# 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 $w^\alpha$ 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 $w^\alpha$ 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 $a^\alpha b^\nu c_\alpha u_\beta v^\beta$, $\alpha$ and $\beta$ are dummy indices with summation implied, and $\nu$ 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 $u^\alpha$ and $u_\alpha$? 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 $u^\alpha$ and $u_\alpha$ 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 $\epsilon_{ijk}$ is defined in the following subsection. Also, $\gamma$ 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 $\gamma$, 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**, $\epsilon_{ijk}$, 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 $\epsilon$ 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 $\epsilon$ components are 1 and 3. Set $j = 1$; then $k$ must be 3, again to have $\epsilon$ 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, $w^2 = w_2 = \epsilon_{231} u^3 v^1 + \epsilon_{213} u^1 v^3 = u^3 v^1 - u^1 v^3$. In the same way, $w^1 = u^2 v^3 - u^3 v^2$, and $w^3 = u^1 v^2 - u^2 v^1$, 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 $3 \times 3$ matrix whose elements are given by $u^i v^j$. An antisymmetric version of this product, called the exterior product $\mathbf{u} \wedge \mathbf{v}$, is easily constructed; it has components $u^i v^j - u^j v^i$ 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: $(\mathbf{u} \times \mathbf{v})^i = (\mathbf{u} \times \mathbf{v})_i = \frac{1}{2}\epsilon_{ijk}(u^j v^k - u^k v^j)$ in a Cartesian basis.

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

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

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

It will happen that *two* $\epsilon$'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 $\epsilon$'s (which position does not matter, because to go from one to the other, the permutation rules are applied twice, once to each $\epsilon$). *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: $\int\limits_{\text{line}}^{b} (\boldsymbol{\nabla} f) \cdot \mathrm{d}\mathbf{l} = f(b) - f(a)$;
2. Curl theorem (Stokes): $\int\limits_{\text{surface}} (\boldsymbol{\nabla} \times \mathbf{u}) \cdot \mathrm{d}\mathbf{a} = \oint\limits_{\text{line}} \mathbf{u} \cdot \mathrm{d}\mathbf{l}$;
3. Divergence theorem (Green, Gauss): $\int\limits_{\text{volume}} \boldsymbol{\nabla} \cdot \mathbf{u} \, \mathrm{d}^3 x = \oint\limits_{\text{surface}} \mathbf{u} \cdot \mathrm{d}\mathbf{a}$.

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: $\oint\limits_{\text{line}} \mathbf{u} \cdot \mathrm{d}\mathbf{l}$ is called the **circulation** of $\mathbf{u}$ around a closed path, whereas $\int\limits_{\text{surface}} \mathbf{u} \cdot \mathrm{d}\mathbf{a}$ 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**:

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

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, $(r, \theta, \phi)$, 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 $\phi$ as the polar angle and $\theta$ 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 $u_r$, $u_\theta$ and $u_\phi$. 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:

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

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:

| | $\hat{\mathbf{r}}$ | $\hat{\boldsymbol{\theta}}$ | $\hat{\boldsymbol{\phi}}$ |
|---|---|---|---|
| $\partial_r$ | $0$ | $0$ | $0$ |
| $\partial_\theta$ | $\hat{\boldsymbol{\theta}}$ | $-\hat{\mathbf{r}}$ | $0$ |
| $\partial_\phi$ | $\hat{\boldsymbol{\phi}}\,\sin\,\theta$ | $\hat{\boldsymbol{\phi}}\,\cos\,\theta$ | $-\hat{\mathbf{r}}\,\sin\,\theta - \hat{\boldsymbol{\theta}}\,\cos\,\theta$ |

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:

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

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 $\nabla^2\mathbf{u} = (\nabla^2 u_x)\hat{\mathbf{x}} + (\nabla^2 u_y)\hat{\mathbf{y}} + (\nabla^2 u_z)\hat{\mathbf{z}}$, it is *not* equal to $(\nabla^2 u_r)\hat{\mathbf{r}} + (\nabla^2 u_\theta)\hat{\boldsymbol{\theta}} + (\nabla^2 u_\phi)\hat{\boldsymbol{\phi}}$, 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 $4\pi$ 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

[†]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: $\partial_i(\sqrt{g}u^i)/\sqrt{g}$.

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:

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

from which it is obvious that the $\delta$-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 $\delta$-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 $4\pi$, 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

A function is said to have **compact support** if it goes to zero faster than $1/r^2$ as $r\to\infty$.

The **Helmholtz theorem** asserts that a vector function $\mathbf{F}(\mathbf{x})$ that vanishes at infinity is uniquely determined over space if its divergence and curl are known and if these have compact support. In that case:

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

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 $\xi$ 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 $\Phi$ into $\nabla^2\Phi = \partial_x^2\Phi + \partial_y^2\Phi + \partial_z^2\Phi$ and divide by $\Phi$ to obtain:

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

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:

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

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 $\alpha^2$ and $\beta^2$ 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 $\alpha^2$, 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 $\alpha$ 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:

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

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:

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

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

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

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 $\alpha^2$ and $\beta^2$ 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 $\epsilon_{ijk}$ 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 obviously 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:

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

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 $[\nabla^2,\, \mathbf{L}] = 0$.

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

$$[L_x,\, L_y] \;=\; \mathrm{i}\,L_z, \qquad [L_y,\, L_z] \;=\; \mathrm{i}\,L_x, \qquad [L_z,\, L_x] \;=\; \mathrm{i}\,L_y, \qquad [\mathbf{L}^2,\, \mathbf{L}] \;=\; 0 \tag{0.31}$$

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

The importance of eq.(0.31) 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:

$$[J_i,\, J_j] \;=\; \mathrm{i}\,\epsilon_{ijk}\,J_k, \tag{0.32}$$

Just from these properties, it is possible to derive the eigenvalues $\lambda$ of $\mathbf{J}^2$, ie. such that $\mathbf{J}^2 g = \lambda\, g$, and the eigenvalues $m$ of $J_z$, such that $J_z f = m\, f$, where $\lambda$ and $m$ are expected to be *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 as $f$. Since the eigenvalues of $J_z$ are non-degenerate, $\mathbf{J}^2 f$ must be 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 $\lambda$.

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

$$\mathbf{J}^2 \;=\; J_\pm\, J_\mp \;+\; J_z^2 \;\mp\; J_z \tag{0.33}$$

For a given value of $\lambda$, 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 (0.33), 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, $\lambda$, 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. (0.33), 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:

$$J_\pm f_{jm} \;=\; \sqrt{j(j+1) \;-\; m(m \pm 1)}\, f_{j,m\pm1} \tag{0.34}$$

### 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, $l := m_{\max}$ 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 (0.32).

The $\theta$ 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:

$$-\left[\frac{1}{\sin\theta}\,\mathrm{d}_\theta\left(\sin\theta\,\mathrm{d}_\theta\right) - \frac{m^2}{\sin^2\theta}\right] F(\theta) = l(l+1)\,F(\theta) \tag{0.35}$$

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**:

$$Y_l^m(\theta,\phi) = \frac{(-1)^m}{2^l l!}\sqrt{\frac{2l+1}{4\pi}\,\frac{(l-m)!}{(l+m)!}}\,(1-x^2)^{m/2}\left[\mathrm{d}_x^{l+m}(x^2-1)^l\right]\mathrm{e}^{\mathrm{i}m\phi} \qquad x = \cos\theta \tag{0.36}$$

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

$$Y_{l,-m}^*(\theta,\phi) = (-1)^m Y_{lm}(\theta,\phi) \tag{0.37}$$

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:

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}^*(\theta',\phi') Y_{lm}(\theta,\phi) = \delta(x-x')\delta(\phi-\phi') \qquad (x = \cos\theta) \tag{0.38}$$

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

$$P_l(x) \;=\; \frac{1}{2^l l!} d_x^l (x^2-1)^l \qquad\qquad (-1 \le x \le 1) \tag{0.39}$$

They satisfy the **Legendre equation**, which is simply eq. (0.35) 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 $\theta$ or $\phi$! See section 3.4 in Jackson for modifications needed when $\theta$ is restricted to a range smaller than $\pi$ in a solution independent of $\phi$.

### 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 $\nabla^2$ derived earlier, eq. (0.29), we write:

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

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

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

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: $\displaystyle \frac{1}{r^2}\,\mathrm{d}_r\big(r^2\,\mathrm{d}_r R_l(r)\big) \,+\, \left(\gamma(r) \,-\, \frac{l(l+1)}{r^2}\right) R_l(r) \;=\; 0$

When $\gamma(r) = 0$, this is the radial part of the Laplace equation which becomes, after the change of variable $r = \mathrm{e}^x$, $\mathrm{d}_x^2 R + \mathrm{d}_x R - l(l+1)R = 0$. Inserting a solution of the form $\mathrm{e}^{px}$ 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:

$$\Psi(r,\theta,\phi) \;=\; \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left(A_{lm}\,r^l \,+\, \frac{B_{lm}}{r^{l+1}}\right) Y_l^m(\theta,\phi) \tag{0.40}$$

If $\Psi$ 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:

$$A_{lm}^{r<a} \;=\; \frac{1}{a^l}\oint \Psi(a,\theta,\phi)Y_{lm}^*(\theta,\phi)\,\mathrm{d}\Omega \tag{0.41}$$

For $r \ge a$, $A_{lm}^{r>a} = 0$ to avoid a divergence as $r \to \infty$, and:

$$B_{lm}^{r>a} = a^{l+1}\oint \Psi(a,\theta,\phi)Y_{lm}^*(\theta,\phi)\,\mathrm{d}\Omega \tag{0.42}$$

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:

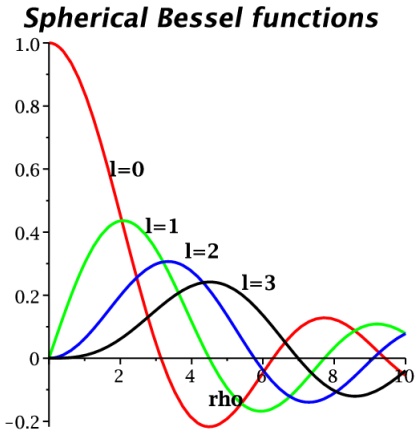
$$\mathrm{d}_x^2 u_l(x) \,+\, \frac{1}{x}\,\mathrm{d}_x u_l(x) \,+\, \left(1 \,-\, \frac{(l+1/2)^2}{x^2}\right) u_l(x) \;=\; 0$$

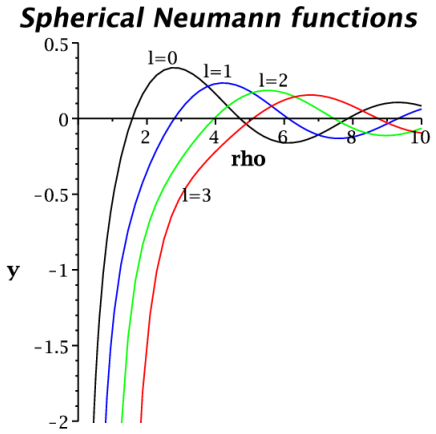
which is a form of the Bessel equation. The solutions are the **spherical Bessel functions** of the first and second (Neumann) kind, usually written as (see also Jackson's *Classical Electrodynamics*, section 9.6):

$$j_l(x) \;=\; \sqrt{\frac{\pi}{2x}}\,J_{l+1/2}(x) \;=\; (-x)^l \left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^l \left(\frac{\sin x}{x}\right) \quad \sim \begin{cases} x^l & x \ll (1,l) \\ \dfrac{1}{x}\sin(x - l\pi/2) & x \gg l \end{cases} \tag{0.43}$$

$$n_l(x) \;=\; \sqrt{\frac{\pi}{2x}}\,N_{l+1/2}(x) \;=\; -\,(-x)^l \left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^l \left(\frac{\cos x}{x}\right) \quad \sim \begin{cases} -\dfrac{1}{x^{l+1}} & x \ll (1,l) \\ -\dfrac{1}{x}\cos(x - l\pi/2) & x \gg l \end{cases} \tag{0.44}$$

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:

$$\begin{aligned} \mathrm{d}_t f(\mathbf{x}_0(t), t) \;&=\; \partial_t f \,+\, \left[\mathrm{d}_t\mathbf{x}_0 \cdot \boldsymbol{\nabla}\right] f \\ &=\; \partial_t f \,+\, (\mathbf{u} \cdot \boldsymbol{\nabla})\, f \end{aligned} \tag{0.45}$$

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,

$$\mathrm{d}_t\mathbf{f}(\mathbf{x}_0(t), t) \;=\; \partial_t f \,+\, \mathbf{u}\,\boldsymbol{\nabla} \cdot \mathbf{f} \,-\, \boldsymbol{\nabla} \times (\mathbf{u} \times \mathbf{f}) \tag{0.46}$$