Power is defined as work done per unit time, ie the rate at which work is done. Power is measured in *Watts* (W = J/s). The total work done in moving a particle subject to forces $\vec{f}$ along a path $\mathcal{C}$ is $\int_{\mathcal{C}} \vec{f} \cdot d\vec{s}$. When a particle moves with velocity $\vec{v}$ along a path $\mathcal{C}$ subject to forces $\vec{f}$, the total power is $\int_{\mathcal{C}} \vec{f} \cdot \vec{v}\ ds$.

proof: A steady current $\vec{I}$ along a loop $\mathcal{C}$ has total energy:

$$\frac{1}{2}LI^2 = \frac{1}{2}\Phi I = \frac{I}{2}\int_{\mathcal{C}} \vec{A}\cdot d\vec{s} = \frac{1}{2}\int_{\mathcal{C}} \vec{A}\cdot\vec{I}\,ds$$

since $\vec{I}ds = Id\vec{s}$ and $\nabla\times\vec{A} = \vec{B}$. By superposition, a steady current density $\vec{J}$ has total energy:

$$\frac{1}{2}\int_{\mathbb{R}^3} \vec{J}\cdot\vec{A}\,dV.$$

To obtain the expression involving $\vec{B} = \nabla\times\vec{A}$, we use Ampere: $\vec{J} = \frac{\nabla\times\vec{B}}{\mu_o}$ as well as the product rule, $\nabla\cdot(X\times Y) = Y\cdot(\nabla\times X) - X\cdot(\nabla\times Y)$, to obtain:

$$\frac{1}{2}\int_{B_r} \vec{J}\cdot\vec{A}\,dV = \frac{1}{2\mu_o}\left(\int_{B_r} |\vec{B}|^2\,dV + \int_{\partial B_r} (\vec{B}\times\vec{A})\cdot d\vec{S}\right)$$

where B_r is a solid ball of radius r. Letting $r\to\infty$ the boundary term goes to zero (for potentials and fields vanishing at infinity, eg from bounded current densities). $\square$

We may intepret the magnetic field itself as storing energy, with:

$$\mathcal{W}_{mag} := \frac{1}{2\mu_o}|\vec{B}|^2$$

the (field) *energy density* for the current distribution.

Total energy in statics	
$Q = CV$	$\Phi = LI$
$W = \frac{1}{2}\frac{Q^2}{C}$	$W = \frac{1}{2}LI^2$
$W = \frac{1}{2}\int_{\mathbb{R}^3}\rho\varphi\,dV = \frac{\varepsilon_o}{2}\int_{\mathbb{R}^3}|\vec{E}|^2\,dV$	$W = \frac{1}{2}\int_{\mathbb{R}^3}\vec{J}\cdot\vec{A}\,dV = \frac{1}{2\mu_o}\int_{\mathbb{R}^3}|\vec{B}|^2\,dV$

Table 8. Total energy in static electromagnetism.

It is important to remark that in statics we have two distinct natural choices for 'energy densities'. Namely either $\rho\varphi$ or $\frac{\varepsilon_o}{2}|\vec{E}|^2$ may be regarded as energy densities in electrostatics. The choice we have made for the energy density involving the fields is the standard convention as, we see below, it appears more naturally in the general dynamic (non-static) case.

Now we apply similar energy considerations in the general case. Let $\vec{E},\vec{B}$ be a solution to Maxwell's equations with sources $\rho,\vec{J}$.
First we examine how much work is done by the electric and magnetic fields on the sources. By Lorentz, the force density (force per volume) is:

$$\mathcal{F} = \rho(\vec{E} + \vec{v}\times\vec{B}) = \rho\vec{E} + \vec{J}\times\vec{B}$$

where $\vec{v}$ is the velocity of the charges, ie $\vec{J} = \rho\vec{v}$. As the sources, ρ, move with velocity $\vec{v}$ (generating the currents $\vec{J} = \rho\vec{v}$), the rate work is done by the fields on the sources per volume, ie the *power density* is:

$$\mathcal{P} := \mathcal{F}\cdot\vec{v} = \vec{E}\cdot\vec{J}.$$

We will now use Maxwell's equations to arrange this power density in the form of a conservation law, similar to conservation of charge. First, we substitute $\vec{J} = \frac{\nabla\times\vec{B}}{\mu_o} - \varepsilon_o\partial_t\vec{E}$, to obtain:

$$\mathcal{P} = -\partial_t\frac{\varepsilon_o|\vec{E}|^2}{2} + \frac{1}{\mu_o}\vec{E}\cdot(\nabla\times\vec{B}).$$

Now, we use the product rule, $\nabla \cdot (X \times Y) = Y \cdot (\nabla \times X) - X \cdot (\nabla \times Y)$, along with $\nabla \times \vec{E} = -\partial_t \vec{B}$, to obtain:

$$\mathscr{P} = -\partial_t \left(\frac{\varepsilon_o |\vec{E}|^2}{2} + \frac{|\vec{B}|^2}{2\mu_o} \right) - \frac{1}{\mu_o} \nabla \cdot (\vec{E} \times \vec{B}).$$

The quantities:

$$\mathscr{W}_f := \frac{\varepsilon_o |\vec{E}|^2}{2} + \frac{|\vec{B}|^2}{2\mu_o}, \quad \vec{S} := \frac{\vec{E} \times \vec{B}}{\mu_o}$$

are called the *field energy density* and *Poynting vector* respectively.

The interpretation of the field energy density, from the static cases, is the amount of work required to produce the fields $\vec{E}, \vec{B}$, ie the energy stored in the electric and magnetic fields. The power density is the rate work is done by the fields on the sources. We write, $\mathscr{P} = \partial_t \mathscr{W}_s$, as the time derivative of the work done by the fields on the sources.

The Poynting vector may be interpreted as the direction of energy flow in the following sense. With our notations above [1]:

$$\mathscr{P} + \partial_t \mathscr{W}_f = \partial_t(\mathscr{W}_s + \mathscr{W}_f) = -\nabla \cdot \vec{S}.$$

Now, let Ω be a region with boundary $\partial\Omega$. There is energy stored in Ω through the work required to generate the fields $\vec{E}, \vec{B}$ in Ω as well as through the amount of work done by the fields on the sources in Ω. We set:

$$W(t) := \int_\Omega (\mathscr{W}_s + \mathscr{W}_f)\ dV$$

for the total energy stored in Ω at time t. Then the rate the total energy in Ω is changing is:

$$\dot{W}(t) = \int_\Omega \mathscr{P} + \partial_t \mathscr{W}_f\ dV = -\int_\Omega \nabla \cdot \vec{S}\ dV = -\int_{\partial\Omega} \vec{S} \cdot d\vec{S}.$$

Thus, if the total energy in Ω is changing, conservation of energy dictates energy is flowing through the boundary. According to the above formula, this rate energy flows through the boundary, $\int_{\partial\Omega} \vec{S} \cdot d\vec{S}$, is the flux of $\vec{S}$ through $\partial\Omega$. As this holds for any region Ω, we interpret the Poynting vector as an energy flux density: the energy is flowing in the direction $\vec{S}$ and the rate at which energy flows across a surface Σ is given by $\int_\Sigma \vec{S} \cdot d\vec{S}$.

Example:

- We consider energy in the simplest electromagnetic waves: the monochromatic plane waves. Consider a monochromatic plane wave ($\omega = ck$):

  $$\vec{E} = E \cos(\omega t - kx)\ \mathbf{j}, \quad \vec{B} = \frac{E}{c} \cos(\omega t - kx)\ \mathbf{k}$$

  traveling along the x-axis. We examine how energy is carried by such a wave.

  The Poynting vector is:

  $$\vec{S} = \frac{\vec{E} \times \vec{B}}{\mu_o} = \varepsilon_o c\ E^2 \cos^2(\omega t - kx)\ \mathbf{i},$$

  signifying that the wave is carrying energy along its direction of propagation (the x-axis) at the rate $\varepsilon_o c\ E^2 \cos^2(\omega t - kx)$. Since the wave propagates at velocity c, the energy density in the wave should then be $\varepsilon_o E^2 \cos^2(\omega t - kx)$. Indeed, the field energy density is (using $c^2 = \frac{1}{\mu_o \varepsilon_o}$):

  $$\frac{\varepsilon_o}{2} |\vec{E}|^2 + \frac{1}{2\mu_o} |\vec{B}|^2 = \varepsilon_o E^2 \cos^2(\omega t - kx).$$

  Over one cycle of the wave, $0 \leq t \leq T = \frac{1}{f} = \frac{2\pi}{\omega}$, the average rate of energy transmission is:

  $$\frac{\omega}{2\pi} \int_0^{2\pi/\omega} \varepsilon_o c\ E^2 \cos^2(\omega t - kx)\ dt = \frac{\varepsilon_o c\ E^2}{2}$$

[1] Compare to the expression for conservation of charge: $\partial_t \rho = -\nabla \cdot \vec{J}$.

called the *intensity*, while the average energy over a cycle of the wave is $\frac{\varepsilon_o E^2}{2}$.

In summary, we have the table:

Energy	
power density (rate work is done by fields on sources)	$\mathscr{P} := \vec{E} \cdot \vec{J}$
field energy density	$\mathscr{W}_f := \frac{\varepsilon_o \lvert\vec{E}\rvert^2}{2} + \frac{\lvert\vec{B}\rvert^2}{2\mu_o}$
energy flux density (Poynting vector)	$\vec{S} := \frac{\vec{E}\times\vec{B}}{\mu_o}$
continuity equation (conservation of energy)	$\mathscr{P} + \partial_t \mathscr{W}_f = -\nabla \cdot \vec{S}$

Table 9. Conventional energy densities in electromagnetism.

Similar remarks as in the static case apply. Namely, the field density and energy flux density we have 'found' above are note unique in their property that $\vec{E} \cdot \vec{J} + \partial_t u = -\nabla \cdot \vec{X}$. Essentially, one may 'trade' terms in u for terms in X, however the Poynting vector and field density above are in some sense the 'simplest' or most natural choices and are by convention considered to represent the local densities and directions of flow for the energy of the system. See ch. 8 of Griffiths on energy and conserved quantities in electromagnetism as well as §27 of Feynman.

§9 Radiation

The focus of this section is that:

an accelerating charge emits (radiates) energy.

We see some particular situations in which this statement may be made more quantitative.

Qualitatively, the phenomenon is understood by first noting an electromagnetic wave carries energy. If a charge initially at rest is accelerated and moved to a new position there will be a 'kink' between the initial and final 'Coulomb' electromagnetic fields propagating at speed c from the charges final position. That is an electromagnetic wave, carrying energy, is produced propagating away from the accelerated charge.

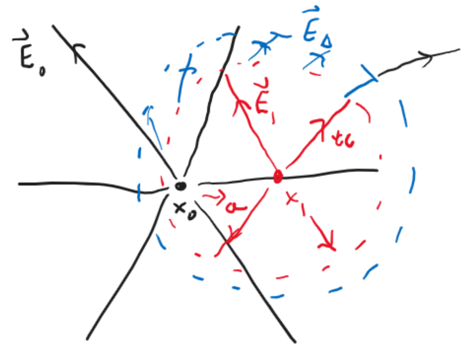

Figure 43. A charge initially at rest at x_o, for $-\infty < t < 0$, produces a static Coulomb field $\vec{E}_o$ centered at x_o. If the charge is moved to a new position x_1 (accelerated from rest), in a time Δt, then for $t > \Delta t$ it produces a Coulomb field $\vec{E}_1$ centered at x_1 and propogated outwards to a sphere of radius ct. In the region between these two static fields, there is a 'kink' in the fields propogating outwards from x_1, ie an electromagnetic wave propogating away from the source is produced when the charge is acclerated. See this link for some better figures and consequences.

The general approach to describe energy *radiated* by sources, is to first determine the Poynting vector $\vec{S}$. One then considers, say, a sphere, S^2_r, of radius r surrounding the sources and determines the flux:

$$P(r) := \int_{S^2_r} \vec{S} \cdot \nu \, dA$$

for the rate of energy (power) passing through the sphere. Now if $P(r) \neq 0$ for some r, it does not mean energy is lost from the sources. It may at a later time re-enter through the sphere. Let:

$$P_{rad} := \lim_{r\to\infty} P(r).$$

If $P_{rad} \neq 0$, then this is a rate energy is transmitted outwards from the system at *all* distances. There is no 'sphere' or surface for such energy to re-enter the system.

As usual we consider bounded charges with fields vanishing at infinity. Since $\vec{S} = \frac{\vec{E}\times\vec{B}}{\mu_o}$, and spheres have area $4\pi r^2$, the relevant terms of $\vec{E}, \vec{B}$ for determining P_{rad} are those with strength $\sim \frac{1}{r}$.

Example:

- A *Hertzian dipole*, consists of charge oscillating between two conducting spheres seperated by a distance ℓ. Let

  $$q(t) = q_o \sin \omega t$$

  be the charge on the 'upper' sphere at time t (so $-q_o \sin \omega t$ is the charge on the lower sphere at time t). The charge transfer is driven by the current $I(t) = q_o\omega \cos \omega t$ along the axis.

We align the axis between the spheres with the z-axis and set:

$$p_o = \ell q_o, \quad p(t) = \ell q(t), \quad \vec{p}(t) = p(t)\ \mathbf{k}$$

for the dipole moments of the charges.

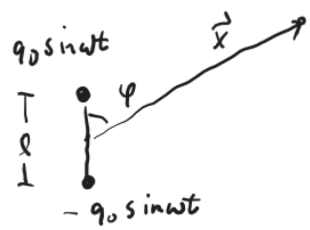

Figure 44. A Hertzian dipole consists of charge oscillating between two 'close' points seperated by distance ℓ.

We would like to describe the energy radiated by these oscillating charges.

First, the magnetic potential of the Lorenz gauge is:

$$\vec{A}(\vec{x}, t) = \frac{\mu_o}{4\pi} \int_{-\ell/2}^{\ell/2} \frac{I(t - r'/c)}{r'}\ dz\ \mathbf{k}$$

where r' is the distance from $\vec{x}$ to $(0, 0, z)$.

Now, we make an expansion to determine the dominant term of $\vec{A}$. As with a dipole, we assume we are at a relatively large distance relative to the seperation between the spheres:

$$\ell << r,$$

where $r := |\vec{x}|$. Then $r' = r(1 + O(\frac{\ell}{r}))$, and $\frac{1}{r'} = \frac{1}{r} + O(\frac{\ell}{r})$. Thus:

$$\vec{A}(\vec{x}, t) \approx \frac{\mu_o \omega q_o}{4\pi r} \int_{-\ell/2}^{\ell/2} \cos \omega(t - r'/c)\ dz\ \mathbf{k}.$$

To simplify the remaining integral, we observe that:

$$\cos \omega(t - r'/c) = \cos \omega(t - r/c) + O(\frac{\omega \ell}{c})$$

and we will further assume that the frequency of oscillation satisfies:

$$\ell << \lambda$$

where $\lambda = \frac{c}{f} = \frac{2\pi c}{\omega}$ is the wavelength. The vector potential is then approximated by:

$$\vec{A}(\vec{x}, t) \approx \frac{\mu_o \omega p_o}{4\pi r}\ \cos \omega(t - r/c)\ \mathbf{k} = \frac{\mu_o}{4\pi r}\ \dot{\vec{p}}(t - r/c)$$

where $r = |\vec{x}|$. The scalar potential φ, may be determined similarly, with the additional condition:

$$\lambda << r$$

to obtain

$$\varphi(\vec{x}, t) \approx \frac{\mu_o c}{4\pi r^2} z\ \dot{p}(t - r/c).$$

Now the electromagnetic field approximations may be obtained by computing $\vec{B} = \nabla \times \vec{A}, \quad \vec{E} = -\nabla\varphi - \vec{A}_t$, resulting in (spherical coordinates)

$$\vec{E} \approx -\frac{\mu_o \omega^2}{4\pi r}\ p(t - r/c)\sin\phi\ \mathbf{e}_\phi,$$

$$\vec{B} \approx -\frac{\mu_o \omega^2}{4\pi r c}\ p(t - r/c)\sin\phi\ \mathbf{e}_\theta,$$

and Poynting vector

$$\vec{S}(\vec{x}, t) \approx \frac{\mu_o \omega^4}{16\pi^2 c} p^2(t - r/c)\ \sin^2\phi\ \frac{\vec{x}}{r^3}$$

where $r = |\vec{x}|$ and ϕ is the angle between $\vec{x}$ and $\mathbf{k}$ (the z-axis).

Note that the Poynting vector is radial. The average over one period of oscillation of the rate of radial energy transfer per area in the direction of the ray with angle ϕ from the z-axis:

$$\langle S \rangle := \frac{1}{T}\int_0^T \vec{S} \cdot \frac{\vec{x}}{|x|}\ dt = \frac{\mu_o \omega^4 p_o^2}{32\pi^2 c}\ \frac{\sin^2\phi}{r^2}.$$

Thus, the average rate of energy flow across a sphere of radius r, over one period, is:

$$P_{rad} = \int_{S_r^2} \langle S \rangle\ dA = \frac{\omega^4 p_o^2}{12\pi\varepsilon_o c^3}.$$

In summary, we have:

Hertzian dipole radiation: suppose charge oscillates with frequency $f = \frac{\omega}{2\pi}$ between two close points seperated by distance $\ell << \lambda$, where $\lambda = \frac{c}{f}$ is the wavelength. Then over one period it emits energy at the rate:

$$P_{rad} = \frac{\mu_o \omega^4 (q_o \ell)^2}{12\pi c}.$$

One may apply similar considerations to determine radiating effects at large distances from bounded moving charges moving at low speeds, $|\vec{v}| << c$ (for instance the Larmor formula below). To deal with the more general case, we will sketch a different more exact approach and state some relevant formulas.

Recall that we have 'solved' Maxwell's equation using the retarded potentials and the Lorenz gauge. More completely, the fields themselves in this general situation are given by:

Jefimenko's equations: the electromagnetic fields generated by a charge and current distribution, $\rho(x,t), \vec{J}(x,t)$, are:

$$\vec{E}(x,t) = \frac{1}{4\pi\varepsilon_o}\int_{y\in\mathbb{R}^3} \left(\frac{\rho(y,t')}{r^3} + \frac{\rho_t(y,t')}{cr^2}\right)(x-y) - \frac{\vec{J}_t(y,t')}{c^2 r}\ dV$$

$$\vec{B}(x,t) = \frac{\mu_o}{4\pi}\int_{y\in\mathbb{R}^3} \left(\frac{\vec{J}(y,t')}{r^3} + \frac{\vec{J}_t(y,t')}{cr^2}\right) \times (x-y)\ dV$$

with $r = |x-y|$, $t' = t - r/c$.

proof: these formulas are derived directly from the Lorenz gauge potentials, $\varphi(x,t) = \frac{1}{4\pi\varepsilon_o}\int_y \frac{\rho(y,t')}{r}\ dV$ and $\vec{A}(x,t) = \frac{\mu_o}{4\pi}\int_y \frac{\vec{J}(y,t')}{r}\ dV$, with $\vec{E} = -\nabla\varphi - \vec{A}_t,\ \vec{B} = \nabla \times \vec{A}$. $\square$

As a special case, one has the following formulas for the fields generated by a point charge:

Liénard-Wiechert potentials: the scalar and vector potentials for a point charge, q, located at $x(t) \in \mathbb{R}^3$, moving with velocity $\vec{v}(t) = \dot{x}(t)$ are:

$$\varphi(x,t) = \frac{q}{4\pi\varepsilon_o r}\ \frac{1}{\vec{u}\cdot\vec{r}/r^2}$$

$$\vec{A}(x,t) = \frac{\vec{v}(t')}{c^2}\varphi(x,t)$$

where $\vec{r} = x - x(t'), r = |\vec{r}|, t' = t - r/c$ and $\vec{u} = \vec{r} - \frac{r\vec{v}(t')}{c}$. The electromagnetic fields are given by:

$$\vec{E}(x,t) = \frac{q}{4\pi\varepsilon_o r^3}\frac{1}{\gamma^2(\vec{u}\cdot\vec{r}/r^2)^3}\left(\vec{u} + \frac{\gamma^2}{c^2}\vec{r}\times(\vec{u}\times\vec{a}(t'))\right)$$

$$\vec{B}(x,t) = \frac{\vec{r}\times\vec{E}(x,t)}{cr}$$

setting $\gamma^2 := \frac{1}{1-\frac{|\vec{v}(t')|^2}{c^2}}$ (γ is called the *Lorentz factor*).

proof: This follows from the formula for the retarded potentials. The computation is subtle and involved (see eg Feynman §21 for the potential computation). ☐

An alternate way one may write the fields in this point charge case is:

Heaviside-Feynman formulas: the electromagnetic fields generated by a point charge, q, located at $x(t) \in \mathbb{R}^3$ are:

$$\vec{E}(x,t) = \frac{q}{4\pi\varepsilon_o}\left(\frac{\vec{r}}{r^3} + \frac{r}{c}\frac{d}{dt}\frac{\vec{r}}{r^3} + \frac{1}{c^2}\frac{d^2}{dt^2}\frac{\vec{r}}{r}\right)$$

$$\vec{B}(x,t) = \frac{\vec{r}\times\vec{E}(x,t)}{cr}$$

where $\vec{r} = x - x(t'), r = |\vec{r}|$ and $t' = t - r/c$.

proof: A direct calculation shows these fields are the same as those following from the Liénard-Wiechert potentials. See §26-34 in vol. 1 of Feynman's lectures for some applications of these last formulas. ☐

We finish with a formula for radiation by an accelerating point charge,

Larmor formula: if a point charge q, initially at rest, undergoes an acceleration a, it emits energy:

$$P_{rad} = \frac{\mu_o q^2 a^2}{6\pi c}.$$

proof: One may derive this by computing the Poynting vector, $\vec{S}(t,x)$, according to the electromagnetic fields of the Liénard-Wiechert potentials, with $v(0) = 0$, and integrating $\vec{S}(t,x)$, over a sphere of radius $r = tc$ centered on the charges initial position as $r \to \infty$. ☐

The Larmor formula gives a good approximation as well for charges moving at low velocities, $|\vec{v}| << c$. It may be sharpened by carrying out the same steps we just sketched for the Larmor formula starting from the Liénard-Wiechert potentials (the computation in this general case is much more involved) without assumptions on the velocity. One obtains *Lienard's generalization*:

$$P_{rad} = \frac{\mu_o q^2\gamma^6}{6\pi c}\left(|\vec{a}|^2 - \frac{|\vec{v}\times\vec{a}|^2}{c^2}\right)$$

where $\gamma^2 = \frac{1}{1-|\vec{v}|^2/c^2}$ is the Lorentz factor.

One implication of these radiating considerations is the following discrepancy with experimental results [1]. Experimental studies, based on the principles of Maxwell's equations, find an atom consists of a positively charged nucleus and negatively charged electrons orbiting the nucleus at a certain distance.

[1]Electrons were discovered by J.J. Thompson in 1897. The atomic structure, a positive nucleus with negatively charged orbiting electrons was found in 1911 by E. Rutherford.

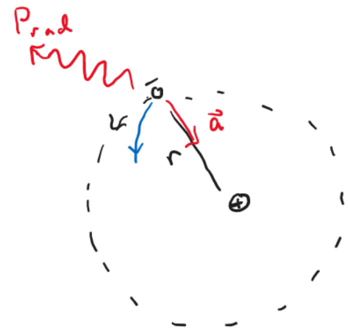

Figure 45. An electron orbiting a proton (nucleus) is accelerated by the Coulomb force from the proton. Hence it radiates, ie loses, energy causing its distance to the proton to decrease. This eventual collision with the proton predicted by the classical electromagnetism is *not* observed, rather the electron remains at a distance from the nucleus.

According to the 'classical' framework, these negatively charged orbiting electrons are accelerating, and hence emitting energy. On the other hand, the negative energy of an orbiting point charge is inversely proportional to its distance from the nucleus. Hence as the electrons emit energy their distance to the nucleus decreases! Maxwell's equations predict they eventually collide with the nucleus, in contradiction to the observed situation.

Let us apply the Larmor formula to estimate the time to collision of a single electron orbiting a single proton (Hydrogen atom).

An electron at distance r orbits the nucleus due to the Coulomb force with strength:

$$f = \frac{e^2}{4\pi\varepsilon_o r^2}$$

where $e \approx 10^{-19}$ C is the (absolute) charge of an electron and proton.

The energy of such a circular motion at radius r is (kinetic plus potential):

$$(*) \qquad W = -\frac{e^2}{8\pi\varepsilon_o r}.$$

Now, the orbiting electron is subject to acceleration:

$$a = \frac{f}{m_e}$$

where $m_e \approx 10^{-32}$ kg is the mass of an electron. Thus, according to the Larmor formula, it radiates or loses energy at the rate:

$$\dot{W} = -\frac{\mu_o e^2}{6\pi c}\left(\frac{e^2}{4\pi\varepsilon_o r^2 m_e}\right)^2.$$

Differentiating $(*)$, and equating to the Larmor formula, the radial distance to the nucleus satisfies:

$$\dot{r} = -k/r^2, \quad k = \frac{e^4}{m_e^2}\frac{\mu_o^2 c}{12\pi^2}.$$

According to experiments, the atomic radius of the Hydrogen atom is $r_o \approx 5 \times 10^{-11}$ m. Integrating, the ode for $r(t)$ with $r(0) = r_o$, we find the time, t_c, until the electron collides with the nucleus is:

$$t_c = -\frac{1}{k}\int_{r_o}^{0} r^2 \, dr = \frac{r_o^3}{3k} = \frac{4\pi^2 m_e^2 r_o^3}{c\mu_o^2 e^4} \approx 10^{-11} s.$$

using $\mu_o \approx 10^{-6} H/m, c \approx 10^8 m/s$ (actually here we used the more precise values found in the 'units' section at the end of this document).

Quite the discrepancy between theory and observation! This 'paradox' of classical electromagnetism is resolved in the setting of quantum mechanics.

See ch. 10, 11 of Griffiths as well as §21 of Feynman on radiation.

Exercises:

1. From the general form, $u(x,t) = f(x-vt) + g(x+vt)$, of solutions to the 1d wave equation, derive d'Alembert's formula:

$$u(x,t) = \frac{u_o(x-vt) + u_o(x+vt)}{2} + \frac{1}{2v}\int_{x-vt}^{x+vt} \dot{u}_o(y)\ dy$$

for a solution with initial conditions $u(x,0) = u_o(x), \partial_t u(x,0) = \dot{u}_o(x)$.

2. (a) If $u(x,t)$ is a solution to the 1d-wave equation, show that so too are $u(-x,t), -u(x,t), -u(-x,t)$.

(b) Consider a semi-infinite string with a fixed endpoint: solutions $u(x,t)$ to the wave equation on $x \geq 0$ with boundary condition $u(0,t) = 0$. Describe the solution[1] with initial conditions $u(x,0) = u_o(x), \partial_t u(x,0) = \dot{u}_o(x)$ on $x \geq 0$ with $u_o(0) = 0, \dot{u}_o(0) = 0$.

(c) Sketch the evolution of a solution to (b) with initial condition $u_o(x)$ a bump function in $x > 0$ and $\dot{u}_o(x) = 0$.

3. Determine the formula for the Laplacian in spherical coordinates, (r, θ, φ).

4. Let $\vec{A}(x,t) = \frac{\mu_o}{4\pi}\int_{y\in\mathbb{R}^3} \frac{\vec{J}(y,t-r/c)}{r}\ dV$, $\varphi(x,t) = \frac{1}{4\pi\varepsilon_o}\int_{y\in\mathbb{R}^3} \frac{\rho(y,t-r/c)}{r}\ dV$ where $r = |x-y|$ and $c = \frac{1}{\sqrt{\mu_o\varepsilon_o}}$. Show that if $\nabla\cdot\vec{J} = -\partial_t\rho$ (charge conservation), then:

$$\nabla\cdot\vec{A} = -\mu_o\varepsilon_o\partial_t\varphi$$

so that the Potentials satisfy the condition to be a Lorenz gauge.

5. Let $f(x), x\in\mathbb{R}$ be a smooth function vanishing outside a compact interval. For $\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ikx}\ dx$ the Fourier transform of f, show that $ik\hat{f}(k)$ is the Fourier transform of $f'(x)$.

6. Consider the 1d wave equation, $v^2 u_{xx} = u_{tt}$ and let $\hat{u}(k,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u(x,t)e^{-ikx}\ dx$ be the Fourier transform of a solution u.

(a) Show that $\hat{u}$ satisfies:

$$-(vk)^2\hat{u} = \hat{u}_{tt}.$$

(b) Deduce from (a) that $\hat{u}(k,t) = a(k)e^{ivkt} + b(k)e^{-ivkt}$ for some functions $a(k), b(k)$.

7. (a) Let $u_o(x) = u(x,0), \dot{u}_o(x) = \partial_t u(x,0)$ be the initial conditions of a solution to the 1d wave equation with Fourier transforms $\hat{u}_o, \hat{\dot{u}}_o$. Using that $u(x,t) = \int_{-\infty}^{\infty}\hat{u}(k,t)e^{ikx}\ dk$ deduce:

$$u(x,t) = \int_{-\infty}^{\infty}\left(\hat{u}_o(k)\cos(vkt) + \frac{\hat{\dot{u}}_o(k)}{vk}\sin(vkt)\right)e^{ikx}\ dk$$

(b) Show the integrals in (7a) may be written as d'Alembert's formula.

8. In spherical coordinates, $r = |x|$, on $\mathbb{R}^3$, and $k\in\mathbb{R}$ a constant, consider the function:

$$u(r) := \frac{e^{-ikr}}{4\pi r}.$$

Show that

$$\int_{\partial B_R(0)} \nabla u\cdot\nu\ dA + \int_{B_R(0)} k^2 u\ dV = -1$$

for any solid ball of radius R centered at the origin [2].

[1] Suggestion: use the symmetry of part (a) to extend the initial conditions as odd functions on the whole x-axis.
[2] This establishes that $\Delta u + k^2 u = -\delta_o$.

9. Let $\hat{u}(x,k) := \frac{1}{2\pi}\int_{-\infty}^{\infty} u(x,t)e^{-ikt}\,dt$ be the (time-coordinate) Fourier transform of $u(x,t)$.

   For $u(x,t)$ (with $x \in \mathbb{R}^3$) a solution of the 3d-wave equation, $\Delta u - \frac{1}{v^2}u_{tt} = -s$, show that $\hat{u}(x,k)$ satisfies:

$$\Delta\hat{u} + \left(\frac{k}{v}\right)^2 \hat{u} = -\hat{s}.$$

10. Use the results of the previous two exercises to derive that

   (a) $\hat{u}(x,k) = \int_{y\in\mathbb{R}^3} \frac{\hat{s}(y,k)e^{-ikr/v}}{4\pi r}\,dV$ with $r = |x-y|$, is a solution to

$$\Delta\hat{u} + \left(\frac{k}{v}\right)^2 \hat{u} = -\hat{s}.$$

   (b) $u(x,t) = \int_{y\in\mathbb{R}^3} \frac{s(y,t-r/v)}{4\pi r}\,dV$ with $r = |x-y|$, is a solution to

$$\Delta u - \frac{1}{v^2}u_{tt} = -s.$$

11. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous (and not necessarily differentiable function). For $\varphi : \mathbb{R}^2 \to \mathbb{R}$ smooth and vanishing outside some compact set, show that [1]:

$$\int_{(x,t)\in\mathbb{R}^2} f(x-t)\left(\varphi_{xx}(x,t) - \varphi_{tt}(x,t)\right)\,dxdt = 0.$$

12. Let $\mathcal{C}_1, \mathcal{C}_2$ be two closed curves in $\mathbb{R}^3$.

   (a) Suppose a steady current I_1 runs in $\mathcal{C}_1$, producing a magnetic field $\vec{B}_1$. Let Φ_{12} be the flux of $\vec{B}_1$ through $\mathcal{C}_2$. Show that there is a constant, L_{12}, such that:

$$\Phi_{12} = L_{12}I_1.$$

   (b) Let L_{21} be the constant of proportionality associated to running a current I_2 in $\mathcal{C}_2$ and the resulting magnetic flux through $\mathcal{C}_1$. Show that:

$$L := L_{12} = L_{21}.$$

   (this constant is called the *mutual inductance* of the pair of loops).

[1] This establishes that $f(x-t)$ (or $f(x+t)$) with f not necessarily differentiable may be considered still as 'weak' solutions to the 1d wave equation (for $v = 1$).

§11 Circuits

See for example §22 of Feynman vol. 2, or ch. 4, 8 of Purcell.

§12 Optics

See for example §26-36 of Feynman vol. 1, or ch. 9 of Griffiths.

§13 Relativity

In this section we consider how Maxwell's equations (1864) motivate [1] the modification of classical (Galilean) relativity (1632) to (Einstein) special relativity (1905).

Upto this point, we have (or should have) always prefaced stating our equations that we are 'in an inertial reference frame'. A reference frame $(x,t) \in \mathbb{R}^3 \times \mathbb{R}$ is called inertial if a particle placed at $x_o \in \mathbb{R}^3$ at rest and subject to no external forces remains at rest. Intuitively, one may think of an inertial frame as coordinates associated to a 'good' laboratory in which one performs experiments, such that the effects of external influences are negligible and the results reproducible or 'universal'.

A fundamental difference between classical relativity and special relativity is that in classical relativity velocities are additive, whereas in special relativity:

the speed of light (electromagnetic waves) in vacuum is constant in all inertial reference frames.

Nowadays, one might be accustomed to accept this claim, having heard it various times. However it is important to realize that it is a very bold and non-intuitive statement.

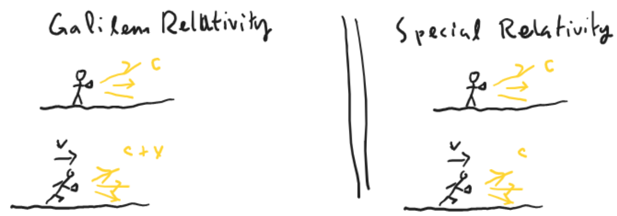

Figure 46. In Galilean relativity velocities are additive. In special relativity, the velocity of light (in vacuum) is always constant.

Now that we have spent time studying Maxwell's equations, the statement seems apparent. After all, in an inertial frame, one experiments with charges/currents/magnets to determine the constants ε_o, μ_o and so the speed:

$$c = \frac{1}{\sqrt{\varepsilon_o \mu_o}}$$

of electromagnetic waves (in vacuum) in any inertial frame.

However, abandoning such long held principles (eg the additive property of velocities) requires strong arguments and supporting evidence. After all, the additive principle of velocities fits our intuition, and the framework of classical mechanics upon which it was built has been used successfully since its more or less formal introduction by Galileo in the 1600's. At the time, it is reasonable to assume the classical framework of mechanics was sound, and one merely needed to 'understand' or formulate more precisely the relatively new (1864) electromagnetic theory of Maxwell's equations to fit the classical framework.

Before exploring the framework and some consequences for the special theory of relativity, we will first explain a chain of ideas which lead to accepting that it is the classical (Galilean) theory of relativity and not Maxwell's equations which should be modified. Let us imagine the following conversation between us (**A**), inquisitive and who have just learned Maxwell's equations, and a late 19th century physicist (**B**).

First, we propose essentially the argument we have outlined above. Namely, we have deduced Maxwell's equations from careful experiments involving measuring forces produced by charges/currents/etc. In particular we produce the constants ε_o, μ_o and so speed c for propagation of electromagnetic waves. There seems to be no reason to believe that anyone else performing the same careful experiments in an inertial frame should arrive at different results. Thus, c should be the same in all inertial frames.

[1] There is a famous quote of Newton, *"If I have seen a little further it is by standing on the shoulders of giants."* regarding his work. When Einstein was asked if he 'stood on the shoulders of Newton', he replied, *"I stand not on the shoulders of Newton, but on the shoulders of James Clerk Maxwell."*

B: This argument is not sound, you are assuming that Maxwell's equations are 'laws of physics'.

A: How so?

B: *Laws of physics are the same in all inertial frames.* However, there are many well understood physical phenomena obeying wave equations. Such phenomena are only described the most simply in *certain* inertial frames. For example the propagation of sound [1]. The speed of sound in air *at rest* is constant, however if an observer moves relative to the air the speed of sound they measure will be changed. The phenomenon is well understood, and in accordance with hearing diffeent pitches from moving objects emitting sound.

A: I am embarrassed to ask, but what is an inertial reference frame? I have always felt iffy on this concept.

B: It is a subtle concept. *An inertial reference frame is one in which the laws of physics hold.* That is $F = ma$, where m, the mass of the object is a constant and F are the forces applied to the object while a is the resulting acceleration of the object.

A: This first claim seems circular: the laws of physics hold in all inertial frames, and an inertial frame is one in which the laws of physics hold? Which are we defining?

B: The laws of physics were stated by Newton. We use them to define inertial frames. These laws have been confirmed by countless careful experiments and groundbreaking and accurate predictions. Why doubt them?

A: That's true, I like classical mechanics and am impressed by its predictions, but still, can't we say the same for Maxwell's equations? That is, they have been confirmed in many experiments and made many profound predictions?

B: Well yes. However they have withstood less of a 'test of time'. However it is always possible that our 'laws of physics' will need to be updated. Physics is a continual process to describe phenomenon by the simplest possible principles.

A: So now I am confused, if Maxwell's equations are not 'laws of physics', what are they? And why are they so useful and in agreement with experiments?

B: Consider again the sound analogy. Sound obeys a wave equation in a frame of reference for which the air is *stationary*. The Maxwell equations, and their resulting electromagnetic waves, obey the corresponding wave equation in which the medium, which we call the *aether*, through which they are propagating is stationary. The Maxwell equations agree closely with experiments because the speed of their propagation, $c = \frac{1}{\sqrt{\varepsilon_o \mu_o}}$, in the stationary aether is so large. Thus, as we move at our common 'everyday' speeds, we do not detect any discrepancies.

A: Ah, that certainly seems to sort things out. However I am quite curious about this 'aether'. Haven't we established that light is an electromagnetic wave, and light propagates across the vacuum of space? So aether is a medium that fills the vacuuum? Also, we on the earth are certainly moving around the sun. Hasn't anyone done experiments to determine our movement through this aether? How would Maxwell's equations change for an observer moving through this aether?

B: We don't really know much 'what' this aether is. The concept of vacuum is rather strange. If we vacate a region of all material, the region is still there, so this aether might be thought of as the material which consti-

[1]**B** would be aware that this is only serves as a useful analogy. A fundamental difference between sound waves and electromagnetic waves is that sounds waves (oscillations in air pressure) are *longitudinal*: the oscillations in pressure occur in the same direction of their direction. Whereas electromagnetic waves are *transverse*: the oscillations in the electromagnetic fields occur in perpendicular directions to their direction. The main point still stands: Maxwell's equations are compatible with Galilean relativity provided we assume there is some medium through which electromagnetic waves propagate.

tutes space itself, although honestly I admit we have really just named something we don't yet understand. As for your second question, what an excellent idea to test by experiment! However the implementation of such an experiment will be difficult, since the speed of light is so large and difficult to measure precisely.

At this point in the 'conversation', we will assume this experiment [1], has been carried out, and the conversation resumes.

A: Did you hear of Michelson-Morley's recent experimental results? What does this imply for Maxwell's equations and the aether?

B: The experiment is impressive, rigorous and precise. The situation seems to be like that of sound when one is travelling in a closed environment. Namely, the medium (air) is carried along with you by friction. Thus for example as one is travelling rapidly in a closed car, the air around you remains stationary. This aether seems then to exhibit similar properties.

A: But wait, we are not in a closed car or closed environment? So this aether has a certain 'stickiness' to it that drags it along with us?

B: Indeed.

A: But then if the aether is a continuous medium 'dragged' along with us as we move, it is like a fluid. As we move, we create ripples in this 'aether fluid'. Wouldn't this produce all kinds of strange optical effects for the light waves propagating through it? Come to think of it, wouldn't this 'drag' take energy from planetary orbits causing them to slowly fall into the sun?

B: Indeed these are serious issues. Fluid mechanics problems such as these are complicated, with complicated equations. Many are working on it, but no-one has come up with a satisfactory theory with predictions that have been verified or fit experiments. There has been a simpler proposal [2] to explain the results of the Michelson-Morley experiment: as objects move through the aether, they are subject to a pressure, contracting their length. Intuitively the drag of the aether compresses objects. Since the speeds of light in the Michelson-Morley experiment where derived by measuring distances, Lorentz has determined that the compressed length, L', of an object initially of length L and moving with velocity v along its axis would be:

$$L' = \sqrt{1 - v^2/c^2}L,$$

in order to fit the results of the Michelson-Morley experiment to motion through the aether.

A: What a strange equation, as we move our length contracts, apparently due to drag on us by this aether. It's quite curious, such terms also end up appearing in the Liénard-Wiechert solutions of Maxwell's equations. Also once $v = c$, the length is 'squashed' to zero and with $v > c$ the formula gives complex numbers!

B: It is a strange equation. However it was derived only to describe a discrepancy in one experiment. Perhaps we should be careful to trust it too far for now. There remains the possibility that one might work out a classical 'fluid theory of aether' to explain these experimental results, including this Lorentz equation as a 'linearized' limiting case.

A: How is that going?

B: It's complicated...

[1] The Michelson-Morley experiment (1887) attempted to measure the speed of the earth relative to this 'aether' by measuring the speed of light on the earth in various directions. It was found to be constant!

[2] H. Lorentz, *Michelson's interference experiment.* The Principle of Relativity. Dover Books on Physics. June 1 (1952): 1-7.

At this stage that one admits, via an 'Occam's razor' type argument, that it is simpler to proceed on the assumption that Maxwell's equations are true 'laws of physics', with the striking acceptance that the speed of electromagnetic waves (light) in vacuum is a constant of nature.

Implicit in the discussion above is the relation:

An 'inertial reference frame' is one in which the 'laws of physics' are valid.

A 'law of physics' is that which is valid in all 'inertial reference frames'.

Thus describing laws physics go hand in hand with describing inertial reference frames. The notion of inertial reference frame is closely tied to that of symmetry, in the sense that the reference frames which are called inertial are related by certain coordinate transformations: a group of coordinate transformations under which the equations giving these 'laws of physics' remain the same.

In this way, one arrives to the fundamental assumption of special relativity: that the speed of light (electromagnetic waves) in vacuum is the same in all inertial frames. We finish this section by examining some more precise situations.

Length contraction: we consider a simplified description of the Michelson-Morley experiment to derive the relativistic formula for (length) *Lorentz contraction*. The situation is as follows. We are an observer moving at constant velocity v through the stationary aether. How might we measure v?

Figure 47. To measure our speed relative to the aether, we may carry with us a device to measure the speed of light in various directions.

We construct two perpendicular 'arms' of identical lengths L. Then we reflect light signals along these arms and measure the time it takes for the light signals to return. By rotating the arms we will obtain various times $\Delta t', \Delta t$, for the time it takes light to reflect back and forth across the arms.

When the first arm is alligned with our direction of motion, we would measure a time:

$$\Delta t' = \frac{L}{c-v} + \frac{L}{c+v} = \frac{2L}{c(1-\frac{v^2}{c^2})} = \frac{2L}{c}\gamma^2$$

for the light to return along the 'horizontal' arm, where

$$\gamma := \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

is the *Lorentz factor*.

Along the 'vertical' arm, we would measure a time:

$$\Delta t = \frac{2L}{c}\gamma.$$

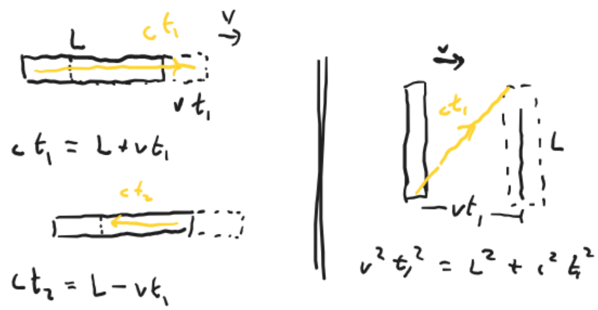

Figure 48. We may measure our speed relative to the aether by measuring the time it takes light to bounce back and forth between the two arms of the apparatus of the Michelson-Morley experiment. In the actual experiment, it was arranged so that any difference between the times would be observed as an interference pattern between the two 'bouncing' light rays.

Hence, according to light propagating at constant speed through the *stationary* aether, we could measure our velocity v through this medium by solving for v based on the measurements of $\Delta t', \Delta t$.

This impressive experiment was performed by Michelson-Morley, with exactly this goal in mind: aiming to determine our speed as observers on the moving earth as we mover through this aether. The result of the experiment was negative: no discrepancy between $\Delta t', \Delta t$ was found! Refinements of this experiment form the basis for the most precise measurements of the speed of light yet obtained [1].

The following 'classical' explanation for these results was proposed by Fitzgerald (1889) and Lorentz (1892). An object moving through the aether is imagined to be subject to a pressure, compressing its length. For the times measured in the Michelson-Morley experiment to be the same, this length contraction should be:

$$L' = \sqrt{1 - v^2/c^2}L = \frac{L}{\gamma}$$

where L' is the compressed length of the bar when moving with velocity v through this aether medium, and L its length 'at rest', or moving perpendicular to the medium.

Next, we consider a 'paradox' in electricity and magnetism explained by Lorentz contraction.

Electromagnetism and length contraction: consider a (straight line) neutral wire carrying current. The wire is neutral, so we consider it as a linear density of equal positive and negative charges, $\lambda_+ = -\lambda_-$, with say the positive charges moving with velocity $\vec{v}$ along the wire generating the current $\vec{I} = \lambda_+\vec{v}$.

Now, a point charge q moving initially with the same velocity $\vec{v}$ parallel to the wire, will experience a magnetic force,

$$\vec{f_1} = q\vec{v} \times \vec{B}$$

due to the current in the wire generating the magnetic field $\vec{B}$.

If we now view the situation in a reference frame moving with velocity $\vec{v}$ with the charge, q, the charge will be initially at rest. The positive charges will as well be at rest, however the negative charges will now appear to be moving (in the direction $-\vec{v}$), producing a current and magnetic field. However the wire appears to still be neutral, and the charge has no initial velocity. Yet, provided the laws of physics are the same in all reference frames, the charge should still move as it did in the other frame of reference!

What force could be causing the charge to move? According to our analysis above, the wire is still neutral, so there is no electric field, and the charge is not initially moving, so according to the Lorentz force, the charge experiences no force and remains stationary in this moving frame.

[1]Eventually, such experiments became so precise, that it was the definition of the meter that produced the most error. Today the meter is defined by taking the speed of light (in vacuum) as given by Maxwell's equations, and the meter defined in terms of this speed and a definition of a unit of time.

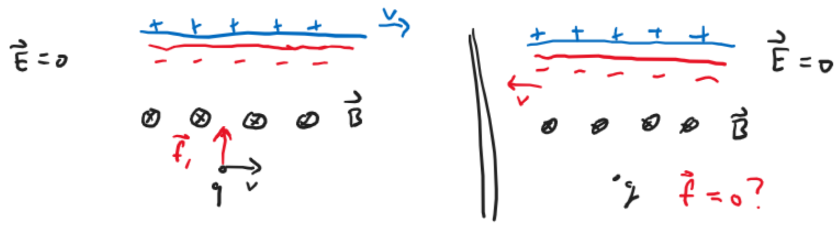

Figure 49. In one frame of reference, a charge moving with velocity $\vec{v}$, is attracted to a current carrying wire due to the magnetic force. The charge falls towards the wire. In a frame moving with the charge with velocity $\vec{v}$, the charge is at rest, apparently the wire is still neutral and there is only a magnetic field present, so no force on the charge, and the charge stays fixed.

We may explain this 'paradox' – that the point charge q moves towards the wire in the stationary frame, whereas remains fixed in the moving frame– by considering length contraction.

In the reference frame moving with the charge, the positive charge density is dilated to the density:

$$\lambda'_+ = \frac{\lambda_+}{\gamma}$$

while the negative charges are contracted to the density:

$$\lambda'_- = \gamma\lambda_-.$$

Thus although the wire was neutrally charged in the reference frame at rest with the wire, in the moving frame, it has charge:

$$\lambda' = \lambda'_+ - \lambda'_- \neq 0$$

and generates an electric field, causing the charge q initially at rest in the moving frame to be as well attracted to the wire.

Before considering some more general situations, let us state another consequence of Maxwell's equations in the previous example.

Time dilation: in the previous example, we have qualitatively explained why the change of reference frame explains the movement of the charge q towards the wire. Let us consider more precisely the resulting forces.

Let d be the initial distance from the point charge q to the wire.

In the first reference frame, the moving point charge q is subject to a (magnetic) force of magnitude:

$$f_1 = q\frac{\mu_o\lambda_+v^2}{2\pi d}$$

since the wire carries current $I = \lambda_+v$.

On the other hand, in the reference frame moving with the point charge q, it is subject to an initial force of magnitude:

$$f_2 = q\frac{|\lambda'|}{2\pi\varepsilon_o d}$$

where $\lambda' = \lambda_+(\frac{1}{\gamma} - \gamma)$, and $\gamma = \frac{1}{\sqrt{1-v^2/c^2}}$. These forces are *not* the same, one computes:

$$f_2 = \gamma\ f_1.$$

Similarly to the Lorentz length contraction, this discrepancy between forces may be balanced by assuming that the mass of an object and the time intervals one measures depend on ones velocity, according to:

$$m_1 = \gamma m_2, \quad dt_1 = \gamma dt_2.$$

The quantities m_1, t_1 are of the point charge q in the reference frame where q moves with velocity v, while m_2, t_2 in the reference frame where q is (initially) at rest.

At this point, these rescalings of mass and time should seem ad hoc. For instance why do we not only rescale mass and not time? Or just time and not mass?

However it is clear that once the rescaling of time is determined, so too is the required corresponding rescaling of mass in order that Maxwell's equations continue to hold in the moving frame.

To understand the rescaling of time (and resulting rescaling of mass) we have proposed above, consider the following situation. We use the constancy of c from Maxwell's equations to build a clock: time is measured by bouncing a beam of light between a known distance L.

FIGURE

Our clock measures time intervals $\Delta t = \frac{2L}{c}$. If we are moving at a constant velocity, $\vec{v}$, relative to another observer, then *they* observe our clock to measure time intervals:

$$\Delta t' = \frac{2L}{\sqrt{c^2 - v^2}} = \gamma \Delta t.$$

Since the length L we use to construct such a clock is arbitrary, this relation holds between arbitrary time intervals. In summary, Maxwell's equations have led us to the rescalings:

$$\gamma L' = L, \quad \Delta t' = \gamma \Delta t, \quad m' = \gamma m$$

of lengths, time intervals, and masses for objects moving relative to us.

Now we will consider in more detail inertial frames, and the consequences of the Michelson-Morley experiment.

Lorentz transformation: First, we consider an example of the classical Galilean notion of relativity in coordinates.

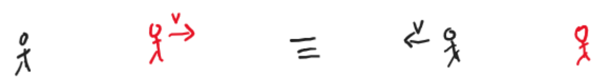

Figure 50. In Galilean (and special) relativity, two observers moving at constant velocity respect to one another are both equally valid to take the point of view that they are fixed, and the other is moving.

Suppose we, an observer, are at the origin of an inertial frame of reference. Another observer, who is moving relative to us at constant velocity $\vec{v}$ may also consider themselves as the origin of an inertial frame of reference. Let us take our x-axis along the motion $\vec{v}$, and $v := |\vec{v}|$ as the relative speed between us.

Assume that when we coincide, we synchronize our clocks. Then in classical relativity, our position and time measurements, (x, t), are related to those of the other observer (x', t') by the following (Galilean) transformation:

$$x' = x - vt, \quad t' = t.$$

As an example of a law of physics, we may consider Hooke's law for a spring, which in our frame reads:

$$\ddot{x} = -k(x - x_o)$$

where x_o is the equilibrium position of the spring. For the other observer, the equilibrium position of the spring is time dependent, at $x'_o = x_o - vt$. However the displacement they see from the equilibrium is the same: $x' - x'_o = x - x_o$. As well the acceleration they observe is the same: $\ddot{x}' = \ddot{x}$, since $\ddot{vt} = 0$. Hence the observer moving relative to us observes the same law of physics for the spring:

$$\ddot{x}' = -k(x' - x'_o).$$

Now, we apply similar considerations according to the Michelson-Morley experiment that when an observer moves with velocity v 'through the stationary aether', they observe electromagnetic waves to obey the same wave equation:

$$c^2\ u_{xx} = u_{tt}$$

as we do. If one performs the Galilean transformation, $x' = x - vt, t' = t$, the wave equation becomes:

$$(c^2 - v^2)\ u_{x'x'} = u_{t't'} - 2v\ u_{x't'}$$

and evidently, an observer moving with constant velocity relating his coordinates to us via a classical Galilean transformation will not observe the same speed of these waves. We ask then, how may the transformation relating our coordinates to an observer moving at speed v relative to us be given so that the wave equation remains the same? We seek linear transformations,

$$x' = Ax + Bt, \quad t' = Cx + Dt$$

preserving the wave equation, ie $\partial_x^2 - \frac{1}{c^2}\partial_t^2 = \partial_{x'}^2 - \frac{1}{c^2}\partial_{t'}^2$, and find the conditions:

$$c^2A^2 - B^2 = c^2, \quad D^2 - c^2C^2 = 1, \quad c^2AC = BD.$$

Now, if these new coordinates $((x', t'))$ represent an observer moving with velocity v relative to us $((x, t))$, then we impose:

$$x = vt \Rightarrow x' = 0$$

which implies $B = -Av$. From this condition, the equations above may be solved, yielding the *Lorentz transformation*:

$$x' = \gamma(x - vt), \quad t' = \gamma(t - \frac{v}{c^2}x)$$

where $\gamma = \frac{1}{\sqrt{1-v^2/c^2}}$. Note that we recover time dilation and length contraction by considering differences. As well, when $v << c$, the Lorentz transformation closely approximates the Galilean transformation.

Poincaré group: we have found above a special case of a Lorentz transformation (a change of coordinates preserving the wave equation with wave velocity c). All such transformations, preserving the wave equation, form a group. They may be described efficiently using linear algebra. It is convenient to normalize out the constant velocity c, by considering the transformations to act on ct rather than t.

First, we observe by chain rule, that a linear transformation:

$$\begin{pmatrix} x' \\ ct' \end{pmatrix} = L \begin{pmatrix} x \\ ct \end{pmatrix}$$

preserves the wave equation, $\Box u = 0$, exactly when the matrix L^T preserves the quadratic form:

$$Q(x, ct) = |x|^2 - c^2t^2.$$

That is, in matrix form, if we represent Q by the matrix, $\begin{pmatrix} I & 0 \\ 0 & -1 \end{pmatrix}$, then:

$$LQL^T = Q.$$

Including translations, $(x, ct) \mapsto (x + x_o, c(t + t_o))$, such transformations form a group, the *Poincaré group*. More compactly, we may describe them as the affine transformations of a 4-dimensional space, $\mathbb{R}^{3,1} \ni (x, ct)$, the *Minkowski space*, which preserve the Minkowski inner product:

$$\langle (x_1, ct_1), (x_2, ct_2) \rangle = x_1 \cdot x_2 - c^2t_1t_2.$$

Such was the structure in place around the end of the 19th century as developed by scientists such as Lorentz, Larmor, and Poincaré. To appreciate the contribution of Einstein [1], we remark that these transformations and rescalings were still thought of as corrections caused by the movement through the stationary aether. That is the underlying framework was still classical relativity, and these correction terms were merely seen as necessary ad hoc additions to add in order to fit the experimental results, describing some sort of linearized version of the true effects from movement through the aether.

The point of view taken by Einstein may be considered a more practical and aesthetic approach. The goal of a physical theory, is to propose a model whose predictions are compatible with experiments. Ideally, the assumptions of the model should be simple. Now, we may observe that all the correction terms, which until this point had been assumed to be causes of this aether medium, have been derived essentially by requiring that the wave equation (or more strongly Maxwell's equations), remain invariant. Einstein proposed that we take this as our starting point, rather than classical Galilean relativity, and showed how all the above rescalings may be derived from the two axioms:

1. *A reference frame moving at constant velocity relative to an inertial frame is inertial.*
2. *The speed of light is constant in all inertial reference frames.*

The resulting theory is called Einstein's *special relativity.* It has as consequence the Lorentz transformations and Poincaré group, relating inertial reference frames. The classical description of an inertial frame as a reference frame for which an object subject to no forces (a free particle), moves in a straight line is still used. These are the only axioms of the theory, and whether or not there is some aether medium as their cause is irrelevant to make predictions fitting experiments.

We comment as well, that Einstein also contributed the relativistic formulas for energy and momentum:

$$\vec{p} = \gamma m_o \vec{v}, \quad E^2 = |c\vec{p}|^2 + (m_o c^2)^2$$

where m_o is the rest mass of the particle, and $\vec{p}$ is its (relativistic) momentum when moving with velocity $\vec{v}$. Note, if an object is not moving, $\vec{v} = \vec{p} = 0$, its energy is given by the famous:

$$E = m_o c^2.$$

Let us finish with a brief more abstract description of the setting of classical and special relativity.

In classical relativity, the space-time is a 4-dimensional affine space, $\mathcal{M}$, and time is a one dimensional (directed) affine line, $\mathcal{T}$. Time is *universal*, in the sense that there is only one time line. We have:

$$\mathbb{E}^3 \to \mathcal{M} \to \mathcal{T}$$

where the Euclidean space fibers consist of events taking place occuring at a given time.

A free motion is a line in the affine space $\mathcal{M}$, projecting onto $\mathcal{T}$ (the position of the particle at each given time).

The laws of physics are Newton's equations $F = ma$, where F is a force law invariant under the Galilean group: the set of affine transformations of $\mathcal{M}$ which preserve events taking place at the same times, and the distances on the Euclidean spaces of events at each given time.

In special relativity, the space-time is a 4-dimensional affine space, $\mathcal{M}$, and there is no universal time line. On the space time, $\mathcal{M}$, is a Minkowski metric:

$$\mathcal{M}, \|\cdot\|$$

A free motion is a time-like line in the affine space $\mathcal{M}$, ie $\|m_1 - m_2\|^2 < 0$ for any two points on the line. An object moves at the speed of light c if it moves along a null line, $\|m_1 - m_2\| = 0$, for any two points on the line.

The laws of physics are those which are invariant under the Poincaré group: the set of affine transformations of $\mathcal{M}$ which preserve the Minkowski norm. Maxwell's equations are invariant under the Poincaré group, so may be considered as the first relativistically correct laws of physics.

[1] A. Einstein, *On the electrodynamics of moving bodies.* Annalen der physik 17.10 (1905): 891-921.

§14 Bundles

The equations of electromagnetism may be formulated concisely using differential forms.

See for example, part 2 of Baumberg and Sternberg's *A course in mathematics for students of physics*. Or as well the more advanced [1] [2].

In this form, Maxwell's equations may be written on manifolds, and in a way compatible not only with special relativity but also general relativity (the *Kaluza-Klein theory*). See for example ch. 12 (12.2.2) of R. Montgomery's *A tour of sub-Riemannian geometry.*

[1] S. Sternberg, *On the role of field theories in our physical conception of geometry.* Differential Geometrical Methods in Mathematical Physics II. Springer, Berlin, Heidelberg, 1978. 1-80.

[2] V. Guillemin, S. Sternberg. *Symplectic techniques in physics.* Cambridge university press, 1990.

§15 Experiments

We will describe in this section some physical considerations and important experiments.

Measuring mass and force: Suppose we have two objects and wish to compare their masses, m, M. According to Newton's 2nd law, if we apply the *same* force, f, to each object we may then measure their resulting accelerations, a, A and will then have:

$$ma = f = MA \Rightarrow \frac{a}{A} = \frac{M}{m}.$$

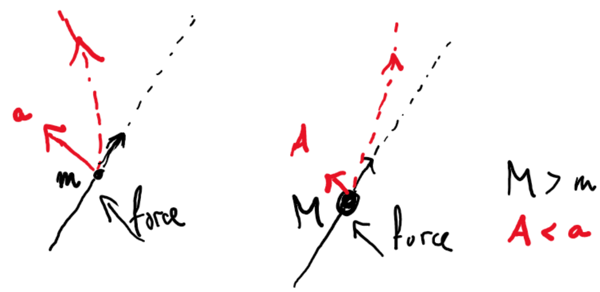

Figure 51. Comparing masses of objects.

Thus for instance if the mass m accelerates twice as fast as the mass M, then the second object is twice as massive as the first object: $M = 2m$.

The mass of any object may be measured in units by electing a standard 'unit mass' object with which to compare. The *kilogram* is defined classically as the mass of one liter (1000 cubic centimeters) of water.

Note that essential to measuring mass is exerting objects to the *same* force. This is done in practice by preparing –as best one can– identical 'laboratory' conditions in which to place the objects. For instance, one may hold the objects at a given fixed position over the surface of the earth (if the objects are placed on a lever, mass ratios relate to distance ratios for the balance point), or one may attach both objects to identical springs stretched to equal lengths.

Once one may measure masses of objects, measuring forces essentially follows by their definition in Newton's 2nd law. Namely, to measure the force present at a given point, one places an object of known mass m at this point and measures the resulting acceleration, $\vec{a}$. Then $\vec{f} = m\vec{a}$. The standard units of force are thus expressed in $kg \cdot m/s^2 = N$, which are called *Newtons*.

Measuring mass and force requires measurements of accelerations, ie lengths and times. Intuitively, we have a good sense of how length and time measurements may be defined. Lengths are measured by comparison with some elected standard or 'unit length'. In this way the *meter* is classically defined as 10^{-7}'th the distance from Earths' equator to north pole through the latitude passing through Paris. The good folks in Paris, to save us the trouble of long voyages, then produced a standard bar of this length and replicas of this are then kindly distributed around the world as rulers or measuring sticks for everyday use. Likewise time is measured by comparison with some elected standard regular repeating occurence. Classically the *second* is defined astronomically as $1/(24 \cdot 60 \cdot 60)$ of an average day on earth. Clocks are then designed which reproduce this time interval for ones everyday use.

These classical definitions of units of measurement should be intuitive, however as one eventually finds in finer experiments, they are lacking in precision. We will explain some of these motivations for redefining the meter, kilogram and second in a later section [1].

Levers, work, energy: To motivate the definition of 'work' we will consider a lever – a (homogeneous) rigid bar free to pivot about an interior point or 'fulcrum'.

[1] See, here for the modern definitions.

First we ask: when a force f_1 is applied perpendicularly to one end of the lever, what force f_2 applied to the other end will exactly cancel the force f_1, ie result in no movement of the lever? One finds the relation between these strengths is:

$$(*) \quad f_1 r_1 = f_2 r_2$$

with r_j the distances of the ends to the pivot point.

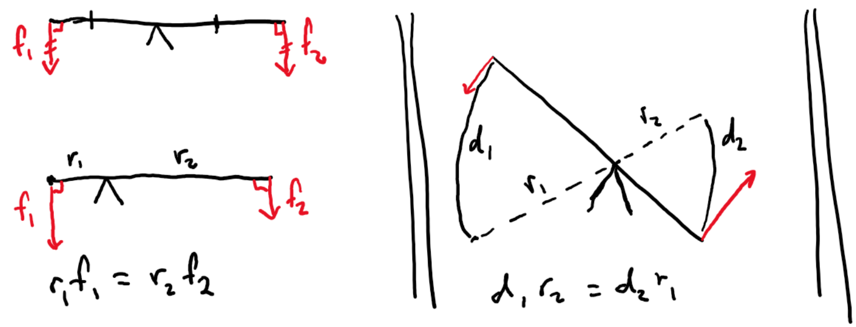

Figure 52. Forces of strength f_j applied perpendicularly to a levers ends 'balance' when $f_1 r_1 = f_2 r_2$. This may be derived by considering 'balance of torques' or as an application of d'Alembert's principle for static equilibrium. When a force f_1 is applied to one end over a distance d_1 resulting in a force f_2 on the other end moving over a distance d_2 one has 'work balance' $f_1 d_1 = f_2 d_2$.

The relation $(*)$ makes the lever a useful tool for 'force magnification'. Indeed, if there is a large force f_2 being exerted at one end of the lever then by exerting any force slightly stronger than $\frac{r_2}{r_1} f_2$ on the other end we may overcome this large force and cause the lever to move –lifting the other end. Placing the fulcrum so that $\frac{r_2}{r_1}$ is very small only a small force need be required to overcome the large force f_2 on the other end.

As well the equilibrium condition $(*)$ is the principle behind using a balance to compare masses of objects. The gravitational force at the surface of the earth on a mass m is $f = mg$ with $g \approx 9.8\ m/s^2$ a constant so that when two masses are placed at the ends of a horizontal lever, the lever will balance when $m_1 r_1 = m_2 r_2$ and so ratios of the masses may be found by measuring ratios of distances to their balance point.

In moving objects with a lever, although one may magnify the strength of ones force applied, it comes at the cost of moving the other end over a different distance. Indeed, if we apply the force f_1 at one end and move this end over a distance d_1 then the other end is subject to a force $f_2 = \frac{r_1}{r_2} f_1$ and moves over a distance d_2 with:

$$d_1 r_2 = d_2 r_1$$

hence if we take $\frac{r_2}{r_1}$ very small to magnify the resulting force at the other end, the distance the the other end moves, $d_2 = \frac{r_2}{r_1} d_1$, will be reduced. Combining these relations we have:

$$f_1 d_1 = f_2 d_2$$

on the two ends of the lever. This relation expresses the balance of 'work' done on the two ends of the lever, as –in general– the *work* done by the force field $\vec{f}$ when an object is moved along an (oriented) curve $\mathcal{C}$ is defined as:

$$W_{\mathcal{C}} = \int_{\mathcal{C}} \vec{f} \cdot T \ ds.$$

The standard units to measure work are then $N \cdot m = kg \cdot m^2/s^2 = J$ and are called *Joules*.

The *energy* of an objects current *state* –its current position and velocity– is the amount of 'stored work' in this state, ie the amount of work required to get the object to this state from a given 'reference state'. The choice of reference state typically amounts to a shift in the energy by a constant. To see that there is 'stored work' in the objects velocity, consider that to accelerate an object from rest to a given velocity v requires applying a force to the object over some distance, ie inputting work. This work required to get the object from rest to its current velocity is called its *kinetic energy*, and one can compute it is given – for a particle with mass m – by:

$$\frac{m}{2}|v|^2.$$

Energy, like work, is measured in Joules.

Foucalt's pendulum: As an illustration of non-inertial reference frames, we will consider the behaviour of a pendulum – a point mass 'bob' is fixed to the end of a string and released from a certain height.

Figure 53. In an inertial frame, the bob of a pendulum subject to vertical forces oscillates back and forth in a fixed vertical plane. For constant vertical acceleration, the angle from the vertical satisfies $\ddot{\theta} = -\frac{g}{\ell}\sin\theta$.

If the earth were an inertial frame of reference, then the only force present would be due to the gravity of the earth and we may apply Newton's laws to see that the bob oscillates back and forth in a *fixed* vertical plane. On the other hand, if the earth were rotating (and not the universe rotating around the earth), one would observe different behaviour. For simplicity, consider that the pendulum is located over the north pole of the earth's rotational axis. Then in the inertial frame 'fixed to the stars' the pendulum would swing in a fixed vertical plane while the earth moves underneath it. Over the time scale of a day an observer with their reference frame 'fixed to the earth' would then observe the vertical plane of the pendulum complete one rotation, 'precessing'.

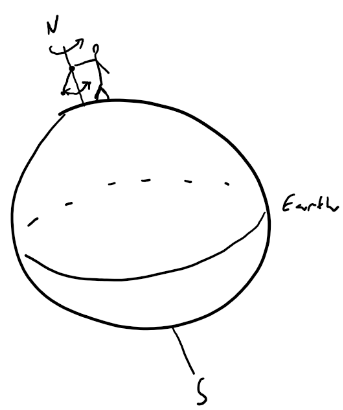

Figure 54. One may test whether a reference frame fixed to the earth is inertial by considering a pendulum over say the north pole. If the earth is rotating then in earth's reference frame, the vertical plane containing the pendulum will be observed to precess. Such effects due to choice of non-inertial reference frames are called *ficititious forces*. See for example this video.

To obtain a more qualitative description –for a pendulum at a general latitude λ– see for example §27 of Arnold's Classical mechanics.

Measuring electromagnetic quantities: We first describe electric *charge*. Qualitatively, charge is a property of an object which causes it to attract or repel other charged objects. One may observe these effects upon rubbing certain materials [1].

A qualitative measurement may be assigned to charge from the following experimental results. It is helpful to imagine the charge of an object as represented by a fluid [2] –similar to how one imagines heat as a fluid. When certain objects are rubbed together, this charge 'fluid' transfers from one object to another leading to an excess of charge on one object and deficit on the other. Certain materials are more susceptible to this transfer of charge through them –conductors (free flow of charge) and insulators (no flow of charge) being the extremes.

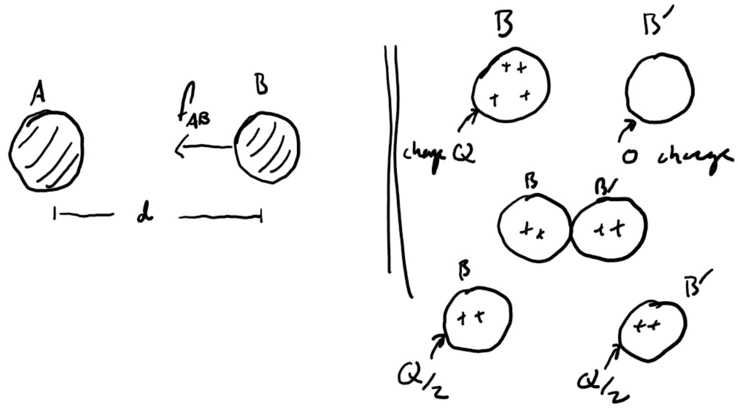

Figure 55. One may measure the force, f_{AB}, between two charged objects located at a given distance. As the charge on one of the objects is varied, one finds the force varies linearly with charge.

Now, to compare charges one may fix a charged object A and place another object B at a given fixed distance from A and measure how the force on B due to A varies as the charge of B is varied. One thus needs a manner in which to vary the charge on B in a precise way. To do this one could take two identical conducting spheres, B and B', where B has been charged –say with charge $Q \in \mathbb{R}-$ and B' has no charge (ie feels no force due to A). If B and B' are placed in contact then by symmetry the charge will flow to equal amounts in both B and B' while still totalling the same amount Q –this is the principle of *charge conservation*– so that one now has two objects each with charge $Q/2$. Similarly one may vary a given charge $Q \to Q/n$ for $n \in \mathbb{N}$ and in this way it is found that the force, f_{AB}, of a fixed charged object A on an object B with charge Q varies linearly in Q.

In this way one establishes that two charged objects located at a *fixed distance* from eachother experience a force of strength

$$f \sim q_1 q_2$$

proportional to the product of their charges. Two charges, q, q', may then be compared by placing them both at say 1 meter from a fixed charged object and measuring the resulting forces, f, f', so that

$$\frac{f}{f'} = \frac{q}{q'}.$$

One might then define a unit of charge, $\hat{q}$, by requiring that when two objects with charge $\hat{q}$ are seperated by a distance of 1 meter they experience a fixed force. The standard unit of charge, the *Coulomb* (C), in this way would be defined as a quantity of charge so that when two charges of 1 C each are seperated by 1

[1]The material amber (elektron in greek) was especially susceptible to charge by rubbing, and is from where the name electricity originates.

[2]The more precise physical description is by the number of elementary charged particles: electrons and protons.

meter (in vacuum) they exert a force of 9×10^9 N on eachother [1]. Finally we emphasize that charges are *signed*, depending on whether they attract (positive) or repel (negative) a given deemed negative charge (eg rubbed amber is called *negative*).

Once one has quantified charge, the electric field at a given point is measured by placing a test charge q at the point and measuring the resulting force, $\vec{f}$, on the test charge with $\vec{E} = \vec{f}/q$ the electric field. Its strength is measured in N/C.

Similarly, the electric potential difference, $\delta\varphi$, between two points x_o, x_1, is measured by first finding the work done by the electric field to move a test charge q from x_o to x_1: $\delta W = \int_C q\vec{E} \cdot d\vec{s}$ where C is a curve from x_o to x_1. In practice, one would measure the work *you* do, $\delta W'$, to move the test charge and take $\delta W = -\delta W'$. A fundamental assumption is that this work is independent of the path. The electric potential difference is then:

$$\delta\varphi := -\delta W/q$$

and is measured [2] in *Volts* (V = J/C). In many situations there is a natural choice of 'base-point' x_o from which to measure potential differences and one speaks then of *the* electric potential at a point as the potential difference between this point and the fixed basepoint.

The ability to measure charges and potential differences, allows one as well to measure *capacitance* of a capacitor. One would place known quantities Q and $-Q$ of charge on the two conductors and measure the resulting potential difference, V, to find:

$$C = \frac{Q}{V}.$$

Capacitance is measured in *Farads* (F = C/ V).

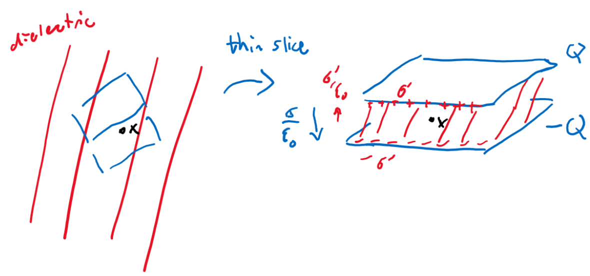

Figure 56. The susceptibility, χ, of a dielectric is related to change in capacitance when a capacitor is filled with the material.

The measurement of capacitance may be used to measure the *susceptibility* of a dielectric material. One may take a thin slice of the material to fill the region between a parallel plate capacitor. When the plates are charged to $Q, -Q$ then without the material there they will have a capacitance $C = \frac{A\varepsilon_o}{d} = \frac{Q}{V}$. However when the material is present between the plates, one will in general measure a different capacitance: $C' = Q/V'$. This potential difference, V', is due to the total electric field, $\vec{E}$, produced by the charged plates and polarization of the dielectric, so that:

$$\vec{E}(x) \cdot \nu \; d \approx V' = \frac{C}{C'}\frac{\sigma d}{\varepsilon_o}$$

$$\Rightarrow \varepsilon_o \vec{E}(x) \cdot \nu \approx \sigma \frac{C}{C'}$$

where ν is the unit normal to the plates (and x an interior point to the plates). On the other hand, this change in capacitance is caused by the accumulation of bound charges, $\sigma' \approx \vec{P}(x) \cdot \nu$, on the surface of the

[1] This is an *enormous* force, and is not the units in which charge was first measured (units of charge were not standardized until rather later). The motivation for this definition of the Coulomb was related to currents. Another unit of charge is the *Franklin* or *statacoulomb* which for 'everyday' charges has less extreme values. Let us also remark that there are more elaborate devices for measuring charge more accurately, called *electrometers*.

[2] A *voltmeter* or *potentiometer* may be used for more accurate measurements of potential differences.

dielectric strip. Thus, as well:

$$\frac{C}{C'}\frac{\sigma d}{\varepsilon_o} = V' \approx \frac{d}{\varepsilon_o}(\sigma - \sigma')$$

$$\Rightarrow \vec{P}(x)\cdot\nu \approx \sigma(1 - \frac{C}{C'}).$$

For isotropic and homogeneous materials, $\vec{P} = \varepsilon_o \chi \vec{E}$, we then have: $\chi \frac{C}{C'} = 1 - \frac{C}{C'}$, so that: [1]

$$C' = (1+\chi)C.$$

Since C and C' may be measured, so too may the susceptibility $\chi(x)$ through this last relation (in general, one would need to take slices in various directions of the material to determine possible isotropy). In the following table we give susceptibilities of some materials (in 'normal' conditions, susceptibility may depend on various factors, eg the temperature).

material	susceptibility (χ)
air	$\approx .0006$
plastic	≈ 4
glass	≈ 6
alcohol	≈ 24
water	≈ 81

Next we consider measurements related to magnetism. Before magnetic effects were related to currents – *unified* with the theory of electricity in *electromagnetism* – one had two seperate theories: one of static electricity and another of static magnetism. Before explaining this connection, let us describe the fundamental ideas of this 'pre-unification' theory of magnetism.

Qualitatively, certain objects may have the property of being *magnetic*: eg they attract certain metals, allign themselves with the north and south (magnetic) poles of the earth, and repel or attract other magnetic objects. Like electric charge, these properties may be thought to be caused by a quantity of 'magnetic charge' of the object that came in two types: North and South. Unlike electric charge, no magnetic substance without equal parts of north and south magnetic charge has yet been found: magnetic monopoles have never been encountered, only magnetic dipoles.

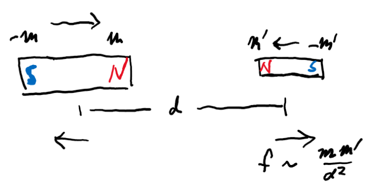

Figure 57. Strengths of magnets may be assigned a quantitative measure based on an inverse sqaure law.

Nonetheless, similarly to the Coulomb law for electric charges, it was found experimentally that magnetic charge also obeyed an analogous inverse square law [2]. From this inverse square law one could measure quantitatively the magnetic charge on the north and south poles of a given magnet. The theory of magnetostatics thus proceeded similarly to the theory of electrostatics, but, with a north and south magnetic charge replacing positive and negative electric charge and –in practice– *only involving dipoles*. So objects could posses three seperate properties: mass, charge, and magnetic charge. Each determining the effects –fields– they produced via inverse square laws.

[1] Alternately, $\varepsilon_o C' = \varepsilon C$.

[2] Also by the same Coulomb: C. Coulomb, *Mémoires sur l'Electricité et le Magnétisme.* Histoire de l'Académie Royale des Sciences 569 (1785).

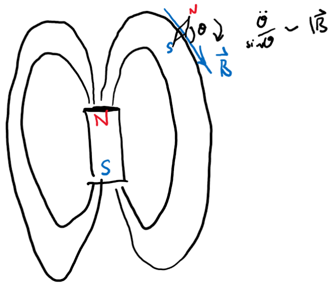

Figure 58. A magnet produces a magnetic field, which may be measured via its torque on a compass: a magnet free to pivot about a fixed point.

A given magnet then produces a magnetic field $\vec{B}$, whose direction at a given point could be measured by placing a unit 'test magnet' (eg a compass) and seeing how the compass alligned its north and south poles. The strength of the magnetic field (in units of N/magnetic charge) could then be measured by measuring the torque on the test magnet (eg how quickly the test magnet alligned with the direction of the magnetic field).

Now, let us explain how the theory of magnetism was unified with electricity through the study of current.

First we describe electric *current*: moving electric charges.

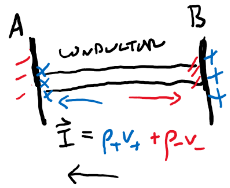

Figure 59. When a charged capacitor is discharged, by connecting the two plates by a conducting wire, there is a short transient current: movement of charges between the plates.

Consider first a capacitor, with charges $-Q, Q$ on the two plates, A and B. If the capacitors are connected by a conducting wire, the free charges in the wire will move to the plates creating –for a short time– a transient current in the wire. After this short time the charges of the conducting wire will redistribute towards a static equilibrium with negative charges collected on plate A and net positive charges on plate B.

To create a lasting current one may connect the wire to a *battery*: a device which moves negative charges from a termimal B to A (as well, one may say the battery maintains A and B at a constant potential difference). Then, when negative charges are transported from B to A by the battery, the conducting wire will respond by redistributing its negative charge to accumulate on plate B, which will then be transported by the battery to A and so on.

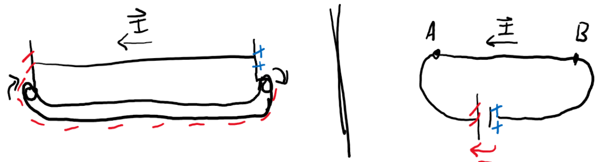

Figure 60. A lasting current may be established using a battery: a device which moves charges from one plate to the other. Alternately, the battery maintains the plates at a constant potential difference. One may have a mechanical battery, using the friction between the plates and a belt to transport charges. More efficient batteries operate using chemical reactions.

Thus in this situation, the charges in the wire are constantly moving in response to the battery. The rate of charge transfer, ie the quantity of charge which moves across a point on the wire per second is the electric current in the wire. Current is measured in *Amperes* (A = C/s).

As we may measure charge and time, as well we may measure current. This might be done by for instance placing a conductor in contact with the current at a given point for a fixed short time. This conductor will 'pick up' some charges passing by, and after we remove the conductor it will have some net charge. This charge is then proportional to the current in the wire at this point.

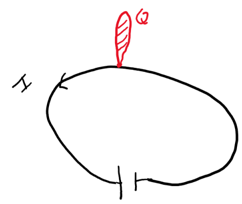

Figure 61. Current in a wire might be measured by placing a conductor in contact with the wire for a short time interval. The ratio between the amount of charge, Q, collected on the conductor during this time interval and the time interval is proportional to the current at this point of the wire.

In the above situation one finds that the current along the wire is *steady*, ie constant. Moreover, for a wire made of a certain (homogeneous and isotropic) conducting material, the current measured is found to be proportional to the potential difference created by the battery: *Ohm's law.* This constant of proportionality is called the *resistance* of the conducting circuit and measured in *Ohms* ($\Omega = $ V/A).

Just as the notion of charge leads to the notion of charge densities, so too may we consider current densities. Suppose in general that some charges are moving in space. Given a surface, Σ, we let I_Σ be the rate of charge passing through Σ. Then the *current density*, $\vec{J}$, measured in A/m^2, is:

$$\vec{J}(x) \cdot n := \lim_{Area(\Sigma) \to 0} \frac{I_\Sigma}{Area(\Sigma)}$$

where $x \in \Sigma$ and n is the unit normal to Σ at x. So, the flux of the current density through a surface is the current through the surface:

$$I_\Sigma = \int_\Sigma \vec{J} \cdot \nu \, dA.$$

As with charge, we may also consider surface current densities (denoted $\vec{K}$ (A/m)) and linear current densities (denoted $\vec{I}$ (A)) [1].

Oersted connected magnetism to currents by observing: a current produces a magnetic field. Namely, if current is flowing through a wire then one may measure a magnetic field around the current by using a compass (the compass needle will be deflected).

Thus moving charges exert forces on magnets. Conversely, magnets exert forces on moving charges. The relation is given by the Lorentz force law:

$$\vec{f_m} = q(\vec{v} \times \vec{B})$$

where $\vec{B}$ is the magnetic field and $\vec{f_m}$ the force on a test particle of charge q moving with velocity $\vec{v}$. From this experimental result the magnetic field may be measured not with magnets but as well by 'throwing' charged particles with various velocities (three independent velocities would suffice). In classic electromagnetism any magnetic effect may be equivalently produced by a certain configuration of currents, so that magnetism

[1] So if a charge distribution σ over a surface is moving along the surface, one has $\int_{\mathcal{C}} \vec{K} \cdot n \, ds = I_{\mathcal{C}}$ is the rate of charge passing through the curve $\mathcal{C}$ on the surface. In general a charge density, ρ, moving with velocity v has current density $\vec{J} = \rho v$. A surface density σ moving with velocity v (tangent to the surface) has surface current density $\vec{K} = \sigma v$. A linear density λ moving with velocity v tangent to the curve has linear current density $\vec{I} = \lambda v$.

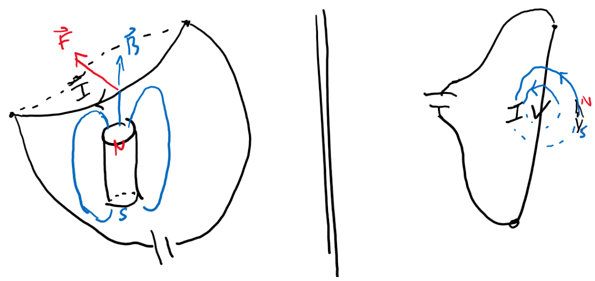

Figure 62. A magnetic field influences moving charges (a wire with current is deflected in the presence of a magnetic field). Similarly moving charges produce a magnetic field (a compass is deflected in the presence of a current).

becomes a study of moving electric charges. The unit in which the strength of the magnetic field is measured are called *Teslas* (T = N/A·m).

The *permeability*, μ, of a material (or too its *magnetic susceptibility*, $\mu = \mu_o(1 + \chi_m)$) may be measured based on a solenoid (replacing the parallel plate capacitor).

Consider a solenoidal magnetic field (produced say by running current through a tightly wrapped helical wire), approximated as uniform: when no material is present, a uniform (vertical) magnetic field of strength $B_o = \mu_o n I$ is prduced inside the cylinder (and at its ends) and a vanishing magnetic field outside.

If we now fill this solenoid with a tube of the material, we may measure the magnetic field strength B_{meas} at the solenoids ends, and have:

$$\frac{\mu}{\mu_o} \approx \frac{B_{meas}}{B_o} \Rightarrow \mu \approx \frac{B_{meas}}{nI}.$$

Since B_{meas}, I, n (recall n is the number of turns per unit length of the wrapping, assumed to be tight) may be measured, so too may μ, and χ_m. To test whether a material is homogeneous and isotropic, one might perform the measurements using different 'slices' by tubes at various points and in various directions of the material.

In the following table we list the magnetic susceptibilities of some materials (in 'normal' conditions, like electric susceptibility, its value may depend on various factors). Materials are called *paramagnetic* when $\chi_m > 0$ and *diamagnetic* when $\chi_m < 0$. Notably materials are often non-linear, called *ferromagnetic*. Iron is the only ferromagnetic in our table (whose value here is only valid for *small* ambient magnetic fields).

material	magnetic susceptibility (χ_m)
air	$\approx 4 \times 10^{-7}$
aluminum	$\approx 2 \times 10^{-5}$
glass	$\approx -13 \times 10^{-6}$
water	$\approx -9 \times 10^{-6}$
Iron	$\approx 2 \times 10^{6}$

Measuring devices: the above 'in principle' description of measurements of electromagnetic quantities in the previous section is meant to be intuitive or heuristic. In practice, measurements are carried out more precisely by using certain instruments.

...

Units

Properties of physical objects are quantities measured in units. A physical understanding of the equations and concepts consists in large part in knowing how (in principle) such quantities are measured (see §15).

*length**	meter (m)	*force*	Newton (N)
*time**	second (s)	*work, energy*	Joule (J)
*mass**	gram (g)	*power*	Watt (W)

charge	Coulomb (C)
electric potential	Volt (V)
capacitance	Farad (F)
*current**	Ampere (A)
magnetic field strength	Tesla (T)
resistance	Ohm (Ω)
inductance	Henry (H)

The starred items above are *fundamental units* of the SI (standard international) system –they are defined physically. All other items on this list are *derived units*– determined as certain ratios or products of the fundamental units. Here are some physical constants in these units:

universal gravitational constant	$G = 6.67 \times 10^{-11}\ m^3/(kg \cdot s^2)$
speed of light	$c = 2.998 \times 10^8\ m/s$
permittivity of free space	$\varepsilon_o = 0.885 \times 10^{-11}\ F/m$
permeability of free space	$\mu_o = 1.257 \times 10^{-6}\ H/m$
elementary charge	$e = 1.602 \times 10^{-19}\ C$
rest mass of electron	$m_e = 9.109 \times 10^{-31}\ kg$
rest mass of proton	$m_p = 1.673 \times 10^{-27}\ kg$

References

A standard reference is the textbook of Griffiths [1]. Some additional references:

- Feynman's lectures [2], are excellent. Purcell [3] is another standard and good introductory textbook. Also the more elementary [4], has more discussion.
- You will learn good things from essentially any book of Arnold. Here, his lectures on partial differential equations [5] are concise and relevant. See [6] for some other more comprehensive references on pdes.
- ...

[1] D. Griffiths, *Introduction to electrodynamics*. New Jersey: Prentice Hall, 1962.

[2] R. Feynman, , R. Leighton, M. Sands, *The Feynman lectures on physics; vol. II.* American Journal of Physics 33.9 (1965). Available online here.

[3] E. Purcell, *Electricity and magnetism. Vol. 2.* New York: McGraw-Hill, 1965.

[4] I. Asimov, *Understanding Physics. 3 Vols.* Scientia, Rivista di Scienza 62.103 (1968).

[5] V.I. Arnold, *Lectures on partial differential equations.* Moscow, (2004).

[6] L. Evans, *Partial differential equations.* Vol. 19. American Mathematical Soc., 2010. M. Taylor, *Partial differential equations I-III.* Springer Science & Business Media, 2013. Y. Egorov, M. Shubin. *Foundations of the classical theory of partial differential equations, Vol. I.* Springer Science & Business Media, 1998.