

Additional Mathematical Tools: Detail

September 9, 2015

The material here is not required, but gives more detail on the additional mathematical tools: coordinate systems, rotations, the Dirac delta function and the proof of the Helmholtz theorem.

1 Cylindrical coordinates

Cylindrical (or polar) coordinates keep the Cartesian z unchanged, but label points in the xy -plane with a radius, ρ , and an angle φ measured counterclockwise (looking down the z -axis) from the x -axis. The relationship of cylindrical (ρ, φ, z) to Cartesian coordinates (x, y, z) is therefore

$$\begin{aligned}x &= \rho \cos \varphi \\y &= \rho \sin \varphi \\z &= z\end{aligned}$$

with inverse relations

$$\begin{aligned}\rho &= \sqrt{x^2 + y^2} \\ \varphi &= \tan^{-1}\left(\frac{y}{x}\right) \\ z &= z\end{aligned}$$

$$\begin{aligned}\frac{1}{\cos^2 \varphi} d\varphi &= \frac{dy}{x} - \frac{y dx}{x^2} \\ d\varphi &= \frac{\cos^2 \varphi}{\rho \cos \varphi} dy - \frac{\rho \cos^2 \varphi \sin \varphi dx}{\rho^2 \cos^2 \varphi} \\ d\varphi &= \frac{\cos \varphi}{\rho} dy - \frac{\sin \varphi}{\rho} dx\end{aligned}$$

We may define unit vectors in the direction of increase of each of these coordinates. The easiest way to develop the relationships we need is to express them in terms of the *constant* Cartesian basis. Since $\hat{\rho}$ lies in the xy plane,

$$\begin{aligned}\hat{\rho} &= \frac{x\hat{\mathbf{i}} + y\hat{\mathbf{j}}}{\sqrt{x^2 + y^2}} \\ &= \frac{\rho \cos \varphi \hat{\mathbf{i}} + \rho \sin \varphi \hat{\mathbf{j}}}{\sqrt{\rho^2}} \\ &= \cos \varphi \hat{\mathbf{i}} + \sin \varphi \hat{\mathbf{j}}\end{aligned}$$

and $\hat{\varphi}$ is orthogonal to this,

$$\hat{\varphi} = -\sin \varphi \hat{\mathbf{i}} + \cos \varphi \hat{\mathbf{j}}$$

while in the z direction we still have $\hat{\mathbf{k}}$. For any vector we may expand in either basis,

$$\mathbf{v} = v_\rho \hat{\boldsymbol{\rho}} + v_\varphi \hat{\boldsymbol{\varphi}} + v_z \hat{\mathbf{k}}$$

Inverting the relations we also have

$$\begin{aligned}\hat{\mathbf{i}} &= \cos \varphi \hat{\boldsymbol{\rho}} - \sin \varphi \hat{\boldsymbol{\varphi}} \\ \hat{\mathbf{j}} &= \sin \varphi \hat{\boldsymbol{\rho}} + \cos \varphi \hat{\boldsymbol{\varphi}}\end{aligned}$$

We can find the length, dl , of infinitesimal displacements in two ways. The difficult way is to change coordinates from the Cartesian distance,

$$dl^2 = dx^2 + dy^2 + dz^2$$

where we can find the differentials,

$$\begin{aligned}dx &= \cos \varphi d\rho - \rho \sin \varphi d\varphi \\ dy &= \sin \varphi d\rho + \rho \cos \varphi d\varphi \\ dz &= dz\end{aligned}$$

Substituting,

$$\begin{aligned}dl^2 &= (\cos \varphi d\rho - \rho \sin \varphi d\varphi)^2 + (\sin \varphi d\rho + \rho \cos \varphi d\varphi)^2 + dz^2 \\ &= (\cos^2 \varphi d\rho^2 - 2\rho \cos \varphi \sin \varphi d\rho d\varphi + \rho^2 \sin^2 \varphi d\varphi^2) \\ &\quad + (\sin^2 \varphi d\rho^2 + 2\rho \sin \varphi \cos \varphi d\rho d\varphi + \rho^2 \cos^2 \varphi d\varphi^2) + dz^2 \\ &= (\cos^2 \varphi + \sin^2 \varphi) d\rho^2 + \rho^2 (\sin^2 \varphi + \cos^2 \varphi) d\varphi^2 + dz^2 \\ &= d\rho^2 + \rho^2 d\varphi^2 + dz^2\end{aligned}$$

This is just the length of the infinitesimal displacement vector

$$d\mathbf{l} = d\rho \hat{\boldsymbol{\rho}} + \rho d\varphi \hat{\boldsymbol{\varphi}} + dz \hat{\mathbf{k}}$$

The second way to find the expression for dl^2 is to draw a picture. In the xy -plane, we can draw two vectors from the origin that are separated by a small angle $d\varphi$ and differ in length by $d\rho$. Then, since these are infinitesimals, we can use the Pythagorean theorem to write down the distance between the ends of the two vectors. The angular separation is by an arc-length $\rho d\varphi$ and the lengths differ by $d\rho$, so the Pythagorean theorem gives the length of the diagonal as

$$(diag)^2 = d\rho^2 + \rho^2 d\varphi^2$$

Adding the Cartesian third direction, we have

$$dl^2 = d\rho^2 + \rho^2 d\varphi^2 + dz^2$$

We also need derivatives. We can transform using the chain rule,

$$\begin{aligned}\frac{\partial}{\partial x} &= \frac{\partial \rho}{\partial x} \frac{\partial}{\partial \rho} + \frac{\partial \varphi}{\partial x} \frac{\partial}{\partial \varphi} + \frac{\partial z}{\partial x} \frac{\partial}{\partial z} \\ &= \frac{x}{\rho} \frac{\partial}{\partial \rho} - \frac{\sin \varphi}{\rho} \frac{\partial}{\partial \varphi} \\ &= \cos \varphi \frac{\partial}{\partial \rho} - \sin \varphi \frac{1}{\rho} \frac{\partial}{\partial \varphi}\end{aligned}$$

For y ,

$$\begin{aligned}\frac{\partial}{\partial y} &= \frac{\partial \rho}{\partial y} \frac{\partial}{\partial \rho} + \frac{\partial \varphi}{\partial y} \frac{\partial}{\partial \varphi} + \frac{\partial z}{\partial y} \frac{\partial}{\partial z} \\ &= \sin \varphi \frac{\partial}{\partial \rho} + \cos \varphi \frac{1}{\rho} \frac{\partial}{\partial \varphi}\end{aligned}$$

so the gradient in cylindrical coordinates is

$$\begin{aligned}\nabla f &= \hat{\mathbf{i}} \frac{\partial f}{\partial x} + \hat{\mathbf{j}} \frac{\partial f}{\partial y} + \hat{\mathbf{k}} \frac{\partial f}{\partial z} \\ &= (\cos \varphi \hat{\rho} - \sin \varphi \hat{\varphi}) \left(\cos \varphi \frac{\partial f}{\partial \rho} - \sin \varphi \frac{1}{\rho} \frac{\partial f}{\partial \varphi} \right) + (\sin \varphi \hat{\rho} + \cos \varphi \hat{\varphi}) \left(\sin \varphi \frac{\partial f}{\partial \rho} + \cos \varphi \frac{1}{\rho} \frac{\partial f}{\partial \varphi} \right) + \hat{\mathbf{k}} \frac{\partial f}{\partial z} \\ &= \hat{\rho} \left(\cos^2 \varphi \frac{\partial f}{\partial \rho} - \sin \varphi \cos \varphi \frac{1}{\rho} \frac{\partial f}{\partial \varphi} + \sin^2 \varphi \frac{\partial f}{\partial \rho} + \sin \varphi \cos \varphi \frac{1}{\rho} \frac{\partial f}{\partial \varphi} \right) \\ &\quad + \hat{\varphi} \left(-\sin \varphi \cos \varphi \frac{\partial f}{\partial \rho} + \sin^2 \varphi \frac{1}{\rho} \frac{\partial f}{\partial \varphi} + \sin \varphi \cos \varphi \frac{\partial f}{\partial \rho} + \cos^2 \varphi \frac{1}{\rho} \frac{\partial f}{\partial \varphi} \right) + \hat{\mathbf{k}} \frac{\partial f}{\partial z} \\ &= \hat{\rho} (\cos^2 \varphi + \sin^2 \varphi) \frac{\partial f}{\partial \rho} + \hat{\varphi} (\sin^2 \varphi + \cos^2 \varphi) \frac{1}{\rho} \frac{\partial f}{\partial \varphi} + \hat{\mathbf{k}} \frac{\partial f}{\partial z} \\ &= \hat{\rho} \frac{\partial f}{\partial \rho} + \hat{\varphi} \frac{1}{\rho} \frac{\partial f}{\partial \varphi} + \hat{\mathbf{k}} \frac{\partial f}{\partial z}\end{aligned}$$

We can compute the divergence and the curl in the same way, writing out the Cartesian expressions and changing coordinates. The process is lengthy. Fortunately, there is a much easier way to do it that depends on the *metric*.

2 General formulas for gradient, divergence and curl

In *any* coordinates, if we can write the infinitesimal distance between two points as

$$dl^2 = f^2 dx_1^2 + g^2 dx_2^2 + h^2 dx_3^2$$

then we can think of the distance as a double sum over the displacement vector (dx_1, dx_2, dx_3) ,

$$dl^2 = \sum_{i,j=1}^3 g_{ij} dx_i dx_j$$

where

$$g_{ij} = \begin{pmatrix} f^2 & 0 & 0 \\ 0 & g^2 & 0 \\ 0 & 0 & h^2 \end{pmatrix}$$

with inverse

$$\bar{g}_{ij} = \begin{pmatrix} \frac{1}{f^2} & 0 & 0 \\ 0 & \frac{1}{g^2} & 0 \\ 0 & 0 & \frac{1}{h^2} \end{pmatrix}$$

We can even include off-diagonal terms in g_{ij} as long as it is symmetric, i.e., if you flip the matrix across the diagonal, you get the same matrix back. In Euclidean space, there always exist coordinates where g_{ij} is diagonal. Let $\hat{\mathbf{e}}_i$ be three orthonormal basis vectors in the x_i directions. The formulas below hold for any symmetric metric g_{ij} , not just diagonal ones.

We need one more thing, that looks a lot like the square root of the metric. If the metric is diagonal, it actually *is* the square root. It's called a triad or a dreibein,

$$h_{ij} = \begin{pmatrix} f & 0 & 0 \\ 0 & g & 0 \\ 0 & 0 & h \end{pmatrix}$$

$$\bar{h}_{ij} = \begin{pmatrix} \frac{1}{f} & 0 & 0 \\ 0 & \frac{1}{g} & 0 \\ 0 & 0 & \frac{1}{h} \end{pmatrix}$$

The general definition of h_{ij} is that it satisfies

$$g_{ij} = \sum_{k,m=1}^3 h_{ik} h_{jm} \delta_{km}$$

$$= \sum_{k=1}^3 h_{ik} h_{jk}$$

Let the determinant of h_{ij} be h . Then we can write everything we need in terms of h_{ij} and its determinant. I have used advanced techniques to derive each of these within a few lines. The techniques are not beyond your ability, but it would take us too far afield to discuss them in detail now.

Gradient:

$$\nabla = \sum_{i,j=1}^3 \hat{\mathbf{e}}_i \bar{h}_{ij} \frac{\partial}{\partial x_j}$$

Divergence:

$$\nabla \cdot \mathbf{v} = \sum_{i,j=1}^3 \frac{1}{h} \frac{\partial}{\partial x_i} (h \bar{h}_{ij} v_j)$$

Curl:

$$\nabla \times \mathbf{v} = \frac{1}{h} \sum_{i,n,j,k,m} \hat{\mathbf{e}}_i h_{in} \varepsilon_{njkm} \partial_j (h_{km} v_m)$$

Laplacian:

$$\nabla^2 f = \sum_{i,j,k=1}^3 \frac{1}{h} \frac{\partial}{\partial x_k} \left[h \bar{h}_{ki} \bar{h}_{ij} \frac{\partial f}{\partial x_j} \right]$$

Example: cylindrical (or polar) coordinates

Let's compute these for cylindrical and spherical coordinates. The metric is the matrix

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

so that

$$h_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

with determinant $h = \rho$. The basis vectors are $\hat{\mathbf{e}}_i = (\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\varphi}}, \hat{\mathbf{k}})$

Gradient:

$$\begin{aligned} \nabla &= \sum_{i,j=1}^3 \hat{\mathbf{e}}_i \bar{h}_{ij} \frac{\partial}{\partial x_j} \\ &= \hat{\boldsymbol{\rho}} \frac{\partial}{\partial \rho} + \hat{\boldsymbol{\varphi}} \frac{1}{\rho} \frac{\partial}{\partial \varphi} + \hat{\mathbf{k}} \frac{\partial}{\partial z} \end{aligned}$$

Divergence:

$$\begin{aligned} \nabla \cdot \mathbf{v} &= \sum_{i,j=1}^3 \frac{1}{h} \frac{\partial}{\partial x_i} (h \bar{h}_{ij} v_j) \\ &= \frac{1}{\rho} \frac{\partial}{\partial x_i} (\rho v_\rho) + \frac{1}{\rho} \frac{\partial}{\partial \varphi} v_\varphi + \frac{\partial}{\partial z} v_z \end{aligned}$$

Curl:

$$\begin{aligned} \nabla \times \mathbf{v} &= \frac{1}{h} \sum_{i,n,j,k,m} \hat{\mathbf{e}}_i h_{in} \varepsilon_{njk} \partial_j (h_{km} v_m) \\ &= \frac{1}{\rho} \sum_{j,k,m} \hat{\boldsymbol{\rho}} \varepsilon_{1jk} \partial_j (h_{km} v_m) + \frac{1}{\rho} \sum_{j,k,m} \hat{\boldsymbol{\varphi}} \rho \varepsilon_{2jk} \partial_j (h_{km} v_m) + \frac{1}{\rho} \sum_{j,k,m} \hat{\mathbf{k}} \varepsilon_{3jk} \partial_j (h_{km} v_m) \\ &= \hat{\boldsymbol{\rho}} \left(\frac{1}{\rho} \partial_\varphi v_z - \partial_z v_\varphi \right) + \hat{\boldsymbol{\varphi}} (\partial_z v_\rho - \partial_\rho v_z) + \hat{\mathbf{k}} \left(\frac{1}{\rho} \partial_\rho (\rho v_\varphi) - \frac{1}{\rho} \partial_\varphi v_\rho \right) \end{aligned}$$

Laplacian:

$$\begin{aligned} \nabla^2 f &= \sum_{i,j,k=1}^3 \frac{1}{h} \frac{\partial}{\partial x_k} \left[h \bar{h}_{ki} \bar{h}_{ij} \frac{\partial f}{\partial x_j} \right] \\ &= \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial f}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \varphi^2} + \frac{\partial^2 f}{\partial z^2} \end{aligned}$$

Example: spherical coordinates

Now consider spherical coordinates, r, θ, φ . The metric is the matrix

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix}$$

so that

$$h_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & r \sin \theta \end{pmatrix}$$

with inverse

$$\bar{h}_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r} & 0 \\ 0 & 0 & \frac{1}{r \sin \theta} \end{pmatrix}$$

and determinant $h = r^2 \sin \theta$. The basis vectors are $\hat{\mathbf{e}}_i = (\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}})$.

Gradient:

$$\begin{aligned}\nabla &= \sum_{i,j=1}^3 \hat{\mathbf{e}}_i \bar{h}_{ij} \frac{\partial}{\partial x_j} \\ &= \hat{\mathbf{r}} \bar{h}_{11} \frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}} \bar{h}_{22} \frac{\partial}{\partial \theta} + \hat{\boldsymbol{\varphi}} \bar{h}_{33} \frac{\partial}{\partial \varphi} \\ &= \hat{\mathbf{r}} \frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\boldsymbol{\varphi}} \frac{1}{r \sin \theta} \frac{\partial}{\partial \varphi}\end{aligned}$$

Divergence:

$$\begin{aligned}\nabla \cdot \mathbf{v} &= \sum_{i,j=1}^3 \frac{1}{h} \frac{\partial}{\partial x_i} (h \bar{h}_{ij} v_j) \\ &= \frac{1}{r^2 \sin \theta} \left(\sin \theta \frac{\partial}{\partial r} (r^2 v_r) + r \frac{\partial}{\partial \theta} (\sin \theta v_\theta) + r \frac{\partial}{\partial \varphi} v_\varphi \right) \\ &= \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 v_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta v_\theta) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \varphi} v_\varphi\end{aligned}$$

Curl:

$$\begin{aligned}\nabla \times \mathbf{v} &= \frac{1}{h} \sum_{i,n,j,k,m} \hat{\mathbf{e}}_i h_{in} \varepsilon_{njkm} \frac{\partial}{\partial x_j} (h_{km} v_m) \\ &= \frac{1}{r^2 \sin \theta} \sum_{j,k,m} \left(\hat{\mathbf{r}} \varepsilon_{1jk} \frac{\partial}{\partial x_j} (h_{km} v_m) + \hat{\boldsymbol{\theta}} r \varepsilon_{2jk} \frac{\partial}{\partial x_j} (h_{km} v_m) + \hat{\boldsymbol{\varphi}} r \sin \theta \varepsilon_{3jk} \frac{\partial}{\partial x_j} (h_{km} v_m) \right) \\ &= \frac{1}{r^2 \sin \theta} \left(\hat{\mathbf{r}} \left(\frac{\partial}{\partial \theta} (h_{33} v_3) - \frac{\partial}{\partial \varphi} (h_{22} v_2) \right) + \hat{\boldsymbol{\theta}} r \left(\frac{\partial}{\partial \varphi} (h_{11} v_1) - \frac{\partial}{\partial r} (h_{33} v_3) \right) + \hat{\boldsymbol{\varphi}} r \sin \theta \left(\frac{\partial}{\partial r} (h_{22} v_2) - \frac{\partial}{\partial \theta} (h_{11} v_1) \right) \right) \\ &= \frac{1}{r^2 \sin \theta} \left(\hat{\mathbf{r}} \left(\frac{\partial}{\partial \theta} (r \sin \theta v_\varphi) - \frac{\partial}{\partial \varphi} (r v_\theta) \right) + \hat{\boldsymbol{\theta}} r \left(\frac{\partial}{\partial \varphi} (v_r) - \frac{\partial}{\partial r} (r \sin \theta v_\varphi) \right) + \hat{\boldsymbol{\varphi}} r \sin \theta \left(\frac{\partial}{\partial r} (r v_\theta) - \frac{\partial}{\partial \theta} v_r \right) \right) \\ &= \left(\hat{\mathbf{r}} \frac{1}{r \sin \theta} \left(\frac{\partial}{\partial \theta} (\sin \theta v_\varphi) - \frac{\partial}{\partial \varphi} v_\theta \right) + \hat{\boldsymbol{\theta}} \left(\frac{1}{r \sin \theta} \frac{\partial}{\partial \varphi} (v_r) - \frac{1}{r} \frac{\partial}{\partial r} (r v_\varphi) \right) + \hat{\boldsymbol{\varphi}} \left(\frac{1}{r} \frac{\partial}{\partial r} (r v_\theta) - \frac{1}{r} \frac{\partial}{\partial \theta} v_r \right) \right)\end{aligned}$$

Laplacian:

$$\begin{aligned}\nabla^2 f &= \sum_{i,j,k=1}^3 \frac{1}{h} \frac{\partial}{\partial x_k} \left(h \bar{h}_{ki} \bar{h}_{ij} \frac{\partial f}{\partial x_j} \right) \\ &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \varphi^2}\end{aligned}$$

3 Rotations

We wish to describe a rotation through an angle θ around an axis in $\hat{\mathbf{n}}$ direction. Positive θ is to correspond to a counterclockwise rotation when looking down at the tip of $\hat{\mathbf{n}}$. Let the rotation, O , act on an arbitrary vector \mathbf{v} to give the rotated vector \mathbf{v}' . To describe what happens, we decompose \mathbf{v} into parts parallel and perpendicular to $\hat{\mathbf{n}}$. The parallel part had magnitude $(\hat{\mathbf{n}} \cdot \mathbf{v})$ and direction $\hat{\mathbf{n}}$ so we define

$$\mathbf{v}_{\parallel} = (\hat{\mathbf{n}} \cdot \mathbf{v}) \hat{\mathbf{n}}$$

The part of \mathbf{v} perpendicular to $\hat{\mathbf{n}}$ has *magnitude* $|\hat{\mathbf{n}} \times \mathbf{v}| = v \sin \alpha$ where α is the angle between the two vectors. However, the vector $\hat{\mathbf{n}} \times \mathbf{v}$ is perpendicular to both $\hat{\mathbf{n}}$ and \mathbf{v} . We need a vector with this magnitude which is perpendicular to $\hat{\mathbf{n}}$ but lies in the $\hat{\mathbf{n}}\mathbf{v}$ -plane. We can get it by taking another cross product with $\hat{\mathbf{n}} \times \mathbf{v}$. Using the BAC-CAB rule,

$$\begin{aligned}\hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{v}) &= \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \mathbf{v}) - \mathbf{v} (\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}) \\ &= -[\mathbf{v} - \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \mathbf{v})]\end{aligned}$$

so we can write

$$\mathbf{v} = \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \mathbf{v}) - \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{v})$$

The two vectors on the right are perpendicular to one another and lie in the $\hat{\mathbf{n}}\mathbf{v}$ -plane. We define the perpendicular component of \mathbf{v} to be

$$\begin{aligned}\mathbf{v}_\perp &= -\hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{v}) \\ \mathbf{v}_\perp &= -\hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{v})\end{aligned}$$

Now we can easily write the rotated vector, \mathbf{v}' , because the parallel part is unchanged while the part lying in the $\hat{\mathbf{n}}\mathbf{v}$ -plane is a simple 2-dimensional rotation. We therefore have

$$\begin{aligned}\mathbf{v}'_\parallel &= \mathbf{v}_\parallel \\ \mathbf{v}'_\perp &= \mathbf{v}_\perp \cos \theta + (\hat{\mathbf{n}} \times \mathbf{v}) \sin \theta\end{aligned}$$

Adding these,

$$\begin{aligned}\mathbf{v}' &= \mathbf{v}_\parallel + \mathbf{v}_\perp \cos \theta + (\hat{\mathbf{n}} \times \mathbf{v}) \sin \theta \\ &= (\hat{\mathbf{n}} \cdot \mathbf{v}) \hat{\mathbf{n}} + (\mathbf{v} - \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \mathbf{v})) \cos \theta + (\hat{\mathbf{n}} \times \mathbf{v}) \sin \theta \\ &= (\hat{\mathbf{n}} \cdot \mathbf{v}) \hat{\mathbf{n}} (1 - \cos \theta) + \mathbf{v} \cos \theta + (\hat{\mathbf{n}} \times \mathbf{v}) \sin \theta\end{aligned}$$

To find the rotation matrix, O , we can look at the three Cartesian unit vectors, $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$. Since $\hat{\mathbf{i}} = (1, 0, 0)$, the matrix product $O\hat{\mathbf{i}}$ gives the first column of O , $O\hat{\mathbf{j}}$ gives the second column, and $O\hat{\mathbf{k}}$ the third. The first column is therefore

$$\begin{aligned}O\hat{\mathbf{i}} &= (\hat{\mathbf{n}} \cdot \hat{\mathbf{i}}) \hat{\mathbf{n}} (1 - \cos \theta) + \hat{\mathbf{i}} \cos \theta + (\hat{\mathbf{n}} \times \hat{\mathbf{i}}) \sin \theta \\ &= (n_x \hat{\mathbf{i}} + n_y \hat{\mathbf{j}} + n_z \hat{\mathbf{k}}) n_x (1 - \cos \theta) + \hat{\mathbf{i}} \cos \theta + (n_z \hat{\mathbf{j}} - n_y \hat{\mathbf{k}}) \sin \theta \\ &= (n_x^2 (1 - \cos \theta) + \cos \theta) \hat{\mathbf{i}} + (n_y n_x (1 - \cos \theta) + n_z \sin \theta) \hat{\mathbf{j}} + (n_z n_x (1 - \cos \theta) - n_y \sin \theta) \hat{\mathbf{k}}\end{aligned}$$

Similarly,

$$\begin{aligned}\mathbf{v}' &= \mathbf{v}_\parallel + \mathbf{v}_\perp \cos \theta + (\hat{\mathbf{n}} \times \mathbf{v}) \sin \theta \\ &= (\hat{\mathbf{n}} \cdot \mathbf{v}) \hat{\mathbf{n}} + (\mathbf{v} - \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \mathbf{v})) \cos \theta + (\hat{\mathbf{n}} \times \mathbf{v}) \sin \theta \\ O\hat{\mathbf{j}} &= (\hat{\mathbf{n}} \cdot \hat{\mathbf{j}}) \hat{\mathbf{n}} (1 - \cos \theta) + \hat{\mathbf{j}} \cos \theta + (\hat{\mathbf{n}} \times \hat{\mathbf{j}}) \sin \theta \\ &= n_x n_y (1 - \cos \theta) \hat{\mathbf{i}} + n_y^2 (1 - \cos \theta) \hat{\mathbf{j}} + n_y n_z (1 - \cos \theta) \hat{\mathbf{k}} + \hat{\mathbf{j}} \cos \theta + n_x \sin \theta \hat{\mathbf{k}} - n_z \sin \theta \hat{\mathbf{i}} \\ &= (n_x n_y (1 - \cos \theta) - n_z \sin \theta) \hat{\mathbf{i}} + (n_y^2 (1 - \cos \theta) + \cos \theta) \hat{\mathbf{j}} + (n_y n_z (1 - \cos \theta) + n_x \sin \theta) \hat{\mathbf{k}} \\ O\hat{\mathbf{k}} &= (\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}) \hat{\mathbf{n}} (1 - \cos \theta) + \hat{\mathbf{k}} \cos \theta + (\hat{\mathbf{n}} \times \hat{\mathbf{k}}) \sin \theta \\ &= n_z (n_x \hat{\mathbf{i}} + n_y \hat{\mathbf{j}} + n_z \hat{\mathbf{k}}) (1 - \cos \theta) + \hat{\mathbf{k}} \cos \theta + n_y \sin \theta \hat{\mathbf{i}} - n_x \sin \theta \hat{\mathbf{j}} \\ &= (n_x n_z (1 - \cos \theta) + n_y \sin \theta) \hat{\mathbf{i}} + (n_y n_z (1 - \cos \theta) - n_x \sin \theta) \hat{\mathbf{j}} + (n_z^2 (1 - \cos \theta) + \cos \theta) \hat{\mathbf{k}}\end{aligned}$$

As a matrix, $[O]$ is therefore,

$$[O]_{ij} = \begin{bmatrix} n_x^2(1 - \cos \theta) + \cos \theta & n_x n_y(1 - \cos \theta) - n_z \sin \theta & n_x n_z(1 - \cos \theta) + n_y \sin \theta \\ n_y n_x(1 - \cos \theta) + n_z \sin \theta & n_y^2(1 - \cos \theta) + \cos \theta & n_y n_z(1 - \cos \theta) - n_x \sin \theta \\ n_z n_x(1 - \cos \theta) - n_y \sin \theta & n_y n_z(1 - \cos \theta) + n_x \sin \theta & n_z^2(1 - \cos \theta) + \cos \theta \end{bmatrix}$$

Example:

Consider a rotation of $\frac{2\pi}{3} = 120^\circ$ about the direction $\hat{\mathbf{n}} = \frac{1}{\sqrt{3}}(1, 1, 1)$. Substituting the vector components with $\cos \frac{2\pi}{3} = -\frac{1}{2}$, $\sin \frac{2\pi}{3} = \frac{\sqrt{3}}{2}$ The first column given above is then

$$\begin{aligned} O\hat{\mathbf{i}} &= (\hat{\mathbf{n}} \cdot \hat{\mathbf{i}}) \hat{\mathbf{n}}(1 - \cos \theta) + \hat{\mathbf{i}} \cos \theta + (\hat{\mathbf{n}} \times \hat{\mathbf{i}}) \sin \theta \\ &= \frac{1}{\sqrt{3}} \hat{\mathbf{n}}(1 - \cos \theta) + \hat{\mathbf{i}} \cos \theta + (\hat{\mathbf{n}} \times \hat{\mathbf{i}}) \sin \theta \end{aligned}$$

$$\begin{aligned} O\hat{\mathbf{i}} &= (n_x^2(1 - \cos \theta) + \cos \theta) \hat{\mathbf{i}} + (n_y n_x(1 - \cos \theta) + n_z \sin \theta) \hat{\mathbf{j}} + (n_z n_x(1 - \cos \theta) - n_y \sin \theta) \hat{\mathbf{k}} \\ &= \left(\frac{1}{3} \left(1 + \frac{1}{2} \right) - \frac{1}{2} \right) \hat{\mathbf{i}} + \left(\frac{1}{3} \left(1 + \frac{1}{2} \right) + \frac{1}{\sqrt{3}} \frac{\sqrt{3}}{2} \right) \hat{\mathbf{j}} + \left(\frac{1}{3} \left(1 + \frac{1}{2} \right) - \frac{1}{\sqrt{3}} \frac{\sqrt{3}}{2} \right) \hat{\mathbf{k}} \\ &= \hat{\mathbf{j}} \end{aligned}$$

so we recover the result we found easily from the symmetry argument.

4 Dirac delta function

The Dirac delta function (actually a *distribution*) is defined as the limit of a sequence of functions,

$$\delta(x - x_0) = \lim_{n \rightarrow \infty} h_n(x)$$

such that

$$\begin{aligned} \delta(x - x_0) &= 0 \quad \text{for } x \neq x_0 \\ \int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx &= f(x_0) \quad \text{for any test function } f(x) \end{aligned}$$

By setting $f(x) = 1$, it follows from the integral expression that

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx = 1$$

and also that the integral of $\delta(x - x_0)$ over any interval not containing the point x_0 will give zero.

It is useful to think of an analogy with the Kronecker delta,

$$\begin{aligned} v_i &= \sum_{j=1}^3 \delta_{ij} v_j \\ f(x_0) &= \int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx \end{aligned}$$

where the discrete sum in the first expression is replaced by a continuous “sum” over x in the second.

One such limit is a sequence of Gaussians,

$$h_n(x) = \frac{n}{\sqrt{2\varphi}} \exp\left[-\frac{1}{2}n^2(x-x_0)^2\right]$$

where we choose each Gaussian to have unit integral. In the limit as $n \rightarrow \infty$, the exponential suppresses the n in the amplitude, except when $x = x_0$. At this point the limit distribution diverges. We can prove the basic integral property by considering small n .

$$\int_{-\infty}^{\infty} f(x) \delta(x-x_0) dx = \lim_{n \rightarrow \infty} \frac{n}{\sqrt{2\varphi}} \int_{-\infty}^{\infty} f(x) \exp\left[-\frac{1}{2}n^2(x-x_0)^2\right] dx$$

For any point $x \neq x_0$, we know that the limit goes to zero, so we can restrict the range of integration to a small neighborhood of x_0 ,

$$\int_{-\infty}^{\infty} f(x) \delta(x-x_0) dx \approx \lim_{n \rightarrow \infty} \frac{n}{\sqrt{2\varphi}} \int_{x_0-\varepsilon}^{x_0+\varepsilon} f(x) \exp\left[-\frac{1}{2}n^2(x-x_0)^2\right] dx$$

with the \approx becoming exact in the limit. We can set $\varepsilon = \frac{1}{n^2}$ and choose n large enough that we may approximate $f(x)$ by the first terms of its Taylor series,

$$f(x) \approx f(x_0) + f'(x_0)(x-x_0)$$

Since $(x-x_0) < \frac{1}{n^2}$ the integral for large n is approximately

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) \delta(x-x_0) dx &\approx \lim_{n \rightarrow \infty} \frac{n}{\sqrt{2\varphi}} \int_{x_0-\varepsilon}^{x_0+\varepsilon} (f(x_0) + f'(x_0)(x-x_0)) \exp\left[-\frac{1}{2}n^2(x-x_0)^2\right] dx \\ &= f(x_0) \lim_{n \rightarrow \infty} \frac{n}{\sqrt{2\varphi}} \int_{x_0-\varepsilon}^{x_0+\varepsilon} \exp\left[-\frac{1}{2}n^2(x-x_0)^2\right] dx \\ &\quad + f'(x_0) \lim_{n \rightarrow \infty} \frac{n}{\sqrt{2\varphi}} \int_{x_0-\varepsilon}^{x_0+\varepsilon} (x-x_0) \exp\left[-\frac{1}{2}n^2(x-x_0)^2\right] dx \\ &< f(x_0) + f'(x_0) \lim_{n \rightarrow \infty} \frac{n}{\sqrt{2\varphi}} \int_{x_0-\varepsilon}^{x_0+\varepsilon} \frac{1}{n^2} \exp\left[-\frac{1}{2}n^2(x-x_0)^2\right] dx \\ &= f(x_0) + f'(x_0) \lim_{n \rightarrow \infty} \frac{1}{n^2} \left(\frac{n}{\sqrt{2\varphi}} \int_{x_0-\varepsilon}^{x_0+\varepsilon} \exp\left[-\frac{1}{2}n^2(x-x_0)^2\right] dx \right) \\ &= f(x_0) \end{aligned}$$

where we use the fact that the Gaussian is normalized and in the finiteness of the second integral.

We can evaluate integrals of expressions containing derivatives of a Dirac delta function as well, using integration by parts. Remembering that

$$\frac{d}{dx}(fg) = \frac{df}{dx}g + f\frac{dg}{dx}$$

we may write

$$\frac{d}{dx}(f\delta(x-x_0)) = \frac{df}{dx}\delta(x-x_0) + f\frac{d}{dx}\delta(x-x_0)$$

If we integrate this over all x ,

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{d}{dx}(f\delta(x-x_0)) &= \int_{-\infty}^{\infty} \frac{df}{dx}\delta(x-x_0) + \int_{-\infty}^{\infty} f\frac{d}{dx}\delta(x-x_0) \\ [f\delta(x-x_0)]_{-\infty}^{\infty} &= \frac{df}{dx}(x_0) + \int_{-\infty}^{\infty} f\frac{d}{dx}\delta(x-x_0) \end{aligned}$$

Since the delta function vanishes at $\pm\infty$ (as does $f(x)$), the left side vanishes and we have

$$\int_{-\infty}^{\infty} f\frac{d}{dx}\delta(x-x_0) = -\frac{df}{dx}(x_0)$$

In higher dimensions, we simply multiply Dirac delta functions

$$\delta^3(\mathbf{x} - \mathbf{x}_0) = \delta(x - x_0)\delta(y - y_0)\delta(z - z_0)$$

This allows us to write a complete expression for charge densities even when they are idealized. For example, an ideal point charge at the origin has charge density

$$\rho(\mathbf{x}) = q\delta^3(\mathbf{x})$$

while a uniformly charged, infinitesimally thick ring of total charge Q and radius R , lying in the xy plane and centered at the origin, has charge density

$$\sigma(\mathbf{x}) = \frac{Q}{2\pi R}\delta(z)\delta(\rho - R)$$

in polar coordinates, since

$$\begin{aligned} \int_0^{\infty} \int_{-\infty}^{\infty} \int_0^{2\pi} \sigma(\mathbf{x}) \rho d\rho dz d\varphi &= \frac{Q}{2\pi R} \int_0^{\infty} \rho d\rho \int_{-\infty}^{\infty} dz \int_0^{2\pi} d\varphi \delta(z)\delta(\rho - R) \\ &= \frac{Q}{2\pi R} 2\pi \int_0^{\infty} \rho d\rho \int_{-\infty}^{\infty} dz \delta(z)\delta(\rho - R) \\ &= \frac{Q}{2\pi R} 2\pi \int_0^{\infty} \rho d\rho \delta(\rho - R) \\ &= \frac{Q}{2\pi R} 2\pi R \\ &= Q \end{aligned}$$

4.1 Example: spherical shell

Write a charge density using the Dirac delta function, $\delta(x)$, and/or the unit step function

$$\Theta(x-a) = \begin{cases} 1 & x-a > 0 \\ 0 & x-a < 0 \end{cases}$$

Example 1: Spherical shell, radius R , total uniform charge Q :

$$\begin{aligned}\rho(\mathbf{x}) &= \rho(r, \theta, \varphi) \\ &= A\delta(r - R)\end{aligned}$$

Fix the constant by computing the total charge, by integrating the density over all space,

$$\begin{aligned}Q &= \int \rho d^3x \\ &= A \int \delta(r - R) r^2 \sin\theta dr d\theta d\varphi \\ &= A \int_0^\infty \delta(r - R) r^2 dr \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\varphi \\ &= A \int_0^\infty \delta(r - R) r^2 dr \cdot 2 \cdot 2\pi \\ &= 4\pi R^2 A\end{aligned}$$

so $A = \frac{Q}{4\pi R^2}$ and we have

$$\rho(\mathbf{x}) = \frac{Q}{4\pi R^2} \delta(r - R)$$

5 Helmholtz theorem

Suppose we have a vector field, \mathbf{v} , in some region or all space, where we know the values of \mathbf{v} on the entire boundary or at infinity, and know its divergence and its curl, $\nabla \cdot \mathbf{v}$ and $\nabla \times \mathbf{v}$, everywhere. Then we may find \mathbf{v} .

To begin, separate \mathbf{v} into a divergence-free part and a curl-free part, with a possible remainder

$$\mathbf{v} = \mathbf{u} + \mathbf{w} + \mathbf{s}$$

We can do this if we can find \mathbf{u} satisfying

$$\begin{aligned}\nabla \cdot \mathbf{u} &= \nabla \cdot \mathbf{v} \\ \nabla \times \mathbf{u} &= 0 \\ \mathbf{u}|_S &= 0\end{aligned}$$

and \mathbf{v} satisfying

$$\begin{aligned}\nabla \times \mathbf{w} &= \nabla \times \mathbf{v} \\ \nabla \cdot \mathbf{w} &= 0 \\ \mathbf{w}|_S &= 0\end{aligned}$$

Then \mathbf{s} must satisfy

$$\begin{aligned}\nabla \cdot \mathbf{v} &= \nabla \cdot \mathbf{u} + \nabla \cdot \mathbf{w} + \nabla \cdot \mathbf{s} \\ &= \nabla \cdot \mathbf{v} + \nabla \cdot \mathbf{s} \\ \nabla \cdot \mathbf{s} &= 0\end{aligned}$$

and

$$\begin{aligned}\nabla \times \mathbf{v} &= \nabla \times \mathbf{u} + \nabla \times \mathbf{w} + \nabla \times \mathbf{s} \\ &= \nabla \times \mathbf{v} + \nabla \times \mathbf{s} \\ \nabla \times \mathbf{s} &= 0\end{aligned}$$

and the boundary condition. This gives

$$\begin{aligned}\nabla \cdot \mathbf{s} &= 0 \\ \nabla \times \mathbf{s} &= 0 \\ \mathbf{s}|_S &= \mathbf{v}|_S\end{aligned}$$

We set about solving these three sets of equations. First, for \mathbf{s} , we take a second curl

$$\begin{aligned}0 &= \nabla \times (\nabla \times \mathbf{s}) \\ &= \nabla (\nabla \cdot \mathbf{s}) - \nabla^2 \mathbf{s} \\ &= -\nabla^2 \mathbf{s}\end{aligned}$$

Therefore, \mathbf{s} satisfies the Laplace equation with boundary conditions:

$$\begin{aligned}\nabla^2 \mathbf{s} &= 0 \\ \mathbf{s}|_S &= \mathbf{v}|_S\end{aligned}$$

We will show in our study of electrostatics that this problem has a unique solution. Therefore, \mathbf{s} is determined. As a special case, \mathbf{s} vanishes if $\mathbf{v}|_S = 0$.

Next, we seek \mathbf{u} satisfying

$$\begin{aligned}\nabla \cdot \mathbf{u} &= \nabla \cdot \mathbf{v} \\ \nabla \times \mathbf{u} &= 0 \\ \mathbf{u}|_S &= 0\end{aligned}$$

This time, we may find a potential. Consider the line integral

$$\begin{aligned}\int_{C_1} \mathbf{u} \cdot d\mathbf{l} &= \int_{C_1; \mathcal{P}}^{\mathbf{x}} \mathbf{u} \cdot d\mathbf{l} \\ \int_{C_2} \mathbf{u} \cdot d\mathbf{l} &= \int_{C_2; \mathcal{P}}^{\mathbf{x}} \mathbf{u} \cdot d\mathbf{l}\end{aligned}$$

along any curves C_1, C_2 from a fixed point \mathcal{P} to some point \mathbf{x} . The difference of these gives a closed loop integral, and by Stokes' theorem gives

$$\begin{aligned}\int_{C_1} \mathbf{u} \cdot d\mathbf{l} - \int_{C_2} \mathbf{u} \cdot d\mathbf{l} &= \oint_{C_1 - C_2} \mathbf{u} \cdot d\mathbf{l} \\ &= \iint_S \hat{\mathbf{n}} \cdot (\nabla \times \mathbf{u}) d^2x \\ &= 0\end{aligned}$$

since the curl of \mathbf{u} vanishes. This means that we do not need to specify a curve when we integrate to \mathbf{x} ; the value is unique and defines a function,

$$f(\mathbf{x}) \equiv - \int_{\mathcal{P}}^{\mathbf{x}} \mathbf{u} \cdot d\mathbf{l}$$

and by the fundamental theorem of calculus,

$$u = -\nabla f$$

The negative sign is a convention. Rewriting the conditions for \mathbf{u} in terms of f , the vanishing curl is now automatic and f satisfies the Poisson equation with vanishing derivative on the boundary:

$$\begin{aligned}\nabla^2 f &= -\nabla \cdot \mathbf{v} \\ \nabla f|_S &= 0\end{aligned}$$

This system again has a unique solution.

Finally, consider \mathbf{w} :

$$\begin{aligned}\nabla \times \mathbf{w} &= \nabla \times \mathbf{v} \\ \nabla \cdot \mathbf{w} &= 0 \\ \mathbf{w}|_S &= 0\end{aligned}$$

We claim there is a vector field \mathbf{A} such that

$$\mathbf{w} = \nabla \times \mathbf{A}$$

and show that \mathbf{A} exists. Indeed, if one does there are many because we may set

$$\mathbf{A}' = \mathbf{A} + \nabla h$$

and we still have $\mathbf{w} = \nabla \times \mathbf{A}'$. Now, either \mathbf{A} or \mathbf{A}' must satisfy,

$$\begin{aligned}\nabla \times (\nabla \times \mathbf{A}) &= \nabla \times \mathbf{v} \\ \nabla \cdot (\nabla \times \mathbf{A}) &\equiv 0 \\ \nabla \times \mathbf{A}|_S &= 0\end{aligned}$$

The first equation becomes

$$\nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = (\nabla \times \mathbf{v})$$

where the divergence of \mathbf{A} , $\nabla \cdot \mathbf{A}$ is some unknown function. This means that the divergence of \mathbf{A}' is given by

$$\nabla \cdot \mathbf{A}' = \nabla \cdot \mathbf{A} + \nabla^2 h$$

so if we can solve

$$\nabla^2 h = -\nabla \cdot \mathbf{A}$$

then $\nabla \cdot \mathbf{A}' = 0$. This is once again the Poisson equation, and therefore has a solution, so we now have

$$\begin{aligned}\nabla(\nabla \cdot \mathbf{A}') - \nabla^2 \mathbf{A}' &= \nabla \times \mathbf{v} \\ \nabla^2 \mathbf{A}' &= -\nabla \times \mathbf{v}\end{aligned}$$

Letting $\mathbf{J} \equiv \nabla \times \mathbf{v}$, \mathbf{A}' now satisfies the Poisson equation

$$\nabla^2 \mathbf{A}' = -\mathbf{J}$$

with $\nabla \times \mathbf{A}'|_S = 0$ as boundary condition.

Combining these results, we have written our original vector \mathbf{v} as

$$\begin{aligned}\mathbf{v} &= \mathbf{u} + \mathbf{w} + \mathbf{s} \\ &= -\nabla f + \nabla \times \mathbf{A}' + \mathbf{s}\end{aligned}$$

where:

$$\begin{aligned}\nabla^2 \mathbf{s} &= 0 \\ \mathbf{s}|_S &= \mathbf{v}|_S\end{aligned}$$

and

$$\begin{aligned}\nabla^2 f &= -\nabla \cdot \mathbf{v} \\ \nabla f|_S &= 0\end{aligned}$$

and

$$\begin{aligned}\nabla^2 \mathbf{A}' &= -\nabla \times \mathbf{v} \\ \nabla \times \mathbf{A}'|_S &= 0\end{aligned}$$

This is the Helmholtz theorem.