Contents

0 Short Review of some Mathematics, Nomenclature and Notation

Vectors are often naïvely thought of as quantities defined at each point in three-dimensional space and endowed with a **magnitude** and a **direction**, whereas a quantity with magnitude but without direction is called a **scalar**. Whilst necessary, the mere possession of these two attributes is by no means sufficient for a quantity to qualify as a vector. And a quantity described by a single number is not necessarily a scalar!

For the moment, let us note that *neither magnitude nor direction depend on the choice of coordinate system*. This is obvious in the well-known geometric representation of vectors using arrows. When we use algebraic expressions, wesometimes need a notation that makes no explicit reference to coordinate systems. After all, vectors enter in many fundamental equations in physics, and these should not change just because we transform from, say, Cartesian to spherical coordinates, or because the coordinate system has been rotated or translated. We need a more sophisticated characterisation of vectors.

0.1 Vector Spaces

Definition 0.1. A **vector space** $\mathcal{V}$ is a (possibly infinite) set of objects that is closed under addition and multiplication by a scalar:, ie. for which any two elements (**vectors**) $\mathbf{u}$ and $\mathbf{v}$ satisfy:

$$(a+b)(\mathbf{u}+\mathbf{v}) \;=\; (a\,\mathbf{u} \,+\, a\,\mathbf{v} \,+\, b\,\mathbf{u} \,+\, b\,\mathbf{v}) \;\in\; \mathcal{V}$$

$\forall\, a,\ b$ elements of some field; in what follows, we restrict these to be real. Also, $(ab)\,\mathbf{u} \,=\, a\,(b\,\mathbf{u})$. This operation is both commutative and associative.

Definition 0.2. If $\mathcal{W}$ is a subspace of $\mathcal{V}$, ie. $\mathcal{W} \subseteq \mathcal{V}$, and if any $\mathbf{w} \in \mathcal{W}$ can be written as a linear combination:

$$\mathbf{w} \;=\; \sum_{\alpha} w^{\alpha}\,\mathbf{e}_{\alpha} \tag{0.1}$$

of a set $\{\mathbf{e}_\alpha \in \mathcal{W}\}$, then that set is said to **span**, or to be a **set of generators** of, $\mathcal{W}$.

If, furthermore, this set is **linearly independent**, in the sense that demanding that $\mathbf{w} = 0$ forces all coefficients w^α in eq. (0.1) to vanish, then it is a **basis** of $\mathcal{W}$. The number n of vectors in the largest linearly independent set defines the dimension of $\mathcal{W}$, and we often write $\mathcal{W}^n$. Conversely, the number of elements of every basis of $\mathcal{W}^n$ is dimension n of $\mathcal{W}^n$, and the sum in eq. (0.1) then runs from 1 to n.

The set $\{\mathbf{e}_\alpha\}$ of basis vectors is said to be orthonormal if $\mathbf{e}_\alpha \cdot \mathbf{e}_\beta = 1$ if $\alpha = \beta$ and 0 otherwise. The operation represented by the dot will be defined below.

The (real) coefficients w^α in eq. (0.2) are called the **components** of the vector $\mathbf{w}$ in this basis. This one-to-one correspondence between $\mathcal{W}^n$ and $\mathbb{R}^n$ can be represented by a $n \times 1$ matrix:

$$\mathbf{w} \;\longmapsto\; \begin{pmatrix} w^1 \\ w^2 \\ \vdots \\ w^n \end{pmatrix}$$

In eq. (0.1), the left-hand side is explicitly basis-independent; we shall call this notation **index-free**, or **geometric**. The right-hand side, in so-called **index notation**, makes explicit reference to a basis even though, taken *as a whole*, it must still be basis-independent because of the equality.. Both notations have advantages and disadvantages. Fluency in both is highly recommended.

Warning! $\mathbf{w}$ *and its components are different beasts and should never be confused*. Also, always remember that the index on $\mathbf{e}_\alpha$ identifies the *basis vector*, not a component of the vector.

Example 0.1. $\mathbb{R}^n$, the set of all ordered n-tuples of real numbers is one of the most important vector spaces.

One popular basis of $\mathbb{R}^3$ is the Cartesian (**rectangular**, or **standard**) basis. Its three vectors are **fixed** and attached to a point arbitrarily chosen as the **origin** $(0,0,0)$.

$$\begin{aligned} \mathbf{e}_1 \equiv \hat{\mathbf{x}} \equiv \mathbf{i} &= (1,0,0)^{\mathrm{T}} \\ \mathbf{e}_2 \equiv \hat{\mathbf{y}} \equiv \mathbf{j} &= (0,1,0)^{\mathrm{T}}) \\ \mathbf{e}_3 \equiv \hat{\mathbf{z}} \equiv \mathbf{k} &= (0,0,1)^{\mathrm{T}}) \end{aligned} \tag{0.2}$$

where the label T denotes the transpose.

0.1.1 Einstein's notation for summed indices

Now is an appropriate time to introduce the **Einstein summation convention**: any index which occurs *twice* in a product term must be summed over. Thus, $u^\mu v_\mu = u_\nu v^\nu \equiv \sum_\mu u^\mu v_\mu = u^1 v_1 + u^2 v_2 + u^3 v_3$. Such an index is often called a dummy index and *any* letter can be used for it **so long as that letter is not used for any other index in the same product term**!! Therefore, a dummy index appears twice and only twice in any product term.

Remaining indices, which appear only once, are called free indices. Example: in $a^\alpha b^\nu c_\alpha u_\beta v^\beta$, α and β are dummy indices with summation implied, and ν is a free index.

I am sure you have noticed that I have written some component indices as subscripts instead of superscripts. What is the relationship between u^α and u_α? Well, in a Euclidean space with Cartesian coordinates, ie. one in which the distance between infinitesimally close points can be written $\mathrm{d}\mathbf{l}^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + \ldots$, we have $u^\alpha = u_\alpha$! In the four-dimensional **spacetime** introduced in a proper relativistic treatment of electromagnetism, this is no longer true, as we shall see later, and u^α and u_α will represent different types of components of the same vector $\mathbf{u}$.

0.1.2 Operations on 3-dim vectors in index and index-free notations

From the orthonormal property of the basis vectors, it can be shown that in Cartesian coordinates (**and only in these coordinates**!), the following operations on 3-dim vectors are defined:

	Index-free	Index components	Properties
Addition:	$\mathbf{w} = \mathbf{u} + \mathbf{v}$	$w^i = u^i + v^i$	commutative, distributive
Scalar product:	$\mathbf{u} \cdot \mathbf{v} = uv\cos\gamma$	$= u^i v_i$	commutative, distributive
Vector product:	$\mathbf{w} = \mathbf{u} \times \mathbf{v} = (uv\sin\gamma)\hat{\mathbf{n}}$	$w_i = \epsilon_{ijk} u^j v^k$	distributive
tensor product	$\mathbf{u} \otimes \mathbf{v}$	$u^i\, v^j$	commutative, distributive

where $u_i = u^i$, the components of the vector $\mathbf{u}$, *only in a Euclidean space with a Cartesian basis*. This means that any index expression that contains components with subscripts will look more complicated in curvilinear bases if we insists on expressing it in terms of *vector* components (u^i) only! The relation between components with superscripts and those with subscripts will be discussed near the end of the course. The symbol ϵ_{ijk} is defined in the following subsection. Also, γ is the angle between $\mathbf{u}$ and $\mathbf{v}$, and $\hat{\mathbf{n}}$ is a unit vector perpendicular to the plane defined by $\mathbf{u}$ and $\mathbf{v}$, and whose direction is given conventionally by the right-hand rule. The magnitude (or norm) of a vector is simply $u = (u^i u_i)^{1/2}$; like γ, it is independent of the coordinate system in which the vector components are expressed. Now, everything that pertains to three dimensional space in this course will assume that it is Euclidean (no curvature, unlike for instance on a sphere), but often we will wish to work in a curvilinear basis, so some care will have to be exercised.

Note that the vector product is neither commutative nor associative. In fact, it is *antisymmetric*. Also, there is no simple expression for the tensor product in terms of magnitude and angle.

0.1.3 Meet your friend, the Levi-Civita symbol

The above expression for the vector product is much more compact than the one which uses a determinant. In three dimensions the **Levi-Civita symbol**, ϵ_{ijk}, is defined as a 27-component object which obeys the following rules:

$$\begin{aligned}\epsilon_{ijk} &= -\,\epsilon_{jik} = -\epsilon_{ikj} \\ &= \epsilon_{kij} = \epsilon_{jki}\end{aligned}$$

The first line means that ϵ is zero whenever any two of its indices are the same. Now, set $\epsilon_{123} = 1$ by convention. Then the other non-vanishing components can only be 1 or -1. We say that the Levi-Civita symbol is totally antisymmetric in its indices. Because of this all-important property, it has only *six* non-zero components, and only one *independent* component

Using the above rules, it is easy to work out the Cartesian vector product. For instance, suppose you want $w_y = w^2$. Set $i = 2$. Then the only values for j and k that give non-vanishing ϵ components are 1 and 3. Set $j = 1$; then k must be 3, again to have ϵ non zero. This gives the first term in the sum. Now, set $j = 3$; then $k = 1$, giving the second and only other term in the sum. Thus, $w^2 = w_2 = \epsilon_{231}u^3v^1 + \epsilon_{213}u^1v^3 = u^3v^1 - u^1v^3$. In the same way, $w^1 = u^2v^3 - u^3v^2$, and $w^3 = u^1v^2 - u^2v^1$, and we have regained the familiar expression for the components of the vector product. The third kind of product between two vectors, the tensor product, is most easily written in component form. Indeed, $\mathbf{u} \otimes \mathbf{v}$ can be represented by a 3×3 matrix whose elements are given by u^iv^j. An antisymmetric version of this product, called the exterior product $\mathbf{u} \wedge \mathbf{v}$, is easily constructed; it has components $u^iv^j - u^jv^i$ in any basis. In a Cartesian basis, these look like the components of the vector product, except that they belong to an object with two indices (rank 2), not just one! The components of the 3-dim exterior product and those of the vector product are said to be **dual** to each other, and are connected via the Levi-Civita symbol: $(\mathbf{u} \times \mathbf{v})^i = (\mathbf{u} \times \mathbf{v})_i = \frac{1}{2}\epsilon_{ijk}(u^jv^k - u^kv^j)$ in a Cartesian basis.

What is also nice about this coordinate notation for the scalar and vector products is that it allows relations involving them to be proved without sweat. Thus:

$$\begin{aligned}\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= \epsilon_{ijk}u^iv^jw^k \\ &= w^k\epsilon_{kij}u^iv^j \qquad \text{(cyclic permutation of indices on } \epsilon) \\ &= \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}).\end{aligned}$$

Et voilà! For comparison, you should try the method that works out all components explicitly.

It will happen that *two* ϵ's are multiplied, with one index summed over. No panic! Use the following (unproven) rule:

$$\epsilon_{ijk}\epsilon^{lnk} = \delta_i{}^l\delta_j{}^n - \delta_i{}^n\delta_j{}^l. \tag{0.3}$$

The index (here k) that gets summed over must appear *in the same position* on both ϵ's (which position does not matter, because to go from one to the other, the permutation rules are applied twice, once to each ϵ). *Note how each **free** index on the left-hand side also appears in **each** of the two terms on the right-hand side*. This rule applies to indices taken as algebraic symbols, not to *particular values* that they can take, as is obvious from the above computation of the components of the vector product!

0.2 Review of Differential Calculus in Three Dimensions

In physics, as in many other pursuits, we are chiefly interested in how things change in space or time. In one dimension, the change of a (scalar) function, $f(x)$, when its argument changes by $\mathrm{d}x$, is $\mathrm{d}f = (\mathrm{d}f/\mathrm{d}x)\mathrm{d}x$, where

$\mathrm{d}f/\mathrm{d}x$ is the slope of f at x. Ask a similar question in three dimensions, namely, what is the change $df(\mathbf{x})(\mathbf{u})$ of a function $f(\mathbf{x})$ in the direction of a vector, $\mathbf{u}$, under an infinitesimal change of its argument in a Cartesian basis? The answer is straightforward:

$$\mathrm{d}f(\mathbf{x})(\mathbf{u}) \;=\; u^i \partial_i f \tag{0.4}$$

where $\partial_i f \equiv \partial f/\partial x^i$, the partial derivatives of $f(\mathbf{x})$ in index notation, can be viewed as the Cartesian components of an object called the **gradient** of f, whose vector version is usually denoted by $\boldsymbol{\nabla} f$. Only in a Cartesian basis are the components of this gradient also $\partial_i f$, that is:

$$\boldsymbol{\nabla} f = \hat{\mathbf{x}}\,\partial_x f + \hat{\mathbf{y}}\,\partial_y f + \hat{\mathbf{z}}\,\partial_z f \tag{0.5}$$

In index-free notation, we can write the directional change of the function $\boldsymbol{\nabla} f \cdot \mathbf{u}$. If $\mathbf{u}$ is a vector in the plane tangent at some point to a curve over which f remains constant, then the gradient must be perpendicular to that plane. Also, the change in the function is maximum in the direction of $\boldsymbol{\nabla} f$.

If $\boldsymbol{\nabla} f$ can be viewed as vector, what then is $\boldsymbol{\nabla}$? It is a **vector operator**, the **gradient operator**, which can be treated as a vector with the understanding that it *acts* on f.

Besides acting on scalar fields, the gradient operator can also act on vectors. Thus, $\boldsymbol{\nabla}$ can act on:

- a scalar f to yield a *vector*: $\boldsymbol{\nabla} f$, with components $\partial_i f$ (**gradient** of f);
- a vector $\mathbf{u}$ to yield a *scalar*: $\boldsymbol{\nabla} \cdot \mathbf{u}$, or $\partial_i u^i$ (**divergence** of $\mathbf{u}$);
- a vector to yield a **vector**: $\boldsymbol{\nabla} \times \mathbf{u}$, with components $\epsilon^{ijk}\partial_j u_k$ (**curl** of $\mathbf{u}$).

Once again, the expressions written in index notation are only valid in a Cartesian basis!

The front cover of Jackson contains useful product rules for $\boldsymbol{\nabla}$. EXERCISE: if you wish to sharpen your index-manipulation skills, you can try proving some or all those product rules using index notation.

Applying $\boldsymbol{\nabla}$ yet again to the gradient, divergence, and curl yields:

1. the **Laplacian** of $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} f) \equiv \nabla^2 f$, or $\partial_j \partial^j f$ in index notation and in a Cartesian basis,
2. $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} f) \equiv 0$,
3. $\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{u})$,
4. $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{u}) \equiv 0$,
5. $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{u}) \equiv \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{u}) - \nabla^2 \mathbf{u}$.

(2), (4), and (5) are **identities**, in the sense that they provide *no information whatsoever* about f and $\mathbf{u}$, respectively, since they are *always* satisfied. Identity (5) just gives two equivalent ways of writing the same thing, and is best viewed as a *definition* of $\nabla^2 \mathbf{u}$.

The only really useful quantity involving second-order spatial derivatives is the Laplacian (1).

0.3 Review of Integral Calculus

Three important theorems govern integrals of gradients, divergences, and curls:

1. Gradient theorem: $\displaystyle\int_{\text{line}}^{b} (\boldsymbol{\nabla} f) \cdot \mathrm{d}\mathbf{l} = f(b) - f(a)$;

2. Curl theorem (Stokes): $\int_{\text{surface}} (\boldsymbol{\nabla} \times \mathbf{u}) \cdot \mathrm{d}\mathbf{a} = \oint_{\text{line}} \mathbf{u} \cdot \mathrm{d}\mathbf{l}$;

3. Divergence theorem (Green, Gauss): $\int_{\text{volume}} \boldsymbol{\nabla} \cdot \mathbf{u} \, \mathrm{d}^3x = \oint_{\text{surface}} \mathbf{u} \cdot \mathrm{d}\mathbf{a}$.

By convention the surface element $\mathrm{d}\mathbf{a}$ is a vector normal to the surface and points *outward* when the surface is closed; in the curl theorem, it points in the direction of the thumb when the other fingers of your right hand curl around the direction of circulation in the line integral.

Some of the integrals in the curl and divergence theorems have names: $\oint_{\text{line}} \mathbf{u} \cdot \mathrm{d}\mathbf{l}$ is called the **circulation** of $\mathbf{u}$ around a closed path, whereas $\int_{\text{surface}} \mathbf{u} \cdot \mathrm{d}\mathbf{a}$ is called the **flux** of $\mathbf{u}$ through a surface.

Notice that in each of the theorems, the left-hand side is, loosely speaking, the integral of the derivative of an object over a (one-, two-, or three-dimensional) region, whereas the right-hand side is the integral of the object over the *boundary* of the same region. The three theorems are just particular forms of a fundamental theorem in differential geometry. As a consequence, when a region (line interval, 2-dimensional surface) is embedded in a higher-dimensional space, as is the case in the gradient and curl theorems, the integrals on the left are equal to integrals over *any* region which has the *same* boundary! Furthermore, when these regions are closed, the right-hand side vanishes.

Since $\oint \boldsymbol{\nabla} f \cdot \mathrm{d}\mathbf{l} = 0$, it follows that *any* vector field $\mathbf{u}$ which is the gradient of some scalar field must have zero circulation around any closed path: $\oint \mathbf{u} \cdot \mathrm{d}\mathbf{l} = 0$. We say that $\mathbf{u}$ is **conservative**. The converse also holds: if $\oint \mathbf{u} \cdot \mathrm{d}\mathbf{l} = 0$ for any closed path in a region, Stoke's theorem demands that $\boldsymbol{\nabla} \times \mathbf{u} = 0$ at all points in the region, and $\mathbf{u}$ must be the gradient of some scalar function (I omit the proof since it is rather fussy). The most useful statement that emerges from the discussion in this paragraph is that if $\boldsymbol{\nabla} \times \mathbf{u} = 0$ over a simply-connected region (no doughnuts or surfaces with holes!), $\mathbf{u}$ is conservative and there exists a scalar field f such that $\mathbf{u} = \boldsymbol{\nabla} f$.

Green's Identities (section J1.8)

By putting $\mathbf{u} = f\boldsymbol{\nabla} g$, where f and g are scalar fields, in the divergence theorem, we can use the seventh product rule inside Jackson's front cover to show **Green's first identity**:

$$\int_{\text{volume}} [f\,\nabla^2 g + (\boldsymbol{\nabla} f) \cdot (\boldsymbol{\nabla} g)] \, \mathrm{d}^3x \;\equiv\; \oint_{\text{surface}} f \, \frac{\partial g}{\partial n} \, \mathrm{d}a \tag{0.6}$$

where $\partial/\partial n$ is the derivative in the direction normal and outward to the surface enclosing the volume. Next, we can write the first identity with f and g interchanged, subtract, and prove **Green's second identity**, sometimes (including by Jackson) confusingly called Green's theorem:

$$\int_{\text{volume}} (f\,\nabla^2 g - g\,\nabla^2 f) \, \mathrm{d}^3x \;\equiv\; \oint_{\text{surface}} \left(f \, \frac{\partial g}{\partial n} \,-\, g \, \frac{\partial f}{\partial n} \right) \mathrm{d}a \tag{0.7}$$

Green's identities will be very useful when we review boundary-value problems and introduce Green function methods.

0.4 Vectors in Curvilinear Coordinates

Other than Cartesian bases, the two most useful types of coordinate system are **spherical** and **cylindrical**.

In a spherical system, the components of a *position* vector are one distance and two angles, (r, θ, ϕ), illustrated in fig. J3.1. The transformations *from* spherical *to* Cartesian position coordinates are:

$$x = r \sin\,\theta\,\cos\,\phi, \qquad y = r \sin\,\theta\,\sin\,\phi, \qquad z = r\cos\,\theta \tag{0.8}$$

Beware: mathematicians often use ϕ as the polar angle and θ for the azimuthal angle in spherical coordinates! This is certainly the case in the symbolic manipulation software *Maple* when it knows that it is working in those coordinates.

As for the inverse transformations, the most useful ones relate the unit vectors:

$$\begin{aligned}\hat{\mathbf{r}} &= \hat{\mathbf{x}}\,\sin\theta\,\cos\phi + \hat{\mathbf{y}}\,\sin\theta\,\sin\phi + \hat{\mathbf{z}}\,\cos\theta \\ \hat{\boldsymbol{\theta}} &= \hat{\mathbf{x}}\,\cos\theta\,\cos\phi + \hat{\mathbf{y}}\,\cos\theta\,\sin\phi - \hat{\mathbf{z}}\,\sin\theta \\ \hat{\boldsymbol{\phi}} &= -\hat{\mathbf{x}}\,\sin\phi + \hat{\mathbf{y}}\,\cos\phi\end{aligned} \tag{0.9}$$

How are the spherical and Cartesian components of a vector $\mathbf{u}$ related? We can always write:

$$\begin{aligned}\mathbf{u} &= u_x\,\hat{\mathbf{x}} + u_y\,\hat{\mathbf{y}} + u_z\,\hat{\mathbf{z}} \\ &= u_r\,\hat{\mathbf{r}} + u_\theta\,\hat{\boldsymbol{\theta}} + u_\phi\,\hat{\boldsymbol{\phi}}\end{aligned}$$

Insert the above expressions for the spherical unit vectors in $u_r = \hat{\mathbf{r}}\cdot\mathbf{u} = u_x\,\hat{\mathbf{r}}\cdot\mathbf{x} + u_y\,\hat{\mathbf{r}}\cdot\mathbf{y} + u_z\,\hat{\mathbf{r}}\cdot\mathbf{z}$, $u_\theta = \hat{\boldsymbol{\theta}}\cdot\mathbf{u}$, and $u_\phi = \hat{\boldsymbol{\phi}}\cdot\mathbf{u}$, where $\mathbf{u}$ is given in terms of u_x, u_y and u_z, and work out the scalar products to obtain the transformation law:

$$\begin{pmatrix} u_r \\ u_\theta \\ u_\phi \end{pmatrix} = \begin{pmatrix} \sin\theta\,\cos\phi & \sin\theta\,\sin\phi & \cos\theta \\ \cos\theta\,\cos\phi & \cos\theta\,\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \end{pmatrix} \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} \tag{0.10}$$

The inverse transformation can be computed either by inverting the matrix or by working out $u_x = \mathbf{x}\cdot\hat{\mathbf{u}}$, $u_y = \hat{\mathbf{y}}\cdot\mathbf{u}$, and $u_z = \hat{\mathbf{z}}\cdot\mathbf{u}$, with $\mathbf{u}$ given this time in terms of u_r, u_θ and u_ϕ. If $\mathbf{u}$ were dependent on the coordinates, ie. if it were a vector *field*, the coordinates themselves would also have to be transformed. EXERCISE: transform $\mathbf{u}(x,y,z) = (-2,-1,3)$, where $(x,y,z) = (2,-1,0)$, to a spherical basis.

It is not hard to see that the infinitesimal displacement vector is given in spherical coordinates by:

$$\mathrm{d}\mathbf{l} = \mathrm{d}r\,\hat{\mathbf{r}} + (r\,\mathrm{d}\theta)\hat{\boldsymbol{\theta}} + (r\,\sin\theta\,\mathrm{d}\phi)\hat{\boldsymbol{\phi}} \tag{0.11}$$

and that the volume element is just the product of the infinitesimal displacements in the three orthogonal directions: $\mathrm{d}^3x = r^2\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi$. There is no general expression for the surface element, $\mathrm{d}\mathbf{a}$, as this depends on its orientation in space.

Once we know the form of $\mathrm{d}\mathbf{l}$ in spherical coordinates (or any other system for that matter) it is easy to find the components of the gradient of a scalar field f. Just write:

$$\begin{aligned}\mathrm{d}f &= (\partial_r f)\mathrm{d}r + (\partial_\theta f)\mathrm{d}\theta + (\partial_\phi f)\mathrm{d}\phi \qquad \text{(chain rule)} \\ &= \boldsymbol{\nabla} f\cdot\mathrm{d}\mathbf{l} \qquad \text{(coordinate-free form)}\end{aligned}$$

Working out the last line after inserting the above expression for $\mathrm{d}\mathbf{l}$ gives:

$$(\boldsymbol{\nabla} f)_r = \partial_r f, \qquad (\boldsymbol{\nabla} f)_\theta = \frac{1}{r}\partial_\theta f, \qquad (\boldsymbol{\nabla} f)_\phi = \frac{1}{r\sin\theta}\partial_\phi f \tag{0.12}$$

from which one reads off the components of the gradient operator $\boldsymbol{\nabla}$.

One might think that the divergence and the curl of a vector can now be easily found. Not quite! Unlike in a Cartesian basis, the unit vectors of a spherical basis are not fixed at the origin; rather, they are attached to the particular point where our vector is defined. As soon as we move away from this point, the set of basis vectors changes its orientation unless the motion is in the radial direction. So we expect that, unlike in a Cartesian basis, the spherical unit vectors will have *non-zero* spatial derivatives. These can be calculated easily from the above expressions relating $(\hat{\mathbf{r}},\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\phi}})$ to $(\mathbf{x},\mathbf{y},\mathbf{z})$. In tabular form, we get:

	$\hat{\mathbf{r}}$	$\hat{\boldsymbol{\theta}}$	$\hat{\boldsymbol{\phi}}$
∂_r	0	0	0
∂_θ	$\hat{\boldsymbol{\theta}}$	$-\hat{\mathbf{r}}$	0
∂_ϕ	$\hat{\boldsymbol{\phi}}\sin\theta$	$\hat{\boldsymbol{\phi}}\cos\theta$	$-\hat{\mathbf{r}}\sin\theta - \hat{\boldsymbol{\theta}}\cos\theta$

Now it is a straightforward—if somewhat tedious—exercise to calculate the divergence[†] of a vector $\mathbf{u}$ in spherical coordinates. Expand:

$$\boldsymbol{\nabla}\cdot\mathbf{u} \;=\; \left(\hat{\mathbf{r}}\partial_r \;+\; \hat{\boldsymbol{\theta}}\,\frac{1}{r}\partial_\theta \;+\; \hat{\boldsymbol{\phi}}\,\frac{1}{r\sin\theta}\partial_\phi\right)\cdot\left(\hat{\mathbf{r}}\,u_r \;+\; \hat{\boldsymbol{\theta}}\,u_\theta \;+\; \hat{\boldsymbol{\phi}}\,u_\phi\right)$$

using the product rule for derivatives and the table above. You can then put the five terms you will get in the elegant form:

$$\boldsymbol{\nabla}\cdot\mathbf{u} \;=\; \frac{1}{r^2}\partial_r(r^2u_r) \;+\; \frac{1}{r\sin\theta}\left[\partial_\theta(\sin\theta\,u_\theta) \;+\; \partial_\phi u_\phi\right] \tag{0.13}$$

An even more tedious calculation yields the expression for $\boldsymbol{\nabla}\times\mathbf{u}$ given inside the back cover of Jackson.

A word of caution: although one uses a Cartesian basis to derive it, the identity $\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{u}) \equiv \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{u}) - \nabla^2\mathbf{u}$ does hold in any coordinate system. But it is of limited interest in any basis other than Cartesian, because of the term $\nabla^2\mathbf{u}$. Although $\nabla^2\mathbf{u} = (\nabla^2u_x)\hat{\mathbf{x}} + (\nabla^2u_y)\hat{\mathbf{y}} + (\nabla^2u_z)\hat{\mathbf{z}}$, it is *not* equal to $(\nabla^2u_r)\hat{\mathbf{r}} + (\nabla^2u_\theta)\hat{\boldsymbol{\theta}} + (\nabla^2u_\phi)\hat{\boldsymbol{\phi}}$, because the derivatives of the unit vectors are non-zero. In fact, $\nabla^2\mathbf{u}$ is best evaluated *from* the identity.

Clearly, divergences and curls of *vetors* are much more complicated in spherical (and cylindrical) bases than in Cartesian ones. Why then bother with non-Cartesian bases? Because in some important situations, those with a *symmetry*, these awful expressions collapse down to very simple ones. In spherical coordinates, for instance, any vector without angular dependence, and therefore spherically symmetric, will have a one-term divergence and zero curl, just by inspection of the relevant expressions.

Consider the vector field $\hat{\mathbf{r}}/r^2$, which plays a very important rôle in electromagnetism. Its divergence in spherical coordinates is very simple to calculate: it vanishes everywhere *except* at $r = 0$, where $\hat{\mathbf{r}}/r^2$ diverges. By contrast, more work is needed to find the same result in a Cartesian basis, with $r^2 = x^2 + y^2 + z^2$. The problem at $r = 0$ is not an artefact of the spherical coordinates since $\hat{\mathbf{r}}/r^2$ diverges in *any* basis. How do we know what the divergence is at the origin then? Well, if we integrate it over a sphere of radius R centered at the origin and use the divergence theorem to convert the volume integral to the flux of $\hat{\mathbf{r}}/r^2$ through the spherical surface, we get 4π for this flux, no matter how small we choose R!. This non-zero result can therefore only come from the origin since we know that the divergence vanishes everywhere else. But how should we write the divergence at the origin?

0.5 Dirac delta-"function"

A very useful object in physics (and invented by P.A.M. Dirac, a physicist trained as an engineer) is the Dirac delta-function, $\delta(x)$, which is zero everywhere on the x-axis, except at $x = 0$ where it is infinite. Clearly. this is not a function in the ordinary sense; to a mathematician, it belongs to a class of objects called distributions. In spite of its strange behaviour, it can be represented in terms of perfectly mundane objects. In one dimension:

$$\delta(x) = \lim_{g\to\infty}\frac{1}{\pi}\frac{\sin gx}{x} \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\,\mathrm{d}k$$

[†]For those who know such things, it is much simpler first to identify the non-zero components g_{ij} of the metric tensor for spherical coordinates which are the coefficients of the $\mathrm{d}r^2$, $\mathrm{d}\theta^2$ and $\mathrm{d}\phi^2$ terms in the norm of the line element given by eq. (0.11), computing $\sqrt{g}$, where $g = r^4\sin^2\theta$ is the determinant of the diagonal g_{ij} matrix, and using the very general formula for the divergence of a vector $\mathbf{u}$ in any space and any basis: $\partial_i(\sqrt{g}u^i)/\sqrt{g}$.

from which it is obvious that the δ-function has units of inverse x.

The integral of the delta-function over its whole range is well defined: $\int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = 1$. Since $f(x)\delta(x-a)$ is for all practical purposes $f(a)\delta(x-a)$, it follows that:

$$f(a) = \int_{-\infty}^{\infty} f(x)\delta(x-a)\,\mathrm{d}x \tag{0.14}$$

for any "well-behaved" function $f(x)$. Thus, the delta-function can be used to pick out a particular value of a function.

Another useful property:

$$\delta(ax) \;=\; \frac{1}{|a|}\,\delta(x) \tag{0.15}$$

These properties are readily extended to the 3-dimensional delta-function, $\delta(\mathbf{x}) = \delta(x)\delta(y)\delta(z)$. Indeed, we have:

$$f(\mathbf{x}_0) = \int\limits_{\text{all space}} f(\mathbf{x})\delta(\mathbf{x}-\mathbf{x}_0)\mathrm{d}^3x$$

See p. 26 in Jackson for other useful properties of the δ-function, in particular, the important representation:

$$\delta(x-x') \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k(x-x')}\,\mathrm{d}k \tag{0.16}$$

The result of problem 1.2 is also important in spherical or cylindrical coordinates!

Now we know our divergence of the previous section: $\boldsymbol{\nabla}\cdot(\mathbf{x}/x^3) = 4\pi\delta(\mathbf{x})$. This vanishes $\forall\,\mathbf{x}\neq 0$ and has a volume integral over all space equal to 4π, consistent with the flux calculated above. More generally:

$$\boldsymbol{\nabla}\cdot\left(\frac{\mathbf{x}-\mathbf{x}'}{|\mathbf{x}-\mathbf{x}'|^3}\right) \;=\; 4\pi\,\delta(\mathbf{x}-\mathbf{x}') \tag{0.17}$$

where the differentiation is with respect to $\mathbf{x}$, $\mathbf{x}'$ being fixed. Now since $\boldsymbol{\nabla}(1/R) \;=\; -(\mathbf{x}-\mathbf{x}')/R^3$, where $R = |\mathbf{x}-\mathbf{x}'|$, we immediately find a result that will prove very useful:

$$\nabla^2\left(\frac{1}{|\mathbf{x}-\mathbf{x}'|}\right) \;=\; -4\pi\,\delta(\mathbf{x}-\mathbf{x}') \tag{0.18}$$

0.5.1 Helmholtz theorem

The **Helmholtz theorem** asserts that a vector function $\mathbf{F}(\mathbf{x})$ that vanishes at infinity faster than $1/|\mathbf{x}-\mathbf{x}'|$ is uniquely determined over space if its divergence and curl are known. In that case:

$$\mathbf{F} \;=\; \boldsymbol{\nabla}u \;+\; \boldsymbol{\nabla}\times\mathbf{w} \tag{0.19}$$

where

$$u(\mathbf{x}) \;=\; -\frac{1}{4\pi}\int \frac{\boldsymbol{\nabla}'\cdot\mathbf{F}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x' \qquad\qquad \mathbf{w}(\mathbf{x}) \;=\; \frac{1}{4\pi}\int \frac{\boldsymbol{\nabla}'\times\mathbf{F}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x'$$

and it is understood that the integration must extend over all space.

Also, vector fields whose curl vanishes everywhere in a simply-connected space can be written as the gradient of some scalar field, the latter being determined only up to a constant. Similarly, vector fields whose divergence vanishes *everywhere* in a simply-connected space ($\hat{\mathbf{r}}/r^2$ is not one of them!) can be written as the curl of some vector, the latter being determined only up to the gradient of a scalar.

0.6 Expansion of Functions in a Complete Set of Orthonormal Functions (section J2.8)

We will be interested in classes of functions with two important properties. First, they are **separable**, that is, written as products of functions, each depending on only one coordinate, in some coordinate system. Secondly, these functions—call them $f(\xi)$, where ξ is some coordinate—can be expanded over sets of functions $U_n(\xi)$ which are (see also pp. 67-69 in Jackson):

1. **complete**: $f(\xi) = \sum_{n=1}^{\infty} c_n U_n(\xi) \;\; \forall f(\xi)$, which means that the $U_n(\xi)$ must satisfy:

$$\sum_{n=1}^{\infty} U_n^*(\xi')\, U_n(\xi) \;=\; \delta(\xi' - \xi) \tag{0.20}$$

2. **orthonormal** over some interval (a, b):

$$\int_a^b U_m^*(\xi)\, U_n(\xi)\, \mathrm{d}\xi \;=\; \delta_{mn} \tag{0.21}$$

which allows to find the expansion coefficients:

$$c_n \;=\; \int_a^b U_n^*(\xi)\, f(\xi)\, \mathrm{d}\xi \tag{0.22}$$

When the interval (a, b) is infinite, $(-\infty, \infty)$, some complete and orthonormal sets may become continuous. A well-known example is the discrete set over the interval $(-a/2, a/2)$:

$$U_n(\xi) \;=\; \frac{1}{\sqrt{a}}\, \mathrm{e}^{\mathrm{i}k_n\xi} \qquad\qquad (k_n = \frac{2\pi n}{a},\; n \text{ integer})$$

defining the **Fourier series** expansion:

$$\begin{aligned} f(\xi) \;&=\; \frac{1}{\sqrt{a}} \sum_{n=-\infty}^{\infty} c_n\, \mathrm{e}^{\mathrm{i}k_n\xi} \\ c_n \;&=\; \frac{1}{\sqrt{a}} \int_{-a/2}^{a/2} f(\xi')\, \mathrm{e}^{-\mathrm{i}k_n\xi'}\, \mathrm{d}\xi' \end{aligned} \tag{0.23}$$

When $a \to \infty$, the sum over modes becomes an integral over k, and the Fourier series goes over to a **Fourier integral**:

$$\begin{aligned} f(\xi) \;&=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k)\, \mathrm{e}^{\mathrm{i}k\xi}\, \mathrm{d}k \\ A(k) \;&=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\xi)\, \mathrm{e}^{-\mathrm{i}k\xi}\, \mathrm{d}\xi \end{aligned} \tag{0.24}$$

with:

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}(k-k')\xi}\, \mathrm{d}\xi \;&=\; \delta(k - k') \qquad\qquad \text{(orthonormality)} \\ \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k(\xi-\xi')}\, \mathrm{d}k \;&=\; \delta(\xi - \xi') \qquad\qquad \text{(completeness)} \end{aligned} \tag{0.25}$$

0.7 Solution of the Laplace Equation in Rectangular Coordinates (sections J2.9–2.11)

Perhaps the most useful method of solving the Laplace equation, $\nabla^2\Phi = 0$, is to express its solutions as expansions over sets of orthonormal functions (see section 0.6 above, or section J2.8). To solve the equation, we must decide first on our coordinate basis. It is an arbitrary choice, but a non-trivial one as it can make the difference between success and failure. It should be guided by the shape of the closed surface over which the boundary conditions (B.C.) are given. We shall consider flat and spherical surfaces in turn.

With flat boundaries, especially if they are mutually perpendicular, it makes eminent sense to use a rectangular basis oriented along the surfaces. This allows the B.C. to be written in their simplest form.

We seek **separable** solutions of the form $\Phi(x,y,z) = X(x)Y(y)Z(z)$. To find out if and under which conditions this **ansatz** works, substitute this form for Φ into $\nabla^2\Phi = \partial_x^2\Phi + \partial_y^2\Phi + \partial_z^2\Phi$ and divide by Φ to obtain:

$$\frac{1}{X}\mathrm{d}_x^2X(x) \;+\; \frac{1}{Y}\mathrm{d}_y^2Y(y) \;+\; \frac{1}{Z}\mathrm{d}_z^2Z(z) \;=\; 0$$

Each additive term depends on only one coordinate. Since the coordinates can take *arbitrary* values inside the volume considered, the equation can be satisfied only if *each term is constant*, giving the three *ordinary* differential equations:

$$\mathrm{d}_x^2X(x) \;=\; \alpha^2X(x), \qquad \mathrm{d}_y^2Y(y) \;=\; \beta^2Y(y), \qquad \mathrm{d}_z^2Z(z) \;=\; \gamma^2Z(z) \qquad\qquad (\alpha^2+\beta^2+\gamma^2=0)$$

with (*not completely general*!) solutions:

$$\begin{aligned} X(x) &= A\,\mathrm{e}^{\alpha x} + B\,\mathrm{e}^{-\alpha x} \\ Y(y) &= C\,\mathrm{e}^{\beta y} + D\,\mathrm{e}^{-\beta y} \\ Z(z) &= F\,\mathrm{e}^{\sqrt{\alpha^2+\beta^2}\,z} + G\,\mathrm{e}^{-\sqrt{\alpha^2+\beta^2}\,z} \end{aligned} \tag{0.26}$$

where α^2 and β^2 are constants, and the constants A, B, etc., must be determined from the B.C. If this can be done with some linear superposition of each of X, Y, and Z, then the product of these superpositions will be *the* solution.

Example 0.2. A two-dimensional example

Many boundary-value problems can be reduced to two-dimensional problems, where the potential does not change along (say) the z direction, being entirely determined in terms of x and y. Then $\Phi(x,y) = X(x)Y(y)$, and $\alpha^2 = -\beta^2$. Unlike Jackson in section J2.10, however, I make no educated guess about the sign of α^2, in order to show you how it comes out in the wash.

In this example $\Phi(0,y) = 0$ and $\Phi(a,y) = 0$. The first condition forces $B = -A$, so that $\Phi(x,y) = A(C\,\mathrm{e}^{\beta y} + D\,\mathrm{e}^{-\beta y})(\mathrm{e}^{\alpha x} - \mathrm{e}^{-\alpha x})$. The second condition demands that $e^{2a\alpha} = 1$, which cannot be satisfied if α is real and non-zero. But it is satisfied if $\alpha = \mathrm{i}\,k$, with $k = n\pi/a$ (n an integer). Absorbing the overall constant A into C and D, recalling that $\beta^2 = -\alpha^2 = -k^2 > 0$, and implementing $\Phi(x,\infty) \to 0$, we have the solution:

$$\Phi_n(x,y) = \mathrm{e}^{-k_ny}\sin\,k_nx \qquad\qquad (k_n = n\pi/a, \quad n>0)$$

At this point we realise that it is impossible to fit the other B.C., $\Phi(x,0) = V$, with just this solution for *arbitrary* values of x. The only way to save the situation is to construct a more general solution, making use of the linearity of the Laplace equation and the fact that any positive integer n gives a solution to write:

$$\Phi(x,y) \;=\; \sum_{n=1}^{\infty} A_n\,\mathrm{e}^{-k_ny}\sin\,k_nx$$

The B.C. now takes the form $V = \sum_{n=1}^{\infty} A_n \sin\, k_n x$, an expansion over the complete orthonormal set $\{\sin k_n x\}$. From this, we determine the (Fourier) coefficients $A_n = (2/a)\int_0^a V \sin k_n x \,\mathrm{d}x$. There comes:

$$\Phi(x,y) \;=\; \frac{4V}{\pi}\sum_{n\ \text{odd}} \frac{1}{n}\,\mathrm{e}^{-n\pi y/a}\,\sin(n\pi x/a) \;=\; \frac{2V}{\pi}\,\tan^{-1}\left(\frac{\sin(\pi x/a)}{\sinh(\pi y/a)}\right) \tag{0.27}$$

where the last equality is derived on pp. 74-75 in Jackson.

Example 0.3. Three-dimensional extension

In section J2.9, Jackson looks at the case of a hollow rectangular box with dimensions (a,b,c) in the (x,y,z) directions, with all sides at zero potential, except for the side $z = c$. Going back to the 3-dimensional solution of eq. (0.26), we can move more quickly by noting that α^2 and β^2 may be chosen to be negative and dropping overall constants in each of the three functions.

Then the vanishing of the potential at $x = 0$ and $y = 0$ forces both X and Y to be sines, while the conditions at $x = a$ and $y = b$ demand that $\alpha_n = n\pi/a$, and $\beta_n = n\pi/b$, so that $\gamma_{nm} = \pi\,(n^2/a^2 + m^2/b^2)^{1/2}$ and the condition at $z = 0$ is satisfied if Z is a sinh. Then:

$$\Phi(x,y,z) \;=\; \sum_{n,m=1}^{\infty} A_{nm}\sin(n\pi x/a)\,\sin(n\pi y/b)\,\sinh(\gamma_{nm} z) \tag{0.28}$$

with the Fourier coefficients A_{nm} determined by imposing $\Phi(x,y,c) = V(x,y)$.

0.8 Laplace and Helmholtz Equations in Spherical Coordinates (sections J3.1, J3.2, J3.5, J9.6)

The Laplacian operator is ubiquitous in physics. To study it, we first separate it into two convenient parts by introducing the self-adjoint vector operators $-\mathrm{i}\boldsymbol{\nabla}$ and $\mathbf{L} = -\mathrm{i}\,\mathbf{x}\times\boldsymbol{\nabla}$, or $L_i = -\mathrm{i}\epsilon_{ijk}x^j\partial^k$, where ϵ_{ijk} is the completely antisymmetric Levi-Civita symbol, and summation over repeated indices is implied. With the identity: $\epsilon_{ijk}\epsilon^{imn} = \delta_j{}^m\delta_k{}^n - \delta_j{}^n\delta_k{}^m$, the scalar product of $\mathbf{L}$ with itself is, in Cartesian coordinates:

$$\mathbf{L}\cdot\mathbf{L} \;=\; -\epsilon_{ijk}\epsilon^{imn}\,x^j\partial^k\,x_m\partial_n \;=\; -x^j\left(\partial_j + x_j\partial^k\partial_k - 3\partial_, - x_k\partial^k\partial_j\right) \;=\; -\,x^j x_j\,\partial^k\partial_k \;+\; x^j\,\partial_j \;+\; x^j\partial_j x^k\partial_k$$

Extracting the Laplacian and reverting to coordinate-free notation, there comes with $x^j x_j = r^2$:

$$\nabla^2 \;=\; -\frac{\mathbf{L}^2}{r^2} \;+\; \frac{1}{r}\left[\partial_r \;+\; \partial_r(r\,\partial_r)\right] \tag{0.29}$$

The distance r to the origin can be expressed in any coordinates we wish, yet this expression wants to single out the direction along $\mathbf{x} = r\,\hat{\mathbf{n}}$ from the other two. Also, it would be nice if $\mathbf{L}$ only involved derivatives in directions perpendicular to $\hat{\mathbf{n}}$. This is most easily realised in a spherical coordinate system, since its radial coordinate naturally correpsonds to the direction along $\mathbf{x}$; the other two coordinates are *angular*. By transforming the Cartesian components of $\mathbf{L}$ to spherical coordinates $(r,\,\theta,\,\phi)$, we obtain:

$$\begin{aligned}
L_x \;&=\; -\,\mathrm{i}\,(y\partial_z \;-\; z\partial_y) \;=\; -\,\mathrm{i}\,(-\sin\phi\,\partial_\theta \;-\; \cot\theta\,\cos\phi\,\partial_\phi)\\
L_y \;&=\; -\,\mathrm{i}\,(z\partial_x \;-\; x\partial_z) \;=\; -\,\mathrm{i}\,(\cos\phi\,\partial_\theta \;-\; \cot\theta\,\sin\phi\,\partial_\phi)\\
L_z \;&=\; -\,\mathrm{i}\,(x\partial_y \;-\; y\partial_x) \;=\; -\,\mathrm{i}\,\partial_\phi
\end{aligned}$$

The derivatives with respect to r have cancelled out, leaving only angular coordinates! We also find that:

$$\mathbf{L}^2 \;=\; L_x^2 + L_y^2 + L_z^2 \;=\; -\left[\frac{1}{\sin\theta}\,\partial_\theta\,(\sin\theta\,\partial_\theta) \;+\; \frac{1}{\sin^2\theta}\,\partial_\phi^2\right] \tag{0.30}$$

It is also useful to have the spherical components of $\mathbf{L}$. Obviously, $L_r = 0$ since $\mathbf{r} \cdot \mathbf{L}$. Using the transformations: $\hat{\boldsymbol{\theta}} = \mathbf{i}\,\cos\theta\,\cos\phi + \mathbf{j}\,\cos\theta\,\sin\phi - \mathbf{k}\,\sin\theta$, and $\hat{\boldsymbol{\phi}} = -\mathbf{i}\,\sin\phi + \mathbf{j}\,\cos\phi$, there comes:

$$\mathbf{L} \;=\; -\,\mathrm{i}\left(\hat{\boldsymbol{\phi}}\,\partial_\theta \;-\hat{\boldsymbol{\theta}}\,\frac{1}{\sin\theta}\,\partial_\phi\right) \tag{0.31}$$

Also, eq. (0.29) makes it obvious that the **commutator** $[\nabla^2,\,\mathbf{L}^2] \coloneqq \nabla^2\mathbf{L}^2 - \mathbf{L}^2\nabla^2 = 0$, so that $[\nabla^2,\,\mathbf{L}] = 0$.

We establish five vector identities involving $\mathbf{L}$ and ∇^2. With $\mathbf{r}$ the radial coordinate in a spherical system:

$$\mathbf{r}\cdot(\boldsymbol{\nabla}\times\mathbf{V}) \;=\; (\epsilon_{ijk}x^i\partial^j)\,V^k \;=\; (\mathbf{r}\times\boldsymbol{\nabla})\cdot\mathbf{V} \;=\; \mathrm{i}\,\mathbf{L}\cdot\mathbf{V} \tag{0.32}$$

$$\nabla^2(\mathbf{r}\cdot\mathbf{V}) \;=\; \partial_i\partial^i(x_jV^j) \;=\; \partial_i(\delta^i{}_jV^j + x_j\partial^i\,V^j) \;=\; \partial_iV^i + \partial_iV^i + x_j\partial_i\partial^iV^j \;=\; 2\boldsymbol{\nabla}\cdot\mathbf{V} + \mathbf{r}\cdot\nabla^2\mathbf{V} \tag{0.33}$$

$$\boldsymbol{\nabla}\cdot(\mathbf{r}\times\mathbf{V}) \;=\; \epsilon_{ijk}\partial^i(x^jV^k) \;=\; \epsilon_{ijk}(\delta^{ij}V^k \;+\; x^j\partial^iV^k) \;=\; -\,\mathbf{r}\cdot(\boldsymbol{\nabla}\times\mathbf{V}) \;=\; -\mathrm{i}\,\mathbf{L}\cdot\mathbf{V} \tag{0.34}$$

Taking $\mathbf{V} \longrightarrow \boldsymbol{\nabla}\times\mathbf{V}$, this last identity reads: $\mathbf{L}\cdot\boldsymbol{\nabla}\times\mathbf{V} = \mathrm{i}\,\boldsymbol{\nabla}\cdot\big[\mathbf{r}\times(\boldsymbol{\nabla}\times\mathbf{V})\big]$. Using an identity in Jackson's front left cover, we have, with $\boldsymbol{\nabla}\times\mathbf{r} = 0$: $\mathbf{r}\times(\boldsymbol{\nabla}\times\mathbf{V}) = \boldsymbol{\nabla}(\mathbf{r}\cdot\mathbf{V}) - (\mathbf{r}\cdot\boldsymbol{\nabla})\mathbf{V} - (\mathbf{V}\cdot\boldsymbol{\nabla})\mathbf{r}$. Then there comes:

$$\mathbf{L}\cdot\boldsymbol{\nabla}\times\mathbf{V} = \mathrm{i}\left[\nabla^2(\mathbf{r}\cdot\mathbf{V}) - \partial_i(x^j\partial_j)V^i) \;-\; \boldsymbol{\nabla}\cdot\mathbf{V}\right] = \mathrm{i}\left[\nabla^2(\mathbf{r}\cdot\mathbf{V}) - 2\boldsymbol{\nabla}\cdot\mathbf{V} - (\mathbf{r}\cdot\boldsymbol{\nabla})\boldsymbol{\nabla}\cdot\mathbf{V}\right] \tag{0.35}$$

Our last identity involves $\mathrm{i}\,\boldsymbol{\nabla}\times\mathbf{L} = \boldsymbol{\nabla}\times(\mathbf{r}\times\boldsymbol{\nabla})$ or, in index notation: $\epsilon_{ijk}\epsilon^{lmk}\partial^j(x_l\partial_m) = (\delta_i{}^l\delta_j{}^m - \delta_i{}^m\delta_j{}^l)\partial^j(x_l\partial_m)$. Expanding yields: $\partial^m(x_i\partial_m) - \partial^l(x_l\partial_i) = x_i\partial^m\partial_m - 2\partial^i - (x_l\partial^l)\partial_i = x_i\partial^m\partial_m - \partial^i - \partial_i(x_l\partial^l)$. When acting on a scalar function f, this gives in coordinate-free notation:

$$\mathrm{i}\,\boldsymbol{\nabla}\times\mathbf{L}\,f \;=\; \mathbf{r}\,\nabla^2 f \;-\; \boldsymbol{\nabla}(f \;+\; r\partial_r f) \tag{0.36}$$

0.8.1 Eigenvalues of $\mathbf{L}^2$ and L_z

Now one readily shows that the following important commutation relations hold:

$$[L_x,\,L_y] \;=\; \mathrm{i}\,L_z, \qquad [L_y,\,L_z] \;=\; \mathrm{i}\,L_x, \qquad [L_z,\,L_x] \;=\; \mathrm{i}\,L_y, \qquad [\mathbf{L}^2,\,\mathbf{L}] \;=\; 0 \tag{0.37}$$

The first three can be written in the compact form: $\mathbf{L}\times\mathbf{L} = \mathrm{i}\,\mathbf{L}$. The importance of these relations cannot be overstated. It says that $\mathbf{L}$ belongs to the class of self-adjoint operators $\mathbf{J}$ whose Cartesian components satisfy the canonical commutation relations:

$$[J_i,\,J_j] \;=\; \mathrm{i}\,\epsilon_{ijk}\,J_k, \tag{0.38}$$

Just from these properties, it is possible to derive the eigenvalues λ of $\mathbf{J}^2$, such that $\mathbf{J}^2 g \;=\; \lambda\,g$, and the eigenvalues m of J_z, such that $J_z f = m\,f$, where λ and m are *real* numbers since the operators are self-adjoint.

Introduce the **ladder** operators $J_\pm = J_x \pm \mathrm{i}J_y$. with coomutation relations $[\mathbf{J}^2,\,J_\pm] = 0$. $[J_z,\,J_\pm] = \pm\,J_\pm$, and $[J_+,\,J_-] = 2J_z$.

$$J_z(J_\pm f) \;\equiv\; [J_z,\,J_\pm]\,f \;+\; J_\pm J_z f \;=\; (m \;\pm\; 1)\,(J_\pm f)$$

Now f is also an eigenfunction of $\mathbf{J}^2$, and $J_z\mathbf{J}^2 f = \mathbf{J}^2 J_z f = m\mathbf{J}^2 f$, so that $\mathbf{J}^2 f$ is an eigenfunction of J_z with the same eigenvalue. Since the eigenvalues of J_z are non-degenerate, $\mathbf{J}^2 f$ is a multiple of f: $\mathbf{J}^2 f = \lambda f$. Also:

$$\mathbf{J}^2(J_\pm f) \;=\; J_\pm(\mathbf{J}^2 f) \;=\; \lambda\,(J_\pm f)$$

These results tell us what $J_\pm$ do for a living: they raise (J_+) or lower (J_-) the eigenvalues of J_z by 1, whence their name. In other words, if f is an eigenfunction of J_z with eigenvalue m, so is $J_\pm f$, but with eigenvalue $(m \pm 1)$. We also have found that *all* the eigenfunctions of J_z reachable with the ladder operators are eigenfunctions of $\mathbf{J}^2$ as well, with the *same* eigenvalue λ.

We can also use the definition of $J_\pm$ to show the identity:

$$\mathbf{J}^2 \;=\; J_\pm\, J_\mp \;+\; J_z^2 \;\mp\; J_z \tag{0.39}$$

For a given value of λ, we expect that m should have a maximum value, $m_{\text{max}} \equiv j$, as well as a minimum value, $m_{\text{min}} \equiv j'$. Now act with $\mathbf{J}^2$ on the eigenfunction of J_z with the maximum value of m, which we call f_j. Then $\mathbf{J}_+ f_j = 0$ and, from the identity (0.39), we find:

$$\mathbf{J}^2 f_j \;=\; J_z^2 f_j \;+\; J_z f_j \;=\; j(j+1)\, f_j \;=\; \lambda f_j$$

Similarly, act with $\mathbf{J}^2$ on the eigenfunction of $\mathbf{J}_z$ with the minimum value of m, $f_{j'}$, keeping in mind that $f_{j'}$ is also an eigenfunction of $\mathbf{J}^2$ with the *same* eigenvalue, λ, as f_j:

$$\mathbf{J}^2 f_{j'} \;=\; (j')^2 f_{j'} \;-\; j' f_{j'} \;=\; j'(j'-1)\, f_{j'} \;=\; \lambda f_{j'}$$

Comparing yields $\lambda \;=\; j(j+1) \;=\; j'(j'-1)$, and thus $j' = -j$. It follows that m goes from $-j$ to j in N integer steps, ie, $j = -j + N$, so $j = N/2$.

We conclude that:

- The eigenvalues of $\mathbf{J}^2$ are $j(j+1)$, where j is a positive integer or a half-integer.
- For a given value of j, m can take $2j+1$ values, from $-j$ to j.

With the help of eq. (0.39), we can now exhibit the full action of J_- on a normalised eigenfunction f_{jm} of $\mathbf{J}^2$ and J_z. Let $J_- f_{jm} = c_-\, f_{j,m-1}$. Then, using the rules for taking adjoints, and with $(f,\, g)$ the inner product of f and g:

$$(f_{jm},\, J_+ J_-\, f_{jm}) \;=\; (J_- f_{jm},\, J_-\, f_{jm}) \;=\; (c_- f_{j,m-1},\, c_- f_{j,m-1}) \;=\; |c_-|^2\, (f_{j,m-1},\, f_{j,m-1}) \;=\; |c_-|^2$$

But since $J_\pm\, J_\mp = \mathbf{J}^2 - J_z^2 \,\pm\, J_z$, we also have that:

$$(f_{lm},\, J_+ J_-\, f_{jm}) \;=\; (f_{jm},\, (\mathbf{J}^2 - J_z^2 + \, J_z)\, f_{jm}) \;=\; j(j+1) \;-\; m^2 \;+\; m$$

Comparing yields c_- up to an unimportant exponential phase factor which we put equal to 1. We find the coefficient in $J_+ f_{jm} = c_+\, f_{j,m+1}$ in a strictly analogous way. The results for both ladder operators are:

$$J_\pm f_{jm} \;=\; \sqrt{j(j+1) \;-\; m(m \,\pm 1)}\, f_{j,m\pm1} \tag{0.40}$$

0.8.2 Eigenfunctions of $\mathbf{L}^2$ and $\mathbf{L}_z$

To find the common eigenfunctions for $\mathbf{J}$ and $\mathbf{J}^2$ operators, we must know what they look like. Here, we will be interested in the $\mathbf{L}$ operator whose form we do know and which makes up the angular part of the Laplacian in spherical coordinates.

The eigenfunctions of L_z are readily obtained by solving the differential equation:

$$L_z f(\theta,\phi) \;=\; -\mathrm{i}\, \partial_\phi f(\theta,\phi) \;=\; m\, f(\theta,\phi)$$

With a separation ansatz: $f(\theta,\phi) = F(\theta)G(\phi)$, the solution for G is: $G(\phi) = \mathrm{e}^{\mathrm{i}m\phi}$. Require that G (and f) be single-valued, that is, $G(\phi + 2\pi) = G(\phi)$ leads to:

$$\mathrm{e}^{\mathrm{i}m(\phi+2\pi)} \;=\; \mathrm{e}^{\mathrm{i}m\phi} \quad \Longrightarrow \mathrm{e}^{2\mathrm{i}m\pi} \;=\; \cos 2m\pi \;+\; \mathrm{i}\, \sin 2m\pi \;=\; 1$$

which constrains m to be any *integer*. Therefore, $l \coloneqq m_{\max}$ must also be an integer. Thus, we find that the particular form $\mathbf{L} = -\mathrm{i}\mathbf{x} \times \boldsymbol{\nabla}$ rules out the possibility of half-integer values of j allowed for a self-adjoint $\mathbf{J}$ that satisfies the canonical commutation relations (0.38).

The θ dependence of the eigenfunctions must be derived from the eigenvalue equation for $\mathbf{L}^2$. Call $f(\theta,\phi) = Y_l^m(\theta,\phi) = F(\theta)G(\phi)$; these must satisfy:

$$-\left[\frac{1}{\sin\theta}\,\partial_\theta\left(\sin\theta\,\partial_\theta\right) \;+\; \frac{1}{\sin^2\theta}\,\partial_\phi^2\right] Y_l^m(\theta,\phi) \;=\; l(l+1)\,Y_l^m(\theta,\phi)$$

Inserting $Y_l^m(\theta,\phi) = F(\theta)\mathrm{e}^{\mathrm{i}m\phi}$ into this equation leaves:

$$-\left[\frac{1}{\sin\theta}\,\mathrm{d}_\theta\left(\sin\theta\,\mathrm{d}_\theta\right) \;-\; \frac{m^2}{\sin^2\theta}\right] F(\theta) \;=\; l(l+1)\,F(\theta) \tag{0.41}$$

Instead of solving this equation by brute force, we use a clever technique involving the ladder operators $L_\pm$:

$$L_\pm \;=\; \pm\,\mathrm{e}^{\mathrm{i}\phi}\left(\partial_\theta \;\pm\; \mathrm{i}\,\cot\theta\,\partial_\phi\right)$$

Now, when $m = l$, we have:

$$L_+ Y_l^l \;=\; \mathrm{e}^{\mathrm{i}\phi}\left(\partial_\theta \;+\; \mathrm{i}\,\cot\theta\,\partial_\phi\right) Y_l^l(\theta,\phi) \;=\; 0$$

Inserting $Y_l^l = F(\theta)\mathrm{e}^{\mathrm{i}l\phi}$, this reduces to the much simpler

$$\mathrm{d}_\theta F(\theta) \;-\; l\,\cot\theta\,F(\theta) \;=\; 0$$

whose solution is $F(\theta) = (\sin\theta)^l$. Therefore, $Y_l^l = (\sin\theta)^l\mathrm{e}^{\mathrm{i}l\phi}$. Applying L_- the requisite number of times generates the other Y_l^m $(0 < m < l)$: $Y_l^m \propto L_-^{l-m} Y_l^l$. When normalised, these are the **spherical harmonics**:

$$Y_l^m(\theta,\phi) \;=\; \frac{(-1)^m}{2^l l!}\sqrt{\frac{2l+1}{4\pi}\,\frac{(l-m)!}{(l+m)!}}\,(1-x^2)^{m/2}\left[\mathrm{d}_x^{l+m}(x^2-1)^l\right]\mathrm{e}^{\mathrm{i}m\phi} \qquad x = \cos\theta \tag{0.42}$$

Spherical harmonics are tabulated on p. 109 of Jackson for $l \leq 3$. They obey:

$$Y_{l,-m}^*(\theta,\phi) \;=\; (-1)^m Y_{lm}(\theta,\phi) \tag{0.43}$$

They satisfy other useful relations which can be found in several references (Arfken's *Mathematical Methods for Physicists* and *Handbook of Mathematical Functions* by Abramowitz and Stegun are two popular ones). They occur in the solution to many problems in physics: in quantum mechanics for instance, when the potential in the Schrödinger equation is spherically-symmetric, the angular dependence of the wave functions is always given by spherical harmonics. And we shall see later how useful they are in electrostatics.

The spherical harmonics form a complete set of orthonormal functions, in the sense that any function $g(\theta,\phi)$ can be written as:

$$g(\theta,\phi) \;=\; \sum_{l=0}^{\infty}\sum_{m=-l}^{l} A_{lm}Y_{lm}(\theta,\phi)$$

where

$$A_{lm} \;=\; \oint f(\theta,\phi)\,Y_{lm}^*(\theta,\phi)\,\mathrm{d}\Omega$$

with $\mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi$, and where we have used the orthonormality condition:

$$\int_0^\pi \sin\theta\,\mathrm{d}\theta\int_0^{2\pi} Y_{l'm'}^*(\theta,\phi)Y_{lm}(\theta,\phi)\,\mathrm{d}\phi \;=\; \delta_{l'l}\delta_{m'm}$$

The spherical harmonics satisfy the completeness relation:

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}^{*}(\theta',\phi')Y_{lm}(\theta,\phi) \;=\; \delta(x-x')\delta(\phi-\phi') \qquad\qquad (x=\cos\theta) \tag{0.44}$$

In axisymmetric situations, with the the z-axis chosen along the axis of symmetry, there is no ϕ dependence, which means that $m-0$. Then the spherical harmonics collapse to the **Legendre polynomials** of order l, given by the **Rodrigues formula**:

$$P_l(x) \;=\; \frac{1}{2^l l!} d_x^l (x^2-1)^l \qquad\qquad (-1\le x\le 1) \tag{0.45}$$

They satisfy the **Legendre equation**, which is simply eq. (0.41) with $m=0$. They are given on p. 97 of Jackson for $l\le 4$. They satisfy a number of relations, some of which are given on p. J100.

One word of caution here: this treatment presupposes *no restriction* on either θ or ϕ! See section 3.4 in Jackson for modifications needed when θ is restricted to a range smaller than π in a solution independent of ϕ.

0.8.3 General Separable Solution of the Laplace and Helmholtz Equations

Suppose we are presented with the equation $\left[\nabla^2+\gamma(\mathbf{x})\right]\Psi(\mathbf{x})=0$. Work in spherical coordinates, and make the **separation ansatz**: $\Psi(\mathbf{x})=R(r)F(\theta,\phi)$. Using the form for ∇^2 derived earlier, eq. (0.29), we write:

$$\begin{aligned}\nabla^2\Psi \;+\; \gamma(\mathbf{x})\Psi \;&=\; -\frac{\mathbf{L}^2\Psi}{r^2} \;+\; \frac{1}{r}\left[\partial_r\Psi \;+\; \partial_r(r\,\partial_r\Psi)\right] \;+\; \gamma(\mathbf{x})\,\Psi \\ &= -\,R(r)\frac{\mathbf{L}^2F(\theta,\phi)}{r^2} \;+\; \frac{F(\theta,\phi)}{r}\left[\mathrm{d}_rR(r) \;+\; \mathrm{d}_r(r\,\mathrm{d}_rR(r))\right] \;+\; \gamma(\mathbf{x})\,R(r)F(\theta,\phi)\end{aligned}$$

Dividing the second line by $R(r)F(\theta,\phi)$ and multiplying by r^2, we see that the equation is separable provided $\gamma(\mathbf{x})=\gamma(r)$:

$$\mathbf{L}^2F(\theta,\phi) \;=\; \lambda\,F(\theta,\phi) \qquad\qquad \mathrm{d}_rR(r) \;+\; \mathrm{d}_r(r\,\mathrm{d}_rR(r)) \;+\; r\,\gamma(r)\,R(r) \;=\; \lambda\,\frac{R(r)}{r}$$

The first equation is the eigenvalue equation for $\mathbf{L}^2$, whose eigenvalues are $\lambda=l(l+1)$ ($l\ge 0\in\mathbb{Z}$) with the spherical harmonics $Y_l^m(\theta,\phi)$ as eigenfunctions.

The radial equation can thus be written: $\dfrac{1}{r^2}\,\mathrm{d}_r\!\left(r^2\,\mathrm{d}_rR_l(r)\right) \;+\; \left(\gamma(r) \;-\; \dfrac{l(l+1)}{r^2}\right) R_l(r) \;=\; 0$

When $\gamma(r)=0$, this is the radial part of the Laplace equation which becomes, after the change of variable $r=\mathrm{e}^x$, $\mathrm{d}_x^2R+\mathrm{d}_xR-l(l+1)R=0$. Inserting a solution of the form e^{px} turns the equation into $p^2+p-l(l+1)=0$, that is, $p=l$ or $p=-(l+1)$, which leads to $R=A\mathrm{e}^{lx}\;+\;B\mathrm{e}^{-(l+1)x}=Ar^l+Br^{-(l+1)}$. Therefore, the general solution to the Laplace equation in spherical coordinates is:

$$\Psi(r,\theta,\phi) \;=\; \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left(A_{lm}\,r^l \;+\; \frac{B_{lm}}{r^{l+1}}\right) Y_l^m(\theta,\phi) \tag{0.46}$$

If Ψ is specified over a surface of constant $r=a$, the coefficients A_{lm} and B_{lm} can be found. Very often, one looks for solutions inside and outside the sphere, $\Psi_{r<a}$ and $\Psi_{r>a}$. For $r\le a$, $B_{lm}^{r<a}=0$ so as to prevent a divergence at $r=0$, and:

$$A_{lm}^{r<a} \;=\; \frac{1}{a^l}\oint \Psi(a,\theta,\phi)Y_{lm}^{*}(\theta,\phi)\,\mathrm{d}\Omega \tag{0.47}$$

For $r \geq a$, and assuming that $\Psi \to 0$ at infinity, $A_{lm}^{r>a} = 0$ to avoid a divergence as $r \to \infty$, and:

$$B_{lm}^{r>a} = a^{l+1} \oint \Psi(a,\theta,\phi) Y_{lm}^*(\theta,\phi)\, \mathrm{d}\Omega \tag{0.48}$$

If either one of these two solutions is known, we can use the **matching condition** $\Psi_{r<a}(a,\theta,\phi) = \Psi_{r>a}(a,\theta,\phi)$ to relate the coefficients: $B_{lm}^{r>a} = a^{2l+1} A_{lm}^{r<a}$.

Clearly, if this solution is to be regular, and if it holds *everywhere*, it must vanish. In other words, if the Laplace equation is valid everywhere, it has no non-vanishing regular solution. For a non-trivial solution, there must be a region of space where there exists an inhomogeneous term acting as a *source*.

Note, however, that the general solution holds at any point where there is no source. The effect of sources is encoded in the coefficients A_{lm} and B_{lm}.

When $\gamma(r) = k^2 > 0$, we get the radial part of the Helmholtz equation in spherical coordinates:

$$\mathrm{d}_r^2 R_l(r) + \frac{2}{r}\,\mathrm{d}_r R_l(r) + \left(k^2 - \frac{l(l+1)}{r^2}\right) R_l(r) = 0$$

The substitutions $R_l = u_l/\sqrt{r}$ and $x = kr$ readily transform it into:

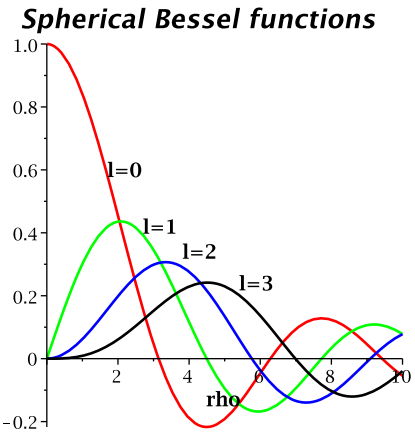
$$\mathrm{d}_x^2 u_l(x) + \frac{1}{x}\,\mathrm{d}_x u_l(x) + \left(1 - \frac{(l+1/2)^2}{x^2}\right) u_l(x) = 0$$

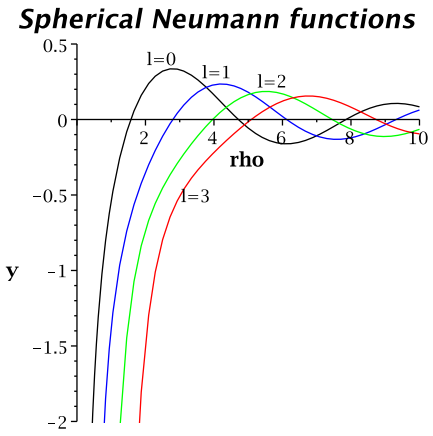
which is a form of the Bessel equation. The solutions are the **spherical Bessel functions** of the first and second (Neumann) kind, usually written as (see also Jackson's *Classical Electrodynamics*, section 9.6):

$$j_l(x) = \sqrt{\frac{\pi}{2x}}\, J_{l+1/2}(x) = (-x)^l \left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^l \left(\frac{\sin x}{x}\right) \sim \begin{cases} x^l & x \ll (1,l) \\ \dfrac{1}{x}\sin(x - l\pi/2) & x \gg l \end{cases} \tag{0.49}$$

$$n_l(x) = \sqrt{\frac{\pi}{2x}}\, N_{l+1/2}(x) = -(-x)^l \left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^l \left(\frac{\cos x}{x}\right) \sim \begin{cases} -\dfrac{1}{x^{l+1}} & x \ll (1,l) \\ -\dfrac{1}{x}\cos(x - l\pi/2) & x \gg l \end{cases} \tag{0.50}$$

Here are a few spherical Bessel and Neumann functions as plotted on *Maple*, with $\rho = x$:





The n_l diverge at the origin and thus are excluded from any solution regular at the origin.

(Spherical) Bessel functions of the third kind, aka **Hankel** functions of the first and second kind, sometimes come in handy: $h_l^{(1,2)}(x) = j_l(x) \pm \mathrm{i}\, n_l(x)$. One can express the general solution of the Helmholtz equation in terms of the j_l and n_l, or in terms of the $h_l^{(1,2)}$.

0.9 The Convective Derivative

It can happen that we wish to know how some quantity $f(\mathbf{x}, t)$ varies with time *if* $\mathbf{x}$ *is a position* $\mathbf{x}_0$ *that moves at velocity* $\mathbf{u} = \mathrm{d}_t\mathbf{x}$. Then $f(\mathbf{x}_0(t), t)$ becomes a function of t only, and we can take its ordinary (or **convective/total**) derivative:

$$\begin{aligned} \mathrm{d}_t f(\mathbf{x}_0(t), t) &= \partial_t f + \left[\mathrm{d}_t\mathbf{x}_0 \cdot \boldsymbol{\nabla}\right] f \\ &= \partial_t f + (\mathbf{u} \cdot \boldsymbol{\nabla})\, f \end{aligned} \tag{0.51}$$

Using identities on the left of the inside front cover of Jackson, we see that if f is a scalar, $(\mathbf{u} \cdot \boldsymbol{\nabla})\, f = \boldsymbol{\nabla} \cdot (\mathbf{u}\, f)$. But if f is a vector,

$$\mathrm{d}_t\mathbf{f}(\mathbf{x}_0(t), t) = \partial_t f + \mathbf{u}\,\boldsymbol{\nabla} \cdot \mathbf{f} - \boldsymbol{\nabla} \times (\mathbf{u} \times \mathbf{f}) \tag{0.52}$$

1 Review of Electrostatics

The first 11 sections (you can omit 1.12 and 1.13) of this chapter in Jackson should only be revision; little is done there that you shouldn't have seen before (presumably...). Therefore, I will assign it for you to read carefully, and will only go over the salient points here.

1.1 Electrostatic Field (sections J1.1–J1.4)

The **Coulomb electric field**, whose **source** is a point-charge q, located at position $\mathbf{x}'$ with respect to an arbitrarily chosen origin, is given by:

$$\mathbf{E}(\mathbf{x}) \;=\; k_{\rm e}\, q\, \frac{\mathbf{x}-\mathbf{x}'}{|\mathbf{x}-\mathbf{x}'|^3} \tag{1.1}$$

where $\mathbf{x}-\mathbf{x}'$ is the distance vector from the source at $\mathbf{x}'$ to the **observation (or field) point** $\mathbf{x}$, In SI units, the **Coulomb constant** $k_{\rm e} = 1/4\pi\epsilon_0 = 8.99\times10^9\ \mathrm{N\cdot m^2/C^2}$; in CGS units it is simply equal to 1. In the first seven chapters of these notes, whenever I deviate from writing the general $k_{\rm e}$, I will adopt SI units as is done in the first ten chapters of Jackson.

The electric field obeys the **principle of superposition**. When the source is a collection of N point-charges q_i sitting at their respective positions $\mathbf{x}_i$, their electric fied at $\mathbf{x}$ is:

$$\mathbf{E}(\mathbf{x}) \;=\; k_{\rm e}\sum_{\rm i=1}^{N} q_{\rm i}\, \frac{\mathbf{x}-\mathbf{x}'_{\rm i}}{|\mathbf{x}-\mathbf{x}'_{\rm i}|^3}$$

This is easily extended to **continuous** source charge distributions. We only give the expression for a distribution extending over a volume:

$$\mathbf{E}(\mathbf{x}) \;=\; k_{\rm e}\int\limits_{\rm volume} \rho(\mathbf{x}')\frac{\mathbf{x}-\mathbf{x}'}{|\mathbf{x}-\mathbf{x}'|^3}\,\mathrm{d}^3x' \tag{1.2}$$

where ρ is the **charge density** in the distribution. Note that only Cartesian components can be used in the integral, because only then are the unit vectors in $\mathbf{x}-\mathbf{x}'$ constant with respect to integration, so that they can be pulled out of the integral. The remaining three integrals can then be evaluated in any coordinate basis you wish.

If a charge Q is placed at the observation point $\mathbf{x}$, it experiences a force $\mathbf{F} = Q\mathbf{E}$ due to the electric field, whatever the source of the field may be.

Note also that you can always write a discrete distribution of N point-charges at positions x_i as:

$$\rho(\mathbf{x}) \;=\; \sum_{\rm i=1}^{N} q_{\rm i}\,\delta(\mathbf{x}-\mathbf{x}_{\rm i})$$

The Coulomb field, as well as *any* electric field, satisfies **Gauss' law** in differential form:

$$\boldsymbol{\nabla}\cdot\mathbf{E} \;=\; 4\pi k_{\rm e}\,\rho \quad (=\ \rho/\epsilon_0 \text{ in SI units}) \tag{1.3}$$

or, via the divergence theorem, in the equivalent integral form:

$$\oint\limits_{\rm surface} \mathbf{E}\cdot\mathrm{d}\mathbf{a} \;=\; 4\pi k_{\rm e}\, q_{\rm in} \tag{1.4}$$

where $q_{\rm in} = \int \rho(\mathbf{x}')\,\mathrm{d}^3x'$ is the total net charge enclosed by the surface.

The field of any *static* charge distribution can be written as a superposition of Coulomb fields. We call such fields **electrostatic**.

1.2 Scalar Potential (sections J1.5–J1.7)

Because it is a central field, the Coulomb field has vanishing curl. Therefore, by the superposition principle, *all electrostatic fields satisfy* $\boldsymbol{\nabla} \times \mathbf{E} = 0$ *everywhere*. This is equivalent to saying that $\oint \mathbf{E} \cdot d\mathbf{l} = 0$ *for any closed integration path*.

Because $\boldsymbol{\nabla} \times \mathbf{E} = 0$ for any electrostatic field $\mathbf{E}$, there exists associated with this vector field a *scalar* field $\Phi(\mathbf{x})$ such that $\mathbf{E} = -\boldsymbol{\nabla}\Phi$. From the fundamental theorem for gradients (section 0.3):

$$\int_a^b \mathbf{E} \cdot d\mathbf{l} \;=\; -\int_a^b (\boldsymbol{\nabla}\Phi) \cdot d\mathbf{l} \;=\; \Phi(a) \;-\; \Phi(b) \tag{1.5}$$

Therefore, the **electrostatic potential** at any point P is $\Phi(P) = -\int_{\mathcal{O}}^{P} \mathbf{E} \cdot \mathrm{d}\mathbf{l}$, where $\mathcal{O}$ is an arbitrary reference point, usually chosen where we want the potential to vanish.

The **Coulomb potential** of a point-charge q is $\Phi(\mathbf{x}) = k_{\mathrm{e}} q/|\mathbf{x} - \mathbf{x}'|$.

The potential Φ, like the field $\mathbf{E}$, obeys the superposition principle. Then the potential of a *localised* charge distribution ρ is:

$$\Phi(\mathbf{x}) \;=\; k_{\mathrm{e}} \int \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}\, \mathrm{d}^3 x' \tag{1.6}$$

where the reference point is taken at infinity, hence the restriction to localised distributions.

Combining $\mathbf{E} = -\boldsymbol{\nabla}\,\Phi$ with Gauss' law yields the **Poisson equation**:

$$\nabla^2 \Phi \;=\; -\,4\pi k_{\mathrm{e}}\, \rho \tag{1.7}$$

In vacuum, this becomes the **Laplace equation**, $\nabla^2 \Phi = 0$.

1.3 Matching Conditions on Electrostatic E and Φ

Let $\hat{\mathbf{n}}$ be a unit vector normal to the boundary of some volume, directed from the "in" region to the " out" region. The electrostatic-field component tangent to any surface is *continuous* across the surface:

$$(\mathbf{E}_{\text{out}} - \mathbf{E}_{\text{in}}) \times \hat{\mathbf{n}} \;=\; 0 \tag{1.8}$$

For *conductors*, on the other hand, the normal component is discontinuous when there is a local surface charge density σ:

$$(\mathbf{E}_{\text{out}} \;-\; \mathbf{E}_{\text{in}}) \cdot \hat{\mathbf{n}} \;=\; =\; 4\pi k_{\mathrm{e}}\, \sigma \tag{1.9}$$

and the local field *due* to σ has magnitude $2\pi k_{\mathrm{e}}\sigma$, so that the *total* field has to jump by $4\pi k_{\mathrm{e}}\sigma$ across the surface.

The potential is continuous across the boundary of a conductor, but its gradient in the direction normal to the surface inherits the discontinuity in $\mathbf{E}$:

$$\partial_n \Phi_{\text{out}} \;-\; \partial_n \Phi_{\text{in}} \;=\; -\,4\pi k_{\mathrm{e}}\, \sigma \tag{1.10}$$

where the normal derivative , $\partial_n \Phi = (\boldsymbol{\nabla}\,\Phi) \cdot \hat{\mathbf{n}}$, is evaluated at the surface.

For non-conductors, we will have to be a bit careful (see chapter 4 in Jackson) with the normal components.

1.4 Boundary-Value Problem with Green Functions; Uniqueness of a Solution (sections J1.9–J1.10)

In realistic electrostatic problems, we rarely know the whole charge distribution that produces the potential observed at a point. What we most often do know (and often can control) is the *potential* over specific surfaces such as conductors. These surfaces can enclose regions with or without charge inside. The proper approach then involves solving the Poisson or Laplace equation, subject to **boundary conditions** (B.C.).

There are two very important questions that must be answered:

- What kind of information do we need to supply in order to obtain a solution to the Poisson equation?
- If we find a solution, is it **unique**?

1.4.1 Uniqueness of the solution

It is convenient to address the first question before the second: is a solution of the Poisson equation that satisfies boundary conditions (B.C.) on a closed surface unique?

Suppose there exist two solutions, Φ_1 and Φ_2, of Poisson's equation that satisfy the same B.C., on a closed surface. Define $\Phi_3 \equiv \Phi_2 - \Phi_1$. From Green's first identity (eq. 0.6) with $f = g = \Phi_3$, we have:

$$\int_{\text{volume}} [\Phi_3 \nabla^2 \Phi_3 \,+\, (\nabla \Phi_3)^2] \, \mathrm{d}^3 x' \;\equiv\; \oint_{\text{surface}} \Phi_3 \, \partial_{n'} \Phi_3 \, \mathrm{d}a'$$

The surface integral is zero because either $\Phi_3 = 0$ (because we have specified the potential) or $\partial_n \Phi_3 = 0$ on the surface (because we have specified the normal derivative). Also, $\nabla^2 \Phi_3 = 0$ inside the volume. Therefore, $\int (\nabla \Phi_3)^2 \mathrm{d}^3 x = 0$, which can only be true if $\nabla \Phi_3 = 0$ inside the volume, so that Φ_3 is a constant there. Then, if $\Phi_3 = 0$ on the surface, it vanishes everywhere inside, whereas Φ_3 can be a non-zero constant inside if $\partial_n \Phi$ is specified on the surface.

We conclude that $\Phi_1 = \Phi_2$ inside (up to a possible irrelevant additive constant), and that the electrostatic field is *uniquely* determined. The importance of this result cannot be exaggerated. It means that any function that satisfies Poisson's equation and the B.C. is *the* solution, no matter how it was found! It also means that if Φ is specified on the surface, its normal derivative is also determined and cannot be specified arbitrarily; if it is $\partial_n \Phi$ that is specified, then we lose almost all control over Φ which is determined up to a constant.

1.4.2 Solution of boundary-value problems with Green functions (J1.8–10)

To answer the first question, we introduce functions $G(\mathbf{x}, \mathbf{x}')$, called **Green functions**.

A Green function G for a linear differential operator L is defined as a solution of the equation $\mathsf{L}\, G(\mathbf{x}, \mathbf{x}') = -4\pi\delta(\mathbf{x} - \mathbf{x}')$, where $\mathbf{x}$ and $\mathbf{x}'$ are two points in the manifold on which the functions are defined. Thus, in 3-dimensional space, the Green function for the Laplacian satisfies:

$$\nabla^2 G(\mathbf{x}, \mathbf{x}') \;=\; -\,4\pi\delta(\mathbf{x} - \mathbf{x}') \tag{1.11}$$

Now take $f = \Phi$ (Φ is the potential) and $g = G$ in Green's second identity (eq. (0.7)):

$$\int_{\text{volume}} \left(\Phi \, \nabla^2 G - G \, \nabla^2 \Phi\right) \mathrm{d}^3 x' \;\equiv\; \oint_{\text{surface}} \left(\Phi \, \partial_{n'} G \,-\, G \, \partial_{n'} \Phi\right) \mathrm{d}a'$$

After using the Poisson equation in the second volume integral we obtain:

$$\int_V -\,4\pi \left[\Phi(\mathbf{x}')\delta(\mathbf{x} - \mathbf{x}') \,+\, k_{\mathrm{e}} \rho(\mathbf{x}') G(\mathbf{x}, \mathbf{x}')\right] \mathrm{d}^3 x' \;=\; \oint_S \left(\Phi \, \partial_{n'} G \,-\, G \, \partial_{n'} \Phi\right) \mathrm{d}a'$$

If $\mathbf{x}$ lies outside the arbitrary volume, the first volume integral is zero. If $\mathbf{x}$ lies inside the volume, we can extend the first volume integral to cover all space without changing anything, and integrate. Rearranging then yields:

$$\Phi(\mathbf{x}) \;=\; k_{\rm e}\int_V \rho(\mathbf{x}')G(\mathbf{x},\mathbf{x}')\,{\rm d}^3x' \;+\; \frac{1}{4\pi}\oint_S \left(G\,\partial_{n'}\Phi \;-\; \Phi\,\partial_{n'}G\right){\rm d}a' \tag{1.12}$$

where ρ is the charge density and S is the surface boundary of volume V. The normal derivatives are to be evaluated *on* S before integrating.

You may think that we have answered the first question: specifying ρ over a volume and Φ and its normal derivative as B.C. over the surface enclosing the volume gives the potential everywhere via the above equation. There is a subtle and important point here, however: as we have seen above, Φ and $\partial_n\Phi$ are *not independent on the surface*. So we are not free to specify them both arbitrarily as such values will in general be inconsistent. We may only specify one *or* the other! In that sense, our expression for Φ should not be considered a solution yet, but just an integral equation. After all, the surface integrals come from an identity.

Specifying Φ on the surface gives **Dirichlet B.C.**, whereas specifying $\partial_n\Phi$ gives **Neumann B.C.** In the case of conducting surfaces, the latter are equivalent to specifying the surface charge density $\sigma = -\epsilon_0\partial_n\Phi$.

How do we get a solution for Φ then? In principle, this is simple. To a particular solution of the defining equation (1.11), we can add any function $F(\mathbf{x},\mathbf{x}')$ that satisfies the Laplace equation: $\nabla^2 F(\mathbf{x},\mathbf{x}') = 0$, to obtain the most general Green function for the Laplacian. "All" we have to do is find a F that eliminates one of the two surface integrals.

For instance, uppose we wish to specify Φ freely on the surface S (Dirichlet problem). Then we should have $G_{\rm D}(\mathbf{x},\mathbf{x}') = 0 \;\forall\,\mathbf{x}\in\; S$. The *solution* for Φ would then be:

$$\Phi(\mathbf{x}) \;=\; k_{\rm e}\int_V \rho(\mathbf{x}')G_{\rm D}(\mathbf{x},\mathbf{x}'){\rm d}^3x' \;-\; \frac{1}{4\pi}\oint_S \Phi(\mathbf{x}')\partial_{n'}G_{\rm D}(\mathbf{x},\mathbf{x}')\,{\rm d}a' \tag{1.13}$$

This means that inside a charge-free region, the potential is determined once its values on the boundary of the region are specified.

Note also that if the volume is all space, the surface integral at infinity vanishes if $\Phi(\mathbf{x}) \to 0$ at least as fast as $1/|\mathbf{x}-\mathbf{x}'|$ (since ${\rm d}a \sim R^2$). In that case we can take $F = 0$ and, comparing eq. (1.11) and (0.18), $G_{\rm D}(\mathbf{x},\mathbf{x}') = 1/|\mathbf{x}-\mathbf{x}'|$ because this already satisfies $G_{\rm D} = 0$ at infinity, and we end up with our earlier Coulomb expression: $\Phi(\mathbf{x}) = k_{\rm e}\int {\rm d}^3x'\rho(\mathbf{x}')/|\mathbf{x}-\mathbf{x}'|$.

Similar considerations apply with Neumann B.C., ie. when $\partial_n\Phi$ is known on the boundary. Then, if the volume is bounded by two surfaces, one closed and finite and the other at infinity (see Jackson p. 39), we have:

$$\Phi(\mathbf{x}) \;=\; k_{\rm e}\int_V \rho(\mathbf{x}')G_{\rm N}(\mathbf{x},\mathbf{x}')\,{\rm d}^3x' \;+\; \frac{1}{4\pi}\oint_S G_{\rm N}(\mathbf{x},\mathbf{x}')\partial_{n'}\Phi(\mathbf{x}')\,{\rm d}a' \tag{1.14}$$

So our first question has now been answered: supply the charge density inside a volume, together with Dirichlet or Neumann B.C. on the boundary of the volume; then Φ inside the volume can be found.

Incidentally, a useful property of Dirichlet Green functions for the Laplacian (or any other differential operator) follows directly by taking $f = G(\mathbf{x},\mathbf{x}'')$ and $g = G(\mathbf{x}',\mathbf{x}'')$ in Green's second identity, eq. (0.7):

$$\int_{\rm volume}\left[G(\mathbf{x},\mathbf{x}'')\,\nabla^2 G(\mathbf{x}',\mathbf{x}'') - G(\mathbf{x}',\mathbf{x}'')\,\nabla^2 G(\mathbf{x},\mathbf{x}'')\right]{\rm d}^3x'' \;=\; \oint_{\rm surface}\left[G(\mathbf{x},\mathbf{x}'')\,\partial_n G(\mathbf{x}',\mathbf{x}'') - G(\mathbf{x}',\mathbf{x}'')\,\partial_n G(\mathbf{x},\mathbf{x}'')\right]{\rm d}a''$$

The right-hand side vanishes because $G = 0$ for Dirichlet B.C. on the boundary surface. Since the volume is arbitrary, implementing the defining equation (1.11) for Green functions yields, :

$$G_{\rm D}(\mathbf{x}',\mathbf{x}) \;=\; G_{\rm D}(\mathbf{x},\mathbf{x}') \tag{1.15}$$

This symmetry of Dirichlet Green functions in their two arguments corresponds to the interchangeability of source point $\mathbf{x}'$ and observation point $\mathbf{x}$. It provides a good check on candidate Green functions for a Dirichlet problem.

1.5 Electrostatic Energy (section J1.11)

1.5.1 Potential (*interaction*) energy

If a point-charge q is sitting at position $\mathbf{x}$ where there exists an electrostatic potential $\Phi(\mathbf{x})$, work had to be done to bring it there from infinity. This work is $W = q\Phi(\mathbf{x})$, and is done by an *external* agent *against* the force associated with the potential acting on the charge q.

We know that the potential produced at a point $\mathbf{x}_j$ (where there is no charge) by a collection of $j-1$ discrete point-charges is:

$$\Phi(\mathbf{x}_j) \;=\; k_{\rm e} \sum_{i\neq j}^{j-1} \frac{q_{\rm i}}{|\mathbf{x}_i - \mathbf{x}_j|}$$

Then, to bring one extra charge, q_j, to $\mathbf{x}_j$ from infinity requires an energy $q_j\Phi(\mathbf{x}_j)$. By virtue of being located at $\mathbf{x}_j$ in the potential of the *other* $j-1$ charges, q_j now has a **potential energy** equal to the work done by the external agent to bring it there, namely $q_j\Phi(\mathbf{x}_j)$. So the total work needed to build a configuration of N charges, each at point $\mathbf{x}_{\rm j}$, is:

$$\begin{aligned} W \;&=\; \sum_{\rm j=1}^{N} \sum_{i<j}^{N} k_{\rm e}\, q_{\rm j}\, \frac{q_i}{|\mathbf{x}_j - \mathbf{x}_i|} \\ &=\; \frac{1}{2} \sum_{j=1}^{N} \sum_{\substack{i=1 \\ i\neq j}}^{N} k_{\rm e}\, \frac{q_{\rm j} q_{\rm i}}{|\mathbf{x}_i - \mathbf{x}_j|} \\ &=\; \frac{1}{2} \sum_{j=1}^{N} q_{\rm j}\, \Phi(\mathbf{x}_j) \end{aligned} \tag{1.16}$$

where $\Phi(\mathbf{x}_j)$ is the potential at $\mathbf{x}_j$ from all charges *other* than the charge sitting there.

By the same token, W is the potential energy of the distribution — in other words, the energy required to assemble it from charges initially located at infinity.

For a continuous volume distribution, this generalises to:

$$W \;=\; \frac{k_{\rm e}}{2} \iint\limits_{V} \frac{\rho(\mathbf{x})\,\rho(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\, \mathrm{d}^3x\, \mathrm{d}^3x' \;=\; \frac{1}{2} \int \rho(\mathbf{x})\, \Phi(\mathbf{x})\, \mathrm{d}^3x \tag{1.17}$$

except that there is a subtle but very important difference with the potential energy of a discrete distribution. In eq. (1.16), the potential at the points where the charges are sitting does not contribute, whereas in eq. (1.17) the integral receives contributions from the potential at *all* points in the volume. The latter, but not the former, includes the "self-energy" of the charge distribution.

1.5.2 Energy density in an electric field

In terms of the field, we have, in SI units:

$$\begin{aligned} W &= \frac{1}{2}\int \rho(\mathbf{x})\,\Phi(\mathbf{x})\,\mathrm{d}^3x \qquad (1.18)\\ &= \frac{\epsilon_0}{2}\int (\boldsymbol{\nabla}\cdot\mathbf{E})\,\Phi(\mathbf{x})\,\mathrm{d}^3x \\ &= \frac{\epsilon_0}{2}\left(\int \boldsymbol{\nabla}\cdot(\mathbf{E}\,\Phi)\,\mathrm{d}^3x \;-\; \int \mathbf{E}\cdot\boldsymbol{\nabla}\,\Phi\,\mathrm{d}^3x\right) \\ &= \frac{\epsilon_0}{2}\left(\oint\limits_{S\to\infty} (\mathbf{E}\,\Phi)\cdot\mathrm{d}\mathbf{a} \;+\; \int E^2\,\mathrm{d}^3x\right) \\ &= \int \frac{\epsilon_0}{2}\,E^2\,\mathrm{d}^3x \qquad (1.19) \end{aligned}$$

Note that even if if the initial volume integral in effect includes only the region where $\rho \neq 0$, starting on the third line, it must extend *over all space* so as to capture all the energy! In this form, W can be interpreted as the energy stored in the field, with $\epsilon_0 E^2/2$ the local **energy density**. Note that this is always positive, and that, therefore, the energy given by eq. (1.17) must also be positive, something which is not obvious there. As mentioned above, the apparent discrepancy with the energy of a discrete distribution, which can be negative, stems from the fact that the latter cannot include the self-energy of the charges. See also the example involving two point-charges on p. 42 in Jackson.

Because the electrostatic energy density is quadratic in the field, it does not obey the superposition principle; only the field does. To find the total energy density from more than one field contribution, you must *first* add the fields before using $\epsilon_0 E^2/2$.

Finally, we define the **capacitance** of a conductor as the amount of charge per unit potential that sits on it, when all other conductors are maintained at zero potential. If this latter condition is not met, the relationship between charge and potential, while still linear, is more complicated:

$$Q_i \;=\; \sum_{j=1}^{N} C_{ij} V_j$$

where V_j is the potential on the j$^{\text{th}}$ conductor, and the C_{ij} $(i \neq j)$ are called induction coefficients.

2 Boundary-Value Problem in Electrostatics

2.1 Method of Images and Associated Green Functions (sections J2.1, J2.6)

One very famous method for finding solutions of the Laplace equation replaces conducting surfaces (or equipotentials) by "virtual" charge distributions that, together with actual charges also present, generate a potential that satisfies the specified Dirichlet B.C.

This so-called **method of images** can be viewed as a clever use of the uniqueness theorem for electrostatic potentials. But there are no hard and fast rules for finding image charges and their location, except that they should always be placed outside the region where one wants the potential, otherwise they would change the problem. It is mostly a matter of guessing, so the method only works for simple enough distributions, such as a point-charge in the example below and in Jackson. If one can find the appropriate image charges and where they are located, then quantities other than the potential can also be calculated with little difficulty:

- the charge density induced on the conducting surface, via $\sigma = -\partial\Phi/\partial n/4\pi k_{\rm e}$;
- the force between the actual charge distribution and the conducting surface, which is simply the force between the real and image charges;
- the energy of the original and induced charge configuration, which can be taken to be the energy of the real plus image configuration *so long as the energy of the fields outside the region of interest is not included*!

The simplest image problems involve a point-charge q in the presence of various conducting objects. Obviously, the point-charge redistributes charge on the conductor so that its surface remains an equipotential. An elementary example is with a grounded conducting plate, where an image charge $-q$ placed behind the plane opposite q and at equal distance will produce the required B.C. on the plane. Here, the volume is bounded by the plane and an hemisphere at infinity *on the side of the real charge*.

Less trivial is Jackson's first worked-out example involving a grounded conducting sphere of radius a, the only one we will take up here using instead our Dirichlet solution (1.13) that, as pointed out (somewhat cryptically) in Jackson's section 2.6, has an interesting connection with the method of images. Indeed, we have already noted in section 1.4 above that the Dirichlet Green function for the Laplacian can be written:

$$G_{\rm D}(\mathbf{x},\mathbf{x}') = \frac{1}{|\mathbf{x}-\mathbf{x}'|} + F(\mathbf{x},\mathbf{x}')$$

where $\mathbf{x}$ ($r = |\mathbf{x}| \geq a$) denotes the position of the observation point, and F is any function that satisfies $\nabla^2 F(\mathbf{x},\mathbf{x}') = 0$. This means that within the volume between the sphere and infinity, G can be viewed as the electrostatic potential due to a point-charge $q = 1/k_{\rm e}$ lying at some point $\mathbf{x}'$ *outside* the sphere, plus another potential due to a source lying at some point $\mathbf{x}''$ *inside* the sphere, since that latter potential will indeed satisfy the Laplace equation *outside* the sphere. This is precisely how the method of images works for a point-charge q in the presence of an object with B.C. specified on its surface!

At $r = a$ we can write:

$$G_{\rm D}(\mathbf{x} = a\hat{\mathbf{n}},\mathbf{x}') = \frac{1}{a\sqrt{1+\dfrac{r'^2}{a^2}-2\dfrac{r'}{a}\hat{\mathbf{n}}\cdot\hat{\mathbf{n}}'}} + \frac{g}{r''\sqrt{1+\dfrac{a^2}{r''^2}-2\dfrac{a}{r''}\hat{\mathbf{n}}\cdot\hat{\mathbf{n}}''}}$$

where $\hat{\mathbf{n}} = \mathbf{x}/r$, etc., and g and $\mathbf{x}'' = r''\hat{\mathbf{n}}''$ are to be fixed by the "'universal" B.C.: $G_{\rm D}(\mathbf{x} = a\hat{\mathbf{n}},\mathbf{x}') = 0$. By inspection, we see that if this B.C. is to be satisfied for $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}}'$ in arbitrary directions, we must have $\hat{\mathbf{n}}\cdot\hat{\mathbf{n}}' = \hat{\mathbf{n}}\cdot\hat{\mathbf{n}}''$ ($\hat{\mathbf{n}}'$ and $\hat{\mathbf{n}}''$ collinear), $1/a = -g/r''$, and $r'/a = a/r''$, or:

$$g = -\frac{a}{r'}, \qquad r'\,r'' = a^2 \tag{2.1}$$

As expected, $r'' < a$ if $r' > a$. Our Dirichlet Green function has been found:

$$G_{\rm D}(\mathbf{x},\mathbf{x}') = \frac{1}{\sqrt{r^2+r'^2-2rr'\cos\gamma}} - \frac{1}{\sqrt{r^2r'^2/a^2+a^2-2rr'\cos\gamma}} \tag{2.2}$$

where $\cos\gamma = \hat{\mathbf{n}}\cdot\hat{\mathbf{n}}'$. $G_{\rm D}$ is symmetric in $\mathbf{x}$ and $\mathbf{x}'$, as it should be, and it vanishes at $r = a$, as required for a Dirichlet problem.

The surface integral in (1.13) does not contribute for a grounded conductor. Writing the volume integral as:

$$\int \rho(\mathbf{y}) G_{\rm D}(\mathbf{x},\mathbf{y})\,{\rm d}^3y$$

and inserting $\rho(\mathbf{y}) = q\,\delta(\mathbf{y}-\mathbf{x}')$ yields the potential exterior to the grounded conducting sphere:

$$\Phi(\mathbf{x}) \;=\; k_{\rm e}q\left[\frac{1}{\sqrt{r^2+r'^2-2rr'\cos\gamma}} - \frac{1}{\sqrt{r^2r'^2/a^2+a^2-2rr'\cos\gamma}}\right] \tag{2.3}$$

The surface charge density on the sphere is easily obtained from:

$$\sigma \;=\; -\frac{1}{4\pi k_{\rm e}}\frac{\partial\Phi}{\partial r}\Big|_{r=a} \;=\; -\frac{q}{4\pi a^2}\left[\frac{a(r'^2-a^2)}{(r'^2+a^2-2ar'\cos\gamma)^{3/2}}\right] \tag{2.4}$$

Green functions are much more powerful than the method of images! The Green function we have found is valid for a point-charge in the presence of *any* sphere and it does not care which particular B.C. for Φ is specified on the sphere. When $\Phi(r'=a)\neq 0$, all we have to do is add the surface integral in eq. (1.13); the volume integral remains the same since it is independent of the B.C. on Φ. Noting that $\hat{\mathbf{n}}'$ for the exterior volume points in the negative radial direction, the partial derivative of the Green function in the surface integral is simply:

$$\frac{\partial G_{\rm D}}{\partial n'} \;=\; -\frac{\partial G_{\rm D}}{\partial r'}\Big|_{r'=a} \;=\; -\frac{r^2-a^2}{a\,(r^2+a^2-2ar\cos\gamma)^{3/2}}$$

In our example, the potential in the presence of a point-charge outside the sphere, with $\Phi(a,\theta,\phi)$ specified on the sphere, is:

$$\begin{aligned}\Phi(\mathbf{x}) \;=\; & k_{\rm e}q\left[\frac{1}{\sqrt{r^2+r'^2-2rr'\cos\gamma}} - \frac{1}{\sqrt{r^2r'^2/a^2+a^2-2rr'\cos\gamma}}\right] \\ & + \frac{1}{4\pi}\int \Phi(r'=a)\,\frac{r^2-a^2}{a\,(r^2+a^2-2ar\cos\gamma)^{3/2}}\,{\rm d}a'\end{aligned} \tag{2.5}$$

If this is evaluated in spherical coordinates centered on the sphere, then $\Phi(r'=a) = \Phi(a,\theta',\phi')$, ${\rm d}a' = a^2{\rm d}\Omega'$, and $\cos\gamma = \cos\theta\,\cos\theta' + \sin\theta\,\sin\theta'\,\cos(\phi-\phi')$.

As a bonus, this easily generalises to arbitrary charge distributions outside the sphere, simply by evaluating the volume integral in eq. (1.13) with the above Green function unchanged. We can even remove any charge outside the sphere, and be left with:

$$\Phi(\mathbf{x}) \;=\; \frac{1}{4\pi}\int \Phi(a,\theta',\phi')\,\frac{a(r^2-a^2)}{(r^2+a^2-2ar\cos\gamma)^{3/2}}\,{\rm d}\Omega' \tag{2.6}$$

which is the solution to the Laplace equation, valid outside the sphere. To use this expression, we don't need to know the charge distribution inside the sphere; all we need is $\Phi(a,\theta',\phi')$ on its surface.

2.2 Boundary-Value Problems with Azimuthal Symmetry (section J3.3)

There is an important class of boundary-value problems in which the potential is independent of the azimuthal angle ϕ. They are said to exhibit **azimuthal symmetry** (not to be confused with cylindrical symmetry). When $m = 0$, the general spherical solution of the Laplace equation, eq. (0.46), has no dependence on the azimuthal angle, and it collapses to:

$$\Phi(r,\theta) \;=\; \sum_{l=0}^{\infty} \left[A_l\, r^l \;+\; B_l\, r^{-(l+1)}\right] P_l(x) \qquad\qquad (x = \cos\theta) \tag{2.7}$$

with the A_l and B_l determined by matching and boundary conditions.

2.2.1 A clever method to find potentials with azimuthal symmetry (pp. J102-103)

If the observation point lies on the symmetry axis ($\theta = 0$ or π), eq. (2.7) becomes:

$$\Phi(r,0) \;=\; \sum_{l=0}^{\infty} \left[A_l\, r^l \;+\; B_l\, r^{-(l+1)}\right]$$

The coefficients A_l and B_l in eq. (2.7) are *independent* of the location of the observation point. Therefore, if we can obtain the on-axis potential by some other means, all we have to do is expand it in a/r and r/a, where a is a characteristic length for the problem, and read off the coefficients from the previous expression. (see pp. 102-103 in Jackson for examples).

2.2.2 An important expansion for $1/|\mathbf{x} - \mathbf{x}'|$

We can also use the same reasoning to obtain a very important result. If we ask for the potential, at an observation point $\mathbf{x}$, of a point-charge $q = 1/k_{\rm e}$ situated at position $\mathbf{x}'$, we know that the answer is the Coulomb expression $\Phi = 1/|\mathbf{x} - \mathbf{x}'|$. This expression is coordinate-free, but nothing prevents us from working in a spherical coordinate basis and orienting the axes so that q lies on the $+z$-axis, thus creating a situation with azimuthal symmetry. Then, from eq. (2.7):

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} \;=\; \sum_{l=0}^{\infty} \left(A_l\, r^l + B_l\, r^{-(l+1)}\right) P_l(\cos\theta)$$

Note that the angle γ between $\mathbf{x}'$ and $\mathbf{x}$ is now the spherical coordinate θ (see figure J3.3).

First, let $r' < r < \infty$. Then $A_l = 0\ \forall\, l$ since $\Phi \to 0$ as $r \to \infty$, Then use our little trick and put the observation point also on the $+z$-axis. We have:

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} \;=\; \frac{1}{r - r'} \;=\; \frac{1}{r}\left(1 - \frac{r'}{r}\right)^{-1} \;\approx\; \frac{1}{r}\left[1 + \frac{r'}{r} + \left(\frac{r'}{r}\right)^2 + \ldots\right]$$

But

$$\sum_{l=0}^{\infty} B_l\, r^{-(l+1)} \;=\; \frac{1}{r}\left[B_0 + \frac{B_1}{r} + \frac{B_2}{r^2} + \ldots\right]$$

By inspection of the two expansions, $B_l = r'^l$, so that when $\mathbf{x}'$ is along the z-axis, with $r > r'$:

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} \;= \frac{1}{r}\sum_{l=0}^{\infty}\left(\frac{r'}{r}\right)^l P_l(\cos\,\theta) \qquad\qquad \forall\,\theta! \quad \theta' = 0$$

When $r' = 0$, $1/|\mathbf{x} - \mathbf{x}'| = 1/r$, as expected.

We can also use the same reasoning to show that, for $r < r'$, $A_l = 1/r'^{l+1}$, $B_l = 0$, and:

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} \;=\; \frac{1}{r'} \sum_{l=0}^{\infty} \left(\frac{r}{r'}\right)^l P_l(\cos\,\theta) \qquad r < r',\; \theta' = 0$$

Now what happens when $\mathbf{x}'$ lies off the z-axis, removing the azimuthal symmetry? Very little, actually: just copy the above expressions to coordinate-free form by noting that γ is *still* the angle between $\mathbf{x}'$ and $\mathbf{x}$, ie. $\mathbf{x}' \cdot \mathbf{x} = rr' \cos\gamma$. Thus, $\cos\theta \to \cos\gamma$.

Finally, the two expressions for $r > r'$ and $r < r'$ can be combined into one, which is the general expansion sought:

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} \;=\; \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} P_l(\cos\gamma) \tag{2.8}$$

where $r_<$ is the smaller of r and r', and $r_>$ is the larger of r and r'.

Apart from a choice of origin, this expression is coordinate-free, but often we wish to work in spherical coordinates. We must invoke the **Addition Theorem for spherical harmonics** proved in section J3.6, using $\cos\gamma = \cos\theta\,\cos\theta' + \sin\theta\,\sin\theta'\,\cos(\phi - \phi')$:

$$P_l(\cos\gamma) \;=\; \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi) \tag{2.9}$$

to write an expansion with completely factorised terms, useful when integrating over primed or unprimed coordinates:

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} \;=\; 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{1}{2l+1} \frac{r_<^l}{r_>^{l+1}} Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi) \tag{2.10}$$

Example 2.1. Application to a volume integral of the field (pp. J148-149)
The expansion of $1/|\mathbf{x} - \mathbf{x}'|$ in terms of Legendre polynomials can be put to work to calculate the integral of the electric field over a sphere of radius R:

$$\begin{aligned}
\int_{r<R} \mathbf{E}(\mathbf{x})\,\mathrm{d}^3x \;&=\; -\int_{r<R} \boldsymbol{\nabla}\Phi(\mathbf{x})\,\mathrm{d}^3x \\
&=\; -\oint_{r=R} R^2\,\Phi(\mathbf{x})\,\hat{\mathbf{n}}\,\mathrm{d}\Omega \\
&=\; -\,k_\mathrm{e}\,R^2 \int \mathrm{d}^3x'\,\rho(\mathbf{x}') \oint_{r=R} \frac{\hat{\mathbf{n}}}{|\mathbf{x} - \mathbf{x}'|}\,\mathrm{d}\Omega
\end{aligned} \tag{2.11}$$

where the second line is obtained from the second theorem on the right in Jackson's front cover. As usual $\hat{\mathbf{n}} = \mathbf{x}/R$ is a unit vector out of the sphere centered at the origin, which reads:

$$\begin{aligned}
\hat{\mathbf{n}} \;=\hat{\mathbf{r}} \;&=\; \mathbf{i}\,\sin\theta\,\cos\phi \;+\; \mathbf{j}\,\sin\theta\,\sin\phi \;+\; \mathbf{k}\,\cos\theta \\
&=\; \sqrt{\frac{2\pi}{3}} \Big(\mathbf{i}\,(Y_{1,-1}^* - Y_{11}^*) \;-\; \mathbf{j}\,\mathrm{i}\,(Y_{1,-1}^* \;+\; Y_{11}^*) \;+\mathbf{k}\,\sqrt{2}\,Y_{10}^*\Big)
\end{aligned} \tag{2.12}$$

The purpose of the last line was to show that $\hat{\mathbf{n}}$ can be expressed as a linear combination of Y_{1m}^*. Therefore, when we insert expansion (2.10) in the surface integral, orthogonality eliminates all terms

with $l \neq 1$, leaving the same combination of spherical harmonics as in eq. (2.12), but with (θ', ϕ') as arguments, that is: $\hat{\mathbf{n}}' = \mathbf{x}'/r'$. We are left with:

$$\begin{aligned}\int_{r<R} \mathbf{E}(\mathbf{x})\,\mathrm{d}^3x \;&=\; -4\pi k_\mathrm{e}\,\frac{R^2}{3}\int \frac{r_<}{r_>^2}\,\hat{\mathbf{n}}'\,\rho(\mathbf{x}')\,\mathrm{d}^3x' \\ &=\; -4\pi k_\mathrm{e}\,\frac{R^2}{3}\int_{r'<R}\frac{r'}{R^2}\,\hat{\mathbf{n}}'\,\rho(\mathbf{x}')\,\mathrm{d}^3x' \;-\; 4\pi k_\mathrm{e}\,\frac{R^2}{3}\int_{R<r'<\infty}\frac{R}{r'^2}\,\hat{\mathbf{n}}'\,\rho(\mathbf{x}')\,\mathrm{d}^3x'\end{aligned} \tag{2.13}$$

Two interesting possibilities can occur:

a) The charged source lies entirely outside the sphere. Then the first integral vanishes, and comparison with eq. (1.2) reveals that the second is just $-R\,\mathbf{E}(0)/4\pi k_\mathrm{e}$, yielding:

$$\int_{r<R} \mathbf{E}(\mathbf{x})\,\mathrm{d}^3x \;=\; \frac{4\pi R^3}{3}\,\mathbf{E}(0) \tag{2.14}$$

Therefore, the average of an electrostatic electric field over a sphere containing no charge is equal to the value of the field at its centre.

b) The sources lie entirely within the sphere. In this case, the second integral vanishes, and:

$$\int_{r<R} \mathbf{E}(\mathbf{x})\,\mathrm{d}^3x \;=\; -\frac{4\pi k_\mathrm{e}}{3}\int \mathbf{x}'\,\rho(\mathbf{x}')\,\mathrm{d}^3x' \tag{2.15}$$

where $\hat{\mathbf{n}}' = \mathbf{x}'/r'$ has been used.

We shall see in the next chapter what meaning can be ascribed to the integral on the right in this last equation.

2.3 Expansion of Green Functions in Spherical Coordinates (section J3.9)

Go back to our solution of the Dirichlet problem with Green functions:

$$\Phi(\mathbf{x}) \;=\; k_\mathrm{e}\int \rho(\mathbf{x}')G_\mathrm{D}(\mathbf{x},\mathbf{x}')d^3x' \;-\; \frac{1}{4\pi}\oint_S \Phi(\mathbf{x}')\,\partial_{n'}G_\mathrm{D}(\mathbf{x},\mathbf{x}')\,\mathrm{d}a'$$

We did obtain the Dirichlet Green function for a sphere, eq. (2.2), but it may not lead to tractable integrals. Fortunately, it is possible to find expansions for Dirichlet Green functions with factorised terms, that vanish on spherical or cylindrical surfaces. Here, we only consider the former.

In spherical coordinates, Green functions for the Laplacian operator all satisfy:

$$\begin{aligned}\nabla_x^2 G(\mathbf{x},\mathbf{x}') \;&=\; -4\pi\delta(\mathbf{x}-\mathbf{x}') \\ &=\; -\frac{4\pi}{r^2}\delta(r-r')\sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}^*(\theta',\phi')Y_{lm}(\theta,\phi)\end{aligned} \tag{2.16}$$

where eq. (0.44) has been invoked.

We shall look for a separable expansion of the form:

$$G(\mathbf{x},\mathbf{x}') \;=\; \sum_{l=0}^{\infty}\sum_{m=-l}^{l} g_l(r,r')\,Y_{lm}^*(\theta',\phi')Y_{lm}(\theta,\phi)$$

Inserting into eq. (2.16), we immediately find that $g_l(r,r')$ must satisfy the radial equation:

$$\frac{1}{r^2}\,\mathrm{d}_r\big(r^2\,g_l(r,r')\big)\;-\;l(l+1)\,g_l(r,r')\;=\;-4\pi\,\delta(r-r')$$

Consider the case $r\neq r'$. Then $g_l(r,r')$ satisfies the Laplace radial equation whose solutions are given by (0.46):

$$g_l(r,r')\;=\;\begin{cases}A(r')\,r^l\;+\;B(r')\,r^{-l-1} & r>r'\\ A'(r')\,r^l\;+\;B'(r')\,r^{-l-1} & r<r'\end{cases}$$

Now, at $r=r'$, $\mathrm{d}_r g_l$ has a discontinuity equal to $-4\pi/r'^2$. To see this, integrate the radial differential equation over r from $r'-\epsilon$ to $r'+\epsilon$. Continuity of g_l at $r=r'$ eliminates the integral of the last term on the left, yielding the result quoted.

Implementing the continuity of g_l and the discontinuity in $\mathrm{d}_r g_l$ at $r=r'$ allows us (EXERCISE) to eliminate A and B:

$$A\;=\;A'\;+\;\frac{4\pi}{2l+1}\,\frac{1}{r'^{l+1}},\qquad\qquad B\;=\;B'\;-\;\frac{4\pi}{2l+1}\,r'^{l}$$

We determine A' and B' for the most important case of a volume enclosed within two concentric spheres of radius a and b, ($a<b$). Since we want to solve the Dirichlet problem, $G(\mathbf{x},\mathbf{x}')$ must vanish on the spheres, so that $g_l(r,r')=0$ for both $r=a$ and $r=b$. This can be used to find $B'(r')$ and $A'(r')$. Enforcing $g_l(a,r')=0$ in the region $r<r'$ yields $B'=-a^{2l+1}A'$.

In the $r>r'$ region, the radial Green function is then:

$$g_l(r,r')\;=\;A'(r')\,r^l\,\left(1\;-\;\frac{a^{2l+1}}{b^{2l+1}}\right)\;-\;\frac{4\pi}{2l+1}\,\left(\frac{r^l}{r'^{l+1}}\;-\;\frac{r'^l}{r^{l+1}}\right)$$

Implementing $g_l(b,r')=0$ leads to (EXERCISE):

$$A`(r')\;=\;\frac{4\pi}{(2l+1)\,[1-(a/b)^{2l+1}]}\,\left(\frac{1}{r'^{l+1}}\;-\;\frac{r'^l}{b^{2l+1}}\right)$$

Then switching r and r' and noting that $g_l(r,r')=g_l(r',r)$, there comes:

$$g_l(r,r')\;=\;\frac{4\pi}{(2l+1)\,[1-(a/b)^{2l+1}]}\,\left(r_<^l\;-\;\frac{a^{2l+1}}{r_<^{l+1}}\right)\left(\frac{1}{r_>^{l+1}}\;-\;\frac{r_>^l}{b^{2l+1}}\right)$$

The complete expansion for Dirichlet Green functions in spherical coordinates becomes:

$$G_{\mathrm{D}}(\mathbf{x},\mathbf{x}')\;=\;4\pi\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{Y_{lm}^*(\theta',\phi')Y_{lm}(\theta,\phi)}{(2l+1)\,[1-(a/b)^{2l+1}]}\,\left(r_<^l\;-\;\frac{a^{2l+1}}{r_<^{l+1}}\right)\left(\frac{1}{r_>^{l+1}}\;-\;\frac{r_>^l}{b^{2l+1}}\right)\tag{2.17}$$

Two important cases:

$$G_{\mathrm{D}}(\mathbf{x},\mathbf{x}')\;=\;4\pi\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{Y_{lm}^*(\theta',\phi')Y_{lm}(\theta,\phi)}{(2l+1)}\,r_<^l\,\left(\frac{1}{r_>^{l+1}}\;-\;\frac{r_>^l}{b^{2l+1}}\right)\qquad(a=0)\tag{2.18}$$

$$G_{\mathrm{D}}(\mathbf{x},\mathbf{x}')\;=\;4\pi\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{Y_{lm}^*(\theta',\phi')Y_{lm}(\theta,\phi)}{(2l+1)}\,\frac{1}{r_>^{l+1}}\,\left(r_<^l\;-\;\frac{a^{2l+1}}{r_<^{l+1}}\right)\qquad(b\to\infty)\tag{2.19}$$

The first expression is useful when finding the potential inside a sphere; the second one, outside a sphere. As expected, when $a=0$ and $b\to\infty$ we recover expansion (2.10).

When $0 \leq r \leq b$ (interior case) we can immediately rewrite (EXERCISE) the term in the surface integral in eq. (1.13) as:

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left[\int \Phi(b, \theta', \phi')\, Y^*_{lm}(\theta', \phi')\, \mathrm{d}\Omega' \right] \left(\frac{r}{b}\right)^l Y_{lm}(\theta, \phi)$$

where $\Phi(b, \theta', \phi')$ is the potential specified on the surface $r' = b$. The normal derivative of the Green function on the surface, $\partial_{n'} G = \partial_{r'} G \Big|_{r'=b}$, has been evaluated for $r_< = r$ and $r_> = r'$ since $r < r' = b$. Also, the surface element on a sphere of radius b is $\mathrm{d}a' = b^2 \mathrm{d}\Omega'$.

This, however, is no other than the interior solution of the Laplace equation we obtained from eq. (0.46) by putting $B_{lm} = 0$: a nice check on our calculations! And we have found another way of calculating the A_{lm} coeffcients.

Now, for simplicity, we concentrate on a Dirichlet problem in which there is no surface term in eq. (1.13) and only the volume integral contributes. We must express the charge density $\rho(\mathbf{x}')$ in spherical coordinates. Very often, δ-functions can be used to localise the source.

Example 2.2. Charged line inside a grounded sphere (J3.10)

Consider a uniform line of total charge Q on the z axis inside and along the whole diameter of a hollow grounded sphere of radius b. The appropriate expression for the *volume* charge density is:

$$\rho(\mathbf{x}') \;=\; \frac{\lambda}{2\pi r'^2} \left[\delta(\cos\theta' - 1) \;+\; \delta(\cos\theta' + 1)\right]$$

where the *linear* charge density is $\lambda = Q/2b$. The factor in front of the δ-functions is chosen so that the volume integral of the charge density over all space gives Q. Thanks to the azimuthal symmetry of the problem, only the $m = 0$ terms in the first equation of eq. (2.18) contribute. The spherical harmonics Y_{l0} are just Legendre polynomials multiplied by $\sqrt{(2l+1)/4\pi}$. After integrating over all solid angles, we find for the potential inside the sphere:

$$\begin{aligned} \Phi(r, \theta) \;&=\; k_\mathrm{e} \int \rho(\mathbf{x}') G_\mathrm{D}(\mathbf{x}, \mathbf{x}')\, r'^2 \,\sin\theta'\, \mathrm{d}r'\, \mathrm{d}\theta'\, \mathrm{d}\phi' \\ &=\; \frac{k_\mathrm{e} Q}{2b} \sum_{l=0}^{\infty} \left[1 + (-1)^l\right] P_l(\cos\theta) \int_0^b r_<^l \left(\frac{1}{r_>^{l+1}} \;-\; \frac{r_>^l}{b^{2l+1}} \right) \mathrm{d}r' \end{aligned} \tag{2.20}$$

where $P_l(1) = 1$ and $P_l(-1) = (-1)^l$ have been used. The radial integral requires some care. First, we evaluate it for $l = 0$, breaking up the interval into a part where $r' < r$ and a part where $r' > r$. The integral is $\ln(b/r)$.

Then the radial integral is evaluated for $l \neq 0$, as in eq. (J3.134) in Jackson. We note that the result is indeterminate when $l = 0$, which is why we had to evaluate the $l = 0$ term on its own. The result is eq. J3.136:

$$\Phi(\mathbf{x}) \;=\; \frac{k_\mathrm{e} Q}{b} \left[\ln b/r \;+\; \sum_{n=1}^{\infty} \frac{4n+1}{2n\,(2n+1)} \left[1 \;-\; (r/b)^{2n}\right] P_{2n}(\cos\theta) \right]$$

Another interesting example involving a charged ring inside a grounded sphere can be found in Jackson.

3 Multipole Expansion and Ponderable Media

3.1 The Multipole Expansion (section J4.1)

3.1.1 Multipole moments

In the absence of boundary conditions other than the one at infinity, the potential outside a sphere that encloses a localised charge distribution $\rho(\mathbf{x})$ is often most usefully expressed using the expansion in spherical coordinates obtained in eq. (2.10), with $r_< = r'$ and $r_> = r$:

$$\begin{aligned}\Phi(\mathbf{x}) &= k_\mathrm{e} \int \frac{\rho(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}^3x' \\ &= 4\pi k_\mathrm{e} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{1}{2l+1} \frac{q_{lm}}{r^{l+1}} Y_{lm}(\theta,\phi)\end{aligned} \tag{3.1}$$

where

$$q_{lm} = \int Y^*_{lm}(\theta',\phi')\, r'^l\, \rho(\mathbf{x}')\, \mathrm{d}^3x' \tag{3.2}$$

are constant coefficients called **spherical multipole moments**. They inherit a property of the spherical harmonics:

$$q_{l,-m} = (-1)^m q^*_{lm} \tag{3.3}$$

The second line of eq. (3.1) is the famous **multipole expansion**. Quite often, the first few non-vanishing terms provide a description of the potential accurate enough for most purposes.

If, on the other hand, we go back to the coordinate-free expansion (2.8) of $1/|\mathbf{x}-\mathbf{x}'|$ and put $\cos\gamma = \mathbf{x}\cdot\mathbf{x}'/rr'$, the potential immediately becomes:

$$\Phi(\mathbf{x}) = \frac{k_\mathrm{e}}{r}\int \rho(\mathbf{x}')\,\mathrm{d}^3x' + \frac{k_\mathrm{e}}{r^3}\int \mathbf{x}\cdot\mathbf{x}'\,\rho(\mathbf{x}')\,\mathrm{d}^3x' + \frac{k_\mathrm{e}}{r^5}\frac{1}{2}\int \left[3(\mathbf{x}'\cdot\mathbf{x})^2 - (\mathbf{x}\cdot\mathbf{x})(\mathbf{x}'\cdot\mathbf{x}')\right]\rho(\mathbf{x}')\,\mathrm{d}^3x' + \ldots$$

This expression is the multipole expansion in coordinate-free notation. Now define the **electric dipole moment**:

$$\mathbf{p} = \int \mathbf{x}'\rho(\mathbf{x}')\,\mathrm{d}^3x' \tag{3.4}$$

and the *symmetric*, *traceless* **electric quadrupole moment** tensor, both in Cartesian coordinates:

$$Q_{ij} = \int (3x'_i x'_j - \delta_{ij} r'^2)\rho(\mathbf{x}')\,\mathrm{d}^3x' \tag{3.5}$$

If (*and only if!*) we work in Cartesian coordinates, we can then rewrite the coordinate-free expansion as:

$$\Phi(\mathbf{x}) = \frac{k_\mathrm{e} q}{r} + k_\mathrm{e}\frac{\mathbf{x}\cdot\mathbf{p}}{r^3} + \frac{k_\mathrm{e}}{2}\frac{x^i x^j}{r^5} Q_{ij} + \ldots \tag{3.6}$$

where q is the total charge of the distribution and a sum over repeated indices is implied. This last expression gives all its meaning to the term "multipole expansion". The potential can be written as the sum of the potentials of a monopole, of a dipole, of a quadrupole, etc. But we must remember that it is valid in Cartesian coordinates only, although we are free to evaluate any *one rectangular component* of the dipole vector or quadrupole tensor in whatever coordinates we wish.

While the spherical moments q_{lm} are written above in spherical coordinates, writing them in a Cartesian basis is also very instructive. We use:

$$\mathrm{e}^{\pm\mathrm{i}\phi'} = \frac{1}{r'\sin\theta'}(x' \mp \mathrm{i}y'), \qquad \cos\theta' = \frac{z'}{r'}$$

to write:

$$q_{00} = \frac{1}{\sqrt{4\pi}}\,q \qquad q_{10} = \sqrt{\frac{3}{4\pi}}\,p_z \qquad q_{1,\pm1} = \mp\sqrt{\frac{3}{8\pi}}\,(p_x \mp \mathrm{i}\,p_y)$$
$$q_{20} = \sqrt{\frac{5}{16\pi}}\,Q_{33} \qquad q_{2,\pm1} = -\sqrt{\frac{5}{24\pi}}\,(Q_{13} \mp \mathrm{i}\,Q_{23}) \qquad q_{2,\pm2} = \sqrt{\frac{5}{96\pi}}\,(Q_{11} \mp 2\mathrm{i}\,Q_{12} - Q_{22}) \tag{3.7}$$

where q is the total charge in the distribution.

A nice feature of the multipole expansion is that it breaks the source into physically intuitive pieces (monopole, dipole, quadrupole, and so on). Because Maxwell's equations for electromagnetism are *linear* in the fields, the potential is the sum of the potentials of each of these pieces. The same holds for *weak* gravitational fields: a mass distribution can be resolved into monopole, quadrupole, ... moments (why is there no dipole moment?), each of which has its own characteristic field. But this is no longer true for *strong* gravitational fields which obey the *nonlinear* equations of Einstein's theory of gravitation.

It is important to remember that r in eq. (3.6) is the distance from an origin inside the distribution to the observation point $\mathbf{x}$ as if the monopole, dipole, quadrupole, etc., were located at that origin. If we wish to put the origin somewhere else, say $\mathbf{x}_0$, without changing the moments, we must replace $\mathbf{x}$ by $\mathbf{x}-\mathbf{x}_0$ and r by $|\mathbf{x}-\mathbf{x}_0|$.

A spherical-multipole expansion of the electric field is easily obtained by differentiating each lm term in (3.1):

$$\mathbf{E}_{lm} = \frac{4\pi k_\mathrm{e}}{2l+1}\,\frac{q_{lm}}{r^{l+2}}\,Y_{lm}(\theta,\phi)\left(\hat{\mathbf{r}} + \hat{\boldsymbol{\theta}}\,\partial_\theta \ln Y_{lm}(\theta,\phi) + \hat{\boldsymbol{\phi}}\,\frac{\mathrm{i}\,m}{\sin\theta}\right) \tag{3.8}$$

3.1.2 Calculating the multipole moments

Symmetric charge distributions can have very simple moments. Consider for instance $\rho(r,\theta)$, which has azimuthal symmetry. Then in eq. (3.2) $m=0$, and inspection of eq. (3.7) immediately yields $p_x = p_y = 0$, and:

$$p_z = 2\pi \int\int r'^3 \cos\theta' \sin\theta' \rho(r',\theta')\,\mathrm{d}r'\,\mathrm{d}\theta' \tag{3.9}$$

Also, $Q_{12} = Q_{13} = Q_{23} = 0$, and $Q_{11} = Q_{22}$. Q_{ij} being traceless, we must have $Q_{11} = Q_{22} = -Q_{33}/2 \equiv -eQ/2$, where:

$$Q = \frac{2\pi}{e}\int\int (3z'^2 - r'^2)\,\rho(r',\theta')\,r'^2 \sin\theta'\,\mathrm{d}r'\,\mathrm{d}\theta' \tag{3.10}$$

is the electric **quadrupole moment**, with units of length squared. Finally, in eq. (3.6), $x^i x^j Q_{ij} = [-(x^2+y^2)/2 + z^2]\,eQ$ (EXERCISE), and if the observation point lies on the z-axis, only $2z^2Q$ survives.

In general the multipole moment integrals depend upon the choice of origin. But the first non-zero moment is *independent of the choice of origin*. This means that when you calculate that moment, and only that one, you are at liberty to translate (but not rotate!) the axes if this results in an easier calculation. It is straightforward to show that this applies to the quadrupole term when $\mathbf{p}=0$; just replace x'^i by $x'^i + a^i$ in the Cartesian expression for Q_{ij}, and expand.

In problems with a point-charge distribution, you can always write $\rho(\mathbf{x}) = \sum q_i \delta(\mathbf{x}-\mathbf{x}'_i)$, where $\mathbf{x}'_i$ is the position vector of charge q_i. The moment integrals must extend over all space, and in this case it is easiest to work them out in Cartesian coordinates.

You can split any charge distribution into sub-distributions whose potentials add. In particular, a point-charge distribution can often be split into sums of cancelling dipoles (by vector addition) plus something else which may turn out to be much simpler than the original distribution.

The components, p_i, of the dipole moment must go like charge times distance; the components, Q_{ij}, of the quadrupole moment each must go like charge times distance squared. This serves as one check on your answers.

3.1.3 Field of the electric dipole moment

The electric field associated with the dipole potential $k_{\rm e}\mathbf{p}\cdot\mathbf{x}/r^3$ is also important, and we can derive it in coordinate-free form:

$$\begin{aligned}\mathbf{E}_{\rm dip} &= -\boldsymbol{\nabla}\Phi_{\rm dip}\\ &= -k_{\rm e}\left[\mathbf{p}\cdot\boldsymbol{\nabla}\right]\left(\frac{\mathbf{x}}{r^3}\right)\\ &= \frac{k_{\rm e}}{r^3}\left[3\hat{\mathbf{n}}\left(\hat{\mathbf{n}}\cdot\mathbf{p}\right) - \mathbf{p}\right]\end{aligned} \tag{3.11}$$

where the third-from-last identity on the left of Jackson's front cover has been invoked. If the dipole is oriented along the z-axis, the familiar expressions for a dipole field are recovered.

Actually, this is not quite the whole story. Back in chapter 2, we derived eq. (2.15), an expression for the integral of *any* electric field over a sphere when this sphere contains charge and there is no charge outside. We now recognise the integral on the right-hand side of that equation as the dipole moment of the charge distribution:

$$\int_{r<R}\mathbf{E}(\mathbf{x})\,\mathrm{d}^3x \;=\; -\frac{4\pi}{3}\,k_{\rm e}\,\mathbf{p} \tag{3.12}$$

The problem is that if, for example, we integrate using eq. (2.12) the field of a dipole as given by eq. (3.11), the angular integrations make all three components of the integral vanish (EXERCISE). So we arrive at $\mathbf{p}=0$, a contradiction.

Well, the differentiation that leads to eq. (3.11) yields unambiguous results everywhere except at the position of the dipole, here the origin. This means that to be consistent with eq. (3.12), we must add a term to our field which vanishes everywhere except at the origin. Generalising to the field of a dipole located at some position $\mathbf{x}_0$, and with $\hat{\mathbf{n}}$ pointing from $\mathbf{x}_0$ to $\mathbf{x}$, we write:

$$\mathbf{E}_{\rm dip}(\mathbf{x}) \;=\; k_{\rm e}\left[\frac{3\hat{\mathbf{n}}\left(\hat{\mathbf{n}}\cdot\mathbf{p}\right) - \mathbf{p}}{|\mathbf{x}-\mathbf{x}_0|^3} - \frac{4\pi}{3}\,\mathbf{p}\,\delta(\mathbf{x}-\mathbf{x}_0)\right] \tag{3.13}$$

3.2 Energy of a Charge Distribution in an External Potential (section J4.2)

The *interaction* energy of a charge distribution with an *external* potential is:

$$W \;=\; \int\rho(\mathbf{x})\,\Phi_{\rm ext}(\mathbf{x})\,\mathrm{d}^3x$$

Unlike in eq. (1.17), there is no factor 1/2 since $\Phi_{\rm ext}$ is not the potential of the distribution!

Assuming that the potential varies slowly over the distribution, perform a Taylor expansion of the potential around some chosen origin $\mathbf{x}=0$, with $\mathbf{E}(\mathbf{x})=-\boldsymbol{\nabla}\Phi(\mathbf{x})$:

$$\begin{aligned}\Phi_{\rm ext}(\mathbf{x}) &= \Phi_{\rm ext}(0) + \mathbf{x}\cdot\boldsymbol{\nabla}\Phi_{\rm ext}\Big|_0 + \frac{1}{2}x^ix^j\,\partial_i\partial_j\Phi_{\rm ext}\Big|_0 + \ldots\\ &= \Phi_{\rm ext}(0) - \mathbf{x}\cdot\mathbf{E}_{\rm ext}(0) - \frac{1}{6}\left(3\,x^ix_j - r^2\delta^i{}_j\right)\partial_iE^j_{\rm ext}\Big|_0 + \ldots\end{aligned}$$

where the last term, $r^2\,\delta^i{}_j\partial_iE^j_{\rm ext}(0) = r^2\,\partial_jE^j_{\rm ext}(0) = r^2\boldsymbol{\nabla}\cdot\mathbf{E}_{\rm ext}\big|_0 = 0$, is introduced so that when the expansion is inserted in the energy, we can write:

$$W \;=\; q\,\Phi_{\rm ext}(0) - \mathbf{p}\cdot\mathbf{E}_{\rm ext}(0) - \frac{1}{6}Q_{ij}\,\partial_iE^j_{\rm ext}\Big|_0 + \ldots \tag{3.14}$$

This tells us that so far as interaction energy is concerned, the total charge of the distribution couples to the external potential, its dipole moment to the external field, and its quadrupole tensor to the gradient of the external field.

We can use our energy expansion to write the interaction energy of two dipoles at $\mathbf{x}_1$ and $\mathbf{x}_2$. From eq. (3.13), this is:

$$W_{12} \;=\; -\,\mathbf{p}_1 \cdot \mathbf{E}_2(\mathbf{x}_1) \;=\; k_\mathrm{e}\,\frac{\mathbf{p}_1\cdot\mathbf{p}_2 \;-\; 3(\hat{\mathbf{n}}\cdot\mathbf{p}_1)(\hat{\mathbf{n}}\cdot\mathbf{p}_2)}{|\mathbf{x}_1-\mathbf{x}_2|^3} \tag{3.15}$$

where $\mathbf{x}_1 \neq \mathbf{x}_2$ and $\hat{\mathbf{n}}$ points from $\mathbf{x}_2$ to $\mathbf{x}_1$.

3.3 Electrostatics in Ponderable Media: Polarisation (section J4.3)

Consider a small volume ΔV of a medium that still contains a very large number of molecules of various type. ΔV may also contain a **free** charge density $\rho_\mathrm{free}(\mathbf{x})$. Normally, over this volume the *bound* charges in each type of molecule average out to zero; so do their electric dipole moments. Assuming that ΔV is located at a variable macroscopically small, smeared "point" $\mathbf{x}'$ and is small enough that any free ρ is uniform over it, we know that the potential at some point $\mathbf{x}$ outside ΔV is:

$$\Delta\Phi(\mathbf{x},\mathbf{x}') \;=\; k_\mathrm{e}\,\frac{\rho_\mathrm{free}(\mathbf{x}')\,\Delta V}{|\mathbf{x}-\mathbf{x}'|}$$

When an external field is applied to the molecules in ΔV, their average charge remains zero, but now the average[†] total dipole moment per unit volume, or **electric polarisation**, $\mathbf{P} = N\langle\mathbf{p}\rangle$, no longer vanishes for a given type of molecule with number density N. The field tries to orient the dipoles in its direction. Even when the molecules have no intrinsic dipole moment, they acquire one under the influence of the field.

Now there is a new dipole contribution to the potential. It is given by the dipole term in the multipole expansion, eq. (3.6), with $\mathbf{x}'$ taken as the location of the multipole sources. That is:

$$\begin{aligned}
\Delta\Phi(\mathbf{x},\mathbf{x}') \;&=\; k_\mathrm{e}\left[\frac{\rho_\mathrm{free}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} \;+\; \frac{\mathbf{P}(\mathbf{x}')\cdot(\mathbf{x}-\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|^3}\right]\Delta V \\
&=\; k_\mathrm{e}\left[\frac{\rho_\mathrm{free}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} \;+\; \mathbf{P}(\mathbf{x}')\cdot\boldsymbol{\nabla}'\left(\frac{1}{|\mathbf{x}-\mathbf{x}'|}\right)\right]\Delta V \\
&=\; k_\mathrm{e}\left[\frac{\rho_\mathrm{free}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} \;-\; \frac{\boldsymbol{\nabla}'\cdot\mathbf{P}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} \;+\; \boldsymbol{\nabla}'\cdot\left(\frac{\mathbf{P}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\right)\right]\Delta V
\end{aligned}$$

At this point, we let $\Delta V \to \mathrm{d}^3x'$ and integrate over *all space*. By the divergence theorem the integral of the divergence vanishes because $\mathbf{P}/|\mathbf{x}-\mathbf{x}'|$ vanishes at infinity. There comes:

$$\Phi(\mathbf{x}) \;=\; k_\mathrm{e}\int_{\text{all space}} \frac{\left[\rho_\mathrm{free}(\mathbf{x}') \;-\; \boldsymbol{\nabla}'\cdot\mathbf{P}(\mathbf{x}')\right]}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x' \tag{3.16}$$

This treatment raises two questions:

1. We started with an equation for the potential outside a charge distribution, and we have come up with an expression that we claim is the potential everywhere, not only outside, but also *inside* the medium. We justify this by mentally removing any charge sitting at $\mathbf{x}$; doing this should not significantly affect the potential at that point, produced by all other charges located elsewhere.

2. More seriously, we have used the multipole expansion and neglected all multipole moments higher than the dipole. How valid can this be when the microscopic fields inside the medium are probably extremely complicated, and we are certainly not far from the sources? The answer lies with eq. (3.12), which, as will be seen in the next section, proves that the *average* field from all the molecules inside an arbitrary sphere is entirely accounted for by the dipole moment of the molecules!

[†]For a careful discussion of this averaging process, see section 6.6 in Jackson.

We see from eq. (3.16) that the net potential is generated by an effective *volume* charge density $\rho_{\text{eff}} = \rho_{\text{free}} + \rho_{\text{pol}} = \rho_{\text{free}} - \boldsymbol{\nabla}\cdot\mathbf{P}$, which means that Gauss' law in differential form, valid for macroscopic media, can now be written as:

$$\boldsymbol{\nabla}\cdot\mathbf{D} \;=\; \rho_{\text{free}} \tag{3.17}$$

where the **electric displacement** $\mathbf{D} \equiv \epsilon_0\mathbf{E}_{\text{net}} + \mathbf{P}$ in SI units. We call $\rho_{\text{pol}} = -\boldsymbol{\nabla}\cdot\mathbf{P}$ the **polarisation charge density** (some call it the **bound** charge density).

Does this equation determine $\mathbf{D}$, in the same way that Gauss' law determines $\mathbf{E}$ because $\boldsymbol{\nabla}\times\mathbf{E} = 0$? In general, no: $\boldsymbol{\nabla}\times\mathbf{P} = 0$ only in certain situations, such as when polarisation is uniform in the medium, or when there is spherical symmetry and the fields are all central.

An immediate consequence of eq. (3.17) is that the boundary condition between two media, eq. (1.9), must be replaced by:

$$(\mathbf{D}_2 \;-\; \mathbf{D}_1)\cdot\hat{\mathbf{n}} = \;\sigma_{\text{free}} \tag{3.18}$$

which is immediately put in the much more useful form in terms of *net* fields:

$$(\mathbf{E}_2 \;-\; \mathbf{E}_1)\cdot\hat{\mathbf{n}} = \;\frac{\sigma_{\text{free}} \;+\; \sigma_{\text{pol}}}{\epsilon_0} \tag{3.19}$$

where $\sigma_{\text{pol}} = -\left(\mathbf{P}_2 \;-\; \mathbf{P}_1\right)\cdot\hat{\mathbf{n}}$, and $\hat{\mathbf{n}}$ is taken to point from medium 1 to medium 2. In terms of the potential, this becomes:

$$\partial_n\Phi_2 \;-\; \partial_n\Phi_1 \;=\; -\,\frac{\sigma_{\text{free}} \;+\; \sigma_{\text{pol}}}{\epsilon_0} \tag{3.20}$$

At the surface of a dielectric placed in air,or in vacuum, the polarisation $\mathbf{P}$ is produced by a surface polarisation (or bound) charge density $\sigma_{\text{pol}} = \mathbf{P}\cdot\hat{\mathbf{n}}$, where in this case $\hat{\mathbf{n}}$ points *out* of the medium.

The tangential components of $\mathbf{E}$ still match at the surface, as does the potential.

There is a subtle yet important point that must be made here. This surface charge density that appears in eq. (3.19) acts as a source for the net field, and therefore the potential calculated from eq. (3.16), yet it does not appear in that equation! To resolve this potential contradiction, we must examine more closely the second term in the integral for the potential. This assumes that $\mathbf{P}$ is a differentiable function of position known everywhere in space. In other words, there can be no *sharp* boundaries where the polarisation vector is discontinuous. The problem is that in the vicinity of the surface of a dielectric (or near an interface between two dielectrics), $\mathbf{P}$ can change very rapidly, for instance decreasing from some finite value down to zero, in a continuous but unknown fashion, so that we cannot in practice calculate the *volume* charge density ρ_{pol} inside that thin layer, and we cannot calculate its contribution to the potential via eq. (3.16).

What we *can* do is to replace that thin layer by a sharp boundary surface where $\mathbf{P}$ is modelled as discontinuous. Then our expression (3.16) cannot include the contribution of the boundary, if only because ρ_{pol} diverges there. To compensate we can integrate over the volume of the dielectric instead of over all space. Then the surface integral that was zero at infinity does not vanish anymore, and our potential becomes:

$$\Phi(\mathbf{x}) \;=\; k_{\text{e}}\int_V \frac{\left[\rho(\mathbf{x}') \;-\; \boldsymbol{\nabla}'\cdot\mathbf{P}(\mathbf{x}')\right]}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x' \;+\; k_{\text{e}}\oint_S \frac{\sigma_{\text{pol}}}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}a' \tag{3.21}$$

where S is the surface of the dielectric, and $\sigma_{\text{pol}} = \mathbf{P}\cdot\mathbf{n}$, as was found before. This is the more useful expression for the potential of a dielectric.

To solve such an electrostatic problem, we need to know the polarisation $\mathbf{P}$. Fortunately, unless applied external fields are very large, the response of the vast majority of media is **linear**, in the sense that the magnitude of $\mathbf{P}$ is linear in the applied field. Most media are also **isotropic**, so that the magnitude of the induced polarisation does not depend on the direction of the applied field and is in the same direction as that field. It is customary to write this so-called **constitutive relation** in terms of the *net* field $\mathbf{E}$, which is the sum of the applied and induced fields, as:

$$\mathbf{P} \;=\; \epsilon_0\chi_{\text{e}}\mathbf{E}_{\text{net}} \tag{3.22}$$

where the dimensionless quantity χ_e bears the suggestive name of **electric susceptibility**. Then the displacement and the net field are related by:

$$\mathbf{D} \;=\; \epsilon_0 \epsilon_r \, \mathbf{E} \;=\; \epsilon \, \mathbf{E} \tag{3.23}$$

where $\epsilon_r = 1 + \chi_e \geq 1$ is called the **dielectric constant** of the medium. (From now on in this discussion, we write $\mathbf{E} = \mathbf{E}_{\text{net}}$.)

At this stage, we have the differential Gauss law $\boldsymbol{\nabla} \cdot (\epsilon \mathbf{E}) = \rho_{\text{free}}$, with ϵ a constant only for homogeneous (or uniform) media. In this very frequent case, $\boldsymbol{\nabla} \cdot \mathbf{E} = \rho_{\text{free}}/\epsilon$, and everything that happens inside the medium can be described by the same expressions we have derived in chapters 1–3 by letting $\epsilon_0 \to \epsilon$. We also see that since $\epsilon > \epsilon_0$, $\mathbf{E}$ is decreased compared to what it would be without the dielectric.

For linear and homogeneous media the polarisation is given by:

$$\mathbf{P} \;=\; (\epsilon_r - 1)\, \epsilon_0 \, \mathbf{E} \tag{3.24}$$

and the matching equation (3.20) then becomes, using $\sigma_{\text{pol}} = -\,(\mathbf{P}_2 \;-\; \mathbf{P}_1) \cdot \hat{\mathbf{n}}$:

$$\epsilon_{r2}\, \partial_n \Phi_2 \;-\; \epsilon_{r1}\, \partial_n \Phi_1 \;=\; -\,\frac{\sigma_{\text{free}}}{\epsilon_0} \tag{3.25}$$

In the absence of free charge, the problem for linear and homogeneous media always reduces to solving the Laplace equation with appropriate B.C., something we know how to do from chapter 2 in these notes..

3.4 Relating Susceptibility to Molecular Polarisability; Models for the Latter (sections J4.5, 4.6)

The electric susceptibility, χ_e, is a *macroscopic* parameter; somehow, it must be related to the microscopic properties of the medium at the molecular level. We can find this relation in the case of non-polar molecules, but first we must discover what field acts on such a molecule.

Consider a sphere of radius R containing lots of molecules, small enough that the macroscopic field $\mathbf{E}$ is roughly uniform over it. The actual field acting on a molecule at the centre of the sphere, however, which we will call $\mathbf{E}_{\text{local}}$, is not really $\mathbf{E}$, especially in dense media! To find $\mathbf{E}_{\text{local}}$, which is a field at the microscopic scale, we subtract from the macroscopic net field $\mathbf{E}$ the contribution $\mathbf{E}_{\text{pol}}$ from the other molecules inside the sphere, calculated according to our polarisation model, and add back what we hope is a more accurate field $\mathbf{E}_{\text{near}}$ from these nearby molecules: $\mathbf{E}_{\text{local}} = \mathbf{E} - \mathbf{E}_{\text{pol}} + \mathbf{E}_{\text{near}}$. We have already found a powerful expression, eq. (3.12), for the integral over the sphere of the field due to the charge distribution inside:

$$\int_{r<R} \mathbf{E}(\mathbf{x})\, \mathrm{d}^3 x \;=\; -\,\frac{\mathbf{p}_{\text{tot}}}{3\epsilon_0}$$

where, here, $\mathbf{p}_{\text{tot}}$ is the total induced dipole moment of all the molecules for $r < R$. Then the average field inside the sphere due to that induced dipole moment is:

$$\mathbf{E}_{\text{pol}} \;=\; -\,\frac{\mathbf{p}_{\text{tot}}}{3\epsilon_0}\,\frac{3}{4\pi R^3} \;=\; -\,\frac{\mathbf{P}}{3\epsilon_0} \tag{3.26}$$

where the use of a uniform polarisation (or dipole moment density) follows from the assumption of roughly uniform $\mathbf{E}$ over the sphere.

Now we need a good estimate for $\mathbf{E}_{\text{near}}$. An argument, due to Lorentz and summarised by Jackson on pp. 160–161, shows that in cubic crystals the field at the centre of the sphere from nearby dipoles vanishes. Therefore[†], $\mathbf{E}_{\text{local}} = \mathbf{E} - \mathbf{E}_{\text{pol}} + \mathbf{E}_{\text{near}} = \mathbf{E} + \mathbf{P}/3\epsilon_0$.

[†] In so-called ferroelectric crystals, the net field of the nearest neighbours, far from cancelling, can be so large as to cause spontaneous polarisation over a small volume, even without an external field! The net polarisation over macrosopic regions still averages to zero, except when an external field aligns these polarisations. The resulting effect can be very large and can persist after the external field is turned off.

Next, introduce the **molecular polarisability**, $\gamma_{\rm mol}$, defined so that, on average, the induced dipole moment of a molecule is:

$$\langle \mathbf{p}_{\rm mol} \rangle \;=\; \frac{\mathbf{P}}{N} \;=\; \gamma_{\rm mol}\,\epsilon_0\,\mathbf{E}_{\rm local} \;=\; \gamma_{\rm mol}\,(\epsilon_0\,\mathbf{E} \,+\, \mathbf{P}/3) \tag{3.27}$$

Recall that the polarisation is related to the *macroscopic* field by $\mathbf{P} \;=\; \epsilon_0 \chi_{\rm e} \mathbf{E}$. Comparing the two expressions yields:

$$\chi_{\rm e} \;=\; \frac{N\,\gamma_{\rm mol}}{1 \,-\, N\,\gamma_{\rm mol}/3} \tag{3.28}$$

This can be inverted to obtain the **Clausius-Mossotti equation** (aka **Lorentz-Lorenz formula**) for the (microscopic) molecular polarisability, $\gamma_{\rm mol}$, in terms of the (macroscopic) dielectric constant, $\epsilon_{\rm r}$:

$$\gamma_{\rm mol} \;=\; \frac{3}{N}\,\frac{\epsilon_{\rm r} - 1}{\epsilon_{\rm r} + 2} \tag{3.29}$$

It is clear that $\gamma_{\rm mol}$ has dimensions of a volume. Since the only relevant volume is that of the molecule, we estimate that $\gamma_{\rm mol} \lesssim 10^{-29}\,{\rm m}^3$. Thus, we predict that the susceptibility will be very small ($\chi_{\rm e} \lesssim 10^{-3}$) for dilute subtances such as gases ($N \approx 10^{25}$), and of order unity for liquids and solids ($N \approx 10^{28} \,-\, 10^{29}$). These predictions agree well with experiment.

Next, we try to calculate $\gamma_{\rm mol}$ at the microscopic level. We take charges e of mass m to be bound to a molecular or atomic core by a force $\mathbf{F} = -m\omega_0^2\mathbf{x}$, ω_0 being a characteristic frequency. We immerse them in a field $\mathbf{E} = E\,\mathbf{k}$ which causes the equilibirium distance to change by $\delta\mathbf{x} \;=\; e\mathbf{E}/m\omega_0^2$, resulting in an induced dipole moment $\mathbf{p}_{\rm mol} = e\delta\mathbf{x} = e^2\mathbf{E}/m\omega_0^2$. Then, comparing with eq. (3.27), $\gamma_{\rm mol} = e^2/(\epsilon_0 m\omega_0^2)$.

But we should worry about the possible effect of thermal fluctuations. In our first calculation, the molecules or atoms have no permanent dipole moment, so that the Hamiltonian of the charges is (note: $\mathbf{p}$ here is momentum!)

$$H \;=\; \frac{\mathbf{p}^2}{2m} \,+\, \frac{1}{2}m\omega_0^2\mathbf{x}^2 \,-\, eEz \;=\; \frac{\mathbf{p}^2}{2m} \,+\, \frac{1}{2}m\omega_0^2\mathbf{x}'^2 \,-\, \frac{e^2E^2}{2m\omega_0^2}$$

where $\mathbf{x}' = \mathbf{x} - eE\mathbf{k}/m\omega_0^2$.

From statistical mechanics, the average dipole moment in the z direction is:

$$\begin{aligned}
\langle p_{\rm mol} \rangle \;&=\; \frac{\int {\rm d}^3p\,{\rm d}^3x\; ez\,{\rm e}^{-H/kT}}{\int {\rm d}^3p\,{\rm d}^3x\,{\rm e}^{-H/kT}} \\
&=\; \frac{\int {\rm d}^3p\,{\rm d}^3x'\,(ez' \,+\, e^2E/m\omega_0^2)\,{\rm e}^{-H/kT}}{\int {\rm d}^3p\,{\rm d}^3x'\,{\rm e}^{-H/kT}} \\
&=\; \frac{e^2E}{m\omega_0^2}
\end{aligned} \tag{3.30}$$

To obtain the last line, we have noticed that the integrand with the z' term is odd since H is even in z', making the integral of the first term vanish. This result shows that thermal fluctuations do not affect the average dipole moment, and there comes

$$\gamma_{\rm mol} \;=\; \frac{\langle p_{\rm mol} \rangle}{\epsilon_0\,E} \;=\; \frac{1}{\epsilon_0}\frac{e^2}{m\omega_0^2} \tag{3.31}$$

The resonant frequency for helium is measured to be $\omega_0 = 3.71 \times 10^{16}$/s; for STP conditions, $N = 2.69 \times 10^{25}/{\rm m}^3$. Inserting these data into the last equation and into eq. (3.28) yields a susceptibility of 6.21×10^{-5} about 10% off the measured value of 6.84×10^{-5}. To do better we would need quantum mechanics.

When molecules have a permanent dipole moment, say $\mathbf{p}_0$, the external field tries to orient the dipoles along its direction, but this ordering effect is impeded by thermal fluctuations. Also, the orientations of nearby dipoles become correlated in a way that invalidates the Clausius-Mossotti equation.

The relevant Hamiltonian for such a dipole in an external field is $H = H_0 - p_0\,E\cos\theta$, where θ is the angle between the dipole and the field $\mathbf{E} = E\mathbf{k}$. Because of the azimuthal symmetry, the average of the permanent moment components in directions other than z must vanish. After integrating over irrelevant variables, we arrive at the **Langevin formula**:

$$\begin{aligned}\langle p_{\text{mol}}\rangle &= \frac{\int \mathrm{d}\Omega\, p_0\cos\theta\, \mathrm{e}^{p_0 E\,\cos\theta/kT}}{\int \mathrm{d}\Omega\, \mathrm{e}^{p_0 E\cos\theta/kT}} \\ &= p_0\,\mathrm{d}_u\,\ln\int_0^\pi \sin\theta\,\mathrm{e}^{u\,\cos\theta}\,\mathrm{d}\theta \;=\; p_0\,\mathrm{d}_u\,\ln\left(\frac{2\sinh u}{u}\right) \\ &= p_0\,(\operatorname{cotanh} u\;-\;1/u)\end{aligned}$$

where $u = p_0\,E/kT$. Now, except at low temperatures where quantum effects will in any case invalidate this classical calculation, u is very small, and $\operatorname{cotanh} u \approx (1+u^2/3)/u$. Thus, the average induced dipole moment is:

$$\langle p_{\text{mol}}\rangle \;\approx\; \frac{1}{3}\frac{p_0^2}{kT}\,E \tag{3.32}$$

Combining induced and permanent contributions, the general dependence of molecular polarisability on temperature is of the form $a + b/T$, in agreement with experimental data.

Example 3.1. An Example of a Boundary-Value Problem with Dielectric Media (section J4.4)

We consider the second example treated by Jackson in this section, that of an uncharged sphere of radius a and dielectric constant ϵ_{r}, with a uniform applied field $\mathbf{E}_{\text{app}} = E_0\mathbf{k}$ and associated potential $\Phi_{\text{app}} = -E_0\,z$. Since this problem has azimuthal symmetry with no free charge, the solution must be of the form:

$$\Phi_{\text{in}} \;=\; \sum_{l=0}^{\infty} A_l\,r^l\,P_l(\cos\theta) \qquad\qquad \Phi_{\text{out}} \;=\; \sum_{l=0}^{\infty}\left[B_l\,r^l \;+\; C_l\,r^{-l-1}\right]P_l(\cos\theta)$$

Matching the exterior and interior potentials at $r = a$ brings:

$$A_l \;=\; B_l \;+\; \frac{C_l}{a^{2l+1}}$$

whereas condition (3.25) is:

$$\epsilon_r\,A_l \;=\; B_l \;-\; \frac{l+1}{l}\,\frac{C_l}{a^{2l+1}}$$

This system can be rewritten as:

$$A_l \;=\; B_l\,\frac{2l+1}{l(\epsilon_{\text{r}}+1)+1} \qquad\qquad C_l \;=\; B_l\,\frac{\epsilon_{\text{r}}-1}{l(\epsilon_{\text{r}}+1)+1}\,a^{2l+1}$$

We expect the induced potential to go to zero at infinity, so that $\Phi_{\text{out}} \to -E_0\,z$. This can happen only if $B_l = -\,\delta_{1l}E_0$. Then our coefficients reduce to:

$$A_1 \;=\; -\,\frac{3}{\epsilon_{\text{r}}+2}\,E_0 \qquad\qquad C_1 \;=\; \frac{\epsilon_{\text{r}}-1}{\epsilon_{\text{r}}+2}\,a^3\,E_0$$

leading to the potential:

$$\begin{aligned}\Phi_{\text{in}} &= -\,\frac{3}{\epsilon_{\text{r}}+2}\,E_0\,z \\ \Phi_{\text{out}} &= -\,E_0\,z \;+\;\frac{\epsilon_{\text{r}}-1}{\epsilon_{\text{r}}+2}\left(\frac{a}{r}\right)^3 E_0\,z\end{aligned} \tag{3.33}$$

For a conductor ($\epsilon_r \to \infty$), these potentials reduce to the correct ones for an uncharged spherical conductor immersed in a uniform electric field.

The potential induced inside the sphere is:

$$\Phi_{\text{pol}} \;=\; \Phi_{\text{in}} \;-\; \Phi_{\text{app}} \;=\; \frac{\epsilon_r - 1}{\epsilon_r + 2}\, E_0\, z$$

corresponding to an induced field directed *opposite* the applied field. The net field inside is:

$$\mathbf{E}_{\text{in}} \;=\; \frac{3}{\epsilon_r + 2}\, \mathbf{E}_0 \tag{3.34}$$

and so the polarisation is:

$$\mathbf{P} \;=\; (\epsilon_r - 1)\,\epsilon_0\, \mathbf{E}_{\text{in}} \;=\; 3\epsilon_0 \left(\frac{\epsilon_r - 1}{\epsilon_r + 2}\right) \mathbf{E}_0 \;=\; -3\epsilon_0\, \mathbf{E}_{\text{pol}} \tag{3.35}$$

Note that the polarisation is uniform. This means that the induced contribution to the potential in eq. (3.16) comes only from a thin layer near the surface of the sphere, over which the polarisation decreases to zero. As we have argued, the whole effect can be modelled as being the result of the induced *surface* charge density:

$$\sigma_{\text{pol}} \;=\; \mathbf{P} \cdot \hat{\mathbf{n}} \;=\; 3\epsilon_0 \left(\frac{\epsilon_r - 1}{\epsilon_r + 2}\right) E_0\, \cos\theta \tag{3.36}$$

whose θ dependence is typical of polarisation charge.

3.5 Electrostatic Energy in a Dielectric Medium (section J4.7)

Suppose there exists a macroscopic, localised charge density $\rho(\mathbf{x})$. The energy associated with a small change $\delta\rho_{\text{free}}$ in the *free*-charge distribution, in the presence of an applied potential Φ, reads:

$$\delta W \;=\; \int \delta\rho_{\text{free}}(\mathbf{x})\, \Phi(\mathbf{x})\, \mathrm{d}^3x \tag{3.37}$$

Using $\boldsymbol{\nabla} \cdot (\delta\mathbf{D}) = \delta\rho_{\text{free}}$ and integrating by parts as we have done in our previous treatment of energy in free space, we find that:

$$\delta W \;=\; \int \mathbf{E} \cdot \delta\mathbf{D}\, \mathrm{d}^3x \tag{3.38}$$

This counts *all* the energy, including what is required to polarise the medium.

For linear, isotropic media, however, $\mathbf{E} \cdot \delta\mathbf{D} = \frac{1}{2}\delta(\mathbf{E} \cdot \mathbf{D})$, and the work becomes:

$$\delta W \;=\; \delta\left[\frac{1}{2}\int \mathbf{E} \cdot \mathbf{D}\, \mathrm{d}^3x\right] \;=\; \delta\left[\int \frac{1}{2}\epsilon\, E^2\, \mathrm{d}^3x\right] \tag{3.39}$$

so that the energy density stored in the total charge configuration is $\epsilon E^2/2$. This fits with our earlier assertion that in linear and homogeneous media, substituting ϵ for ϵ_0 allows the use of free-space results. It is also equivalent to:

$$W \;=\; \frac{1}{2}\int \rho_{\text{free}}(\mathbf{x})\, \Phi(\mathbf{x})\, \mathrm{d}^3x$$

showing that this latter expression is only valid for linear media. Otherwise the more general eq. (3.38) must be used to calculate the energy of the dielectric medium. (But it is not always clear what meaning can be ascribed to "energy" in those non-linear systems for which the work done to produce the final state often depends on the path taken!)

4 A. Magnetostatics (sections J5.1–5.12)

4.1 Sources of the Magnetic Field

Electrical phenomena involve two kinds of charges, positive and negative. Similarly, magnetostatics involves two poles, usually dubbed North and South. The all-important difference, however, is that those poles cannot be isolated like electric charges. So far as we know, **magnetic monopoles** do not exist. At bottom, magnetism arises from electric charges *in motion*. This motion gives rise to a **current density**, $\mathbf{J}$, with units of charge density times velocity, or charge per unit area per unit time.

Now electric charge is conserved and, like anything that is conserved, its density and current density are related by a **continuity equation**:

$$\partial_t \rho \;+\; \boldsymbol{\nabla}\cdot\mathbf{J} \;=\; 0 \tag{4.1}$$

The physical meaning of this equation is clearer when it is integrated over an arbitrary volume and the divergence theorem used to convert the $\mathbf{J}$ integral to a flux integral. Then we recognise that any increase/decrease in the charge density within the volume corresponds to a flux of charge entering/leaving the volume through the surface enclosing the volume.

4.2 Biot and Savart, Ampère Laws

The fundamental field $\mathbf{B}$ in magnetism is called **magnetic flux density or magnetic induction field** by Jackson (and many others), and we shall do the same for consistency. Then, according to the **Law of Biot and Savart**:

$$\mathbf{B}(\mathbf{x}) \;=\; k_{\rm m}\int \frac{I\,\mathrm{d}\mathbf{l}\times(\mathbf{x}-\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|^3} \tag{4.2}$$

where $I = \int \mathbf{J}\cdot\mathrm{d}\mathbf{a}$ is the **current** (or amount of charge per unit of time) flowing along a length element $\mathrm{d}\mathbf{l}$ situated at $\mathbf{x}'$ and pointing in the direction of the current, assuming that the surface (the cross-section of a wire, for instance) through which I flows is small enough. $k_{\rm m} \equiv \mu_0/4\pi = 10^{-7}\,\mathrm{N/A^2}$ in SI units. It is remarkable that this expression holds even when the motion of the charges is relativistic.

Lines of magnetic induction are *always* closed, and they are perpendicular both to the current element and to the distance between the current element and the point of observation, which means that they are concentric around the current element.

In terms of current density $\mathbf{J}$, the Law of Biot and Savart is:

$$\begin{aligned}\mathbf{B}(\mathbf{x}) \;&=\; k_{\rm m}\int \mathbf{J}(\mathbf{x}')\times\frac{\mathbf{x}-\mathbf{x}'}{|\mathbf{x}-\mathbf{x}'|^3}\,\mathrm{d}^3x' \;=\; -\,k_{\rm m}\int \mathbf{J}(\mathbf{x}')\times\boldsymbol{\nabla}\frac{1}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x' \\ &=\; k_{\rm m}\,\boldsymbol{\nabla}\times\int\frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x'\end{aligned} \tag{4.3}$$

One piece of information we draw from the continuity equation (4.1) is that when $\partial_t\rho = 0$ everywhere, as happens in the **steady state**, $\boldsymbol{\nabla}\cdot\mathbf{J} = 0$. Look at this as a constraint on magnetic sources in magnetostatics. A moving point-charge does not give rise to a steady-state current density as its fields cannot be static.

The last form for $\mathbf{B}$ in eq. (4.3) is often much more useful than the first.

Example 4.1. Consider a straight wire of length $2L$ with a uniform $\mathbf{J} = J\,\mathbf{k}$. Put the origin of the z axis mid-way down the wire and the observation point in the xy plane at distance ρ from the wire. With $I = J\delta(x-x')\delta(y-y')$:

$$\begin{aligned}\int\frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x' \;&=\; I\,\mathbf{k}\int_{-L}^{L}\frac{1}{\sqrt{\rho^2+z'^2}}\,\mathrm{d}z' \\ &=\; 2I\,\mathbf{k}\,\big[-\ln\rho\;+\;\ln(L+\sqrt{\rho^2+L^2})\big]\end{aligned}$$

As seen from the x-y plane, the source has cylindrical symmetry, and its field is found (EXERCISE) by taking the curl in cylindrical coordinates. When L is very large, the second term goes to a constant, and $\mathbf{B} = (2k_{\rm m}I/\rho)\hat{\boldsymbol{\phi}}$.

Although the vector structure of the $\mathbf{B}$ field is rather different from that of the electric field $\mathbf{E}$, there are often circumstances in which $\mathbf{B}$ turns out to be analogous to $\mathbf{E}$ for similar source geometry, as we have seen in the above example.

4.2.1 Gauss' Law for B and Ampère's Law (section J5.3)

According to eq. (4.3), the magnetic induction field can be written as a curl, which means that its divergence vanishes. On the other hand, we could argue directly that the absence of magnetic monopoles anywhere means that $\mathbf{B}$ must be divergenceless:

$$\boldsymbol{\nabla}\cdot\mathbf{B} \;=\; 0 \tag{4.4}$$

This is Gauss' Law for magnetism in differential form, and it shows that the Law of Biot and Savart is consistent with the non-existence of magnetic charge. Gauss' Law imposes a constraint on any $\mathbf{B}$, but not on the source $\mathbf{J}$.

To obtain a differential form for Biot and Savart's Law, we must take $\boldsymbol{\nabla}\times\mathbf{B}$, then use $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{V} = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{V}) - \nabla^2\mathbf{V}$ and replace $\boldsymbol{\nabla}$ by $-\boldsymbol{\nabla}'$ after bringing it into one integral:

$$\boldsymbol{\nabla}\times\mathbf{B} \;=\; k_{\rm m}\,\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\int\frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,{\rm d}^3x' \;=\; k_{\rm m}\,\boldsymbol{\nabla}\left[\int\mathbf{J}(\mathbf{x}')\cdot\boldsymbol{\nabla}\frac{1}{|\mathbf{x}-\mathbf{x}'|}\,{\rm d}^3x'\right] - k_{\rm m}\int\mathbf{J}(\mathbf{x}')\cdot\nabla^2\frac{1}{|\mathbf{x}-\mathbf{x}'|}\,{\rm d}^3x'$$

In the first integral, replace $\boldsymbol{\nabla}$ with $-\boldsymbol{\nabla}'$ and obtain a total divergence using $\boldsymbol{\nabla}'\cdot\mathbf{J}(\mathbf{x}') = 0$. When turned into a surface integral at infinity, this vanishes. Then eq. (0.18) yields:

$$\boldsymbol{\nabla}\times\mathbf{B} \;=\; 4\pi\,k_{\rm m}\,\mathbf{J} \tag{4.5}$$

Ampère's Law, as this is known, guarantees the conservation law $\boldsymbol{\nabla}\cdot\mathbf{J} = 0$ as an identity on its left-hand side.

An integral form for Ampère's Law is readily obtained by taking its flux through some open, not necessarily planar, surface and invoking Stoke's theorem to arrive at

$$\oint\mathbf{B}\cdot{\rm d}\mathbf{l} \;=\; 4\pi\,k_{\rm m}\,I \tag{4.6}$$

where I is the current in the loop enclosing the surface. $\oint\mathbf{B}\cdot{\rm d}\mathbf{l}$ is called the **circulation** of $\mathbf{B}$ around the loop.

4.3 The Vector Potential A (section J5.4)

We have seen that we could deduce $\boldsymbol{\nabla}\cdot\mathbf{B} = 0$ from the Biot and Savart Law, but starting instead from the divergence is very instructive. If its divergence vanishes *everywhere* (no magnetic monopoles), $\mathbf{B}$ can be written as the curl of a vector:

$$\mathbf{B} \;=\; \boldsymbol{\nabla}\times\mathbf{A} \tag{4.7}$$

where, to be consistent with Biot and Savart, $\mathbf{A}(\mathbf{x})$ must be:

$$\mathbf{A}(\mathbf{x}) \;=\; k_{\rm m}\int\frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,{\rm d}^3x' \;+\; \boldsymbol{\nabla}\Psi(\mathbf{x})$$

with Ψ an arbitrary scalar function. Gauss' Law for magnetism is then an *identity* on $\mathbf{A}(\mathbf{x})$, so contains no information about it. Such information is to be found in Ampère's Law:

$$\boldsymbol{\nabla}\times\mathbf{B} \;=\; \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{A}) \;\equiv\; \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) \;-\; \nabla^2\mathbf{A} \;=\; 4\pi\,k_{\rm m}\,\mathbf{J}$$

But, as we saw in the last section when deriving eq. (4.5), the divergence of the integral term in $\mathbf{A}$ vanishes, and $\boldsymbol{\nabla}\cdot\mathbf{A} = \nabla^2\Psi$ is *undetermined*, free for us to choose as we wish. A convenient choice here is $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$. Then Ψ is no longer completely arbitrary, but must satisfy the Laplace equation *everywhere*, and thus can only be zero. So we arrive at the **vector potential** *due to* a localised $\mathbf{J}$ distribution:

$$\mathbf{A}(\mathbf{x}) \;=\; k_{\mathrm{m}} \int \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x' \tag{4.8}$$

which satisfies both $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$ and a Poisson equation for each *Cartesian* component of $\mathbf{A}$:

$$\nabla^2\mathbf{A} \;=\; -\,4\pi\,k_{\mathrm{m}}\,\mathbf{J} \tag{4.9}$$

This argument, of course, is fully consistent with Biot and Savart's Law, but it serves to reinforce an important point: $\mathbf{A}$ is not fully determined by the theory. The sources nail down its curl, but its divergence is arbitrary!

As usual with vector integration, the integral in eq. (4.8) must be written first in Cartesian coordinates, after which we are free to transform each component to another coordinate system if we so wish.

Sometimes it is easier to find the field first from the integral form of Ampère's law, $\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0 I_{\mathrm{encl}}$, and then exploit its formal analogy with:

$$\oint \mathbf{A}\cdot\mathrm{d}\mathbf{l} \;=\; \int (\boldsymbol{\nabla}\times\mathbf{A})\cdot\mathrm{d}\mathbf{a} \;=\; \int \mathbf{B}\cdot\mathrm{d}\mathbf{a}$$

If there is sufficient symmetry and if $\mathbf{B}$ is known, one finds the flux of $\mathbf{B}$ and then proceeds as with Ampère's law. For instance, in the case of the infinite solenoid, the form $\mathbf{A} = k_{\mathrm{m}}\int \mathbf{J}/|\mathbf{x}-\mathbf{x}'|\,\mathrm{d}^3x'$ cannot be used because $\mathbf{J}$ is not localised. But it is not difficult to obtain $\mathbf{B}$ directly, and then $\mathbf{A}$ is only one step away (EXERCISE).

Now that we have introduced the vector potential, we can find useful volume integrals of the magnetic induction field, just like we did for the electrostatic field in example 2.1. The argument (see p. J187-188) proceeds along the same lines as with the latter, so we will only quote the results:

There are two possible outcomes to the integral of a magnetic induction field over a sphere of radius R:

a) The source currents lie entirely outside the sphere. Then:

$$\int_{r<R} \mathbf{B}(\mathbf{x})\,\mathrm{d}^3x \;=\; \frac{4\pi R^3}{3}\,\mathbf{B}(0) \tag{4.10}$$

Therefore, the average magnetic induction field over a sphere containing no current density is equal to the value of the field at its centre. (Note the typo in eq. J5.63)).

b) The current density lies entirely within the sphere. In this case:

$$\int_{r<R} \mathbf{B}(\mathbf{x})\,\mathrm{d}^3x \;=\; \frac{\mu_0}{3}\int \mathbf{x}'\times\mathbf{J}(\mathbf{x}')\,\mathrm{d}^3x' \tag{4.11}$$

4.4 Dipole Approximation to Fields of a Localised Current Distribution (section J5.6)

For each component of the vector potential, Eq. (4.8) is formally the same as the Coulomb equation for the electrostatic scalar potential. Therefore we expect to be able to express it as a multipole expansion similar to the one we found in chapter 3. That expansion, of course, is an expansion of $1/|\mathbf{x}-\mathbf{x}'|$, of which we already have obtained a couple of useful versions in terms of spherical harmonics. Here, we expand the i^{th} component of $\mathbf{A}$ (really it is $1/|\mathbf{x}-\mathbf{x}'|$ that is being expanded —see J5.50) in a Taylor series, keeping only the first two terms for the moment:

$$A^i(\mathbf{x}) \;=\; k_{\mathrm{m}}\left[\frac{1}{|\mathbf{x}|}\int J^i(\mathbf{x}')\,\mathrm{d}^3x' \;+\; \frac{\mathbf{x}}{|\mathbf{x}|^3}\cdot\int \mathbf{x}'\,J^i(\mathbf{x}')\,\mathrm{d}^3x'\right] \;+\; \ldots$$

If $\mathbf{J}$ is localised, then $\int \boldsymbol{\nabla}' \cdot (\mathbf{J}\, x_i')\mathrm{d}^3x' = 0$. But this is also $\int [\mathbf{J} \cdot \boldsymbol{\nabla}' x_i' + x_i' \boldsymbol{\nabla}' \cdot \mathbf{J}]\mathrm{d}^3x' = \int J^{k'} \partial_{k'} x_i' \mathrm{d}^3x' = \int J^i \mathrm{d}^3x'$. The second term in A^i, which never vanishes, becomes the dominant contribution, and it can be rewritten so that:

$$\mathbf{A}(\mathbf{x}) \;=\; k_\mathrm{m} \frac{\mathbf{m} \times \mathbf{x}}{|\mathbf{x}|^3} \;+\; \ldots \tag{4.12}$$

where identity J5.52 has been invoked, and:

$$\mathbf{m} \;=\; \frac{1}{2} \int \mathbf{x}' \times \mathbf{J}(\mathbf{x}')\, \mathrm{d}^3x' \tag{4.13}$$

is the **magnetic dipole moment** of a localised current distribution.

In the important case of a closed circuit, part of the integration is over the cross-section of the loop, and

$$\mathbf{m} \;=\; \frac{I}{2} \oint \mathbf{x}' \times \mathrm{d}\mathbf{l}' \tag{4.14}$$

where I is the current flowing throught he circuit and $\mathrm{d}\mathbf{l}'$ is a length element situated at $\mathbf{x}'$ along the loop with respect to the origin. Furthermore, if the loop is planar, $\mathbf{m}$ is perpendicular to the loop, and $\big|\mathbf{x} \times \mathrm{d}\mathbf{l}\big|/2 = \mathrm{d}a$, where $\mathrm{d}a$ is the area of the solid angle subtended by $\mathrm{d}l$. So the loop integral is just the area of the loop.

Thus, the magnetic moment of a planar loop has a magnitude equal to the current flowing in it multiplied by the loop's area.

To obtain the magnetic induction, take the curl of eq. (4.12) and use two identities on the left of Jackson's front cover:

$$\begin{aligned} \mathbf{B}_\mathrm{dip}(\mathbf{x}) \;&=\; k_\mathrm{m}\, \boldsymbol{\nabla} \times \left(\frac{\mathbf{m} \times \mathbf{x}}{r^3} \right) \\ &=\; k_\mathrm{m}\, (\mathbf{m} \cdot \boldsymbol{\nabla}) \frac{\mathbf{x}}{r^3} \\ &=\; k_\mathrm{m} \left[\frac{3\hat{\mathbf{n}}\, (\hat{\mathbf{n}} \cdot \mathbf{m}) \;-\; \mathbf{m}}{r^3} \right] \end{aligned} \tag{4.15}$$

where, as usual, $\hat{\mathbf{n}} = \mathbf{x}/r$. Again, we encounter the same form in electrostatics, for the dipole field given by eq. (3.11).

The dipole induction field splits nicely into components along, and perpendicular to, the direction of observation:

$$\mathbf{B}_\mathrm{dip} \;=\; k_\mathrm{m} \frac{m}{r^3} \left(2\cos\theta\, \hat{\mathbf{r}} \;+\; \hat{\boldsymbol{\theta}}\, \sin\theta \right) \tag{4.16}$$

where θ is the angle from the dipole moment to the direction of observation.

With eq. (4.13), our result of eq. (4.11) for the volume integral of $\mathbf{B}$ over a sphere of radius R containing currents, with none outside the sphere, becomes:

$$\int_{r<R} \mathbf{B}\, \mathrm{d}^3x \;=\; \frac{2\mu_0}{3}\, \mathbf{m} \tag{4.17}$$

As in the electrostatic case, however, the volume integral of a dipole magnetic induction over a sphere vanishes due to the spherical symmetry of the integrand about the origin of the dipole. To be fully consistent, we must write:

$$\mathbf{B}_\mathrm{dip}(\mathbf{x}) \;=\; k_\mathrm{m} \left[\frac{3\hat{\mathbf{n}}\, (\hat{\mathbf{n}} \cdot \mathbf{m}) \;-\; \mathbf{m}}{r^3} \;+\; \frac{8\pi}{3}\, \mathbf{m}\, \delta(\mathbf{x}) \right] \tag{4.18}$$

4.5 Magnetic Force and Energy for Localised Currents (section J5.7)

The magnetic force on a current element $I\mathrm{d}\mathbf{l}$ is $I\mathrm{d}\mathbf{l}\times\mathbf{B}$, so the total force exerted by an external $\mathbf{B}$ on a current distribution is just $\int[\mathbf{J}(\mathbf{x}')\times\mathbf{B}(\mathbf{x}')]\,\mathrm{d}^3x'$.If $\mathbf{B}$ varies slowly enough over the volume where $\mathbf{J}$ exists, we can expand it around some origin inside, so that:

$$F_i \;=\; \epsilon_{ijk}\left[B^k(0)\int J^j(\mathbf{x}')\,\mathrm{d}^3x' \;+\; \int \mathbf{x}'\cdot\boldsymbol{\nabla}B^k\Big|_{\mathbf{x}=0} J^j(\mathbf{x}')\,\mathrm{d}^3x' \;+\; \ldots\right]$$

As before, the first integral vanishes, and the second one can be written in terms of the magnetic moment by noticing that since $\boldsymbol{\nabla}B^k\big|_{\mathbf{x}=0}$ is a constant vector which can be taken out of the integral, replacing $\mathbf{x}$ with it in the expansion of $\mathbf{A}$ in section 4.4 leads, by analogy to eq. (4.12), to::

$$\begin{aligned} F_i \;&=\; \epsilon_{ijk}\left[(\mathbf{m}\times\boldsymbol{\nabla})^j\,B^k\right]_{\mathbf{x}=0} \\ &=\; \epsilon_{ijk}\epsilon^{jln}\,m_l\,\partial_n B^k \;=\; -\left(\delta_i{}^l\delta_k{}^n \;-\; \delta_i{}^n\delta_k{}^l\right)m_l\,\partial_n B^k \\ &=\; \partial_i(\mathbf{m}\cdot\mathbf{B}) \;-\; m_i\boldsymbol{\nabla}\cdot\mathbf{B} \end{aligned}$$

So, to lowest-order:

$$\mathbf{F} \;=\; \boldsymbol{\nabla}(\mathbf{m}\cdot\mathbf{B}) \tag{4.19}$$

Using the idea of the force as the negative gradient of potential energy, we immediately write the *interaction* energy (which is not the total energy!) of a dipole in an external magnetic induction:

$$U \;=\; -\,\mathbf{m}\cdot\mathbf{B} \tag{4.20}$$

4.6 Macroscopic Equations, Matching Conditions (section J5.8)

Just as in electrostatics, magnetic phenomena inside matter are complicated by the response of the matter to applied fields. For instance, existing atomic dipoles will tend to align with the applied field, creating their own field. Current sources will in general fluctuate rapidly at the atomic scale. Thus the need to average over a microscopically large but macroscopically small volume.

Define the **magnetisation** as the average magnetic moment density over a macroscopically small volume containing many dipoles and situated at point $\mathbf{x}'$:

$$\mathbf{M}(\mathbf{x}') \;=\; \sum_i N_i\,\langle\mathbf{m}_i\rangle$$

where N is the dipole number density. We combine the vector potential from free currents, eq. (4.8), with the matter contribution derived from eq. (4.12) to obtain, after integrating over all space:

$$\mathbf{A}(\mathbf{x}) \;=\; k_\mathrm{m}\left[\int\frac{\mathbf{J}_\mathrm{free}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x' \;+\; \int\mathbf{M}(\mathbf{x}')\times\frac{\mathbf{x}-\mathbf{x}'}{|\mathbf{x}-\mathbf{x}'|^3}\,\mathrm{d}^3x'\right]$$

The geometrical factor in the second integrand is now rewritten as $\boldsymbol{\nabla}'(1/|\mathbf{x}-\mathbf{x}'|)$. Using an identity in Jackson (left front cover), the integrand is converted to a curl plus $\left(\boldsymbol{\nabla}'\times\mathbf{M}(\mathbf{x}')\right)/|\mathbf{x}-\mathbf{x}'|)$. With the integral $\int\boldsymbol{\nabla}\times\mathbf{C}\,\mathrm{d}^3x = -\int_\mathrm{S}\mathbf{C}\times\hat{\mathbf{n}}'\,\mathrm{d}a$ on the right front cover, this gives:

$$\mathbf{A}(\mathbf{x}) \;=\; k_\mathrm{m}\left[\int\frac{\mathbf{J}_\mathrm{free}(\mathbf{x}') \;+\; \boldsymbol{\nabla}'\times\mathbf{M}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x' \;+\; \oint_S\frac{\mathbf{M}(\mathbf{x}')\times\hat{\mathbf{n}}'}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}a\right] \tag{4.21}$$

If $\mathbf{M}$ is localised and goes smoothly to zero (no discontinuity), the surface integral does not contribute because the volume integral can be taken over all space. Like in electrostatics, however, there can be interfaces near which the magnetisation changes very quickly in an unknowable way over a thin layer near the interface. So we replace this

thin layer by a surface across which $\mathbf{M}$ is discontinuous and thus not differentiable. The missing contribution is taken up by the surface integral, where S is now the interface, and there is a surface current, $\mathbf{M}(\mathbf{x}') \times \hat{\mathbf{n}}'$. In the absence of free currents inside the volume and if the magnetisation is uniform within, the whole vector potential is generated by these surface currents.

Because of the presence of the extra magnetisation current density, $\boldsymbol{\nabla} \times \mathbf{M}$, Ampère's Law must be modified as follows:

$$\boldsymbol{\nabla} \times \mathbf{H} \;=\; \mathbf{J}_{\text{free}} \tag{4.22}$$

where we have introduced the **magnetic field**:

$$\mathbf{H} \;=\; \frac{1}{4\pi k_{\text{m}}}\,\mathbf{B} \;-\; \mathbf{M} \tag{4.23}$$

$\mathbf{H}$ plays the same rôle in magnetism that $\mathbf{D}$ plays in electricity. Both take into account the contributions from atomic charges and currents to sources.

In terms of the vector potential, Ampère's Law becomes, with the choice $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$:

$$\nabla^2 \mathbf{A} \;=\; -\,4\pi\, k_{\text{m}}\,(\mathbf{J}_{\text{free}} \;+\; \boldsymbol{\nabla} \times \mathbf{M}) \tag{4.24}$$

This completely general Poisson-like equation has for solution eq. (4.21).

Unlike the response of most materials to applied electric fields, which is most often linear, the response to applied magnetic fields is typically much more complicated. Only in isotropic diamagnetic and paramegnetic materials can one write the simple analog to eq. (3.23), in SI units:

$$\mathbf{B} \;=\; \mu_{\text{r}}\,\mu_0 \mathbf{H} \;=\mu\,\mathbf{H} \tag{4.25}$$

where μ_{r} is the **relative permeability** of the material. In these cases, $1-\mu_{\text{r}} \approx \pm 10^{-5}$. For ferromagnetic materials, the relationship between the magnetic field $\mathbf{H}$ and the induction $\mathbf{B}$ is typically nonlinear and depends on history (hysteresis).

In linear and homogeneous magnetic media, the vector potential satisfies

$$\nabla^2 \mathbf{A} \;=\; -\,\mu_0(\mu_{\text{r}}\,\mathbf{J}_{\text{free}}) \tag{4.26}$$

Another difference with electric phenomena in matter is that, whereas electric polarisation results in a net field that is smaller than the applied field, in magnetism, when $\mu_{\text{r}} > 1$, the effect is the opposite: magnetisation results in a larger induction field than the external induction field.

4.6.1 Magnetostatic matching conditions

By using Gauss' Law for magnetism in its integral form, $\oint \mathbf{B} \cdot \mathrm{d}\mathbf{a} \;=\; 0$, it is easy to show that the component of $\mathbf{B}$ *normal* to any surface is continuous across the surface. The tangential components of $\mathbf{H}$, however, are discontinuous if there is a free surface current density $\mathbf{K}$, as can be derived from Ampère's law in matter. Therefore:

$$\begin{aligned} (\mathbf{B}_2 - \mathbf{B}_1) \cdot \hat{\mathbf{n}} \;&=\; 0 \\ \hat{\mathbf{n}} \times (\mathbf{H}_2 - \mathbf{H}_1) \;&=\; \mathbf{K} \end{aligned} \tag{4.27}$$

where, as before, $\hat{\mathbf{n}}$ points from medium 1 to medium 2.

The vector potential is continuous across a surface, but its normal derivative inherits the discontinuity in the *tangential* components of $\mathbf{H}$.

Example 4.2. Magnetic Potential and Induction of a Circular Current Loop (section J5.5)

The most important example of a magnetic induction field is that of a circular loop, say of radius a, carrying a counterclockwise current I, and with coordinate axes oriented so that it lies in the xy plane.

The source is $\mathbf{J} = J_\phi \hat{\boldsymbol{\phi}} = -J_\phi \sin\phi' \,\mathbf{i} + J_\phi \cos\phi' \,\mathbf{j}$, where $J_\phi = I\,\delta(\cos\theta')\,\delta(r'-a)/a$. With this, $\mathbf{J}$ is *perpendicular* to the x axis when $\phi' = 0$.

Without loss of generality, the azimuthal symmetry of the problem allows us to put the observation point in the xz plane, at $\phi = 0$. We have, with $\hat{\mathbf{x}}\cdot\hat{\mathbf{x}}' = \cos\theta\,\cos\theta' + \sin\theta\,\sin\theta'\,\cos\phi'$:

$$\mathbf{A}(r,\theta) \;=\; k_{\rm m}\int \frac{-J_\phi\,\sin\phi'\,\mathbf{i} + J_\phi\,\cos\phi'\,\mathbf{j}}{\sqrt{r^2+r'^2-2rr'\,(\cos\theta\,\cos\theta' + \sin\theta\,\sin\theta'\,\cos\phi')}}\,\mathrm{d}^3x' \tag{4.28}$$

The integral along the x axis vanishes by symmetry, leaving the y component, which is also the ϕ component at $\phi = 0$. Then, putting in J_ϕ and doing the δ-function integrals in spherical coordinates leaves:

$$A_\phi(r,\theta) \;=\; k_{\rm m}\,Ia\int_0^{2\pi} \frac{\cos\phi'\,\mathrm{d}\phi'}{\sqrt{r^2+a^2-2ar\,\sin\theta\,\cos\phi'}}$$

This can be recast in terms of elliptic integrals and so is a closed-form solution, but it is not so intuitive...

Instead, expand the inverse distance factor inside the integral (4.28) using eq. (2.10) with $\phi = 0$. Write $\cos\phi'$ in terms of complex exponential $\mathrm{e}^{\pm\mathrm{i}\phi'}$, and notice that this will restrict the m sum to its $m = \pm1$ terms[†], leaving $Y^*_{l,\pm1}(\theta',\phi')$, because $\int \mathrm{e}^{\mathrm{i}(m\pm1)\phi'}\,\mathrm{d}\phi' = 0$. For the $m = \pm1$ terms the integrands are no longer ϕ'-dependent, and the integral over ϕ' just yields 2π. After integrating over the δ-functions as before, we get:

$$A_\phi \;=\; 8\pi^2\,k_{\rm m}\,Ia\sum_{l=1}^{\infty}\frac{Y_{l,1}(\theta,0)}{2l+1}\,\frac{r_<^l}{r_>^{l+1}}\,Y_{l,1}(\pi/2,0)$$

where $r_<$ is the smaller of a and r, and $r_>$ the larger of the two.

Eq. (0.42) then leads to:

$$A_\phi(r,\theta) \;=\; -\,\pi\,k_{\rm m}\,Ia\sum_{n=0}^{\infty}(-1)^{n+1}\frac{(2n-1)!!}{2^n\,(n+1)!}\,\frac{r_<^{2n+1}}{r_>^{2(n+1)}}\,\sin\theta\,\mathrm{d}_x P_{2n+1}(x) \tag{4.29}$$

where $x = \cos\theta$ and $(2n-1)!! = ((2n-1)(2n-3)\ldots\times 3\times 1$. The magnetic induction is then obtained by taking the curl:

$$\mathbf{B}(r,\theta) \;=\; -\frac{1}{r}\left[\hat{\mathbf{r}}\,\partial_x\big(\sqrt{1-x^2}\,A_\phi\big) \;+\; \hat{\boldsymbol{\theta}}\,\partial_r(r\,A_\phi)\right] \tag{4.30}$$

The derivatives are straightfoward and one arrives at eq. J5.48 and 5.49, which we will not bother to reproduce here. Rather we will extract the asymptotic behaviour, far from the loop at $r >> a$. Then $r_< = a$ and $r_> = r$. The $n = 0$ term dominates the series. Since $\mathrm{d}_x P^1(x) = 1$, the potential is, to a good approximation:

$$A_\phi(r,\theta) \underset{r\gg a}{\longrightarrow} \pi\,k_{\rm m}\,I\,\frac{a^2}{r^2}\,\sin\theta$$

with a magnetic induction:

$$\mathbf{B}(r,\theta) \underset{r\gg a}{\longrightarrow} \frac{k_{\rm m}\,I\,(\pi a^2)}{r^3}\,(\hat{\mathbf{r}}\,2\cos\theta \;+\; \hat{\boldsymbol{\theta}}\,\sin\theta) \tag{4.31}$$

Alternatively, and more simply, we can derive (EXERCISE) the dipole approximation directly from eq. (4.12) and (4.18) in the previous section, recognising that here the area of the loop is πa^2.

[†]Note how azimuthal symmetry does not lead to $m = 0$, because the *Cartesian* components of the source current density have a ϕ dependence.

4.7 Magnetostatic Boundary-Value Problems with no Free Current (section J5.9)

A useful simplification can be made in a simply-connected region with *no free currents*. Indeed, in such a region, $\boldsymbol{\nabla} \times \mathbf{H} = 0$ allows us to define a magnetostatic *scalar* potential, $\Phi_{\rm m}$, with $\mathbf{H} = -\boldsymbol{\nabla}\Phi_{\rm m}$. In each linear region with uniform $\mu_{\rm r}$, this scalar potential obeys the Laplace equation, $\nabla^2\Phi_{\rm m} = 0$, which can be solved as a boundary-value problem, just as in chapter 2.

What can we say when there are no free currents but the magnetic material is non-linear? Sometimes, such as in substances whose magnetisation is more or less insensitive to applied fields (so long as these are not large), information about $\mathbf{M}$ may still be recoverable from the magnetostatic scalar potential. In general, Gauss' Law for magnetism, written in terms of $\mathbf{H}$ and $\mathbf{M}$, yields:

$$\nabla^2\Phi_{\rm m} \;=\; -\,\boldsymbol{\nabla}\cdot\mathbf{H} \;=\; \boldsymbol{\nabla}\cdot\mathbf{M} \;\equiv\; -\,\rho_{\rm m} \tag{4.32}$$

Here we have introduced an *effective* magnetic charge density. One form of the general solution of this Poisson equation is:

$$\Phi_{\rm m} \;=\; -\,\frac{1}{4\pi}\int \frac{\boldsymbol{\nabla}'\cdot\mathbf{M}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,{\rm d}^3x' \;+\; \frac{1}{4\pi}\oint_S \frac{\hat{\mathbf{n}}'\cdot\mathbf{M}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,{\rm d}a' \tag{4.33}$$

where the surface integral contributes only when $\mathbf{M}$ goes to zero in discontinuous fashion at a surface. The origin of this integral can be traced to the existence of an *effective* magnetic surface charge density, $\sigma_{\rm m} = \hat{\mathbf{n}}\cdot\mathbf{M}$, which is easiest to exhibit by taking the volume integral of $\boldsymbol{\nabla}\cdot\mathbf{M} = -\rho_{\rm m}$ over a small cylinder enclosing a surface element, invoking the divergence theorem and evaluating the surface integral when $M = 0$ on one side of the surface.

The presence of $\sigma_{\rm m}$ will cause some lines of $\mathbf{H}$ (but not of $\mathbf{B}$!) to terminate on the surface.

An alternative fornulation of the solution starts with just the first term of eq. (4.33), ignoring any possible discontinuity. So long as $\mathbf{M}$ is localised, an integration by parts (or use of an identity on Jackson's front left cover) produces:

$$\begin{aligned}\Phi_{\rm m} \;&=\; \frac{1}{4\pi}\int_{\text{all space}} \mathbf{M}(\mathbf{x}')\cdot\boldsymbol{\nabla}'\frac{1}{|\mathbf{x}-\mathbf{x}'|}\,{\rm d}^3x' \\ &=\; -\,\frac{1}{4\pi}\,\boldsymbol{\nabla}\cdot\int \frac{\mathbf{M}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,{\rm d}^3x'\end{aligned} \tag{4.34}$$

where in the last line $\boldsymbol{\nabla}'$ was first turned into $-\boldsymbol{\nabla}$ and then taken outside. The advantage of this particular form of the solution is that it can handle *any* $\mathbf{M}$, even a discontinuous one. So, if one chooses to use it, there is no need for a surface integral.

In the first line in the last equation, we could instead have calculated the gradient and taken the large distance limit: $|\mathbf{x}-\mathbf{x}'| \approx |\mathbf{x}| = r$, so that far from any source:

$$\Phi_{\rm m} \;\approx\; \frac{\mathbf{m}\cdot\mathbf{x}}{4\pi r^3} \tag{4.35}$$

where $\mathbf{m} = \int \mathbf{M}(\mathbf{x}'){\rm d}^3x'$ is the total dipole moment of the distribution. We recognise this as a dipole scalar potential giving rise to the same kind of field as the electrostatic dipole field, but with $\mathbf{p} \to \mathbf{m}$.

Example 4.3. Uniformly Magnetised Sphere (section J5.10)

Consider a sphere of radius a and permanent magnetisation $\mathbf{M} = M_0\,\mathbf{k}$, embedded in a non-magnetic medium. Because of the uniform magnetisation and discontinuity at $r = a$, only the surface integral in eq. (4.33) contributes, with $\sigma_{\rm m} = M_0\,\cos\theta$. This has azimuthal symmetry, whence:

$$\Phi_{\rm m} \;=\; \frac{M_0 a^2}{4\pi}\int \frac{\cos\theta'}{|\mathbf{x}-\mathbf{x}'|}\,{\rm d}\Omega'$$

We have encountered this kind of integral before in section 2.1, where we made use of the expansion (2.10). Since $\cos\theta' = P_1(\cos\theta')$, only the $l = 1$ term survives the integration. We find:

$$\Phi_{\rm m}(r,\theta) \;=\; \begin{cases} \dfrac{1}{3}\,M_0\,z & (r<a) \\[2ex] \dfrac{m}{4\pi r^3}\,z & (r\geq a) \end{cases}$$

where the total dipole moment of the sphere is $\mathbf{m} = 4\pi a^3\mathbf{M}/3$. We see that the exterior potential is that of a dipole not only asymptotically, but *everywhere*. Inside the sphere, the magnetic field and induction are:

$$\mathbf{H}_{\rm in} \;=\; -\frac{1}{3}\,\mathbf{M} \qquad\qquad \mathbf{B}_{\rm in} \;=\; \frac{2\mu_0}{3}\,\mathbf{M} \tag{4.36}$$

We observe that $\mathbf{B}$ is parallel, and $\mathbf{H}$ antiparallel, to $\mathbf{M}$.

We could also have found the vector potential with eq. (4.21). This is a bit more involved because of the vector character of the integrand, but the derivation of the fields proceeds pretty much along the same lines and with the same type of argument, as can be seen in section J5.10.

Example 4.4. Magnetised Sphere in an External Field (section J5.11)

We can certainly take the interior fields found in the last example and superpose them with any uniform external field $\mathbf{B}_0$. Suppose that instead of being permanent, $\mathbf{M}$ is entirely *induced* by the external field, as in a paramagnetic or diamagnetic material. That induced uniform magnetisation still generates the fields given in eq. (4.36), to which the external field is added to produce the total fields $\mathbf{B}_{\rm in}$ and $\mathbf{H}_{\rm in}$, But we must also have: $\mathbf{B}_{\rm in} = \mu\,\mathbf{H}_{\rm in}$, that is:

$$\mathbf{B}_0 \,+\, \frac{2\mu_0}{3}\,\mathbf{M} \;=\; \mu\left(\frac{1}{\mu_0}\,\mathbf{B_0} \,-\, \frac{1}{3}\,\mathbf{M}\right)$$

which yields for the magnetisation inside the sphere:

$$\mathbf{M} \;=\; \frac{3}{\mu_0}\left(\frac{\mu_{\rm r}\,-\,1}{\mu_{\rm r}\,+\,2}\right)\mathbf{B}_0 \tag{4.37}$$

Observe the analogy with the electric polarisation in a dielectric sphere in chapter 3.

In a ferromagnetic material, however, we must seek another relation between $\mathbf{B}_{\rm in}$ and $\mathbf{H}_{\rm in}$. We can always eliminate the magnetisation from the expressions for each of $\mathbf{B}_{\rm in}$ and $\mathbf{H}_{\rm in}$ to obtain (for a uniformly magnetised sphere only!):

$$\mathbf{B}_{\rm in} \;+\; 2\mu_0\,\mathbf{H}_{\rm in} \;=\; 3\mathbf{B}_0$$

To go further, we need the **hysteresis curve**. If, for instance, the external field is increased until $\mathbf{B}_{\rm in}$ reaches saturation and is then brought back to zero, one should look at the resulting hysteresis curve (see fig. J5.12) for its intersection with the straight line of slope -2 with $3B_0$ as intercept.

Example 4.5. An Example of Magnetic Shielding (section J5.12)

Consider a spherical shell of inner radius a and outer radius b, with a linear permeability μ, immersed in a constant and uniform induction $\mathbf{B}_0 = B_0\,\hat{\mathbf{k}}$. There are no free currents. Then our remarks at the beginning this section (or J5.9B) lead to the existence of a scalar potential $\Phi_{\rm m}$ which obeys the Laplace equation everywhere and which goes to $-H_0 z$ as $r \to \infty$. The following general solutions are regular:

$$\Phi_{\rm m} = \begin{cases} -H_0 z + \displaystyle\sum_{l=0} \alpha_l\, r^{-(l+1)}\, P_l(\cos\theta) & b \le r < \infty \\ \displaystyle\sum_{l=0} \left(\beta_l\, r^l + \gamma_l\, r^{-(l+1)}\right) P_l(\cos\theta) & a \le r \le b \\ \displaystyle\sum_{l=0} \kappa_l\, r^l\, P_l(\cos\theta) & 0 \le r \le a \end{cases} \tag{4.38}$$

Continuity of $\Phi_{\rm m}$ at $r = b$ and $r = a$ gives, respectively:

$$\begin{aligned} \alpha_l - \beta_l\, b^{2l+1} - \gamma_l &= H_0\, \delta_{l1} b^{l+2} \\ \beta_l a^{2l+1} + \gamma_l - \kappa_l\, a^{2l+1} &= 0 \end{aligned}$$

$\mu_{\rm r2}\,\partial_n \Phi_{\rm m2} = \mu_{\rm r1}\,\partial_n \Phi_{\rm m1}$ at $r = b$ and $r = a$ gives, respectively:

$$\begin{aligned} (l+1)\alpha_l + \mu_{\rm r}\, l\, \beta_l b^{2l+1} - \mu_{\rm r}(l+1)\gamma_l &= -H_0 \delta_{l1} b^{l+2} \\ \mu_{\rm r}\, l\, \beta_l a^{2l+1} - \mu_{\rm r}(l+1)\gamma_l - l\,\kappa_l\, a^{2l+1} &= 0 \end{aligned}$$

When $l \neq 1$, the system has only a trivial solution: all coefficients vanish. Only $l = 1$ terms contribute. Give the system with $l = 1$ to ***Maple***; which outputs:

```
> solve(sys_eq,{alpha[1],beta[1],gamma[1],kappa[1]});
```

$$\kappa_1 = -9\,\frac{\mu_r H_0 b^3}{4\,\mu_r a^3 - 2\,a^3 + 5\,b^3\mu_r + 2\,b^3 - 2\,\mu_r^2 a^3 + 2\,b^3\mu_r^2},\ \gamma_1 = -3\,\frac{a^3 H_0 b^3(\mu_r - 1)}{4\,\mu_r a^3 - 2\,a^3 + 5\,b^3\mu_r + 2\,b^3 - 2\,\mu_r^2 a^3 + 2\,b^3\mu_r^2}$$

$$\alpha_1 = \frac{H_0 b^3\left(-2\,\mu_r^2 a^3 + \mu_r a^3 + 2\,b^3\mu_r^2 - b^3\mu_r + a^3 - b^3\right)}{4\,\mu_r a^3 - 2\,a^3 + 5\,b^3\mu_r + 2\,b^3 - 2\,\mu_r^2 a^3 + 2\,b^3\mu_r^2},\ \beta_1 = -3\,\frac{(2\,\mu_r + 1)\,H_0 b^3}{4\,\mu_r a^3 - 2\,a^3 + 5\,b^3\mu_r + 2\,b^3 - 2\,\mu_r^2 a^3 + 2\,b^3\mu_r^2}$$

We are interested in the cavity surrounded by the shell, where $\Phi_{r<a} = \kappa_1\, z$ and $H_{r<a} = -\kappa_1$. Force ***Maple*** to simplify the denominator of the coefficients nicely:

```
> factor(selectremove(has,expand(denom(rhs(op(1,%)))/b^3),a)[1]) +
> factor(selectremove(has,expand(denom(rhs(op(1,%)))/b^3),a)[2]);
```

$$-2\,\frac{a^3\,(\mu_r - 1)^2}{b^3} + (\mu_r + 2)\,(2\,\mu_r + 1)$$

Then:

$$\kappa_1 = -\left[\frac{9\mu_{\rm r}}{(2\mu_{\rm r}+1)(\mu_{\rm r}+2) - 2\,\frac{a^3}{b^3}(\mu_{\rm r}-1)^2}\right] H_0 \underset{\mu_{\rm r}\gg 1}{\longrightarrow} \kappa_1 \approx -\frac{9}{2\mu_{\rm r}\,(1 - a^3/b^3)}\,H_0$$

So long as the shell is not too thin, κ_1 can be quite small and the interior field can be much smaller than the external one. The spherical shell thus acts as a magnetic shield that bends field lines around the inner cavity.

4 B. Time-dependent Magnetic Phenomena

4.8 Faraday's Law of Induction (section J5.15)

When a conducting loop is moved in a static magnetic field, it is an experimental fact that a current is set up in the loop proportional to the rate of change of the magnetic induction flux through the loop. This current multiplied by the loop's resistance (assuming Ohm's law holds) is a quantity called the **electromotive force**, or **emf** (often dubbed **motional emf**), a quite misleading name really since its dimensions are those of a force per unit charge times length. The phenomenon is due to the magnetic force acting on the free charges in the conductor, and it can be shown that the emf is indeed:

$$\mathcal{E}_{\text{motion}} \;=\; -\mathrm{d}_t \int \mathbf{B}\cdot \mathrm{d}\mathbf{a} \tag{4.39}$$

Where things get interesting is that, as was discovered by Faraday, if the loop is kept at rest in the magnetic induction but the latter varied, the same phenomenon occurs: when the induction field is generated by a current flowing in another loop, only the *relative* motion of the two loops matters. This is quite striking, because when the magnetic source is moving, there is no magnetic force on the charges in the other loop since they are at rest. The only possible explanation is that an *electric* field is induced which in turn produces the emf that sets up the current. So we write:

$$\mathcal{E} \;=\; \oint \mathbf{E}'\cdot \mathrm{d}\mathbf{l} \;=\; -\,\mathrm{d}_t F$$

$\mathbf{E}'$ is the electric field measured *in the loop's rest-frame*. This electric field is different from an electrostatic field: its sources are not monopoles and its circulation does not vanish.

There is a catch, however. If the source loop is moving with respect to the observer, the magnetic flux is not the same as if it were at rest! The time derivative (see section 0.9) is in fact a *convective* derivative. In order to be able to bring the time derivative inside as only a partial derivative, without extra terms, both the electric and the magnetic induction fields must be understood to be in the same reference frame, eg., the lab frame. We arrive at **Faraday's induction law**:

$$\oint \mathbf{E}\cdot \mathrm{d}\mathbf{l} \;=\; -\int \partial_t \mathbf{B}\cdot \mathrm{d}\mathbf{a} \tag{4.40}$$

The direction of integration around the loop is given by the right-hand rule, with the thumb pointing in the direction of $\mathrm{d}\mathbf{a}$. The direction of integration is essentially the direction of the current that would be set up if there were a wire loop to carry it. The direction of this current is usually determined by **Lenz's Law**, according to which the direction is such that the flux of the magnetic induction of this current opposes the *change* in flux of the external field through the loop. It is best viewed as preventing runaway buildups of emf, acting as a sort of inertia.

When we convert the left-hand side of eq. (4.40) to a flux integral and note that the path and surface are arbitrary, we obtain the important differential form:

$$\boldsymbol{\nabla}\times\mathbf{E} \;+\partial_t\mathbf{B} \;=\; 0 \tag{4.41}$$

If we put $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$ into the differential Faraday law, we see that now it is the curl of $\mathbf{E}+\partial_t\mathbf{A}$ that vanishes everywhere. So $\mathbf{E}+\partial_t\mathbf{A}$ must be the gradient of some scalar function, and we obtain the most general form for $\mathbf{E}$:

$$\mathbf{E} = -\,\partial_t\mathbf{A} \;-\; \boldsymbol{\nabla}\Phi \tag{4.42}$$

where Φ is just the electrostatic potential. When the time-varying magnetic field is the only source for the electric field (no free charges anywhere), we can drop the electrostatic term.

Note that we have expressed the six components of the electric and magnetic fields in terms of the *four* in the $\mathbf{A}$ and Φ potentials. Clearly, the field components are not all independent. In fact, writing $\mathbf{E}$ and $\mathbf{B}$ in terms of the potentials takes the place of the two homogeneous equations, Gauss' Law for magnetism and Faraday's Law. A burning question now arises: just how many independent variables (or **degrees of freedom**) are there in electromagnetism? The question will be answered in chapter 5, and we will learn why it is vital that it be answered.

4.9 Quasi-static Magnetic Fields in Conductors (section J5.18)

Faraday's induction law prompts the question: if changing magnetic fields induce electric fields, is the converse also true? As we shall see in the next chapter , it is, and Ampère's Law will have to be modified accordingly. Except when we are dealing with electromagnetic waves, however, there will be many situations where this modification can be neglected and the magnetostatic equations still hold. In particular, whenever characteristic lengths are small compared with cT, with c the speed of light and T the characteristic time scale of the fields, this **quasi-static approximation** is valid. In conductors,

$$\boldsymbol{\nabla} \times \mathbf{H} \;=\; \mathbf{J}_{\text{free}} \;=\; \sigma \mathbf{E}$$

where σ is the conductivity of the material assumed to obey Ohm's Law. We note that even if there are no free sources of electrostatic $\mathbf{E}$, there are still free currents since these are the only source of the magnetic field. In homogeneous, linear magnetic media:

$$\boldsymbol{\nabla} \times \mathbf{B} \;=\; \mu\, \mathbf{J}_{\text{free}} \;=\; \mu\sigma \mathbf{E} \tag{4.43}$$

Differentiating with respect to time and using Faraday's Law; recalling that $\boldsymbol{\nabla} \cdot \mathbf{E} = 0$ when there are no free charges as sources of $\mathbf{E}$, we obtain a diffusion equation for the electric field:

$$\nabla^2 \mathbf{E} \;=\; \mu\sigma\, \partial_t \mathbf{E}$$

Writing eq. (4.43) in terms of the vector potential $\mathbf{A}$ shows that it obeys the same diffusion equation.

4.10 Magnetic Field Energy (section J5.16)

To find the magnetic energy of a current-field distribution, we must include the energy needed to set up the currents. In a magnetic material, we must also factor in the energy associated with the response of the atoms to the applied magnetic field. So we must ask how much work has to be done to change the *macroscopic* fields while incorporating the response of the material.

Start with some free current density $\mathbf{J}$ which generates an induction field $\mathbf{B}$. Change the current by an amount $\delta\mathbf{J}$, thereby producing a change $\delta\mathbf{B}$ in a time δt long enough that the quasi-static spproximation is still valid. By Faraday's Law, there will be an induced electric field $\mathbf{E}$ which, by Lenz's Law, opposes the change in the current, doing work on it at the the rate $\int \mathbf{J} \cdot \mathbf{E}\, \mathrm{d}^3x$. By conservation of energy, work to change the current density is done at the rate:

$$\delta W/\delta t \;=\; -\int \mathbf{J} \cdot \mathbf{E}\, \mathrm{d}^3x$$

Since $\mathbf{E} \;=\; -\delta\mathbf{A}/\delta t$, and $\mathbf{J}$ is a *free* current, we obtain:

$$\begin{aligned} \delta W = \int \delta\mathbf{A} \cdot \mathbf{J}\, \mathrm{d}^3x \;&=\; \int \delta\mathbf{A} \cdot \boldsymbol{\nabla} \times \mathbf{H}\, \mathrm{d}^3x \\ &= \int \mathbf{H} \cdot \boldsymbol{\nabla} \times \delta\mathbf{A}\, \mathrm{d}^3x \;+\; \int \boldsymbol{\nabla} \cdot (\mathbf{H} \times \delta\mathbf{A})\, \mathrm{d}^3x \end{aligned}$$

As usual, we turn the second integral into a flux integral vanishing at infinity. We are left with:

$$\delta W \;=\; \int_{\text{all space}} \mathbf{H} \cdot \delta\mathbf{B}\, \mathrm{d}^3x \tag{4.44}$$

This expression is completely general and applies to all types of magnetic media. It is the analog of eq. (3.38) in electrostatics.

If the material is linear, we can go further and write $\mathbf{H} \cdot \delta\mathbf{B} = \frac{1}{2}\delta(\mathbf{H} \cdot \mathbf{B})$. To bring the fields up from zero to their final values necessitates an energy:

$$W \;=\; \frac{1}{2} \int \mathbf{H} \cdot \mathbf{B}\, \mathrm{d}^3x \;=\; \frac{1}{2} \int \mathbf{J} \cdot \mathbf{A}\, \mathrm{d}^3x \tag{4.45}$$

In a non-linear material, we can use the hysteresis curve to find a relationship between $\mathbf{B}$ and $\mathbf{H}$.

By an argument entirely analogous to the one used in electrostatics, one finds that when a magnetic material is inserted in a magnetic field whose sources are fixed, the resulting change of energy is:

$$W \;=\; \frac{1}{2}\int \mathbf{M}\cdot\mathbf{B}\,\mathrm{d}^3x \tag{4.46}$$

The sign is different from the electrostatic result, due to the fact that, contrary to what happens in electrostatics, the fixed magnetic sources must also do work against the induced emf. We should also keep in mind that this is the total energy needed to produce the final configuration, not just to rotate the magnetic moments in an external field, ie. the interaction energy of eq. (4.20).

Example 4.6. Example of the use of the quasistatic approximation (section J5.18)

We consider a harmonic, uniform magnetic field $\mathbf{H} = (H_0\,\cos\omega t)\,\mathbf{i}$ defined below the xy plane, with the space above the xy plane entirely occupied by a conductor with uniform conductivity σ and permeability μ. What is the field inside the conductor?

First, matching at the interface shows that the field for $z>0$ is also in the x direction. Taking the curl of eq. (4.43) shows that $\mathbf{H}$ also obeys the same diffusion equation as $\mathbf{E}$ and $\mathbf{A}$:

$$\nabla^2\mathbf{H} \;=\; \mu\sigma\,\partial_t\mathbf{H}$$

Inserting the complexified inhomogoneous solution $H_x(z,t) \;=\; h(z)\,\mathrm{e}^{-\mathrm{i}\omega t}$ gives:

$$\left(\mathrm{d}_z^2 \;+\; \mathrm{i}\mu\sigma\omega\right)h(z) \;=\; 0$$

whose solution is $h(z) \;=\; A\,\mathrm{e}^{\mathrm{i}kz} \;+\; C\,\mathrm{e}^{-\mathrm{i}kz}$, where $k \;=\; (1+\mathrm{i})/\delta$, with the **skin depth** $\delta \;=\; \sqrt{2}/\sqrt{\mu\sigma\omega}$. Setting $C \;=\; 0$ to quell a divergence at infinity, and $A \;=\; H_0$ for correct matching at $z=0$, we finally get:

$$H_x(z,t) \;=\; \Re\left[H_0\,\mathrm{e}^{-z/\delta}\,\mathrm{e}^{\mathrm{i}(z/\delta-\omega t)}\right]$$

or

$$\mathbf{H}(z,t) \;=\; \mathbf{i}\,H_0\,\mathrm{e}^{-z/\delta}\,\cos(z/\delta-\omega t) \tag{4.47}$$

The time-variation of $\mathbf{B}$ generates an electric field inside the conductor; it is most easily found from Ampère's Law, eq. (4.43), using the complex exponential form of $\mathbf{H}$:

$$\begin{aligned}\mathbf{E} \;=\; \frac{1}{\sigma}\,\mathrm{d}_z\mathbf{H}\,(\mathbf{k}\times\mathbf{i}) \;&=\; \mathbf{j}\,\Re\left[\frac{-1+\mathrm{i}}{\sigma\delta}\,H_0\,\mathrm{e}^{-z/\delta}\,\mathrm{e}^{\mathrm{i}(z/\delta-\omega t)}\right]\\ &=\; \mathbf{j}\,\frac{\mu\omega\delta}{\sqrt{2}}\,H_0\,\mathrm{e}^{-z/\delta}\,\cos(z/\delta-\omega t+3\pi/4)\end{aligned} \tag{4.48}$$

We should check whether this electric field is consistent with the quasi-static approximation which we have been using. The relevant dimensionless ratio to take is:

$$\frac{E/c}{\mu\,H} \;=\; \mathrm{O}(\omega\delta/c)$$

For the quasi-static approximation to be valid, this should be much smaller than 1. For copper, a good conductor, $\omega\delta/c = 1.2\times10^{-8}\,\sqrt{\nu}$ with ν the field frequency. For sea-water, the ratio is $3.7\times10^{-5}\,\sqrt{\nu}$. This puts an upper limit on the frequency of the field that will allow the above treatment.

The time-averaged power dissipated per unit volume, $\langle\mathbf{J}\cdot\mathbf{E}\rangle = \sigma\,\langle E^2\rangle$, is:

$$P \;=\; \frac{1}{2}\mu\omega H_0^2\,\mathrm{e}^{-2z/\delta} \;=\; \frac{\sigma}{\delta^2}\,H_0^2\,\mathrm{e}^{-2z/\delta} \tag{4.49}$$

5 Maxwell's Equations and Conservation laws

5.1 First and Second Order Maxwell Field Equations (section J6.1)

When we allow electric and magnetic fields to vary in time and do not require that $\partial_t \rho = 0$, Ampère's Law, $\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{J}$, is no longer compatible with the continuity equation. This inconsistency led Maxwell to propose instead:

$$\boldsymbol{\nabla} \times \mathbf{H} \;=\; \mathbf{J}_{\text{free}} \;+\; \partial_t \mathbf{D} \tag{5.1}$$

The new term, which he called the **displacement current**, restores consistency via Gauss' Law $\boldsymbol{\nabla} \cdot \mathbf{D} = \rho$, and makes it clear that changing electric fields are also sources of the magnetic field. The two equivalent full sets:

$$\begin{array}{llll} \boldsymbol{\nabla} \cdot \mathbf{D} \;=\; \rho_{\text{free}} & \boldsymbol{\nabla} \times \mathbf{H} \;=\; \mathbf{J}_{\text{free}} \;+\; \partial_t \mathbf{D} & \boldsymbol{\nabla} \cdot \mathbf{E} \;=\; 4\pi k_{\text{e}} \rho_{\text{total}} & \boldsymbol{\nabla} \times \mathbf{B} \;=\; 4\pi k_{\text{m}}\, \mathbf{J}_{\text{total}} \;+\; \dfrac{1}{c^2} \partial_t \mathbf{E} \\ \boldsymbol{\nabla} \cdot \mathbf{B} \;=\; 0 & \boldsymbol{\nabla} \times \mathbf{E} \;+\; \partial_t \mathbf{B} \;=\; 0 & \boldsymbol{\nabla} \cdot \mathbf{B} \;=\; 0 & \boldsymbol{\nabla} \times \mathbf{E} \;+\; \partial_t \mathbf{B} \;=\; 0 \end{array} \tag{5.2}$$

constitute **Maxwell's macroscopic (on the left) and microscopic (on the right) field equations**. With the relations: $\mathbf{D} = 4\pi k_{\text{e}} \mathbf{E} + \mathbf{P}$ and $\mathbf{H} = \mathbf{B}/4\pi k_{\text{m}} - \mathbf{M}$, they provide a complete classical description of electromagnetic phenomena and do contain the continuity equation for the sources, whether free or total (including polarisation $\mathbf{P}$ and magnetisation $\mathbf{M}$):

$$\partial_t \rho_{\text{free}} \;=\; \boldsymbol{\nabla} \cdot\, \partial_t \mathbf{D} \;=\; -\,\boldsymbol{\nabla} \cdot \mathbf{J}_{\text{free}} \qquad\qquad \partial_t \rho_{\text{total}} \;=\; \frac{1}{4\pi k_{\text{e}}}\, \boldsymbol{\nabla} \cdot\, \partial_t \mathbf{E} \;=\; -\,\boldsymbol{\nabla} \cdot \mathbf{J}_{\text{total}} \tag{5.3}$$

By taking the curl of Faraday's Law and the curl of Ampère's Law in the microscopic version, we arrive (EXERCISE) at the **wave equations**:

$$\begin{aligned} \Box\, \mathbf{E} \;&=\; -\,4\pi k_{\text{e}} \left(\boldsymbol{\nabla} \rho_{\text{total}} \;+\; \frac{1}{c^2} \partial_t \mathbf{J}_{\text{total}} \right) \\ \Box\, \mathbf{B} \;&=\; 4\pi k_{\text{m}}\, \nabla \times \mathbf{J}_{\text{total}} \end{aligned} \tag{5.4}$$

where $\Box \;=\; (1/c^2)\partial_t^2 \,-\, \nabla^2$ is the **d'Alembertian** operator (Jackson's definition is consistent with his metric convention in chapter 11), and $c = \sqrt{k_{\text{e}}/k_{\text{m}}}$ in any units ($c = 1/\sqrt{\epsilon_0 \mu_0}$ in SI units). Equivalent equations for $\mathbf{D}$ and $\mathbf{H}$, with source terms now involving the polarisation and magnetisation, may be obtained from the macroscopic first-order equations.

So called *free fields* obey the source-free first-order equations:

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{E} \;&=\; 0 & \qquad \boldsymbol{\nabla} \times \mathbf{B} \;-\; \frac{1}{c^2}\, \partial_t \mathbf{E} \;&=\; 0 \\ \boldsymbol{\nabla} \cdot \mathbf{B} \;&=\; 0 & \qquad \boldsymbol{\nabla} \times \mathbf{E} \;+\; \partial_t \mathbf{B} \;&=\; 0 \end{aligned} \tag{5.5}$$

and the second-order equations:

$$\Box\, \mathbf{E} \;=\; 0 \qquad\qquad \Box\, \mathbf{B} \;=\; 0 \tag{5.6}$$

Such free fields have field lines that do not end on electric charges ($\mathbf{E}$) and do not enclose currents ($\mathbf{B}$). It is important always to remember that these second-order equations hold for each component of the field only in Cartesian coordinates. If one did insist on using curvilinear coordinates, the action of the Laplacian operator on the vector would have to be defined using the identity: $\nabla^2 \mathbf{E} = \nabla(\boldsymbol{\nabla} \cdot \mathbf{E}) - \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}$, and *only then* would each *curvilinear* component of this object enter in the d'Alembertian to form a scalar wave equation.

5.2 Electromagnetic Potentials (section J6.2)

The two homogeneous equations 5.2 are still equivalent to:

$$\mathbf{E} \;=\; -\,\partial_t \mathbf{A} \;-\, \boldsymbol{\nabla}\Phi \qquad\qquad \mathbf{B} \;=\; \boldsymbol{\nabla}\times\mathbf{A} \tag{5.7}$$

In unpolarised, unmagnetised media, the inhomogeneous Maxwell equations then become (exercise) *second-order* equations for the potentials:

$$\begin{aligned} \nabla^2\Phi \;+\; \partial_t(\boldsymbol{\nabla}\cdot\mathbf{A}) \;&=\; -4\pi\, k_{\mathrm{e}}\rho \\ \Box\,\mathbf{A} \;+\; \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\mathbf{A} \;+\; \frac{1}{c^2}\partial_t\Phi\right) \;&=\; 4\pi\, k_{\mathrm{m}}\,\mathbf{J} \end{aligned} \tag{5.8}$$

These two equations constitute Maxwell's equations for the scalar and vector potentials with *free* sources.

5.2.1 Transverse and longitudinal projections of Maxwell's equations for the potentials

When we try to solve for the vector potential $\mathbf{A}$, we expect Maxwell's theory to provide us with equations, either first-order or second-order, for $\boldsymbol{\nabla}\times\mathbf{A}$ and $\boldsymbol{\nabla}\cdot\mathbf{A}$. It certainly does for $\boldsymbol{\nabla}\times\mathbf{A}$. However, neither of Maxwell's equations (5.8) for the potentials Φ and $\mathbf{A}$ can determine $\boldsymbol{\nabla}\cdot\mathbf{A}$. A very instructive way to see this is to note that, according to Helmholtz's theorem (0.19), any 3-dim vector field $\mathbf{A}$ may be written as the sum of two vectors:

$$\mathbf{A} \;=\; \mathbf{A}_{\mathrm{L}} \;+\; \mathbf{A}_{\mathrm{T}} \;=\; \boldsymbol{\nabla}u \;+\; \boldsymbol{\nabla}\times\mathbf{w} \tag{5.9}$$

The first term, $\mathbf{A}_{\mathrm{L}} \coloneqq \boldsymbol{\nabla}u$, whose curl vanishes *identically*, is called the **longitudinal** (irrotational) part, or projection, of $\mathbf{A}$; the second, $\mathbf{A}_{\mathrm{T}} \coloneqq \boldsymbol{\nabla}\times\mathbf{w}$, whose divergence vanishes *identically*, is called the **transverse** (solenoidal) part of $\mathbf{A}$. Each of the three components of $\mathbf{A}$ potentially carries energy as a wave, all the way to infinity, in which case we say that it is a **dynamical degree of freedom**. By inspection, $\mathbf{A}_{\mathrm{L}}$ has one independent component, whereas $\mathbf{A}_{\mathrm{T}}$ has two independent components (the third is determined by $\boldsymbol{\nabla}\cdot\mathbf{A}_{\mathrm{T}} = 0$).

Observe that $\boldsymbol{\nabla}\cdot\mathbf{A}$ is really $\boldsymbol{\nabla}\cdot\mathbf{A}_{\mathrm{L}}$ and contains no information about $\mathbf{A}_{\mathrm{T}}$; also, $\boldsymbol{\nabla}\times\mathbf{A}$ is really $\boldsymbol{\nabla}\times\mathbf{A}_{\mathrm{T}}$ and contains no information about $\mathbf{A}_{\mathrm{L}}$.

Project the second equation (5.8). The transverse projection immediately gives:

$$\Box\,\mathbf{A}_{\mathrm{T}} \;=\; 4\pi\, k_{\mathrm{m}}\mathbf{J}_{\mathrm{T}} \tag{5.10}$$

where we have used the fact that a gradient is a longitudinal object.

Now take the divergence of the longitudinal projection of the second equation (5.8):

$$\begin{aligned} \boldsymbol{\nabla}\cdot\left[\Box\,\mathbf{A}_{\mathrm{L}} \;+\; \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\mathbf{A}_{\mathrm{L}} \;+\; \frac{1}{c^2}\partial_t\Phi\right) \;-\; 4\pi\, k_{\mathrm{m}}\,\mathbf{J}\right] \;&=\; \Box(\boldsymbol{\nabla}\cdot\mathbf{A}_{\mathrm{L}}) \;+\; \nabla^2(\boldsymbol{\nabla}\cdot\mathbf{A}_{\mathrm{L}}) \;+\; \frac{\partial_t\nabla^2\Phi}{c^2} \;-\; 4\pi\, k_{\mathrm{m}}\boldsymbol{\nabla}\cdot\mathbf{J} \\ &=\; \frac{1}{c^2}\,\partial_t\left[\partial_t(\boldsymbol{\nabla}\cdot\mathbf{A}_{\mathrm{L}}) \;+\; \nabla^2\Phi \;+\; 4\pi\, k_{\mathrm{e}}\rho\right] \end{aligned}$$

But the terms in the square bracket on the second line are just the first of equations (5.8). Therefore, the longitudinal projection of the second equation Maxwell equation for the 3-vector potential contains no information about $\boldsymbol{\nabla}\cdot\mathbf{A}$ that is not in the first equation. But that is really an equation for Φ with $\boldsymbol{\nabla}\cdot\partial_t\mathbf{A}$ (more precisely, $\boldsymbol{\nabla}\cdot\partial_t\mathbf{A}_{\mathrm{L}}$) as a source together with ρ. Therefore, Maxwell's theory cannot determine the divergence of the 3-vector potential.

5.2.2 Choices of the divergence of A

Since the theory does not know $\boldsymbol{\nabla}\cdot\mathbf{A}$, *we* have to tell it what it is by making an arbitrary choice. If we choose $\boldsymbol{\nabla}\cdot\mathbf{A}$ to vanish (**Coulomb condition**), the vector potential is pure transverse, and the Φ equation becomes a Poisson-type equation with solution:

$$\Phi(\mathbf{x},t) \;=\; k_{\rm e}\int\frac{\rho(\mathbf{x}',t)}{|\mathbf{x}-\mathbf{x}'|}\,{\rm d}^3x' \tag{5.11}$$

This looks innocuous enough until we realise that any change in the source is *instantaneously* reflected in the scalar potential. The Coulomb condition leads to acausal behaviour, which is also a reflection of the fact that the condition is not relativistically covariant, in the sense that it is not necessarily the same in all inertial frames. But the equation for Φ is not a classical wave equation, and Φ does not really propagate as a wave, so one should not expect proper causal behaviour from it. As for $\mathbf{A}$, it now obeys a wave equation equivalent (EXERCISE) to eq. (5.10).

When manifestly causal behaviour for the potentials is desired, we can choose instead the Lorenz[†] condition: $\boldsymbol{\nabla}\cdot\mathbf{A} = -\partial_t\Phi/c^2$, which also happens to turn eq. (5.8) into nice-looking, decoupled wave equations for each component of $\mathbf{A}$, of the type $\Box(\text{potential}) = $ source with causal solution eq. (5.18). Then one can calculate the energy radiated to infinity following standard treatments (eg. chapter 14 in Jackson) and find that the scalar potential does make a *mathematical* contribution to the energy radiated to infinity by accelerated charges. But we should not attach any physical significance to that contribution: it arises simply out of consistency with the choice of the Lorenz condition for $\boldsymbol{\nabla}\cdot\mathbf{A}$.

Note also that, under the Lorenz condition, the energy radiated to infinity by an oscillating system can be calculated (see chapter 7 here) solely in terms of $\mathbf{A}$, since Φ is determined.

5.3 Initial Value Problem in Maxwell's Theory: First-order Cauchy problem

The Initial Value Problem (IVP) of a theory consists in finding which data must be specified at a given time for the time evolution of variables of the theory to be uniquely determined by their equations of "motion". By **initial data**, one means the state of the system of variables everywhere in space at $t = t_0$. The IVP together with the evolution equations constitute the **Cauchy Problem** of the theory. If it can be solved, the dynamical behaviour of the system can be uniquely predicted from its initial data.

Most often, the equations of "motion" take the form of a set of wave equations, each of the form $\Box f = F$. If they always told the whole story, the Cauchy problem would be solved by specifying the value of f and its first-order time derivatives at $t = t_0$. Things are not so simple, however when there are inherent, built-in **constraints** on the initial data. Those constraint equations must be discovered and solved. Also, we must find which initial data we are allowed to specify freely.

Eq. (5.4) do look like standard wave equations for six decoupled quantities. But the *first-order*, coupled, field equations (5.2) must also be satisfied!

The two divergence equations contain no time derivatives and are thus constraints on $\mathbf{E}$ and $\mathbf{B}$ at all time, including $t = t_0$. The one involving $\mathbf{E}$ ($\mathbf{E}_{\rm L}$ in effect) can be rewritten $\nabla^2 u = \rho$ for a scalar field u, a Poisson-type equation which can in principle be solved for u *at initial time* so long as ρ falls off faster than $1/r$ at infinity. As for $\mathbf{B}$, because its divergence vanishes, it is always a manifestly transverse object. Thus, we have no control over the longitudinal components of the fields; the two transverse components for each are the only ones for which initial data can be freely specified at this stage.

Next, look at the two Maxwell first-order equations for $\mathbf{E}$ and $\mathbf{B}$ which contain time derivatives. Suppose we specify $\mathbf{E}$ and $\partial_t\mathbf{E}$ at $t = t_0$, which are needed to solve the $2^{\rm nd}$-order equation for $\mathbf{E}$ in eq. (5.4). Then the two

[†]L. Lorenz, *On the Identity of the Vibrations of Light with Electrical Currents*, Philosophical Magazine and Journal of Science, **34**, July-December, 1867, pp. 287–301 (translated from Annalen der Physik und Chemie, June 1867)

transverse components of $\mathbf{B}$ are determined by the generalised Ampère Law; $\partial_t \mathbf{B}$ is determined by Faraday's Law, also at $t = t_0$. Once we have specified the two transverse components of $\mathbf{E}$ and their time derivatives, Maxwell's equations take over and determine the others at $t = t_0$. Alternatively, we could have started with the two transverse components of $\mathbf{B}$; specifying them and their time derivatives at $t = t_0$ constrains all the other field components and time derivatives.

You can also use the transverse/longitudinal projections of the first-order equations (5.2) to show (EXERCISE) that in source-free space, only the transverse components of $\mathbf{E}$ and $\mathbf{B}$ obey a classical wave equation.

One of the advantages of this Cauchy analysis is that it does not rely on any particular solution, but is valid for *any* electromagnetic field. Later we shall reformulate the theory in the much more powerful and compact language of the four-vector relativistic formalism which will allow for much easier manipulation. We note that since $\boldsymbol{\nabla} \cdot \mathbf{A}$, which determines the longitudinal component of $\mathbf{A}$, is arbitrary, only the two transverse components controlled by $\boldsymbol{\nabla} \times \mathbf{A}$ correspond to physical *dynamical* degrees of freedom, consistent with the Cauchy-data analysis of the first-order Maxwell equations (although the latter could not tell us which initial field data could be freely specified, only their number).

5.4 Green Functions for the d'Alembertian Operator (section J6.4)

With the Fourier integral representation (note the normalisation!):

$$g(\mathbf{x},t) \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} g(\mathbf{x},\omega)\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}\omega \qquad\qquad g(\mathbf{x},\omega) \;=\; \int_{-\infty}^{\infty} g(\mathbf{x},t)\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t \tag{5.12}$$

we can transform a typical wave equation, with $\Box = (1/c^2)\partial_t^2 - \nabla^2$:

$$\Box\Psi(\mathbf{x},t) \;=\; 4\pi\, f(\mathbf{x},t)$$

where $f(\mathbf{x},t)$ is a known source, to its so-called Helmholtz form for a frequency-dependent function $\Psi(\mathbf{x},\omega)$:

$$(\nabla^2 \,+\, k^2)\Psi(\mathbf{x},\omega) \;=\; -\,4\pi\, f(\mathbf{x},\omega) \tag{5.13}$$

where k can be taken as a short form for ω/c. In a non-dispersive medium, $\omega(k) = kc$ is actually the dispersion relation, with **k** the **wave vector** and k the wave number.

Just as for the Laplacian operator, there exist Green functions for $\nabla^2 + k^2$; they satisfy:

$$(\nabla^2 \,+\, k^2)\, G(\mathbf{x},\mathbf{x}') \;=\; -4\pi\,\delta(\mathbf{x}-\mathbf{x}') \tag{5.14}$$

Jackson does not really explain why one must take $G(\mathbf{x},\mathbf{x}') = G(\mathbf{x}-\mathbf{x}')$; like the Laplacian, the k^2 dependence on the wave vector picks no preferred direction, so the solutions have spherical symmetry with the centre of symmetry at the source point $\mathbf{x}'$. Therefore $G(\mathbf{x}-\mathbf{x}') = G(R)$, with $R = |\mathbf{x}-\mathbf{x}'|$.

A solution of the inhomogeneous equation (5.14) for $G(R)$ is:

$$G^{(\pm)}(R) \;=\; \frac{1}{R}\,\mathrm{e}^{\pm\mathrm{i}kR} \tag{5.15}$$

Indeed:

$$\begin{aligned}(\nabla^2 + k^2)\left(\frac{\mathrm{e}^{\pm\mathrm{i}kR}}{R}\right) \;&=\; \frac{1}{R}\nabla^2\mathrm{e}^{\pm\mathrm{i}kR} \,+\, \mathrm{e}^{\pm\mathrm{i}kR}\,\nabla^2\left(\frac{1}{R}\right) \,+\, 2\,\boldsymbol{\nabla}\left(\frac{1}{R}\right)\cdot\boldsymbol{\nabla}\mathrm{e}^{\pm\mathrm{i}kR} \,+\, k^2\,\frac{\mathrm{e}^{\pm\mathrm{i}kR}}{R} \\ &=\; \left[-\frac{k^2}{R} \,\pm\, \frac{2\mathrm{i}k}{R^2} \,-\, 4\pi\,\delta(\mathbf{x}-\mathbf{x}') \,\mp\, \frac{2\mathrm{i}k}{R^2} \,+\, \frac{k^2}{R}\right]\mathrm{e}^{\pm\mathrm{i}kR} \;=\; -\,4\pi\,\delta(\mathbf{x}-\mathbf{x}')\end{aligned}$$

Now we are ready to find the full Green functions for the d'Alembertian operator, which satisfy:

$$\Box_x G(\mathbf{x},t;\mathbf{x}',t') \;=\; 4\pi\,\delta(\mathbf{x}-\mathbf{x}')\,\delta(t-t') \tag{5.16}$$

The Fourier transform of the source term, $f(\mathbf{x},t) = 4\pi\,\delta(\mathbf{x}-\mathbf{x}')\,\delta(t-t')$, is $f(\mathbf{x},\omega) = 4\pi\,\delta(\mathbf{x}-\mathbf{x}')\,\mathrm{e}^{\mathrm{i}\omega t'}$. In the frequency domain, then, we have:

$$(\nabla_x^2 \,+\, k^2)G(\mathbf{x},\mathbf{x}',\omega,t') \;=\; -\,4\pi\,\delta(\mathbf{x}-\mathbf{x}')\,\mathrm{e}^{\mathrm{i}\omega t'}$$

Assume separable solutions of the form $G(\mathbf{x},\mathbf{x}')\mathrm{e}^{\mathrm{i}\omega t'}$; inserting into this equation, we get from (5.14) the solutions $G^{\pm}(\mathbf{x},\mathbf{x}',\omega,t') = \mathrm{e}^{\mathrm{i}(\pm kR+\omega t')}/R$. Then, transforming back to the time domain and using the representation (0.16) for the δ-function yields the time-dependent Green functions in a nondispersive medium ($k=\omega/c$):

$$\begin{aligned} G^{(\pm)}(\mathbf{x},t;\mathbf{x}',t') \;&=\; \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\mathrm{e}^{\mathrm{i}\omega[\pm R/c-(t-t')]}}{R}\,\mathrm{d}\omega \\ &=\; \frac{\delta\big(t' \,-\, [t\mp R/c]\big)}{R} \end{aligned} \tag{5.17}$$

Using eq. (5.16), we recognise that:

$$\Box_x\int_{-\infty}^{\infty} G^{(\pm)}(\mathbf{x},t;\mathbf{x}',t')\,f(\mathbf{x}',t')\,\mathrm{d}^3x'\,\mathrm{d}t' \;=\; \int_{-\infty}^{\infty} f(\mathbf{x}',t')\,\Box_x G^{(\pm)}(\mathbf{x},t;\mathbf{x}',t')\,\mathrm{d}^3x'\,\mathrm{d}t' \;=\; 4\pi\,f(\mathbf{x},t)$$

has the generic form $\Box\,\Psi(\mathbf{x},t) = 4\pi f(\mathbf{x},t)$, which shows that the general solution of a wave equation with sources *localised in time and space* can be written either as the **retarded** solution:

$$\begin{aligned} \Psi(\mathbf{x},t) \;&=\; \Psi_{\text{in}}(\mathbf{x},t) \;+\; \iint_{-\infty}^{\infty} G^{(+)}(\mathbf{x},t;\mathbf{x}',t')\,f(\mathbf{x}',t')\,\mathrm{d}^3x'\,\mathrm{d}t' \\ &=\; \Psi_{\text{in}}(\mathbf{x},t) \;+\; \int\frac{f(\mathbf{x}',t'_{\text{ret}})}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x' \end{aligned} \tag{5.18}$$

or, equivalently, as the **advanced** solution:

$$\begin{aligned} \Psi(\mathbf{x},t) \;&=\; \Psi_{\text{out}}(\mathbf{x},t) \;+\; \iint_{-\infty}^{\infty} G^{(-)}(\mathbf{x},t;\mathbf{x}',t')\,f(\mathbf{x}',t')\,\mathrm{d}^3x'\,\mathrm{d}t' \\ &=\; \Psi_{\text{out}}(\mathbf{x},t) \;+\; \int\frac{f(\mathbf{x}',t'_{\text{adv}})}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x' \end{aligned} \tag{5.19}$$

where the suffixes *ret* and *adv* stand for the fact that t' must be evaluated at the **retarded time** $t'_{\text{ret}} = t - R/c$, or the **advanced time** $t'_{\text{adv}} = t + R/c$. This ensures the proper causal behaviour of the solutions, in the sense that, eg., the solution at time t is only influenced by the behaviour of the source point $\mathbf{x}'$ at time $t - R/c$. Ψ_{in} and Ψ_{out} are possible plane-wave solutions of the *homogeneous* wave equation for Ψ. Most often they can be taken to be zero. Note, however, that this breaks the time-reversal symmetry of theories such as Maxwell's electromagnetism.

5.5 Retarded Solutions for the Potentials and Fields of a Localised Source (section J6.5)

Nothing prevents the source $f(\mathbf{x}',t'_{\text{ret}})$ in eq. (5.18) from having terms that contain Ψ. Thus, the inhomogeneous causal solution to Maxwell's equation (5.8) for the potential $\mathbf{A}$ is:

$$\mathbf{A}(\mathbf{x},t) \;=\; \int\frac{-\boldsymbol{\nabla}'(\boldsymbol{\nabla}'\cdot\mathbf{A}(\mathbf{x}',t') \,+\, \partial_{t'}\Phi(\mathbf{x}',t')/c^2)/4\pi\big|_{\text{ret}} \,+\, k_{\text{m}}\,\mathbf{J}(\mathbf{x}',t'_{\text{ret}})}{R}\,\mathrm{d}^3x'$$

One advantage of this result (and, therefore, of the Green-function method) is that whatever condition we impose on $\boldsymbol{\nabla}\cdot\mathbf{A}$ will *automatically* be satisfied by the solution. If we impose the Lorenz condition, the electromagnetic potentials now obey wave equations whose retarded solutions are:

$$\begin{aligned} \Phi(\mathbf{x},t) \;&=\; k_{\text{e}}\int\mathrm{d}^3x'\,\frac{\rho(\mathbf{x}',t'_{\text{ret}})}{R} \\ \mathbf{A}(\mathbf{x},t) \;&=\; k_{\text{m}}\int\mathrm{d}^3x'\,\frac{\mathbf{J}(\mathbf{x}',t'_{\text{ret}})}{R} \end{aligned} \tag{5.20}$$

where $R = |\mathbf{x} - \mathbf{x}'_{\rm ret}|$. For the potentials, and *only* for the potentials, the solutions have the same dependence as the static potentials, except that the sources must be evaluated at retarded time before integrating. This similarity does not hold for the fields, however! In media without polarisation or magnetisation, eq. (5.4) becomes:

$$\begin{aligned} \Box\mathbf{E} &= -4\pi k_{\rm e}\left(\boldsymbol{\nabla}\rho + \frac{1}{c^2}\partial_t\mathbf{J}\right) \\ \Box\mathbf{B} &= 4\pi k_{\rm m}\,\boldsymbol{\nabla}\times\mathbf{J} \end{aligned} \tag{5.21}$$

where the sources are now free. Again, according to the results of the last section:

$$\begin{aligned} \mathbf{E}(\mathbf{x},t) &= -k_{\rm e}\int {\rm d}^3x'\,\frac{\left[\boldsymbol{\nabla}'\rho(\mathbf{x}',t') + \frac{1}{c^2}\partial_{t'}\mathbf{J}\right]_{\rm ret}}{R} \\ \mathbf{B}(\mathbf{x},t) &= k_{\rm m}\int {\rm d}^3x'\,\frac{\left[\boldsymbol{\nabla}'\times\mathbf{J}(\mathbf{x}',t')\right]_{\rm ret}}{R} \end{aligned} \tag{5.22}$$

We would like to transfer the spatial derivatives to the $1/R$ factor with an integration by parts, but to do this we must have derivatives of functions which have *first* been evaluated at $t' = t - R/c$, keeping t' fixed . In the integrands, however, we must also take into account the dependence of R on $\mathbf{x}'$ when differentiating. From the chain rule, we have:

$$\boldsymbol{\nabla}' f(\mathbf{x}', t - R/c) = \left[\boldsymbol{\nabla}' f(\mathbf{x}',t')\right]_{t'=t-R/c} + \partial_{t'} f\Big|_{t'=t-R/c}\boldsymbol{\nabla}'(t - R/c)$$

that is, $\left[\boldsymbol{\nabla}' f(\mathbf{x}',t')\right]_{\rm ret} = \boldsymbol{\nabla}' f(\mathbf{x}', t - R/c) - \dfrac{\hat{\mathbf{R}}}{c}\,\partial_{t'} f\Big|_{\rm ret}$, where we have used $\nabla' R = -(\mathbf{x}-\mathbf{x}')/|(\mathbf{x}-\mathbf{x}'| = -\hat{\mathbf{R}}$.

A similar argument for the curl of some vector $\mathbf{C}$ leads to:

$$\left[\boldsymbol{\nabla}'\times\mathbf{C}(\mathbf{x}',t')\right]_{\rm ret} = \boldsymbol{\nabla}'\times\mathbf{C}(\mathbf{x}', t - R/c) + \frac{\hat{\mathbf{R}}}{c}\times\partial_{t'}\mathbf{C}\Big|_{\rm ret}$$

Now we can insert these expressions into the field solutions and integrate by parts the term with the spatial derivative to get:

$$\begin{aligned} \mathbf{E}(\mathbf{x},t) &= k_{\rm e}\int {\rm d}^3x'\,\left[\rho(\mathbf{x}',t_{\rm ret})\,\frac{\mathbf{R}}{R^3} + \frac{\mathbf{R}}{R^2}\,\partial_{ct'}\rho\Big|_{\rm ret} - \frac{1}{cR}\,\partial_{ct'}\mathbf{J}\Big|_{\rm ret}\right] \\ \mathbf{B}(\mathbf{x},t) &= k_{\rm m}\int {\rm d}^3x'\,\left[\mathbf{J}(\mathbf{x}',t_{\rm ret})\times\frac{\mathbf{R}}{R^3} + \partial_{ct'}\mathbf{J}\Big|_{\rm ret}\times\frac{\mathbf{R}}{R^2}\right] \end{aligned} \tag{5.23}$$

The first term in each expression is clearly the straightforward generalisation of the static field, evaluated at retarded time; but note the extra terms, all with time derivatives.

These **Jefimenko** solutions are not often used because it is normally easier to find the retarded potentials first and then obtain the fields via eq. (5.7). But they provide some insight on how the time-dependent solutions go to their static limit. And by expanding $\mathbf{J}$ in a Taylor series around t and, assuming that it varies slowly enough, keeping only first-order terms in $t_{\rm ret} - t$, one can show that we recover the Biot-Savart Law, $\mathbf{B} = k_{\rm m}\int {\rm d}^3x'\mathbf{J}(\mathbf{x},t)\times\hat{\mathbf{R}}/R^3$ which is thus seen to have a wider domain of validity than one might have supposed.

5.6 The Hertz Superpotential Formulation of Maxwell's Theory (section J6.13)

In eq. (5.2) we wrote a set of Maxwell field equations in terms of the $\mathbf{E}$ and $\mathbf{B}$ fields:

$$\begin{aligned} \boldsymbol{\nabla}\cdot\mathbf{E} &= 4\pi k_{\rm e}\,\rho_{\rm tot} & \qquad \boldsymbol{\nabla}\times\mathbf{B} &= 4\pi k_{\rm m}\,\mathbf{J}_{\rm tot} + \frac{1}{c^2}\partial_t\mathbf{E} \\ \boldsymbol{\nabla}\cdot\mathbf{B} &= 0 & \qquad \boldsymbol{\nabla}\times\mathbf{E} + \partial_t\mathbf{B} &= 0 \end{aligned} \tag{5.24}$$

where the *total* sources are now involved, not just the free ones. These so-called **microscopic** equations have general validity and are the ones to be used to find the fields of a localised source. Under the Lorenz condition:

$$\boldsymbol{\nabla}\cdot\mathbf{A} + \frac{1}{c^2}\partial_t\Phi = 0$$

they are equivalent to the second-order equations for the scalar (Φ) and vector ($\mathbf{A}$) potentials:

$$\Box\Phi = 4\pi k_{\rm e}\,\rho_{\rm tot} \qquad \Box\mathbf{A} = 4\pi k_{\rm m}\,\mathbf{J}_{\rm tot} \tag{5.25}$$

But these equations do not "know" how they have been derived, and (unless they are solved by the method of sections 5.4 and 5.5, we must check that any of their solutions satisfy the Lorenz condition. While this is not hard to do for plane-wave solutions, it can get messy in more general cases.

In a 1889 study of the fields radiated by an oscillating dipole, Hertz was led to introduce a vector **superpotential** (often referred to nowadays as a polarisation potential or a Hertz vector) which lately has received some attention as a useful tool. The idea is to express the potentials as derivatives of this superpotential, which always exists as will easily be shown later (in chapter 11) with the four-vector formalism. Here, we will just review a non-relativistic version.

In his very cursory treatment of the polarisation potential, Jackson (J6.13) uses *macroscopic* potential equations in polarised and magnetised matter, with no free sources but with external sources of polarisation and magnetisation. While there is nothing wrong with this, it fails to exhibit the full power of the superpotential. Instead we follow (with some alterations!) a paper by J. J. Sein[†]. This paper also cites many useful references.

The Hertz superpotential $\boldsymbol{\Pi}$ is defined so as to satisfy: $\Phi = -\boldsymbol{\nabla}\cdot\boldsymbol{\Pi}$. To find $\mathbf{A}$ in terms of $\boldsymbol{\Pi}$, we take advantage of our choice of the Lorenz condition. In terms of $\mathbf{A}$ and $\boldsymbol{\Pi}$, this is:

$$\boldsymbol{\nabla}\cdot\left(\mathbf{A} - \frac{1}{c^2}\,\partial_t\boldsymbol{\Pi}\right) = 0$$

whose solution is $\mathbf{A} = (1/c^2)\partial_t\boldsymbol{\Pi} + \boldsymbol{\nabla}\times\mathbf{V}$, with $\mathbf{V}$ an arbitrary vector field. Then eq. (5.25) for the potentials are readily converted to the following equations for $\boldsymbol{\Pi}$:

$$\boldsymbol{\nabla}\cdot\Box\boldsymbol{\Pi} = -4\pi k_{\rm e}\,\rho_{\rm tot} \qquad \frac{1}{c^2}\,\partial_t\Box\boldsymbol{\Pi} + \boldsymbol{\nabla}\times\Box\mathbf{V} = 4\pi k_{\rm m}\,\mathbf{J}_{\rm tot}$$

Now apply the following transformation to the superpotential: $\delta\,\boldsymbol{\Pi} = \boldsymbol{\nabla}\times\mathbf{G}$, where G is arbitrary. The transformation induced on the potentials leaves (EXERCISE) the Lorenz condition, as well as the divergence equation, invariant. Then we simply choose $\mathbf{G}$ so as to cancel the $\mathbf{V}$ term in the second equation. This is equivalent to setting $\mathbf{V} = 0$, and we are left with:

$$\boldsymbol{\nabla}\cdot\Box\boldsymbol{\Pi} = -4\pi k_{\rm e}\,\rho_{\rm tot} \qquad \partial_t\,\Box\,\boldsymbol{\Pi} = 4\pi k_{\rm e}\,\mathbf{J}_{\rm tot} \tag{5.26}$$

As a check, taking the time derivative of the first equation and subtracting the divergence of the second, the continuity equation $\boldsymbol{\nabla}\cdot\mathbf{J}_{\rm tot} + \partial_t\rho_{\rm tot} = 0$ is seen to be satisfied, as expected.

[†] Am. J. Phys. **57**, 834(1989) [https://aapt.scitation.org/doi/10.1119/1.15905]

The superpotential approach is most useful for harmonic sources: $\rho_{\text{tot}}(\mathbf{x},t) = \rho_{\text{tot}}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}$ and $\mathbf{J}_{\text{tot}}(\mathbf{x},t) = \mathbf{J}_{\text{tot}}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}$. This in effect turns time derivatives into $-\mathrm{i}\omega = -\mathrm{i}ck$, where $k = \omega/c$. Then the continuity equation becomes: $\boldsymbol{\nabla}\cdot\mathbf{J}_{\text{tot}} - \mathrm{i}ck\rho_{\text{tot}} = 0$, and we can eliminate ρ_{tot} from our evolution equations (5.26) for $\boldsymbol{\Pi}$ to obtain:

$$\boldsymbol{\nabla}\cdot\left[\nabla^2\boldsymbol{\Pi} + k^2\boldsymbol{\Pi} - 4\pi\mathrm{i}\,\frac{k_{\text{e}}}{ck}\,\mathbf{J}_{\text{tot}}\right] = 0 \qquad\qquad \nabla^2\boldsymbol{\Pi} + k^2\boldsymbol{\Pi} - 4\pi\mathrm{i}\,\frac{k_{\text{e}}}{ck}\mathbf{J}_{\text{tot}} = 0$$

The first equation automatically holds since the second one is satisfied everywhere, and we arrive at a simple equation for $\boldsymbol{\Pi}$:

$$(\nabla^2 + k^2)\,\boldsymbol{\Pi} = 4\pi\mathrm{i}\,\frac{k_{\text{e}}}{ck}\,\mathbf{J}_{\text{tot}} \tag{5.27}$$

This is an inhomogeneous Helmholtz equation whose solution has been found in eq. (5.20), with $\mathbf{J}_{\text{tot}}(\mathbf{x},t'_{\text{ret}}) = \mathbf{J}_{\text{tot}}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t'} = \mathbf{J}_{\text{tot}}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{e}^{\mathrm{i}\omega R/c}$, and $R = |\mathbf{x}-\mathbf{x}'|$:

$$\boldsymbol{\Pi}(\mathbf{x}) = \mathrm{i}\,\frac{k_{\text{e}}}{ck}\int \mathbf{J}_{\text{tot}}(\mathbf{x}')\,\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\,\mathrm{d}^3x' \tag{5.28}$$

where the integration runs over all space. We recognise $\mathrm{e}^{\mathrm{i}kR}/R$ as the Green function for the Helmholtz operator. This equation may not look so different from the equivalent equation for the vector potential $\mathbf{A}$, but it is. For one thing, *any* of its solutions will be consistent with the Lorenz condition—no need to impose it anymore! Also, it holds in arbitrary media since these may be treated as a bunch of localised sources in vacuum. It all comes down to specifying $\boldsymbol{J}_{\text{tot}}$. The formalism also applies equally well to *macroscopic* variables, eg. in a conducting, magnetic dielectric where $\boldsymbol{J}_{\text{tot}} = \boldsymbol{J}_{\text{free}} - \mathrm{i}\omega\mathbf{P} + \boldsymbol{\nabla}\times\mathbf{M}$; $\mathbf{P}$ and $\mathbf{M}$ are the polarisation and magnetisation vectors.

Once the superpotential has been obtained, the harmonic potentials and the harmonic electric and magnetic fields can be found, using the fields' expressions in terms of the potentials:

$$\Phi = -\boldsymbol{\nabla}\cdot\boldsymbol{\Pi}, \qquad \mathbf{A} = -\mathrm{i}\,\frac{k}{c}\,\boldsymbol{\Pi}; \qquad\qquad \mathbf{E} = k^2\boldsymbol{\Pi} + \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{\Pi}), \qquad \mathbf{B} = -\mathrm{i}\,\frac{k}{c}\,\boldsymbol{\nabla}\times\boldsymbol{\Pi} \tag{5.29}$$

5.7 Energy and Momentum Conservation for a Field-Particle System (section J6.7)

5.7.1 Conservation of Energy

Let u be the energy density (electric and magnetic) associated with electromagnetic fields in some volume of a *linear*, nondispersive material, which we assume to be the same as for static fields (sections 3.5 and 4.10):

$$u_{\text{field}} = \frac{1}{2}\,(\mathbf{E}\cdot\mathbf{D} + \mathbf{H}\cdot\mathbf{B})$$

If there are no free charges within the volume, we can write a continuity equation:

$$\partial_t u_{\text{field}} + \boldsymbol{\nabla}\cdot\mathbf{S} = 0$$

where $\mathbf{S}$ is an energy current-density vector. As usual this expresses the fact that any change in the amount of energy inside the volume must correspond to a flux of energy into, or out of, the volume through its boundaries.

If there are free charges or currents in the volume, however, we expect the electric field to do work on them (the magnetic field does no work) and energy to be transferred between the fields and the charges. To take this into account, we should write a continuity equation of the form:

$$\partial_t u + \boldsymbol{\nabla}\cdot\mathbf{S} = 0$$

where now $u = u_{\text{field}} + u_{\text{mech}}$, with u_{mech} the mechanical energy density of the charge configuration.

To find $\mathbf{S}$, assume that the medium is both linear and non-dispersive (otherwise, see the last section in this chapter). Take the time derivative of u_{field}:

$$\begin{aligned}\partial_t u_{\text{field}} &= \mathbf{E}\cdot\partial_t\mathbf{D} + \mathbf{H}\cdot\partial_t\mathbf{B} \\ &= \mathbf{E}\cdot\boldsymbol{\nabla}\times\mathbf{H} - \mathbf{H}\cdot\boldsymbol{\nabla}\times\mathbf{E} - \mathbf{E}\cdot\mathbf{J} \\ &= -\boldsymbol{\nabla}\cdot(\mathbf{E}\times\mathbf{H}) - \mathbf{E}\cdot\mathbf{J}\end{aligned}$$

where we have also used Faraday's Law and the generalised Ampère Law, together with a vector calculus identity.

It is now clear that we can write the energy current density as the **Poynting vector**:

$$\mathbf{S} = \mathbf{E}\times\mathbf{H} \tag{5.30}$$

if we recognise $\mathbf{E}\cdot\mathbf{J}$ as the rate at which the fields do work on the sources per unit volume, and therefore as the rate of change of the sources' mechanical energy density, $\partial_t u_{\text{mech}}$. $\mathbf{S}$ has the expected dimensions of energy per unit area per unit time:

$$\partial_t u + \boldsymbol{\nabla}\cdot(\mathbf{E}\times\mathbf{H}) = 0 \tag{5.31}$$

with u the total energy density (fields *and* sources) is often known as the **Poynting Theorem**.

5.7.2 Conservation of Momentum

If there is energy in the fields, it should come as no surprise that there is also momentum, and that the fields can exchange it with the sources. From the Lorentz force law, the total force exerted by the fields on sources contained inside a volume V is:

$$\mathbf{F} = \mathrm{d}_t\mathbf{P}_{\text{mech}} = \int_{\mathrm{V}}(\rho\,\mathbf{E}+\mathbf{J}\times\mathbf{B})\,\mathrm{d}^3x$$

where $\mathbf{P}_{\text{mech}}$ is the total momentum of all the charges (free *and* bound) in the volume. Now we eliminate the sources with Maxwell's equations for the microscopic fields $\mathbf{E}$ and $\mathbf{B}$. This yields:

$$\rho\,\mathbf{E}+\mathbf{J}\times\mathbf{B} = \frac{1}{4\pi k_{\mathrm{e}}}\left[\mathbf{E}\,(\boldsymbol{\nabla}\cdot\mathbf{E}) + c^2\,\mathbf{B}\,(\boldsymbol{\nabla}\cdot\mathbf{B}) - \mathbf{E}\times(\boldsymbol{\nabla}\times\mathbf{E}) - c^2\,\mathbf{B}\times(\boldsymbol{\nabla}\times\mathbf{B})\right] - \frac{1}{4\pi k_{\mathrm{e}}}\,\partial_t(\mathbf{E}\times\mathbf{B})$$

where we have inserted a $\boldsymbol{\nabla}\cdot\mathbf{B}$ term and written:

$$(\partial_t\mathbf{E})\times\mathbf{B} = \partial_t(\mathbf{E}\times\mathbf{B}) - \mathbf{E}\times\partial_t\mathbf{B} = \partial_t(\mathbf{E}\times\mathbf{B}) + \mathbf{E}\times(\boldsymbol{\nabla}\times\mathbf{E})$$

Now that we have a symmetrical-looking form, we will work out the electric part and transpose the results to the magnetic part. With the identities in Jackson's left front cover we eliminate the curl term for $\mathbf{E}$. Thus:

$$\mathbf{E}\,(\boldsymbol{\nabla}\cdot\mathbf{E}) - \mathbf{E}\times(\boldsymbol{\nabla}\times\mathbf{E}) = \mathbf{E}\,(\boldsymbol{\nabla}\cdot\mathbf{E}) + (\mathbf{E}\cdot\boldsymbol{\nabla})\,\mathbf{E} - \frac{1}{2}\boldsymbol{\nabla}(E^2)$$

The right-hand side is easier to manipulate if we write it in index notation:

$$E^i\,(\partial_j E^j) + (E^j\partial_j)\,E^i - \frac{1}{2}\partial_i E^2 = \partial_j\left[E^iE^j - \frac{1}{2}\delta_j{}^i\,E^2\right]$$

If then we define the components of the **Maxwell stress tensor** as:

$$T_{ij} \coloneqq \frac{1}{4\pi k_{\mathrm{e}}}\left[E_iE_j + c^2\,B_iB_j - \frac{1}{2}\delta_{ij}\,(E^2 + c^2B^2)\right] \tag{5.32}$$

we obtain the law of conservation of momentum:

$$\begin{aligned} \partial_t(\mathbf{P}_{\text{mech}} + \mathbf{P}_{\text{field}})^i &= \int_{\text{V}} \partial_j T^{ij}\, \mathrm{d}^3x \\ &= \oint_{\text{S}} T^{ij}\, n_j\, \mathrm{d}a \end{aligned} \tag{5.33}$$

where, in SI units:

$$\mathbf{P}_{\text{field}} = \int_{\text{V}} \epsilon_0(\mathbf{E}\times\mathbf{B})\, \mathrm{d}^3x = \frac{1}{c^2}\int_{\text{V}} \mathbf{S}\, \mathrm{d}^3x \tag{5.34}$$

If no charge crosses the boundary of the volume, we interpret T^{ij} as the i$^{\text{th}}$ component of the flow of field momentum per unit area, or **momentum flux density**, in the j^{th} direction across the boundary. We can also see it as the force per unit area acting across the boundary on the field-particle system contained in the volume. On the other hand $\mathbf{S}/c^2$ is identified as the **momentum density** in the fields.

The *Cartesian* components of the total electromagnetic force on a charge distribution are given by:

$$F^i = -\frac{1}{c^2}\,\mathrm{d}_t \int S^i\, \mathrm{d}^3x + \oint_{\text{S}} T^{ij} n_j\, \mathrm{d}a \tag{5.35}$$

where any volume that encloses the whole charge distribution can be used. In static cases, the first term on the right-hand side vanishes.

5.8 Poynting Theorem for Harmonic Fields (section J6.9)

Often the fields we deal with are harmonic, and it is useful to derive the form of the Poynting theorem relevant to those fields.

The time dependence of the harmonic fields and sources is of the form $\mathrm{e}^{-\mathrm{i}\omega t}$, eg.:

$$\mathbf{E}(\mathbf{x},t) = \frac{1}{2}\left(\mathbf{E}(\mathbf{x})\,\mathrm{e}^{-\mathrm{i}\omega t} + \mathbf{E}^*(\mathbf{x})\,\mathrm{e}^{\mathrm{i}\omega t}\right)$$

Maxwell's macroscopic field equations for the position-dependent part of harmonic fields then become:

$$\begin{aligned} \boldsymbol{\nabla}\cdot\mathbf{D} &= \rho & \qquad \boldsymbol{\nabla}\times\mathbf{H} &= \mathbf{J} - \mathrm{i}\,\omega\,\mathbf{D} \\ \boldsymbol{\nabla}\cdot\mathbf{B} &= 0 & \qquad \boldsymbol{\nabla}\times\mathbf{E} - \mathrm{i}\,\omega\,\mathbf{B} &= 0 \end{aligned} \tag{5.36}$$

To understand the information contained in the results we are about to derive, we note that scalar products of real vectors can be written as:

$$\begin{aligned} \mathbf{A}(\mathbf{x},t)\cdot\mathbf{B}(\mathbf{x},t) &= \frac{1}{4}\left[\mathbf{A}(\mathbf{x})\,\mathrm{e}^{-\mathrm{i}\omega t} + \mathbf{A}^*(\mathbf{x})\,\mathrm{e}^{\mathrm{i}\omega t}\right]\cdot\left[\mathbf{B}(\mathbf{x})\,\mathrm{e}^{-\mathrm{i}\omega t} + \mathbf{B}^*(\mathbf{x})\,\mathrm{e}^{\mathrm{i}\omega t}\right] \\ &= \frac{1}{2}\,\Re\left[\mathbf{A}^*(\mathbf{x})\cdot\mathbf{B}(\mathbf{x}) + \mathbf{A}(\mathbf{x})\cdot\mathbf{B}(\mathbf{x})\,\mathrm{e}^{-2\mathrm{i}\omega t}\right] \end{aligned}$$

Therefore, the time-averaged scalar product is:

$$\left\langle\, \mathbf{A}(\mathbf{x},t)\cdot\mathbf{B}(\mathbf{x},t)\,\right\rangle = \frac{1}{2}\,\Re\left[\mathbf{A}^*(\mathbf{x})\cdot\mathbf{B}(\mathbf{x})\right] \tag{5.37}$$

Now use Maxwell's equations to write sources in terms of fields in the following expression:

$$\begin{aligned}\frac{1}{2}\int_{\mathrm{V}} \mathbf{J}^*\cdot\mathbf{E}\,\mathrm{d}^3x &= \frac{1}{2}\int_{\mathrm{V}} \mathbf{E}\cdot\left[\boldsymbol{\nabla}\times\mathbf{H}^* - \mathrm{i}\omega\,\mathbf{D}^*\right]\mathrm{d}^3x \\ &= \frac{1}{2}\int_{\mathrm{V}}\left[-\boldsymbol{\nabla}\cdot(\mathbf{E}\times\mathbf{H}^*) - \mathrm{i}\omega\,(\mathbf{E}\cdot\mathbf{D}^* - \mathbf{B}\cdot\mathbf{H}^*)\right]\mathrm{d}^3x \\ &= 2\mathrm{i}\,\omega\int_{\mathrm{V}}(w_{\mathrm{m}}-w_{\mathrm{e}})\,\mathrm{d}^3x - \oint_{\mathrm{S}}\mathbf{S}\cdot\hat{\mathbf{n}}\,\mathrm{d}a \end{aligned} \tag{5.38}$$

the last line being the complex Poynting theorem for harmonic fields, where:

$$\mathbf{S} = \frac{1}{2}\,\mathbf{E}\times\mathbf{H}^* \tag{5.39}$$

is defined as the complex Poynting vector, and:

$$w_{\mathrm{e}} = \frac{1}{4}\,\mathbf{E}\cdot\mathbf{D}^* \qquad\qquad w_{\mathrm{m}} = \frac{1}{4}\,\mathbf{B}\cdot\mathbf{H}^* \tag{5.40}$$

are the electric and magnetic field energy density, respectively.

Thus, $\langle\,\mathbf{J}(\mathbf{x},t)\cdot\mathbf{E}(\mathbf{x},t)\,\rangle$ is just the real part of the left-hand side of the complex Poynting theorem. We see that the real part of eq. (5.38) is just a statement of energy conservation for *time-averaged* harmonic fields. Also, the real part of the right-hand side of eq. (5.39) is to be interpreted as the time-averaged Poynting vector for harmonic fields.

In the complex Poynting theorem, the term that depends on the field energies contributes to the real part only when these energies are complex, ie. in the case of the lossy dielectrics treated in the next section.

5.9 Poynting Theorem for Lossy, Dispersive Linear Media (section J6.8) — (optional)

In realistic media, even when linear and isotropic, other phenomena are present, such as energy dissipation and dependence of the permittivity ϵ on frequency (**dispersion**), that make energy conservation somewhat more complicated to analyse in the macrosopic formalism.

As we shall see when we solve Maxwell's wave equations, energy losses translate into a complex ϵ. The way to introduce this into our analysis while at the same time allowing for dispersion is to write a time-domain to frequency-domain Fourier integral of the fields. For instance:

$$\begin{aligned}\mathbf{E}(\mathbf{x},t) &= \int_{-\infty}^{\infty}\mathbf{E}(\mathbf{x},\omega)\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}\omega \\ \mathbf{D}(\mathbf{x},t) &= \int_{-\infty}^{\infty}\mathbf{D}(\mathbf{x},\omega)\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}\omega\end{aligned}$$

where linearity and isotropy mean that $\mathbf{D}(\mathbf{x},\omega)=\epsilon(\omega)\,\mathbf{E}(\mathbf{x},\omega)$, with $\mathbf{E}(\mathbf{x},\omega)$ and $\mathbf{D}(\mathbf{x},\omega)$ *complex* fields. Take the complex conjugate of these expressions and note that the fields on the left-hand side are real. Then consistency demands that $\mathbf{E}(\mathbf{x},-\omega)=\mathbf{E}^*(\mathbf{x},\omega)$ and $\epsilon(-\omega)=\epsilon^*(\omega)$.

Now, in our derivation of the Poynting theorem (5.31), when differentiating the electric energy density, $\frac{1}{2}\mathbf{D}\cdot\mathbf{E}$, we assumed that the medium was non-dispersive. We can calculate a more general form of the theorem in terms of our Fourier integrals. Substituting $\omega'=-\omega$ in the $\mathbf{E}$ integral, $\int\mathbf{E}(\mathbf{x},-\omega')\mathrm{e}^{\mathrm{i}\omega' t}\,\mathrm{d}\omega' = \int\mathbf{E}^*(\mathbf{x},\omega')\mathrm{e}^{\mathrm{i}\omega' t}\,\mathrm{d}\omega'$, and there comes (dropping the $\mathbf{x}$ dependence to minimise clutter):

$$\begin{aligned}\mathbf{E}\cdot\partial_t\mathbf{D} &= \int\mathrm{d}\omega\int\mathrm{d}\omega'\,\mathbf{E}^*(\omega')\cdot\left[-\mathrm{i}\,\omega\epsilon(\omega)\right]\mathbf{E}(\omega)\,\mathrm{e}^{-\mathrm{i}(\omega-\omega')t} \\ &= \frac{1}{2}\int\mathrm{d}\omega\int\mathrm{d}\omega'\,\mathbf{E}^*(\omega')\cdot\left[-\mathrm{i}\,\omega\epsilon(\omega) + \mathrm{i}\,\omega'\epsilon^*(\omega')\right]\mathbf{E}(\omega)\,\mathrm{e}^{-\mathrm{i}(\omega-\omega')t}\end{aligned}$$

where the first line can be recovered (exercise) from the second by substituting $\omega \to -\omega'$ and $\omega' \to -\omega$ in its second term. At this stage, all this is just rewriting.

To make further progress, we assume that $\mathbf{E}(\mathbf{x},\omega)$ has a narrow frequency spectrum (it comes as a long pulse), ie., it is negligible outside a fairly narrow range of frequencies over which $\epsilon(\omega)$ changes only slowly. If this true, then an expansion of the terms inside the square bracket around $\omega' = \omega$ gives:

$$-\mathrm{i}\,\omega\epsilon(\omega) \,+\, \mathrm{i}\,\omega'\epsilon^*(\omega') \;\approx\; 2\,\omega\,\Im[\epsilon(\omega)] \;-\; \mathrm{i}\,(\omega-\omega')\,\mathrm{d}_\omega\Big[\omega\,\epsilon^*(\omega)\Big]$$

Inserting this into our previous expression and noticing that the second term can be written as a time derivative, we finally get:

$$\begin{aligned}\mathbf{E}\cdot\partial_t\mathbf{D} \;=\; & \int \mathrm{d}\omega \int \mathrm{d}\omega'\,\mathbf{E}^*(\omega')\cdot\mathbf{E}(\omega)\,\omega\,\Im[\epsilon(\omega)]\,\mathrm{e}^{-\mathrm{i}(\omega-\omega')t} \\ & +\; \partial_t\,\frac{1}{2}\int \mathrm{d}\omega \int \mathrm{d}\omega'\,\mathbf{E}^*(\omega')\cdot\mathbf{E}(\omega)\,\mathrm{d}_\omega\big[\omega\,\epsilon^*(\omega)\big]\,\mathrm{e}^{-\mathrm{i}(\omega-\omega')t}\end{aligned} \tag{5.41}$$

When ϵ is independent of frequency (no dispersion), $\epsilon^*(\omega) = \epsilon(\omega) = \epsilon$, so $\Im[\epsilon(\omega)] = 0$ and the first term vanishes. In that case the second term is simply the change in electric energy density, $\partial_t u_{\text{field}}$, as in our previous treatment of the Poynting theorem which we now realise assumed a non-dispersive medium.

Evidently, in the dispersive case, the term that is being differentiated with respect to time must be some sort of effective energy density u_{eff}. We can write it in a less complicated form when $\mathbf{E}(\mathbf{x},t) = \tilde{\mathbf{E}}(\mathbf{x},t)\cos(\omega_0 t)$, where ω_0 is the central frequency of the narrow frequency range over which the field is significant, and $\tilde{\mathbf{E}}(\mathbf{x},t)$ is a function with slow time variation. Averaging over a period corresponding to ω_0 yields:

$$\big\langle\,\mathbf{E}\cdot\partial_t\mathbf{D}\,\big\rangle \;=\; 2\omega_0\,\Im\big[\epsilon(\omega_0)\,\big]\big\langle\,\mathbf{E}^2(\mathbf{x},t)\,\big\rangle \;+\; \partial_t\left(\Re\big[\mathrm{d}_\omega\big[\omega\,\epsilon(\omega)\big]\Big|_{\omega_0}\big]\,\big\langle\,\mathbf{E}^2(\mathbf{x},t)\,\big\rangle\right)$$

An exactly analogous expression obtains for $\mathbf{H}\cdot\partial_t\mathbf{B}$. In the end, the Poynting theorem is modified in the following way:

$$\begin{aligned}\partial_t u_{\text{eff}} \,+\, \boldsymbol{\nabla}\cdot\mathbf{S} \;=\; & -\,\mathbf{J}\cdot\mathbf{E} \;-\; 2\omega_0\,\Im\big[\epsilon(\omega_0)\big]\,\big\langle\,\mathbf{E}^2(\mathbf{x},t)\,\big\rangle \\ & -\; 2\omega_0\,\Im\big[\mu(\omega_0)\big]\,\big\langle\,\mathbf{H}^2(\mathbf{x},t)\,\big\rangle\end{aligned} \tag{5.42}$$

where u_{eff} is the average energy stored in the fields:

$$u_{\text{eff}} \;=\; \Re\left[\mathrm{d}_\omega\big[\omega\,\epsilon(\omega)\big]\Big|_{\omega_0}\right]\,\big\langle\,\mathbf{E}^2(\mathbf{x},t)\,\big\rangle \;+\; \Re\left[\mathrm{d}_\omega\big[\omega\,\mu(\omega)\big]\Big|_{\omega_0}\right]\,\big\langle\,\mathbf{H}^2(\mathbf{x},t)\,\big\rangle \tag{5.43}$$

The new terms on the right of the improved Poynting theorem represent non-ohmic dissipation of field energy, ie. energy absorbed by the material. Only in non-dipersive media do they vanish, since then the permittivity and permeability are real. Dissipation can occur only when there is dispersion. As for u_{eff}, we see that it has the expected dependence on the fields, but with $\epsilon/2$ replaced by the **Brillouin correction**, $\Re\Big[\mathrm{d}_\omega\big[\omega\,\epsilon(\omega)\big]\Big|_{\omega_0}\Big]$, and similarly for the magnetic term.

6 Plane Electromagnetic Waves

6.1 Plane Waves in Nonconducting media (section J7.1)

There are very well known solutions to the wave equation for Maxwell fields in vacuum, eq. (5.6): they are the solutions to the source-free scalar wave equation. To find those **plane-wave** solutions, we look for vector functions that depend only on time and one spatial variable, say x. Then we use an elegant trick from monsieur d'Alembert to change variables to the two independent variables $\xi_\pm = x \pm ct$, so that the wave equation becomes $\partial^2_{\xi_+\xi_-}\mathbf{F} = 0$, where $\mathbf{F} = \{\mathbf{E},\, \mathbf{B}\}$, whose general solution drops out immediately:

$$\mathbf{F} \;=\; \mathbf{f}_-(x - ct) \;+\; \mathbf{f}_+(x + ct)$$

The arbitrary functions $\mathbf{f}_\mp(0, t)$ are propagated in opposite directions at velocity $\pm c\hat{\mathbf{x}}$, with $\mathbf{F}_\pm(x \pm ct, y, z)$ uniform on a plane perpendicular to the direction of propagation corresponding to a constant value of $x \pm ct$, thus justifying their name "plane wave".

Instead of a wave propagating in a one-dmensional vacuum, consider a wave propagating in a 3-dim medium without free charges and currents, and specialise to time-harmonic fields. According to eq. (5.36), Maxwell's equations for the position-dependent part of *harmonic* fields in a uniform, linear and isotropic medium without free sources are:

$$\begin{aligned} \boldsymbol{\nabla} \times \mathbf{B} \;+\; \mathrm{i}\,\omega\mu\epsilon\,\mathbf{E} \;&=\; 0 \\ \boldsymbol{\nabla} \times \mathbf{E} \;-\; \mathrm{i}\,\omega\,\mathbf{B} \;&=\; 0 \end{aligned} \tag{6.1}$$

These two equations are sufficient to determine the fields since they contain the divergence equations as identities. To decouple them, take the curl of one and combine it with the other to obtain Helmholtz equations ($k^2 = \omega^2\mu\epsilon$):

$$\begin{aligned} (\nabla^2 \;+\; k^2)\,\mathbf{E} \;&=\; 0 \\ (\nabla^2 \;+\; k^2)\,\mathbf{B} \;&=\; 0 \end{aligned} \tag{6.2}$$

The complex representation of the solution of the Helmholtz equation (with the harmonic time dependence tagged on) is, in Cartesian coordinates:

$$\begin{aligned} \mathbf{E}(\mathbf{x}, t) \;&=\; \mathbf{E}_0\, \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x} - \omega t)} \\ \mathbf{B}(\mathbf{x}, t) \;&=\; \mathbf{B}_0\, \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x} - \omega t)} \end{aligned}$$

where we have also folded in the time dependence, and:

$$\mathbf{k}\cdot\mathbf{k} \;=\; \mu\epsilon\,\omega^2 \;=\; \mu_\mathrm{r}\epsilon_\mathrm{r}\,\omega^2/c^2$$

with the field strengths $\mathbf{E}_0$ and $\mathbf{B}_0$ complex constants, and $\mathbf{k}$ the wave vector in the direction of propagation. The solutions exhibit the magic $\mathbf{x} \pm \mathbf{v}t$ dependence on $\mathbf{x}$ and t seen in d'Alembert's solution that identifies them as progressive waves traveling at the **phase velocity** of magnitude $c_\mathrm{m} = \omega/k = c/n$, where $n = \sqrt{\epsilon_\mathrm{r}\mu_\mathrm{r}}$ is known as the **index of refraction** of the medium.

In a nondispersive medium, in which ϵ_r and μ_r are independent of frequency, we can superpose solutions with different values of k to construct any pulse shape that propagates at the same phase velocity as its components. In a dispersive medium, each component propagates at its own speed and this causes the shape to change as it propagates.

Now the Helmholtz equation is a second-order differential equation and, as we know, not all its solutions will satisfy the first-order Maxwell equations! Inserting the solutions into the divergence and curl equations gives:

$$\mathbf{k}\cdot\mathbf{E}_0 \;=\; 0 \qquad\qquad \mathbf{k}\cdot\mathbf{B}_0 \;=\; 0 \qquad\qquad \mathbf{B}_0 \;=\; \sqrt{\mu_\mathrm{r}\epsilon_\mathrm{r}}\,\hat{\mathbf{n}} \times \frac{\mathbf{E}_0}{c}$$

where $\hat{\mathbf{n}} = \mathbf{k}/k$ is a unit vector in the direction of propagation. This means that both $\mathbf{E}$ and $\mathbf{B}$ are perpendicular to each other and to the direction of propagation (hence the name **transverse wave**).

The energy density (5.40) in the fields is:

$$\begin{aligned} u &= \frac{1}{4}\left(\epsilon\,\mathbf{E}\cdot\mathbf{E}^* + \frac{1}{\mu}\,\mathbf{B}\cdot\mathbf{B}^*\right) \\ &= \frac{\epsilon}{2}\,|E_0|^2 = \frac{1}{2\mu}\,|B_0|^2 \end{aligned} \tag{6.3}$$

the last line being true only if $\mathbf{k}$ is real.

The energy flux density is given by the real part of the Poynting vector from eq. (5.39), averaged over time to get rid of oscillations. If $\mathbf{k}$ is real, we have:

$$\mathbf{S} = \frac{1}{2Z}\,|E_0|^2\,\hat{\mathbf{n}} = u\,\mathbf{v} \tag{6.4}$$

where $Z = \sqrt{\mu/\epsilon}$, with units of ohms, is called the **impedance** of the medium. We also see that the speed of the energy flow is $v = c/n$ as expected. [Note the missing factor of 1/2 in eq. (J7.13).]

When $\mathbf{k}$ is complex, things get a bit more, well, complex. If we write $\Re[\hat{\mathbf{n}}] \equiv \hat{\mathbf{n}}_\mathrm{R}$, and $\Im[\hat{\mathbf{n}}] \equiv \hat{\mathbf{n}}_\mathrm{I}$, then the requirement that $\hat{\mathbf{n}}\cdot\hat{\mathbf{n}} = 1$ no longer means that $\mathbf{n}$ has unit length! Instead:

$$\begin{aligned} \hat{\mathbf{n}}_\mathrm{R}^2 - \hat{\mathbf{n}}_\mathrm{I}^2 &= 1 \\ \hat{\mathbf{n}}_\mathrm{R}\cdot\hat{\mathbf{n}}_\mathrm{I} &= 0 \end{aligned}$$

The orthogonality of the real and imaginary parts of $\mathbf{k}$ provides one natural choice for the x and y coordinate axes. Also, we can write $|\hat{\mathbf{n}}_\mathrm{R}| = \cosh\theta$ and $|\hat{\mathbf{n}}_\mathrm{I}| = \sinh\theta$, with θ a real parameter. Therefore:

$$\hat{\mathbf{n}} = \hat{\mathbf{e}}_1\cosh\theta + \mathrm{i}\,\hat{\mathbf{e}}_2\sinh\theta \qquad (\hat{\mathbf{e}}_1 = \mathbf{i},\ \hat{\mathbf{e}}_2 = \mathbf{j})$$

Then the constraint $\mathbf{k}\cdot\mathbf{E}_0 = k\hat{\mathbf{n}}\cdot\mathbf{E}_0 = 0$ can be satisfied by:

$$\mathbf{E}_0 = A\,(\mathrm{i}\,\sinh\theta\,\hat{\mathbf{e}}_1 - \cosh\theta\,\hat{\mathbf{e}}_2) + A'\,\hat{\mathbf{e}}_3 \tag{6.5}$$

Also, when $\mathbf{k}$ is complex, the amplitude of the fields undergoes exponential decay, and so do the energy density and flux.

6.2 Linear and Elliptical Polarisation (section J7.2)

When $\mathbf{k}$ is real, we shall find it useful to introduce a set of basis vectors ($\hat{\boldsymbol{\epsilon}}_1$, $\hat{\boldsymbol{\epsilon}}_2$, $\hat{\mathbf{n}}$), called the **linear polarisation** basis.

Consider two superposed waves of equal frequency with orthogonal electric fields $\mathbf{E}_\mathrm{i} = \hat{\boldsymbol{\epsilon}}_\mathrm{i}\,\tilde{E}_\mathrm{i}\,\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{x} - \mathrm{i}\omega t)$ (i=1, 2), where $\tilde{E}_i \equiv a_\mathrm{i}\,\mathrm{e}^{\mathrm{i}\delta_\mathrm{i}}$ are complex amplitudes. The associated magnetic fields in the waves are: $\mathbf{B}_\mathrm{i} = \sqrt{\mu_\mathrm{r}\epsilon_\mathrm{r}}\,\hat{\mathbf{n}}\times\mathbf{E}_\mathrm{i}/c$.

The most general form for the superposition of waves that have a phase difference is:

$$\begin{aligned} \mathbf{E}(\mathbf{x},t) &= (\hat{\boldsymbol{\epsilon}}_1\,\tilde{E}_1 + \hat{\boldsymbol{\epsilon}}_2\,\tilde{E}_2)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)} \\ &= \left(\hat{\boldsymbol{\epsilon}}_1\,a_1 + \hat{\boldsymbol{\epsilon}}_2\,a_2\,\mathrm{e}^{-\mathrm{i}(\delta_1-\delta_2)}\right)\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t+\delta_1)} \end{aligned} \tag{6.6}$$

Two cases of interest arise:

- **Linear polarisation** ($\delta_1 = \delta_2$ or $\delta_1 = \delta_2 \pm \pi$)

  If the waves are exactly in phase, we can write $a_1 = E_0 \cos\theta$, $a_2 = E_0 \sin\theta$. Then $\mathbf{E}$ is linearly polarised, with its direction at an angle such that $\tan\theta = a_2/a_1$.

- **Elliptical polarisation** ($\delta_1 - \delta_2 \neq 0, \neq \pm\pi$)

  When the two perpendicular waves are out of phase, we can recast eq. (6.6) in a more transparent form by choosing the x and y axes along $\hat{\boldsymbol{\epsilon}}_1$ and $\hat{\boldsymbol{\epsilon}}_2$, respectively. Then $\mathbf{k}\cdot\mathbf{x} = kz$, and the field components are obtained from the real part:

$$\begin{aligned} E_x(\mathbf{x},t) &= a_1 \cos(kz - \omega t + \delta_1) \\ E_y(\mathbf{x},t) &= a_2 \left[\cos(\delta_1 - \delta_2)\cos(kz - \omega t + \delta_1) + \sin(\delta_1 - \delta_2)\sin(kz - \omega t + \delta_1)\right] \end{aligned} \tag{6.7}$$

  Squaring and eliminating the time-dependent harmonic functions, we arrive at:

$$\frac{E_x^2}{a_1'^2} + \frac{E_y^2}{a_2'^2} - 2\cos(\delta_1 - \delta_2)\frac{E_x}{a_1'}\frac{E_y}{a_2'} = 1$$

  where $a_{\rm i}' = a_{\rm i}\sin(\delta_1 - \delta_2)$. This shows that the tip of the field vector traces an ellipse, and the wave is said to be elliptically polarised. The cross-term indicates that the ellipse is tilted with respect to the x and y axes. When $\delta_1 - \delta_2 = \pm\pi/2$, the principal axes are oriented along the coordinate axes. If in addition $a_1 = a_2$, we speak of circular polarisation.

  Viewed from a *fixed* point and facing the oncoming wave, the electric field vector rotates counterclockwise (positive helicity) when $0 < \delta_1 - \delta_2 < \pi$, or clockwise (negative helicity) when $\pi < \delta_1 - \delta_2 < 2\pi$.

  Elliptical polarisation can also be described by transforming to a different orthonormal basis, the **circular polarisation** basis $\hat{\boldsymbol{\epsilon}}_\pm = (\hat{\boldsymbol{\epsilon}}_1 \pm {\rm i}\,\hat{\boldsymbol{\epsilon}}_2)/\sqrt{2}$, in which:

$$\mathbf{E}(\mathbf{x},t) = (\hat{\boldsymbol{\epsilon}}_+ \tilde{E}_+ + \hat{\boldsymbol{\epsilon}}_- \tilde{E}_-)\,{\rm e}^{{\rm i}(\mathbf{k}\cdot\mathbf{x} - \omega t)} \tag{6.8}$$

  where

$$\tilde{E}_\pm = a_\pm\,{\rm e}^{{\rm i}\delta_\pm} = \frac{1}{\sqrt{2}}(\tilde{E}_1 \pm {\rm i}\,\tilde{E}_2)$$

We close this short discussion of polarisation by mentioning a useful parametrisation that allows to find the polarisation of a wave from intensity measurements. In the linear polarisation basis, they are:

$$\begin{aligned} s_0 &= |\hat{\boldsymbol{\epsilon}}_1\cdot\mathbf{E}_1|^2 + |\hat{\boldsymbol{\epsilon}}_2\cdot\mathbf{E}_2|^2 = a_1^2 + a_2^2 \\ s_1 &= |\hat{\boldsymbol{\epsilon}}_1\cdot\mathbf{E}_1|^2 - |\hat{\boldsymbol{\epsilon}}_2\cdot\mathbf{E}_2|^2 = a_1^2 - a_2^2 \\ s_2 &= 2\,\Re[(\hat{\boldsymbol{\epsilon}}_1\cdot\mathbf{E}_1)^*(\hat{\boldsymbol{\epsilon}}_2\cdot\mathbf{E}_2)] = 2a_1 a_2\cos(\delta_1 - \delta_2) \\ s_3 &= 2\,\Im[(\hat{\boldsymbol{\epsilon}}_1\cdot\mathbf{E}_1)^*(\hat{\boldsymbol{\epsilon}}_2\cdot\mathbf{E}_2)] = 2a_1 a_2\sin(\delta_1 - \delta_2) \end{aligned} \tag{6.9}$$

These **Stokes parameters**, as they are known, obey the constraint $s_0^2 - s_1^2 - s_2^2 - s_3^2 = 0$.

Similar expressions can be written in the circular basis — see eq. (J7.28).

6.3 Reflection and Refraction at a Plane Interface Between Dielectrics (section J7.3)

We wish to find out what happens to an electromagnetic wave incident on a flat interface between two linear, isotropic and homogeneous dielectrics. A monochromatic plane wave:

$$\mathbf{E}(\mathbf{x},t) = \mathbf{E}_0{\rm e}^{{\rm i}(\mathbf{k}\cdot\mathbf{x} - \omega t)}, \qquad \mathbf{B}(\mathbf{x},t) = \frac{1}{c_{\rm m}}(\hat{\mathbf{n}}\times\mathbf{E}) \tag{6.10}$$

is incident on a plane interface (see figure J7.5) , giving rise to a transmitted wave:

$$\mathbf{E}'(\mathbf{x},t) \;=\; \mathbf{E}'_0 \mathrm{e}^{\mathrm{i}(\mathbf{k}'\cdot\mathbf{x}-\omega t)}, \qquad \mathbf{B}'(\mathbf{x},t) \;=\; \frac{1}{c_{\mathrm{m}}}(\hat{\mathbf{n}}'\times\mathbf{E}') \tag{6.11}$$

and a reflected wave:

$$\mathbf{E}''(\mathbf{x},t) \;=\; \mathbf{E}''_0 \mathrm{e}^{\mathrm{i}(\mathbf{k}''\cdot\mathbf{x}-\omega t)}, \qquad \mathbf{B}''(\mathbf{x},t) \;=\; \frac{1}{c''_{\mathrm{m}}}(\hat{\mathbf{n}}''\times\mathbf{E}') \tag{6.12}$$

All three waves have the same frequency, ω, and their wave numbers are related by $\omega = k\, c_{\mathrm{m}} = k''\, c_{\mathrm{m}} = k' c'_{\mathrm{m}}$, with $c_{\mathrm{m}} = 1/\sqrt{\mu\epsilon}$.

The combined incident and reflected fields in the medium of incidence must match the fields transmitted into the other medium at the interface. These boundary conditions will have the general form:

$$(\text{stuff})\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}} + (\text{stuff})\mathrm{e}^{\mathrm{i}\,\mathbf{k}''\cdot\mathbf{x}} \;=\; (\text{stuff})\mathrm{e}^{\mathrm{i}\,\mathbf{k}'\cdot\mathbf{x}} \tag{6.13}$$

We'll worry about the "stuff" later. What is important in all the matching conditions is that the position and time dependence reside *in the exponentials*, not in the stuff which is made of constants. Since the relations must hold everywhere on the interface and at all times, the exponentials must be equal! This means that, on the interface, $\mathbf{k}\cdot\mathbf{x} = \mathbf{k}'\cdot\mathbf{x} = \mathbf{k}''\cdot\mathbf{x}$, that is: $(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x} = (\mathbf{k}-\mathbf{k}'')\cdot\mathbf{x} = 0$. Since $\mathbf{x}$ is not restricted other than being on the interface, this can hold only if the components of all three $\mathbf{k}$ vectors are *separately equal*.

Conclusion:

a) The incident, reflected, and transmitted wave vectors form a plane, the **plane of incidence**.

The equality of the components of the wave vectors implies that $k\sin\, i = k'\sin\, r = k''\sin\, r'$, where i is the **angle of incidence**, r' is the **angle of reflection**, and r is the **angle of transmission** (or refraction). Then:

b) the angles of reflection and incidence are equal:

$$i \;=\; r' \tag{6.14}$$

c) and the Snell-Descartes law governs refraction:

$$n\sin\, i \;=\; n'\sin\, r \tag{6.15}$$

These three statements are the fundamental laws of geometrical optics. Notice that they arise from the fact that we are dealing with waves. The electromagnetic nature of these waves has not been used. Other kinds of waves (eg. sound) will obey the same laws!

Going back to the general form of the matching conditions (6.13), we can now cancel the exponentials and deal with the "stuff". This is where the electromagnetic nature of the waves comes into play.

The matching of the normal components of $\mathbf{D}$ and $\mathbf{B}$ gives:

$$\left[\epsilon\,(\mathbf{E}_0 \,+\, \mathbf{E}''_0) \;-\; \epsilon'\,\mathbf{E}'_0\right]\cdot\hat{\mathbf{n}} \;=\; 0 \qquad\qquad \left[\mathbf{B}_0 \,+\, \mathbf{B}''_0 \,-\, \mathbf{B}'_0\right]\cdot\hat{\mathbf{n}} \;=\; 0 \tag{6.16}$$

Next, matching of the tangential (parallel to the interface) components of $\mathbf{E}$ and $\mathbf{H}$ leads to:

$$\left[\mathbf{E}_0 \,+\, \mathbf{E}''_0 \;-\; \mathbf{E}'_0\right]\times\hat{\mathbf{n}} \;=\; 0 \qquad\qquad \left[\mathbf{H}_0 \,+\, \mathbf{H}''_0 \,-\, \mathbf{H}'_0\right]\times\hat{\mathbf{n}} \;=\; 0 \tag{6.17}$$

Note that since $\mathbf{H} = \mathbf{B}/\mu = \sqrt{\epsilon\mu}\,\hat{\mathbf{n}}\times\mathbf{E}/\mu$, $H_0 = E_0/Z$ from the definition of the impedance: $Z = \sqrt{\mu/\epsilon}$.

Two cases must be considered: when the polarisation of the incident wave is parallel (p-polarised) to the plane of incidence, and when it is perpendicular (senkrecht in German), or s-polarised, to the plane of incidence. In both cases, one normal condition is trivial and the other one is equivalent to one of the tangential conditions after implementing the Snell-Descartes law in the form $\epsilon'\sin r = (Z/Z')\epsilon\sin i$. Then we shall only use eq. (6.17).

- For **E** parallel to the plane of incidence (p-polarisation), eq. (6.17) projected on the interface becomes:

$$\cos i\,(E_0 - E_0'') - \cos r\,E_0' = 0$$
$$\frac{1}{Z}(E_0 + E_0'') - \frac{1}{Z'}E_0' = 0$$

Solving for the relative amplitudes with the Snell-Descartes law, we obtain the **Fresnel relations** for p-polarisation:

$$\frac{E_0'}{E_0} = \frac{2\cos i}{(Z/Z')\cos i + \sqrt{1-\frac{n^2}{n'^2}\sin^2 i}}$$
$$\frac{E_0''}{E_0} = \frac{(Z/Z')\cos i - \sqrt{1-\frac{n^2}{n'^2}\sin^2 i}}{(Z/Z')\cos i + \sqrt{1-\frac{n^2}{n'^2}\sin^2 i}} \tag{6.18}$$

- For s-polarised **E**, perpendicular to the plane of incidence (ie. parallel to the interface), eq. (6.17) becomes:

$$E_0 + E_0'' - E_0' = 0$$
$$\frac{1}{Z}(E_0 - E_0'')\cos i - \frac{1}{Z'}E_0'\cos r = 0$$

Solving for the relative amplitudes, using the Snell-Descartes law, we obtain the **Fresnel relations** for s-polarisation:

$$\frac{E_0'}{E_0} = \frac{2\cos i}{\cos i + (Z/Z')\sqrt{1-\frac{n^2}{n'^2}\sin^2 i}}$$
$$\frac{E_0''}{E_0} = \frac{\cos i - (Z/Z')\sqrt{1-\frac{n^2}{n'^2}\sin^2 i}}{\cos i + (Z/Z')\sqrt{1-\frac{n^2}{n'^2}\sin^2 i}} \tag{6.19}$$

Note the formal resemblance with the "p" case, except for the location of the impedance ratio. At normal incidence, the two sets become:

$$\frac{E_0'}{E_0} = \frac{2}{1+Z/Z'} \underset{\mu=\mu'}{=} \frac{2}{1+n'/n}$$
$$\frac{E_0''}{E_0} = \pm\frac{1-Z'/Z}{1+Z'/Z} \underset{\mu=\mu'}{=} \pm\frac{1-n/n'}{1+n/n'} \tag{6.20}$$

where the $+/-$ sign applies to perpendicular/parallel polarisation, respectively.

If $n' > n$, the reflected electric field has its phase reversed with respect to the incident field; it is in phase if $n' < n$. The transmitted field is always in phase with the incident one.

6.4 Polarisation by Reflection and Total Internal Reflection (section J7.4)

For p-polarisation, the reflected amplitude vanishes when the numerator does, ie. when:

$$\tan i = \sqrt{\frac{(Z/Z')^2 - 1}{1-(n/n')^2}} \underset{\mu=\mu'}{=} \frac{n'}{n} \tag{6.21}$$

This magic angle of incidence is called the **Brewster angle**, $\theta_{\rm B}$. For s-polarisation, demanding that the reflected amplitude in eq. (6.19) vanish gives $\tan i = \sqrt{-1}$: there is always some reflection for any angle of incidence.

For an air to water ($n = 4/3$) interface, $\theta_{\rm B} \approx 50°$. Since s-polarised waves always undergo some reflection, unpolarised sunlight incident on a water surface at an angle of about $50°$ is reflected mostly s-polarised, *ie. parallel to the horizontal water surface*. This is why polaroid sunglasses with a vertical transmission axis help to reduce the intensity of sunlight reflected from water.

The Brewster effect can also be put to effective use whenever one wishes to extract the s-polarised component of a beam, or to maximise transmission of p-polarised light at an interface. The efficiency of gas lasers is routinely maximised by the use of so-called Brewster windows, through which the p-polarised component of the light can pass with negligible loss.

The Snell-Descartes law shows that when the incidence angle i has a value i_0 called the **critical angle**, such that $\sin i_0 = n'/n$ with $n > n'$, $r = \pi/2$ and the transmitted wave in fact propagates along the surface as an **evanescent wave** which is exponentially attenuated with distance into the second medium. Indeed, when $i > i_0$, $\sin r = (n/n')\sin i = \sin i/\sin i_0$, so that $\cos r = \mathrm{i}[(\sin i/\sin i_0)^2 - 1]^{1/2}$ is imaginary. However:

$$\mathrm{e}^{\mathrm{i}\mathbf{k}'\cdot\mathbf{x}} \;=\; \mathrm{e}^{\mathrm{i}k'(z\cos r + x\sin r)} \;=\; \mathrm{e}^{-k'|\cos r|z}\,\mathrm{e}^{\mathrm{i}k'(\sin i/\sin i_0)x}$$

Then, if we calculate the energy transmitted into the the second medium, we find, using eq. (5.37):

$$\begin{aligned}
\langle\, \mathbf{S}\cdot\hat{\mathbf{z}}\,\rangle \;&=\; \frac{1}{2}\,\Re\big[\hat{\mathbf{z}}\cdot(\mathbf{E}'\times\mathbf{H}'^{*})\big] \\
&=\; \frac{1}{2\omega\,\mu'}\,\Re\big[(\hat{\mathbf{z}}\cdot\mathbf{k}')|\mathbf{E}'_0|^2\big] \\
&=\; \frac{1}{2\omega\,\mu'}\,\Re\big[(k'\,\cos\,r)|\mathbf{E}'_0|^2\big] \\
&=\; 0
\end{aligned}$$

since $\cos\,r$ is imaginary. This shows that no energy is transmitted into the medium, even though the fields there are not zero. If, however, the thickness of the medium is of the order of a few wavelengths, a wave can be observed travelling on the other side of the medium, as in transmission through a finite-height potential barrier in quantum mechanics when the incident energy is smaller than the height of the barrier.

Interestingly enough, though, when a narrow beam is incident at an angle larger than the critical angle, it emerges slightly shifted laterally, as though it has reflected from a depth of about $\delta = \left[k\sqrt{\sin^2 i - \sin^2 i_0}\right]^{-1}$ beyond the interface. This is known as the Goos-Hänchen effect (See Jackson and references quoted therein).

6.5 The Hertz Superpotential and the Interaction of a Plane Wave with a Medium

In this section, we would like to present an illustration of the use of the Hertz superpotential formulation developed in section 5.6 to study the interaction of a plane wave with a bounded medium. We recall that the general solution for macroscopic electric and magnetic harmonic fields is, from eq. (5.29) and (5.28):

$$\begin{aligned}
\mathbf{\Pi}(\mathbf{x}) \;&=\; \mathrm{i}\,\frac{k_{\rm e}}{ck}\int \mathbf{J}_{\rm tot}(\mathbf{x}')\,\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\,\mathrm{d}^3x' \\
\mathbf{E} \;&=\; k^2\mathbf{\Pi} + \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{\Pi}), \qquad \mathbf{B} \;=\; -\,\mathrm{i}\,\frac{k}{c}\,\boldsymbol{\nabla}\times\mathbf{\Pi}
\end{aligned}$$

where $\mathbf{J}_{\rm tot} = \mathbf{J}_{\rm free} - \mathrm{i}\omega\mathbf{P} + \boldsymbol{\nabla}\times\mathbf{M}$, $\mathbf{P}$ and $\mathbf{M}$ being the polarisation and magnetisation vectors. As before, the integrand is to be evaluated at retarded time. The use of macroscopic variables simplifies the treatment considerably, because it bypasses the need to care about things like the "local field", and the molecular polarisibility models of section 3.4.

For the purposes of our illustration we will assume no free charges in a homogeneous, linear and isotropic dielectric medium, so that $\mathbf{J}_{\rm tot}(\mathbf{x}) = -\mathrm{i}\omega\mathbf{P}(\mathbf{x})$. Using the constitutive relation that relates $\mathbf{P}$ and the electric field $\mathbf{E}'$ in such dielectrics: $\mathbf{P} = \epsilon_0\chi_{\rm e}\mathbf{E}'$, it is clear that the above are integro-differential equations for the fields. In this view, the electric field at any point in space is the superposition of the retarded electric fields from every source point. In particular, inside a bounded medium, the net field must be the sum of the incident fields from sources outside the medium and those produced by free (when present) and induced sources in the medium.

So as to be self-consistent, the sources must generate a field that cancels the incident field inside the medium and replaces it by the transmitted field inside the medium..

Similarly, the reflected (scattered) field outside the medium is produced by sources *at every point* inside the medium, not just from the surface as in the standard analysis given in section 6.3.

Let the dielectric occupy the infinite half-space $x > 0$. We will first solve the wave equation (5.29) for the superpotential:

$$\nabla^2\mathbf{\Pi} + k^2\mathbf{\Pi} = 4\pi\mathrm{i}\frac{k_{\rm e}}{ck}\mathbf{J}_{\rm tot} = 4\pi k_{\rm e}\,\mathbf{P} \tag{6.22}$$

We will work with the polarisation vector, and assume that it is a linearly polarised plane wave travelling in the direction perpendicular to the interface: $\mathbf{P}(\mathbf{x}) = \mathbf{P}_0\mathrm{e}^{\mathrm{i}k'x} = \mathbf{P}_0\mathrm{e}^{\mathrm{i}nkx}$, where $\mathbf{P}_0 = P_0\hat{\mathbf{z}}$, and n is a dimensionless constant. Then eq. (5.28) becomes:

$$\mathbf{\Pi}(\mathbf{x}) = \mathrm{i}\,\frac{k_{\rm e}}{ck}\int \mathbf{J}_{\rm tot}(\mathbf{x}')\,\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\,\mathrm{d}^3x' = k_{\rm e}\mathbf{P}_0\int \mathrm{e}^{\mathrm{i}nkx'}\,\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\,\mathrm{d}^3x' = k_{\rm e}\,\mathrm{e}^{\mathrm{i}nkx}\mathbf{P}_0\int \mathrm{e}^{\mathrm{i}nk(x'-x)}\,\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\,\mathrm{d}^3x'$$

The integral of the Green function in the y'-z' plane is most easily done in polar coordinates (ρ', ϕ'), with the change of variable: $\rho'^2 = R^2 - |x' - x|^2$, suggested by the geometry of the calculation:

$$\int_{\rho'=0}^{\infty}\int_{\phi'=0}^{2\pi}\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\,\rho'\mathrm{d}\rho'\,\mathrm{d}\phi' = 2\pi\int_{|x'-x|}^{\infty}\mathrm{e}^{\mathrm{i}kR}\,\mathrm{d}R = 2\pi\lim_{\alpha\to 0}\int_{|x'-x|}^{\infty}\mathrm{e}^{(\mathrm{i}k-\alpha)R}\,\mathrm{d}R$$

The regulator $\alpha > 0$ is introduced to compensate for the unphysical uniformity of the plane wave in the plane perpendicular to the direction of propagation. Then even ***Maple*** is happy to evaluate it:

```
> int(exp((I*k -alpha)*R),R=a..infinity) assuming alpha>0;
```

$$-\frac{\mathrm{e}^{a(\mathrm{I}k-\alpha)}}{\mathrm{I}k-\alpha}$$

Taking the limit and inserting into our expression for $\mathbf{\Pi}$ leaves an integral over x'. Without loss of generality, take the point of observation to lie on the x axis, either outside ($x < 0$) or inside $x > 0$) the medium. In SI units:

$$\mathbf{\Pi}(\mathbf{x}) = \mathbf{\Pi}^{(0)}(\mathbf{x}) + \mathrm{i}\,\frac{\mathbf{P}_0}{2k\epsilon_0}\,\mathrm{e}^{\mathrm{i}nkx}\int_0^{\infty}\mathrm{e}^{\mathrm{i}nk(x'-x)}\,\mathrm{e}^{\mathrm{i}k|x'-x|}\,\mathrm{d}x' \tag{6.23}$$

where we have added a formal homogeneous solution, $\mathbf{\Pi}^{(0)}$, to the particular solution of eq. (6.22) we have just found. It is clear that $\mathbf{\Pi}^{(0)}$ is associated with the incident electric field $\mathbf{E}$, with $\mathbf{E} = \mathbf{E}_0\mathrm{e}^{\mathrm{i}kx}$.

Now, in eq. (5.29), $\boldsymbol{\nabla}\cdot\mathbf{\Pi} = 0$ because the superpotential has no component in the longitudinal direction, which carries the x dependence. Therefore, $\mathbf{E}' = k^2\mathbf{\Pi}'$ for $x > 0$ (inside the medium), and $\mathbf{E}'' = k^2\mathbf{\Pi}''$ for $x < 0$ (outside the medium).

- $x > 0$

  When the observation point lies inside the medium, we have:

$$\begin{aligned}\mathbf{\Pi}'(\mathbf{x}) &= \mathbf{\Pi}^{(0)}(\mathbf{x}) + \mathrm{i}\,\frac{\mathbf{P}_0}{2k\epsilon_0}\,\mathrm{e}^{\mathrm{i}nkx}\left[\int_0^{x}\mathrm{e}^{\mathrm{i}(n-1)k(x'-x)}\,\mathrm{d}x' + \int_x^{\infty}\mathrm{e}^{\mathrm{i}(n+1)k(x'-x)}\,\mathrm{d}x'\right]\\ &= \mathbf{\Pi}^{(0)}(\mathbf{x}) + \frac{\mathbf{P}_0}{k^2\epsilon_0}\left[\frac{1}{n^2-1}\,\mathrm{e}^{\mathrm{i}nkx} - \frac{1}{2(n-1)}\,\mathrm{e}^{\mathrm{i}kx}\right]\end{aligned}$$

The second integral has been made to converge with a regulator in the same way as above. Now it is time to implement consistency with the constitutive relation $\mathbf{P} = \epsilon_0 \chi_e \mathbf{E}'$:

$$\mathbf{P} \;=\; \mathbf{P}_0\, \mathrm{e}^{inkx} \;=\; \epsilon_0 \chi_e (k^2\, \mathbf{\Pi'}) \;=\; \chi_e \left[\epsilon_0 \mathbf{E}_0 \mathrm{e}^{\mathrm{i}kx} + \mathbf{P}_0 \left(\frac{1}{n^2-1}\, \mathrm{e}^{inkx} - \frac{1}{2(n-1)}\, \mathrm{e}^{\mathrm{i}kx} \right) \right]$$

or, rearranging:

$$\mathbf{P}_0 \left(1 - \frac{\chi_e}{n^2-1} \right) \mathrm{e}^{inkx} \;=\; \chi_e \left(\epsilon_0 \mathbf{E}_0 - \frac{\mathbf{P}_0}{2(n-1)} \right) \mathrm{e}^{\mathrm{i}kx}$$

For this equality to hold everywhere, the coefficients in front of the exponentials on both sides must vanish separately. This immediately yields:

$$n^2 \;=\; 1 + \chi_e \tag{6.24}$$

the standard result for the index of refraction, leading directly to the phase velocity in the medium. Also:

$$\mathbf{E}_0 - \frac{1}{n-1} \frac{\mathbf{P}_0}{2\epsilon_0} \;=\; 0$$

represents the cancellation of the incident field inside the medium by a field emitted by the sources induced by that very same incident field. This is usually known as the **extinction theorem**.

Indeed, with this condition, the transmitted field becomes:

$$\mathbf{E}'(\mathbf{x}) \;=\; \frac{1}{\cancel{2(n-1)}} \frac{\mathbf{P}_0}{\cancel{\epsilon_0}} \cancel{\mathrm{e}^{\mathrm{i}kx}} + \frac{\mathbf{P}_0}{\epsilon_0} \left(\frac{1}{n^2-1}\, \mathrm{e}^{inkx} - \frac{1}{\cancel{2(n-1)}} \cancel{\mathrm{e}^{\mathrm{i}kx}} \right) \;=\; \frac{\mathbf{P}_0}{\epsilon_0} \frac{1}{n^2-1}\, \mathrm{e}^{inkx} \;=\; \mathbf{E}'_0\, \mathrm{e}^{inkx} \tag{6.25}$$

- $x < 0$

  When the observation point lies outside the medium, we obtain the reflected field emitted by the induced sources inside:

$$\mathbf{E}''(\mathbf{x}) \;=\; k^2\, \mathbf{\Pi}''(\mathbf{x}) \;=\; \mathrm{i}\, k\, \frac{\mathbf{P}_0}{2\epsilon_0}\, \mathrm{e}^{-\mathrm{i}kx} \int_0^\infty \mathrm{e}^{\mathrm{i}(n+1)kx'}\, \mathrm{d}x' \;=\; -\frac{\mathbf{P}_0}{2\epsilon_0} \frac{1}{n+1}\, \mathrm{e}^{-\mathrm{i}kx} \;=\; \mathbf{E}''_0\, \mathrm{e}^{-\mathrm{i}kx} \tag{6.26}$$

With $\mathbf{P}_0 = 2\epsilon_0 (n-1) \mathbf{E}_0$ from the extinction theorem, the reflection and transmission coefficients are, respectively:

$$\frac{E''_0}{E_0} \;=\; -\frac{n-1}{n+1} \qquad\qquad \frac{E'_0}{E_0} \;=\; \frac{2}{n+1} \tag{6.27}$$

which are consistent with the Fresnel relations obtained in eq. (6.20). Note how, in the superpotential formulation, there is no need for the matching conditions at the interface, unlike in the more standard treatment. Everything, including the phase velocity in the medium, derives from the self-consistency of the constitutive relation for the medium.

6.6 Dispersion in Dielectrics, Conductors and Plasmas (section J7.5)

6.6.1 Simple Drude-Lorentz model for the frequency-dependent dielectric constant

In section 4.3, we wrote down an an expression, eq. (4.5), relating the polarisation vector $\mathbf{P}$ to the molecular polarisability $\gamma_{\rm mol}$ of a species of molecules with number density N and the local electric field:

$$\mathbf{P} \;=\; N\,\langle \mathbf{p}_{\rm mol}\rangle \;=\; N\,\gamma_{\rm mol}\,\epsilon_0\,\mathbf{E}_{\rm local}$$

In trying to calculate $\gamma_{\rm mol}$, we assumed that bound electrons behaved as damped harmonic oscillators. We now delve deeper into this model. If we write $\mathbf{p} = -e\mathbf{x}$, where $\mathbf{x}$ is the position of an electron of mass m and charge $-e$, then the dipole moment that it contributes obeys:

$$\mathrm{d}_t^2\mathbf{p} \;+\; \gamma\,\mathrm{d}_t\mathbf{p} \;+\; \omega_0^2\,\mathbf{p} \;=\; -\,\frac{e^2}{m}\,\mathbf{E}(\mathbf{x},t) \tag{6.28}$$

where $\mathbf{E}$ is the applied electric field which for our purposes here we take to be harmonic and roughly equal to the local field. The constant γ (not to be confused with $\gamma_{\rm mol}$) represents a damping effect. This equation has the well-known solution:

$$\mathbf{p} \;=\; \frac{e^2/m}{\omega_0^2-\omega^2-\mathrm{i}\gamma\,\omega}\,\mathbf{E} \tag{6.29}$$

Now, $\mathbf{P} = (\epsilon_{\rm r}-1)\,\epsilon_0\,\mathbf{E}$ (eq. (3.24)). If we multiply the solution for $\mathbf{p}$ by the number of electrons per molecule, f_j, that have natural frequency ω_j, and then by N, and sum over j, we obtain an equation for $\mathbf{P}$ which can be compared with the relation involving ϵ_r, yielding:

$$\epsilon_{\rm r}(\omega) \;=\; 1+\frac{N\,e^2}{\epsilon_0\,m}\sum_j\frac{f_j}{\omega_j^2-\omega^2\;-\;\mathrm{i}\,\gamma_j\omega} \;=\; 1+\frac{N\,e^2}{\epsilon_0\,m}\sum_j\frac{f_j}{\omega_j^2}\left[\frac{1-\omega^2/\omega_j^2\;+\;\mathrm{i}\gamma_j/\omega_j}{(1-\omega^2/\omega_j^2)^2\;+\;(\gamma_j^2/\omega_j^2)(\omega^2/\omega_j^2)}\right] \tag{6.30}$$

with $\sum f_j \;= Z$ if there are Z electrons in one molecule. The f_j are often called **oscillator strengths**.

6.6.2 Anomalous dispersion, resonant absorption

In general, $\gamma_j/\omega_{\rm j} \ll 1$ and, at frequencies below the lowest resonant frequency, ϵ_r as given by eq. (6.30) is approximately real, and it increases with frequency: we say that the medium exhibits **normal dispersion**. When $\omega \approx \omega_1$ (see figure J7.8), however, the imaginary part of ϵ_r manages to dominate, giving rise to **resonant absorption** of the wave. Indeed, we know that $\mathbf{k}\cdot\mathbf{k} \;=\; \mu_{\rm r}\epsilon_{\rm r}\,\omega^2/c^2$. Writing $k \equiv \beta + \mathrm{i}\alpha/2$, this equation becomes, for non-magnetic media:

$$\frac{\omega^2}{c^2}\,\Re[\epsilon_{\rm r}] \;=\; \beta^2 \;-\; \frac{\alpha^2}{4} \qquad\qquad \frac{\omega^2}{c^2}\,\Im[\epsilon_{\rm r}] \;=\; \alpha\,\beta$$

Eliminating α yields:

$$\beta^2 \;=\; \frac{1}{2}\frac{\omega^2}{c^2}\,\Re[\epsilon_{\rm r}]\left[1\;+\;\sqrt{1\;+\;4\,(\Im[\epsilon_{\rm r}]/\Re[\epsilon_{\rm r}])^2}\right] \;\approx\; \begin{cases}\dfrac{\omega^2}{c^2}\,\Re[\epsilon_{\rm r}] & \Im[\epsilon_{\rm r}] \ll \Re[\epsilon_{\rm r}] \\[2ex] \dfrac{\omega^2}{c^2}\,\Im[\epsilon_{\rm r}] & \Im[\epsilon_{\rm r}] \gg \Re[\epsilon_{\rm r}]\end{cases} \tag{6.31}$$

The case $\Im[\epsilon_{\rm r}] \ll \Re[\epsilon_{\rm r}]$ illustrates the direct link between the **absorption coefficient** α that attenuates the wave intensity and $\Im[\epsilon_{\rm r}]$: $\alpha \;\approx\; (\Im[\epsilon_{\rm r}]/\Re[\epsilon_{\rm r}])\,\beta$.

As calculated from eq. (6.30), $\mathrm{d}_\omega\Re[\epsilon_r] < 0$ and $\Re[\epsilon_r] > 0$ decreases (**anomalous dispersion**) in the range $\omega \approx \omega_j$.

Example 6.1. Low frequencies, conductivity

When $\omega \ll \omega_1$, eq. (6.30) goes over to:

$$\epsilon_{\rm r}(\omega) \;=\; 1 \,+\, \frac{N\,e^2}{\epsilon_0\,m}\sum_j \frac{f_{\rm j}}{\omega_j^2} \;>\; 1 \tag{6.32}$$

if, as happens in a dielectric, the lowest resonant frequency does not vanish.

If a fraction f_0 of a molecule's electrons are *free*, however, $\omega_0 = 0$ and we must treat their contribution to the dielectric constant in eq. (6.30) separately, with an extra term that diverges at $\omega = 0$:

$$\mathrm{i}\,\frac{N\,f_0\,e^2}{\epsilon_0\,m\omega(\gamma_0 \,-\, \mathrm{i}\,\omega)}$$

Free electrons, as we know, make a substance a conductor, so we would expect their contribution to be related somehow to the conductivity σ. A quick and dirty connection is provided by the classical Drude model, which writes Ampère's generalised law as: $\boldsymbol{\nabla}\times\mathbf{H} = \mathbf{J} - \mathrm{i}\omega\mathbf{D} = -\mathrm{i}\,\omega\big(\epsilon + \mathrm{i}\,\sigma/\omega\big)\mathbf{E}$, using Ohm's law, $\mathbf{J} = \sigma\mathbf{E}$.

But we could just as well put $\mathbf{J} = 0$ and identify the quantity in round brackets with the *total* permittivity which would already include the contribution from conducting electrons. This would give:

$$\sigma \;=\; \frac{f_0\,N\,e^2}{m(\gamma_0 \,-\, \mathrm{i}\,\omega)} \tag{6.33}$$

Example 6.2. High-frequency limit, plasma frequency

At the other end of the spectrum, far above any resonant frequency, eq. (6.30) assumes the simple form:

$$\epsilon_{\rm r} \;\approx\; 1 \;-\; \frac{NZ\,e^2}{\epsilon_0\,m}\,\frac{1}{\omega^2} \;=\; 1 \;-\; \frac{\omega_{\rm p}^2}{\omega^2} \tag{6.34}$$

with $\omega_{\rm p}$ the **plasma frequency**. At high frequencies, the permittivity is independent of the details of the model, such as the resonant fequencies and damping. In dielectrics, $\epsilon_{\rm r} \approx 1$, and this equation holds only when $\omega \gg \omega_{\rm p}$.

In a tenuous plasma, where electrons are mostly free and γ can be neglected, eq. (6.34) holds even at low frequency, as can be seen from eq. (6.30) with $\gamma_{\rm j}$ neglected. Remembering that $k^2 = \epsilon_{\rm r}\,\omega^2/c^2$ in a non-magnetic material, we arrive at the dispersion relation:

$$\omega(k) \;=\; \sqrt{\omega_{\rm p}^2 \,+\, k^2\,c^2} \tag{6.35}$$

which shows that $\omega_{\rm p}$ is also a cut-off frequency, below which there is no wave propagation.

When $\omega < \omega_{\rm p}$, the wave vector is purely imaginary and the wave turns into exponentially decreasing fields. Such a plasma reflects all incident waves of frequency lower than its plasma frequency.

Example 6.3. Absorption coefficient of water

A final interesting case is provided by the dramatic behaviour of the absorption coefficient of water. At microwave frequencies, water is very absorptive (microwave ovens!), but the coefficient crashes by more than seven orders of magnitude in the range $4\times10^{14} < \nu < 8\times10^{14}$ Hz, what we call the visible region...

6.7 Simplified Model of Propagation in the Ionosphere (section J7.6)

Wave propagation in a tenuous plasma can be strongly affected by the presence of a magnetic induction field, such as that of the Earth. Consider a uniform, static induction $\mathbf{B}_0$, with transverse waves propagating in its direction.

As we have done before, we ignore collisions as well as the influence of the magnetic induction of the waves themselves on the motion of the electrons, to write the equation of motion for a dipole $\mathbf{p} = -e\mathbf{x}$:

$$\ddot{\mathbf{p}} \;-\; \frac{e}{m}\left(\mathbf{B}_0 \times \dot{\mathbf{p}} \;+\; e\mathbf{E}\,\mathrm{e}^{-\mathrm{i}\omega t}\right) \;=\; 0$$

(We are introducing the useful standard notation $\dot{p} = \mathrm{d}_t p$.)

Restrict to circular polarisation: $\mathbf{E}_\pm = (\hat{\boldsymbol{\epsilon}}_1 \,\pm \mathrm{i}\,\hat{\boldsymbol{\epsilon}}_2)\,E$. Now $\mathbf{B}_0$ is parallel to the wave vector $\mathbf{k}$, and so must be perpendicular to both $\hat{\boldsymbol{\epsilon}}_1$ and $\hat{\boldsymbol{\epsilon}}_2$. We look for solutions of the form: $\mathbf{p}_\pm = (\hat{\boldsymbol{\epsilon}}_1 \,\pm \mathrm{i}\,\hat{\boldsymbol{\epsilon}}_2)\,p\,\mathrm{e}^{-\mathrm{i}\omega t}$.

Inserting this into the differential equation, and using $\hat{\mathbf{B}}_0 \times (\hat{\boldsymbol{\epsilon}}_1 \pm \mathrm{i}\,\hat{\boldsymbol{\epsilon}}_2) = \mp\mathrm{i}\,(\hat{\boldsymbol{\epsilon}}_1 \pm \mathrm{i}\,\hat{\boldsymbol{\epsilon}}_2)$, we see that:

$$-\,\omega^2\,\mathbf{p}_\pm \;\pm\; \omega\,\omega_\mathrm{B}\,\mathbf{p}_\pm \;=\; e^2\,\mathbf{E}_\pm/m$$

where $\omega_\mathrm{B} = eB_0/m$. Solving for $\mathbf{p}$, we find that the dipole moment of an electron obeys:

$$\mathbf{p}_\pm \;=\; -\,\frac{e^2\,\mathbf{E}_\pm}{m\omega\big(\omega \,\mp\, \omega_\mathrm{B}\big)}$$

Multiplying by the electron concentration NZ and comparing again with $\mathbf{P} = (\epsilon_\mathrm{r} - 1)\,\epsilon_0\,\mathbf{E}$, we arrive at:

$$\epsilon_{\mathrm{r}\mp} \;=\; 1 \;-\; \frac{\omega_\mathrm{p}^2}{\omega\big(\omega \,\mp\, \omega_\mathrm{B}\big)} \;=\; 1 \;-\; \frac{(\omega_\mathrm{p}/\omega_\mathrm{B})^2}{\dfrac{\omega}{\omega_\mathrm{B}}\left(\dfrac{\omega}{\omega_\mathrm{B}} \,\mp\, 1\right)} \tag{6.36}$$

where the upper(lower) sign applies to a positive(negative) helicity wave. This means that the medium is **birefringent**: the dependence of the dielectric constant (and the index of refraction) depends on the helicity of the incident wave.

At low frequencies ($\omega < \omega_\mathrm{B}$), $\epsilon_{\mathrm{r}-}$ is always positive, so waves with *positive* helicity always propagate. But $\epsilon_{\mathrm{r}+}$ turns negative below a certain frequency that depends on $\omega_\mathrm{p}/\omega_\mathrm{B}$. Wave components with negative helicity are then reflected by the plasma.

For $\omega > \omega_\mathrm{B}$, the positive-helicity components cannot propagate until a certain frequency, again dependent on $\omega_\mathrm{p}/\omega_\mathrm{B}$, is reached. At high enough frequencies, both helicities always propagate and there is no reflection. See figure J7.10 for illustrations.

Since a linearly polarised wave can be viewed as the sum of two waves with opposite helicities, there will be ranges of incident frequencies for which one helicity is totally reflected while the other is partially reflected, changing the helicity mix to produce an elliptically polarised reflection. By probing layers in the ionosphere from the ground with suitably chosen frequencies, it is possible to map the dependence of the electron concentration on altitude. Indeed, NZ, which increases with altitude, goes like ω_p^2, so that when ω_p is large enough for a given incident frequency, one helicity component will see a negative dielectric constant and will be reflected. By timing the arrival of the reflected wave, the altitude of the corresponding concentration can be determined (see fig. J7.11).

Such reflections off ionospheric layers explains why VHF signals, like TV, are sometimes received much further away than line of sight propagation would allow.

6.8 One-dimensional Wave Superposition, Group Velocity (section J7.8)

Fourier synthesis allows us to build pretty much any shape—finite-length wave train, pulse or what have you—out of a superposition of harmonic waves of different frequencies, each with infinite duration and extent, eg., in one dimension:

$$\begin{aligned} \psi(x,t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k)\, \mathrm{e}^{\mathrm{i}[kx-\omega(k)t]}\, \mathrm{d}k \\ A(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x,0)\, \mathrm{e}^{-\mathrm{i}kx}\, \mathrm{d}x \end{aligned} \tag{6.37}$$

The relation that gives the frequency $\omega(k) = \omega(-k)$ as a function of **mode** or wave number is the dispersion relation, one example of which we saw in the previous section. If dissipation occurs, k is complex.

The intensity of the signal is proportional to $|\psi(x)|^2$, with $|A(k)|^2$ the intensity in k-space. Define the average of a function $f(x)$ as:

$$\langle\, f(x)\,\rangle = \frac{\int f(x)\, |\psi(x)|^2\, \mathrm{d}x}{\int |\psi(x)|^2\, \mathrm{d}x}$$

where all integrals in this derivation extend from $-\infty$ to ∞. We also define the mean-square deviation as $\Delta f^2 = \left\langle \left[f - \langle\, f\,\rangle \right]^2 \right\rangle = \langle\, f^2\,\rangle - \langle\, f\,\rangle^2$.

Similarly, we can define the average of k by integrating over all k. But we can also express this average with an integral over x. First, insert a δ-function in its integral representation (0.16):

$$\langle\, k\,\rangle = \frac{\int k\, |A(k)|^2\, \mathrm{d}k}{\int |A(k)|^2\, \mathrm{d}k} = \frac{\iint k'\, A^*(k)\, A(k')\, \delta(k'-k)\, \mathrm{d}k\, \mathrm{d}k'}{\int |A(k)|^2\, \mathrm{d}k} = \frac{1}{2\pi} \frac{\iiint k'\, A^*(k)\, A(k')\, \mathrm{e}^{\mathrm{i}(k'-k)x}\, \mathrm{d}x\, \mathrm{d}k\, \mathrm{d}k'}{\int |A(k)|^2\, \mathrm{d}k}$$

With the Parseval-Plancherel theorem applied to the denominator, this can be written as:

$$\langle\, k\,\rangle = \frac{1}{\sqrt{2\pi}} \frac{\int \mathrm{d}x\, \psi^*(x) \int k'\, A(k')\, \mathrm{e}^{\mathrm{i}\,k'x}\, \mathrm{d}k}{\int |\psi(x)|^2\, \mathrm{d}x} = \frac{\int \psi^*(x)\, [-\mathrm{i}\, \mathrm{d}_x \psi(x)]\, \mathrm{d}x}{\int |\psi(x)|^2\, \mathrm{d}x}$$

In the same manner:

$$\langle\, k^2\,\rangle = -\frac{\int \psi^*(x)\, \mathrm{d}_x^2 \psi(x)\, \mathrm{d}x}{\int |\psi(x)|^2\, \mathrm{d}x}$$

Then:

$$\Delta k^2 = \frac{\int \left| \left(-\,\mathrm{i}\, \mathrm{d}_x - \langle\, k\,\rangle \right) \psi(x) \right|^2 \mathrm{d}x}{\int |\psi(x)|^2\, \mathrm{d}x}$$

represents the spread of the shape/pulse in k-space, just as Δx^2 represents its spread in position.

Next, consider the positive definite ratio:

$$r = \frac{\int |g(x)|^2\, \mathrm{d}x}{\int |\psi(x)|^2\, \mathrm{d}x}$$

where

$$g(x) = \left(\left[x - \langle\, x\,\rangle \right] - \mathrm{i}\,\eta \left[-\,\mathrm{i}\, \mathrm{d}_x - \langle\, k\,\rangle \right] \right) \psi(x)$$

with η an *arbitrary* parameter. Now we proceed to evaluate r. First, by definition:

$$\frac{\int \psi^*(x) \left(x - \langle\, x\,\rangle \right)^2 \psi(x)\, \mathrm{d}x}{\int |\psi(x)|^2\, \mathrm{d}x} = \Delta x^2$$

Also, as we have written above:

$$\eta^2 \, \frac{\int \left| \left(- \, \mathrm{i} \, \mathrm{d}_x \, - \, \langle \, k \, \rangle \right) \psi(x) \right|^2 \mathrm{d}x}{\int |\psi(x)|^2 \, \mathrm{d}x} \; = \; \eta^2 \, \Delta k^2$$

The cross-terms give:

$$\begin{aligned}
& \eta \, \frac{\int \left[\left(x - \langle \, x \, \rangle \right) \psi^* \left(\mathrm{d}_x \, - \, \mathrm{i} \cancel{\langle \, k \, \rangle} \right) \psi \, + \, \left(\mathrm{d}_x \psi^* \, + \, \mathrm{i} \cancel{\langle \, k \, \rangle} \, \psi^* \right) \left(x - \langle \, x \, \rangle \right) \psi \right] \, \mathrm{d}x}{\int |\psi(x)|^2 \, \mathrm{d}x} \\
& = \; \eta \, \frac{\int \left(x - \langle \, x \, \rangle \right) \left[\psi^* \, \mathrm{d}_x \psi \, + \, \psi \, \mathrm{d}_x \psi^* \right] \mathrm{d}x}{\int |\psi(x)|^2 \, \mathrm{d}x} \\
& = \; \eta \, \frac{\int \left(x - \cancel{\langle \, x \, \rangle} \right) \psi^* \, \mathrm{d}_x \psi \, - \, \psi^* \, \mathrm{d}_x \left[\left(x - \cancel{\langle \, x \, \rangle} \right) \psi \right] \mathrm{d}x}{\int |\psi(x)|^2 \, \mathrm{d}x} \\
& = \; - \, \eta
\end{aligned}$$

where the next-to-last line is obtained by integrating the term with $\mathrm{d}_x \psi^*$ in the previous one by parts. So $r = \Delta x^2 \, + \, \eta^2 \, \Delta k^2 \, - \, \eta \, \geq 0$, with η an arbitrary parameter that we can choose so as to minimise r. Setting $\mathrm{d}_\eta r$ equal to zero, we find $\eta = [2 \Delta k^2]^{-1}$. Then:

$$r_{\mathrm{min}} \; = \; \Delta x^2 \; - \; \frac{1}{4 \, \Delta k^2} \; \geq 0$$

or

$$\Delta x \, \Delta k \; \geq \; 1/2 \tag{6.38}$$

A similar derivation shows that:

$$\Delta t \, \Delta \omega \; \geq \; 1/2 \tag{6.39}$$

Thus, for a given spread in k-space or frequency space (bandwith), the spread of the signal in position or time is bounded from below. We should realise that this important result applies to any kind of wave, not only electromagnetic waves.

We can also address the behaviour of a pulse or finite wave train as it propagates. If its bandwidth is fairly narrow, or if the dependence of frequency on wave number (given by the dispersion relation) is not too fast, then $\omega(k)$ can be expanded around the central value k_0 which dominates the spectrum:

$$\omega(k) \; = \; \omega_0 \; + \mathrm{d}_k \omega \big|_{k_0} \, (k - k_0) \; + \; \ldots$$

When inserted into eq. (6.37), this yields:

$$\begin{aligned}
\psi(x,t) \; &= \; \frac{\mathrm{e}^{\mathrm{i}[k_0 \mathrm{d}_k \omega - \omega_0] t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k) \, \mathrm{e}^{\mathrm{i} k [x - \mathrm{d}_k \omega \, t]} \, \mathrm{d}k \\
&= \; \mathrm{e}^{\mathrm{i}[k_0 \mathrm{d}_k \omega - \omega_0] t} \, \psi(x - v_{\mathrm{g}} \, t \, , \, 0)
\end{aligned}$$

where we have introduced the **group velocity**:

$$v_{\mathrm{g}} \; = \; \mathrm{d}_k \omega \big|_0 \tag{6.40}$$

which is in general different from the phase velocity $v_{\mathrm{ph}} = \omega(k)/k = c/n(k)$. Since the energy density goes like $|\psi|^2$, we see that it also propagates at the group velocity. Also, the pulse itself propagates undistorted.

Written in terms of n, the group velocity becomes:

$$v_{\rm g} \;=\; c\,{\rm d}_k(k/n)\big|_0 \;=\; \frac{c}{n}\big(1 \,-\, \frac{k}{n}\,v_{\rm g}\,{\rm d}_\omega n\big) \qquad\Longrightarrow\qquad v_{\rm g} \;=\; \frac{c}{n(\omega) \,+\, \omega\,{\rm d}_\omega n} \tag{6.41}$$

where we have used $v_{\rm g}\,{\rm d}_\omega n = {\rm d}_k n$ to express $v_{\rm g}$ in terms of $n(\omega)$.

In the case of normal dispersion, ${\rm d}_\omega n > 0$, and the group velocity is smaller than the phase velocity. When dispersion is anomalous, however, the negative ${\rm d}_\omega n$ causes the group velocity to be larger; in extreme cases, we could even get $v_{\rm g} > c$ (see figure J7.14) and $v_{\rm g} < 0$! Before we get all excited about this, however, we note that eq. (6.41) was obtained under assumptions which do not hold when dispersion is anomalous as n then decreases quickly with frequency. Although one can contrive experiments in which the peak of a pulse emerges from a thin absorber before the peak of the incident pulse has entered it ($v_{\rm g} < 0$), no information can be transmitted this way. One of our faculty, Aephraim Steinberg, has been heavily involved in such experiments—see reference cited in Jackson.

6.9 Causality and the Relation between D and E - Kramers-Kronig Relations (section J7.10)

6.9.1 Non-locality in time

The response of a medium to a field cannot be instantaneous. When a field is turned on, it must take some time for the electrons and atoms to respond (after all, they have mass!); when the field is turned off, it takes time for the medium to relax back to its previous state. How can we incorporate this property in our formalism?

In section 6.8, when obtaining the Poynting theorem a dispersive medium, we already used a Fourier decomposition of the time dependence of $\mathbf{E}$ and $\mathbf{D}$ over frequency space:

$$\begin{aligned}\mathbf{D}(\mathbf{x},t) \;&=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathbf{D}(\mathbf{x},\omega)\,{\rm e}^{-{\rm i}\omega t}\,{\rm d}\omega \\ \mathbf{D}(\mathbf{x},\omega) \;&=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathbf{D}(\mathbf{x},t')\,{\rm e}^{{\rm i}\omega t'}\,{\rm d}t'\end{aligned}$$

with $\mathbf{D}(\mathbf{x},\omega) = \epsilon(\omega)\,\mathbf{E}(\mathbf{x},\omega)$, where $\epsilon(\omega) = \epsilon_0\big(1+\chi(\omega)\big)$. Thus, we can write:

$$\begin{aligned}\mathbf{D}(\mathbf{x},t) \;&=\; \frac{1}{2\pi}\int_{-\infty}^{\infty}{\rm d}\omega\,\epsilon(\omega)\,{\rm e}^{-{\rm i}\omega t}\int_{-\infty}^{\infty}{\rm d}t'\,{\rm e}^{{\rm i}\omega t'}\,\mathbf{E}(\mathbf{x},t') \\ &=\; \epsilon_0\,\mathbf{E}(\mathbf{x},t) \;+\; \epsilon_0\int_{-\infty}^{\infty}G(\tau)\,\mathbf{E}(\mathbf{x},t-\tau)\,{\rm d}\tau \qquad (\tau = t-t')\end{aligned}$$

where the **kernel**, or **response function**, is defined as:

$$G(\tau) \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty}\chi(\omega)\,{\rm e}^{-{\rm i}\omega\tau}\,{\rm d}\omega \tag{6.42}$$

We can now identify the term containing the kernel with the polarisation vector $\mathbf{P}$, which provides a constitutive relation between the response of the material and the field. We see that $\mathbf{D}(t)$ depends on the whole history of $\mathbf{E}$, and that the frequency-domain formalism provides non-locality in time (but not in space) in a natural way (see J7.10B for a discussion of the time scales involved, and on how far we can push our implicit assumption of locality in *space*). But there is a problem when $\tau \le 0$, or $t' \ge t$: the response precedes the stimulus...

6.9.2 Causality and analyticity of $\chi(\omega)$

If we impose causality, however, only values of the field *before* t can contribute to $\mathbf{P}(t)$. Thus, $G(\tau \leq 0) = 0$, and we arrive at a very general relation:

$$\mathbf{P}(\mathbf{x}, t) \;=\; \epsilon_0 \int_0^\infty G(\tau)\, \mathbf{E}(\mathbf{x}, t-\tau)\, \mathrm{d}\tau \tag{6.43}$$

with the frequency dependence of the susceptibility given by:

$$\chi(\omega) \;=\; \int_0^\infty G(\tau)\, \mathrm{e}^{\mathrm{i}\omega\tau}\, \mathrm{d}\tau \tag{6.44}$$

Some very general information about the behaviour of the susceptibility can be derived by postulating that (1) $G(\tau)$ and its derivatives go to 0 as $\tau \to \infty$, (2) that $G(\tau)$ is bounded everywhere, as well as (3) that $G(\tau \leq 0) = 0$ (causality).

Now we can show, using contour integration, that $\chi(\omega)$ has no pole in the upper half of the complex-ω plane. If a function $f(\omega)$ is analytic inside a contour C, the Cauchy-Goursat Theorem asserts that:

$$\oint_{\mathrm{C}} f(\omega)\, \mathrm{d}\omega \;=\; 0$$

Conversely, let us evaluate the integral for $G(\tau)$ in eq. (6.42) over a contour enclosing the whole upper half-plane:

$$\oint_{\mathrm{C}} \chi(\omega)\, \mathrm{e}^{-\mathrm{i}\omega\tau}\, \mathrm{d}\omega \;=\; \int_{-\infty}^{\infty} \chi(\omega)\, \mathrm{e}^{-\mathrm{i}\omega\tau}\, \mathrm{d}\omega \;+\; \int_{R\to\infty} \chi(\omega)\, \mathrm{e}^{-\mathrm{i}\omega\tau}\, \mathrm{d}\omega$$

Since we can choose any value of τ, let $\tau < 0$. Then the first term on the right, which is just $G(\tau)$ (up to 2π), vanishes because of causality. The integrand in the second term is bounded by $|\chi|\mathrm{e}^{\Im(\omega)\tau}$. Because here $\Im(\omega) > 0$, the integrand vanishes as $R \to \infty$ when $\tau < 0$, and the contribution from the semi-circle at infinity also vanishes. From complex analysis, the vanishing of the contour integral means that either $\chi(\omega)$ has no singularities in the upper half-plane, or else their contributions cancel out. But differentiate eq. (6.44) n times and note that the resulting integrand remains bounded when $\tau > 0$ and $\Im(\omega) > 0$. On the real ω axis, $g(\omega)$ is bounded, and any branch point can be bypassed without changing anything. Because a function whose n^{th}-order derivative exists $\forall\, n$ must be analytic, this establishes that $\chi(\omega)$ has no singularities in the upper half-plane when causality is imposed. In other words, causality implies analyticity!

Moreover, if we repeatedly integrate the right-hand side of eq. (6.44) by parts, we obtain:

$$\chi(\omega) \;\sim\; \frac{\mathrm{i}}{\omega}\, G(0) \;-\; \frac{1}{\omega^2}\, \mathrm{d}_\tau G\big|_{\tau=0} \;+\; \ldots$$

assuming that G and its derivatives vanish at infinity. The first term vanishes from causality, and we find that:

$$\Re[\chi(\omega)] \underset{\omega\to\infty}{\sim} \mathrm{O}(1/\omega^2) \qquad\qquad \Im[\chi(\omega)] \underset{\omega\to\infty}{\sim} \mathrm{O}(1/\omega^3)$$

which shows that $|\chi(\omega)| \to 0$ as $\omega \to \infty$, an important result which will be useful shortly.

We saw in section 6.6.2 that at high frequency the susceptibility may be written (eq. (6.34)):

$$\chi(\omega) \;=\; -\,\frac{\omega_{\mathrm{p}}^2}{\omega^2}$$

This result, which was derived in the context of a specific model (eq. (6.30)), is now seen to have more general validity, and may be used to *define* the plasma frequency as:

$$\omega_{\mathrm{p}}^2 \;=\; -\lim_{\omega\to\infty} \omega^2\, \chi(\omega) \tag{6.45}$$

6.9.3 Kramers-Kronig Dispersion Relations

Because $\chi(\omega)$ is analytic in the upper ω half-plane, the Cauchy Theorem can be used to relate its real and imaginary parts at *real* frequencies..

First, consider the integral of the function $\chi(\omega')/(\omega'-\omega)$ over a closed counter-clockwise path (contour) in the complex ω' plane, consisting of a half-circle of infinite radius $R=|\omega'|$ in the upper half-plane. centered on the origin, with its base along the real axis. $\chi(\omega')$ is analytic on and within the contour, but $\chi(\omega')/(\omega'-\omega)$ has a singularity at ω on the real axis, so *it* is not analytic on the contour.

A strategy to make sense of this integral is to "avoid" ω along a semi-circle C_δ of radius δ in the upper half-plane, centered on ω, whose diameter extends from $\omega-\delta$ to $\omega+\delta$. Since now $\chi(\omega')/(\omega'-\omega)$ is analytic everywhere within and on the closed path we have chosen, the contour integral vanishes by the Cauchy-Goursat theorem, and we end up with:

$$0 \;=\; \int_{-\infty}^{\omega-\delta} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega' \;+\; \int_{\omega+\delta}^{\infty} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega' \;+\; \int_{C_\delta} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega' \;+\; \int_{R\to\infty} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega' \tag{6.46}$$

The contribution from the upper half-circle vanishes[†] because of the fall-off of the integrand at infinity guaranteed by the fact that $|\chi(\omega')|\to 0$ at infinity, as we have shown in section 6.9.2.

A corollary of the Cauchy-Goursat Theorem says that since $\chi(\omega')$ is analytic at ω, the integral of $\chi(\omega')/(\omega'-\omega)$ on a counter-clockwise circular arc C_δ of radius δ, centered at ω and intercepting an angle α, is:

$$\int_{\mathrm{C}_\delta} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega' \;=\; \alpha\,\mathrm{i}\,\chi(\omega)$$

Here $\alpha=-\pi$ (C_δ is traversed *clockwise*), and:

$$\int_{\mathrm{C}_\delta} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega' \;=\; -\mathrm{i}\,\pi\,\chi(\omega)$$

Now we let $\delta\to 0$. Then the first two terms in eq. (6.46) are called the **principal value integral**:

$$\mathcal{P}\int_{-\infty}^{\infty} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega' \;\equiv\; \lim_{\delta\to 0}\int_{-\infty}^{\omega-\delta} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega' \;+\; \lim_{\delta\to 0}\int_{\omega+\delta}^{\infty} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega' \tag{6.47}$$

Eq. (6.46) relates this object to the susceptibility evaluated at ω:

$$\chi(\omega) \;=\; \frac{1}{\mathrm{i}\,\pi}\,\mathcal{P}\int_{-\infty}^{\infty} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega'$$

The i in front of the right-hand side works its magic when we separate out the real and imaginary parts:

$$\begin{aligned} \Re[\chi(\omega)] \;&=\; \frac{1}{\pi}\,\mathcal{P}\int_{-\infty}^{\infty} \frac{\Im[\chi(\omega')]}{\omega'-\omega}\,\mathrm{d}\omega' \\ \Im[\chi(\omega)] \;&=\; -\frac{1}{\pi}\,\mathcal{P}\int_{-\infty}^{\infty} \frac{\Re[\chi(\omega')]}{\omega'-\omega}\,\mathrm{d}\omega' \end{aligned} \tag{6.48}$$

These relations are examples of **Hilbert transforms**. If that principal value causes discomfort, the following manipulation gets rid of it, by making use of the relation, which holds on the real ω axis:

$$\mathcal{P}\int_{-\infty}^{\infty} \frac{\chi(\omega')}{\omega'-\omega}\,\mathrm{d}\omega' \;=\; \mathcal{P}\int_{-\infty}^{\infty} \frac{\chi(\omega')\;-\;\chi(\omega)}{\omega'-\omega}\,\mathrm{d}\omega'$$

[†]For the detailed calculation see section 4.5 in http://www.physics.utoronto.ca/~phy1540f/p154019_cumullect.pdf.

Now, because χ is analytic on the real axis, it is differentiable at ω, and we can drop the $\mathcal{P}$ symbol in the right-hand side. Then the Hilbert transforms become:

$$\Re[\chi(\omega)] \;=\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\Im[\chi(\omega')]-\Im[\chi(\omega)]}{\omega'-\omega}\,\mathrm{d}\omega', \qquad \Im[\chi(\omega)] \;=\; -\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\Re[\chi(\omega')]-\Re[\chi(\omega)]}{\omega'-\omega}\,\mathrm{d}\omega' \tag{6.49}$$

With a little extra manipulation, we can rewrite the Hilbert transforms as integrals over positive frequencies. To do this, we must use the fact, already established back in section 6.8 for $\epsilon(\omega)$, that $\mathbf{E}(\mathbf{x},t)\in\mathbb{R}^3$ and $\mathbf{D}(\mathbf{x},t)\in\mathbb{R}^3$ demands that $\chi(-\omega)=\chi^*(\omega)$. In terms of real and imaginary parts, we have the **crossing relations**:

$$\Re[\chi(-\omega)] \;=\; \Re[\chi(\omega)] \qquad\qquad \Im[\chi(-\omega)] \;=\; -\,\Im[\chi(\omega)] \tag{6.50}$$

Therefore, with a change of variable $\omega'\to-\omega'$ over the interval $(-\infty,0)$, there comes::

$$\begin{aligned}\Re[\chi(\omega)] \;&=\; \frac{1}{\pi}\,\mathcal{P}\left[\int_0^{\infty}\frac{\Im[\chi(-\omega')]}{-\omega'-\omega}\,\mathrm{d}\omega' \;+\; \int_0^{\infty}\frac{\Im[\chi(\omega')]}{\omega'-\omega}\,\mathrm{d}\omega'\right]\\ &=\; \frac{1}{\pi}\,\mathcal{P}\left[\int_0^{\infty}\frac{\Im[\chi(\omega')]}{\omega'+\omega}\,\mathrm{d}\omega' \;+\; \int_0^{\infty}\frac{\Im[\chi(\omega')]}{\omega'-\omega}\,\mathrm{d}\omega'\right]\\ &=\; \frac{2}{\pi}\,\mathcal{P}\int_0^{\infty}\frac{\omega'\,\Im[\chi(\omega')]}{\omega'^2-\omega^2}\,\mathrm{d}\omega'\end{aligned}$$

$\Im[\chi(\omega)]$ can be rewritten in the same way, and we obtain the **Kramers-Kronig dispersion relations**:

$$\begin{aligned}\Re[\chi(\omega)] \;&=\; \frac{2}{\pi}\,\mathcal{P}\int_0^{\infty}\frac{\omega'\,\Im[\chi(\omega')]}{\omega'^2-\omega^2}\,\mathrm{d}\omega'\\ \Im[\chi(\omega)] \;&=\; -\,\frac{2\,\omega}{\pi}\,\mathcal{P}\int_0^{\infty}\frac{\Re[\chi(\omega')]}{\omega'^2-\omega^2}\,\mathrm{d}\omega'\end{aligned} \tag{6.51}$$

This type of relation, which can be derived with a minimum of assumptions (causality), can be very useful and exists in other areas of physics (it was often used in particle physics in the early sixties, for instance). Similar expressions exist between the real and imaginary parts of the index of refraction $n(\omega)=\sqrt{1+\chi(\omega)}$. Provided $\chi(\omega)\ll 1$, $n(\omega)=\approx 1+\chi(\omega)/2$, and:

$$\begin{aligned}\Re[n(\omega)] \;&=\; 1 \;+\; \frac{1}{\pi}\,\mathcal{P}\int_0^{\infty}\frac{\omega'\,\Im[n(\omega')]}{\omega'^2-\omega^2}\,\mathrm{d}\omega'\\ \Im[n(\omega)] \;&=\; -\,\frac{\omega}{\pi}\,\mathcal{P}\int_0^{\infty}\frac{\Re[n(\omega')]}{\omega'^2-\omega^2}\,\mathrm{d}\omega'\end{aligned} \tag{6.52}$$

Useful so-called **sum rules** can also be easily derived from the Kramers-Kronig relations: see eq. J7.122 and J7.123 for examples.

6.10 A Constraint on the Speed of a Signal Through a Dispersive Medium (section J7.11)

We consider a plane wave train at normal incidence from vacuum to a nonpermeable medium of refraction index $n(\omega)$. Then the electric field inside the medium ($x>0$) is, using Fresnel's equation (6.20),

$$E_{\mathrm{r}}(x,t) \;=\; \int_{-\infty}^{\infty}\left(\frac{2}{1+n(\omega)}\right)\,A_{\mathrm{i}}(\omega)\,\mathrm{e}^{\mathrm{i}\omega[n(\omega)x/c-t]}\,\mathrm{d}\omega$$

where

$$A_{\mathrm{i}}(\omega) \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty}E_{\mathrm{i}}(0,t)\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t$$

is the Fourier transform of the real incident field at $x = 0^-$, just outside the medium.

The integral for E_r can be evaluated by contour integration in the complex ω plane. The integral enclosing the whole upper half-plane vanishes because the integrand is analytic there. Also, we saw in section 6.9.2 that $\chi(\omega) \to 0$ when $|\omega| \to \infty$ (even if causality is not imposed!). Then $n(\omega) = \sqrt{1+\chi(\omega)} \to 1$ and the argument of the exponential becomes $\mathrm{i}\,\omega[x - ct]/c$, so that the contribution from the semi-circle at infinity also vanishes if $x > ct$. Then the contribution along the real axis must also vanish, and there is no electric field amplitude for $x > ct$. Without any specific knowledge of $n(\omega)$, analyticity does not allow a signal to propagate faster than c in any medium. This provides evidence that Maxwell's theory is consistent with Special Relativity.

6.11 Spreading of a Propagating Pulse (section J 7.9)

When we wrote our general expression for a propagating shape in eq. (6.37), we did not bother about the fact that since this is a solution of second-order differential equation, two pieces of initial data must be given for the initial-value problem to be well defined. Now we must be more careful, and write the *general* solution for $\psi(x,t)$:

$$\psi(x,t) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[A(k)\,\mathrm{e}^{\mathrm{i}[kx-\omega(k)t]} \;+\; B(k)\,\mathrm{e}^{-\mathrm{i}[kx-\omega(k)t]}\right] \mathrm{d}k$$

Consistency demands that $B(k) = A^*(k)$ if $\psi(x,t)$ is to be real.. Then renormalising $A(k) \longrightarrow \frac{1}{2}A(k)$ gives:

$$\psi(x,t) \;=\; \frac{1}{2}\left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k)\,\mathrm{e}^{\mathrm{i}[kx-\omega(k)t]}\,\mathrm{d}k \;+\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A^*(k)\,\mathrm{e}^{-\mathrm{i}[kx-\omega(k)t]}\,\mathrm{d}k\right]$$

Write $A(k)$ in terms of $\psi(x,0)$ and $\partial_t\psi(x,t)\big|_0$. Now: $\frac{1}{2}\left[(\psi(x,0) + \mathrm{i}\partial_t\psi(x,t)\big|_{t=0}/\omega(k)\right] = \int_{-\infty}^{\infty} A(k)\,\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k$, leading to:

$$A(k) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[\psi(x,0) \;+\; \frac{\mathrm{i}}{\omega(k)}\,\partial_t\psi(x,t)\Big|_{t=0}\right] \mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x$$

To illustrate this, choose an oscillation with a Gaussian envelope at $t = 0$: $\psi(x,0) \;=\; \Re\left[\mathrm{e}^{-x^2/2L^2}\,\mathrm{e}^{\mathrm{i}k_0x}\right]$.

L is the width of the envelope, in the sense that $\Delta x \;=\; L$. To simplify manipulations, we choose $\partial_t\psi(x,t)\Big|_{t=0} = 0$. This can only happen if $\psi(x,0)$ is in fact a superposition of two identical pulses travelling at equal speed in opposite directions. From above, we then find:

```
> Int(exp(-x^2/2/L^2)*cos(k0*x)*exp(-I*k*x),x=-infinity..infinity)*
> (1/sqrt(2*Pi))=factor(combine(expand((1/sqrt(2*Pi))*int(exp(-x^2/2/L^2)*
> cos(k0*x)*exp(-I*k*x),x=-infinity..infinity)))) assuming L>0;
```

$$1/2\,\sqrt{2}\int_{-\infty}^{\infty} \mathrm{e}^{-1/2\,\frac{x^2}{L^2}}\mathrm{e}^{-ikx}\,\cos\left(k0\,x\right)dx\left(\sqrt{\pi}\right)^{-1} = 1/2\,L\left(\mathrm{e}^{-1/2\,(k-k0)^2L^2} + \mathrm{e}^{-1/2\,(k+k0)^2L^2}\right)$$

Next, we need a dispersion relation for $\omega(k)$. We take the one obtained in example 6.2 for a tenuous plasma:

$$\omega \;=\; \omega_\mathrm{p}\sqrt{1+(ck/\omega_\mathrm{p})^2} \;\approx\; \omega_\mathrm{p}\left[1 \;+\; \frac{1}{2}\,\frac{c^2k^2}{\omega_\mathrm{p}^2}\right]$$

which leads to the group velocity:

$$v_\mathrm{g} \;=\; \frac{\mathrm{d}\omega}{\mathrm{d}k}\Big|_{k_0} \;=\; \frac{c^2\,k_0/\omega_\mathrm{p}}{\sqrt{1+(c\,k_0/\omega_\mathrm{p})^2}}$$

To leading order around k_0, this is:

$$v_\mathrm{g} \;\approx\; \omega_\mathrm{p}\,\frac{c^2\,k_0}{\omega_\mathrm{p}^2} \;=\; \omega_\mathrm{p}\,\frac{\lambda_\mathrm{p}^2}{4\pi^2}\,k_0 \tag{6.53}$$

where $\lambda_{\rm p}$ is the incident wavelength around which dispersive effects are expected to become important.

To make contact with Jackson's notation, just set $\omega_{\rm p} = \nu$ and $c/\omega_{\rm p} = \lambda/2\pi = a$. Also, define dimensionless quantities: $\xi = x/L$, $\kappa = kL$, and $\alpha = a/L$. Inserting $A(k)$ and ω in our expression for $\psi(x,t)$ then yields:

$$\psi(x,t) \;=\; \frac{1}{\sqrt{8\pi}}\,\Re\left[\int_{-\infty}^{\infty}\left[\mathrm{e}^{-(\kappa-\kappa_0)^2/2} \;+\; \mathrm{e}^{-(\kappa+\kappa_0)^2/2}\right]\mathrm{e}^{\mathrm{i}\kappa\xi - \mathrm{i}\nu t[1+\alpha^2\kappa^2/2]}\,\mathrm{d}\kappa\right] \tag{6.54}$$

Now concentrate on the first term. We can either expand it to:

$$\frac{1}{\sqrt{8\pi}}\,\Re\left\{\mathrm{e}^{-\mathrm{i}\nu t-\kappa_0^2/2}\int_{-\infty}^{\infty}\exp\left[-\kappa^2\,\frac{1+\mathrm{i}\,\nu\,t\,\alpha^2}{2} \;+\; \kappa\,(\mathrm{i}\,\xi \;+\; \kappa_0)\right]\,\mathrm{d}\kappa\right\}$$

and, noticing that we have an integrand of the form $\mathrm{e}^{-(a\,\kappa^2 - b\,\kappa)} \;=\; \mathrm{e}^{-(a\,-b/2a)^2}\mathrm{e}^{b^2/4\alpha}$, perform the change of variables $y = \sqrt{a}\,(\kappa - b/2a)$, so that the term becomes:

$$\frac{1}{\sqrt{8\pi}}\,\Re\left\{\frac{\sqrt{2}}{\left(1+\mathrm{i}\nu t\alpha^2\right)^{1/2}}\,\exp\left[\frac{(\mathrm{i}\,\xi+\kappa_0)^2}{2(1+\mathrm{i}\nu t\alpha^2)} - \frac{\kappa_0^2}{2}\right]\,\mathrm{e}^{-\mathrm{i}\nu t}\int_{-\infty}^{\infty}\mathrm{e}^{-y^2}\,\mathrm{d}y\right\}$$

where the integral is just $\sqrt{\pi}$. Or, more simply, just give the initial integral to ***Maple***:

```
> int(exp(-(kappa-kappa0)^2/2)*exp(I*(kappa*xi -nu*t*kappa^2*lambda^2/2)),
> kappa=-infinity..infinity):  Int(exp(-(kappa-kappa0)^2/2)*
> exp(I*(kappa*xi -nu*t*kappa^2*lambda^2/2)),kappa=-infinity..infinity)
> =simplify(op(2,value(%))/sqrt(8*Pi),symbolic);
```

$$\int_{-\infty}^{\infty} e^{-1/2\,(\kappa-\kappa_0)^2}\,e^{i(\kappa\,\xi - 1/2\,\nu\,t\kappa^2\alpha^2)}d\kappa = 1/2\,e^{\frac{1/2\,i\left(-\kappa_0^2\nu\,t\alpha^2+2\,\kappa 0\,\xi+i\xi^2\right)}{1+i\nu\,t\alpha^2}}\,\frac{1}{\sqrt{1+i\nu\,t\alpha^2}}$$

Therefore the first term in $u(x,t)$ is:

$$\frac{1}{2}\,\Re\left\{\frac{1}{\left(1+\mathrm{i}\nu t\alpha^2\right)^{1/2}}\,\exp\left[\frac{-\xi^2 + 2\mathrm{i}\kappa_0\xi - \mathrm{i}\,\kappa_0^2\nu t\alpha^2}{2(1+\mathrm{i}\nu t\alpha^2)}\right]\,\mathrm{e}^{-\mathrm{i}\nu t}\right\}$$

The rest is "merely" a matter of extracting $\mathrm{i}(\kappa_0\xi - \nu t\kappa_0^2\alpha^2/2)$ in the argument of the big exponential. With $v_{\rm g} = \nu\kappa_0\alpha L$, a few lines of algebra lead to:

$$\psi(x,t) \;=\; \frac{1}{2}\,\Re\left\{\frac{1}{\left(1+\mathrm{i}\nu t\alpha^2\right)^{1/2}}\,\exp\left[-\frac{(x-v_{\rm g}\,t)^2}{2L^2\left(1+\mathrm{i}\nu t\alpha^2\right)}\right]\,\mathrm{e}^{\mathrm{i}k_0 x - \mathrm{i}\nu t[1+k_0^2a^2/2]} \;+\; (k_0 \longrightarrow -\,k_0)\right\} \tag{6.55}$$

where we have added the result for the second term which follows immediately from the form of eq. (6.54).

Admittedly, this expression is still somewhat obscure because of all the complex factors. But a few more lines of manipulation show that it can be rewritten as:

$$\psi(x,t) \;=\; \frac{1}{\left(1+\nu^2t^2a^4/L^4\right)^{1/4}}\,\mathrm{e}^{(x-v_{\rm g}t)^2/2L^2(t)}\,\cos(k_0x - \omega_0 t + \theta) \;+\; (k_0 \longrightarrow -\,k_0) \tag{6.56}$$

where ω_0 and θ are complicated quantities irrelevant for our purpose. What is important is that we still have Gaussian pulses travelling at speed $v_{\rm g}$, but whose width is now:

$$L(t) \;=\; L\left[1 \;+\; \left(\nu t a^2/L^2\right)^2\right]^{1/2} \tag{6.57}$$

Thus, the pulses spread out as they propagate. The smaller the initial width compared to a (wavelength at which dispersion becomes noticeable), the faster the spreading.

The Parseval-Plancherel theorem tells us that $\psi(x,t)$ and its Fourier transform, $A'(k,t) = A(k)\,\mathrm{e}^{-\mathrm{i}\omega(k)t}$, have the same norm. Therefore:

$$\begin{aligned}\int \left|\psi(x,t)\right|^2 \mathrm{d}x &= \int \left|A'(k,t)\right|^2 \mathrm{d}k \\ &= \int \left|A(k,0)\right|^2 \mathrm{d}k \\ &= \int \left|\psi(x,0)\right|^2 \mathrm{d}x\end{aligned}$$

Since $\int |\psi(x,t)|^2\,\mathrm{d}x$ is a measure of the energy of the pulse, we see that this energy is constant in time in the absence of dissipative effects.

Finally, this phenomenon of spreading of signals in dispersive media is very general, and the width always increases with time.

7 Radiating Systems

7.1 Fields of a Localised Harmonic Source (sections J9.1 and J9.6)

7.1.1 General results

In section 5.5, we wrote down the retarded solution for the vector potential, eq. (5.20), in the absence of boundary surfaces:

$$\mathbf{A}(\mathbf{x},t) \;=\; k_{\rm m}\int \mathrm{d}^3x'\,\frac{\mathbf{J}(\mathbf{x}',t'_{\rm ret})}{R}$$

with retarded time $t'_{\rm ret}=t-R/c$, $R=|\mathbf{x}-\mathbf{x}'|$. This solution was obtained by imposing the Lorenz condition.

We continue to assume that sources, potentials and fields are harmonic, ie. $\mathbf{J}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t'} = \mathbf{J}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{e}^{\mathrm{i}\omega R/c}$, with ω the same for the three components, although their relative phase can be different. Then the time-oscillating factors in t on both sides cancel out, and with $k=\omega/c$, we have:

$$\mathbf{A}(\mathbf{x}) \;=\; k_{\rm m}\int \mathbf{J}(\mathbf{x}')\,\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\,\mathrm{d}^3x' \tag{7.1}$$

When $\mathbf{A}$ has been found, the harmonic fields in vacuum immediately follow from:

$$\mathbf{H} \;=\; \frac{1}{\mu_0}\boldsymbol{\nabla}\times\mathbf{A} \qquad\qquad \mathbf{E} \;=\; \mathrm{i}\,\frac{Z_0}{k}\,\boldsymbol{\nabla}\times\mathbf{H} \;=\; \mathrm{i}\,\frac{c}{k}\left[\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) \;-\; \nabla^2\mathbf{A}\right] \tag{7.2}$$

where $Z_0=\sqrt{\mu_0/\epsilon_0}$.

Alternatively, this expression for harmonic electric fields in terms of the vector potential can be derived from the Lorenz condition, $\boldsymbol{\nabla}\cdot\mathbf{A}+\partial_t\Phi/c^2=0$, by eliminating the scalar potential from $\mathbf{E}=-\partial_t\mathbf{A}-\boldsymbol{\nabla}\Phi$. Indeed:

$$\Phi(\mathbf{x}) \;=\; -\,\mathrm{i}\,\frac{c}{k}\,\boldsymbol{\nabla}\cdot\mathbf{A}(\mathbf{x})$$

Thus, the electric field *everywhere* can be found without the need for the scalar potential:

$$\mathbf{E}(\mathbf{x}) \;=\; \mathrm{i}\,\omega\mathbf{A} \;+\; \mathrm{i}\,\frac{c}{k}\,\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) \;=\; \mathrm{i}\,\frac{c}{k}\left[\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) \;+\; k^2\mathbf{A}\right] \tag{7.3}$$

7.1.2 A useful expansion of the vector potential of a localised source

Henceforth we will take the point of observation to be sufficiently far from the source that $kr\gg kr'$. Approximating $R\approx\; r-\hat{\mathbf{n}}\cdot\mathbf{x}'$ and $1/R\approx(1/r)[1+\hat{\mathbf{n}}\cdot\mathbf{x}'/r]$, the potential in eq. (7.1) becomes, to leading order:

$$\mathbf{A}(\mathbf{x}) \;=\; k_{\rm m}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\int \mathbf{J}(\mathbf{x}')\,\mathrm{e}^{-\mathrm{i}k\hat{\mathbf{n}}\cdot\mathbf{x}'}\,\mathrm{d}^3x' \tag{7.4}$$

In this approximation, the source is localised in a small (compared to r) volume around the origin. The r dependence resides entirely in the $\mathrm{e}^{\mathrm{i}kr}/r$ factor; the integral depends only on angles in $\hat{\mathbf{n}}\cdot\mathbf{x}'=r'\cos\gamma$.

Most often, the integral in eq. (7.1) — or even the one for localised sources in in eq. (7.4) — cannot be evaluated in closed form. To extract information from it, recall that the troublesome factor $\mathrm{e}^{\mathrm{i}kR}/R$ in the integrand was found in section 5.4 to be the Green function for the Helmholtz operator ∇^2+k^2. We now obtain an expansion of this factor in spherical coordinates. Following the method used in sections 0.8.3 and 2.3 above (section J3.9) for the Laplacian operator, we write:

$$G(\mathbf{x},\mathbf{x}') \;=\; \sum_{l=0}^{\infty}\sum_{m=-l}^{l} g_l(r,r')\,Y^*_{lm}(\theta',\phi')Y_{lm}(\theta,\phi)$$

Inserting into $(\nabla^2 + k^2)G = -4\pi\delta(\mathbf{R})$, we immediately find that $g_l(r, r')$ must satisfy the radial equation:

$$\left[\mathrm{d}_r^2 \,+\, \frac{2}{r}\,\mathrm{d}_r \,+\, k^2 \,-\, \frac{l(l+1)}{r^2}\right] g_l(r,r') \;=\; -\,\frac{4\pi}{r^2}\,\delta(r-r')$$

The homogeneous radial equation is known as the **spherical Bessel equation**, and its solutions are the spherical Bessel functions $j_l(x)$ and $n_l(x)$ introduced in section 0.8.3. More properties of these functions are discussed in Mathematical Handbooks, and the most important are summarised on pp. J426–427.

The solution of the inhomogeneous equation that is finite at the origin is:

$$g_l(r,r') \;=\; 4\pi\,\mathrm{i}\,k\,j_l(kr_<)\,h_l^{(1)}(kr_>)$$

where l is an integer and A can be shown to be equal to $\mathrm{i}\,k$, and the $h_l^{(1)} = j_l + \mathrm{i}n_l$ are Hankel functions of the first kind. Our spherical Green function expansion is thus:

$$\frac{\mathrm{e}^{\mathrm{i}kR}}{R} \;=\; 4\pi\,\mathrm{i}\,k\sum_{l=0}^{\infty} j_l(kr_<)\,h_l^{(1)}(kr_>)\sum_{m=-l}^{l} Y_{lm}^*(\theta',\phi')Y_{lm}(\theta,\phi) \tag{7.5}$$

Inserting this into eq. (7.1) yields, for $\mathbf{x}$ outside the sources ($r_> = r$, $r_< = r'$):

$$\mathbf{A}(\mathbf{x}) \;=\; 4\pi\,k_\mathrm{m}\,\mathrm{i}\,k\sum_{l,m} h_l^{(1)}(kr)\,Y_{lm}(\theta,\phi)\int \mathbf{J}(\mathbf{x}')\,j_l(kr')\,Y_{lm}^*(\theta',\phi')\,\mathrm{d}^3x' \tag{7.6}$$

This expression is still exact, with no approximation. Eq. (J9.86) for j_l and n_l becomes, in terms of the spherical Hankel functions:

$$\begin{aligned} h_l^{(1)}(x) \;&=\; -\,\mathrm{i}\,(-x)^l\left[\frac{\mathrm{d}_x}{x}\right]^l\left(\frac{\mathrm{e}^{\mathrm{i}x}}{x}\right) \\ &=\; -\,\mathrm{i}\,(2l-1)!!\,\frac{\mathrm{e}^{\mathrm{i}x}}{x^{l+1}}\left(1 \,+\, a_1\,(\mathrm{i}\,x) \,+\, a_2\,(\mathrm{i}\,x)^2 \,+\, \ldots \,+\, a_l\,(\mathrm{i}\,x)^l\right) \end{aligned}$$

where the a_i are known coefficients.

Now we assume that the dimensions of the source are much smaller than the characteristic length of the potential (wavelength $\lambda = 2\pi c/\omega$) (**long-wavelength approximation**), In particular, $kr' \ll 1$ in all that follows. But nothing more about kr is assumed at this stage than $kr > kr'$! Then we can use eq. (J9.88) which gives the small x limit of $j_l(x)$; we keep only the leading order, $x^l/(2l+1)!!$. Inserting the low x behaviour of $j_l(kr')$ and our expression for $h_l^{(1)}(kr)$ in our expansion for the vector potential of a localised source $\mathbf{A}(\mathbf{x})$ gives:

$$\mathbf{A}(\mathbf{x}) \;=\; k_\mathrm{m}\sum_{l,m}\frac{4\pi\,Y_{lm}(\theta,\phi)}{2l+1}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r^{l+1}}\left[1 \,+\, a_1\,(\mathrm{i}\,kr) \,+\, \ldots \,+\, a_l\,(\mathrm{i}\,kr)^l\right]\int \mathbf{J}(\mathbf{x}')\,r'^l\,Y_{lm}^*(\theta',\phi')\,\mathrm{d}^3x' \tag{7.7}$$

Often, we will find it convenient to use the Addition Theorem for spherical harmonics (section J3.6) to write:

$$\mathbf{A}(\mathbf{x}) \;=\; k_\mathrm{m}\sum_{l}\frac{\mathrm{e}^{\mathrm{i}kr}}{r^{l+1}}\left[1 \,+\, a_1\,(\mathrm{i}\,kr) \,+\, \ldots \,+\, a_l\,(\mathrm{i}\,kr)^l\right]\int \mathbf{J}(\mathbf{x}')\,r'^l\,P_l(\cos\gamma)\,\mathrm{d}^3x' \tag{7.8}$$

where $\cos\gamma = \hat{\mathbf{n}}\cdot\hat{\mathbf{n}}'$.

Two limits are of interest:

- **Near-field region** ($kr' \ll kr \ll 1$)

  Inspection of eq. (7.7) shows that when the distance to the observation point is much shorter than the wavelength, ie. when $kr \ll 1$, we recover the static (or, more precisely, quasi-static) case, since $\mathrm{e}^{\mathrm{i}kr} \longrightarrow 1$:

$$\mathbf{A}(\mathbf{x}) \underset{kr\ll 1}{\approx} k_\mathrm{m} \sum_{l,m} \frac{4\pi}{2l+1} \frac{Y_{lm}(\theta,\phi)}{r^{l+1}} \int \mathbf{J}(\mathbf{x}')\, r'^l\, Y^*_{lm}(\theta',\phi')\, \mathrm{d}^3x' \tag{7.9}$$

  which also follows more directly by inserting eq. (2.10) into eq. (7.1) in the same limit. The corresponding near-fields display a leading $1/r^2$ dependence, as we shall find explicitly below in section 7.2.

- **Far-field region** ($kr \gg 1$)

  Then the term $a_l\,(\mathrm{i}\,kr)^l$ in the square bracket in eq. (7.8) dominates. Also, $P_l(\cos\gamma)$ is just a polynomial of order l in $\hat{\mathbf{n}}\cdot\hat{\mathbf{n}}' = \hat{\mathbf{n}}\cdot\mathbf{x}'/r'$. Eq. (7.8) becomes:

$$\mathbf{A}(\mathbf{x}) \underset{kr\gg 1}{\approx} k_\mathrm{m} \frac{\mathrm{e}^{\mathrm{i}kr}}{r} \sum_l a_l\,(\mathrm{i}\,k)^l \int \mathbf{J}(\mathbf{x}')\, r'^l\, P_l(\hat{\mathbf{n}}\cdot\hat{\mathbf{n}}')\, \mathrm{d}^3x' \tag{7.10}$$

  Just as this potential, the far-fields have a leading $1/r$ dependence. By inserting the expression for the vector potential of a localised source into eq. (7.2), applying vector calculus identities and keeping only the leading $1/r$ terms at each stage, one shows (EXERCISE) that the far-field magnetic field is:

$$\mathbf{H} \;=\; \frac{\mathrm{i}\,k}{\mu_0}\,\hat{\mathbf{n}}\times\mathbf{A} \tag{7.11}$$

  To find the electric field in the far zone, we go back to the general expression $\mathbf{E} = \mathrm{i}kc\mathbf{A} - \boldsymbol{\nabla}\Phi$ and calculate the gradient directly, starting with the far-zone expression for Φ derived from eq. (5.20) and carrying out the calculation using vector calculus identities and the continuity equation. The result is (EXERCISE):

$$\mathbf{E} \;=\; \mathrm{i}\,kc\left[\mathbf{A} \;-\; \hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\mathbf{A})\right] \;=\; Z_0\,\mathbf{H}\times\hat{\mathbf{n}} \tag{7.12}$$

  This shows that not only plane-wave fields, but all electromagnetic fields far from an oscillating source are perpendicular to each other and transverse to the direction of observation.

  Calculating the time-averaged Poynting vector for the far-fields from eq. (5.39) is now straightforward:

$$\mathbf{S} \;=\; \frac{1}{2}(\mathbf{E}\times\mathbf{H}^*) \;=\; \frac{1}{2}Z_0\,|\mathbf{H}|^2\,\hat{\mathbf{n}} \;=\; \frac{\mu_0}{2}c\,|\mathbf{H}|^2\,\hat{\mathbf{n}}$$

  We see that electromagnetic energy is being propagated at the speed of light in the direction of observation, and that only the transverse component of $\mathbf{A}$ can drive this propagation. We can also write $\mathbf{S}$ directly in terms of the vector potential:

$$\mathbf{S} \;=\; \frac{1}{2}\,k^2\frac{Z_0}{\mu_0^2}\,|\hat{\mathbf{n}}\times\mathbf{A}|^2\,\hat{\mathbf{n}} \;=\; \frac{1}{2\mu_0}\,k^2c\,|\mathbf{A}|^2\,\sin^2\theta\,\hat{\mathbf{n}} \tag{7.13}$$

  where θ is the angle between the vector potential and the direction of observation (and propagation).

  Remembering that the Poynting vector represents the amount of energy flowing through a surface element $R^2\mathrm{d}\Omega$ per unit of time, we can ask how much power the source emits per unit solid angle $\mathrm{d}\Omega$. The **angular power distribution** is:

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} \;=\; |\mathbf{S}\cdot\hat{\mathbf{n}}|\,R^2 \;\approx\; |\mathbf{S}\cdot\hat{\mathbf{n}}|\,r^2 \tag{7.14}$$

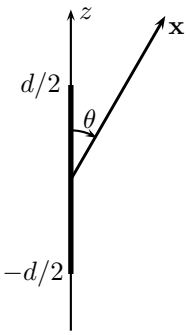
  Inserting all the information from above in turn yields:

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} \;=\; \frac{1}{2\mu_0}\,(kr)^2c\,|\mathbf{A}|^2\,\sin^2\theta \;=\; \frac{1}{8\pi\,k_\mathrm{m}}\,k^2c\,\sin^2\theta\left|\int \mathbf{J}(\mathbf{x}')\,\mathrm{e}^{-\mathrm{i}k\hat{\mathbf{n}}\cdot\mathbf{x}'}\,\mathrm{d}^3x'\right|^2 \tag{7.15}$$

  with angular information about the source stored in the integral. Since this expression has no r dependence, only energy propagated by far-fields (aka radiation) reach arbitrarily large distances without attenuation.

Example 7.1. The Linear Centre-Fed Antenna (section J9.4)

Sometimes we get lucky, and the current density is sufficiently simple that the integral for the potential $\mathbf{A}$ can be done in closed form, and we don't have to use the expansion in eq. (7.6).



Consider a linear antenna rod of length d, with a circular cross-section of radius a. Orient the z axis so that the antenna lies along it, with its centre at the origin, where there is a small gap that serves to excite the antenna.
Make the following assumptions:

1) The antenna is a perfect conductor. This means that on its surface the z component of the electric field vanishes, because it is tangential: in cylindrical coordinates, $E_z(\rho = a) = 0$.

2) a is sufficiently smaller than λ and d that any current in the antenna is along the$z$ axis and flows only at the surface. Then $\mathbf{A} = A_z\,\hat{\mathbf{z}}$, and the fields are azimuthally symmetric.

As usual, we work with harmonic potentials.

Because $E_z(\rho = a) = 0$, this means that on the surface, the z component of eq. (7.3) is:

$$\left(\partial_z^2 \,+\, k^2\right) A_z(\rho = a, z) \;=\; 0$$

and the vector potential is harmonic in space *on the surface*, even when $\mathbf{J}(\mathbf{x})$ is not!

We cannot go much further without making more assumptions. As Jackson mentions (section J9.4B), finding the current flowing through the antenna is a messy affair which depends on the method of excitation. We cannot even take the current as given since it is modified by the emission of radiation...

The next assumption states that:

2) the antenna becomes infinitely thin: $a \to 0$. Just $a \ll d$ may not be sufficient since significant corrections for a finite antenna cross-section exist even for ratios a/d of the order of one thousandth.

Only with this approximation can the current density be taken as spatially harmonic (or sinusoidal). Because of the method of excitation, the current density in one half of the antenna is in phase with the other half, ie. even in the position dependence about the centre:

$$\mathbf{J}(\mathbf{x}) \;=\; I\,\sin\Big(kd/2 - k\,\big|z\big|\Big)\,\delta(x)\,\delta(y)\,\hat{\mathbf{z}}$$

With this expression, the current vanishes at the ends, and the current at the central gap is $I_0 = I\sin(kd/2)$.

We are interested in the far fields. This current density has the particularity that if we insert it in eq. (7.1), we can solve *exactly* for the vector potential without making the long-wavelength ($kd \ll 1$) approximation, just by putting $R \approx r - \hat{\mathbf{n}}\cdot\mathbf{x}' = r - z'\cos\theta$. Dropping the primes, we find:

$$\mathbf{A} \;=\; \hat{\mathbf{z}}\,k_{\mathrm{m}}\,I\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\int_{-d/2}^{d/2}\sin\Big(kd/2 - k\,\big|z'\big|\Big)\,\mathrm{e}^{-\mathrm{i}kz'\cos\theta}\,\mathrm{d}z'$$

The integral, while "straightforward" (says Jackson), still requires a modicum of care because of the

absolute value. Notice that the imaginary part vanishes because it is odd in z, leaving:

$$\int_{-d/2}^{d/2} \sin\left(kd/2 - k\,|z'|\right) \mathrm{e}^{-\mathrm{i}kz'\cos\theta}\,\mathrm{d}z' \;=\; 2\int_0^{d/2} \sin\left(kd/2 - k\,z'\right)\,\cos\left(kz'\,\cos\theta\right)\mathrm{d}z'$$
$$= \int_0^{d/2} \Big[\sin\left[kd/2 - k\,(1-\cos\theta)z\right]\mathrm{d}z \;+\; \sin\left[kd/2 - k\,(1+\cos\theta)\,z\right]\Big]\,\mathrm{d}z'$$

The substitutions $u = kd/2 - k\,(1 \pm \cos\theta)\,z$ allow for easy integration, and we arrive at:

$$\mathbf{A} \;=\; \hat{\mathbf{z}}\,2\,k_{\mathrm{m}}\,I\,\frac{\mathrm{e}^{\mathrm{i}kr}}{kr}\left[\frac{\cos\left(\frac{kd}{2}\,\cos\theta\right) \;-\; \cos\left(\frac{kd}{2}\right)}{\sin^2\theta}\right] \tag{7.16}$$

Of course. this is easier with ***Maple*** — except perhaps for how it thinks the left-hand side should be written:

```
> Int(sin(k*d/2 -k*abs(z))*exp(-I*k*z*cos(theta)), z=-d/2..d/2):
> %=convert(value(%),trig) assuming d>0;
```

$$\int_{-d/2}^{d/2} -\sin\left(-1/2\,kd + k\,|z|\right)e^{-ikz\cos(\theta)}dz = \frac{2\,\cos\left(1/2\,dk\cos\left(\theta\right)\right) - 2\,\cos\left(1/2\,kd\right)}{\left(\sin\left(\theta\right)\right)^2\,k}$$

In the radiation zone, eq. (7.15) leads directly ($Z_0 \;=\; 4\pi k_{\mathrm{m}} c$) to the time-averaged angular power distribution:

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} \;=\; \frac{Z_0}{8\pi^2}\,I^2\left|\frac{\cos\left(\frac{kd}{2}\,\cos\theta\right) \;-\; \cos\left(\frac{kd}{2}\right)}{\sin\theta}\right|^2 \tag{7.17}$$

We note that the distribution depends on kd, ie. on the ratio d/λ. The total power is difficult to find analytically, except in the long wavelength limit. When $kd \ll 1$, the numerator in the bracket expanded to second order is $(kd/2)^2(\sin^2\theta)/2$, and $I_0 = I\sin(kd/2) \approx Ikd/2$, so that:

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega}\Big|_{kd\ll 1} \;=\; \frac{Z_0\,I_0^2}{128\,\pi^2}\,(kd)^2\,\sin^2\theta \tag{7.18}$$

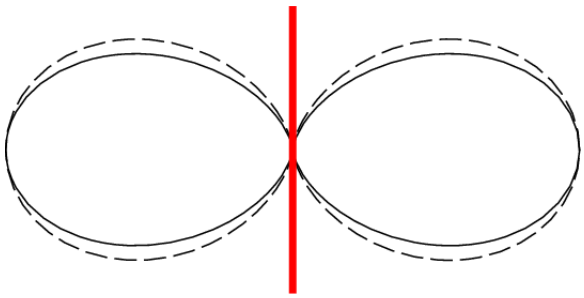
which is the pattern emitted by a current $I(z) = I_0\big(1 - 2|z|/d\big)$ — the long-wavelength limit of the sinusoidal current we have used for the exact result — in the dipole approximation (see eq. (7.20) below). The total power emitted in the long-wavelength limit is then:

$$P \underset{kd\ll 1}{\approx} \frac{1}{48\,\pi}\,Z_0\,I_0^2\,(kd)^2 \tag{7.19}$$

So long as the frequency does not increase too much, the power emitted goes like the square of the frequency.

The dependence of eq. (7.17) on d/λ is well illustrated by comparing two special cases: the **half-wave** ($d = \lambda/2$) antenna and the **full-wave** ($d = \lambda$) antenna. Here is a comparison, obtained with ***Maple***, between the exact result (solid line) and dipole approximation (dashed line) for a "half-wave", for which $d = \lambda/2$. The antenna is represented by the thick rod.

```
> exactradpattern:=plottools[rotate](plots[polarplot]((cos(Pi*cos(theta)/2))^2/
> (sin(theta))^2,theta=0..2*Pi,color=black,axes=none,tickmarks=[0,0]),Pi/2):
> dipoleradpattern:=plottools[rotate](plots[polarplot]((sin(theta))^2,theta=0..2*Pi,
> color=black,axes=none,linestyle=dash,tickmarks=[0,0]),Pi/2):
> antenna:=plottools[rotate](plot(.000001,x=-0.5..0.5,thickness=5),Pi/2):
> plots[display](exactradpattern, dipoleradpattern, antenna);
```



The half-wave antenna's radiation pattern looks, at least qualitatively, like a dipole pattern. The full-wave ($d = \lambda$) figure-of-8 pattern (fig. J9.7) is much more pinched, however. The difference is the contribution from higher multipole moments. Actually, including just one extra moment gives an excellent approximation.

7.2 Electric Dipole Fields from Oscillating Sources

Most often, $\mathbf{A}$ cannnot be found in closed form. We shall be using eq. (7.6), keeping only the leading term ($l = 0$). Using $h_0^{(1)}(x) = \mathrm{e}^{\mathrm{i}x}/\mathrm{i}x$ and $j_0(x) = (\sin x)/x$, as given in eq. (J9.87), this is:

$$\mathbf{A}(\mathbf{x}) \;\approx\; k_\mathrm{m}\left[\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\int \mathbf{J}(\mathbf{x}')\,\frac{\sin kr'}{kr'}\,\mathrm{d}^3x'\right] \underset{kr'\ll 1}{\approx} k_\mathrm{m}\left[\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\int \mathbf{J}(\mathbf{x}')\,\mathrm{d}^3x'\right] \tag{7.20}$$

With the same integration by parts we performed in section 4.4, the integrand $\mathbf{J}(\mathbf{x}')$ is turned into $-\mathbf{x}'(\boldsymbol{\nabla}'\cdot\mathbf{J})$. Then, because of the continuity equation, we find:

$$\begin{aligned}\mathbf{A}_\mathrm{edip}(\mathbf{x}) \;&=\; -\,\mathrm{i}\,k_\mathrm{m}\omega\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\int \mathbf{x}'\,\rho(\mathbf{x}')\,\mathrm{d}^3x' \\ &=\; -\,\mathrm{i}\,k_\mathrm{m}\omega\,\mathbf{p}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\end{aligned} \tag{7.21}$$

where we have recognised the electric dipole moment[†] $\mathbf{p}$ from chapter 3. The magnetic field is:

$$\begin{aligned}\mathbf{H}_\mathrm{edip} \;&=\; \frac{1}{\mu_0}\,\boldsymbol{\nabla}\times\mathbf{A}_\mathrm{edip} \;=\; -\,\mathrm{i}\,\frac{\omega}{\mu_0}\,k_\mathrm{m}\,\boldsymbol{\nabla}\times\left(\mathbf{p}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\right) \\ &=\; -\,\mathrm{i}\,\frac{c\,k}{4\,\pi}\,\hat{\mathbf{n}}\,\partial_r\left(\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\right)\times\mathbf{p} \\ &=\; \frac{c\,k^2}{4\,\pi}\,\hat{\mathbf{n}}\times\mathbf{p}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\left(1\,-\,\frac{1}{\mathrm{i}\,kr}\right)\end{aligned} \tag{7.22}$$

To write the second line we have invoked one of the vector identities on Jackson's front left cover.

The electric field is a little more involved; from one part of the second eq. (7.2) :

$$\mathbf{E}_\mathrm{edip} \;=\; \mathrm{i}\,\frac{Z_0}{k}\,\boldsymbol{\nabla}\times\mathbf{H}_\mathrm{edip} \;=\; \mathrm{i}\,\frac{c\,k\,Z_0}{4\,\pi}\,\boldsymbol{\nabla}\times\left[\hat{\mathbf{n}}\times\mathbf{p}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\left(1\,-\,\frac{1}{\mathrm{i}\,kr}\right)\right]$$

With

$$f(r) \;=\; \frac{\mathrm{e}^{\mathrm{i}kr}}{r}\left(1\,-\,\frac{1}{\mathrm{i}\,kr}\right)$$

[†]Spherically symmetric sources (monopoles) cannot radiate, because the only time-dependent spherically symmetric solution that satisfies Gauss's Law is a $1/r^2$ electric field.

this has the form:

$$\begin{aligned}
\boldsymbol{\nabla}\times\big[\hat{\mathbf{n}}\,f(r)\times\mathbf{p}\big] &= (\mathbf{p}\cdot\boldsymbol{\nabla})\,\hat{\mathbf{n}}\,f(r) \;-\; \mathbf{p}\,\big[\boldsymbol{\nabla}\cdot\big(\hat{\mathbf{n}}\,f(r)\big)\big] \\
&= \frac{f(r)}{r}\,\big[\mathbf{p} \;-\; \hat{\mathbf{n}}\,(\hat{\mathbf{n}}\cdot\mathbf{p})\big] \;+\; \hat{\mathbf{n}}\,(\hat{\mathbf{n}}\cdot\mathbf{p})\,\partial_r f(r) \;-\; \mathbf{p}\,\left[\frac{2}{r}\,f(r) \;+\; \partial_r f(r)\right] \\
&= \hat{\mathbf{n}}\,(\hat{\mathbf{n}}\cdot\mathbf{p})\,r\,\partial_r\left(\frac{f(r)}{r}\right) \;-\; \frac{\mathbf{p}}{r}\,\partial_r\big[r\,f(r)\big] \\
&= -\,\mathrm{i}\,k\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\big[\mathbf{p} \;-\; \hat{\mathbf{n}}\,(\hat{\mathbf{n}}\cdot\mathbf{p})\big] \;-\; \frac{\mathrm{i}}{k}\,\big[3\,\hat{\mathbf{n}}\,(\hat{\mathbf{n}}\cdot\mathbf{p}) \;-\; \mathbf{p}\big]\,\left[\frac{1}{r^3} \;-\; \frac{\mathrm{i}\,k}{r^2}\right]\,\mathrm{e}^{\mathrm{i}kr}
\end{aligned}$$

where we have used several identities from that famous front-left cover. Since $Z_0 = 1/(c\epsilon_0)$, we finally arrive at:

$$\mathbf{E}_{\text{edip}} \;=\; k_{\mathrm{e}}\,\left[k^2\,(\hat{\mathbf{n}}\times\mathbf{p})\times\hat{\mathbf{n}}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r} \;+\; \big[3\,\hat{\mathbf{n}}\,(\hat{\mathbf{n}}\cdot\mathbf{p}) \;-\; \mathbf{p}\big]\,\left(\frac{1}{r^2} \;-\; \frac{\mathrm{i}\,k}{r}\right)\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\right] \tag{7.23}$$

Although they do extend over a short distance, the parts of the fields that fall off faster than $1/r$ *do not carry energy away to infinity*, unlike the $1/r$ parts. In fact, the dominant part of the electric (or longitudinal) near-field is just the electrostatic dipole field when $kr \longrightarrow 0$, and the magnetic near-field is a factor kr/Z_0 smaller than the electric field which therefore dominates.

Far from the source, we have the **radiation fields emitted by an electric dipole**:

$$\begin{aligned}
\mathbf{H}(\mathbf{x},t) &= \frac{c\,k^2}{4\,\pi}\,\hat{\mathbf{n}}\times\mathbf{p}\,\frac{\mathrm{e}^{\mathrm{i}(kr-\omega t)}}{r} \\
\mathbf{E} &= Z_0\,\mathbf{H}\times\hat{\mathbf{n}}
\end{aligned} \tag{7.24}$$

They represent a transverse spherical electromagnetic wave, with $\hat{\mathbf{n}}$ the direction of propagation. Of course, far enough from the source, the waves are well-approximated by plane waves.

If the components of $\mathbf{p}$ all oscillate in phase, so that it has a constant direction, the time-averaged power per unit area (intensity) carried away by the waves that follows directly from eq. (7.15) reads:

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} \;=\; \frac{1}{2}\,\frac{k_{\mathrm{e}}^2}{Z_0}k^4\,\big|\mathbf{p}\big|^2\,\sin^2\theta \tag{7.25}$$

where θ is the angle between $\mathbf{p}$ and the direction of observation. The total power radiated is:

$$P \;=\; \frac{4\pi}{3}\,\frac{k_{\mathrm{e}}^2}{Z_0}\,k^4\,\big|\mathbf{p}\big|^2 \;=\; \frac{c^2\,Z_0\,k^4}{12\,\pi}\,\big|\mathbf{p}\big|^2 \tag{7.26}$$

7.3 Magnetic Dipole and Quadrupole Fields (section J9.3)

Now we consider the $l = 1$ term in eq. (7.6). The m sum is given by the Addtion Theorem for spherical harmonics of eq. (2.9) or (J3.62), converting it to $P_1(\hat{\mathbf{n}}\cdot\hat{\mathbf{x}}') = \hat{\mathbf{n}}\cdot\hat{\mathbf{x}}'$. After looking up $h_1^{(1)}$ and j_1 in eq. (J9.87), we find:

$$3k_{\mathrm{m}}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{(kr)^2}\,(1-\mathrm{i}\,kr)\,k\int\big(\hat{\mathbf{n}}\cdot\hat{\mathbf{x}}'\big)\,\mathbf{J}(\mathbf{x}')\,\left(\frac{\sin kr'}{(kr')^2}-\frac{\cos kr'}{kr'}\right)\mathrm{d}^3x' \underset{kr'\ll 1}{\approx} k_{\mathrm{m}}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r^2}\,(1-\mathrm{i}\,kr)\int\big(\hat{\mathbf{n}}\cdot\mathbf{x}'\big)\mathbf{J}(\mathbf{x}')\,\mathrm{d}^3x'$$

To elucidate the physical meaning of the integral on the right, rewrite the integrand as:

$$\big(\hat{\mathbf{n}}\cdot\mathbf{x}'\big)\,\mathbf{J}(\mathbf{x}') \;=\; \frac{1}{2}\,\big[(\hat{\mathbf{n}}\cdot\mathbf{x}')\,\mathbf{J} \;+\; (\hat{\mathbf{n}}\cdot\mathbf{J})\,\mathbf{x}'\big] \;+\; \frac{1}{2}\,(\mathbf{x}'\times\mathbf{J})\times\hat{\mathbf{n}}$$

7.3.1 Magnetic dipole radiation

We recognise in the integral of the second term the magnetic dipole moment $\mathbf{m} = \frac{1}{2}\int \mathbf{x}' \times \mathbf{J}\, \mathrm{d}^3x'$ defined in eq. (4.13), so we immediately interpret this contribution to the vector potential as that of a magnetic dipole:

$$\mathbf{A}_{\text{magdip}}(\mathbf{x}) \;=\; \mathrm{i}\, k_{\mathrm{m}}\, k\, \hat{\mathbf{n}} \times \mathbf{m}\, \frac{\mathrm{e}^{\mathrm{i}kr}}{r} \left(1 \,-\, \frac{1}{\mathrm{i}\, kr}\right) \tag{7.27}$$

We now notice that if we substitute $c\mathbf{p}$ for $\mathbf{m}$, we have from eq. (7.22):

$$\mathbf{A}_{\text{magdip}} \;=\; \mathrm{i}k_{\mathrm{m}}\, k\, \frac{4\pi}{k^2}\, \mathbf{H}_{\text{edip}}\big|_{\mathbf{p}c\to\mathbf{m}}$$

Then:

$$\mathbf{H}_{\text{magdip}} \;=\; \frac{1}{4\pi\, k_{\mathrm{m}}}\, \boldsymbol{\nabla} \times \mathbf{A}_{\text{magdip}} \;=\; \mathrm{i}\frac{1}{k}\, \boldsymbol{\nabla} \times \mathbf{H}_{\text{edip}}\big|_{\mathbf{p}c\to\mathbf{m}} \;=\; \frac{1}{Z_0}\, \mathbf{E}_{\text{edip}}\big|_{\mathbf{p}c\to\mathbf{m}}$$

where we have used Ampère's Law for harmonic fields: $\boldsymbol{\nabla} \times \mathbf{H} = -(\mathrm{i}k/Z_0)\mathbf{E}$, to relate $\mathbf{H}_{\text{edip}}$ to $\mathbf{E}_{\text{edip}}$. From eq. (7.23), making the substitution $\mathbf{p} \longrightarrow \mathbf{m}/c$, we arrive at:

$$\mathbf{H}_{\text{magdip}} \;=\; \frac{1}{4\,\pi} \left[k^2\, (\hat{\mathbf{n}} \times \mathbf{m}) \times \hat{\mathbf{n}}\, \frac{\mathrm{e}^{\mathrm{i}kr}}{r} \,+\, \left[3\,\hat{\mathbf{n}}\, (\hat{\mathbf{n}} \cdot \mathbf{m}) \,-\, \mathbf{m}\right] \left(\frac{1}{r^3} \,-\, \frac{\mathrm{i}\, k}{r^2}\right) \mathrm{e}^{\mathrm{i}kr}\right] \tag{7.28}$$

We could have written down our result simply by substituting $\mathbf{E}_{\text{edip}} \longrightarrow 4\pi k_{\mathrm{m}} c \mathbf{H}_{\text{magdip}} = Z_0 \mathbf{H}_{\text{magdip}}$ and $\mathbf{p} \longrightarrow \mathbf{m}/c$ in eq. (7.23)!

Similarly, we expect that we should obtain $\mathbf{E}_{\text{magdip}}$ by substituting, in eq. (7.22), $\mathbf{H}_{\text{edip}} \longrightarrow -\mathbf{E}_{\text{magdip}}/Z_0$ and $\mathbf{p} \longrightarrow \mathbf{m}/c$. Then:

$$\mathbf{E}_{\text{magdip}} \;=\; -\,\frac{Z_0}{4\,\pi}\, k^2\, \hat{\mathbf{n}} \times \mathbf{m}\, \frac{\mathrm{e}^{\mathrm{i}kr}}{r} \left(1 \,-\, \frac{1}{\mathrm{i}\, kr}\right) \tag{7.29}$$

Since the radiation's angular dependence goes like $|\mathbf{A}|^2$ (eq. (7.13) with *its* $\theta = \pi/2$), then $|\hat{\mathbf{n}} \times \mathbf{m}|^2 = |\mathbf{m}|^2 \sin^2\theta$, the same as for electric dipoles. At a given angle, however, the magnetic dipole radiation is down by a factor $(kr')^2$ (EXERCISE), where r' is the dimension of the source, and $kr' \ll 1$ by assumption. Because its angular signature is the same as that of electric dipole radiation, magnetic dipole radiation can be very difficult to observe, except in cases where the electric one vanishes for some reason, such as a selection rule.

The polarisation is also different: it is in the $\hat{\mathbf{n}}$-dipole plane for the electric case, and perpendicular to it in the magnetic case.

7.3.2 Electric quadrupole radiation

Let us move on to the $(\mathbf{x}' \times \mathbf{J}) \times \hat{\mathbf{n}}$ integral in the $l = 1$ term of eq. (7.6), which can be written as:

$$\frac{1}{2}\int \left[(\hat{\mathbf{n}} \cdot \mathbf{x}')\, \mathbf{J} \,+\, (\hat{\mathbf{n}} \cdot \mathbf{J})\, \mathbf{x}'\right] \mathrm{d}^3x' \;=\; -\,\frac{\mathrm{i}\,\omega}{2}\int \mathbf{x}'\, (\hat{\mathbf{n}} \cdot \mathbf{x}')\, \rho(\mathbf{x}')\, \mathrm{d}^3x'$$

To prove this, we note that:

$$\int \left[(\hat{\mathbf{n}} \cdot \mathbf{x}')\, \mathbf{J} \,+\, (\hat{\mathbf{n}} \cdot \mathbf{J})\, \mathbf{x}'\right] \mathrm{d}^3x' \;\equiv\; \int \left[(\hat{\mathbf{n}} \cdot \mathbf{x}')\, (\mathbf{J} \cdot \boldsymbol{\nabla}')\, \mathbf{x}' \,+\, \mathbf{x}'\, (\mathbf{J} \cdot \boldsymbol{\nabla}')\, (\hat{\mathbf{n}} \cdot \mathbf{x}')\right] \mathrm{d}^3x' \;=\; \int (\mathbf{J} \cdot \boldsymbol{\nabla}')\, \left[(\hat{\mathbf{n}} \cdot \mathbf{x}')\, \mathbf{x}'\right] \mathrm{d}^3x'$$

Integrating by parts and using the continuity equation $\boldsymbol{\nabla} \cdot \mathbf{J} = \mathrm{i}\omega\rho$ completes the proof. The electric quadrupole potential is:

$$\mathbf{A}_{\text{equad}} \;=\; -\,\frac{k_{\mathrm{m}}}{2}\, c\, k^2\, \frac{\mathrm{e}^{\mathrm{i}kr}}{r} \left(1 \,-\, \frac{1}{\mathrm{i}\, kr}\right) \int \mathbf{x}'\, (\hat{\mathbf{n}} \cdot \mathbf{x}')\, \rho(\mathbf{x}')\, \mathrm{d}^3x' \tag{7.30}$$

To see why we associate this contribution with an electric quadrupole, from the quadrupole tensor Q_{ij} defined in eq. (3.5) construct:

$$Q_i \;\equiv\; Q_{ij}\,n^j \;=\; \int [3x'_i\,n^j\,x'_j \;-\; r'^2\,n_i]\,\rho(\mathbf{x}')\,\mathrm{d}^3x'$$

or

$$\mathbf{Q}(\hat{\mathbf{n}}) \;=\; \int \left[3\,\mathbf{x}'\,(\hat{\mathbf{n}}\cdot\mathbf{x}') \;-\; \hat{\mathbf{n}}\,r'^2\right]\rho(\mathbf{x}')\,\mathrm{d}^3x'$$

Then:

$$\mathbf{A}_{\text{equad}} \;=\; -\,\frac{1}{6}\,k_{\text{m}}\,c\,k^2\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\left(1\;-\;\frac{1}{\mathrm{i}\,kr}\right)\left[\mathbf{Q}(\hat{\mathbf{n}})\;+\;\hat{\mathbf{n}}\int r'^2\,\rho(\mathbf{x}')\,\mathrm{d}^3x'\right]$$

The radiation (far) fields are:

$$\begin{aligned}\mathbf{H}_{\text{equad}} \;&=\; \frac{\mathrm{i}\,k}{4\pi\,k_{\text{m}}}\,\hat{\mathbf{n}}\times\mathbf{A} \;=\; -\,\mathrm{i}\,\frac{c\,k^3}{24\,\pi}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\,\hat{\mathbf{n}}\times\mathbf{Q}(\hat{\mathbf{n}})\\ \mathbf{E}_{\text{equad}} \;&=\; Z_0\,\mathbf{H}\times\hat{\mathbf{n}} \;=\; -\,\mathrm{i}\,\frac{c\,Z_0\,k^3}{24\,\pi}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\left[\hat{\mathbf{n}}\times\mathbf{Q}(\hat{\mathbf{n}})\right]\times\hat{\mathbf{n}}\end{aligned} \tag{7.31}$$

The angular power distribution is:

$$\frac{\mathrm{d}P_{\text{equad}}}{\mathrm{d}\Omega} \;=\; \frac{1}{2Z_0}\,r^2\left|\mathbf{E}\right|^2 \;=\; \frac{c^2\,Z_0\,k^6}{1152\,\pi^2}\left|\left[\hat{\mathbf{n}}\times\mathbf{Q}(\hat{\mathbf{n}})\right]\times\hat{\mathbf{n}}\right|^2 \tag{7.32}$$

In the important case of azimuthal symmetry, the quadrupole tensor is diagonal, and this expression undergoes considerable simplification. Since this tensor is always traceless, we can write $Q_{33} = Q_0$, and $Q_{11} = Q_{22} = -Q_0/2$. Then $\mathbf{Q} = -Q_0(\hat{\mathbf{n}} - 3\hat{\mathbf{z}}\cos\theta)/2$, and $\hat{\mathbf{n}}\cdot\mathbf{Q} = -Q_0(1-3\cos^2\theta)/2 = Q_0\,P_2(\cos\theta)$. A few straightforward steps lead to:

$$\left|\left[\hat{\mathbf{n}}\times\mathbf{Q}(\hat{\mathbf{n}})\right]\times\hat{\mathbf{n}}\right|^2 \;=\; \left|\mathbf{Q}(\hat{\mathbf{n}})\right|^2 \;-\; \left|\hat{\mathbf{n}}\cdot\mathbf{Q}(\hat{\mathbf{n}})\right|^2 \;=\; \frac{9}{4}\,Q_0^2\,\cos^2\theta\,\sin^2\theta$$

and

$$\frac{\mathrm{d}P_{\text{equad}}}{\mathrm{d}\Omega} \;=\; \frac{c^2\,Z_0\,k^6}{512\,\pi^2}\,Q_0^2\,\cos^2\theta\,\sin^2\theta \tag{7.33}$$

Integrating over all solid angles, the total power radiated is found to be:

$$P_{\text{equad}} \;=\; \frac{c^2\,Z_0\,k^6}{960\,\pi}\,Q_0^2 \tag{7.34}$$

In general, however, quadrupole angular distributions can be rather complicated. Yet the total power radiated is surprisingly easy to compute. Indeed, we write again $\left|\left[\hat{\mathbf{n}}\times\mathbf{Q}\right]\times\hat{\mathbf{n}}\right|^2 \;=\; \left|\mathbf{Q}\right|^2 \;-\; \left|\hat{\mathbf{n}}\cdot\mathbf{Q}\right|^2$. This is equivalent to $Q^*_{ij}\,Q_{ik}\,n^j\,n^k \;-\; Q^*_{ij}\,Q_{kl}\,n^i\,n^j\,n^k\,n^l$, and has the effect of storing the angular information in the products of Cartesian components of $\hat{\mathbf{n}}$ and in their index structure.

To integrate the products over all solid angles, we note that $n^j\,n^k$ is symmetric in j and k. Also, when $j \neq k$, $n^j\,n^k$, the integral over ϕ vanishes. This means that $\int n^j n^k\,\mathrm{d}\Omega = \alpha\delta_{jk}$. When $j = k$, we have $\int n^j n^j\,\mathrm{d}\Omega = \int \mathrm{d}\Omega = 4\pi = \alpha\delta_{jj} = 3\alpha$. Therefore, $\alpha = 4\pi/3$, and:

$$\int n^j\,n^k\,\mathrm{d}\Omega \;=\; \frac{4\,\pi}{3}\,\delta_{jk}$$

Similarly, $n^i\,n^j\,n^k\,n^l$ is symmetric in all its indices, and the indices must be equal *in pairs* to avoid an integrand that is odd under reversal of direction. So its integral must be proportional to $\delta_{ij}\,\delta_{kl} + \delta_{ik}\,\delta_{jl} + \delta_{il}\,\delta_{kj}$. We find the proportionality constant by calculating the easiest integral: $\int n_z^4\,\mathrm{d}\Omega = \int \cos^4\theta\,\mathrm{d}\Omega = 4\pi/5$. Therefore:

$$\int n^i\,n^j\,n^k\,n^l\,\mathrm{d}\Omega \;=\; \frac{4\pi}{15}(\delta_{ij}\,\delta_{kl} + \delta_{ik}\,\delta_{jl} + \delta_{il}\,\delta_{kj})$$

The total angular integral is:

$$\int \left| \left[\hat{\mathbf{n}} \times \mathbf{Q}(\hat{\mathbf{n}}) \right] \times \hat{\mathbf{n}} \right|^2 \mathrm{d}\Omega \;=\; \frac{4\,\pi}{3} \left| Q_{ij} \right|^2 \;-\; \frac{4\pi}{15} \left[Q^{*i}_{\;\;i} Q^{*j}_{\;\;j} \;+\; 2 \left| Q_{ij} \right|^2 \right]$$

The quadrupole tensor being traceless, $Q^i_{\;i} = 0$, and:

$$\int \left| \left[\hat{\mathbf{n}} \times \mathbf{Q}(\hat{\mathbf{n}}) \right] \times \hat{\mathbf{n}} \right|^2 \mathrm{d}\Omega \;=\; \frac{4\pi}{5} \left| Q_{ij} \right|^2$$

We obtain for the power radiated:

$$P \;=\; \frac{c^2\, Z_0\, k^6}{1440\,\pi} \left| Q_{ij} \right|^2 \tag{7.35}$$

When the source is azimuthally symmetric, $\left| Q_{ij} \right|^2 = Q_{11}^2 + Q_{22}^2 + Q_{33}^2 = 3Q_0^2/2$, and we regain eq. (7.34).

At this point we have pretty much reached the limit to the usefulness of eq. (7.8). As we have just seen, the terms with $l > 0$ produce expressions in which different electric and magnetic multipole moments are mixed up in a way that makes them very difficult to disentangle, all the more so when $l > 1$. We now introduce a different approach, which has the advantage that multipole moments are kept separated from the beginning,

7.4 Multipole Expansion of Electromagnetic Fields (J9.7, J9.10)

In this approach we work directly with the fields, not the potential. Furthermore, we take their sources to be not only localised, but also harmonic, in the knowledge that more general time behaviour can always be constructed with a Fourier superposition. Thus the time dependence is of the form $\mathrm{e}^{-\mathrm{i}\,\omega t}$. In addition to the charge and current densities ρ and $\mathbf{J}$, we allow for an intrinsic magnetisation $\boldsymbol{\mathcal{M}}$ which we keep separate. Following Jackson, we write Maxwell's equations in SI units and in terms of a magnetic field $\mathbf{H} \coloneqq \mathbf{B}/\mu_0$:

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{E} \;&=\; \rho/\epsilon_0 & \qquad \boldsymbol{\nabla} \times \mathbf{H} \;+\; \mathrm{i}\,\frac{k}{Z_0}\,\mathbf{E} &=\; \mathbf{J} \;+\; \boldsymbol{\nabla} \times \boldsymbol{\mathcal{M}} \\ \boldsymbol{\nabla} \cdot \mathbf{H} \;&=\; 0 & \qquad \boldsymbol{\nabla} \times \mathbf{E} \;-\; \mathrm{i}\,k Z_0\,\mathbf{H} \;&=\; 0 \end{aligned} \tag{7.36}$$

with $k = \omega/c$. The redefinition $\mathbf{E}' = \mathbf{E} \;+\; \mathrm{i}\,\mathbf{J}/\omega\epsilon_0 = \mathbf{E} \;+\; \mathrm{i}\,Z_0 \mathbf{J}/k$ results in two independent equations:

$$\boldsymbol{\nabla} \times \mathbf{E}' \;-\; \mathrm{i}\,k Z_0\,\mathbf{H} \;=\; \frac{\mathrm{i} Z_0}{k}\,\boldsymbol{\nabla} \times \mathbf{J} \qquad\qquad \boldsymbol{\nabla} \times \mathbf{H} \;+\; \mathrm{i}\,\frac{k}{Z_0}\,\mathbf{E}' \;=\; \boldsymbol{\nabla} \times \boldsymbol{\mathcal{M}} \tag{7.37}$$

which guarantee that the divergence of both $\mathbf{H}$ amd $\mathbf{E}'$ vanishes.

As usual, taking the curl of each equation and combining with the other yields second-order *decoupled* Helmholtz equations:

$$\begin{aligned} (\nabla^2 \;+\; k^2)\,\mathbf{H} \;&=\; -\,\boldsymbol{\nabla} \times (\mathbf{J} \;+\; \boldsymbol{\nabla} \times \boldsymbol{\mathcal{M}}) \\ (\nabla^2 \;+\; k^2)\,\mathbf{E}' \;&=\; -\,\mathrm{i}\,Z_0 k\,\boldsymbol{\nabla} \times \left(\boldsymbol{\mathcal{M}} \;+\; \frac{1}{k^2}\,\boldsymbol{\nabla} \times \mathbf{J} \right) \end{aligned} \tag{7.38}$$

Only one of these needs solving, so long as the vanishng of the divergence of the field is implemented. Then the other field follows from the appropriate first-order curl equation.

At this point we could follow the same path that leads to (7.6) for the vector potential $\mathbf{A}$, with $\mathbf{A}$ replaced by the fields and $\mathbf{J}$ by the right-hand sides of eq. (7.38). But we would encounter the same issues with disentangling the multipoles, which are rooted in the Helmholtz operator acting on a *vector*, not a scalar. Instead, like Jackson,

we follow an approach proposed by Bouwkamp and Casimir[†] which consists in writing equations for true *scalars*, not vector components.

Now project the Helmholtz equations along the radial coordinate by taking a scalar product with $\mathbf{r}$. Thanks to identities 0.33 and 0.32, we easily obtain the scalar equations sought:

$$\begin{aligned}(\nabla^2 + k^2)\,(\mathbf{r}\cdot\mathbf{H}) &= -\,\mathrm{i}\,\mathbf{L}\cdot(\mathbf{J} + \boldsymbol{\nabla}\times\boldsymbol{\mathcal{M}})\\ (\nabla^2 + k^2)\,(\mathbf{r}\cdot\mathbf{E}') &= Z_0 k\,\mathbf{L}\cdot\left(\boldsymbol{\mathcal{M}} + \frac{1}{k^2}\,\boldsymbol{\nabla}\times\mathbf{J}\right)\end{aligned} \tag{7.39}$$

where $\mathbf{L} \coloneqq -\mathrm{i}\,\mathbf{r}\times\boldsymbol{\nabla}$. There are also links between $\mathbf{r}\cdot\mathbf{H}$ and the electric field $\mathbf{E}$, and between $\mathbf{r}\cdot\mathbf{E}$ and $\mathbf{H}$, provided by the first-order curl equations in eq. (7.36), with the help of identity 0.32:

$$\begin{aligned}Z_0 k\,\mathbf{r}\cdot\mathbf{H} &= -\,\mathrm{i}\,\mathbf{r}\cdot(\boldsymbol{\nabla}\times\mathbf{E}) = \mathbf{L}\cdot\mathbf{E}\\ k\,\mathbf{r}\cdot\mathbf{E}' &= -\,Z_0\,\mathbf{L}\cdot(\mathbf{H} - \boldsymbol{\mathcal{M}})\end{aligned} \tag{7.40}$$

In section 5.4, we obtained the inhomogeneous retarded solution (5.18):

$$\Psi(\mathbf{x},t) = \int \frac{f(\mathbf{x}',t'_{\text{ret}})}{|\mathbf{x}-\mathbf{x}'|}\,\mathrm{d}^3x'$$

for the equation $\left[\nabla^2 - (1/c^2)\partial_t^2\right]\Psi(\mathbf{x},t) = -4\pi\,f(\mathbf{x},t)$. For harmonic scalar functions and sources of the form $f(\mathbf{x}',t'_{\text{ret}}) = f(\mathbf{x}')\mathrm{e}^{-\mathrm{i}\,\omega t'_{\text{ret}}} = f(\mathbf{x}')\mathrm{e}^{-\mathrm{i}\,\omega(t-R/c)} = f(\mathbf{x}')\mathrm{e}^{-\mathrm{i}\,\omega t}\mathrm{e}^{\mathrm{i}\,\omega R/c}$, the differential equation becomes a Helmholtz equation: $(\nabla^2 + k^2)\Psi(\mathbf{x}) = -4\pi f(\mathbf{x}')\mathrm{e}^{\mathrm{i}\,\omega R/c}$, with solution:

$$\Psi(\mathbf{x}) = \int f(\mathbf{x}')\,\frac{\mathrm{e}^{\mathrm{i}\,kR}}{R}\,\mathrm{d}^3x'$$

where $R = |\mathbf{x}-\mathbf{x}'|$. Comparing with eq. (7.39), we immediately write down the inhomogeneous solutions:

$$\begin{aligned}\mathbf{r}\cdot\mathbf{H}(\mathbf{x}) &= \frac{\mathrm{i}}{4\pi}\int \frac{\mathrm{e}^{\mathrm{i}\,kR}}{R}\,\mathbf{L}'\cdot\left[\mathbf{J}(\mathbf{x}') + \boldsymbol{\nabla}\times\boldsymbol{\mathcal{M}}(\mathbf{x}')\right]\mathrm{d}^3x'\\ \mathbf{r}\cdot\mathbf{E}'(\mathbf{x}) &= -\,\frac{Z_0 k}{4\pi}\int \frac{\mathrm{e}^{\mathrm{i}\,kR}}{R}\,\mathbf{L}'\cdot\left[\boldsymbol{\mathcal{M}}(\mathbf{x}') + \frac{1}{k^2}\,\boldsymbol{\nabla}\times\mathbf{J}(\mathbf{x}')\right]\mathrm{d}^3x'\end{aligned} \tag{7.41}$$

with $\mathbf{L}' = -\mathrm{i}\,\mathbf{r}'\times\boldsymbol{\nabla}'$. We recognise $\mathrm{e}^{\mathrm{i}\,kR}/R$ as the Green function for the Helmholtz operator $\nabla^2 + k^2$, for whch an expansion in terms of spherical Bessel functions was obtained earlier in eq. (7.5). For $\mathbf{x}$ outside the sources, where $\mathbf{E}' = \mathbf{E}$, we arrive at expressions very similar to eq. (7.6):

$$\begin{aligned}\mathbf{r}\cdot\mathbf{H}(\mathbf{x}) &= -\,k\sum_{l,m} h_l^{(1)}(kr)\,Y_{lm}(\theta,\phi)\int j_l(kr')\,Y_{lm}^*(\theta',\phi')\,\mathbf{L}'\cdot\left[\mathbf{J}(\mathbf{x}') + \boldsymbol{\nabla}\times\boldsymbol{\mathcal{M}}(\mathbf{x}')\right]\,\mathrm{d}^3x'\\ \mathbf{r}\cdot\mathbf{E}(\mathbf{x}) &= -\,\mathrm{i}\,Z_0 k^2\sum_{l,m} h_l^{(1)}(kr)\,Y_{lm}(\theta,\phi)\int j_l(kr')\,Y_{lm}^*(\theta',\phi')\,\mathbf{L}'\cdot\left[\boldsymbol{\mathcal{M}}(\mathbf{x}') + \frac{1}{k^2}\,\boldsymbol{\nabla}\times\mathbf{J}(\mathbf{x}')\right]\,\mathrm{d}^3x'\end{aligned} \tag{7.42}$$

In order to make contact witth spherical multipole moments labelled by (l,m), we look at homogeneous solutions of eq. (7.39). These have been quoted in section 0.8.3 in terms of spherical Bessel functions $j_l(kr)$ and $n_l(kr)$ given in eq. (0.49), but often the Hankel functions $h_l^{1,2} \coloneqq j_l \pm \mathrm{i}\,n_l$ are preferred. Then, with a radial solution for

[†] C. J̇Bouwkamp and H. B. G. Casimir, *Physica* **20**, 539 (1954).

a given l of the form: $f_l = A_l^{(1)} h_l^{(1)} + A_l^{(2)} h_l^{(2)}$, the general homogeneous solutions of the Helmholtz equation in spherical coordinates would read:

$$\mathbf{r}\cdot\,\mathbf{H}(\mathbf{x}) \;=\; \frac{1}{k}\sum_{l,m} l(l+1)\,B_{lm}\,g_l(kr)\,Y_{lm}(\theta,\phi)$$

$$\mathbf{r}\cdot\,\mathbf{E}(\mathbf{x}) \;=\; -\,\frac{Z_0}{k}\sum_{l,m} l(l+1)\,C_{lm}\,f_l(kr)\,Y_{lm}(\theta,\phi)$$

where the factors $l(l+1)/k$ and $-Z_0$ have been extracted from the coefficients B_{lm} and C_{lm} for convenience. Ultimately, though, we will be concerned with outgoing waves. The functions $h_l^{(2)}(kr)$ go like $\mathrm{e}^{-\mathrm{i}\,kr}$ at large distance (see eq. (9.89) in Jackson) and correspond to incoming waves. Because of this, we choose $A_l^{(2)} = 0$ and $g_l = f_l = h_l^{(1)}$.

Now, observe that eq. (7.40) must also be satisfied outside sources:

$$\mathbf{L}\cdot\mathbf{E}(\mathbf{x}) \;=\; Z_0 k\,\mathbf{r}\cdot\,\mathbf{H}(\mathbf{x}) \;=\; Z_0\sum_{l,m} l(l+1)\,B_{lm}\,h_l^{(1)}(kr)\,Y_{lm}(\theta,\phi)$$

$$\mathbf{L}\cdot\mathbf{H}(\mathbf{x}) \;=\; -\,\frac{k}{Z_0}\;\mathbf{r}\cdot\,\mathbf{E}(\mathbf{x}) \;=\; \sum_{l,m} l(l+1)\,C_{lm}\,h_l^{(1)}(kr)\,Y_{lm}(\theta,\phi)$$

where we have used identity (0.32). Recalling that the operator $\mathbf{L}$ only acts on the angular dependence of the fields, we see that for $\mathbf{L}$ to yield the same Y_{lm} when acting on a given Y_{lm} in the free-space fields, these must take the form:

$$\mathbf{E} \;=\; Z_0\sum_{l,m}\sqrt{l(l+1)}\,B_{lm}\,h_l^{(1)}(kr)\,\mathbf{X}_{lm}(\theta,\phi) \qquad\qquad \mathbf{H} \;=\; \sum_{l,m}\sqrt{l(l+1)}\,C_{lm}\,h_l^{(1)}(kr)\,\mathbf{X}_{lm}(\theta,\phi)$$

where

$$\mathbf{X}_{lm} \;\coloneqq\; \mathbf{L}Y_{lm}/\sqrt{l(l+1)} \tag{7.43}$$

are called **vector spherical harmonics**. This is because $L^2 Y_{lm} = \mathbf{L}\cdot\mathbf{L}Y_{lm} = l(l+1)Y_{lm}$. Since $\mathbf{L}$ is a self-adjoint operator, we find that the vector spherical harmonics are orthonormal. Indeed:

$$\begin{aligned}\int \mathbf{X}^*_{l'm'}\cdot\mathbf{X}_{lm}\,\mathrm{d}\Omega \;&=\; \int \frac{(\mathbf{L}\,Y_{l'm'})^*\cdot\mathbf{L}\,Y_{lm}}{\sqrt{l'(l'+1)}\sqrt{l(l+1)}}\,\mathrm{d}\Omega \\ &=\; \int \frac{Y^*_{l'm'}\,L^2\,Y_{lm}}{\sqrt{l'(l'+1)}\sqrt{l(l+1)}}\,\mathrm{d}\Omega \;=\; \sqrt{\frac{l(l+1)}{l'(l'+1)}}\int Y^*_{l'm'}\,Y_{lm}\,\mathrm{d}\Omega \;=\; \delta_{l'l}\,\delta_{m'm}\end{aligned}$$

In the same way it can be shown (EXERCISE) that $\int \mathbf{X}^*_{l'm'}\cdot(\mathbf{r}\times\mathbf{X}_{lm})\mathrm{d}\Omega = 0\;\;\forall\, l,\; l', m, m'$.

Computation of the components of the vector spherical harmonics in spherical coordinates with eq. (0.31) is quite straightforward:

$$\mathbf{X}_{lm} \;=\; -\,\frac{1}{\sqrt{l(l+1)}}\left(\hat{\boldsymbol{\theta}}\,\frac{m\,Y_{lm}}{\sin\theta} \;+\; \hat{\boldsymbol{\phi}}\,\mathrm{i}\,\partial_\theta Y_{lm}\right) \tag{7.44}$$

We see that vector spherical harmonics have no radial component and dependence. This follows directly from the fact that $\mathbf{r}\cdot\mathbf{L} = 0$. It guarantees that our free-space fields are divergenceless, as expected. But it also means that these fields have no radial component! In fact, we are in the presence of two types of fields: a transverse (to the radial direction) electric field coupled to a magnetic induction field via Faraday's law, and a *different* transverse

magnetic induction field coupled to an electric field via Ampère's law in vacuum (see eq. (7.37)):

$$\begin{aligned} \mathbf{E} &= Z_0 \sum_{l,m} \sqrt{l(l+1)}\, B_{lm}\, h_l^{(1)}(kr)\, \mathbf{X}_{lm}(\theta,\phi) & \mathbf{H} &= -\,\mathrm{i}\,\frac{1}{Z_0 k}\,\boldsymbol{\nabla}\times\mathbf{E} \\ \mathbf{H} &= -\,Z_0 \sum_{l,m} \sqrt{l(l+1)}\, C_{lm}\, h_l^{(1)}(kr)\, \mathbf{X}_{lm}(\theta,\phi) & \mathbf{E} &= \mathrm{i}\,\frac{Z_0}{k}\,\boldsymbol{\nabla}\times\mathbf{H} \end{aligned}$$

Then the general solution of Maxwell's equations can be written as a linear combintation of the two types. For instance, the electric field is the sum of the electric contribution in the first line, and the electric contribution in the second line:

$$\begin{aligned} \mathbf{E} &= Z_0 \sum_{l,m} \sqrt{l(l+1)}\, \left[B_{lm}\, h_l^{(1)}(kr)\, \mathbf{X}_{lm}(\theta,\phi) \,+\, \frac{\mathrm{i}}{k}\, C_{lm}\, \boldsymbol{\nabla}\times\left(h_l^{(1)}(kr)\, \mathbf{X}_{lm}(\theta,\phi)\right)\right] \\ \mathbf{H} &= \sum_{l,m} \sqrt{l(l+1)}\, \left[C_{lm}\, h_l^{(1)}(kr)\, \mathbf{X}_{lm}(\theta,\phi) \,-\, \frac{\mathrm{i}}{k}\, B_{lm}\, \boldsymbol{\nabla}\times\left(h_l^{(1)}(kr)\, \mathbf{X}_{lm}(\theta,\phi)\right)\right] \end{aligned} \tag{7.45}$$

Note that the l sums starts at $l = 1$; can you see why?

7.4.1 Calculation of the multipole coefficients

With identity (0.32), together with $L^2 Y_{lm} = l(l+1)Y_{lm}$:

$$\mathbf{r}\cdot\mathbf{E} \;=\; -\,\frac{Z_0}{k}\sum_{l,m} C_{lm}\,\frac{1}{\mathrm{i}}\,\mathbf{r}\cdot\boldsymbol{\nabla}\times\left(h_l^{(1)}\,\mathbf{L}Y_{lm}\right) \;=\; -\,\frac{Z_0}{k}\sum_{l,m} C_{lm}\, L^2 h_l^{(1)}\, Y_{lm} \;=\; -\,\frac{Z_0}{k}\sum_{l,m} C_{lm}\, l(l+1)\, h_l^{(1)}\, Y_{lm}$$

Similarly, $\mathbf{r}\cdot\mathbf{H} = \frac{1}{k}\sum B_{lm} l(l+1)\, h_l^{(1)}\, Y_{lm}$. Extracting the coefficients using orthogonality of the Y_{lm}, we finally obtain the standard forms:

$$\begin{aligned} a_{lm}^{\mathrm{E}} &:= \sqrt{l(l+1)}\, C_{lm} \;=\; -\,\frac{k}{\sqrt{l(l+1)}}\,\frac{1}{h_l^{(1)}(kr)}\int Y_{lm}^*(\theta,\phi)\,\mathbf{r}\cdot\mathbf{E}\,\mathrm{d}\Omega \\ a_{lm}^{\mathrm{M}} &:= \sqrt{l(l+1)}\, B_{lm} \;=\; \frac{k}{\sqrt{l(l+1)}}\,\frac{1}{h^{(1)}(kr)}\int Y_{lm}^*(\theta,\phi)\,\mathbf{r}\cdot\mathbf{H}\,\mathrm{d}\Omega \end{aligned} \tag{7.46}$$

from which we see that knowledge of $\mathbf{r}\cdot\mathbf{E}$ and $\mathbf{r}\cdot\mathbf{H}$ from eq. (7.41) determines the fields. From eq. (7.5) the factor $\mathrm{e}^{\mathrm{i}kR}/R$ in eq. (7.41) is, with $r_< = r'$ and $r_> = r$:

$$\frac{\mathrm{e}^{\mathrm{i}kR}}{R} \;=\; 4\pi\,\mathrm{i}\,k\sum_{l=0}^{\infty} j_l(kr')\, h_l^{(1)}(kr) \sum_{m=-l}^{l} Y_{lm}^*(\theta',\phi') Y_{lm}(\theta,\phi)$$

When, say, $\mathbf{r}\cdot\mathbf{H}$ is inserted into a_{lm}^{M}, the terms can be rearranged so that, using orthogonality, the integral over all solid angles picks out one value each of l and m in the expansion, the values in a_{lm}^{M} :

$$\frac{1}{4\pi}\int Y_{lm}^*(\theta,\phi)\,\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\,\mathrm{d}\Omega \;=\; \mathrm{i}\,k\, h_l^{(1)}(kr)\, j_l(kr')\, Y_{lm}^*(\theta',\phi')$$

Then eq, (7.41) immediately yields for a_{lm}^{M}:

$$a_{lm}^{\mathrm{M}} \;=\; -\,\frac{k^2}{\sqrt{l(l+1)}}\int j_l(kr')\, Y_{lm}^*(\theta',\phi')\,\mathbf{L}'\cdot\left[\mathbf{J}(\mathbf{x}') \,+\, \boldsymbol{\nabla}\times\boldsymbol{\mathcal{M}}(\mathbf{x}')\right]\mathrm{d}^3x' \tag{7.47}$$

As for a^{E}_{lm}, which depends on $\mathbf{r}\cdot\mathbf{E}$, eq. (7.41), with $\mathbf{E}'=\mathbf{E}$ since the point of observation lies outside the source, leads to (EXERCISE):

$$a^{\mathrm{E}}_{lm} \;=\; \mathrm{i}\,\frac{k^3}{\sqrt{l(l+1)}}\int j_l(kr')\,Y^*_{lm}(\theta',\phi')\,\mathbf{L}'\cdot\left[\boldsymbol{\mathcal{M}}(\mathbf{x}')\;+\;\frac{1}{k^2}\,\boldsymbol{\nabla}\times\mathbf{J}(\mathbf{x}')\right]\mathrm{d}^3x' \tag{7.48}$$

There still remains to interpret these coefficients and the fields with which they are associated in eq. (7.45). To this end we first rewrite the source terms. Inserting identities 0.32 and 0.35: $\mathbf{L}\cdot\mathbf{V}=\mathrm{i}\,\boldsymbol{\nabla}\cdot(\mathbf{r}\times\mathbf{V})$, and $\mathbf{L}\cdot\boldsymbol{\nabla}\times\mathbf{V}=\mathrm{i}\left[\nabla^2(\mathbf{r}\cdot\mathbf{V})-2\boldsymbol{\nabla}\cdot\mathbf{V}-(\mathbf{r}\cdot\boldsymbol{\nabla})\boldsymbol{\nabla}\cdot\mathbf{V}\right]=\mathrm{i}\left[\nabla^2(\mathbf{r}\cdot\mathbf{V})-\partial_r(r^2\,\boldsymbol{\nabla}\cdot\mathbf{V})/r\right]$, in our result for a^{E}_{lm}, there comes:

$$a^{\mathrm{E}}_{lm} \;=\; -\,\frac{k^3}{\sqrt{l(l+1)}}\int j_l(kr')\,Y^*_{lm}(\theta',\phi')\,\left[\boldsymbol{\nabla}'\cdot(\mathbf{r}'\times\boldsymbol{\mathcal{M}})\;+\;\frac{1}{k^2}\nabla'^2(\mathbf{r}'\cdot\mathbf{J})\;-\;\frac{\mathrm{i}\,c}{kr'}\,\partial_{r'}((r'^2\,\rho)\right]\mathrm{d}^3x'$$

where, in the last term, the continuity equation: $\boldsymbol{\nabla}\cdot\mathbf{J}\;=\;-\mathrm{i}\,c\rho/k$, has been invoked. Note also that, unlike Jackson, we retain the prime on the integration variables to emphasise that integration is with respect to source variables. The second integral can be transformed with the help of Green's second identity 0.7, in which the surface lies outside the source. Then, since $j_l(kr')\,Y^*_{lm}$ satisfies a homogeneous Helmholtz equation, the integral becomes: $\int j_l(kr')\,Y^*_{lm}\nabla'^2(\mathbf{r}'\cdot\mathbf{J})\mathrm{d}^3x'\;=\;\int(\mathbf{r}'\cdot\mathbf{J})\nabla'^2\left[j_l(kr')\,Y^*_{lm}\right]\mathrm{d}^3x'\;=\;-k^2\int(\mathbf{r}'\cdot\mathbf{J})j_l(kr')\,Y^*_{lm}\mathrm{d}^3x$. Furthermore, the term involving the charge density can be integrated by parts: $\int j_l Y^*_{lm}r'\partial_{r'}(r'^2\rho)\mathrm{d}r'\mathrm{d}\Omega'\;=\;-\int[\partial_{r'}(r'j_l)]\rho Y^*_{lm}r'^2\mathrm{d}r'\mathrm{d}\Omega'$. Also, from the recurrence relations in eq. (J9.90), $\partial_{r'}(r'j_l)=(l+1)j_l-kr'j_{l+1}$. Exactly the same line of argument can be applied (EXERCISE) to a^{M}_{lm}, leading to the final exact form (no approximation!) for the coefficients:

$$a^{\mathrm{E}}_{lm} \;=\; -\,\mathrm{i}\,\frac{k^2}{\sqrt{l(l+1)}}\int \mathrm{d}^3x'\,Y^*_{lm}(\theta',\phi')\begin{cases}\left[\partial_{r'}\big(r'j_l(kr')\big)\rho c\;+\;\mathrm{i}\,j_l(kr')\,k\mathbf{r}'\cdot\big(\mathbf{J}\;+\;\boldsymbol{\nabla}'\times\boldsymbol{\mathcal{M}}\big)\right]\\ \left[j_l(kr')\big[(l+1)\rho\,c\;+\;\mathrm{i}\,k\mathbf{r}'\cdot\big(\mathbf{J}\;+\;\boldsymbol{\nabla}'\times\boldsymbol{\mathcal{M}}\big)\big]\;-\;\rho c\,kr'\,j_{l+1}(kr')\right]\end{cases}$$

$$a^{\mathrm{M}}_{lm} \;=\; \frac{k^2}{\sqrt{l(l+1)}}\int Y^*_{lm}(\theta',\phi')\left[j_l(kr')\left[\boldsymbol{\nabla}'\cdot(\mathbf{r}'\times\mathbf{J})\;-\;k^2\,\mathbf{r}'\cdot\boldsymbol{\mathcal{M}}\right]\;+\;\boldsymbol{\nabla}'\cdot\boldsymbol{\mathcal{M}}\,\partial_{r'}\big(r'\,j_l(kr')\big)\right]\mathrm{d}^3x' \tag{7.49}$$

where the identity (0.34) has been invoked on the $\boldsymbol{\mathcal{M}}$ in a^{E}_{lm}, and we have given two equivalent versions of this coefficient.

Example 7.2. The thin linear centre-fed antenna revisited (section J9.12)

To illustrate how to calculate multipole moments, let us go back to example 7.1, an ultra-thin linear antenna of length d driven at its centre, in which a current $I(z)=I\sin(kd/2-k|z|)$ is set up that is harmonic not only in time but also in position. We will need the corresponding current and charge densities In spherical coordinates:

$$\begin{aligned}\mathbf{J}(\mathbf{x}') \;&=\; \hat{\mathbf{r}}'\,\frac{I(r')}{2\pi\,r'^2}\left[\delta(\cos\theta'-1)\;-\;\delta(\cos\theta'+1)\right]\\ \rho(\mathbf{x}') \;&=\; \frac{\boldsymbol{\nabla}'\cdot\mathbf{J}}{\mathrm{i}kc}\;=\;\frac{1}{\mathrm{i}kc}\,\frac{\mathrm{d}_{r'}I(r')}{2\pi\,r'^2}\left[\delta(\cos\theta'-1)\;-\;\delta(\cos\theta'+1)\right]\end{aligned}$$

The antenna has no intrinsic magnetisation, and the current density is radial, so that the coefficients a^{M}_{lm} all vanish. As for a^{E}_{lm}, we first take care of the straightforward angular integration in eq. (7.49). The ϕ integration gives zero unless $m=0$, as expected from the azimuthal symmetry of the source. The θ integration yields:

$$\begin{aligned}2\pi\int\sqrt{2l+1/4\pi}P_l(\cos\theta)\left[\delta(\cos\theta-1)\;-\;\delta(\cos\theta+1)\right]\mathrm{d}(\cos\theta)\;&=\;\sqrt{\pi(2l+1)}\big(P_l(0)-P_l(\pi)\big)\\ &=\sqrt{\pi(2l+1)}\big(1-(-1)^l\big)\end{aligned}$$

Then, according to the first line for a^{E}_{lm} in eq. (7.49), there comes:

$$\begin{aligned} a^{\mathrm{E}}_{l0} &= \frac{k^2}{2\pi}\sqrt{\frac{\pi(2l+1)}{l(l+1)}}\big(1-(-1)^l\big)\int_0^{d/2}\Big[kr'\,j_l(kr')\,I(r') - \frac{1}{k}\,\mathrm{d}_{r'}I(r')\mathrm{d}_{r'}\big(r'\,j_l(kr')\big)\Big]\,\mathrm{d}r' \\ &= \frac{1}{2\pi}\sqrt{\frac{\pi(2l+1)}{l(l+1)}}\big(1-(-1)^l\big)\int_0^{d/2}\Big[-\mathrm{d}_{r'}\big(kr'\,j_l(kr')\,\mathrm{d}_{r'}I\big) + kr'\,j_l(kr')\,\big(\cancelto{0}{\mathrm{d}^2_{r'}I + k^2 I}\big)\Big]\,\mathrm{d}r' \\ &= -\frac{1}{2\pi}\sqrt{\frac{\pi(2l+1)}{l(l+1)}}\big(1-(-1)^l\big)\,kr'\,j_l(kr')\,\mathrm{d}_{r'}I(r')\Big|_0^{d/2} \\ &= \frac{I}{\pi d}\sqrt{\frac{\pi(2l+1)}{l(l+1)}}\big(1-(-1)^l\big)\left(\frac{kd}{2}\right)^2 j_l(kd/2) \end{aligned} \tag{7.50}$$

where in the second line the harmonic dependence of the current on position has been used to simplify the result, and where $\mathrm{d}_r I(r) = -kI\cos(kd/2-kr)$. This result holds at all distances and for all wavelengths.

The source terms in the integrands on eq. (7.49) justify the names **electric multipole coefficient** for a^{E}_{lm} and **electric multipole field** for the associated contributions to $\mathbf{E}$ and $\mathbf{H}$ in eq. (7.45). Also, we call a^{M}_{lm} a **magnetic multipole coefficient** and its asssociated field a **magnetic multipole field**.

We can find clearer evidence for this interpretation in the long-wavelength approximation, $kr' \ll 1$, appropriate to radiation from atoms and nuclei (see section J9.11 for a detailed discussion). In that limit, the leading term of $j_l(kr')$ as given in eq. (J9.88) is $(kr')^l/(2l+1)!!$, where $(2l+1)!! := (2l+1)(2l-1)(2l-3)\cdots$. Then the term with $\boldsymbol{J}$ and $\boldsymbol{\mathcal{M}}$ in a^{E}_{lm} is down by kr' compared to the first ρ term in the second expression for a^{E}_{lm} in eq. (7.49), and the second ρ term is down by $(kr')^2$. In that limit, then, the electric coefficient is:

$$a^{\mathrm{E}}_{lm} = -\mathrm{i}\frac{k^{l+2}\,c}{(2l+1)!!}\sqrt{\frac{l+1}{l}}\,(Q_{lm} + \tilde{Q}_{lm}) \tag{7.51}$$

with electric multipole moments:

$$Q_{lm} = \int r'^l\,\rho\,Y^*_{lm}\,\mathrm{d}^3x' \qquad \tilde{Q}_{lm} = \frac{\mathrm{i}}{(l+1)\,c}\int r'^l\,k\mathbf{r}'\cdot\big[\mathbf{J}+\boldsymbol{\nabla}'\times\boldsymbol{\mathcal{M}}\big]\,Y^*_{lm}\,\mathrm{d}^3x' \tag{7.52}$$

The Q_{lm} are exactly the electrostatic moments introduced in eq. (3.2)!

For a^{M}_{lm} the long-wavelength approximation means that the $\mathbf{r}\cdot\boldsymbol{\mathcal{M}}$ term in eq. (7.49) is much smaller than the two others, leading to:

$$a^{\mathrm{M}}_{lm} = \mathrm{i}\,\frac{k^{l+2}\,c}{(2l+1)!!}\sqrt{\frac{l+1}{l}}\,(M_{lm} + \tilde{M}_{lm}) \tag{7.53}$$

with magnetic multipole moments:

$$M_{lm} = \frac{1}{l+1}\int r'^l\,\boldsymbol{\nabla}'\cdot(\mathbf{r}'\times\mathbf{J})\,Y^*_{lm}\,\mathrm{d}^3x' \qquad \tilde{M}_{lm} = \int r'^l\,Y^*_{lm}\,\boldsymbol{\nabla}'\cdot\boldsymbol{\mathcal{M}}\,\mathrm{d}^3x' \tag{7.54}$$

We recognise in the integrand of $M_{lm}/2$ the magnetic dipole moment density, $(\mathbf{r}\times\mathbf{J})/2$, defined in eq. (4.13).

For the sake of completeness we give the final expressions for the electric and magnetic fields:

$$\begin{aligned}
\mathbf{E} &= Z_0 \sum_{l,m} \Big[a^{\mathrm{M}}_{lm}\, h^{(1)}_l(kr)\, \mathbf{X}_{lm}(\theta,\phi) \;+\; \frac{\mathrm{i}}{k}\, a^{\mathrm{E}}_{lm}\, \boldsymbol{\nabla} \times \Big(h^{(1)}_l(kr)\, \mathbf{X}_{lm}(\theta,\phi) \Big) \Big] \\
\mathbf{H} &= \sum_{l,m} \Big[a^{\mathrm{E}}_{lm}\, h^{(1)}_l(kr)\, \mathbf{X}_{lm}(\theta,\phi) \;-\; \frac{\mathrm{i}}{k}\, a^{\mathrm{M}}_{lm} \boldsymbol{\nabla} \times \Big(h^{(1)}_l(kr)\, \mathbf{X}_{lm}(\theta,\phi) \Big) \Big] \\
&= \sum_{l,m} \Big[\Big(a^{\mathrm{E}}_{lm}\, \mathbf{X}_{lm}(\theta,\phi) \;-\; \frac{\mathrm{i}}{k}\, a^{\mathrm{M}}_{lm}\, \boldsymbol{\nabla} \times \mathbf{X}_{lm}(\theta,\phi) \Big)\, h^{(1)}_l(kr) \;-\; \frac{\mathrm{i}}{k}\, a^{\mathrm{M}}_{lm} \boldsymbol{\nabla} h^{(1)}_l(kr) \times \mathbf{X}_{lm}(\theta,\phi) \Big]
\end{aligned} \tag{7.55}$$

where the coefficients (or moments) are given by eq. (7.49). These expressions exhibit the manifest separation of the electic and magnetic contributions at all orders lm. It is important to observe (EXERCISE) that in each field the electric contribution is transverse to the magnetic one.

7.4.2 Multipole radiation fields and angular power distribution

The benefits of the multipole expansion of the fields over the potential approach are most evident in the treatment of radiation, ie. when taking $kr \gg 1$ in eq. (7.55). First, focus on the $\boldsymbol{\nabla} \times \mathbf{X}_{lm}$ term that occurs in both fields and is explicit in the second line for $\mathbf{H}$. Invoking identity (0.36), it is proportional to: $\mathrm{i}\boldsymbol{\nabla} \times \mathbf{L} Y_{lm}(\theta,\phi) = \mathbf{r}\nabla^2 Y_{lm} - \boldsymbol{\nabla} Y_{lm}$. Both terms go like $1/r$, and the terms in which the curl appears go like $1/kr$; thus, the curl can always be neglected in the radiation zone. Also, from eq. (J9.89), $h^{(1)}_l(kr) \to (-\mathrm{i})^{l+1} \mathrm{e}^{\mathrm{i}\,kr}/kr$. We keep only the leading order from the gradient of this function, leading to:

$$\begin{aligned}
\mathbf{H}(\mathbf{x},t) &= \frac{\mathrm{e}^{\mathrm{i}\,(kr-\omega t)}}{kr} \sum_{l,m} (-\mathrm{i})^{l+1} \left[a^{\mathrm{E}}_{lm}\, \mathbf{X}_{lm}(\theta,\phi) \;+\; a^{\mathrm{M}}_{lm}\, \hat{\mathbf{n}} \times \mathbf{X}_{lm}(\theta,\phi) \right] \\
\mathbf{E}(\mathbf{x},t) &= Z_0\, \mathbf{H} \times \hat{\mathbf{n}}
\end{aligned} \tag{7.56}$$

where the electric field has been found in eq. (7.12). The fields exhibit the characteristic form of a spherical wave. Recall that the time-averaged Poynting vector is: $\hat{\mathbf{n}} \cdot < \mathbf{S} >= \frac{1}{2}\Re\big[\hat{\mathbf{n}} \cdot (\mathbf{E} \times \mathbf{H}^*)\big] = \frac{1}{2} Z_0 \Re\big[\hat{\mathbf{n}} \cdot [(\mathbf{H} \times \hat{\mathbf{n}}) \times \mathbf{H}^*]\big] = \frac{1}{2} Z_0 \big|\mathbf{H} \times \hat{\mathbf{n}}\big|^2$ using the triple-product rule. Then, from definition (7.14), the angular power distribution is:

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = < \mathbf{S} \cdot \hat{\mathbf{n}} > \, r^2 \;=\; \frac{Z_0}{2k^2} \left| \sum_{l,m} (-\mathrm{i})^{l+1} \left[a^{\mathrm{E}}_{lm}\, \mathbf{X}_{lm}(\theta,\phi) \;\times \hat{\mathbf{n}} \;+\; a^{\mathrm{M}}_{lm}\, \mathbf{X}_{lm}(\theta,\phi) \right] \right|^2 \tag{7.57}$$

where $\hat{\mathbf{n}} \cdot \mathbf{X}_{lm} = 0$ in the far-zone. We see that the electric and magnetic contributions for a given order (lm) have the same angular dependence, and can only be distinguished by their polarisations which are perpendicular. Earlier we found that this was true for dipoles, but now we know that its is a feature of all multipoles.

Example 7.3. Going back to our centre-fed antenna for which we derived multipole moments in example 7.2, the calculation of the angular power distribution is simplified by the azimuthal symmetry and the vanishing of the magnetic moments. Then, using $(\mathbf{X}^*_{l'm'} \times \hat{\mathbf{n}}) \cdot (\mathbf{X}_{lm} \times \hat{\mathbf{n}}) = \mathbf{X}^*_{l'm'} \mathbf{X}_m$, plus eq. (7.44) with $m = 0$, we find:

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} \;=\; \frac{Z_0}{8\pi k^2} \left| \sum_l \sqrt{\frac{2l+1}{l(l+1)}}\, (-\mathrm{i})^l\, a^{\mathrm{E}}_{l0}\, \mathrm{d}_\theta P_l \right|^2 \tag{7.58}$$

As discussed in section J9.12, including just $l = 1$ and $l = 3$ multipoles results in an excellent match to the exact distribution; this is the case even when the long-wavelength approximation, which is invalid here, is used to calculate the multipole coefficients. When the emitters are atoms or nuclei (see sections J9.9 and J9.10), the approximation is usually perfectly valid, and expressions (7.51) and (7.53) for the coefficients can be used.

8 Introduction to Relativity

A Digression on Units

Important note: in what follows, **Gaussian units** will be used so as to be consistent with the textbook. Do not confuse these with electrostatic (esu) units, in which $k_\mathrm{e} = 1$ and $k_\mathrm{m} = 1/c^2$), and which were used in the second edition! The whole subject of units is a bit of a dog's breakfast. But the textbook does have a clear discussion of this vexed topic in its appendix.

A few quick transposition rules from SI to Gaussian units. First, in a SI equation, make the following transformations:

$$\text{Sources:} \qquad q, \rho, \mathbf{J}, I, \mathbf{P} \;\to\; \sqrt{4\pi\epsilon_0}\,(q, \rho, \mathbf{J}, I, \mathbf{P}) \;=\; \frac{1}{\sqrt{k_\mathrm{e}}}\,(q, \rho, \mathbf{J}, I, \mathbf{P}), \qquad \mathbf{M} \;\to\; \sqrt{4\pi/\mu_0}\,\mathbf{M} \;=\; \frac{1}{\sqrt{k_\mathrm{m}}}\,\mathbf{M} \tag{8.1}$$

$$\begin{aligned} \text{Fields:} \qquad \mathbf{E} \;&\longrightarrow\; \frac{\mathbf{E}}{\sqrt{4\pi\epsilon_0}} \;=\; \sqrt{k_\mathrm{e}}\,\mathbf{E}, & \mathbf{D} \;&\longrightarrow\; \sqrt{\frac{\epsilon_0}{4\pi}}\,\mathbf{D} \;=\; \frac{\mathbf{D}}{4\pi\sqrt{k_\mathrm{e}}} \\ \mathbf{B} \;&\longrightarrow\; \sqrt{\frac{\mu_0}{4\pi}}\,\mathbf{B} \;=\; \sqrt{k_\mathrm{m}}\,\mathbf{B} & \mathbf{H} \;&\longrightarrow\; \frac{\mathbf{H}}{\sqrt{4\pi\mu_0}} \;=\; \frac{\mathbf{H}}{4\pi\sqrt{k_\mathrm{m}}} \end{aligned} \tag{8.2}$$

Many times, there will be no ϵ_0 or μ_0 left. If there are, either they will be in the combination $\sqrt{\epsilon_0\mu_0}$, in which case put it equal to $1/c$, as in Faraday's Law for instance. Or if they should occur alone, put them equal to 1.

8.1 A Mathematical Digression on Tensors

I have already mentioned in the preamble that to characterise a vector by the fact that it has a magnitude and a direction is unsatisfactory. Physics is full of quantities which have magnitude and direction, but are *not* vectors: angular displacement, electric current, flux, etc. A more sophisticated approach classifies mathematical objects according to *how they behave under certain transformations*. For instance, take the three numbers, $(x^1, x^2, x^3) = (x, y, z)$, that specify a displacement in a Cartesian coordinate basis. If this displacement is with respect to the origin, we can call it a position.

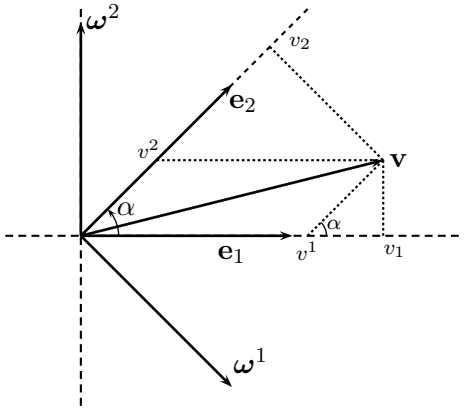
As stated in the mathematical preamble to these notes (section 0), the position vector (and in fact any vector) can be written in terms of a coordinate basis (or system) $\{\mathbf{e}_\mathrm{i}\}$ (i = 1 to 3):

$$\mathbf{x} \;=\; x^i\,\mathbf{e}_\mathrm{i}$$

where the x^i are called **contravariant** components. But to the same vector can be associated a companion, called a **covector**, or 1-form $\boldsymbol{\chi}$, which itself has **covariant** components x_i in a **cobasis** $\{\boldsymbol{\omega}^i\}$:

$$\boldsymbol{\chi} \;=\; x_i\,\boldsymbol{\omega}^\mathrm{i}$$

A basis and its cobasis are related very simply: $\boldsymbol{\omega}^\mathrm{i}(\mathbf{e}_\mathrm{j}) = \delta^i{}_j$, that is, a cobasis vector is perpendicular to all the basis vectors with a different label. Although, strictly speaking, a vector and its covector live in different spaces, it is quite usual to identify them and to put them on the same plot. Here is a two-dimensional example with a general vector $\mathbf{v}$:



With respect to standard orthogonal axes, $\mathbf{e}_1$ has components $\begin{pmatrix}1\\0\end{pmatrix}$ and $\mathbf{e}_2$ components $\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}$. This is an **oblique** basis, such as might be found in cristallography.
Consider the rectangle triangle with side v_1 and $|\mathbf{v}|$ as its hypotenuse.The third side is also a side of the triangle with hypothenuse of length v^2, together with one apex at the tip of $\mathbf{v}$, like the first triangle, and another apex at the point labelled v^1. Then Pythagoras' theorem gives:

$$|\mathbf{v}|^2 \;=\; (v_1)^2 \,+\, \left[(v^2)^2 \,-\, (v_1 \,-\, v^1)^2\right] \;=\; (v^2)^2 \,-\, (v^1)^2 \,+\, 2\,v_1\,v^1$$

Consider the triangle with one apex at the origin, another at point labelled v_2, and the third at the tip of $\mathbf{v}$. A similar argument yields (EXERCISE) another expression for $|\mathbf{v}|^2$ in terms of $v_2\,v^2$. Adding the two, there comes:

$$|\mathbf{v}|^2 \;=\; v_1\,v^1 \,+\, v_2\,v^2$$

We conclude that the length squared of a vector is: $v^i v_i$, not $(v^1)^2 + (v^2)^2 + \ldots$, as many believe. This error arises from the fact that when a Cartesian basis is orthonormal, its cobasis merges with it on the same plot, and $v_i = v^i$.

Now rotate an *orthogonal* basis by an angle θ counterclockwise around the x axis. The components of the vector, (x', y', z'), in the rotated system are given by (see section 1.1.5 of Griffiths' *Introduction to Electrodynamics*):

$$\begin{pmatrix}x'\\y'\\z'\end{pmatrix} \;=\; \begin{pmatrix}1 & 0 & 0\\ 0 & \cos\,\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{pmatrix}\,\begin{pmatrix}x\\y\\z\end{pmatrix} \tag{8.3}$$

or, in index notation:

$$x'^i = R^i{}_j(\theta)\,x^j \tag{8.4}$$

where the $R^i{}_j(\theta)$ are the coefficients of the **rotation matrix**. As one might expect, rotations around an arbitrary axis have a much more complicated rotation matrix; fortunately, we shall pretty much always be able to *choose* the x-axis (or some other Cartesian axis) to be the rotation axis. The above transformation is an example of a *homogeneous linear transformation*.

The set of all rotational transformations form a group, in the sense that:

1. it contains the identity matrix (rotation by zero angle);
2. there is an inverse transformation, $R^{-1}(\theta) = R(-\theta)$: $x^k = (R^{-1})^k{}_i\,x'^i$, where $(R^{-1})^k{}_i R^i{}_j\,x^j = \delta^k{}_j x^j$;
3. two successive rotations are equivalent to a rotation:

$$x''^k \;=\; (R_1)^k{}_i(\phi)\,(R_2)^i{}_j(\theta)\,x'^j \;=\; R^k{}_j\,(\theta \,+\, \phi)\,x^j$$

The components x_i in the cobasis should transform as $x'_i = S_i{}^k x_k$, with $\mathbf{S}$ a matrix to be determined. Then:

$$x'^i x'_i = R^i{}_j\,S_i{}^k\,x^j x_k \;=\; x^k x_k \;=\; \delta^k{}_j x^j x_k$$

so that $(R^i{}_j S_i{}^k - \delta^k{}_j)x^j x_k = 0$. Since there is no restriction on x^j and x_k, we must have $S_i{}^k R^i{}_j = \delta^k{}_j$, or:

$$S_i{}^k \;=\; (R^{-1})^k{}_i$$

so that the x_i must transform as:

$$x'_i \;=\; x_k\,(R^{-1})^k{}_i \tag{8.5}$$

EXERCISE: Show how to obtain the inverse transformation $x_i = x'_k R^k{}_i$.

And now for some definitions:

Definition 8.1. A **3-tensor of rank 0** (scalar) is a one-component object that remains **invariant** (keeps the same numerical value) under a rotation of a coordinate system. Examples: magnitude of a vector, scalar product of two vectors, mass, temperature, etc.

A contravariant **3-tensor of rank 1** in three dimensions (3-vector), $\mathbf{u}$, is an object whose three components transform like those of position under a rotation of a coordinate system:

$$u'^i \;=\; R^i{}_j\,u^j$$

A contravariant **3-tensor of rank 2** in three dimensions, T, is an object whose nine components transform in the following way:

$$T'^{ij} \;=\; R^i{}_k\,R^j{}_l\,T^{kl} \;=\; R^i{}_k\,T^{kl}\,R^j{}_l \;=\; R^i{}_k\,T^{kl}\,R^{\mathrm{T}}{}_l{}^j \qquad\Longleftrightarrow\qquad \mathsf{T}' \;=\; \mathsf{R}\,\mathsf{T}\,\mathsf{R}^{\mathrm{T}} = \mathsf{R}\,\mathsf{T}\,\mathsf{R}^{-1}$$

where the last equality holds because the inverse of the rotation matrix is its transpose. This would not necessarily be true for other transformations.

The entries of the 3×3 matrix T—the components T^{ij} of the tensor—depend on the choice of coordinate basis.

Until further notice, when we speak of a tensor, we mean a 3-tensor according to the above definitions.

These definitions can be extended to a rank-n tensor (ie. a tensor with n indices) in a straightforward manner. All that one has to remember is that the transformation contains n applications of the *same* rotation matrix:

$$T'^{ij\cdots}{}_{kl\ldots} \;=\; (R_k{}^n\,R_l{}^s\,R^i{}_m\,R^j{}_t\cdots)T^{mt\ldots}{}_{ns\ldots} \;=\; (R^{-1})^n{}_k\,(R^{-1})^s{}_l\,R^j{}_t\cdots T^{mt\ldots}{}_{ns\ldots} \tag{8.6}$$

Note that tensor components are generally not scalars (invariant under rotations), and neither are numbers constructed from them! It does happen sometimes that they are, but this must be established properly.

The transformation rules can always be used to determine whether an object is a rank-n tensor. In many cases, however, there is a short cut.

An object is a tensor if it is constructed from other tensors by addition or multiplication by a scalar: if T and Q are tensors of the same rank, $\alpha\mathsf{T} + \beta\mathsf{Q}$ (α, β scalars) is also a tensor of the same rank.

The tensor product of two vectors, $\mathbf{u}\otimes\mathbf{v}$, is a rank-2 tensor with components $u^i v^j$. Beware though: any rank-2 tensor is not necessarily a tensor product of two vectors!

For constructing new tensors, it is useful to have a library of known tensors. For instance, the Kronecker delta is a rank-2 tensor. Indeed:

$$\delta'^i{}_j \;=\; R^i{}_k\,(R^{-1})^n{}_j\,\delta^k{}_n \;=\; R^i{}_k\,(R^{-1})^k{}_j \;=\; \delta^i{}_j$$

Here we learn that there is something more to the Kronecker delta than just being a tensor: its components remain the same under rotations, ie. whatever the basis used! The same is true of the Levi-Civita epsilon: in Cartesian bases, it is a rank-3 tensor whose components themselves do not change under rotations of the bases; in curvilinear bases, however, it is a tensor only if each component is multiplied by a factor (the square root of the absolute value of the determinant of the metric, for the experts) that depends on the coordinates.

There is another way of constructing other tensors from tensors of rank 2 and higher. Start with a tensor T with components $T^i{}_j$. Now **contract over** i **and** j, ie. make the two indices the same letter to get, say, $T^i{}_i$. The object we obtain is called the **trace** of T, **Tr** T in coordinate-free notation. We have reduced the rank of the original

rank-2 tensor by 2 to get a *scalar* equal to the sum of the diagonal coefficients of the matrix representing $\mathbf{T}$. By the same process of contracting over two indices, always one contravariant and one covariant, one can always reduce the rank of a tensor of rank $n > 2$ to one of rank $n - 2$.

Finally, it is quite often useful to decompose rank-2 tensors into their symmetric and antisymmetric parts. In index notation:

$$T^{ij} \;=\; \frac{1}{2}(T^{ij} + T^{ji}) \;+\; \frac{1}{2}(T^{ij} - T^{ji}) \;\equiv\; T^{(ij)} \;+\; T^{[ij]}$$

Clearly, a rank-2 tensor $\mathbf{T}$ is symmetric if $T^{ji} = T^{ij}$, and it is antisymmetric if $T^{ji} = -T^{ij}$.

Strictly speaking, all operations on tensors are understood to be performed at one point. It does not really make sense to add tensors at different points. (Yes, this rule is blatantly violated by differentiation, which involves taking the difference between the values of a tensor at two infinitesimally close points. The fact is that one can get away with this in flat Euclidean space—and even spacetime—but *not* in general curved spaces, such as the ones General Relativity deals with. Then differentiation must be redefined.)

A **field** is simply the collection of all the values of a tensor over all space, given by continuous functions of position $\mathbf{x}$. Fields can be scalar (density $\rho(\mathbf{x})$), vector (velocity $\mathbf{v}$), etc.

How do fields transform? A scalar field $f(\mathbf{x})$ will change to $f'(\mathbf{x}') = f(\mathbf{x}')$. All one has to do is transform the position. In terms of the new coordinates, the *functional form* f' will be different from f, but the number one gets when f' is evaluated at $\mathbf{x}'$ will be the *same* as $f(\mathbf{x})$. Tensor fields of rank 1 and higher need more careful treatment. Again, one must transform the coordinates of $\mathbf{x}$ to those of $\mathbf{x}'$. *But one must also transform the components of the tensor themselves*!

8.2 Galilean Relativity

First, introduce a few definitions.

> **Definition 8.2.** An *event* is characterised by its occurrence at a spacetime point P. The position in spacetime can be specified by three spatial coordinates and one time (we assume that spacetime is four-dimensional). These are given in a **reference frame** which is a set of axes in space and a set of identical, synchronised clocks, one at each point in space. Warning: do not confuse a reference frame with a coordinate basis! You can change coordinate basis while remaining in the same frame. In this chapter we generally use *Cartesian* spatial coordinates.
>
> Reference frames—frames for short—can be in *relative* translational and rotational motion with respect to one another. Of all possible frames, we restrict ourselves to **inertial** frames, ie. unaccelerated frames in which *no* force of any kind is experienced. A rotating frame is not inertial. Consequently, two inertial frames are either *at rest* or in *uniform translational* motion with respect to each other.

Now that we have set the stage, we ask the following. Let an event occur at point P, specified in a frame S by spacetime coordinates (x, y, z, t). What are the coordinates (x', y', z', t') of this event (or of P) as observed in a frame S' ?

In Newtonian mechanics, the relation between $(\mathbf{x}, t)$ and $(\mathbf{x}', t')$ is given by the **Galilean transformations**:

- **Space and time translations**:

  $$\delta\mathbf{x} \;=\; \boldsymbol{\alpha} \;\in\; \mathbb{R}^3 \qquad\qquad \delta t \;=\; \eta \;\in\; \mathbb{R}$$

  Invariance under these is a manifestation of the assumed homogeneity of space and time.

- **Spatial rotations**: $x'^i = R^i{}_j(\phi)\, x^j \iff \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

  Invariance under these is a manifestation of the assumed isotropy of space (no preferred direction).

- **Boosts**:

$$\delta\mathbf{x} \;=\; -\,\mathbf{v}t, \qquad\qquad \delta t = 0 \tag{8.7}$$

where S' is moving at velocity $\mathbf{v}$ with respect to S. From these one immediately derives the **velocity transformations**:

$$\delta\mathbf{u} \;=\; \mathrm{d}(\delta\mathbf{x})/\mathrm{d}t' \;=\; \mathrm{d}(\delta\mathbf{x})/\mathrm{d}t \;=\; -\,\mathbf{v} \tag{8.8}$$

- **Parity, Time-reversal**: $\mathbf{x} \longrightarrow -\,\mathbf{x}, \qquad\qquad t \longrightarrow -\,t$

All these can be combined into the **group of generalised Galilean transformations**:

$$\begin{aligned} x'^i \;&=\; a^i{}_j\, x^j \;+\; \epsilon\, v^i\, t \;+\; \alpha^i \qquad\qquad (i = 1,2,3) \\ \mathbf{x}' \;&=\; \pm\mathbf{x} \qquad\qquad t' \;=\; \pm t \;+\; \eta \end{aligned} \tag{8.9}$$

What is the geometry of spacetime consistent with the Galilean transformations? By geometry, we mean a measure of the distance between two spacetime points, known as the **line element**. Here, the geometry is given by the *absolute* infinitesimal length, $\mathrm{d}l^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = \mathrm{d}x'^2 + \mathrm{d}y'^2 + \mathrm{d}z'^2$, and by the *absolute* infinitesimal time interval, $\mathrm{d}t = \mathrm{d}t'$. Thus, observers in different inertial frames measure the same spatial distance and the same time intervals between any two events. *Space and time are not related*: there are *two* line elements.

The **Galilean relativity principle** restricted to Newtonian mechanics asserts that the laws of Newtonian mechanics hold in all inertial frames, in the sense that their *form* is **invariant** under Galilean transformations. Thus, for instance, the laws of conservation of momentum and kinetic energy:

$$\begin{aligned} \left(\sum_j m_j \mathbf{u}_j\right)_{\text{initial}} \;&=\; \left(\sum_j m_j \mathbf{u}_j\right)_{\text{final}} \quad \text{becomes} \quad \left(\sum_j m_j \mathbf{u'}_j\right)_{\text{initial}} \;=\; \left(\sum_j m_j \mathbf{u'}_j\right)_{\text{final}} \\ \left(\sum_j \frac{1}{2}\, m_j u_j^2\right)_{\text{initial}} \;&=\; \left(\sum_j \frac{1}{2}\, m_j u_j^2\right)_{\text{final}} \quad \text{becomes} \quad \left(\sum_j \frac{1}{2}\, m_j u_j'^2\right)_{\text{initial}} \;=\; \left(\sum_j \frac{1}{2}\, m_j u_j'^2\right)_{\text{final}} \end{aligned} \tag{8.10}$$

provided $\sum_i m_i$ is invariant (mass is conserved) under the transformations.

The Galilean relativity principle precludes the existence of an absolute state of rest, as well as of an absolute, universal, **preferred** frame. But not every equation in Newtonian mechanics satisfies this principle. Indeed, the classical wave equation for a scalar field $f(\mathbf{x}, t)$ is not invariant under Galilean transformations: $(\partial_t^2 - v^2\boldsymbol{\nabla}^2) f(\mathbf{x}, t) \;=\; 0 \;\neq\; (\partial_{t'}^2 - v'^2\nabla'^2) f'(\mathbf{x}', t')$, where v is the wave speed. Should we get upset about this? Not at all, because in mechanics wave propagation takes place in a medium whose rest-frame is a **preferred** frame. For instance, it is easy to tell if we are moving with respect to a stationary sound source in air or water, and in order to determine the frequency of the sound from the observed frequency, we must correct for our motion by using the Doppler shift equation.

8.3 Einstein's Relativity

8.3.1 The need for a new relativity principle

Do electrical phenomena satisfy the Galilean relativity principle? No. Oscillating electric fields involved in light obey a wave equation that is not form-invariant under Galilean transformations. Of course, as in mechanics, this could mean that there exists a preferred frame (often called the ether). Two arguments can be brought against this possibility, however. The first is a theoretical one, put forth by Einstein in his classic paper *On the Electrodynamics of Moving Bodies*: if the vacuum is in fact an ether, one should be able to transform to a frame in which light is at rest. To an observer at rest in this frame, there would be a *static* electric field oscillating as a function of position.

But there is no such static oscillating solution to the Laplace equation, because it does not admit solutions with local extrema.

The other argument rests on experiment: if a preferred frame existed, the speed of light (and therefore of the electric wave) would depend on the speed of the observer with respect to this preferred frame. There would be a sort of "ether wind" for which experimental results must be corrected. But this is not supported by experiment. Starting with the Michelson-Morley experiment in the mid-1880's, ever smaller limits have been put on this wind speed. A very nice experiment [Jaseja, Javan, Murray, Townes, *Physical Review* 133, A 1221 (1964)] uses two laser beams which are allowed to interfere after one has travelled along the direction of motion of the Earth and the other perpendicular to it. The results are compared with measurements made six months later, when the Earth is moving in the opposite direction with a relative speed of about $60\,\mathrm{km/s}$. The absence of any observed change in that experiment puts an upper limit of about $1\,\mathrm{km/s}$ on the speed of the ether wind. More recent experiments give an upper limit [†] of one part in 10^{17}.

If there is no preferred frame for the electric field, and if we believe that Maxwell's equations are correct, only one option is left: to modify the relativity principle itself so as to (1) make it *universal*, applying to *all* fundamental laws of physics, while (2) postulating that the speed of light is *absolute*, with the same value in all inertial frames. Not only is this consistent with the lack of experimental evidence for an ether, but it also defuses the first argument since it makes it impossible to transform to a frame in which light is at rest.

8.3.2 Spacetime interval

To the new relativity principle (let's call it the Einstein principle) must correspond new transformation laws between frames. There are a number of ways of deriving them. Historically, Lorentz did it first, by asking which transformations left the form of Maxwell's wave equations invariant. Here, we will proceed via a route (see problem J11.1) that assumes that *the speed of light is absolute* and that spacetime is *homogeneous* and *isotropic*, meaning that there is no special point or direction in spacetime. These assumptions will allow us first to answer the following question: what is the geometry imposed on spacetime in Einstein's relativity theory? In other words, is there an *invariant* (frame-independent) measure of distance between spacetime points?

Consider two events infinitesimally close in space and time in some frame S, and such that the spatial distance $\mathrm{d}\mathbf{l}$ and time interval $\mathrm{d}t$ between them satisfies $\mathrm{d}\mathbf{l} = \pm c\,\hat{\mathbf{l}}\mathrm{d}t$, where c is the speed of light. Form the object $\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}l^2$. Clearly, for our two events, $\mathrm{d}s^2 = 0$. Equally clearly, since c is the same in another frame S', $\mathrm{d}\mathbf{l}'/dt' = \pm c\hat{\mathbf{l}}$, or $\mathrm{d}\mathbf{l}' = \pm c\hat{\mathbf{l}}'\,\mathrm{d}t'$, and we have $\mathrm{d}s'^2 = 0$. In this case, $\mathrm{d}s^2 = \mathrm{d}s'^2$ for any inertial frame S'.

But what if $\mathrm{d}\mathbf{l} \neq \pm c\hat{\mathbf{l}}\,\mathrm{d}t$? Certainly $\mathrm{d}s^2$ is no longer zero. What is $\mathrm{d}s'^2$ in terms of $\mathrm{d}s^2$, then? To find it, let the two events lie along the x-axis to simplify the formalism. Even without the exact form of the transformations between (x,t) and (x',t'), we expect them to be be homogeneous and linear. Then, quite generally:

$$c^2\mathrm{d}t'^2 \;-\; \mathrm{d}x'^2 = A(x,t)\,c^2\,\mathrm{d}t^2 \;+\; B(x,t)\,\mathrm{d}x^2 \;+\; C(x,t)\,\mathrm{d}x\,\mathrm{d}(ct)$$

But this must be true also for the case we just considered, with $\mathrm{d}x = \pm c\,\mathrm{d}t$ and $\mathrm{d}x' = \pm c\,\mathrm{d}t'$. Inserting this, we have:

$$\begin{aligned} 0 \;&=\; A(x,t)c^2\,\mathrm{d}t^2 \;+\; B(x,t)\,\mathrm{d}x^2 \;\pm\; C(x,t)\mathrm{d}x^2 \\ &=\; c^2\,\mathrm{d}t^2 \;-\; \mathrm{d}x^2 \end{aligned}$$

Thus, $C(x,t) = 0$ and $B(x,t) = -A(x,t)$ in general, since there is no restriction on x and t. We obtain $\mathrm{d}s'^2 = A(x,t)\mathrm{d}s^2$.

[†] Eisele, Ch.; Nevsky, A. Yu.; Schiller, S. (2009), Laboratory Test of the Isotropy of Light Propagation at the 10^{-17} level, Phys. Rev. Lett. 103 (9): 090401;
Herrmann, S., Senger, A., Möhle, K. Nagel, M. Kovalchuk, E. V. Peters, A. (2009), Rotating optical cavity experiment testing Lorentz invariance at the 10^{-17} level, Phys. Rev. D 80 (100): 105011.

Now $A(x,t)$ must be a constant, A; otherwise, one could use it to make a fundamental distinction between spacetime points. The only parameter on which A could depend and which distinguishes between the frames is their relative velocity. Because we have assumed that there is no preferred direction (isotropy), A can only depend on the relative speed, and $\mathrm{d}s'^2 = A(|\mathbf{v}|)\mathrm{d}s^2$. The whole argument can be run in reverse, starting from S', to yield $\mathrm{d}s^2 = A(|-\mathbf{v}|)\mathrm{d}s'^2$. Combining, we get $ds^2 = A^2(v)\mathrm{d}s^2$, so that $A = 1$. $A = -1$ is ruled out, otherwise you could imagine a third frame S'' moving at velocity $\mathbf{v}'$ with respect to S' and at at velocity $\mathbf{v}''$ with respect to S. Then we would be able to write: $\mathrm{d}s''^2 = A(v')\mathrm{d}s'^2 = A(v')A(v)\mathrm{d}s^2 = \mathrm{d}s^2$, while $\mathrm{d}s''^2 = A(v'')\mathrm{d}s^2 = -\mathrm{d}s^2$.

This shows that although $\mathrm{d}\mathbf{l}$ and $\mathrm{d}t$ are expected to change under the transformations, $\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}l^2$ is an invariant. It is the meaningful measure of spacetime distance, otherwise known as the **spacetime interval** or **line element**. No longer do we have two separate geometries as in Newtonian mechanics; there is *one* unified four-dimensional geometry.

8.4 Lorentz transformations (section J11.3)

We can now use the invariance of the line-element $\mathrm{d}s^2$ to derive the so-called boost transformations between frames. Demand that these be linear and homogeneous, so as to be single-valued and not single out the origin as a special point. The transformations must be consistent with the invariant nature of the speed of light and leave the wave equation invariant. They must have an inverse, obtained by: $\mathbf{v} \longrightarrow -\mathbf{v}$.

Write (for relative motion in the x direction):

$$x' \;=\; A\,x \,+\, B(ct) \qquad\qquad ct' \;=\; F(ct) \,+\, G\,x$$

Insert this into $(ct')^2 - x'^2 = (ct)^2 - x^2$. Since there is no restriction on (x,t) or (x',t'), the following must hold:

$$AB - FG = 0, \qquad\qquad A^2 - G^2 = 1, \qquad\qquad F^2 - B^2 = 1$$

Re-parameterise the constants:

$$A \;=\; \cosh\,\psi \qquad\qquad G \;=\; -\,\sinh\,\psi \qquad\qquad F \;=\; \cosh\,\theta \qquad\qquad B \;=\; -\,\sinh\,\theta$$

so that the second and third relations are automatically satisfied, whereas the first yields $\tanh\theta = \tanh\psi$. Since then $\theta = \psi$, $A = F$ and $B = G$.

To find θ, notice that $\tanh\,\theta = -B/A$. Then consider the motion of the origin of S', $x' = 0$, in S with its position given by $x = vt = \beta ct$, where the dimensionless quantity $\beta \equiv v/c$ is the relative speed in units of the speed of light. Since $0 = x - \beta ct$, one has $B/A = -\beta$, so that $x' = A(x - \beta ct)$, $\tanh\,\theta = \beta$. The parameter θ is called the **rapidity** and it ranges from 0 to ∞, as β ranges from 0 to 1; it is also additive. Our boost transformations are:

$$\begin{aligned} x' \;&=\; \gamma\,(x \,-\, \beta\,ct) \qquad & ct' \;&=\; \gamma\,(ct \,-\, \beta\,x) \qquad & y' \;&=\; y \qquad & z' \;&=\; z \\ x \;&=\; \gamma\,(x' \,+\, \beta\,ct') \qquad & ct \;&=\; \gamma\,(ct' \,+\, \beta\,x') \qquad & y \;&=\; y' \qquad & z \;&=\; z' \end{aligned} \tag{8.11}$$

These are the famous **Lorentz transformations** parametrised in terms of β. They can be generalised to the case where $\mathbf{v}$ is in an arbitrary direction:

$$ct' \;=\; \gamma\,(ct \,-\, \boldsymbol{\beta}\cdot\mathbf{x}) \qquad\qquad \mathbf{x}' = \mathbf{x} \,+\, \frac{\gamma - 1}{\beta^2}\,(\boldsymbol{\beta}\cdot\mathbf{x})\,\boldsymbol{\beta} \,-\, \gamma\,\boldsymbol{\beta}\,ct \tag{8.12}$$

According to this expression, the projection of $\mathbf{x}$ on $\hat{\boldsymbol{\beta}}$ transforms as $\hat{\boldsymbol{\beta}}\cdot\mathbf{x}' = \gamma[\hat{\boldsymbol{\beta}}\cdot\mathbf{x} - \beta\,ct]$, as expected. The projection perpendicular to $\hat{\boldsymbol{\beta}}$ transforms as $\mathbf{x}' - (\hat{\boldsymbol{\beta}}\cdot\mathbf{x}')\hat{\boldsymbol{\beta}} = \mathbf{x} - (\hat{\boldsymbol{\beta}}\cdot\mathbf{x})\hat{\boldsymbol{\beta}}$ and therefore does not change.

The Lorentz transformations go over to the Galilean transformations in the so-called nonrelativistic limit, $\beta \ll 1$, or $\gamma \approx 1$. Simultaneous events in S, say at $(x,t) = (0,0)$ and $(d,0)$, are transformed to $(x',t') = (0,0)$ and $(\gamma d, -\gamma\beta d)$ in S', illustrating the relativity of simultaneity. Also, an array of synchronised clocks at rest in S' and observed, say at $t = 0$, by an observer at rest in S, are *not* synchronised in S. Indeed, in S they are observed to read $t' = -\gamma v x$, x being their position in S.

If the S observer observes a *single* clock *at rest* in S', then $\Delta t = \gamma \Delta t'$, since $\Delta x' = 0$. Calling the rest-frame of the clock its **proper frame** (here S'), and $\Delta\tau = \Delta t'$ the **proper time**, we have $\Delta t = \gamma\Delta\tau \geq \Delta\tau$. Moving clocks run slow, by a factor γ. On the other hand, if at time t our S observer measures a stick at rest in S' which has length $\Delta x'$ in that frame, (s)he will find $\Delta x = L_0/\gamma$, where $L_0 \equiv \Delta x'$ is the **proper length** of the stick (the length in its own rest-frame). Here we have length contraction since $\Delta x \leq L_0$.

The proper time of a particle is closely related to the line element, $\mathrm{d}s^2$. Indeed:

$$\mathrm{d}s^2 \;=\; c^2\,\mathrm{d}t^2 \;-\; \mathrm{d}l^2 \;=\; c^2\,\mathrm{d}t^2\left[1 \;-\; \left(\frac{\mathrm{d}l}{\mathrm{d}(ct)}\right)^2\right] \;=\; c^2\,\mathrm{d}t^2\,(1 \;-\; \beta^2) = \frac{1}{\gamma^2}\,c^2\,\mathrm{d}t^2 \tag{8.13}$$

where β is the speed of the particle. Then $\mathrm{d}s = c\mathrm{d}t/\gamma = c\mathrm{d}\tau$, and we find that proper time is also Lorentz-invariant.

Suppose one is observing a particle moving at speed β, where β is not necessarily constant. Then $d\tau$ is the infinitesimal proper time element of this particle (the time in its rest-frame), and dt is the observer's infinitesimal time element (observer time). One can then relate macroscopic observer time to macroscopic proper time of the particle:

$$\Delta\tau = \int_0^{\Delta t} \sqrt{1-\beta^2(t')}\,\mathrm{d}t', \qquad\qquad \Delta t \;=\; \int_0^{\Delta\tau} \gamma(\tau')\,\mathrm{d}\tau' \tag{8.14}$$

Note that time dilation is no longer symmetrical when there is acceleration! This is the solution to the "twin paradox".

8.4.1 Addition of Velocities (section J11.4)

The Lorentz velocity transformations are rather complicated, reflecting the fact that the usual velocity is no longer the best way to describe the rate of change of position. Let $\boldsymbol{\beta}_u = \mathrm{d}_{ct}\mathbf{x} = \mathbf{u}/c$. Its components transform as:

$$\begin{aligned}
\beta'_{ux} &= \frac{\mathrm{d}x'}{\mathrm{d}(ct')} \;=\; \frac{\gamma\,[\mathrm{d}x \;-\; \beta\,\mathrm{d}(ct)]}{\gamma[\mathrm{d}(ct) \;-\; \beta\,\mathrm{d}x]} \;=\; \frac{\mathrm{d}x/\mathrm{d}(ct) \;-\; \beta}{1 \;-\; \beta\,\mathrm{d}x/\mathrm{d}(ct)} \;=\; \frac{\beta_{ux} \;-\; \beta}{1-\beta\beta_{ux}} \\
\beta'_{uy} &= \frac{\mathrm{d}y'}{\mathrm{d}(ct')} \;=\; \frac{1}{\gamma}\,\frac{\mathrm{d}y}{\mathrm{d}(ct) \;-\; \beta\,\mathrm{d}x} \;=\; \frac{1}{\gamma}\,\frac{\mathrm{d}y/\mathrm{d}(ct)}{1 \;-\; \beta\,\mathrm{d}x/\mathrm{d}(ct)} \;=\; \frac{1}{\gamma}\,\frac{\beta_{uy}}{1 \;-\; \beta\beta_{ux}} \\
\beta'_{uz} &= \frac{\mathrm{d}z'}{\mathrm{d}(ct')} \;=\; \frac{1}{\gamma}\,\frac{\beta_{uz}}{1 \;-\; \beta\,\beta_{ux}}
\end{aligned} \tag{8.15}$$

These automatically preserve the speed of light: if $\beta_{ux} = 1$, $\beta'_{ux} = 1$. Warning: These expressions assume that $\boldsymbol{\beta}$ is in the positive x direction. If it is in the negative x direction, all the signs in front of terms containing β must be reversed. Moreover, the sign of u is *implicit* and must be put in.

Another, more compact, formulation for the velocity transformations is:

$$\begin{aligned}
\beta'_{u\parallel} &= \frac{\beta_{u\parallel} \;-\; \beta}{1 \;-\; \boldsymbol{\beta}\cdot\boldsymbol{\beta}_u} \\
\boldsymbol{\beta}'_{u\perp} &= \frac{\boldsymbol{\beta}_{u\perp}}{\gamma_v\left(1 \;-\; \boldsymbol{\beta}\cdot\boldsymbol{\beta}_u\right)}
\end{aligned} \tag{8.16}$$

The inverse transformations are easily obtained by interchanging primed and unprimed u and reversing the sign of $\boldsymbol{\beta}$.

As an EXERCISE, derive the Lorentz transformations for acceleration.

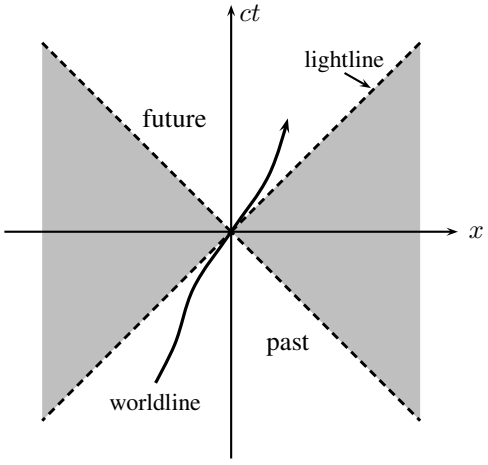
Advice: When solving problems that involve length, distance, or time interval comparisons between frames, it is often best to list all the Lorentz transformations, from S to S' and backward, and scan them to find the one(s) leading to the solution with the least amount of calculations. Also, although which frame you call S and which S' is a purely arbitrary matter, *the proper frame appropriate to the problem isn't*! It is a good idea to identify it first; the other frame will be the **observer** frame. When using the time dilation and length contraction equations, check that you do indeed get the right relative sizes. If you get the opposite, you have probably misidentified the proper frame.

Finally, we rewrite Lorentz boosts in the x direction in a more compact form, in preparation for the four-vector formalism that we will introduce shortly: $x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}$, where the indices $\mu$and ν run from 0 to 3, and we have used the summation convention over repeated indices. More explicitly, with $\tanh\,\theta = \beta$:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cosh\theta & -\sinh\theta & 0 & 0 \\ -\sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{8.17}$$

8.4.2 Causal structure

The importance of the line element cannot be overstated. Indeed, take any spacetime point and put the origin at that point. Plot a two-dimensional cut in the x-ct plane for purposes of illustration, showing also the spacetime trajectories that light can follow from this point.



The two **lightlines** mark the intersection of the **light-cone** with the $x - ct$ plane. The light-cone divides spacetime into two regions: one in which spacetime points can be connected to the event at the origin by signals that travel slower than light (that's the interior of the light-cone, closest to the ct axis), and the other in which points can only be connected to the origin by faster-than-light signals. Also shown is the **worldline** of a particle, giving its position x as a function of ct.

Now look at:

$$(\Delta s)^2 \;=\; (ct)^2 \;-\; x^2 \;=\; (ct')^2 \;-\; x'^2 = (\Delta s')^2$$

If $(\Delta s)^2 = 0$, the interval is said to be **null** or **lightlike**. Notice that the light-cone is *absolute*, frame-independent!

The spacetime region for which $(\Delta s)^2 > 0$ (**timelike interval**) lies inside the light-cone of the event at the origin. For any spacetime point inside the light-cone, it is always possible to find a frame S' in which the events occur at the same spatial position: $x'^2 = 0$. No event inside the light-cone is simultaneous with the event at the origin, and it is impossible to find a frame in which it is.

The spacetime region for which $(\Delta s)^2 < 0$ (**spacelike interval**) lies outside the light-cone. For any spacetime point in that region, it is always possible to find a frame S' in which it is simultaneous with the event at the origin. In fact, there exist frames in which their time ordering is *reversed*.

The implications are clear: if causality is to make sense, ie. if time ordering is to remain invariant under a change of frame, the event at the origin can only influence or be influenced by events that lie *inside* its light-cone. This region is the only one to be causally connected to the origin. Each point in spacetime thus has an absolute future, its forward light-cone made of all the spacetime points it can influence, and an absolute past, its backward light-cone containing all the spacetime points which could have influenced it. In other words, the spacetime interval between causally connected events is always **timelike**.

8.5 Four-dimensional Tensors: the Four-vector Formalism (section J11.6)

The most general Lorentz transformations combine a boost and a spatial rotation; they have a determinant equal to 1: $\|\Lambda^\mu{}_\nu\| = 1$. Here we shall focus on boosts in the x direction only. Recall how we wrote such Lorentz transformations:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \qquad (\mu, \nu = 0, 1, 2, 3)$$

where $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$, and:

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}}, \qquad \boldsymbol{\beta} = \hat{\mathbf{x}}\, v/c$$

In analogy with the 3-vectors at the beginning of the chapter, we introduce the coordinate-free notation $\mathbf{x}$ for the spacetime position with coordinates x^μ. We give ourselves four basis 4-vectors, $\mathbf{e}_{(\mu)}$, for our four-dimensional spacetime, where μ runs from 0 to 3. From now on we use Greek for spacetime indices and Roman for spatial indices $(1, 2, 3)$. Some authors do the reverse, and you may come across older books in which spacetime indices run from 1 to 4, with 4 corresponding to the time index; but it is quite standard these days to use 0 for the time index. Note that the index μ on $e_{(\mu)}$ says which basis 4-vector, not which component! In a basis $\mathbf{e}_{(\mu)}$, then, $\mathbf{x} = x^\mu \mathbf{e}_{(\mu)}$. As before, repeated indices must be summed over.

We take the transformation of the position 4-vector $\mathbf{x}$ as defining all 4-vectors:

Definition 8.3. A Lorentz 4-vector a *is any object whose components, a^μ, transform as $a'^\mu = \Lambda^\mu{}_\nu a^\nu$ under a Lorentz transformation $\mathbf{\Lambda}$.*

8.5.1 Minkowski metric

A **Lorentz scalar** is invariant under a Lorentz transformation. We already know one such scalar: the line element:

$$\begin{aligned} \mathrm{d}s^2 &= (\mathrm{d}x^0)^2 - (\mathrm{d}x^1)^2 - (\mathrm{d}x^2)^2 - (\mathrm{d}x^3)^2 \\ &= \eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu \end{aligned} \tag{8.18}$$

where, in a Cartesian basis:

$$\eta_{\mu\nu} \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{8.19}$$

$\boldsymbol{\eta}$ is known as the **Minkowski tensor**, or **metric tensor for flat spacetime**; that it is a tensor will be shown below. Also, in Cartesian coordinates only, its compnents are themselves invariant under Lorentz boosts. The overall sign of $\boldsymbol{\eta}$ is a matter of convention; I have been consistent with Jackson's (rather unfortunate!) choice. Here, we see one of the two important jobs of a metric tensor: to define distances on a space.

8.5.2 Relation between covariant and contravariant components of a 4-tensor

The components a^μ of a 4-vector $\mathbf{a}$, where $\mathbf{a} = a^\mu \mathbf{e}_{(\mu)}$, are said to be **contravariant**. But it is also useful to form the set $a_\mu = \eta_{\mu\nu} a^\nu$, which are called the **covariant** components of the *same* 4-vector $\mathbf{a}$.

Clearly, in Cartesian coordinates, $a_0 = a^0$, and $a_i = -a^i$. Thus, unlike in three-dimensional space, covariant and contravariant components in spacetime are in general different, and it is crucial to keep track of them! This is due, of course, to the fact that η_{00} always has its sign opposite to the sign of η_{ii}. Because of this, mathematicians say that the metric of flat spacetime is **pseudo-Euclidean** (not Euclidean as would be the case if all the diagonal elements had the same sign, like in three-dimensional space). In the *curved* spacetimes of Einstein's theory of gravitation, contravariant and covariant components differ by much more than a sign. Even in flat (non-curved) Minkowski spacetime, using curvilinear cooridnates brings in a Minkowski tensor with space-space diagonal components different from 1,

The Minkowski tensor, with covariant components $\eta_{\mu\nu}$, defines the spacetime distance; it also provides a link between not only the contravariant and covariant compoents of a vector, as we have just seen, but also those of a general tensor of any rank. covariant and contravariant. The operation that obtains covariant components from contravariant ones is often called **lowering an index**. The contravariant components of a tensor of rank r can also be lowered to yield its covariant components. Take for instance a rank-2 tensor. Then: $T_{\mu\nu} = \eta_{\mu\alpha}\eta_{\nu\beta}T^{\alpha\beta}$. Or only one index can be lowered to produce **mixed** components: $T^\mu{}_\alpha = \eta_{\alpha\beta}T^{\mu\beta}$, $T_\alpha{}^\nu = \eta_{\alpha\beta}T^{\beta\nu}$. Rank-$n$ tensors can have lots of components of different type!

Needless to say, if indices can be lowered, they can also be raised. That is, given covariant components, one can obtain their contravariant counterparts:

$$a^\mu = \eta^{\mu\nu} a_\nu = \eta^{\mu\nu}\eta_{\nu\rho}a^\rho \tag{8.20}$$

But $a^\mu = \delta^\mu{}_\rho a^\rho$, where $\delta^\mu{}_\rho$ is the Kronecker delta in four dimensions. Comparing, we have that $\eta^{\mu\nu}\eta_{\nu\rho} = \delta^\mu{}_\rho$, which means that $\eta^{\mu\nu}$ is the inverse of $\eta_{\nu\rho}$, and the contravariant and covariant *Cartesian* components of $\boldsymbol{\eta}$ are identical. This is also consistent with the fact that $\eta^{\mu\nu} = \eta^{\mu\alpha}\eta^{\nu\beta}\eta_{\alpha\beta}$. Conclusion: we raise indices with $\eta^{\mu\nu}$ and lower them with $\eta_{\mu\nu}$. For a rank-n tensor, we simply apply as many times $\boldsymbol{\eta}$ as the number of indices that we want to raise or lower, being careful to implement all the rules about repeated and free indices that we introduced in three dimensions and which still hold in four.

The **norm** of a 4-vector $\mathbf{a}$ is simply $\mathbf{a}\cdot\mathbf{a}$, or $a^\nu a_\nu = (a^0)^2 - \mathbf{a}^2$ since we often write $\mathbf{a} = (a^0, \vec{a})$. It is a Lorentz invariant (or scalar). So $\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \mathrm{d}x^\nu \mathrm{d}x_\nu$ is the norm of $\mathbf{dx}$.

As could have been gathered from the transformation law that we wrote earlier for a tensor of rank r, covariant and contravariant indices do not transform in quite the same way. How do the covariant components of a 4-vector transform then? Well, certainly we must write (formally) $a'_\mu = a_\beta L^\beta{}_\mu$, but what is $L^\beta{}_\mu$? We invoke the Lorentz invariance of $a^\mu a_\mu$:

$$a'_\mu a'^\mu \;=\; a_\beta L^\beta{}_\mu \Lambda^\mu{}_\nu a^\nu \;=\; a^\beta a_\beta \;=\; \delta^\beta{}_\nu a^\nu a_\beta$$

which we rewrite: $\left(L^\beta{}_\mu \Lambda^\mu{}_\nu \;-\; \delta^\beta{}_\nu\right) a^\nu a_\beta \;=\; 0$. Therefore, $L^\beta{}_\mu \Lambda^\mu{}_\nu = \delta^\beta{}_\nu$, so that $\mathbf{L} = \boldsymbol{\Lambda}^{-1}$.

This can also be used to write the inverse Lorentz transformation in the 4-vector formalism. I claim that $a^\nu = (\Lambda^{-1})^\nu{}_\mu a'^\mu$. Indeed, since $a'^\mu = \Lambda^\mu{}_\rho a^\rho$, we have:

$$a^\nu \;=\; (\Lambda^{-1})^\nu{}_\mu \Lambda^\mu{}_\rho a^\rho \;=\; \delta^\nu{}_\rho a^\rho \;=\; a^\nu$$

Just as in three-dimensional space, we can extend the transformation law to define Lorentz tensors T of higher rank, whose components must transform like:

$$T'^{\mu\nu\cdots}{}_{\lambda\rho\ldots} = (\Lambda^{-1})^\alpha{}_\lambda (\Lambda^{-1})^\beta{}_\rho \Lambda^\mu{}_\gamma \Lambda^\nu{}_\sigma \, T^{\gamma\sigma\cdots}{}_{\alpha\beta\ldots} \tag{8.21}$$

where one Lorentz matrix or its inverse must be applied for each index.

8.5.3 Constructing Lorentz tensors

Like all tensors, Lorentz tensors can be constructed from other Lorentz tensors. For instance, the trace of a tensor T of rank 2 is $T^\mu{}_\mu = T_\nu{}^\nu$. And if $\mathsf{T} = \mathbf{a} \otimes \mathbf{b}$, then $T^\nu{}_\nu = a^\nu b_\nu = a_\nu b^\nu$, the scalar (inner) product $\mathbf{a} \cdot \mathbf{b}$ of two 4-vectors. Note that:

$$\begin{aligned} a^\mu b_\mu \;&=\; a^0 b_0 \,+\, a^1 b_1 \,+\, a^2 b_2 \,+\, a^3 b_3 \\ &=\; \eta_{\mu\nu} a^\mu b^\nu \;=\; a^0 b^0 \,-\, a^1 b^1 \,-\, a^2 b^2 \,-\, a^3 b^3 \end{aligned}$$

The generalisation of the vector product of two 3-vectors is called the **exterior product** of two covectors: $\mathbf{a} \wedge \mathbf{b}$, which is an antisymmetric rank-2 tensor (or 2-form) with components $a_\mu b_\nu - a_\nu b_\mu$. This is valid in any basis! If we insist on writing it in terms of the components of the associated *vectors*, however, we must first raise the indices on both a_μ and b_μ with the metric tensor. The same is true in three dimensions: *only* in a Cartesian basis are the contravariant and covariant components the same. The rank-2 covariant tensor with components $T_{ij} = a_i b_j - a_j b_i$ can then be converted to those of a vector via the operation $\frac{1}{2}\epsilon^{ijk} T_{jk}$, which is an example of a **Hodge dual**..

Here is a property that can be useful when one wishes to establish the tensor character of an object: if $T^{\mu\nu} a_\mu b_\nu$ is a scalar for *any* 4-vectors $\mathbf{a}$ and $\mathbf{b}$, then$\mathsf{T}$ is a tensor. Moreover, if $T^{\mu\nu} a_\mu$ is a 4-vector for *any* 4-vector $\mathbf{a}$, then T is a tensor.

Example 8.1. Under a Lorentz transformation:

$$\begin{aligned} \mathrm{d}s'^2 \;&=\; \mathrm{d}x'_\nu \mathrm{d}x'^\nu \;=\; \eta'_{\mu\nu} \mathrm{d}x'^\mu \mathrm{d}x'^\nu \\ &=\; \eta'_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \, \mathrm{d}x^\alpha \mathrm{d}x^\beta \end{aligned}$$

But $\mathrm{d}s'^2 = \mathrm{d}s^2 = \eta_{\alpha\beta} \mathrm{d}x^\alpha \mathrm{d}x^\beta$. Comparing with the last line gives:

$$\mathrm{d}x^\alpha \mathrm{d}x^\beta (\eta'_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta - \eta_{\alpha\beta}) \;=\; 0$$

Therefore, $\eta_{\alpha\beta} = \eta'_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta$, so that $\boldsymbol{\eta}$ is, as claimed before, a rank-2 covariant Lorentz tensor.

Before putting the four-vector formalism to good use, one may well ask why all this fuss about tensors. Why do we bother with them? The reason lies in the transformation laws for a Lorentz tensor T. Being homogeneous, they ensure that if $\mathsf{T} = 0$ in some inertial frame, $\mathsf{T} = 0$ in *any other* frame. In other words, if you can show that *all* the components of a tensor vanish in some frame, you have shown that they vanish in all frames. Now, any equation can be put in the form: something is equal to zero. If that something is a tensor, then *the equation holds in all frames*! By building the fundamental equations of physics out of Lorentz tensors, we guarantee that they satisfy the Einstein relativity principle.

If $\mathsf{T}(\mathbf{x})$ is a tensor field of rank r, the equation $\mathsf{T}(\mathbf{x}) = 0$ contains in fact 4^r equations, one for each component $T^{\mu\nu\lambda\cdots}$. This is similar to what happens in three-dimensional space, where $\mathsf{T}(\mathbf{x}) = 0$ contains 3^r equations. Each

equation is automatically satisfied in all frames when properly transformed, even if it may not *look* that way. Only $\mathsf{T}(\mathsf{x}) = 0$ itself satisfies the relativity principle in a *manifest* way. Many people say that the tensor equation is **Lorentz-covariant**, but I don't want to use this terminology because the word "covariant" has already been assigned a different meaning. Instead, I will speak of **form-invariant** expressions, meaning that the *form* of the expression does not change under Lorentz transformations.

8.6 Differential Operators in the Four-vector Formalism

Form the object with components $\partial_\mu = \partial/\partial x^\mu = \big(\partial_{ct},\, \partial_i\big)$. How does it behave under Lorentz transformations? This can be established easily from the chain rule for partial differentiation:

$$\frac{\partial}{\partial x'^\mu} \;=\; \frac{\partial x^\nu}{\partial x'^\mu}\,\frac{\partial}{\partial x^\nu}$$

But since $x^\nu = (\Lambda^{-1})^\nu{}_\alpha x'^\alpha$ (inverse Lorentz transformation):

$$\frac{\partial x^\nu}{\partial x'^\mu} \;=\; (\Lambda^{-1})^\nu{}_\alpha\,\frac{\partial x'^\alpha}{\partial x'^\mu} \;=\; (\Lambda^{-1})^\nu{}_\alpha\,\delta^\alpha{}_\mu \;=\; (\Lambda^{-1})^\nu{}_\mu$$

Therefore:

$$\frac{\partial}{\partial x'^\mu} \;=\; (\Lambda^{-1})^\nu{}_\mu\,\frac{\partial}{\partial x^\nu} \tag{8.22}$$

which is the Lorentz transformation for the *covariant* components of a 4-vector—more precisely, the compoenents of a 4-covector. The object we have constructed is the **4-gradient** operator,$\boldsymbol{\nabla}$, with naturally covariant components $\partial_\mu = (\partial_{ct},\, \partial_i)$ and whose contravariant components read: $\partial^\mu = \partial/\partial x_\mu = \eta^{\mu\nu}\partial_\nu = (\partial_{ct},\, -\partial_i)$.

We construct first derivatives of Lorentz tensors:

- The **4-gradient** of a scalar field $f(\mathsf{x})$, the covector or **1-form** $\boldsymbol{\nabla} f$, with components $\partial_\mu f$. Note that $\mathrm{d}f = (\boldsymbol{\nabla} f)\cdot\mathbf{dx} = (\partial_\mu f)\mathrm{d}x^\mu$.

- The **divergence** of a 4-vector field $\mathbf{a}(\mathsf{x})$, $\boldsymbol{\nabla}\cdot\mathbf{a}$, a Lorentz scalar; in index notation, this is written $\partial_\mu a^\mu = \partial^\mu a_\mu = \partial_{ct}a^0 + \boldsymbol{\nabla}^{(3)}\cdot\mathbf{a}$, where $\boldsymbol{\nabla}^{(3)}$ is the 3-dim gradient operator..

- The exterior derivative of a 4-dim 1-form field, $\mathbf{c} = \boldsymbol{\nabla}\wedge\mathbf{a}$, which is an antisymmetric covariant tensor of rank 2. (In three dimensions, the exterior derivative of a covector is also a rank-2 covariant tensor whose form is basis-independent; it is converted to a 3-vector with the Levi-Civita symbol.) The matrix form of the components $c_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$, with $a^i = -a_i$ the components of a 3-vector $\mathbf{a}$, is:

$$\begin{pmatrix} 0 & -(\partial_{ct}a_x + \partial_x a_0) & -(\partial_{ct}a_y + \partial_y a_0) & -(\partial_{ct}a_z + \partial_z a_0) \\ \partial_{ct}a_x + \partial_x a_0 & 0 & -(\boldsymbol{\nabla}^{(3)}\times\mathbf{a})_z & (\boldsymbol{\nabla}^{(3)}\times\mathbf{a})_y \\ \partial_{ct}a_y + \partial_y a_0 & (\boldsymbol{\nabla}^{(3)}\times\mathbf{a})_z & 0 & -(\boldsymbol{\nabla}^{(3)}\times\mathbf{a})_x \\ \partial_{ct}a_z + \partial_z a_0 & -(\boldsymbol{\nabla}^{(3)}\times\mathbf{a})_y & (\boldsymbol{\nabla}^{(3)}\times\mathbf{a})_x & 0 \end{pmatrix} \tag{8.23}$$

- The curl of a 2-form, or covariant antisymmetric tensor field $\mathsf{T}(\mathsf{x})$ of rank 2, $\boldsymbol{\nabla}\wedge\mathsf{T}$, is a covariant antisymmetric tensor of rank 3, a 3-form. In index notation, its components are written as:

$$\partial_\rho T_{\mu\nu} + \partial_\nu T_{\rho\mu} + \partial_\mu T_{\nu\rho} \tag{8.24}$$

A (p+1)-form can always be constructed by taking the exterior derivative of a completely antisymmetric covariant tensor (aka p-form), but when p is also the dimension of the space in which these objects live, the result must vanish (WHY?). So, in four dimensions, only p-forms up to $p = 4$ can exist, with 4-forms being of the form: $f(\mathsf{x})\epsilon_{\mu\nu\rho\lambda}$, in which the index structure is entirely carried by the Levi-Civita symbol.

Among tensors formed from double partial derivatives, only the **d'Alembertian** is nontrivial: $\Box \equiv \partial_\mu \partial^\mu = \partial_{ct}^2 - \nabla^{(3)^2}$. (Note that some authors use the notation $\Box^2$ for the d'Alembertian.) The classical wave equation for a scalar wave propagating in vacuum at the speed of light is:

$$\Box\, f(\mathbf{x}) \;\equiv\; \boldsymbol{\nabla}\cdot\boldsymbol{\nabla} f \;\equiv\; \partial_\mu\partial^\mu f \;=\; \left(\frac{1}{c^2}\partial_t^2 - \nabla^{(3)^2}\right) f = 0 \tag{8.25}$$

Its general solution has the form $f(\mathbf{k}\cdot\mathbf{x})$, where $\mathbf{k}$ is called the **wave 4-vector**, with $k^0 = \omega/c$ and $\mathbf{k} = (2\pi/\lambda)\hat{\mathbf{k}}$, $\hat{\mathbf{k}}$ giving the direction of propagation; ω is the angular frequency and λ the wavelength of the wave. Then the phase $\mathbf{k}\cdot\mathbf{x} = k_\mu x^\mu = \omega t - \mathbf{k}\cdot\mathbf{x}$, and consistency with the wave equation imposes $\mathbf{k}\cdot\mathbf{k} = 0$, or $(\omega/c)^2 - (2\pi/\lambda)^2 = 0$, which gives, as expected, $c = \omega/k$.

It should be stressed that care must be exercised with the spatial derivatives if a non-Cartesian basis is used, because of those pesky unit vectors or, equivalently, the metric coefficients that inevitably contribute when raising indices of components, plus $\sqrt{|g|}$ factors that come from taking Hodge duals. It is safer to stick to Cartesian bases, as we will do, especially when dealing with the d'Alembertian of tensors of rank ≥ 1, like $\Box\, a^\mu$, or $\Box\, T^{\mu\nu}$.

What about identities on these objects? There is only one, but it is of utmost importance: the famous Poincaré lemma, which says that the exterior derivative applied twice on any p-form vanishes identically, and of which $\boldsymbol{\nabla}\times\boldsymbol{\nabla} f = 0$ and $\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\times\mathbf{a} = 0$ are just two 3-dim examples. It follows that if the exterior derivative of a 1-form is zero everywhere, this 1-form is the 4-gradient of some scalar field (assuming a simply connected space).

8.7 Relativistic Kinematics and Dynamics

8.7.1 4-velocity and 4-acceleration

Having introduced the spacetime position 4-vector, we seek to define a velocity that transforms the right way under Lorentz transformations, ie. as a 4- vector. We could try $\mathbf{U} = \mathbf{dx}/\mathrm{d}t$, an object with four components. But because $\mathrm{d}t$ does not transform as a Lorentz tensor, $\mathbf{dx}/dt$ is not a Lorentz tensor. If we could find a time that behaves as a scalar in replacement of $\mathrm{d}t$, we could construct a tensor. But we do have such a beast at hand: the proper time $\mathrm{d}\tau = \mathrm{d}t/\gamma$. So:

$$\mathbf{U} \;=\; \frac{\mathbf{dx}}{\mathbf{d}\boldsymbol{\tau}} \tag{8.26}$$

is a 4-vector. How is this 4-velocity related to the usual velocity $\mathrm{d}\mathbf{x}/\mathrm{d}t$? We write in index notation:

$$U^\mu \;=\; \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \;=\; \gamma\frac{\mathrm{d}x^\mu}{\mathrm{d}t} \;=\; \left(\gamma\frac{\mathrm{d}x^0}{\mathrm{d}t}, \gamma\,\mathbf{u}\right) \;=\; \gamma\,(c, \mathbf{u}) \tag{8.27}$$

where $\mathbf{u} = \mathrm{d}\mathbf{x}/dt$, and $\gamma = 1/\sqrt{1-\beta^2}$ with $\boldsymbol{\beta} = \mathbf{u}/c$. *Note that the components of* $\mathbf{u}$ *must be multiplied by* γ *to obtain the spatial components* U^j *of* $\mathbf{U}$! The 4-velocity transforms properly : $U'^\mu = \Lambda^\mu{}_\nu U^\nu$, unlike $\mathbf{u}$ whose transformation laws are much more complicated because of the $\mathrm{d}t$ in the denominator.

The covariant components of $\mathbf{u}$ are immediately obtained as: $U_\mu = \gamma(c, -\mathbf{u})$, and its norm is $U^\nu U_\nu = c^2$.

While we're at it, we might as well construct the 4-acceleration $\boldsymbol{\Gamma} = \mathbf{dU}/\mathbf{d}\boldsymbol{\tau}$. EXERCISE: Find the components Γ^μ in terms of $\mathbf{u}$ and $\mathbf{a}$, where $\mathbf{a} = \mathrm{d}\mathbf{u}/dt$ is the Newtonian 3-acceleration. Then also find the norm of $\boldsymbol{\Gamma}$ in terms of $\mathbf{u}$ and $\mathbf{a}$, and show that $U^\mu\Gamma_\mu = 0$ (hint: take the proper time derivative of the norm of $\mathbf{U}$).

8.7.2 Energy-momentum four-vector (section J11.5)

Consider the object $\mathbf{P} = m\mathbf{U}$, where m is mass. Now $\mathbf{P}\cdot\mathbf{P} = m^2\mathbf{U}\cdot\mathbf{U} = (mc)^2$, so that we can take m to be a Lorentz scalar and therefore $\mathbf{P}$ a Lorentz 4-vector. The spatial components $P^i = mU^i$ are interpreted as the

components of **relativistic momentum**: $\mathbf{p} = \gamma m \mathbf{u}$, where $\mathbf{u} = \mathrm{d}\mathbf{x}/\mathrm{d}t$ and $\gamma = 1/\sqrt{1-\beta_u^2}$. They go over to the Newtonian definition of momentum in the nonrelativistic limit $\gamma \approx 1$.

What makes momentum so interesting a quantity is the fact that it is conserved in processes during which no external force acts: the total momentum of an isolated system remains constant, ie. the same after and before the process. This conservation law should obey the relativity principle. But Newtonian momentum does not satisfy the Einstein relativity principle, only the Galilean one. This destroys its usefulness since conservation becomes frame-dependent. By contrast, thanks to its tensor nature, the 4-momentum P is conserved whatever the Lorentz frame in which we work. Moreover, since P is a 4-vector, we get a conservation law not only for relativistic momentum $\mathbf{p}$, but for the 0-component of P. For a process involving two bodies, we write this as:

$$\left(\gamma_1 m_1 c^2 + \gamma_2 m_2 c^2\right)_{\text{initial}} = \left(\gamma_1 m_1 c^2 + \gamma_2 m_2 c^2\right)_{\text{final}}$$

We interpret $\gamma m c^2 = P^0 c$ as the total energy of a body. This explains why P is often called the energy-momentum 4-vector instead of 4-momentum. The contravariant components of P are thus $P^\mu = (E/c, \mathbf{p})$, and its covariant components $P_\mu = (E/c, -\mathbf{p})$, where once again $\mathbf{p}$ is the *relativistic* 3-momentum $\gamma m \mathbf{u}$.

If $E = \gamma m c^2$ is the total energy of a body, what kind of energies contribute to it? Notice first that for a body at rest $E = E_0 = mc^2$, which means that a body has energy merely by virtue of having a mass. Now:

$$E - E_0 \;=\; (\gamma - 1) mc^2 \underset{\gamma \to 1}{\longrightarrow} \left[1 + \frac{1}{2}\frac{u^2}{c^2} + \ldots - 1\right] mc^2 \;=\; \frac{1}{2}\, m\, u^2 \tag{8.28}$$

Thus, $E - E_0$ goes over to the Newtonian kinetic energy in the limit $\gamma \approx 1$ and can be interpreted as the *relativistic* kinetic energy. So we find that E contains the rest-energy of a body and its kinetic energy, but not any potential energy it may have as a result of sitting in some field.

Writing the norm of P explicitly in terms of energy and momentum yields a very important and useful relation between E and $\mathbf{p} = \gamma m \mathbf{u}$:

$$E^2 \;=\; (pc)^2 + (mc^2)^2 \tag{8.29}$$

It is also clear that we can regard γ as the total energy of a body in units of its rest-energy; the total energy is often a better parameter of the motion than the velocity $\mathbf{u}$, because of the complicated nonlinear dependence of γ on $\mathbf{u}$. In real life—inasmuch as the word can apply to high-energy physics experiments at accelerators—all particles move at speeds so close to the speed of light that speed is no longer a useful parameter; motion is usually characterized by the energy of the particle.

Relativity allows the existence of particles with *zero* rest-energy (or zero mass), *as long as they move at the speed of light*, such as the photon. In this case, the expression $E = \gamma m c^2$ does not hold since it gives $0/0$, but $E^2 = (pc)^2 + (mc^2)^2$ yields $E = pc$. Usually, it is the energy of the zero-mass particle which is known experimentally, and one can then deduce its momentum via $p = E/c$. Note that for zero-mass particles, $\mathbf{p} \neq \gamma m \mathbf{u}$, again because γm is indeterminate in that case.

There is another useful criterion for deciding whether an object's motion is relativistic. If the energy quoted is smaller than the rest-energy E_0, it obviously refers to kinetic energy. If that is *much* smaller than E_0, the **non-relativistic** approximation, ie. the usual Newtonian expressions for all quantities, may be used. If not, relativistic expressions must be used, as also when the energy quoted is larger than E_0, in which case it should be specified whether this is the total or kinetic energy.

If the energy given is *much* larger than E_0, one can make the **ultrarelativistic** approximation which consists in ignoring E_0, treating the particle as if it were massless. Then $E \approx pc$. (Aside: when you calculate an energy, it is quite permissible to make one of the two approximations if you suspect that it is justified. If your answer is consistent with your approximation, you were right; if not, you must use the full relativistic expressions.)

8.7.3 Relativistic dynamics

Having introduced the 4-acceleration $\mathbf{\Gamma}$, it is but a short step to construct a 4-force, the **Minkowski force** $\mathbf{K} = m\mathbf{\Gamma}$. Let us take a closer look at its components:

$$K^0 \;=\; m\,\frac{\mathrm{d}U^0}{\mathrm{d}\tau} \;=\; \frac{1}{c}\,\frac{\mathrm{d}(\gamma\, mc^2)}{\mathrm{d}\tau} \;=\; \frac{1}{c}\,\frac{\mathrm{d}E}{\mathrm{d}\tau} \tag{8.30}$$

K^0 therefore measures the proper rate at which energy is being transferred to or from the particle by a force acting on it. Which force? Well:

$$K^i \;=\; m\,\frac{\mathrm{d}U^i}{\mathrm{d}\tau} \;=\; \frac{\mathrm{d}P^i}{\mathrm{d}\tau} \;=\; \gamma\,\frac{\mathrm{d}p^i}{\mathrm{d}t} \tag{8.31}$$

So we can still use the usual definition of the force, $\mathbf{F} = \mathrm{d}\mathbf{p}/dt$, but with $\mathbf{p}$ the *relativistic* 3-momentum: $\mathbf{p} = \gamma m\mathbf{u}$. So the K^i are not the components of $\mathbf{F}$, although they are related to them: $K^i = \gamma F^i$.

Next, work out the invariant $\mathbf{K}\cdot\mathbf{U}$:

$$K^\nu U_\nu \;=\; \left(\frac{1}{c}\,\gamma\,\frac{\mathrm{d}E}{\mathrm{d}t}\right)(\gamma\, c) \;-\; (\gamma\,\mathbf{F})\cdot(\gamma\mathbf{u})$$

But we already know that $\mathbf{K}\cdot\mathbf{U} = 0$ since $\mathbf{\Gamma}\cdot\mathbf{U} = 0$. Thus:

$$\frac{\mathrm{d}E}{\mathrm{d}t} \;=\; \mathbf{F}\,\cdot\,\mathbf{u} \tag{8.32}$$

whose form is then Lorentz-invariant, although not manifestly so. Integrating both sides, we arrive at the work-energy theorem $\Delta E = \int \mathbf{F}\cdot\mathrm{d}\mathbf{l}$, the work done by $\mathbf{F}$. The form of the theorem is the same as in Newtonian physics, but with $E = \gamma\, mc^2$ and $\mathbf{F} = \mathrm{d}(\gamma m\mathbf{u})/\mathrm{d}t$.

The vanishing of $\mathbf{K}\cdot\mathbf{U}$ has another interesting consequence: it means that at least one component of $\mathbf{K}$ must depend on $\mathbf{U}$. Assuming that it is linear, the dependence should be written so as to satisfy the relativity principle. The only possible form is $\mathbf{K} = \lambda\,\mathbf{X}\cdot\mathbf{U}$, where $\mathbf{X}$ is a rank-2 tensor and λ is some Lorentz-invariant constant that matches the units on both sides. In index notation, $K^\mu = \lambda\, X^{\mu\nu}U_\nu$. By writing $\mathbf{K}\cdot\mathbf{U}$ in component notation, we immediately show that since $\mathbf{U}$ is arbitrary, the symmetric part of $\mathbf{X}$ must vanish, and $\mathbf{X}$ is an antisymmetric rank-2 tensor.

The relativistic generalisation of the Newtonian expression $F_\mathrm{c} = m\,u^2/R$ for the centripetal force keeping a body in uniform motion at speed u on a circular trajectory of radius R is $F_\mathrm{c} = \gamma m\,u^2/R$.

Relativistic Electrodynamics

8.8 Current-density 4-vector

It seems appropriate to start with the source of the electric field, the charge density $\rho = nq$, where n is the number density of particles all carrying charge q. If all of these move with a common velocity $\mathbf{u}$, there is a current density $\mathbf{j} = \rho\mathbf{u} = nq\mathbf{u}$. Now, calling n_0 the proper number density (density in the rest-frame of the charges), we have $n = \gamma n_0$ from length contraction. Then, since it is an experimental fact that q is a Lorentz invariant (see p. J554), the object $\mathbf{J}$ with contravariant components:

$$J^\mu \;=\; (\rho c, \mathbf{j}) = n_0\, q\,(\gamma c, \gamma\mathbf{u}) = n_0\, q\, U^\mu \tag{8.33}$$

is postulated to be a 4-vector.

We also know that ρ and $\mathbf{j}$ satisfy a conservation law, called the continuity equation, $\partial_t\rho + \boldsymbol{\nabla}^{(3)}\cdot\mathbf{j} = 0$, which does obey the relativity principle, but not *manifestly* so. Indeed:

$$\boldsymbol{\nabla}\cdot\mathbf{J} \;=\; \partial_\mu J^\mu \;=\; \partial_{ct}(\rho c) \;+\; \boldsymbol{\nabla}^{(3)}\cdot\mathbf{j} \;=\; 0$$

We say that **J** is a **conserved 4-current density** (we often omit "density"...). Note that the word "conservation" here does not refer to states before and after a process! The conservation of **J**, *like all conservation laws taken in that specific sense (continuity)*, is expressed in the 4-vector formalism by the vanishing of its 4-divergence, which is a *manifestly* form-invariant statement.

8.9 4-vector Potential

Recall the Poisson equation for the electrostatic potential: $\nabla^2\Phi = -4\pi\rho$ (in Gaussian units!). It is certainly not manifestly form-invariant under Lorentz boosts, although it is under spatial rotations. Indeed, time derivatives arise from the transformation of the Laplacian although one might think that "static" should be a Lorentz-invariant property. Moreover, the charge density, ρ, which behaves as a scalar (a 3-tensor) under spatial rotations, becomes in spacetime the 0^{th} component of a 4-vector. This means that if, as we should, we write our theory in terms of Lorentz tensors, the Poisson equation will be the 0^{th} component of a more general tensor equation taken in the static limit.

Rewrite the Poisson equation in the equivalent form $\nabla^2\Phi = -4\pi\rho = -4\pi J^0/c$. The spatial part of the tensor equation of which the Poisson equation will be the 0^{th} component should contain derivatives of a 3-vector potential, $\mathbf{A}$, and a right-hand side proportional to $\mathbf{j}$. Therefore, define the 4-dim potential as a 1-form $\mathbf{A}$ with components $A_\mu = (\Phi, -\mathbf{A})$ (this is more natural than a vector).

The differential tensor equation we seek for $\mathbf{A}$ must have 4-vector character since its right-hand side is a 4-vector ($\mathbf{J}$). It must be *linear* in $\mathbf{A}$ so as to reduce correctly to the Poisson equation. And it must contain second-order space and time derivatives. The most general equation which meets these criteria is:

$$\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\mathbf{A} \;+\; a\,\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) \;+\; b\,\mathbf{A} \;=\; d\,\mathbf{J}$$

where a, b, and d are constants to be determined. A possible constant in front of the first term can always be absorbed in the other ones. In index notation:

$$\partial_\mu\partial^\mu A_\nu \;+\; a\,\partial_\nu(\partial^\mu A_\mu) \;+\; b\,A_\nu \;=\; d\,J_\nu$$

Consider the static case: $\partial_t\mathbf{A} = 0$. The 0^{th} component ($\nu = 0$) of the tensor equation is $-\nabla^2\Phi + b\Phi = d\,J^0$. Then consistency with the Poisson equation imposes $b = 0$ and $d = 4\pi/c$.

[Aside: if $b \neq 0$, the solution of the resulting equation is $\Phi(r) = e^{-\alpha r}/r$ in vacuum in the spherically-symmetric case, where $\alpha^2 = -b$. (EXERCISE: check this.) This is not the Coulomb potential, but the *short-range* potential associated with a *massive* photon.]

Now take the divergence of the tensor equation to obtain:

$$\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) \;+\; a\,\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) \;=\; \frac{4\pi}{c}\,\boldsymbol{\nabla}\cdot\mathbf{J}$$

that is, $(1+a)\Box\,(\boldsymbol{\nabla}\cdot\mathbf{A}) = 0$, since $\boldsymbol{\nabla}\cdot\mathbf{J} = 0$. Thus, conservation of $\mathbf{J}$ implies that $a = -1$. (If instead $\Box\,(\boldsymbol{\nabla}\cdot\mathbf{A}) = 0$, $\boldsymbol{\nabla}\cdot\mathbf{A}$ would be sourceless; also, A_0 would obey a wave equation; it can be shown that if A_0 propagates, this leads to unacceptable behaviour since A_0 would propagate negative energy to infinity.)

We end up with:

$$\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\,\mathbf{A} \;-\; \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) \;=\; \frac{4\pi}{c}\,\mathbf{J} \tag{8.34}$$

or

$$\partial^\mu\partial_\mu A_\nu \;-\; \partial_\nu(\partial^\mu A_\mu) \;=\; \frac{4\pi}{c}\,J_\nu \tag{8.35}$$

in terms of covariant components.

8.10 Faraday-Maxwell Field Tensor and Maxwell's Equations (section J11.9)

Since we can interchange the partial derivatives in the second term (in flat spacetime only!), we can write:

$$\partial^\mu(\partial_\mu A_\nu \;-\; \partial_\nu A_\mu) \;=\; \frac{4\pi}{c}\, J_\nu$$

which becomes:

$$\partial^\mu F_{\mu\nu} \;=\; \frac{4\pi}{c}\, J_\nu \tag{8.36}$$

(note the order of the indices!), or, if we define $\mathsf{F} \equiv \nabla \wedge \mathbf{A}$:

$$\boldsymbol{\nabla} \cdot \mathsf{F} \;=\; \frac{4\pi}{c}\, \mathsf{J} \tag{8.37}$$

Clearly, being constructed from 4-vectors, F is a 2-form (rank-2 covariant tensor) with components $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, so that F is antisymmetric. Its contravariant components are $F^{0i} = -F_{0i}$, and $F^{ij} = F_{ij}$.

Since F is the exterior derivative of a 1-form, *its* exterior derivative must be identically zero, by Poincaré's lemma: $\boldsymbol{\nabla} \wedge \mathsf{F} = 0$. In index notation:

$$\partial_\rho F_{\mu\nu} \;+\; \partial_\nu F_{\rho\mu} \;+\; \partial_\mu F_{\nu\rho} \;=\; 0.$$

Now we must give a physical interpretation for F. As the first derivative of a potential, it is a field (in the physical sense of the word) and should somehow contain the electric field. Since $A_\mu = (\Phi,\, A_i) = (\Phi,\, -A^i)$, we have:

$$F_{0i} \;=\; \partial_0 A_i \;-\; \partial_i A_0 \;=\; -\,(\boldsymbol{\nabla}^{(3)}\Phi \;+\; \partial_{ct}\mathbf{A}) \tag{8.38}$$

Seeing that this reduces to $\mathbf{E}$ in the static case, we identify the time-space components of F, F_{0i}, with E^i.

The three other independent components of F, F_{ij}, are equivalent to the components of the 3-dim curl of $\mathbf{A}$ in eq. (8.23). The three non-zero independent space-space components of F then form a new 3-dim field $\mathbf{B} \equiv \boldsymbol{\nabla}^{(3)} \times \mathbf{A}$ which we call the **magnetic induction** field. The matrix form of the covariant and contravariant components of the **Faraday-Maxwell electromagnetic field tensor**, F, reads:

$$F_{\mu\nu} \;=\; \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \qquad F^{\mu\nu} \;=\; \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \tag{8.39}$$

In three-dimensional space, $\mathbf{E}$ and $\mathbf{B}$ are perfectly good -tensors (vectors); in spacetime, however, their true nature is revealed: their components are *together the components of a covariant, antisymmetric tensor of rank 2*, the electromagnetic Faraday-Maxwell field.

Let us now write the two equations for F, the source equation $\boldsymbol{\nabla} \cdot \mathsf{F} = 4\pi\mathsf{J}/c$, and $\boldsymbol{\nabla} \wedge \mathsf{F} = 0$, explicitly in terms of time and space derivatives ($3+1$ formalism).

First:

$$\begin{aligned} \partial^i F_{i0} \;&=\; \partial_i F^{i0} \;=\; \frac{4\pi}{c}\, J^0 \\ \Longrightarrow \qquad \partial_i E^i \;&=\; \frac{4\pi}{c}\, \rho\, c \\ \Longrightarrow \qquad \boldsymbol{\nabla}^{(3)} \cdot \mathbf{E} \;&=\; 4\pi\, \rho \qquad\qquad \text{Gauss} \end{aligned} \tag{8.40}$$

and then, considering the spatial components of $\partial_\mu F^{\mu\nu} \;=\; \frac{4\pi}{c}\, J^\nu$:

$$\partial_0 F^{0j} \,+\, \partial_i F^{ij} \;=\; -\,\frac{1}{c}\,\partial_t E^j \,+\, \partial_i \epsilon^{ijk}\, B_k \;=\; \frac{4\pi}{c}\, J^j$$

$$\Longrightarrow \qquad \boldsymbol{\nabla}^{(3)} \times \mathbf{B} \,-\, \frac{1}{c}\,\partial_t \mathbf{E} \;=\; \frac{4\pi}{c}\,\mathbf{j} \qquad\qquad \text{Ampère} \tag{8.41}$$

where we have used[†] $F^{ij} = \epsilon^{ijk} B_k$. These are the **inhomogeneous Maxwell equations** in $3+1$ form (Gaussian units!) .

The other equation, $\boldsymbol{\nabla} \wedge \mathsf{F} = 0$, is somewhat easier to handle if we first introduce the **dual** of F: $\star\mathsf{F}$. Now, instead of this more standard notation, we shall put $\boldsymbol{\mathcal{F}} \equiv \star\mathsf{F}$ in order to be consistent with the textbook. In index notation:

$$\boldsymbol{\mathcal{F}}^{\mu\nu} \;\coloneqq\; \frac{1}{2}\epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta} \tag{8.42}$$

where $\boldsymbol{\epsilon}$ is the skew-symmetric Levi-Civita 4-tensor of rank 4 (Cartesian coordinates!), with properties analogous to those of the 3-dim Levi-Civita symbol and $\epsilon^{0123} = +1$. Thus, F and its dual contain exactly the same information, but rearranged: time-space components become space-space components (up to a sign), and vice-versa. Indeed:

$$\begin{aligned}
\boldsymbol{\mathcal{F}}^{0i} \;&=\; \frac{1}{2}\epsilon^{0ijk} F_{jk} \;=\; \frac{1}{2}\epsilon^{ijk}\,(\partial_j A_k \,-\, \partial_k A_j) \;=\; \epsilon^{ijk}\,\partial_j A_k \;=\; -\,B^i \\
\boldsymbol{\mathcal{F}}^{ij} \;&=\; \frac{1}{2}\epsilon^{ij\alpha\beta}\, F_{\alpha\beta} \;=\; \frac{1}{2}(\epsilon^{ij0k}\, F_{0k} \,+\, \epsilon^{ijk0}\, F_{k0}) \;=\; \epsilon^{ij0k}\, F_{0k} \;=\; \epsilon^{0ijk}\, F_{0k} \;=\; \epsilon^{ijk}\, E_k
\end{aligned} \tag{8.43}$$

The advantage of $\boldsymbol{\mathcal{F}}$ is that because of the way it depends on the potentials, it satisfies a differential identity:

$$\partial_\mu \boldsymbol{\mathcal{F}}^{\mu\nu} \;=\; \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}\,\partial_\mu(\partial_\alpha A_\beta - \partial_\beta A_\alpha) \;\equiv\; 0 \tag{8.44}$$

which is equivalent to $\boldsymbol{\nabla} \wedge \mathsf{F} = 0$ but is easier to expand into components (four instead of 64!):

$$\partial_i \boldsymbol{\mathcal{F}}^{i0} \;= \partial_i B^i \;=\; \boldsymbol{\nabla}^{(3)} \cdot \mathbf{B} \;=\; 0 \qquad\qquad \text{Gauss} \tag{8.45}$$

$$\partial_0 \boldsymbol{\mathcal{F}}^{0j} \,+\, \partial_i \mathcal{F}^{ij} \;=\; -\left(\frac{1}{c}\,\partial_t B^j \,+\, \epsilon^{jik}\partial_i E_k\right) \;=\; 0$$

$$\Longrightarrow \qquad \boldsymbol{\nabla}^{(3)} \times \mathbf{E} \,+\, \frac{1}{c}\,\partial_t \mathbf{B} \;=\; 0 \qquad\qquad \text{Faraday} \tag{8.46}$$

These are the **homogeneous Maxwell equations**, again in Gaussian units. One may object that since they are identities, it should be impossible to retrieve any information from them. But $\boldsymbol{\nabla} \wedge \mathsf{F} = 0$ (or $\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{F}} = 0$) becomes an identity only when one writes F in terms of the potential A. It does not contain any information about A, but if one works only with the electromagnetic field, F, it can (and must) be used to solve for $\mathbf{E}$ and $\mathbf{B}$.

In Lorentz-tensor form, Maxwell's equations are then:

$$\boldsymbol{\nabla} \cdot \mathsf{F} \;=\; \frac{4\pi}{c}\,\mathsf{J}, \qquad\qquad \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{F}} \;=\; 0 \tag{8.47}$$

or:

$$\partial_\mu F^{\mu\nu} \;=\; \frac{4\pi}{c}\, J^\nu \qquad\qquad \partial_\mu \boldsymbol{\mathcal{F}}^{\mu\nu} \;=\; 0$$

Because F and $\boldsymbol{\mathcal{F}}$ are both antisymmetric, the left-hand sides of Maxwell's equations have identically vanishing divergence: $\partial_\mu \partial_\nu F^{\mu\nu} \equiv 0$, from which $\partial_\nu J^\nu = 0$, or $\boldsymbol{\nabla} \cdot \mathsf{J} = 0$. We see that Maxwell's equations incorporate

[†]We will be consistent the unusual convention implied by eq. J11.139: $\epsilon^{123} = 1$, $\epsilon_{123} = -1$. We also have: $\epsilon_{ijk}\epsilon^{ilm} = -(\delta_j{}^l \delta_k{}^m - \delta_j{}^m \delta_k{}^l)$ (note the overall minus sign that comes from Jackson's choice $\eta_{ii} = -1$!). Other useful expressions: $B^i = -\epsilon^{ijk}\partial_j A_k = -\frac{1}{2}\epsilon^{ijk} F_{jk}$, and $(\boldsymbol{\nabla} \times \mathbf{B})^i = -\epsilon^{ijk}\partial_j B_k$.

conservation of the source current as *an identity on the field equations*. This is a general feature of field theories with conserved sources.

If one chooses to work with the electromagnetic potential $\mathbf{A}$, there is but one Maxwell equation (instead of the two for $\mathbf{F}$), and we have already found it:

$$\Box\,\mathbf{A} \;-\; \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) \;=\; \frac{4\pi}{c}\,\mathbf{J} \tag{8.48}$$

Once $\mathbf{A}$ is known, $\mathbf{F} = \boldsymbol{\nabla}\wedge\mathbf{A}$ yields the electromagnetic field tensor.

By taking the curl of Maxwell's equation (8.48) for $\mathbf{A}$, we obtain, invoking the Poincaré lemma:

$$\Box\,(\boldsymbol{\nabla}\wedge\mathbf{A}) \;-\; \boldsymbol{\nabla}\wedge\cancelto{0}{\nabla(\boldsymbol{\nabla}\cdot\mathbf{A})} \;=\; \Box\,\mathbf{F} \;=\; \frac{4\pi}{c}\,\boldsymbol{\nabla}\wedge\mathbf{J} \tag{8.49}$$

or:

$$\Box\,F_{\mu\nu} \;=\; \frac{4\pi}{c}\,(\partial_\mu J_\nu \;-\partial_\nu J_\mu)$$

which is a classical wave equation for $\mathbf{F}$.

8.11 Lorentz Transformations of the Faraday Field Tensor (section J11.10)

Since $\mathbf{F}$ is a rank-2 tensor, its contravariant components transform as:

$$F'^{\mu\nu} \;=\; \Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta F^{\alpha\beta} \tag{8.50}$$

We would like to multiply the three terms on the right hand-side in their matrix form. Before, though, this right-hand side must be rewritten so that it looks like a product of matrices, ie. $F'^{\mu\nu} = \Lambda^\mu{}_\alpha F^{\alpha\beta}\tilde{\Lambda}_\beta{}^\nu$, where $\tilde{\Lambda}_\beta{}^\nu$ is the transpose (rows and columns interchanged) of $\Lambda^\nu{}_\beta$. In other words, $\mathbf{F'} = \boldsymbol{\Lambda}\mathbf{F}\tilde{\boldsymbol{\Lambda}}$. For a Lorentz boost in the x direction, this is fairly simple:

$$\begin{pmatrix} 0 & -E'_x & -E'_y & -E'_z \\ E'_x & 0 & -B'_z & B'_y \\ E'_y & B'_z & 0 & -B'_x \\ E'_z & -B'_y & B'_x & 0 \end{pmatrix}$$

is equivalent to:

$$\begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

This yields immediately:

$$\begin{aligned} E'_x &= E_x & \qquad B'_x &= B_x \\ E'_y &= \gamma\,(E_y \;-\; \beta\,B_z) & \qquad B'_y &= \gamma\,(B_y \;+\; \beta\,E_z) \\ E'_z &= \gamma\,(E_z \;+\; \beta\,B_y) & \qquad B'_z &= \gamma\,(B_z \;-\; \beta\,E_y) \end{aligned} \tag{8.51}$$

Here are the transformed electromagnetic components for a general Lorentz boost $\boldsymbol{\beta}$. They naturally split into components longitudinal ($\|$) and transverse ($\perp$) to the motion:

$$\begin{aligned} \mathbf{E}'_\| &= \mathbf{E}_\|, & \qquad \mathbf{E}'_\perp &= \gamma\,[\mathbf{E}_\perp + \boldsymbol{\beta}\times\mathbf{B}] \\ \mathbf{B}'_\| &= \mathbf{B}_\|, & \qquad \mathbf{B}'_\perp &= \gamma\,[\mathbf{B}_\perp - \boldsymbol{\beta}\times\mathbf{E}] \end{aligned} \tag{8.52}$$

See also eq. (J11.149) for a very useful equivalent form. It is now clear that what we call electric or magnetic fields are frame-dependent objects; the two transform into one another. Notice that, unlike 4-vectors, whose transverse components are unaffected, the *longitudinal* components of rank-2 tensors do not change and the transverse ones do. We immediately see that if one can transform to a frame where $\mathbf{B}' = 0$, the second line forces the boost to be perpendicular to both $\mathbf{B}$ and $\mathbf{E}$, and its magnitude to satisfy $\mathbf{B} = \boldsymbol{\beta} \times \mathbf{E}$, ie., $\beta = B/E < 1$.

Indeed, whenever $\mathbf{E}$ or $\mathbf{B}$ vanish in some frame, the transformations are very simple and can be used to solve problems which otherwise would be rather complicated. Take the electromagnetic field of a point-charge in uniform motion. In the rest-frame of the charge, $\mathbf{B}' = 0$ and $E' = q/r'^2$. Without loss of generality, put the charge at the origin of its rest-frame S' and let S' move along the x axis of S at speed v; locate the observation point P in the x'-y' plane at $(x', y', 0)$ in the charge's frame. Then the electric field has one longitudinal component, $E'_{x'}$, and two transverse components, $E'_{y'}$ and $E'_{z'}$. The electric components in S (lab-frame) are obtained from the ones in S' via the Lorentz transformations, using $E_{y'} = \gamma\,(E'_{y'} + \beta\,B'_{z'})$ and $E_{z'} = \gamma\,(E'_{z'} - \beta\,B'_{y'})$:

$$E_{x'} = E'_{x'} = qx'/r'^3 \qquad E_{y'} = \gamma\,E'_{y'} = \gamma\,qy'/r'^3 \qquad E_{z'} = \gamma\,E'_{z'} = \gamma\,qz'/r'^3 \tag{8.53}$$

These are the fields in the S frame expressed *in terms of S' coordinates*. To complete the transformation, we must also transform the coordinates of the point of observation P from S' to S. Referring to fig. J11.8, but without choosing any special location in the $x = 0$ plane (transverse to the direction of motion) for the observation point P, we have $\sqrt{y'^2 + z'^2} = \sqrt{y^2 + z^2} = r\sin\psi$, and $x' = \gamma(x - \beta ct) = \gamma r_x = -\gamma r\cos\psi$, where r is the distance from the charge's *present* position to the observation point, and ψ is the angle between $\mathbf{r} = r\hat{\mathbf{n}}$ and the x axis. Therefore:

$$r'^2 = x'^2 + y'^2 + z'^2 = r^2\big(\gamma^2\cos^2\psi + \sin^2\psi\big) = \gamma^2\,r^2\big(1 - \beta^2\sin^2\psi\big)$$

Combining everything gives the field in terms of the charge's present position:

$$\mathbf{E} = q\,\frac{1}{\gamma^2}\,\frac{1}{(1 - \beta^2\sin^2\psi)^{3/2}}\,\frac{1}{r^2}\,\hat{\mathbf{n}} \tag{8.54}$$

The electric field points away from the *instantaneous* position of the charge, ie. the point where it is at the time of observation *if it always moves at speed v*. One could well ask what happened to causality here. Clearly, the field registered at P at time t must be the one generated by q at time $t - r/c$, the so-called retarded time (see chapter 6). The fact is that the direction of the field at time t does not say that the charge *must* be at the point where the field is pointing from at that time. If the charge changes direction or speed *after* $t - r/c$, the "news" of this won't be known at P until the resulting change in the field, propagating at the speed of light, reaches P.

We can gain further insight by writing eq. (8.53) as a function of time. For the sake of simplicity, we now choose P to lie at $(x, y) = (0, b)$, as in fig. J11.8). Then $x' = -\gamma vt$, $y' = b$, $r'^2 = b^2 + (vt')^2 = b^2 + (\gamma vt)^2$, and there comes:

$$\mathbf{E}(\mathbf{x}) = \frac{\gamma\,q}{\big(b^2 + \gamma^2v^2t^2\big)^{3/2}}\,(-vt\,\mathbf{i} + b\,\mathbf{j}) \tag{8.55}$$

Plotted against vt, the transverse component E_y looks like a pulse of width b/γ, whereas the longitudinal component E_x is positive for $t < 0$ and negative for $t > 0$, and is appreciable also only for a time interval b/γ. This means that at high γ, only the transverse component is readily observable.

The magnetic field of the point-charge in the lab-frame is easily obtained from $\mathbf{B} = \boldsymbol{\beta} \times \mathbf{E}$:

$$\mathbf{B} = \frac{q}{\gamma^2}\,\frac{1}{(1 - \beta^2\sin^2\psi)^{3/2}}\,\frac{\boldsymbol{\beta} \times \hat{\mathbf{n}}}{r^2} \tag{8.56}$$

With the geometry of fig. J11.8, this means that the only component of $\mathbf{B}$ is $B_z = \beta E_y$.

In the nonrelativistic limit, $\beta \ll 1$, this reduces to the Biot-Savart form:

$$\mathbf{B} = \frac{q}{c}\,\frac{\mathbf{v} \times \hat{\mathbf{n}}}{r^2} \tag{8.57}$$

It is easy to see that B is down by a factor β compared to E. Also, it is perpendicular both to the electric field and to the direction of motion of the source charge.

In the ultrarelativistic case ($\gamma \gg 1$), on the other hand, the electric and magnetic fields of a point-charge look more and more like that of a pulse of linearly polarised radiation propagating in the x direction.

Electromagnetic invariants can be formed from the Maxwell field tensor and its dual; this is left as an EXERCISE (problem J11.14). The sign of these invariants tells us which of the electric or the magnetic components can be transformed away by performing a Lorentz boost to some other frame.

8.12 Gauge Invariance of Maxwell's Theory

The most important property of Maxwell's equations—apart from manifestly obeying the Einstein relativity principle—is their invariance under a completely different type of transformation.

Add to $\mathbf{A}$ the 4-gradient of an arbitrary scalar field f: $\mathbf{A} \rightarrow \mathbf{A} + \boldsymbol{\nabla} f$, or $\delta\mathbf{A} = \boldsymbol{\nabla} f$. Then $\mathbf{F}$ is *manifestly* invariant under this transformation: $\mathbf{F} \rightarrow \boldsymbol{\nabla} \wedge (\mathbf{A} + \boldsymbol{\nabla} f) = \mathbf{F} + \boldsymbol{\nabla} \wedge \boldsymbol{\nabla} f = \mathbf{F}$, or $\delta\mathbf{F} = 0$, since the 4-curl of a 4-gradient is identically zero.

On the other hand, Maxwell's equation (8.48) for $\mathbf{A}$ changes by:

$$\Box(\delta\,\mathbf{A}) \;-\; \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\delta\,\mathbf{A}) \;=\; \Box(\boldsymbol{\nabla} f) \;-\; \boldsymbol{\nabla}\,(\Box f) \;\equiv\; 0 \tag{8.58}$$

We conclude that Maxwell's theory, whether expressed in terms of fields or of potentials, is *invariant under the* **gauge transformation** $\delta\,\mathbf{A} = \boldsymbol{\nabla} f$, where the **gauge function** f is an *arbitrary*, differentiable scalar function of position and time. These transformations are now seen to lie at the origin of the arbitrariness in Maxwell's theory discussed in section 5.2.

. This arbitrariness cancels out of $\mathbf{F} = \boldsymbol{\nabla} \wedge \mathbf{A}$, but not of $\boldsymbol{\nabla} \cdot \mathbf{A}$ which changes by $\Box f$ under a gauge transformation. We can remove the arbitrariness by imposing so-called gauge conditions, of which there are many in use. They can formally be shown, sometimes laboriously, to be accessible via some gauge function. Usually, though, a gauge condition (or just **gauge**) is set directly, without exhibiting the gauge transformation that sends $\mathbf{A}$ into something that satisfies the condition.

One very useful and popular gauge choice is the **Lorenz gauge condition** (do not confuse Lorenz with Lorentz!), $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$. It has the advantage of being fully consistent with Relativity—or **Lorentz covariant**. Also, our 4-potential equation now assumes the simple form:

$$\Box\mathbf{A} \;=\; \frac{4\pi}{c}\,\mathbf{J} \tag{8.59}$$

as we found in chapter 5 of these notes, but this time in Gaussian units. Because it is one relation between the four components of $\mathbf{A}$, the Lorenz condition brings the number of independent degrees of freedom down to three.

It is important to note, however, that there is still some arbitrariness left, since $\boldsymbol{\nabla} \cdot \mathbf{A}$ is invariant under a **restricted** gauge transformation with any gauge function satisfying $\Box f = 0$. When all arbitrariness has been removed (see example below), only two independent degrees of freedom are left, as expected.

Although it is possible to find formal wave solutions for A^0, as in Griffiths' book, for instance, we recall that the propagation of A^0 is an artefact the arbitrary Lorenz gauge choice. In fact, A^0 is really relevant only in static solutions. Its real nature is revealed by working out the zero-component of eq. (8.48) without any gauge choice. We see that the second-order time derivatives cancel out, so A^0 does not really satisfy a wave equation.

Working instead in the Coulomb gauge, $\boldsymbol{\nabla}^{(3)} \cdot \mathbf{A} = 0$, we have shown in our chapter 5 that A^0 and the longitudinal component of $\mathbf{A}$ are not propagating degrees of freedom, leaving only two propagating (or dynamical) degrees of freedom. The Coulomb gauge belongs to the class of so-called **non-covariant** gauges because it is not form-invariant under Lorentz transformations. Consequently, it must be reimposed "by hand" every time one changes frame. By contrast, the Lorenz condition belongs to the class of **covariant** gauge choices since it is

manifestly form-invariant. This does not mean that non-covariant gauges must be shunned, simply that one should use them only when Lorentz form-invariance is not important for what we're doing.

Example 8.2. As noted above, it is possible to prove that Maxwell's theory contains only two dynamical degrees of freedom if one works with the Lorenz condition. First, consider plane-wave solutions to $\Box\,\mathbf{A} = 0$: $\mathbf{A} = \mathbf{e}\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}$, where $\mathbf{e}$ is a constant 4-vector not determined by the wave equation. Inserting this back into the wave equation gives $\mathbf{k}\cdot\mathbf{k} = 0$. We also have $\mathbf{P} = \hbar\mathbf{k}$, which gives $\mathbf{P}\cdot\mathbf{P} = 0$. So the plane-wave solutions correspond to massless particles, called photons, travelling at the speed of light with 4-momentum $\mathbf{P}$.

Unlike the inhomogeneous solution of section 5.5, consistency of honogeneous solutions with $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$ is not automatic. Imposing it leads to: $\mathbf{k}\cdot\mathbf{A} = 0$. Taking the x axis along the direction of propagation, this becomes $k_0 A^0 + k_x A_x = 0$, while $\mathbf{k}\cdot\mathbf{k} = 0$ becomes $k_0 k^0 + k_x^2 = 0$, or $k_0 = \pm k_x \equiv k$, so that $A^0 = \pm A_x$.

Now, as pointed out above, imposing $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$ does not completely remove the arbitrariness in $\mathbf{A}$. The gauge function $f = f_0\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}$, which satisfies $\Box f = 0$, provides a further transformation that removes what is left of the arbitrariness. Indeed, $A_x \to (e_x + \mathrm{i}kf_0)\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}$, and since f_0 is an arbitrary constant, simply take $f_0 = -\mathrm{i}e_x/k$, which removes the longitudinal component of the transformed $\mathbf{A}$, as well as A^0 since $A^0 = \pm A_x$ still holds after this last gauge transformation. Again, only the two transverse components of $\mathbf{A}$ survive. Here, we had to resort to actual solutions in order to establish this, contrary to the reasoning in the Coulomb gauge.

More sophisticated methods to be discussed in the next chapter confirm, without fixing the gauge, that only two of the four quantities $(\Phi, \mathbf{A})$ can really propagate energy, in the sense that they—the transverse components of $\mathbf{A}$—are insensitive to any gauge transformation on $\mathbf{A}$. By contrast, both Φ and the longitudinal component of $\mathbf{A}$ change under gauge transformations, so that any wave character they may possess with one choice of gauge can be made to disappear with a different choice.

8.12.1 Are electromagnetic fields more fundamental than potentials?

In classical electromagnetism, the equation of motion for charges (aka the Lorentz force law) contains fields but no potentials. In fact, there is nothing in classical electromagnetism that cannot be derived also from the fields. Thus, for a long time after they were introduced, potentials were considered useful to simplify some calculations, but non-physical. One very often quoted argument held that potentials could not be physical because, unlike the fields, they were not gauge-invariant. But now we see from the IVP analysis of Maxwell's theory that it is possible to set consistent initial data for the *transverse* part of the vector potential, which could not be true if they were gauge-variant

In a seminal paper published in 1959, Aharonov and Bohm[†] proposed experiments, one of which involved two beams of electrons travelling on opposite sides of a very long, very thin solenoid. When a current flows in the solenoid, the magnetic field outside is strictly zero, but there exists a vector potential $\mathbf{A}$ so as to match the potential inside where there is a magnetic field $\mathbf{B}$. The circulation $\oint \mathbf{A}\cdot\mathrm{d}\mathbf{l}$ around and *outside* the solenoid is the flux $\int \mathbf{B}\cdot\mathrm{d}\mathbf{S}$ of the interior field through the cross-sectional area of the solenoid. Classically, however, this exterior vector potential cannot be observed: no field, so no Lorentz force.

The authors noted that electrons do not really obey the Lorentz force law (except in the classical limit), but instead Schrödinger's equation, which involves the potentials, not the fields. The wave-functions of the electrons would interfere and switching the magnetic field on and off ought to perturb the interference pattern, even though no force acted on the electrons. The effect that they (and others before) discussed was soon observed experimentally. Interestingly enough, however, in the final section of the paper the authors are still a bit confused about the gauge objection, since they think that it could be raised against them:

[†] http://link.aps.org/doi/10.1103/PhysRev.115.485.

> " The main objection that could be raised against the above suggestion is grounded in the gauge invariance of the theory. In other words, if the potentials are subject to the transformation $A_\mu \to A'_\mu = A_\mu + \partial\psi/\partial x_\mu$, where ψ is a continuous scalar function, then all the known physical quantities are left unchanged. As a result, the same physical behavior is obtained from any two potentials, $A_\mu(x)$ and $A'_\mu(x)$, related by the above transformation. This means that insofar as the potentials are richer in properties than the fields, there is no way to reveal this additional richness. It was therefore concluded that the potentials cannot have any meaning, except insofar as they are used mathematically, to calculate the fields."

They are reduced to say that the objection is wrong because counter-examples can be produced. In fact, $\oint \mathbf{A}\cdot \mathrm{d}\,\mathbf{l}$ is gauge-invariant because $\oint(\boldsymbol{\nabla} f)\cdot \mathrm{d}\mathbf{l} = 0$. Since the transverse $\mathbf{A}$ is itself gauge-invariant, only the longitudinal component $\mathbf{A}_\mathrm{L}$ varies, and it is really $\oint \mathbf{A}_\mathrm{L}\cdot \mathrm{d}\mathbf{l}$ that is unaffected by gauge transformations. Moreover, the high degree of symmetry makes knowledge of the curl sufficient to determine $\mathbf{A}$ everywhere in cylindrical coordinates: $\mathbf{A}_\text{out} = (BR^2/2\rho)\hat{\boldsymbol{\phi}}$, and $\mathbf{A}_\text{in} = (B\rho/2)\hat{\boldsymbol{\phi}}$, where B is the magnitude of the field inside and R the radius of the solenoid. The divergence of $\mathbf{A}$ vanishes *everywhere*, which ensures that $\mathbf{A}$ is transverse. This does not contradict the fact that $\boldsymbol{\nabla}\times\mathbf{A} = 0$ outside the solenoid. For $\mathbf{A}$ to be pure longitudinal there, the curl must vanish everywhere in a simply connected space, ie., a space where every closed loop in the space can be contracted to a point, which is certainly not the case here for loops around the solenoid. What *is* important is that electrons in the quantum-mechanical regime are described by a wave-function, with a phase that can be shown to depend on $\mathbf{A}$. Phase differences, of course, are responsible for interference, and when the beams go through a region of zero magnetic field but non-zero potential on *topologically* distinct paths (paths which cannot deformed into one another while leaving their ends fixed), $\mathbf{A}$ disturbs the interference pattern of the beams. (Strictly speaking, the word "path" is not so appropriate, though: in quantum mechanics paths are not well-defined, and this why there can be an interference effect.)

So we must conclude that electromagnetic potentials (or, more precisely, their transverse part) are physical, and that they contain more information than fields. Only in the quantum-mechanical regime, however, can this information be retrieved.

8.13 Duality Properties of Maxwell's Theory (J6.11)

Let us form the complex combination of the Faraday 2-form field strength and its dual: $\mathfrak{F} = \mathsf{F} + \mathrm{i}\mathcal{F}$. Likewise, define the 3-dim complex vector field $\boldsymbol{\mathcal{E}} = \mathbf{E} + \mathrm{i}\mathbf{B}$. Therefore, the complexified Faraday 2-form has components:

$$\mathfrak{F}_{\mu\nu} \;=\; \begin{pmatrix} 0 & \mathcal{E}_x & \mathcal{E}_y & \mathcal{E}_z \\ -\mathcal{E}_x & 0 & \mathrm{i}\,\mathcal{E}_z & -\mathrm{i}\,\mathcal{E}_y \\ -\mathcal{E}_y & -\mathrm{i}\,\mathcal{E}_z & 0 & \mathrm{i}\,\mathcal{E}_x \\ -\mathcal{E}_z & \mathrm{i}\,\mathcal{E}_y & -\mathrm{i}\,\mathcal{E}_x & 0 \end{pmatrix} \tag{8.60}$$

Maxwell's equations can then be written in a very compact way:

$$\partial_\mu \mathfrak{F}^{\mu\nu} \;=\; \frac{4\pi}{c}\, J^\nu_\mathrm{e} \tag{8.61}$$

or, in $3+1$ form:

$$\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{E}} \;=\; 4\pi\,\rho_\mathrm{e} \qquad\qquad \boldsymbol{\nabla}\times\boldsymbol{\mathcal{E}} \;-\; \mathrm{i}\,\partial_{ct}\boldsymbol{\mathcal{E}} \;=\; \mathrm{i}\,\frac{4\pi}{c}\,\mathbf{j}_\mathrm{e} \tag{8.62}$$

In this form, it is easy to see that, *in vacuum*, Maxwell's equations are invariant under the **duality rotation**:

$$\mathfrak{F} \longrightarrow \mathrm{e}^{\mathrm{i}\phi}\mathfrak{F}, \qquad\qquad \boldsymbol{\mathcal{E}} \longrightarrow \mathrm{e}^{\mathrm{i}\phi}\boldsymbol{\mathcal{E}} \tag{8.63}$$

with ϕ any real constant. Choosing $\phi = \pi/2$ gives $\mathbf{E} \longrightarrow -\mathbf{B}$ and $\mathbf{B} \longrightarrow \mathbf{E}$. This unsuspected symmetry is the reason for the similarity between the electric and magnetic fields from sources with the same geometrical configuration..

When sources are put in, however, eq. (8.61) is no longer invariant under duality rotations. Invariance can only be restored by including a magnetic charge density $\rho_{\rm m}$ and a magnetic current density $\mathbf{j}_{\rm m}$ so that now:

$$\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{E}} \;=\; 4\pi\rho \qquad\qquad \boldsymbol{\nabla}\times\boldsymbol{\mathcal{E}} \;-\; \mathrm{i}\,\partial_{ct}\boldsymbol{\mathcal{E}} \;=\; \mathrm{i}\,\frac{4\pi}{c}\,\boldsymbol{\mathcal{J}} \tag{8.64}$$

where now $\rho = \rho_{\rm e} + \mathrm{i}\,\rho_{\rm m}$ and $\boldsymbol{\mathcal{J}} = \mathbf{j}_{\rm e} + \mathrm{i}\,\mathbf{j}_{\rm m}$. when both electric and magnetic sources are included, the duality transformation with $\phi = \pi/2$ exchanges not only the fields, but also the electric and magnetic charges.

If instead one chooses $\tan\phi = -\rho_{\rm m}/\rho_e$, the magnetic charge and current densities are transformed to zero, and this, without changing the physical content of the theory! So, if all particles have the *same* universal ratio of magnetic to electric charge, we may not be able to tell whether magnetic charge exists.

As Jackson points out in his section 6.11, if a magnetic charge g exists, one can define a magnetic fine-structure constant $\alpha_{\rm m} = (g/4\pi)^2/(k_{\rm m}\hbar c)$ analogous to the electric one $k_{\rm e}e^2/\hbar c \approx 1/137$. Dirac has shown that $\alpha_{\rm m} \approx \mathrm{O}(1/\alpha_{\rm e})$, which means that the magnetic coupling is very strong.

Now, in field theory, one often expands in powers of a small coupling constant, but this is impossible for a large coupling constant like $\alpha_{\rm m}$. The previous considerations suggest that electromagnetic duality is in fact a weak-coupling/strong-coupling duality. There are theories such as quantum chromodynamics which have strong couplings. It might be possible, using this idea of duality, to transform to a weak equivalent of the theory and infer some information that remains valid in the strong-coupling regime. These ideas feature prominently in some of the current research on Yang-Mills theories (QCD) and string theories, which have been discovered to be related to each other by analogous duality transformations.

9 Dynamics of Charged Particles and Electromagnetic Fields

9.1 From the Minkowski Force to the Lorentz Force

We make use of the general expression derived earlier that relates the Minkowski force to the 4-velocity of a body on which that force acts: $\mathsf{K} = \lambda\,\mathsf{X}\cdot\mathsf{U}$, where X is some antisymmetric rank-2 tensor that contains the physics, and λ is a constant that normalises for the right units. In our case, one available antisymmetric rank-2 tensor is the Faraday field tensor, F. This means that λ must be equal to charge q/c for correct units. Therefore:

$$K^\mu \;=\; \mathrm{d}_t P^\mu \;=\; \frac{q}{c}\,F^{\mu\nu}\,U_\nu \tag{9.1}$$

where $U_\mu = \gamma(c,\,-\mathbf{u})$. The spatial components of K are:

$$K^i \;=\; \frac{q}{c}\left(F^{i0}\,U_0 \;+\; F^{ij}U_j\right) \;=\; q\left[E^i\,\gamma \;-\; \epsilon^{ijk}B_k\,(\gamma\,u_j/c)\right] \;=\; \gamma\,F^i$$

Then, with $\boldsymbol{\beta} = \mathbf{u}/c$, we can write the components F^i of the 3-force as:

$$\mathbf{F} \;=\; q(\mathbf{E} + \boldsymbol{\beta}\times\mathbf{B}) \tag{9.2}$$

which is the Lorentz force equation. For a charge distribution, multiply $K^\mu \;=\; q\,F^{\mu\nu}U_\nu/c$ by the observer-frame number density $n = \gamma n_0$. The charge density is $\rho = qn$, and, with $\mathsf{J} = \rho\mathsf{U}$, $n\mathsf{K}$ is a force *density* with components $F^{\mu\nu}J_\nu/c$.

As an EXERCISE, work out the meaning of the $0^{\rm th}$ component of the electromagnetic Minkowski-force equation.

Note: We have not really *derived* the equation of motion (9.1)! After all, $\boldsymbol{\mathcal{F}}$ could also have been used... We will now show that it can be obtained from a Lagrangian, using the principle of least action.

9.2 Very Short Review of Classical Lagrangian Formalism

Consider a system with a finite number N of degrees of freedom, $q^n(t)$ ($1 \leq n \leq N$). We assume that the dynamics is described by the **action integral**:

$$S \coloneqq \int_{t_\mathrm{i}}^{t_\mathrm{f}} L(q^n,\, \dot{q}^n)\,\mathrm{d}t \tag{9.3}$$

where the functional L is called the **Lagrangian** of the system, and the **velocities** $\dot{q}^n$ are taken to be independent of the q^n. In other words, to any trajectory $q^n(t)$ in $\mathbb{R}^n$ that takes the initial value $q^n_\mathrm{i} \coloneqq q^n(t_\mathrm{i})$ and the final value $q^n_\mathrm{f} \coloneqq q^n(t_\mathrm{f})$, we associate the *number* S given by the integral.

Time evolution of the system is obtained from the principle of least action: classical trajectories $q^n(t)$ are those for which the action is **stationary**, that is, arbitrary small variations $\delta q^n(t)$ lead to $\delta S = 0$. When we have found the trajectories that minimise S in that way, small deviations from these trajectories do not result in a change in S. Now, according to the calculus of variations, a variation of $L(q^n,\, \dot{q}^n)$ is written as:

$$\delta L \;=\; \frac{\partial L}{\partial q_n}\,\delta q_n + \frac{\partial L}{\partial \dot{q}_n}\,\delta \dot{q}_n \qquad\qquad (\text{summation over } n)$$

where $\delta q_n = \epsilon f_n(t)$ and $\delta \dot{q}_n = \epsilon \dot{f}_n(t)$, with ϵ a small parameter and $f_n(t)$ arbitrary functions. Inserting into $\delta S = \int \delta L \mathrm{d}t$, ϵ cancels out and there comes:

$$\int_{t_\mathrm{i}}^{t_\mathrm{f}} \left[\frac{\partial L}{\partial q_n}\,f_n + \frac{\partial L}{\partial \dot{q}_n}\,\dot{f}_n\right]\mathrm{d}t \;=\; 0 \tag{9.4}$$

Integrating the second term by parts and setting $f_n(t_\mathrm{i}) = f_n(t_\mathrm{f}) = 0$ (to keep the end-points of the trajectories fixed), we arrive at:

$$\int_{t_\mathrm{i}}^{t_\mathrm{f}} \left[\frac{\partial L}{\partial q_n} \;-\; \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_n}\right] f_n\,\mathrm{d}t \;=\; 0 \tag{9.5}$$

Since the f_n are arbitrary functions of time, the expression in the square bracket must itself vanish if it is continuous in time, and we have the **Euler-Lagrange** equations of motion:

$$\partial_n L \;-\; \mathrm{d}_t \pi_n \;=\; 0 \tag{9.6}$$

where the **generalised momenta conjugate to** q_n, $\pi_n = \partial L/\partial \dot{q}_n$, have been introduced.

Now the total time-derivative in the Euler-Lagrange equations can be rewritten as:

$$\dot{\pi}_n \;=\; \ddot{q}_m\,\frac{\partial \pi_n}{\partial \dot{q}_m} \;+\; \dot{q}_m\,\frac{\partial \pi_n}{\partial q_m}$$

This yields an interesting, more detailed form of the Euler-Lagrange equations:

$$\ddot{q}_m\,\frac{\partial^2 L}{\partial \dot{q}_m \partial \dot{q}_n} \;+\; \dot{q}_m\,\frac{\partial^2 L}{\partial q_m \partial \dot{q}_n} \;-\; \partial_n L \;=\; \ddot{q}_m\,\frac{\partial \pi_n}{\partial \dot{q}_m} \;+\; \dot{q}_m\,\frac{\partial \pi_n}{\partial q_m} \;-\; \partial_n L \;=\; 0 \tag{9.7}$$

Because this is a system of second-order differential equations, people speak of it as the **second-order formalism**. That the **Hessian matrix** $\partial^2 L/\partial \dot{q}_m \partial \dot{q}_n = \partial \pi_n/\partial \dot{q}_m$ of L has non-zero determinant is a necesssary and sufficient condition for the equations, and thus the accelerations, to be linearly independent. If the determinant vanishes, we say that the system is singular, in the sense that the matrix is not invertible, and the velocities as functions of the coordinates and the monenta are themselves not invertible. The system is also said to be constrained, because there exist relations of the type $\phi(q, \pi) = 0$ that make manifest the fact that the momenta are not all independent, In addition, $\ddot{q}_n$ is not *uniquely* determined by the coordinates and velocities at any given time, and the solution to the equations of motion contains arbitrary functions of time.

Example 9.1. A simple example will help illustrate this. Consider the Lagrangian $L = (\dot{q}_1 - \dot{q}_2)^2/2$, with conjugate momenta $\pi_1 = \dot{q}_1 - \dot{q}_2$ and $\pi_2 = \dot{q}_2 - \dot{q}_1$. The two momenta are not linearly independent since $\pi_1 + \pi_2 = 0$. As a result of this constraint, the whole velocity space $\dot{q}_1$-$\dot{q}_2$ is mapped to the line $\pi_2 = -\pi_1$ in momentum space. Each point accessible to the system in momentum space maps back to a straight *line* defined by $\dot{q}_2 - \dot{q}_1 =$ constant. At the level of the equations of motion, eq. (9.7) contains only one independent equation: $\ddot{q}_2 - \ddot{q}_1 = 0$.

9.3 Motion of a Charged Particle in an Electromagnetic Field from a Covariant Lagrangian (section J12.1)

We want a Lagrangian L that yields the equation of motion (9.1) and its 3-dim version, the Lorentz force law, eq. (9.2), for a particle of mass m and charge q. It must be a Lorentz scalar built from available 4-vectors, in this case the 4-potential A and the 4-veloocity U (dependence on the 4-position x would break translation invariance which is the foundation of the conservation of energy and momentum).

Now Lagrangians are not unique: for instance, we will not use Jackson's Lagrangian in eq. (J12.31). The reason lies outside the scope of our treatment; let us just say that his L leads to a Hamiltonian (not introduced here) that is not manifestly positive-definite. We take:

$$L \;=\; \frac{1}{2}\,m\,U^\mu U_\mu \;-\; \frac{q}{c}\,U^\nu A_\nu \;=\; \frac{1}{2}\,m\,\eta_{\mu\nu}U^\mu U^\nu \;+\; \frac{q}{c}\,U^\nu A_\nu \tag{9.8}$$

Remember that $U^\mu U_\mu = c^2$. But this constraint is not to be implemented at this stage, unless a Lagrange-multiplier term is added to L:

$$L \;=\; \frac{1}{2}\,m\,c^2 \;-\; \frac{q}{c}\,U^\nu A_\nu \;+\; \lambda(U^\beta U_\beta \;-\; c^2)$$

The covariant form of the Euler-Lagrange equations is: $\partial_\alpha L \;-\; \mathrm{d}_\tau(\partial L/\partial U^\alpha) = 0$, with $\partial_\alpha = \partial/\partial x^\alpha$. Because the kinetic term in L has no x^α dependence, only the electromagnetic term contributes:

$$\partial_\alpha L \;=\; \frac{q}{c}\,U^\beta\,\partial_\alpha A_\beta$$

On the other hand, the components of the conjugate momentum are:

$$\begin{aligned}\pi_\alpha \;=\; \frac{\partial L}{\partial U^\alpha} \;&=\; \frac{m}{2}\,\eta_{\mu\nu}\,\frac{\partial L}{\partial U^\alpha}(U^\mu U^\nu) \;+\; \frac{q}{c}\,\delta_\alpha{}^\beta A_\beta \\ &=\; \frac{m}{2}\,\eta_{\mu\nu}\,(\delta_\alpha{}^\mu U^\nu \;+\; U^\mu \delta_\alpha{}^\nu) \;+\; \frac{q}{c}\,A_\alpha \\ &=\; m\,U_\alpha \;+\; \frac{q}{c}\,A_\alpha \end{aligned} \tag{9.9}$$

With the definition for the 4-momentum of the particle, $\mathsf{P} = m\mathsf{U}$, we see that $c\pi_0 = mU_0 + qA_0 = \gamma mc^2 + qA_0$ is its total energy, and that $\boldsymbol{\pi} = \mathbf{p} + q\mathbf{A}/c$, justifying the often-used term "generalised momentum" .

Differentiating the conjugate momentum components with respect to proper time yields:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L}{\partial U^\alpha}\right) \;=\; \mathrm{d}_\tau P_\alpha \;+\; \frac{q}{c}\,\mathrm{d}_\tau x^\beta\,\partial_\beta A_\alpha \;=\; \mathrm{d}_\tau P_\alpha \;+\; \frac{q}{c}\,U^\beta\,\partial_\beta A_\alpha$$

where we have used the chain rule in the first equality. From this we reconstruct the Euler-Lagrange equations:

$$\mathrm{d}_\tau P_\alpha \;-\; \frac{q}{c}\,(\partial_\alpha A_\beta \;-\; \partial_\beta A_\alpha) \;=\; 0$$

which is indeed eq. (9.1).

Neither our Lagrangian nor Jackson's are *manifestly* gauge-invariant. Under a gauge transformation $\delta\mathbf{A} = \partial_\alpha f$, the Lagrangian changes by:

$$\delta L \;=\; \frac{q}{c}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\,\partial_\alpha f \;=\; \frac{q}{c}\frac{\mathrm{d}f}{\mathrm{d}\tau}$$

This extra term, however, does not change (EXERCISE) the Euler-Lagrange equations. Also, notice that eq. (9.1) depends on the potential through the magic combination $\partial_\alpha A_\beta - \partial_\beta A_\alpha$, which is gauge-independent.

9.4 Some Solutions of the Equation of Motion for a Charged Particle

We consider solutions of eq. (9.1), $K^\mu = \mathrm{d}_\tau P^\mu/m = (q/mc)F^{\mu\nu}U_\nu$, for various cases.

In section 8.7.3 the zeroth component of the Minkowski force on a particle was shown to be:

$$K^0 \;=\; m\frac{\mathrm{d}U^0}{\mathrm{d}\tau} \;=\; \frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}\tau} \;=\; \frac{\gamma}{c}\frac{\mathrm{d}E}{\mathrm{d}t}$$

where $E = \gamma mc^2$ is the energy of the particle. Now, using $U_i = -\gamma u^i$, we have in the electromagnetic: $K^0 = \gamma(q/c)F^{0i}u_i = \gamma(q/c)\mathbf{u}\cdot\mathbf{E}$, where $\mathbf{u}$ is the 3-velocity of the particle, and we are left with:

$$\frac{\mathrm{d}E}{\mathrm{d}t} \;=\; q\,\mathbf{u}\cdot\mathbf{E} \tag{9.10}$$

9.4.1 Uniform, static magnetic field (section J12.2)

Without any electric field, the energy of the particle does not change with time, and therefore neither does γ.

With the dot denoting differentiation with respect to observer time, the Lorentz force equation is:

$$\mathbf{F} \;=\; \gamma m\dot{\mathbf{u}} \;=\; \frac{q}{c}\,\mathbf{u}\times\mathbf{B}$$

or

$$\dot{\mathbf{u}} \;=\; \mathbf{u}\times\boldsymbol{\omega} \qquad\qquad \boldsymbol{\omega} \coloneqq \frac{q\,\mathbf{B}}{\gamma mc} \tag{9.11}$$

Writing this equation of motion as $\dot{u}_i = \epsilon_{ijk}u^j\omega^k$, we see that it couples the velocity components to each other. If we set the three coordinate axes so that the unit vector $\mathbf{e}_3$ is aligned with $\boldsymbol{\omega}$, ie., with the magnetic field, the u^3 component remains constantt, and the linear combinations $u^1 + \mathrm{i}u^2$ obeys a decoupled equation:

$$\mathrm{d}_t(u^1 + \mathrm{i}u^2) \;=\; -\,\mathrm{i}\,(u^1 + \mathrm{i}u^2)\,\omega_\mathrm{B} \qquad\qquad \boldsymbol{\omega} = \omega_\mathrm{B}\,\mathbf{e}_3$$

which has as solution:

$$u^1 + \mathrm{i}\,u^2 \;=\; (u^1_0 + \mathrm{i}\,u^2_0)\,\mathrm{e}^{-\mathrm{i}\omega_\mathrm{B}t}. \tag{9.12}$$

Define the velocity transverse to the direction of the magnetic field: $\mathbf{u}_\perp = u^1\mathbf{e}_1 + u^2\mathbf{e}_2$. A cursory inspection of the solution shows that $u_\perp$ is the constant length of a vector that rotates at angular velocity ω_B, as it remains tangent to a circle of radius $a = u_\perp/\omega_\mathrm{B}$, called the radius of gyration. Write $u^i_0 = \omega_\mathrm{B}ab^i$, with $(b^1)^2 + (b^2)^2 = 1$. Choosing $b^1 = 1$, $b^2 = 0$, we obtain: $\mathbf{u}_\perp = \Re\big[\omega_\mathrm{B}a(\mathbf{e}_1 - \mathrm{i}\,\mathbf{e}_2)\mathrm{e}^{-\mathrm{i}\omega_\mathrm{B}t}\big]$, which gives for the velocity:

$$\mathbf{u} \;=\; u^3\mathbf{e}_3 \;+\; \Re\big[\omega_\mathrm{B}a(\mathbf{e}_1 - \mathrm{i}\,\mathbf{e}_2)\mathrm{e}^{-\mathrm{i}\omega_\mathrm{B}t}\big] \tag{9.13}$$

and for the position:

$$\mathbf{x} \;=\; u^3t\,\mathbf{e}_3 \;+\; \Re\big[\mathrm{i}\,a(\mathbf{e}_1 - \mathrm{i}\,\mathbf{e}_2)\mathrm{e}^{-\mathrm{i}\omega_\mathrm{B}t}\big] \tag{9.14}$$

This is a helix of radius a and pitch $\tan^{-1}(u_3/u_\perp)$.

9.4.2 Combined uniform, static electric and magnetic fields (section J12.3)

In the lab frame, a particle is moving at velocity $\mathbf{u}$ in static, uniform electric and magnetic fields. The problem can be simplified by boosting to a frame in which one of the fields is zero, with eq. (8.52) given here for convenience:

$$\begin{aligned} \mathbf{E}'_{\parallel} &= \mathbf{E}_{\parallel}, & \mathbf{E}'_{\perp} &= \gamma\left[\mathbf{E}_{\perp} + \boldsymbol{\beta}\times\mathbf{B}\right] \\ \mathbf{B}'_{\parallel} &= \mathbf{B}_{\parallel}, & \mathbf{B}'_{\perp} &= \gamma\left[\mathbf{B}_{\perp} - \boldsymbol{\beta}\times\mathbf{E}\right] \end{aligned}$$

where $\boldsymbol{\beta}$ is the boost (not the particle's!) dimensionless velocity, There are two possibilities:

(1) $E < B$

Since $B^2 - E^2$ is Lorentz-invariant, we can transform $\mathbf{E}$ away: $\mathbf{E}' = 0$. From the top line of equations it is clear that $\boldsymbol{\beta}$ *must* be perpendicular to $\mathbf{E}$ (this is not a choice, contrary to what Jackson says). The transformations on the right only involve $\mathbf{B}_\perp$; the top one gives $\boldsymbol{\beta}\times\mathbf{E}_\perp = \beta^2\mathbf{B}_\perp$ which, when inserted in the bottom one, leads to a weaker field in the transformed frame: $\mathbf{B}'_\perp = \mathbf{B}_\perp/\gamma$ or, since $\beta = E_\perp/B_\perp$,

$$\mathbf{B}'_\perp = \sqrt{1 - E^2/B^2}\,\mathbf{B}_\perp$$

At this point we can choose $\mathbf{B}$ to be (1) perpendicular to $\boldsymbol{\beta}$, which remains true in the transformed frame, and (2) perpendicular to $\mathbf{E}$. Then we can write Jackson's expression (12.43):

$$\boldsymbol{\beta} = \frac{\mathbf{E}\times\mathbf{B}}{B^2}$$

In the transformed frame, then, the situation is identical to that discussed in the last section. In the observer frame, the orbit of the particle does not close, even when its velocity has no component along the magnetic field; instead, the orbit drifts in the boost direction, perpendicular to the plane of the crossed fields.

The crossed-fields arrangement also allows an interesting application, best understood from the Lorentz force equation (9.2): when $\mathbf{E} = -\boldsymbol{\beta}\times\mathbf{B}$, where $\boldsymbol{\beta}$ is the dimensionless velocity of the particle, the net force vanishes and there is no deflection. This occurs for a particle velocity $\boldsymbol{\beta} = (\mathbf{E}\times\mathbf{B})/B^2$, or $\beta = E/B$. Such a velocity selector can then be used to remove from a beam all particles with a velocity different from $\boldsymbol{\beta}$ given above, whatever their charge and mass.

(2) $E > B$

This time, we can transform to a frame where the magnetic field $\mathbf{B}' = 0$. The symmetry of the transformations (8.52), with a boost velocity $\boldsymbol{\beta} = (\mathbf{E}\times\mathbf{B})/E^2$, leads to a transformed electric field $\mathbf{E}' = \mathbf{E}/\gamma = \sqrt{1 - B^2/E^2}\mathbf{E}$.

If we align the x axis along the boost velocity and the y axis along the electric field, solving the equations of motion (9.1) yields (EXERCISE) the trajectory:

$$y' = \frac{\gamma_0 c}{\omega}\left(\cosh\frac{\omega x'}{\gamma_0\beta_0 c} - 1\right)$$

where β_0 is the initial speed of the particle assumed to be at the origin at $t = 0$, and $\omega \coloneqq qE/mc$.

9.4.3 Motion parallel to the magnetic flux lines — adiabatic invariants (section J12.5)

As our final example, we consider what happens to the component of the velocity parallel to the magnetic field lines, $\mathbf{u}_\parallel$, when those field lines are not parallel to each other.

In classical mechanics, we learn that a *periodic* coordinate q_i of a system gives rise to the action integral $J_i \coloneqq \oint \pi_i\,\mathrm{d}q_i$, where the integral is over a period of the motion, and π_i is the momentum conjugate to q_i. If the

system experiences changes in its physical parameters that are slow (adiabatic) on the scale set by the period, then J_i is a constant of the motion.

Now, in a uniform, static magnetic field $\mathbf{B}$, we do have periodic motion *transverse* to the field. Inserting the conjugate momenta found in eq. (9.9) leads to:

$$J \;=\; \oint \boldsymbol{\pi}_\perp \cdot \mathrm{d}\mathbf{l} \;=\; \oint \gamma m \mathbf{u}_\perp \cdot \mathrm{d}\mathbf{l} \;+\; \frac{q}{c} \oint \mathbf{A} \cdot \mathrm{d}\mathbf{l}$$

We choose the direction of integration to be counterclockwise as viewed in the direction of $\mathbf{B}$.

The first integral on the right is simply:

$$2\pi\gamma m \omega_{\mathrm{B}} a^2 \;=\; 2\frac{q}{c} B\, \pi a^2$$

with, as before, $u_\perp = \omega_{\mathrm{B}} a$ uniform. The second integral becomes the magnetic flux through the surface enclosed by the circular orbit, with a minus sign coming from the normal to the surface being opposite to $\mathbf{B}$, because of the chosen direction of integration. Combining the two terms gives:

$$J \;=\; \frac{q}{c} B\, \pi a^2 \;=\; \pi\gamma m \omega_{\mathrm{B}} a^2 \;=\; \frac{\pi c}{q} \frac{p_\perp^2}{B} \tag{9.15}$$

From these expressions we can say, equivalently, that the magnetic flux is invariant, or that $p_\perp^2/B$ is invariant, if $\mathbf{B}$ deviates significantly from uniformity only over distances large with respect to the size of the orbit.

The energy of the particle remains constant since $\mathbf{B}$ does no work on it. At low speeds ($u \ll c$), with velocity $\mathbf{u}_0 = \mathbf{u}_\perp + \mathbf{u}_\parallel$ at some initial position, this translates into: $u_0^2 = u_{0\perp}^2 + u_{0\parallel}^2 = u_\perp^2 + u_\parallel^2$. Also, γ remains constant, and the invariance of J can be written as: $u_\perp^2/B = u_{0\perp}^2/B_0$

Now take $\mathbf{B}$ to be along the z axis, with its field lines becoming ever more pinched as z increases. If this pinch effect is gradual enough over distance, invariance of the kinetic energy and of J can be combined, and there comes:

$$u_\parallel^2 \;=\; u_0^2 \;-\; u_{0\perp}^2 \frac{B(z)}{B_0} \tag{9.16}$$

As the parallel component of $\mathbf{u}$ decreases, its transverse component increases; translational energy is converted into rotational energy. From the point of view of flux conservation, as the field intemsity increases with z, the radius of the particle's orbits must decrease. Its trajectory is a spiral that tightens with distance along z. The particle may even stop and be reflected. As shown on pp. J594-595, the same result can be obtained by expanding the equations of motion to first order in the change in $\mathbf{B}$.

This effect has been extensively used in plasma confinement and to construct magnetic traps.

9.5 Electromagnetic Field Lagrangian and Maxwell's Equations (section J12.7)

Continuous fields can also be discussed in the context of the prnciple of least action, and Euler-Lagrange equations derived for them. The action is still defined as $S = \int L\, \mathrm{d}t$, but since each degree of freedom now depends on position as well as on time, the basic object is a function, the **Lagrangian density** $\mathcal{L}$, not only of the field components ϕ^i, but of the components $\partial_\mu \phi^i$ of their 4-gradient, which gives the Lagrangian:

$$L \;=\; \sum_i \int \mathcal{L}(\phi^i, \partial_\mu \phi^i)\, \mathrm{d}^3 x$$

A derivation completely analogous to that in section 9.2 yields the Euler-Lagrange equations for each field component:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i)} \right) \;-\; \frac{\partial \mathcal{L}}{\partial \phi^i} \;=\; 0 \tag{9.17}$$

We require that $\mathcal{L}$ be a Lorentz scalar invariant under spatial inversions, and we expect that it should contain terms quadratic in the velocities, ie., in the derivatives of the 4-vector potential $\mathbf{A}$. The only available object that satisfies these criteria is $F^{\mu\nu}F_{\mu\nu}$, which is entirely constructed from velocities. A term proportional to $A_\mu A^\mu$ is also allowed, but since it corresponds to the photon being massive, we will leave it out in what follows — see Jackson's section 12.8 for a short discussion.

A coupling, or interaction, term containing the current density is also needed, and something proportional to $J^\mu A_\mu$ fits the purpose. Notice, however, that whereas the field term is manifestly gauge-invariant, the coupling term is not. Indeed, under a transformation $\delta\mathbf{A} = \boldsymbol{\nabla} f$, $\delta(J^\mu A_\mu) = J^\mu \partial_\mu f$. Yet integration by parts of the gauge-induced term gives $-(\partial_\mu J^\mu) f$, plus a total divergence which does not contribute when integrated. Then we can think of f as a Lagrange multiplier that enforces the conservation law $\partial_\mu J^\mu = 0$. Or we can decide to live with an action that is not gauge-invariant, so long as gauge invariance holds "on-shell", ie., at the level of the field equations resulting from minimising the action. Our Lagrangian density is:

$$\mathcal{L} \;=\; -\frac{1}{4}\, F^{\alpha\beta} F_{\alpha\beta} \;-\; \frac{4\pi}{c}\, J^\mu A_\mu \tag{9.18}$$

In the so-called **second-order formalism**, the field term is rewritten (EXERCISE) as: $\frac{1}{2}(\eta^{\alpha\lambda}\eta^{\beta\rho} - \eta^{\alpha\rho}\eta^{\beta\lambda})(\partial_\rho A_\lambda)(\partial_\alpha A_\beta)$. Each term is symmetric under interchange of a pair of $\alpha\beta$ indices with a a $\rho\lambda$ pair. Then we can immediately write:

$$\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} \;=\; (\eta^{\alpha\lambda}\eta^{\beta\rho} - \eta^{\alpha\rho}\eta^{\beta\lambda})\, \delta^\nu{}_\alpha\, \delta^\mu{}_\beta\, \partial_\rho A_\lambda \;=\; \partial^\mu A^\nu \;-\; \partial^\nu A^\mu \;=\; F^{\mu\nu} \tag{9.19}$$

The resulting field equations are:

$$\partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} \right) \;-\; \frac{\partial \mathcal{L}}{\partial A_\mu} \;=\; \partial_\nu F^{\mu\nu} \;+\; \frac{4\pi}{c}\, J^\mu \;=\; 0 \tag{9.20}$$

which is identical to eq. (8.36), as expected. Not only is the conservation law $\partial_\nu J^\nu = 0$ implicit as an identity on the field equations, but the equations are manifestly gauge-invariant.

The generalised-momentum components conjugate to A_μ are found simply by setting $\nu = 0$ in eq. (9.19): $\pi_\mu = F^{\mu 0} = F_{0\mu}$. Thus, $\pi_0 = 0$ and $\boldsymbol{\pi} = \mathbf{E} = -\dot{\mathbf{A}} - \boldsymbol{\nabla} A_0$, the electric field. Whereas the transformation from $\dot{\mathbf{A}}$ to π is invertible, that from $\dot{A}_0$ to π_0 is not. There are only three non-zero momenta for the six field components, so at most three degrees of freedom can propagate, and A_0 will never be one of them, even when a mass term for the photon is added to $\mathcal{L}$. Because $\pi_0 = 0$, the determinant of the Hessian matrix of $\mathcal{L}$ also vanishes, and there is some arbitrariness in the time evolution. On the other hand, the vanishing of the Hessian yields a *constraint* equation on the π_i which leaves us with the expected two propagating degrees of freedom.

To understand the role of A_0, it is instructive to express the field term in $\mathcal{L}$ in 3+1 form:

$$\mathcal{L}_{\text{field}} \;=\; \frac{1}{2}\,(E^2 \;-\; B^2) \;=\; \frac{1}{2}\left[\pi_i^2 \;-\; \mathbf{B}\cdot\boldsymbol{\nabla}^{(3)}\times\mathbf{A}\right] \;=\; \frac{1}{2}\left[\pi_i^2 \;-\; \mathbf{A}\cdot\boldsymbol{\nabla}^{(3)}\times\mathbf{B}\right] \tag{9.21}$$

The last equality is up to a total divergence, which we ignore as usual. While Ampère's law is easily retrieved as an Euler-Lagrange equation, we may well wonder what happened to Gauss's law. If it is does not turn up as a field equation, what is it? To find out, we rewrite $\mathcal{L}_{\text{field}}$ in yet another equivalent form, suggested by Schwinger:

$$\mathcal{L}_{\text{field}} \;=\; -\frac{1}{2}\left[F^{\mu\nu}\,(\partial_\mu A_\nu \;-\; \partial_\nu A_\mu) \;-\; \frac{1}{2}\, F^{\mu\nu} F_{\mu\nu}\right]$$

where $\mathbf{F}$ and $\mathbf{A}$ are taken to be *independent*, that is, we do not assume that $\mathbf{F} = \boldsymbol{\nabla} \wedge \mathbf{A}$; but the components of $\mathbf{F}$ are still written in terms of $\mathbf{E}$ and $\mathbf{B}$ as in eq. (8.39). Since only first-order derivatives are explicitly present, this is known as the *first-order* formalism.

Already we can see that it is the variation with respect to $\mathbf{F}$ that will yield $\mathbf{F} = \boldsymbol{\nabla} \wedge \mathbf{A}$. Specifically:

$$\begin{aligned}
\mathcal{L}_{\text{field}} \;&=\; -\,F^{0i}\,(\partial_0 A_i \;-\; \partial_i\, A_0) \;-\; F^{ij}\,(\partial_i A_j \;-\; \partial_j\, A_i) \;-\; \frac{1}{2}(E^2 \;-\; B^2) \qquad i<j \\
&=\; \mathbf{E}\cdot(-\boldsymbol{\nabla} A_0 \;-\; \dot{\mathbf{A}}) \;-\; \mathbf{B}\cdot\boldsymbol{\nabla}\times\mathbf{A} - \frac{1}{2}(E^2 \;-\; B^2)
\end{aligned}$$

Varying with respect to $\mathbf{E}$ and $\mathbf{B}$ does yield $\mathbf{E} = -\boldsymbol{\nabla} A_0 \, - \, \dot{\mathbf{A}}$ and $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$.

Before varying with respect to $\mathbf{A}$, we note that, up to total divergences, $\mathcal{L}_{\text{field}}$ is also:

$$\mathcal{L}_{\text{field}} \; = \; A_0 \, \boldsymbol{\nabla} \cdot \mathbf{E} \; - \; \mathbf{E} \cdot \dot{\mathbf{A}} \; - \; \mathbf{A} \cdot \boldsymbol{\nabla} \times \mathbf{B} - \; \frac{1}{2}(E^2 \; - \; B^2)$$

We recover Ampère's law by varying with respect to $\mathbf{A}$, but now we can see that A_0 acts as a Lagrange multiplier enforcing Gauss's law, $\boldsymbol{\nabla} \cdot \mathbf{E} = \boldsymbol{\nabla} \cdot \boldsymbol{\pi} = 0$, which is revealed to be the constraint on the π_i momenta that we knew must exist.

9.6 Energy-Momentum in Electromagnetism (section J12.10)

We are now ready to discuss energy in electromagnetism making use of the 4-vector formalism; this will lead to a truly unified treatment and will simplify some derivations. Consider the manifestly symmetric rank-2 tensor T with contravariant components:

$$T^{\mu\nu} \; = \; \frac{1}{4\pi} \left(F^{\mu}{}_{\rho} F^{\rho\nu} \; + \; \frac{1}{4} \eta^{\mu\nu} F^{\alpha\beta} F_{\alpha\beta} \right) \tag{9.22}$$

9.6.1 Explicit components of the electromagnetic energy-momentum tensor

Calculate T in terms of the electric and magnetic components of the Faraday tensor F.

The invariant $F^{\alpha\beta} F_{\alpha\beta}$ is already known to be $-2(E^2 - B^2)$. Therefore:

$$\begin{aligned} T^{00} \; &= \; \frac{1}{4\pi} \left[F^0{}_i F^{i0} \; - \; \frac{1}{2} \eta^{00} (E^2 \; - \; B^2) \right] \; = \; \frac{1}{4\pi} \left[E^2 \; - \; \frac{1}{2}(E^2 \; - \; B^2) \right] \\ &= \; \frac{1}{8\pi} \left(E^2 + B^2 \right) \end{aligned} \tag{9.23}$$

$$\text{and} \qquad T^{i0} \; = \; \frac{1}{4\pi} \, F^i{}_j \, F^{j0} \; = \; - \frac{1}{4\pi} \, E^j \epsilon^{ijk} B_k \; = \; \frac{1}{4\pi} \, (\mathbf{E} \times \mathbf{B})^i$$

Finally (see the note at the bottom of p. 114) in section 8.10):

$$\begin{aligned} 4\pi \, T^{ij} \; &= \; F^i{}_0 F^{0j} \; + \; F^i{}_k F^{kj} \; - \; \frac{1}{2} \eta^{ij} (E^2 \; - \; B^2) \\ &= \; - \, E^i E^j \; + \; \epsilon_{ikl} \epsilon^{jkm} B^l B_m \; + \; \frac{1}{2} \delta^{ij} (E^2 \; - \; B^2) \\ &= \; - \, E^i E^j \; - \; (\delta_i{}^j \delta_l{}^m \; - \; \delta_i{}^m \delta_l{}^j) B^l B_m \; + \; \frac{1}{2} \delta^{ij} (E^2 \; - \; B^2) \\ &= \; - \left(E^i E^j \; - \; \frac{1}{2} \, \delta^{ij} E^2 \right) \; - \left(B^i B^j \; - \; \frac{1}{2} \, \delta^{ij} B^2 \right) \end{aligned} \tag{9.24}$$

9.6.2 Energy-momentum tensor and conservation laws

It is not at all obvious at this stage why T might be interesting. Its form is derived from Lagrangian and Hamiltonian field theory (see Jackson for a detailed discussion). Here, we shall discover its physical interpretation by investigating one of its properties: we shall calculate its divergence, $\boldsymbol{\nabla} \cdot \mathsf{T}$:

$$4\pi \, \partial_\mu T^{\mu\nu} \; = \; F^{\rho\nu} \, \partial_\mu F^{\mu}{}_{\rho} + F^{\mu}{}_{\rho} \partial_\mu F^{\rho\nu} + \frac{1}{2} \, F^{\alpha\beta} \partial^\nu F_{\alpha\beta}$$

We now proceed to show that the last two terms cancel each other. Subtract zero in the form of a term containing $\boldsymbol{\nabla} \wedge \mathsf{F}$, thus:

$$F_{\mu\rho} \partial^\mu F^{\rho\nu} + \cancel{\frac{1}{2} F^{\alpha\beta} \partial^\nu F_{\alpha\beta}} \; - \; \frac{1}{2} F_{\alpha\beta} \, (\cancel{\partial^\nu F^{\alpha\beta}} + \partial^\beta F^{\nu\alpha} \; + \; \partial^\alpha F^{\beta\nu})$$

The -last two terms.are equal, and their sum cancels the first term. Therefore:

$$\partial_\mu T^{\mu\nu} \;=\; \frac{1}{4\pi} F^{\rho\nu}\partial^\mu F_{\mu\rho} \;=\; -\,\frac{1}{c}\,F^{\nu\rho}\,J_\rho \tag{9.25}$$

where we have used the inhomogeneous first-order equation for F.

We recognise $F^{\nu\rho}J_\rho/c$ as the components of the force density $\mathbf{f} = n\mathbf{K}$ introduced at the end of section 9.1:

$$\begin{aligned} f^0 \;&=\; \frac{1}{c}\,F^{0i}J_i \;=\; -\,\frac{1}{c}E^i J_i \;=\; \frac{1}{c}\mathbf{E}\cdot\mathbf{J} \\ f^i \;&=\; \frac{1}{c}\,F^{i0}J_0 \,+\, \frac{1}{c}\,F^{ik}J_k = E^i\rho \,+\, \frac{1}{c}\,(\epsilon^{ikj}B_j)\,J_k \end{aligned} \tag{9.26}$$

Indeed, $\mathbf{f} \;=\; \rho\mathbf{E} + \mathbf{J}\times\mathbf{B}/c$ is the electromagnetic force *density* acting on the charges, ie. the rate of change of mechanical momentum per unit volume (momentum density), $\mathrm{d}_t\mathbf{p}_{\text{mech}}$, of these particles. Moreover, $\mathbf{u}\cdot\mathbf{f} = \rho\mathbf{u}\cdot\mathbf{E} = \mathbf{J}\cdot\mathbf{E} = cf^0$,

Now, examine more closely $f^\mu + \partial_\nu T^{\mu\nu} \;=\; 0$ by splitting it into its component equations. First, $c\,f^0 = -\partial_t T^{00} - c\partial_i T^{0i}$. Define $U_{\text{field}} \equiv T^{00}$ and take $c\,T^{0i}$ to be the components of a 3-vector $\mathbf{S}$. Then, since $c\,f^0 = \mathbf{u}\cdot\mathbf{f}$, we have:

$$\mathbf{u}\cdot\mathbf{f} \;=\; -\,\partial_t U_{\text{field}} \,-\, \boldsymbol{\nabla}\cdot\mathbf{S}$$

Integrate this over the volume of the source and use the divergence theorem to obtain:

$$\frac{\mathrm{d}}{\mathrm{d}t}W_{\text{mech}} \;=\; -\,\frac{\mathrm{d}}{\mathrm{d}t}W_{\text{field}} - \oint\limits_{\text{surface}} \mathbf{S}\cdot\mathrm{d}\mathbf{a} \tag{9.27}$$

where we have defined $W_{\text{field}} \equiv \int U_{\text{field}}\,\mathrm{d}^3x$, and $\mathrm{d}W_{\text{mech}}/\mathrm{d}t \equiv \int \mathbf{u}\cdot\mathbf{f}\,\mathrm{d}^3x = \int \mathbf{J}\cdot\mathbf{E}\,\mathrm{d}^3x$ is the rate of change of the mechanical (kinetic, potential) energy of the whole charge distribution. Then:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(W_{\text{field}} + W_{\text{mech}}\right) \;=\; -\oint\limits_{\text{surface}} \mathbf{S}\cdot\mathrm{d}\mathbf{a} \tag{9.28}$$

Obviously, this is a statement of energy conservation which we recognise from eq. (5.31) as Poynting's theorem. U_{field} represents the energy density stored in the electromagnetic field, and $\mathbf{S}$, the Poynting vector, is the flux of this energy per unit time per unit area, ie. the **electromagnetic energy flux density**.

Turn now to the spatial components of $f^\mu + \partial_\nu T^{\mu\nu} \;=\; 0$. The force per unit volume acting on the charge distribution (or mechanical momentum density of the distribution) has components:

$$f^i \;=\; -\,\partial_0 T^{i0} \,-\, \partial_j T^{ij} \;=\; -\,\frac{1}{c^2}\,\partial_t S^i \,-\, \partial_j T^{ij}$$

Integrate over the volume containing the charges, use the divergence theorem, and rearrange:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{P}_{\text{mech}} \,+\, \mathbf{P}_{\text{field}}\right) \;=\; -\oint\limits_{\text{surface}} T^{ij}n_i\,\mathrm{d}a \tag{9.29}$$

where $\mathbf{P}_{\text{mech}}$ is the total mechanical momentum of the charges, and $\mathbf{P}_{\text{field}} = \int \mathbf{p}_{\text{field}}\,\mathrm{d}^3x$, with $\mathbf{p}_{\text{field}} = \mathbf{S}/c^2$). The unit vector $\hat{\mathbf{n}}$ in the surface integral is perpendicular to the surface.

This is identical to the law of conservation of momentum, eq. (5.33), obtained in section 5.7.2, with T^{ij} the components of the Maxwell stress tensor.

Let us review what has been achieved. We have introduced an object, the **energy-momentum tensor**, $\mathbf{T}_{\text{field}}$, which contains all the information about the energy and momentum of the fields:

- T^{00} is the **energy density** stored in the fields.
- cT^{0i} are the components of the **energy flux density**, or Poynting vector $\mathbf{S}$; T^{0i}/c are the components of **momentum density**, $\mathbf{p}_{\text{field}}$.
- T^{ij} corrrespond to **momentum flux density**, and $-T^{ij}$ the **stress** (pressure or shear) exerted by the fields.

There is even more! $\mathbf{f} = -\boldsymbol{\nabla} \cdot \mathsf{T}_{\text{field}}$ controls how the field exchanges energy and momentum with matter. In fact, *starting* with $\mathsf{T}_{\text{field}}$, we can *derive* the force acting on matter. We can even think of $\mathbf{f}$ as the divergence of a mechanical energy-momentum tensor, T_{mech}, for matter itself. Then the *total* energy-momentum tensor, $\mathsf{T} = \mathsf{T}_{\text{mech}} + \mathsf{T}_{\text{field}}$, obeys $\boldsymbol{\nabla} \cdot \mathsf{T} = 0$, which is the conservation law in differential form for the total energy-momentum of the system, including matter *and* fields.

Another important point is that this discussion *applies equally well to other fields that interact with matter*. All that is required is to know the relevant $\mathsf{T}_{\text{field}}$. Its components will have exactly the same physical interpretation, and $\boldsymbol{\nabla} \cdot \mathsf{T} = 0$ will yield the force exerted by that field on matter or, equivalently, conservation laws for the transfer of energy and momentum between matter and fields.

10 Electromagnetic Fields of, and Radiation from, Relativistic Point-charges

10.1 Potentials of a Point-Charge in Arbitrary Motion (section J14.1)

In section 8.11, we made clever use of the Lorentz transformations of the electromagnetic fields to find the fields of a point-charge in uniform motion. Here we present a brief outline of a derivation of the electromagnetic potential $\mathbf{A}$ and Faraday field tensor $\mathbf{F}$ for a point-charge q in arbitrary motion, whose position in an inertial frame S is given by $\mathbf{r}(t)$, and whose instantaneous 3-velocity in S is $\mathbf{v}(t) = \mathrm{d}_t\mathbf{r}$. The charge's position 4-vector $\mathbf{x}$ in S thus has components $\big(ct, \mathbf{r}(t)\big)$, and its 4-velocity $\mathbf{U}$ has components $U^\mu = \gamma(c, \mathbf{v})$.

10.1.1 Lorentz form-invariant expression for the point-source

The charge density ρ and 3-current density $\mathbf{j}$ associated with the point charge are easily expressed as:

$$\rho(\mathbf{x},t) \;=\; q\,\delta^{(3)}\big(\mathbf{x}-\mathbf{r}(t)\big), \qquad\qquad \mathbf{j}(\mathbf{x},t) \;=\; q\mathbf{v}(t)\,\delta^{(3)}\big(\mathbf{x}-\mathbf{r}(t)\big)$$

These, however, do not transform properly under Lorentz transformations. We know that they should be written as the components of the 4-current density $\mathbf{J}$ written in terms of manifestly form-invariant objects. Jackson (eq. J12.139) gives the following expression:

$$J^\mu(\mathbf{x}) \;=\; q\,c\int \mathrm{d}\tau\, U^\mu(\tau)\,\delta^{(4)}\big(\mathbf{x}-\mathbf{r}(\tau)\big) \tag{10.1}$$

where $\delta^{(4)}\big(\mathbf{x}-\mathbf{r}(\tau)\big) = \delta(ct-c\tau)\delta^{(3)}\big(\mathbf{x}-\mathbf{r}(\tau)\big) = (1/c)\delta(t-\tau)\delta^{(3)}\big(\mathbf{x}-\mathbf{r}(\tau)\big)$.

10.1.2 Lorentz form-invariant expressions for the Green functions of the d'Alembertian (section J12.11)

In section 5.4 we obtained the general Green functions (without specified boundary or initial conditions) for the d'Alembertian operator involved in an equation of the type: $\Box\Psi(\mathbf{x},t) \;=\; 4\pi\, f(\mathbf{x},t)$. These Green functions, which satisfy the defining equation: $\Box_x G(\mathbf{x},t;\mathbf{x}',t') \;=\; \delta(\mathbf{x}-\mathbf{x}')\,\delta(t-t')$, are:

$$G^{(\pm)}(\mathbf{x},t;\mathbf{x}',t') \;=\; \frac{1}{4\pi R}\,\delta\big(t' \;-\; [t \mp R/c]\big)$$

where $R = |\mathbf{x}-\mathbf{x}'|$, and $+/-$ correspond to retarded/advanced functions. In the 4-vector formalism, this becomes:

$$G^{(\pm)}(\mathbf{x}-\mathbf{x}') \;=\; \frac{1}{4\pi R}\,\theta\big(\pm(x_0-x_0')\big)\,\delta(x_0 \;-\; x_0' \pm R) \tag{10.2}$$

where the step-function θ ensures that source time is earlier than observation time in the retarded case, and later in the advanced case.

Our Green functions are not *manifestly* form-invariant under Lorentz transformations.. To make the form-invariance manifest, note that $\mathbf{x}\cdot\mathbf{x}$ *is* invariant, and write:

$$\delta(\Delta\mathbf{x}\cdot\Delta\mathbf{x}) \;=\; \delta\big[(\Delta x_0)^2 \;-\; R^2\big] \;=\; \delta\big[(\Delta x_0 - R)(\Delta x_0 + R)\big]$$

Now we use the following identity (not given in the preamble) for the δ-function:

$$\int \mathrm{d}y\, g(y)\,\delta[f(y)] \;=\; \int \mathrm{d}f\,\frac{g(y)}{\mathrm{d}f/\mathrm{d}y}\,\delta(f) \;=\; \sum_i \left|\frac{g(y)}{\mathrm{d}f/\mathrm{d}y}\right|_{f(y_i)}$$

where $f(y_i) = 0$. Without the integrals, this is equivalent to:

$$g(y)\,\delta[f(y)] \;=\; \sum_i \left|\frac{g(y)}{\mathrm{d}f/\mathrm{d}y}\right|_{y_i} \delta(y-y_i) \tag{10.3}$$

Now take $g(y) = 1$, $y = \Delta x_0$, and $f(y) = (\Delta x_0 - R)(\Delta x_0 + R)$, which vanishes at $\Delta x_0 = \pm R$. Then:

$$\delta(\Delta\mathbf{x}\cdot\Delta\mathbf{x}) \;=\; \frac{1}{2R}\left[\delta(\Delta x_0 - R) \,+\, \delta(\Delta x_0 + R)\right] \tag{10.4}$$

leading to:

$$G^{(\pm)}(\mathbf{x}-\mathbf{x}') \;=\; \frac{1}{2\pi}\,\theta\big(\pm(x_0 - x_0')\big)\,\delta(\Delta\mathbf{x}\cdot\Delta\mathbf{x}) \tag{10.5}$$

It should be clear that the retarded Green functions must live on the light-come: $\Delta\mathbf{x}\cdot\Delta\mathbf{x} = 0$, with the retarded one restricted to the forward light-cone with respect to the source point. The step-function is now seen to be needed so as to distinguish between $x_0 > x_0'$ and $x_0 < x_0'$, which the invariant form of the δ function no longer does. Also, the ste-function is invariant when restricted by the δ function.

10.1.3 Lorentz form-invariant expression for the 4-potential (section J14.1)

The form-invariant inhomogeneous retarded solution of the wave equation for the Cartesian components of $\mathbf{A}$ takes the form:

$$\begin{aligned} A_\mu(\mathbf{x}) \;&=\; \frac{4\pi}{c}\int \mathrm{d}^4x'\, G^+(\mathbf{x}-\mathbf{x}')\, J_\mu(\mathbf{x}') \\ &=\; 2q\int \mathrm{d}\tau \int \mathrm{d}^4x'\,\delta(\Delta\mathbf{x}\cdot\Delta\mathbf{x})\,\theta(x_0 - x_0')\,\delta^{(4)}\big(\mathbf{x}' - \mathbf{r}(\tau)\big)\,U_\mu(\tau) \\ &=\; 2q\int \mathrm{d}\tau\,\delta(\Delta\mathbf{x}\cdot\Delta\mathbf{x})\,\theta\big(x^0 - r^0(\tau)\big)\,U_\mu(\tau) \end{aligned} \tag{10.6}$$

where now, because of the 4-dim δ-function, $\Delta\mathbf{x} = \mathbf{x} - \mathbf{r}(\tau)$. Again, as expected, the only contribution to the integral comes from the backward light-cone of $\mathbf{x}$, at $\tau = \tau_0$ defined by $\big[\mathbf{x}-\mathbf{r}(\tau_0)\big]^2 = \big(\mathbf{x}-\mathbf{r}(\tau_0)\big)\cdot\big(\mathbf{x}-\mathbf{r}(\tau_0)\big) = 0$, with $x^0 > r^0(\tau_0)$.

In the integral form of identity (10.3), put $y = \tau$, $f(y) = \Delta\mathbf{x}\cdot\Delta\mathbf{x}$, and $g(y) = U_\mu(\tau)$. Then, since:

$$\mathrm{d}_\tau\big[(\mathbf{x}-\mathbf{r}(\tau))\cdot(\mathbf{x}-\mathbf{r}(\tau))\big] \;=\; 2\big(x^\mu - r^\mu(\tau)\big)\,\mathrm{d}_\tau\big(x_\mu - r_\mu(\tau)\big) \;=\; -2\big(x_\mu - r_\mu(\tau)\big)\,\mathrm{d}_\tau r^\mu(\tau) \;=\; -2\big(x^\mu - r^\mu(\tau)\big)\,U_\mu(\tau)$$

eq. (10.6) becomes:

$$A_\mu(\mathbf{x}) \;=\; \left.\frac{q\,U_\mu(\tau)}{\mathbf{U}\cdot\big(\mathbf{x}-\mathbf{r}(\tau)\big)}\right|_{\tau_0} \tag{10.7}$$

which are called the manifestly form-invariant **Liénard-Wiechert potentials** of a point-charge in arbitrary motion. with the light-cone condition $\big[\mathbf{x} - \mathbf{r}(\tau_0)\big]^2 = 0$, or $\big(x^0 - r^0(\tau_0)\big)^2 = c^2(t-\tau_0)^2 = R^2$ and $\big(x^0 - r^0(\tau_0)\big)^2 = \big(\mathbf{x} - \mathbf{r}(\tau_0)\big)^2 = \mathbf{R}\cdot\mathbf{R}$, the denominator can be rewritten (EXERCISE) as $\mathbf{U}\cdot\mathbf{R} = \gamma cR(1 - \boldsymbol{\beta}\cdot\hat{\mathbf{n}})$, where $\hat{\mathbf{n}}$ is in the direction from the source point at τ_0 to the observation point *at retarded time*, and $R^\mu = (R, \mathbf{R}) = (R, R\hat{\mathbf{n}})$. With $U_\mu = \gamma(c, -\mathbf{v})$, the usual expressions for the Liénard-Wiechert potentials are recovered:

$$\Phi(\mathbf{x},t) \;=\; \frac{q}{R(1-\boldsymbol{\beta}\cdot\hat{\mathbf{n}})}, \qquad\qquad \mathbf{A}(\mathbf{x},t) \;=\; \frac{q\,\boldsymbol{\beta}}{R(1-\boldsymbol{\beta}\cdot\hat{\mathbf{n}})} \tag{10.8}$$

where it is understood that the right-hand sides must be evaluated at retarded time.

Without doing the detailed calculation, it is not hard to see that when the vector potential is differentiated with respect to time, the leading term will go like $\dot{\boldsymbol{\beta}}/R$, so that the radiation fields is expected to depend on the acceleration of the charge. If the charges does not accelerate, the leading dependence will go like $1/R^2$ or $\boldsymbol{\beta}/R^2$, as we discovered in section 8.11, and there can be no radiation from the charge.

10.2 Lorentz form-invariant Expressions for the Fields of a Point-Charge(section J14.1)

To derive the form-invariant fields, we find it convenient to differentiate eq. (10.6)—instead of eq. (10.7)—with respect to the coordinates $\mathbf{x}$ of the observation point:

$$\partial_\mu A_\nu \;=\; 2q \int \mathrm{d}\tau\, U_\nu(\tau)\, \theta\big(x^0 - r^0(\tau)\big)\, \partial_\mu \delta\big(\Delta\mathbf{x}\cdot\Delta\mathbf{x}\big)$$

Another term with $\partial_0 \theta\big(x^0 - r^0(\tau)\big) = \delta\big(x^0 - r^0(\tau)\big)$ vanishes since observation time cannot be the same as source time.

Writing $f = [\mathbf{x} - \mathbf{r}(\tau)]\cdot[\mathbf{x} - \mathbf{r}(\tau)]$ to minimise clutter, the chain rule gives $\partial_\mu \delta(f) = (\mathrm{d}_\tau f)^{-1} \mathrm{d}_\tau \delta(f) \partial_\mu f$. In the last section we already calculated $\mathrm{d}_\tau f$ to be $-2(\mathbf{x}-\mathbf{r})\cdot\mathbf{U}$. Also, $\partial_\mu f = 2(x^\alpha - r^\alpha)\delta^\mu{}_\alpha$, and there comes:

$$\partial_\mu \delta(f) \;=\; -\,\frac{x_\mu - r_\mu}{(\mathbf{x}-\mathbf{r})\cdot\mathbf{U}}\, \mathrm{d}_\tau \delta(f)$$

Now insert this into our expression for $\partial_\mu A_\nu$ and integrate by parts, noting once again that the τ derivative of the step-function vanishes:

$$\partial_\mu A_\nu \;=\; 2q \int \mathrm{d}\tau\, \theta\big(x^0 - r^0(\tau)\big)\, \delta\big(\Delta\mathbf{x}\cdot\Delta\mathbf{x}\big)\, \mathrm{d}_\tau \left[\frac{U_\nu(x_\mu - r_\mu)}{\mathbf{U}\cdot\big(\mathbf{x}-\mathbf{r}(\tau)\big)}\right]$$

Comparing with eq. (10.6) reveals that the two expressions have exactly the same form, with the potential replaced by its derivative, and U_μ by the τ derivative. Accordingly, we can make these same substitutions in the form-invariant expression (10.7), and we arrive at an intermediate result for the components of the Faraday tensor:

$$F_{\mu\nu} \;=\; \partial_\mu A_\nu \;-\; \partial_\nu A_\mu \;=\; \frac{q}{\mathbf{U}\cdot\big(\mathbf{x}-\mathbf{r}(\tau)\big)}\, \mathrm{d}_\tau \left[\frac{(\mathbf{x}-\mathbf{r})_\mu U_\nu \;-\; (\mathbf{x}-\mathbf{r})_\nu U_\mu}{\mathbf{U}\cdot\big(\mathbf{x}-\mathbf{r}(\tau)\big)}\right]_{\tau_0} \tag{10.9}$$

Incidentally, as a by-product of the above expression for $\partial_\mu A_\nu$, we immediately confirm that $\mathbf{A}$ satisfies the Lorenz condition $\partial_\mu A^\mu = 0$, as we know it must from arguments in section 5.5.

There remains to evaluate the τ derivative, for which we now use the dot notation, eg., $\mathrm{d}_\tau \mathbf{U} \equiv \dot{\mathbf{U}}$. We have: $\mathrm{d}_\tau(\mathbf{x}-\mathbf{r})_\mu = -U_\mu$ and $\mathrm{d}_\tau\big[\mathbf{U}\cdot(\mathbf{x}-\mathbf{r})\big] = \dot{\mathbf{U}}\cdot(\mathbf{x}-\mathbf{r}) - \mathbf{U}\cdot\mathbf{U} = \dot{\mathbf{U}}\cdot(\mathbf{x}-\mathbf{r}) - c^2$. Then:

$$F_{\mu\nu} \;=\; \frac{q}{\mathbf{U}\cdot(\mathbf{x}-\mathbf{r})} \left[\frac{\cancel{-U_\mu U_\nu} + (\mathbf{x}-\mathbf{r})_\mu \dot{U}_\nu + \cancel{U_\nu U_\mu} - (\mathbf{x}-\mathbf{r})_\nu \dot{U}_\mu}{\mathbf{U}\cdot(\mathbf{x}-\mathbf{r})} \;-\; \frac{[(\mathbf{x}-\mathbf{r})_\mu U_\nu \;-\; (\mathbf{x}-\mathbf{r})_\nu U_\mu][\dot{\mathbf{U}}\cdot\big(\mathbf{x}-\mathbf{r}\big) - c^2]}{[\mathbf{U}\cdot\big(\mathbf{x}-\mathbf{r}\big)]^2}\right]_{\tau_0}$$

As we did for the Liénard-Wiechert potential, we write $\mathbf{x}-\mathbf{r} = \mathbf{R}$ with contravariant components $(R, \mathbf{R})$ so that the components of the Faraday tensor may be expressed in terms of the distance 4-vector between the retarded spacetime position of the charge and the observation point:

$$F_{\mu\nu} \;=\; qc^2 \left[\frac{R_\mu U_\nu \;-\; R_\nu U_\mu}{(\mathbf{U}\cdot\mathbf{R})^3}\right]_{\tau_0} \;+\; q \left[\frac{R_\mu \dot{U}_\nu - R_\nu \dot{U}_\mu}{(\mathbf{U}\cdot\mathbf{R})^2} \;-\; \frac{(R_\mu U_\nu \;-\; R_\nu U_\mu)\mathbf{R}\cdot\dot{\mathbf{U}}}{(\mathbf{U}\cdot\mathbf{R})^3}\right]_{\tau_0} \tag{10.10}$$

The first term has a $1/R^2$ dependence, whereas the second one goes like $1/R$. The first is velocity-dependent, the second acceleration-dependent. We have obtained a (form-invariant!) decomposition of the field-tensor components into near-fields and far-fields. As claimed before, the latter—the radiation fields—exist only when the charge is accelerating. The near-fields have no acceleration dependence.

From our form-invariant equation for the Faraday components, we can calculate the electric and magnetic fields.

Example 10.1. Let us derive the electric near-field components for the point-charge:

$$E^i_{\text{near}} \;=\; F^{\text{near}}_{0i} \;=\; qc^2\,\left[\frac{R_0U_i \;-\; R_iU_0}{(\mathbf{U}\cdot\mathbf{R})^3}\right]_{\tau_0}$$

where $R_0 = R$, $R_i = -R^i$, $U_0 = \gamma c$, and $U_i = -\gamma\beta^i c$. Also, $\mathbf{U}\cdot\mathbf{R}$ was given in section 10.1.3 as: $\gamma cR(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})$, with $\boldsymbol{\beta}$ the instantaneous velocity of the charge. Putting all this together yields:

$$\mathbf{E}_{\text{near}} \;=\; \frac{q}{\gamma^2R^2}\,\frac{\hat{\mathbf{n}}\;-\;\boldsymbol{\beta}}{(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^3}\bigg|_{\tau_0} \qquad\qquad \hat{\mathbf{n}} \;=\; \frac{\hat{\mathbf{R}}}{R} \tag{10.11}$$

When $\boldsymbol{\beta}$ is constant, this is the whole electric field and should be identical to the result we derived (with much less work!) in section 8.11 via a Lorentz transformation:

$$\mathbf{E} \;=\; q\,\frac{1}{\gamma^2}\,\frac{1}{(1-\beta^2\sin^2\psi)^{3/2}}\,\frac{1}{r^2}\,\hat{\mathbf{n}}$$

The expressions look somewhat different, especially since, in eq. (10.11), $\hat{\mathbf{n}}$ points from the source *at retarded time* and R is the distance between the charge and the point of observation, $\mathbf{x}$, also at retarded time, whereas in our previous result the unit vector and the distance referred to the position of the charge at *observation time*.

Showing the equivalence is left as an EXERCISE. Here are the main milestones. First, write eq. (10.11) as:

$$\mathbf{E}_{\text{near}} \;=\; \frac{q}{\gamma^2R^2}\,\frac{\mathbf{R}\;-\;\boldsymbol{\beta}R}{(R-\mathbf{R}\cdot\boldsymbol{\beta})^3}\bigg|_{\tau_0}$$

Then calculate the retarded time τ_0 as a function of observer time t, starting from the definition: $R = c(t-\tau_0) = |\mathbf{x} - c\boldsymbol{\beta}\tau_0|$:

$$\tau_0 \;=\; \frac{\gamma^2}{c}\,\left[r \;-\; \sqrt{\tilde{r}^2 \;+\; (1-\beta^2)(x^2-c^2t^2)}\right]$$

where $x^2 = \mathbf{x}\cdot\mathbf{x}$, and $\tilde{r} = ct - \boldsymbol{\beta}\cdot\mathbf{x}$. Then, again using the definitions of R and $\mathbf{R}$, show that:

$$(R-\mathbf{R}\cdot\boldsymbol{\beta})\Big|_{\tau_0} \;=\; \tilde{r} \;-\; \frac{c\tau_0}{\gamma^2} \;=\; \sqrt{\tilde{r}^2 \;+\; (1-\beta^2)(x^2-c^2t^2)}$$

Next, expanding the argument of the square root, show that it can be written as $r^2 - \beta^2r^2 + (\boldsymbol{\beta}\cdot\mathbf{r})^2 = r^2 \;-\; \beta^2r^2\sin^2\psi$, where $\mathbf{r} = \mathbf{x} - \boldsymbol{\beta}ct$, the distance vector between the charge and the observer at present time. Finally, evaluating the numerator in eq. (10.11) at retarded time, show that the result is indeed what we obtained with the Lorentz transformation, with now $\hat{\mathbf{n}} = \mathbf{r}/r$. Perhaps now you can appreciate how useful the Lorentz transformation can be in the right circumstances!

We are much more interested in the far-field components of the electric field as given by eq. (10.10):

$$E^i_{\text{far}} \;=\; F^{\text{far}}_{0i} \;=\; q\,\left[\frac{R_0\,\mathrm{d}_\tau U_i - R_i\,\mathrm{d}_\tau U_0}{(\mathbf{U}\cdot\mathbf{R})^2} \;-\; \frac{(R_0\,U_i \;-\; R_i\,U_0)\mathbf{R}\cdot\mathrm{d}_\tau\mathbf{U}}{(\mathbf{U}\cdot\mathbf{R})^3}\right]_{\tau_0}$$

We still have: $\mathbf{U}\cdot\mathbf{R} = \gamma cR(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})$. There only remains to calculate:

$$\mathrm{d}_\tau U_i \;=\; \mathrm{d}_t(\gamma c\beta_i)\,\mathrm{d}_\tau t \;=\; -\,c\gamma^2\big[\dot{\beta}_i \;+\; \gamma^2\beta_i(\boldsymbol{\beta}\cdot\dot{\boldsymbol{\beta}})\big]$$

$$\mathrm{d}_\tau U_0 \;=\; \mathrm{d}_t(\gamma c)\,\mathrm{d}_\tau t \;=\; \gamma^4 c\,\boldsymbol{\beta}\cdot\dot{\boldsymbol{\beta}}$$

where we have used $\mathrm{d}_\tau t = \gamma$ and $\mathrm{d}_t\gamma = \dot\gamma = \gamma^3\boldsymbol{\beta}\cdot\dot{\boldsymbol{\beta}}$. (*Notation alert!* The dot now denotes differentiation with respect to observer time.) From this we find:

$$\mathbf{R}\cdot\mathrm{d}_\tau\mathbf{U} \;=\; \gamma^2 c\big[\gamma^2 R\,\boldsymbol{\beta}\cdot\dot{\boldsymbol{\beta}} - \dot{\boldsymbol{\beta}}\cdot\mathbf{R} - \gamma^2(\boldsymbol{\beta}\cdot\mathbf{R})(\boldsymbol{\beta}\cdot\dot{\boldsymbol{\beta}})\big] \;=\; -\gamma^2 cR\left[\dot{\boldsymbol{\beta}}\cdot\hat{\mathbf{n}} - \gamma^2(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})(\boldsymbol{\beta}\cdot\dot{\boldsymbol{\beta}})\right]$$

Putting everything together, there comes:

$$\begin{aligned}\mathbf{E}_{\text{far}} \;&=\; -\frac{q}{c}\left[\frac{\dot{\boldsymbol{\beta}} + \cancel{\gamma^2(\boldsymbol{\beta}-\hat{\mathbf{n}})(\boldsymbol{\beta}\cdot\dot{\boldsymbol{\beta}})}}{R\,(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^2} + \frac{\boldsymbol{\beta}-\hat{\mathbf{n}}}{R(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^3}\left[\dot{\boldsymbol{\beta}}\cdot\hat{\mathbf{n}} - \cancel{\gamma^2(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})(\boldsymbol{\beta}\cdot\dot{\boldsymbol{\beta}})}\right]\right]_{\tau_0} \\ &=\; \frac{q}{c}\left[\frac{(\hat{\mathbf{n}}-\boldsymbol{\beta})(\hat{\mathbf{n}}\cdot\dot{\boldsymbol{\beta}}) \;-\; \dot{\boldsymbol{\beta}}(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})}{R(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^3}\right]_{\tau_0} \;=\; \frac{q}{c}\left[\frac{\hat{\mathbf{n}}\times\big[(\hat{\mathbf{n}}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}\big]}{R(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^3}\right]_{\tau_0}\end{aligned} \tag{10.12}$$

Which of the two expressions on the last line to use depends on ease of calculation.

As is always the case for far-fields, the magnetic induction field is given by $\mathbf{B} = \hat{\mathbf{n}}\times\mathbf{E}$, with $\hat{\mathbf{n}}$ evaluated at retarded time.

10.3 Power Radiated by a Point-Charge (section J14.2)

Since the components of the Poynting vector are related to components of the energy-momentum tensor by $S^i = cT^{0i}$, we see from eq. (9.23) that in Gaussian units $\mathbf{S} = (c/4\pi)\mathbf{E}\times\mathbf{B}$. Now in the radiation (far) zone $\mathbf{B}_{\text{far}} = \mathbf{n}\times\mathbf{E}_{\text{far}}$, and $\mathbf{E}_{\text{far}}\times(\hat{\mathbf{n}}\times\mathbf{E}_{\text{far}}) = E_{\text{far}}^2\hat{\mathbf{n}} - \mathbf{E}_{\text{far}}(\hat{\mathbf{n}}\cdot\mathbf{E}_{\text{far}})$. Using eq. (10.12) and keeping only $1/R^2$ contributions leaves us with:

$$\mathbf{S}\cdot\hat{\mathbf{n}} \;=\; \frac{c}{4\pi}E_{\text{far}}^2 \;=\; \frac{q^2}{4\pi c}\,\frac{\Big[\hat{\mathbf{n}}\times\big[(\hat{\mathbf{n}}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}\big]\Big]^2}{R^2(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^6}$$

where the right-hand side is evaluated at retarded time, with $\hat{\mathbf{n}}$ pointing from the retarded position of the charge to the observation point. Note that $\mathbf{S}\cdot\hat{\mathbf{n}}$ is the energy per unit area per unit of *observer* time.

The angular power distribution is:

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} \;=\; \mathbf{S}\cdot\hat{\mathbf{n}}\,R^2 \;=\; \frac{q^2}{4\pi c}\,\frac{\Big[\hat{\mathbf{n}}\times\big[(\hat{\mathbf{n}}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}\big]\Big]^2}{(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^6} \tag{10.13}$$

The energy per unit time (power) per unit solid angle detected *at the point of observation* is:

$$\frac{\mathrm{d}E}{\mathrm{d}\Omega\,\mathrm{d}t} \;=\; \frac{\mathrm{d}P}{\mathrm{d}\Omega} \;=\; |\mathbf{S}\cdot\hat{\mathbf{r}}|\,R^2$$

An often more relevant question is: how much power is *emitted* by the charge per unit of *its* time? We will need to know the derivatives of $t' = t - R/c$. First:

$$\begin{aligned}\partial_t t' \;&= 1 - \frac{1}{c}\partial_t\sqrt{\mathbf{R}\cdot\mathbf{R}} \;=\; 1 - \frac{1}{2Rc}\partial_t(\mathbf{x}-\mathbf{r}')^2 \\ &=\; 1 + \frac{1}{c}\,\frac{\mathbf{x}-\mathbf{r}'}{R}\;\cdot\;\partial_{t'}\mathbf{r}'(t')\,\partial_t t' \\ &=\; 1 + \hat{\mathbf{n}}\cdot\boldsymbol{\beta}'\,\partial_t t'\end{aligned}$$

where we have written $R = \sqrt{\mathbf{R}\cdot\mathbf{R}}$ to deal with the absolute value in $R = |\mathbf{x}-\mathbf{r}'|$. Solving for $\partial_t t'$ yields:

$$\partial_t t' \;=\; \frac{1}{1 \;-\; \boldsymbol{\beta}\cdot\hat{\mathbf{n}}} \tag{10.14}$$

There comes:

$$\frac{\mathrm{d}P_\mathrm{r}}{\mathrm{d}\Omega} \;=\; \frac{\mathrm{d}E}{\mathrm{d}\Omega\,\mathrm{d}t'} \;=\; \frac{\mathrm{d}E}{\mathrm{d}\Omega\,\mathrm{d}t}\,\frac{\mathrm{d}t}{\mathrm{d}t'} \;=\; \frac{\mathrm{d}P}{\mathrm{d}\Omega}\,(1\,-\,\boldsymbol{\beta}\cdot\hat{\mathbf{r}})$$

where $\mathrm{d}P_\mathrm{r}/\mathrm{d}\Omega$ is the power *radiated* per unit solid angle.

Instead of integrating eq. (10.13) for the observed angular power distribution, or even the radiated one, it is easier to compute the power directly. This is a bit of a guess, and Jackson presents plausible arguments for such a guess. The result should go over to the expression calculated from the non-relativistic approximation, to eq. (10.13), which is not too hard to integrate.

We will take as our relativistic expression for the power:

$$P \;=\; \frac{2}{3}\,\frac{q^2}{m^2\,c^3}\,\mathrm{d}_\tau P^\mu\,\mathrm{d}_\tau P_\mu \;=\; \frac{2}{3}\,\frac{q^2}{m^2\,c^3}\,\left[-\frac{1}{c^2}\left(\frac{\mathrm{d}E}{\mathrm{d}\tau}\right)^2 \;+\; \left(\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}\tau}\right)^2\right] \tag{10.15}$$

where P is the energy-momentum 4-vector. This would mean that power is a relativistic invariant, which can actually be proved independently. Now, with $\mathrm{d}t = \gamma\mathrm{d}\tau$ and $E = \gamma mc^2$:

$$\mathrm{d}_\tau E \;=\; mc^2\,\mathrm{d}_\tau\gamma \;=\; \frac{1}{2}mc^2\,\mathrm{d}_t\gamma^2 \;=\; mc^2\,\gamma^4\,(\boldsymbol{\beta}\cdot\dot{\boldsymbol{\beta}})$$

and

$$\mathrm{d}_\tau\mathbf{p} \;=\; mc\,\gamma^2\,\dot{\boldsymbol{\beta}} \;+\; mc\,\gamma^4\,\boldsymbol{\beta}\,(\boldsymbol{\beta}\cdot\dot{\boldsymbol{\beta}})$$

Then (EXERCISE):

$$-\,(\mathrm{d}_\tau E/c)^2 \;+\; (\mathrm{d}_\tau\mathbf{p})^2 \;=\; m^2c^2\,\gamma^6\,\left[\dot{\boldsymbol{\beta}}^2 \;-\; (\boldsymbol{\beta}\times\dot{\boldsymbol{\beta}})^2\right]$$

$$P \;=\; \frac{2}{3}\,\frac{q^2}{c}\,\gamma^6\,\left[\dot{\boldsymbol{\beta}}^2 \;-\; (\boldsymbol{\beta}\times\dot{\boldsymbol{\beta}})^2\right] \;=\; \frac{2}{3}\,\frac{q^2}{c}\,\gamma^6\,\left[\dot{\boldsymbol{\beta}}_\parallel^2 \;+\; \frac{1}{\gamma^2}\,\dot{\boldsymbol{\beta}}_\perp^2\right] \tag{10.16}$$

where $\dot{\boldsymbol{\beta}}_\parallel$ and $\dot{\boldsymbol{\beta}}_\perp$ are the components of the acceleration parallel and transverse to the velocity, respectively. The left equality is known as **Liénard's formula**, and it reduces to the correct nonrelativistic expression, as it should. It might be tempting to conclude that acceleration in the same direction as velocity results in γ^2 more power being radiated than with transverse acceleration. But this conclusion is a little hasty! In particle accelerators, for instance, the centripetal acceleration $\dot{\boldsymbol{\beta}}_\perp$ is determined by the velocity and does not change the energy, whereas $\dot{\boldsymbol{\beta}}_\parallel$ is linked to the rate at which the energy changes, so to the power applied to the charge to increase its energy.

To see how to think about this the useful way, we must rewrite the collinear term. In the parallel direction, the above expression for $\mathrm{d}_\tau\mathbf{p}$ becomes $mc\gamma^2\dot{\beta}(1+\gamma^2\beta^2) = mc\gamma^4\dot{\beta}$.

The collinear power contribution to eq. (10.16) becomes the relativistic version of the **Larmor formula**:

$$P^\parallel_\mathrm{rad} \;=\; \frac{2}{3}\,\frac{q^2}{m^2\,c^3}\,\left(\frac{1}{\gamma}\mathrm{d}_\tau\mathbf{p}\right)^2 \;=\; \frac{2}{3}\,\frac{q^2}{m^2\,c^3}\,\dot{\mathbf{p}}^2 \tag{10.17}$$

where $\mathbf{p}$ is the *relativistic* momentum of the charge at retarded time. Since $p = \gamma mc\beta$, an d $E = \gamma mc^2$, then $\dot{p} = \mathrm{d}_{ct}(E\beta)$ which can be written as $\dot{p} = \beta\mathrm{d}_x(E\beta)$. This is especially useful in the case of ultrarelativistic motion, for then $\dot{p} \approx \mathrm{d}_x E$. We obtain an expression for the power radiated away in terms of the change in the energy per unit length of the charge:

$$P^\parallel_\mathrm{rad} \;\approx\; \frac{2}{3}\,\frac{q^2}{m^2\,c^3}\,\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)^2$$

But $\mathrm{d}E/\mathrm{d}x$ results from the application of an external force on the charge. The power $P^\parallel_\mathrm{in}$ fed in by this force is $P_\mathrm{in} = \mathrm{d}_t E = \beta c\mathrm{d}_x E \approx c\mathrm{d}_x E$, and we find the useful ratio:

$$\frac{P^\parallel_\mathrm{rad}}{P_\mathrm{in}} \;=\; \frac{2}{3}\,\frac{q^2/mc^2}{mc^2}\,\frac{\mathrm{d}E}{\mathrm{d}x} \tag{10.18}$$

This ratio is of order 1 if the external force can impart an energy mc^2 over a distance q^2/mc^2. That would mean a force capable of increasing an electron's energy by about 10^8 TeV/m, vastly beyond what is possible now and in the foreseeable future. Therefore, radiation losses are completely negligible in linear electron accelerators. And they are even more negligible when accelerating protons which are 2 000 times more massive.

Is the situation different for a charge moving in a circular orbit of radius r_0? Let us ignore any acceleration due to the driving force and focus on the centripetal acceleration. Here we can use the Liénard formula as is, and there comes:

$$P_{\text{rad}} \;=\; \frac{2}{3}\,\frac{q^2}{c}\,\gamma^4\,\dot{\beta}^2 \;=\; \frac{2}{3}\,\frac{q^2\,c}{r_0^2}\,(\gamma\beta)^4$$

where we have used the relation between centripetal acceleration and speed: $\dot{\beta} = \beta^2 c/r_0$. The energy radiated per revolution is:

$$\Delta E \;=\; \frac{2\pi r_0}{\beta c}\,P_{\text{rad}} \;=\; \frac{4\pi}{3}\,\frac{q^2\,\gamma^4\,\beta^3}{r_0}$$

In ultrarelativistic motion:

$$\Delta E_{\text{rad}} \;\approx\; \frac{4\pi}{3}\,\frac{q^2\,\gamma^4}{r_0} \;=\; \frac{4\pi}{3}\,\frac{q^2}{(mc^2)^4}\,\frac{E^4}{r_0}$$

A useful numerical version of this result for electrons and positrons is:

$$\Delta E_{\text{rad}}[\text{MeV}] \;=\; 8.85\times 10^{-2}\,\frac{E^4[\text{GeV}]}{r_0}$$

At the Large Electron-Positron (LEP) accelerator, the CERN machine of radius 5.5 km which collided electrons with positrons at 100 GeV until the end of the second millennium, 1.6 GeV was lost to radiation for each revolution, and therefore would have had to be fed in just to keep the beam circulating at 100 GeV! Clearly, linear accelerators are immensely less wasteful than circular ones for electrons or positrons. At the Large Hadron Collider, where protons are now accelerated to 7 TeV in the same tunnel where LEP operated, ΔE_{rad} is a much more reasonable (if still non negligible!) 3 keV per revolution. That inverse fourth power of the mass works in favour of protons in circular accelerators.

Finally, for arbitrary motion, the radiation is a superposition of contributions from $\mathbf{a}_\parallel$ and $\mathbf{a}_\perp$. When comparing the two, the Liénard formula gives a misleading impression. Instead, as we did in the collinear case, ask what is the power emitted for a given input from an external force. In circular motion, the power emitted is:

$$P^{\perp}_{\text{rad}} \;=\; \frac{2}{3}\,\frac{q^2}{m^2c^3}\,(\mathrm{d}_\tau \mathbf{p})^2 \;=\; \frac{2}{3}\,\frac{q^2}{m^2c^3}\,\gamma^2\,\dot{\mathbf{p}}^2$$

The factor $\dot{\mathbf{p}}^2$ depends on the external force; thus, for the same external force, the power radiated by the perpendicular contribution is γ^2 larger than the power radiated by the collinear contribution, eq. (10.17), which can then be ignored in the ultrarelativistic regime. We conclude that in ultrarelativistic motion, the power emitted by an accelerated charged particle is the same as if the particle were moving (instantaneously!) at constant speed on a circular trajectory.