

Electromagnetism Part 1: Mathematical Preliminaries

Mike Hughes

1 Introduction

These notes are a reminder of some of the maths that will be required in this part of the module. You have encountered most of the content of this document before, it is not new, but you need to revisit it now, **before** tackling the physics. The notes are a brief summary, you can consult your notes from Stage 1 maths modules for further details or work through the first chapter of the core reading [Griffiths, Introduction to Electrodynamics](#).

2 Scalars and Vectors

Scalars are quantities with only magnitude (e.g. speed) while vectors are quantities with magnitude and direction (e.g. velocity). Electric and magnetic fields are vector fields, we care about direction as well as magnitude, so we will use vectors extensively. You worked with vectors in several places in Stage 1 and we assume by now that you are comfortable with them. If you are not, then the very first thing you should do is review vectors (see the separate PDF on Moodle or your Stage 1 notes).

In this module we will primarily use the $\hat{i}$, $\hat{j}$ and $\hat{k}$ notation, but you should also be happy with column notation. For example,

$$\vec{A} = 3\hat{i} + 6\hat{j} - 2\hat{k} = \begin{bmatrix} 3 \\ 6 \\ -2 \end{bmatrix}$$

You should be comfortable taking the scalar (dot) and vector (cross) product of two vectors, and know what these mean in terms of the angles between the vector (e.g. the scalar product is zero for perpendicular vectors and the vector product is zero for parallel vectors). You should be able to compute the magnitude of a vector and find the unit vector that points in the same direction as a vector. For example, if we have a second vector,

$$\vec{B} = 2\hat{i} - 3\hat{j},$$

then we have,

$$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A} = (3 \times 2) + (6 \times -3) + (-2 \times 0) = -12 \quad (1)$$

$$\vec{A} \times \vec{B} = [(6)(0) - (-2)(-3)]\hat{i} - [(3)(0) - (-2)(2)]\hat{j} + [(3)(-3) - (6)(2)]\hat{k} = -4\hat{j} - 21\hat{k} \quad (2)$$

$$\hat{A} = \frac{\vec{A}}{|\vec{A}|} = \frac{3\hat{i} + 6\hat{j} - 2\hat{k}}{\sqrt{3^2 + 6^2 + (-2)^2}} = \frac{3}{7}\hat{i} + \frac{6}{7}\hat{j} - \frac{2}{7}\hat{k}. \quad (3)$$

If this is not obvious to you then you should start with the separate vectors review.

3 Basic Calculus

Again, you have studied calculus extensively, and you should be happy enough taking derivatives and finding integrals of simple functions such as polynomials, trig functions, e^x etc., remember the product, chain and quotient rules for differentiation, and have some idea how to compute more complex integrals via substitutions or integration by parts. You should remember how to do both indefinite and definite integrals. You should remember that the derivative of a function is zero at a turning point, with the sign of the second derivative telling you if this is a maximum or a minimum. If this is not all familiar, you should review this section of your Mathematics I notes from Stage 1.

4 Partial Derivatives

If we have some function, f that only depends on a single variable, x , then we often write it as $f(x)$. For example, gravitational potential energy is a function that, for an object with a given mass on Earth, only depends on its height, so we might write this as $U(h) = mgh$.

The derivative of the function, $\frac{df}{dx}$ is another function that tells us the slope of $f(x)$ for each value of x .

In general, functions may depend on more than one variable. The partial derivative tells us the slope of the function with respect to one of those variables. For example, suppose you are on a hillside, where your height depends on your x and y locations, i.e. there is some function $f(x, y)$. $\frac{\partial f}{\partial x}$ tells you the slope in the x direction and $\frac{\partial f}{\partial y}$ tells you the slope in the y direction at any point.

Partial derivatives are easy to calculate; simply treat all the other variables as constants and then differentiate in the normal way. For example, if,

$$f(x, y) = x^2 + 2xy^2,$$

then

$$\frac{\partial f}{\partial x} = 2x + 2y^2,$$

and

$$\frac{\partial f}{\partial y} = xy.$$

5 Vector Calculus

In electromagnetism we will be dealing with both scalar and vector fields. In a **scalar field** we define a scalar (i.e. a number) for each point in space. For example, in a weather report, the temperature is a scalar field. In a **vector field** we define a vector for each point in space. In a weather report, the wind is a vector field; we care about direction as well as speed. The electric field is an example of a vector field, while the electric potential is an example of a scalar field.

We use vector calculus to tell us something about how the field changes in space. Below we define the gradient (grad), divergence (div), and curl of a field. These do not take a single value for a field, each of these can be calculated for each point in space - i.e. when we apply them to a field we end up with another field!

The **gradient** of a scalar field tells us the direction of steepest ascent. For example, if our field was a height map, the gradient at each point is the vector pointing up the steepest slope, with the size of the vector proportional to the steepness of the slope (i.e. the gradient).

The **divergence** of a vector field tells us the net flow into or out of each point in space, i.e. the difference between the in-flow and out-flow. This is easiest to picture if we imagine a basin of moving liquid, with some pipes which bring in and take out the liquid, and where our vector field tells us the direction and speed that the fluid is moving in at each point. Points of positive divergence would be where water emerges from a pipe, negative divergence would be where water drains away into a pipe, and everywhere else has zero divergence.

The **curl** of a vector field tells us how much the field tends to rotate around each point in space. Imagine placing a tiny paddle wheel at a point in our basin of liquid. If the wheel starts spinning, there is non-zero curl at that point. The direction of the curl vector indicates the axis about which this rotation occurs and its magnitude measures how strong that rotational tendency is.

5.1 Del Operator

Operators are things that act on functions. In Cartesian (i.e. x, y, z) coordinates, the **del operator**, ∇ is defined as,

$$\nabla = \frac{\partial}{\partial x} \hat{i} + \frac{\partial}{\partial y} \hat{j} + \frac{\partial}{\partial z} \hat{k}.$$

Notice that, by itself, ∇ makes no sense, any more than $\frac{d}{dx}$ or $+$ make sense by themselves, it has to *act on something*.

5.2 Gradient (Grad)

When ∇ acts on a **scalar** function, we call this ‘**grad**’. This returns a vector function, which in Cartesian coordinates is,

$$\nabla F \equiv \frac{\partial F}{\partial x} \hat{i} + \frac{\partial F}{\partial y} \hat{j} + \frac{\partial F}{\partial z} \hat{k}.$$

At each point in the field, we take the derivative w.r.t. x , w.r.t. y and w.r.t. z and use these derivatives to define a new vector field. This field, at every point in space, points in the direction of maximum increase of the function F at that point. The magnitude $|\nabla F|$ gives the *rate* of increase along this direction.

Example: If $F(x, y, z) = 2x^2 + z^3$ then $\nabla F = 4x \hat{i} + 3z^2 \hat{k}$.

5.3 Divergence (Div)

Conceptually similar to taking a dot product, ‘**div**’ applied to a vector field gives a scalar field, in Cartesian coordinates,

$$\vec{\nabla} \cdot \vec{F} \equiv \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}.$$

Notice the difference from grad. We are now acting on a vector field and obtain a scalar field (i.e. the opposite). The scalar field tells us, at each point, how much the vector function is spreading out (or sinking).

Example: If $\vec{F}(x, y, z) = 4x \hat{i} + yz^2 \hat{k}$ then $\vec{\nabla} \cdot \vec{F} = 4 + 0 + 2yz = 4 + 2yz$.

5.4 Curl

Conceptually similar to taking a vector product, ‘**curl**’ applied to a vector field gives another vector field. In Cartesian coordinates,

$$\nabla \times \vec{F} \equiv \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} \equiv \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \hat{i} - \left(\frac{\partial F_z}{\partial x} - \frac{\partial F_x}{\partial z} \right) \hat{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \hat{k}.$$

It is harder to see what is going on here, but in a very loose sense it tells us to what extent the vector field swirls around each point.

Example: If $\vec{F}(x, y, z) = 4x \hat{i} + yz^2 \hat{k}$ then

$$\nabla \times \vec{F} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 4x & 0 & yz^2 \end{vmatrix} = \hat{i} \begin{vmatrix} \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & yz^2 \end{vmatrix} - \hat{j} \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial z} \\ 4x & yz^2 \end{vmatrix} + \hat{k} \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ 4x & 0 \end{vmatrix} = 2z \hat{i} \quad (4)$$

If this example doesn’t make sense, make sure you first know how to take the vector (cross) product of two vectors using the determinant notation (see vectors PDF). Now we are doing something similar, but replacing one of the vectors with the partial derivatives, and where we before were

multiplying we are now applying the partial derivative operators, i.e. for the first 2×2 determinant we have,

$$\frac{\partial}{\partial y}(yz^2) - \frac{\partial}{\partial z}(0) = 2z,$$

while the other two determinants come out to be zero.

5.5 Second Derivatives and the Laplacian Operator

We could apply grad, div and curl in sequence. For example, if we take the gradient of a scalar field F we end up with a vector field, call it $\vec{G}$,

$$\vec{G} = \nabla F$$

and we could now take the divergence or curl of this new field G . There are several other such possibilities, and we could even imagine taking third and fourth derivatives. However, it turns out that the only one that will be useful for us is if we take the divergence of the gradient of a scalar field. We call this the **Laplacian operator** and write it as,

$$\nabla^2 F = \nabla \cdot (\nabla F).$$

Example: Given the scalar field

$$F(x, y, z) = x^2y + 3z^2,$$

the Laplacian is,

$$\nabla^2 F = \frac{\partial^2}{\partial x^2}(x^2y + 3z^2) + \frac{\partial^2}{\partial y^2}(x^2y + 3z^2) + \frac{\partial^2}{\partial z^2}(x^2y + 3z^2) = 2y + 0 + 6 = 2y + 6.$$

We haven't written out the intermediate step of taking the gradient, which would give us a vector, as taking the divergence then gets us straight back to a scalar.

6 Line, Surface and Volume Integrals

Line, surface and volume integrals essentially involve 'adding up' the values of a field at all points along a line, over a surface or throughout a volume. This turns out to be a very useful thing to do in electromagnetism. This can be done for both scalar and vector fields; for scalar fields it is (conceptually) quite straightforward, while for vector fields we also need to think about direction.

6.1 Line Integral

If we have a scalar field, $F(x, y, z)$ then the line integral, I , of the field along a line between two points is,

$$I = \int_a^b F dl$$

Careful here, this is **not** the same as integrating a 1D function between two limits, a and b . Here, a and b are labels for two points in 3D space. We need to know the path that we are taking between the two points in 3D space to calculate the integral. (In a 1D function there is only one path between two points).

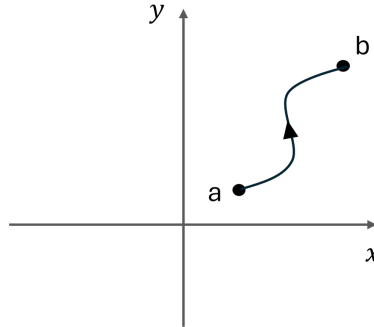


Figure 1: Illustration of a line integral between points a and b .

If we have a vector field, $\vec{F}(x, y, z)$ then the line integral is,

$$I = \int_a^b \vec{F} \cdot d\vec{l}.$$

Here, $d\vec{l}$ is the infinitesimal vector that is tangent to the line at every point. Recall that a dot product between two vectors tells us the extent to which they are parallel, so in a line integral we are going along the line and adding up the field values along the line, but scaled by the extent to which the field is pointing in the same direction as the line at each point. Points where the field and the line are perpendicular do not contribute at all, but the scalar product of two perpendicular vectors is always zero.

For some arbitrary path this can't really be done analytically, we would need to use a computer, but we will encounter some interesting situations where we can calculate it quite easily.

Note that, in general, the result of a line integral of either a scalar or a vector field depends on the specific path taken between the two points, not only the locations of the start and end points.

For example, take a vector field given by $\vec{F}(x, y) = (2x^2 + y)\hat{i} + y^2\hat{j}$, and we are going to compute the line integral from $a = (1, 1)$ and $b = (2, 2)$.

We could go along the horizontal first (C_1), and then up the vertical (C_2), in which case the total integral is,

$$\int_C \vec{F} \cdot d\vec{l} = \int_{C_1} \vec{F} \cdot d\vec{l} + \int_{C_2} \vec{F} \cdot d\vec{l}.$$

Along segment 1, C_1 from $(1, 1)$ to $(2, 1)$, $y = 1$ is constant, so $dy = 0$ and $d\vec{l} = dx\hat{i}$.

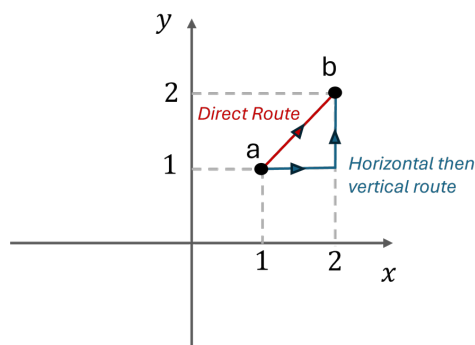


Figure 2: Line integral example.

The vector field becomes

$$\vec{F}(x, 1) = (2x^2 + 1)\hat{i} + 1^2\hat{j} = (2x^2 + 1)\hat{i} + \hat{j}.$$

Because we are going along the x -direction, $dl \cdot \hat{j} = 0$, and so,

$$\vec{F} \cdot d\vec{\ell} = (2x^2 + 1) dx.$$

The integral is,

$$\int_{C_1} \vec{F} \cdot d\vec{\ell} = \int_1^2 (2x^2 + 1) dx = \left[\frac{2}{3}x^3 + x \right]_1^2.$$

Similarly, the integral going up the vertical is,

$$\int_{C_2} \vec{F} \cdot d\vec{\ell} = \int_1^2 y^2 dy = \left[\frac{y^3}{3} \right]_1^2 = \frac{8}{3} - \frac{1}{3} = \frac{7}{3}.$$

and the total integral is,

$$\int_C \vec{F} \cdot d\vec{\ell} = \frac{17}{3} + \frac{7}{3} = \frac{24}{3} = 8.$$

Now, let's do the integral along the direct path from a to b . We can parametrise the straight line from $(1, 1)$ to $(2, 2)$ by

$$\vec{r}(t) = (1 + t)\hat{i} + (1 + t)\hat{j}, \quad 0 \leq t \leq 1.$$

Then

$$\frac{d\vec{r}}{dt} = \hat{i} + \hat{j}, \quad d\vec{\ell} = (\hat{i} + \hat{j}) dt.$$

Evaluate the field along the path:

$$x = 1 + t, \quad y = 1 + t,$$

so

$$F_x = 2x^2 + y = 2(1+t)^2 + (1+t) = 2(1+2t+t^2) + 1+t = 3+5t+2t^2,$$

$$F_y = y^2 = (1+t)^2 = 1+2t+t^2.$$

The integrand is the dot product $\vec{F} \cdot d\vec{\ell} = (F_x + F_y) dt$,

$$F_x + F_y = (3+5t+2t^2) + (1+2t+t^2) = 4+7t+3t^2,$$

and the line integral is,

$$\int_C \vec{F} \cdot d\vec{\ell} = \int_0^1 (4+7t+3t^2) dt = \left[4t + \frac{7}{2}t^2 + t^3 \right]_0^1 = \frac{17}{2}.$$

So we can see in this case that the path we take **does** matter.

However, there is a special class of fields for which the **specific path of the integral doesn't matter**: any path between the same start and end points gives the same result. These are called **conservative fields**. It turns out that the electrostatic field *is* conservative, and so we definitely do care about this 'special case'!

6.2 Surface Integrals

We can extend the idea of the line integral to two dimensions, where we are now integrating a field over a surface rather than along a line. Again, for a scalar field, we write a surface integral S as,

$$S = \iint F dA$$

while for a vector field,

$$S = \iint \vec{F} \cdot \hat{n} dA,$$

where $\hat{n}$ is the unit vectors perpendicular to the surface at each point on the surface (i.e the surface normal). We are taking the dot product of the field, $\vec{F}$, with this surface normal vector at each point, and so the product vanishes at points where the field is parallel to the surface, and is largest where it is perpendicular to the surface. Often we compact the notation by defining,

$$d\vec{a} = \hat{n} dA,$$

and so we write,

$$S = \iint \vec{F} \cdot d\vec{a}.$$

While this is a double integral, for convenience, we often write a single integral sign and then put a subscript S to remind us that it is a 2D surface integral, e.g.,

$$S = \int_S \vec{F} \cdot d\vec{A}.$$

There is an ambiguity in which way we make the vector $d\vec{A}$ or $\hat{n}$ point; for a horizontal surface we could make it point up or down. However, for a closed surface (i.e. one which encloses a volume, with no holes, then there is a clear ‘inside’ and ‘outside’ of the surface. (For example, if we consider the surface of the Earth, we know what we meant by ‘inside’ the Earth’s surface.) For closed surfaces, we always have the vector pointing outwards. If the integral is over a closed surface, we write,

$$S = \oint_S \vec{F} \cdot d\vec{A}.$$

6.3 Volume Integral

Finally, we can also integrate a scalar or vector field over a volume; again we are effectively summing the value of the field at all points within the defined volume. For a scalar field this is,

$$I = \iiint_V F dV = \int_V F dV,$$

where again we simplify notation by using a single integral sign with a V to indicate a triple integral.

We will rarely need to calculate a volume integral of a vector field, but for completeness,

$$I = \int_V \vec{F} dV = \hat{i} \int_V F_x dV + \hat{j} \int_V F_y dV + \hat{k} \int_V F_z dV,$$

i.e. we compute the integral separately for each component of the field, and then use these three components to give us a vector.

Example: We have a scalar field $F = x^2$, integrate this over a cube defined for $x = 0$ to 2 , $y = 0$ to 2 and $z = 0$ to 2 .

$$I = \int_0^2 \int_0^2 \int_0^2 F \, dx \, dy \, dz \quad (5)$$

$$= \int_0^2 \int_0^2 \int_0^2 x^2 \, dx \, dy \, dz \quad (6)$$

$$= \int_0^2 \int_0^2 \frac{8}{3} \, dy \, dz \quad (7)$$

$$= \int_0^2 \frac{16}{3} \, dz \quad (8)$$

$$= \frac{32}{3} \quad (9)$$

Note this is just a scalar, as we expect. We are integrating over the field, so we end up with a number, not a field.

6.4 Divergence Theorem

There is a convenient relation for a closed surface, called the **Divergence Theorem**, that we will make good use of, that relates the volume and the surface,

$$\oint_S \vec{F} \cdot d\vec{A} = \int_V \nabla \cdot \vec{F} \, dV.$$

We will look at this more closely when we study Gauss' Law, but the basic intuition is this: The divergence at each point in the volume measures the flow into and out of that point. But if we consider the neighbouring points, then anything that flows out of the original point must flow into those points. When we do the integral (i.e. essentially a sum), all these divergences end up cancelling out. Where this breaks down is when we get to the edge of the volume (i.e. the surface). So, it turns out that the volume integral of the divergence of a field ends up only depending on the surface integral of the field, i.e. the value of the field inside the surface doesn't matter!

6.5 Stokes' Theorem

Another useful theorem relates surface and line integrals,

$$\oint_{\partial S} \vec{F} \cdot d\vec{\ell} = \iint_S (\nabla \times \vec{F}) \cdot d\vec{A}.$$

This relates the circulation of a vector field around the boundary of a surface to the curl of the field through the surface itself.

The intuition is similar to that of the Divergence Theorem, but now we are concerned with *rotation* rather than *flow*. The curl at each point tells us how much the field rotates around that point. When we integrate the curl over a surface, most of the local rotation cancels out with that of neighbouring points, except at the boundary. The net curl that remains is exactly what shows up in the line integral around the edge.

7 Curvilinear (Non-Cartesian) Coordinate Systems

While a standard Cartesian (x, y, z) coordinates system may seem the most natural, many problems are easier to solve (and concepts easier to understand) if we use different coordinate systems that more naturally reflect the symmetry of the problem.

7.1 Polar Coordinates (r, ϕ)

In 2D we can use polar coordinates, where a point is specified in terms of its distance from the origin, r , and ϕ , which is the angle measured counter-clockwise from the x -axis to a line connecting the point to the origin.

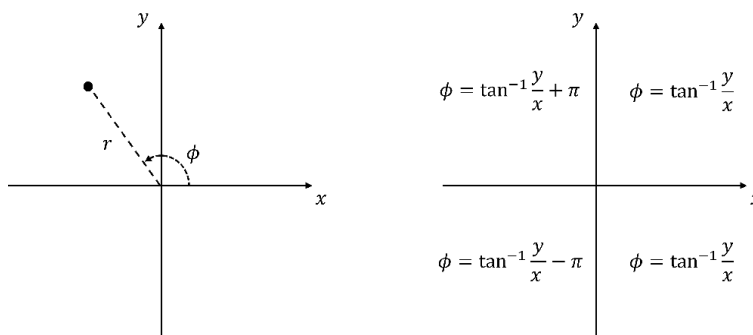


Figure 3: Left: Polar coordinates, Right: Adjustments needed when using arctan to find ϕ .

We have,

$$r = \sqrt{x^2 + y^2} \quad \tan \phi = \frac{y}{x}$$

where $\phi \in [0, 2\pi)$. When calculating ϕ via inverse tangent we have to keep track of the quadrant the point lies in, adding π if it is in the upper left, and subtracting π if it is in the lower left. (The function $\text{atan2}(y, x)$ is provided in programming languages to simplify this).

We also have,

$$x = r \cos \phi, \quad y = r \sin \phi.$$

7.2 Cylindrical Coordinates (ρ, ϕ, z)

In 3D, cylindrical coordinates are useful whenever we have cylindrical symmetry. We define ρ as the radial distance (like the r in polar coordinates) and ϕ as the angle as for polar coordinates, and then add an additional distance z , perpendicular to ρ .

We have,

$$\begin{aligned}
\rho &= \sqrt{x^2 + y^2}, & x &= \rho \cos \phi, \\
\tan \phi &= \frac{y}{x}, & y &= \rho \sin \phi, \\
z &= z, & z &= z.
\end{aligned}$$

where we have highlighted that the cylindrical z coordinates is the same as the cartesian z coordinate.

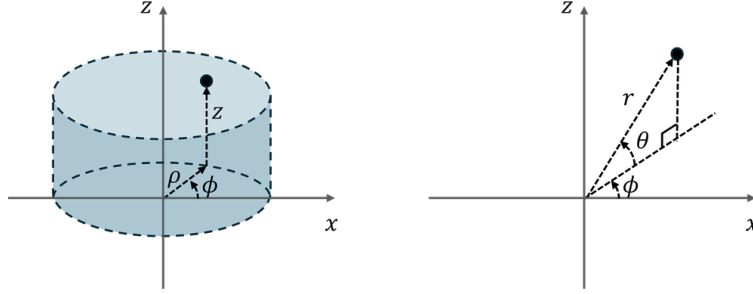


Figure 4: Left: Cylindrical coordinates, Right: Spherical coordinates.

7.3 Spherical Coordinates (r, θ, ϕ)

Finally, an alternative 3D coordinate system uses a distance, r and two angles, ϕ and θ , as shown in the diagram above. We have,

$$\begin{aligned}
r &= \sqrt{x^2 + y^2 + z^2}, & x &= r \sin \theta \cos \phi, \\
\cos \theta &= \frac{z}{r}, & y &= r \sin \theta \sin \phi, \\
\tan \phi &= \frac{y}{x}, & z &= r \cos \theta.
\end{aligned}$$

Note that $\phi \in [0, 2\pi)$, as before, but we only need $\theta \in [0, \pi)$ as we would otherwise have two ways of referencing each point in space (you can check this yourself).

7.4 Infinitesimal Lengths and Area

When working in Cartesian coordinates, we often need to state the infinitesimal lengths, e.g. dx , dy in calculus. In curvilinear coordinates, we need to be careful to think what we are doing.

For example, while in Cartesian coordinates an infinitesimal area is $dA = dx dy$, in polar coordinates, an infinitesimal area is not $dr d\theta$. (This does not even have units of area!). Recalling that the length of an arc is $r\theta$, then the infinitesimal distance along the arc is $r d\theta$ and so the infinitesimal area $dA = r dr d\theta$, as shown in the diagram. Note this actually depends on r , i.e the size of the infinitesimal area is different in different places.

Similarly, while an infinitesimal area of volume in Cartesian coordinates is $dV = dx dy dz$, in spherical coordinates it turns out to be $dV = r^2 \sin \theta dr d\theta d\phi$.

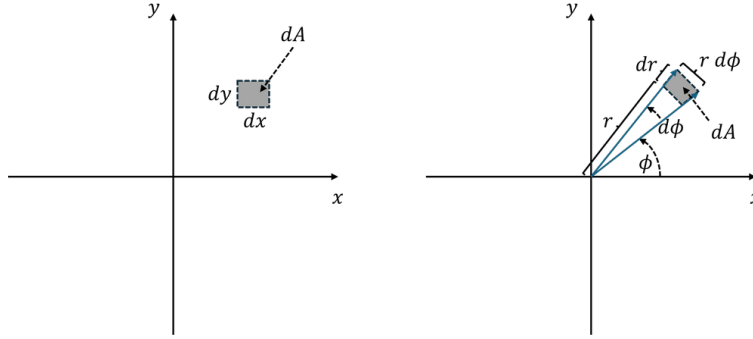


Figure 5: Left: Infinitesimal area in Cartesian coordinates, Right: Infinitesimal area in polar coordinates.

You can look up the infinitesimal length, areas and volumes for the different coordinates systems. There is no need to memorise them all, but you will need to be able to use them; we will do some example in the lectures and workshops, so for now just be aware.

7.5 Unit Vectors

As we have $\hat{i}$, $\hat{j}$ and $\hat{k}$ in Cartesian coordinates, we can similarly define unit vectors such as $\hat{r}$, $\hat{\theta}$, $\hat{\rho}$ and $\hat{\phi}$. Bear in mind that these have different Cartesian direction at different points in space. Again, we will cover this when we come to it, but be aware.

7.6 Vector Calculus

Finally, what happens with grad, div and curl when we use curvilinear coordinates? We could work out what dx , dy and dz are in the different coordinates systems and, with a bit of work, figure out expression for grad div and curl in cylindrical and spherical coordinates. For example, in spherical coordinates, if we have a vector field defined in terms of $\vec{F} = F_r \hat{r} + F_\theta \hat{\theta} + F_\phi \hat{\phi}$, then the divergence turns out to be,

$$\nabla \cdot \vec{F} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 F_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta F_\theta) + \frac{1}{r \sin \theta} \frac{\partial F_\phi}{\partial \phi}.$$

Again, you can look all these up if needed, and we will explore them more when we need to use them.