

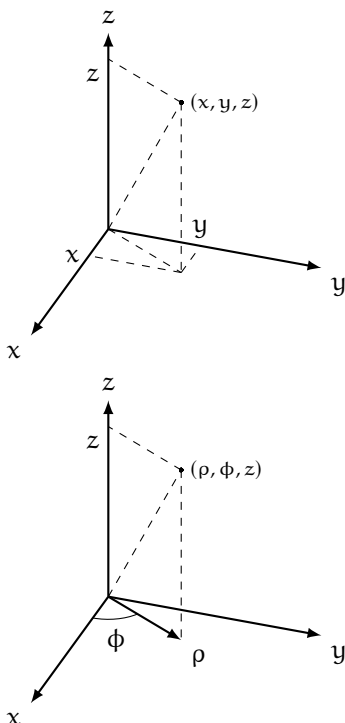
## Contents

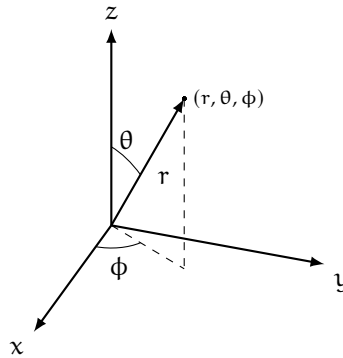
<b>Things You Just Gotta Know</b>	<b>2</b>
Coordinate Systems . . . . .	2
Polar Coordinates . . . . .	3
Spherical and Cylindrical Coordinates . . . . .	5
Scale Factors and Jacobians . . . . .	6
Complex Numbers . . . . .	7
Introduction . . . . .	7
Some Trigonometry with Complex Exponentials . . . . .	8
Index Algebra . . . . .	10
Contractions and Dummy Indices . . . . .	10
Two Special Tensors . . . . .	11
Binomial Theorem . . . . .	14
Infinite Series . . . . .	14
Power Series . . . . .	16
Taylor Series . . . . .	18
Ten Integration Techniques . . . . .	20
Integration by Parts . . . . .	20
Change of Variables . . . . .	21
Even/Odd . . . . .	23
Products and Powers of Sines and Cosines . . . . .	23
Axial and Spherical Symmetry . . . . .	24
Differentiation with Respect to a Parameter . . . . .	26
Gaussian Integral . . . . .	26
Completing the Square . . . . .	27
Series Expansion . . . . .	28
Partial Fractions . . . . .	29
Delta Distribution . . . . .	29
Properties of the Delta Distribution . . . . .	30
<b>Vector Calculus</b>	<b>34</b>
Vector Fields . . . . .	34
Gradient, Divergence, and Curl . . . . .	38
The $\nabla$ Operator . . . . .	38
Applying Vector Identities to the $\nabla$ Operator . . . . .	40
Changing Coordinates . . . . .	42
Understanding $\nabla^2$ , $\nabla \cdot$ , and $\nabla \times$ . . . . .	42
Integrating Scalar and Vector Fields . . . . .	46
Line Integrals . . . . .	46
Surface Integrals . . . . .	48
Circulation . . . . .	48
Flux . . . . .	52
Calculating Flux Integrals with Boundary Geometry . . . . .	56
Divergence Theorem . . . . .	57
Stokes's Theorem . . . . .	59
The Fundamental Theorem of Calculus, Revisited . . . . .	61
Helmholtz Decomposition . . . . .	63
Maxwell's Equations . . . . .	64
Integrating Maxwell . . . . .	64
From Integrals to Derivatives . . . . .	65
Substituting Potentials . . . . .	66

<b>Linear Algebra</b>	<b>67</b>
Superposition . . . . .	67
Vector Spaces . . . . .	68
Essentials of Vector Spaces . . . . .	68
Touching Base(s) . . . . .	68
Kets and Reps . . . . .	70
Inner Products . . . . .	71
Adjoins . . . . .	72
Cauchy–Schwarz Inequality . . . . .	75
Orthonormal Sets and Bases . . . . .	76
Constructing an Orthonormal Basis: Gram–Schmidt Decomposition . . . . .	81
Completeness . . . . .	84
Matrix Representation of Operators . . . . .	86
Rotations . . . . .	88
Introduction to Rotations . . . . .	88
What is a Rotation? . . . . .	89
Improper Rotations . . . . .	93

## Things You Just Gotta Know

### Coordinate Systems



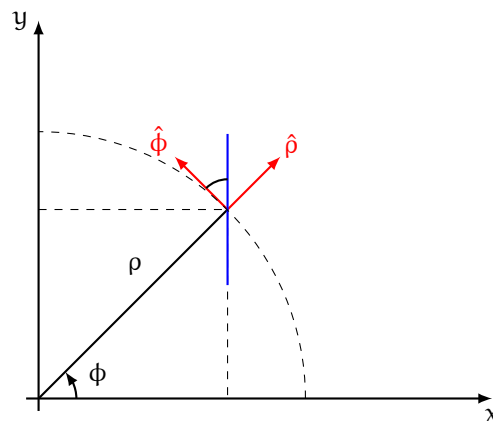


We want to focus on vector-valued functions of coordinates.

$$\vec{V}(\mathbf{r}) = V_x(x, y)\hat{i} + V_y(x, y)\hat{j}.$$

Notice that a vector function uses the coordinate system twice. Once for the function's inputs, once for the vectors themselves.

### Polar Coordinates



We can also express the inputs to  $\vec{V}$  in polar coordinates,  $(\rho, \phi)$ .

$$\vec{V}(\mathbf{r}) = V_\rho(\rho, \phi)\hat{i} + V_\phi(\rho, \phi)\hat{j}.$$

To extract the input functions, we take

$$V_x = \hat{i} \cdot \vec{V}$$

$$V_y = \hat{j} \cdot \vec{V}.$$

Alternatively, we can project  $\vec{V}$  onto the  $\hat{\rho}, \hat{\phi}$  axis:

$$\vec{V}(\mathbf{r}) = V_\rho(\rho, \phi)\hat{\rho} + V_\phi(\rho, \phi)\hat{\phi},$$

and we extract

$$V_\rho = \hat{\rho} \cdot \vec{V}$$

$$V_\phi = \hat{\phi} \cdot \vec{V}.$$

Notice that  $\mathbf{r}$  is an abstract vector; we need to project it onto a basis.

For instance, we can take the position vector and project it onto the cartesian and polar axes:

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

The main reason we avoided using the  $\hat{\rho}, \hat{\phi}$  axis up until this point is that  $\rho$  and  $\phi$  are *position-dependent*, while the  $\hat{i}, \hat{j}$  axis is position-independent.

Now, we must figure out the position-dependence of  $\hat{\rho}$  and  $\hat{\phi}$ :

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial \rho} d\rho + \frac{\partial \mathbf{r}}{\partial \phi} d\phi.$$

If we hold  $\phi$  constant, it must be the case that any change in  $\rho$  is in the  $\hat{\rho}$  direction. Therefore,

$$\begin{aligned}
 \hat{\rho} &= \frac{\frac{\partial \mathbf{r}}{\partial \rho}}{\left\| \frac{\partial \mathbf{r}}{\partial \rho} \right\|} \\
 &= \frac{\cos \phi \hat{i} + \sin \phi \hat{j}}{|\cos \phi \hat{i} + \sin \phi \hat{j}|} \\
 &= \cos \phi \hat{i} + \sin \phi \hat{j}.
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 \hat{\phi} &= \frac{\frac{\partial \mathbf{r}}{\partial \phi}}{\left\| \frac{\partial \mathbf{r}}{\partial \phi} \right\|} \\
 &= \frac{-\rho \sin \phi \hat{i} + \rho \cos \phi \hat{j}}{\|-\rho \sin \phi \hat{i} + \rho \cos \phi \hat{j}\|} \\
 &= -\sin \phi \hat{i} + \cos \phi \hat{j}.
 \end{aligned}$$

Thus, we can see that the  $\hat{\rho}, \hat{\phi}$  axis is orthogonal.

$$\begin{aligned}
 \frac{\partial \hat{\rho}}{\partial \phi} &= -\sin \phi \hat{i} + \cos \phi \hat{j} \\
 &= \hat{\phi}, \\
 \frac{\partial \hat{\phi}}{\partial \phi} &= -\hat{\rho}, \\
 \frac{\partial \hat{\phi}}{\partial \rho} &= 0,
 \end{aligned}$$

and

$$\frac{\partial \hat{\rho}}{\partial \rho} = 0$$

**Example (Velocity).**

$$\mathbf{v} = \frac{d\mathbf{s}}{dt}$$

$$= \frac{d}{dt} (x\hat{i}) + \frac{d}{dt} (y\hat{j}).$$

In the case of cartesian coordinates,  $\hat{i}$  and  $\hat{j}$  are constants.

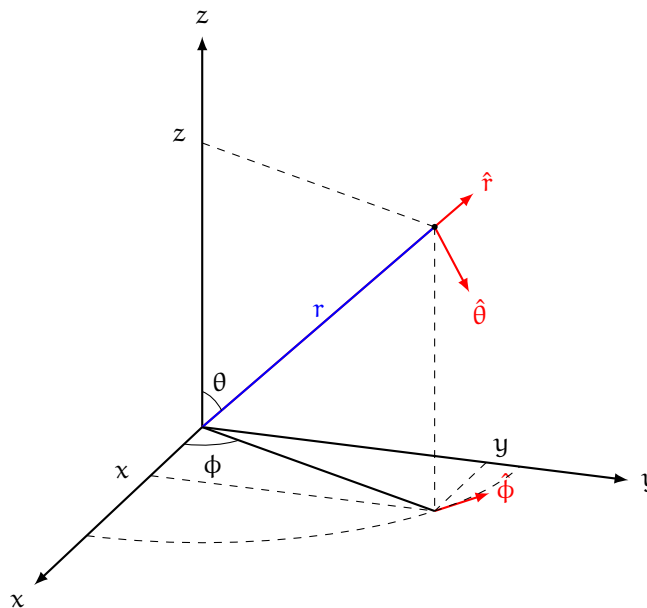
$$= v_x\hat{i} + v_y\hat{j}$$

When we examine polar coordinates, since  $\hat{\rho}$  and  $\hat{\phi}$  are position-dependent, we must use the chain rule.<sup>1</sup>

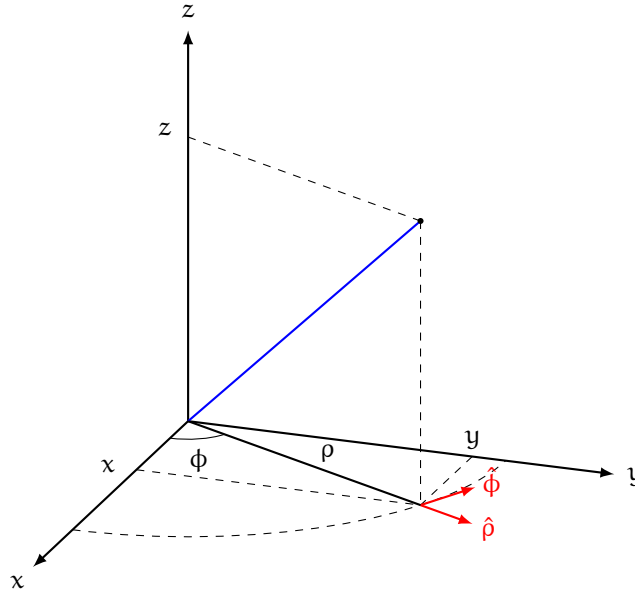
$$\begin{aligned} \mathbf{v} &= \frac{d\mathbf{s}}{dt} \\ &= \frac{d\rho}{dt}\hat{\rho} + \rho \frac{d\hat{\rho}}{dt} \\ &= \frac{d\rho}{dt}\hat{\rho} + \rho \left( \frac{\partial \hat{\rho}}{\partial \rho} \frac{d\rho}{dt} + \underbrace{\frac{\partial \hat{\rho}}{\partial \phi} \frac{d\phi}{dt}}_{=\hat{\phi}} \right) \\ &= \frac{d\rho}{dt}\hat{\rho} + \rho \frac{d\phi}{dt}\hat{\phi} \\ &= \dot{\rho}\hat{\rho} + \rho\dot{\phi}\hat{\phi}. \end{aligned}$$

Notice that  $\dot{\rho}$  is the radial velocity and  $\dot{\phi} = \omega$  is the angular velocity.

### Spherical and Cylindrical Coordinates



<sup>1</sup>Note that  $\hat{\rho} = \hat{\rho}(\rho, \phi)$  and  $\hat{\phi} = \hat{\phi}(\rho, \phi)$ .



Here,<sup>11</sup>  $\phi$  denotes the polar angle and  $\theta$  denotes the azimuthal angle. Notice that  $\phi \in [0, 2\pi)$  and  $\theta \in [0, \pi]$ .

We can see that  $\hat{\rho}$ ,  $\hat{\phi}$ , and  $\hat{\theta}$  in spherical coordinates are also position-dependent.

$$\begin{aligned}\hat{r} &= \frac{\frac{\partial \mathbf{s}}{\partial r}}{\left\| \frac{\partial \mathbf{s}}{\partial r} \right\|} \\ &= \sin \theta \cos \phi \hat{i} + \sin \theta \sin \phi \hat{j} + \cos \theta \hat{k} \\ \hat{\phi} &= \frac{\frac{\partial \mathbf{s}}{\partial \phi}}{\left\| \frac{\partial \mathbf{s}}{\partial \phi} \right\|} \\ &= -\sin \phi \hat{i} + \cos \phi \hat{j} \\ \hat{\theta} &= \frac{\frac{\partial \mathbf{s}}{\partial \theta}}{\left\| \frac{\partial \mathbf{s}}{\partial \theta} \right\|} \\ &= \cos \phi \cos \theta \hat{i} + \cos \phi \sin \theta \hat{j} - \sin \theta \hat{k}\end{aligned}$$

### Scale Factors and Jacobians

In cylindrical coordinates, we can use the chain rule to find the value of  $d\mathbf{r}$ :

$$d\mathbf{r} = \hat{\rho} d\rho + \rho \hat{\phi} d\phi + \hat{k} dz.$$

The extra factor of  $\rho$  in the expression of  $\rho \hat{\phi} d\phi$  is the *scale factor* on  $\phi$ .

Similarly, in spherical coordinates, we have

$$d\mathbf{r} = \hat{r} dr + r \sin \theta \hat{\phi} d\phi + r \hat{\theta} d\theta,$$

with scale factors of  $r \sin \theta$  on  $\hat{\phi} d\phi$  and  $r$  on  $\hat{\theta} d\theta$ .

When we go from line elements (of the form  $d\mathbf{r}$ ) to area elements (of the form  $d\mathbf{a}$ ), we can see that the area element in polar coordinates is  $d\mathbf{a} = \rho d\rho d\phi$  — we need the extra factor of  $\rho$  to account for the fact that the

<sup>11</sup>Physicists amirite?

magnitude of the area element scales with the radius.

Similarly, the volume element in cylindrical coordinates is  $d\tau = r dr d\phi dz$  and the volume element in spherical coordinates is  $r^2 \sin \theta dr d\phi d\theta$ .

Recall that the definition of an angle  $\phi$  that subtends an arc length  $s$  is  $\phi = \frac{s}{r}$ , where  $r$  is the radius of a circle. We can imagine a similar concept on a sphere — a *solid angle* measured in steradians is of the form  $\Omega = \frac{A}{r^2}$ , where  $A$  denotes the surface area subtended by the angle  $\Omega$ . In particular, since  $d\Omega = \frac{dA}{r^2}$ , we find that  $d\Omega = \sin \theta d\phi d\theta$ .

When we are dealing with products of scale factors, we need to use the Jacobian to determine the proper scale factor on any given element:

$$\begin{aligned} d\mathbf{a} &= dx dy \\ &= |J| du dv, \end{aligned}$$

where  $|J|$  denotes the determinant of the Jacobian matrix. We write the Jacobian as follows:

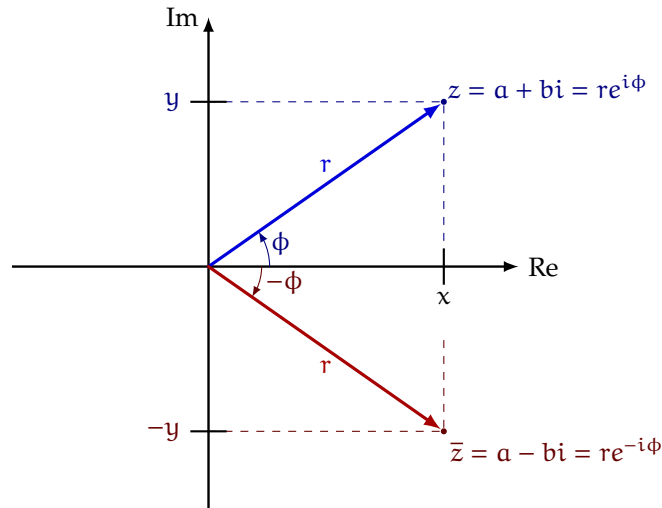
$$\begin{aligned} J &= \frac{\partial(x, y)}{\partial(u, v)} \\ &= \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix}. \end{aligned}$$

We specifically desire the determinant:

$$|J| = \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial y}{\partial u} \frac{\partial x}{\partial v}.$$

## Complex Numbers

### Introduction



A complex number is denoted

$$z = a + bi$$

where  $i^2 = -1$  and  $a, b \in \mathbb{R}$ . This is known as the Cartesian representation. However, we can also imagine  $z$  as the polar representation:

$$z = re^{i\phi},$$

where  $\phi = \arg z$  is known as the argument, and  $r = |z|$  is the modulus. We can see the relation between the Cartesian and polar representations through Euler's identity:<sup>III</sup>

$$r(\cos \phi + i \sin \phi) = re^{i\phi}.$$

We denote the conjugate of  $z$  as  $\bar{z}$ , found by  $\bar{z} = a - bi = re^{-i\phi}$ .

We find  $\text{Re}(z)$  and  $\text{Im}(z)$ , the real and imaginary parts of  $z$ , by

$$\begin{aligned}\text{Re}(z) &= \frac{z + \bar{z}}{2} \\ \text{Im}(z) &= \frac{z - \bar{z}}{2i}.\end{aligned}$$

We say that a complex number of the form  $e^{i\phi}$  is a *pure phase*, as  $|e^{i\phi}| = 1$ .

To find if some complex number  $z$  is purely real or purely imaginary, we can use the following criterion:

$$\begin{aligned}z \in \mathbb{R} &\Leftrightarrow z = \bar{z} \\ z \in i\mathbb{R} &\Leftrightarrow z = -\bar{z}.\end{aligned}$$

**Example** (Real, Imaginary, or Complex?). Consider

$$z_1 = i^i.$$

To find if this is purely real or complex, we take

$$\begin{aligned}\bar{z}_1 &= (-i)^{-i} \\ &= \left(\frac{1}{-i}\right)^i \\ &= i^i.\end{aligned}$$

Thus,  $z_1 \in \mathbb{R}$ . In order to determine the value of  $i^i$ , we substitute the polar form:

$$\begin{aligned}z_1 &= \left(e^{i\frac{\pi}{2}}\right)^i \\ &= e^{-\frac{\pi}{2}}.\end{aligned}$$

### Some Trigonometry with Complex Exponentials

Consider  $z = \cos \phi + i \sin \phi$ . We can see that

$$\begin{aligned}\text{Re}(z) &= \cos \phi \\ &= \frac{(\cos \phi + i \sin \phi) + (\cos \phi - i \sin \phi)}{2} \\ &= \frac{e^{i\phi} + e^{-i\phi}}{2} \\ \text{Im}(z) &= \sin \phi \\ &= \frac{(\cos \phi + i \sin \phi) - (\cos \phi - i \sin \phi)}{2i} \\ &= \frac{e^{i\phi} - e^{-i\phi}}{2i}.\end{aligned}$$

We can actually define  $\sin \phi$  and  $\cos \phi$  with the above derivation.

---

<sup>III</sup>This can be proven relatively easily through substitution into the Taylor series, which is allowed because  $e^z$  is entire.



**Theorem** (De Moivre):

$$\begin{aligned} e^{inx} &= \cos(nx) + i \sin(nx) \\ &= \left(e^{ix}\right)^n \\ &= (\cos x + i \sin x)^n. \end{aligned}$$

**Example** (Finding  $\cos(2x)$  and  $\sin(2x)$ ).

$$\begin{aligned} \cos(2x) + i \sin(2x) &= (\cos x + i \sin x)^2 \\ &= (\cos^2 x - \sin^2 x) + i(2 \sin x \cos x). \end{aligned}$$

Since the real parts and imaginary parts have to be equal, this means

$$\begin{aligned} \cos 2x &= \cos^2 x - \sin^2 x \\ \sin^2 x &= 2 \sin x \cos x. \end{aligned}$$

In particular, we can see that  $e^{in\pi} = (-1)^n$  and  $e^{in\frac{\pi}{2}} = i^n$ .<sup>IV</sup>

Additionally, we can see that for  $z = re^{i\phi}$ ,

$$\begin{aligned} z^{1/m} &= \left(re^{i\phi+2\pi n}\right)^{1/m} \\ &= r^{1/m} e^{i\frac{1}{m}(\phi+2\pi n)}, \end{aligned}$$

where  $n \in \mathbb{N}$  and  $m$  is fixed. For  $r = 1$ , we call these values the  $m$  roots of unity.

**Example** (Waves and Oscillations). Recall that for a wave with spatial frequency  $k$ , angular frequency  $\omega$ , and amplitude  $A$ , the wave is represented by

$$f(x, t) = A \cos(kx - \omega t).$$

The speed of a wave  $v$  is equal to  $\frac{\omega}{k}$ .

Simple harmonic motion is characterized by the solution to the differential equation  $\ddot{x} = -\omega^2 x$ , where  $x$  denotes position. In simple harmonic motion, there is no spatial motion, meaning our function is only of time:

$$\begin{aligned} f(t) &= A \cos \omega t \\ &= \operatorname{Re} \left( A e^{i\omega t} \right). \end{aligned}$$

As a result of the representation of complex numbers in polar form, we can do math entirely in exponentials, then take the real part of our solution to find  $f(t)$ .

Unfortunately, in the real world, there is friction; as a result, our oscillation is damped by an exponential factor.

**Example** (Hyperbolic Sine and Hyperbolic Cosine). We wish to calculate  $\cos ix$  and  $\sin ix$ .

$$\begin{aligned} \cos ix &= \frac{1}{2} \left( e^{i(ix)} + e^{-i(ix)} \right) \\ &= \frac{e^{-x} + e^x}{2} \end{aligned}$$

---

<sup>IV</sup>This will be especially useful when we get to Fourier series.

We define  $\cosh x = \cos(ix)$ . Additionally,

$$\begin{aligned} -i \sin ix &= -i \frac{1}{2i} (e^{i(ix)} - e^{-i(ix)}) \\ &= i \frac{e^x - e^{-x}}{2i} \\ &= \frac{e^x - e^{-x}}{2}. \end{aligned}$$

We define  $\sinh x = -i \sin(ix)$ .

Similar to how  $\cos^2 x + \sin^2 x = 1$ , we can find that  $\cosh^2 x - \sinh^2 x = 1$ .

## Index Algebra

We usually denote vectors by either  $\vec{A}$ ,  $\mathbf{A}$ , or

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix},$$

which is defined by a basis.

If we imagine we are in  $n$ -dimensional space, we can let  $A_i$  where  $i = 1, 2, \dots, n$  denote both

- the  $i$ th component of  $\vec{A}$ ;
- the entire vector  $\vec{A}$  (since  $i$  can be arbitrary).

## Contractions and Dummy Indices

Consider  $C = AB$ , where  $A, B$  are  $n \times m$  and  $m \times p$  matrices respectively.

$$C = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ B_{m1} & B_{m2} & \cdots & B_{mp} \end{pmatrix}.$$

**Definition** (Matrix Multiplication in Index Notation). For matrices  $A$  and  $B$ , where  $A$  is an  $m \times n$  and  $B$  is a  $n \times p$  matrix, we write

$$C_{ij} = \sum_{k=1}^n A_{ik} B_{kj}$$

We say that  $k$  is a dummy index, since  $k$  takes values from 1 to  $n$ . Note that the value we calculate is  $C_{ij}$ ; in other words, in the sum  $\sum_k A_{ik} B_{kj}$ , the indices of the form  $ij$  are the “net indices” from the multiplication.

Note that if  $C = BA$ , then

$$\begin{aligned} C_{ij} &= \sum_{k=1}^n B_{ik} A_{kj} \\ &= \sum_{k=1}^n A_{kj} B_{ik} \end{aligned}$$

$$\neq \sum_{k=1}^n A_{ik} B_{kj}.$$

The corresponding fact is that  $AB \neq BA$  necessarily.

Note that the index that is summed over always appears exactly twice.

**Definition** (Symmetric Matrix). Let  $C$  be a matrix. Then, we say  $C$  is symmetric if

$$C_{ij} = C_{ji}$$

**Definition** (Antisymmetric Matrix). Let  $C$  be a matrix. We say  $C$  is antisymmetric if

$$C_{ij} = -C_{ji}.$$

We can always decompose a random matrix into the sum of a symmetric matrix and an antisymmetric matrix.

## Two Special Tensors

**Definition** (Kronecker Delta). The Kronecker Delta,  $\delta_{ij}$ , is the tensor that denotes the identity matrix.

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

**Example** (Extracting an Index). Consider  $A$  as vector. Then,

$$\sum_i A_i \delta_{ij} = A_j.$$

In other words, the Kronecker Delta collapses the sum to the  $j$ th index.

**Example** (Orthonormal Basis from Kronecker Delta). Let  $\{\hat{e}_i\}_{i=1}^n$  be a basis for some vector space  $V$ . If

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij}$$

for every  $i, j$ , then  $\{\hat{e}_i\}_{i=1}^n$  is an orthonormal basis for  $V$ .

**Definition** (Levi-Civita Symbol). In two dimensions, as a matrix, we write

$$\epsilon_{ij} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

meaning

$$\epsilon_{ij} = \begin{cases} 1 & i = 1, j = 2 \\ -1 & i = 2, j = 1 \\ 0 & \text{else} \end{cases}.$$

The Levi-Civita Symbol is antisymmetric, just as the Kronecker Delta is symmetric.

In three dimensions, we define

$$\epsilon_{ijk} = \begin{cases} 1 & (i, j, k) = (1, 2, 3) \text{ cyclically} \\ -1 & (i, j, k) = (2, 1, 3) \text{ cyclically} \\ 0 & \text{else} \end{cases}.$$

In other words,  $\epsilon_{ijk} = -\epsilon_{jik}$ .

**Exercise** (Relations between  $\delta_{ij}$  and  $\epsilon_{ijk}$ ):

$$\sum_{j,k} \epsilon_{mjk} \epsilon_{njk} = 2\delta_{mn}$$

$$\sum_{\ell} \epsilon_{mn\ell} \epsilon_{ij\ell} = \delta_{mi} \delta_{nj} - \delta_{mj} \delta_{ni}$$

**Definition** (Dot Product). Let  $\{\hat{e}_i\}_{i=1}^n$  be an orthonormal basis for  $V$ . Let  $\mathbf{A} = \sum_i A_i \hat{e}_i$  and  $\mathbf{B} = \sum_i B_i \hat{e}_i$ . Then,

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} &= \sum_{i,j} (A_i \hat{e}_i) \cdot (B_j \hat{e}_j) \\ &= \sum_{i,j} A_i B_j (\hat{e}_i \cdot \hat{e}_j) \\ &= \sum_{i,j} A_i B_j \delta_{ij} \\ &= \sum_i A_i B_i \end{aligned}$$

**Definition** (Cross Product). Let  $\{\hat{e}_i\}_{i=1}^3$  be the standard basis over  $\mathbb{R}^3$ . Let  $\mathbf{A} = \sum_i A_i \hat{e}_i$  and  $\mathbf{B} = \sum_i B_i \hat{e}_i$ . Then,

$$\begin{aligned} \mathbf{A} \times \mathbf{B} &= \sum_{i,j} (A_i \hat{e}_i) \times (B_j \hat{e}_j) \\ &= \sum_{i,j} A_i B_j (\hat{e}_i \times \hat{e}_j) \\ &= \sum_{i,j,k} A_i B_j (\epsilon_{ijk} \hat{e}_k). \end{aligned}$$

Instead of asking about  $\mathbf{A} \times \mathbf{B}$ , we ask about  $(\mathbf{A} \times \mathbf{B})_\ell$ , yielding

$$\begin{aligned} (\mathbf{A} \times \mathbf{B})_\ell &= (\mathbf{A} \times \mathbf{B}) \cdot \hat{e}_\ell \\ &= \left( \sum_{i,j,k} A_i B_j (\epsilon_{ijk} \hat{e}_k) \right) \cdot \hat{e}_\ell \\ &= \sum_{i,j} \epsilon_{ij\ell} A_i B_j. \end{aligned}$$

**Remark:** This notation for  $\mathbf{A} \times \mathbf{B}$  automatically shows us that

$$\begin{aligned} (\mathbf{B} \times \mathbf{A})_\ell &= \sum_{i,j} \epsilon_{ij\ell} B_i A_j \\ &= - \sum_{i,j} \epsilon_{ji\ell} B_i A_j \\ &= - \sum_{i,j} \epsilon_{jil} A_j B_i \\ &= - \sum_{i,j} \epsilon_{ij\ell} A_i B_j \\ &= -(\mathbf{A} \times \mathbf{B})_\ell. \end{aligned}$$

$i, j$  are dummy indices

**Example** (Central Force and Angular Momentum). A central force is defined by

$$\mathbf{F} = f(r) \hat{r},$$

where  $\hat{r}$  is a radial vector.

Angular momentum is defined by

$$\mathbf{L} = \mathbf{r} \times \mathbf{p},$$

where  $\mathbf{r}$  denotes position and  $\mathbf{p}$  denotes momentum. Then,

$$\begin{aligned} \frac{d\mathbf{L}}{dt} &= \frac{d}{dt} (\mathbf{r} \times \mathbf{p}) \\ &= \left( \frac{d}{dt} \mathbf{r} \times \mathbf{p} \right) + \mathbf{r} \times \left( \frac{d\mathbf{p}}{dt} \right) \\ &= m \left( \frac{d}{dt} \mathbf{r} \times \frac{d}{dt} \mathbf{r} \right) + \mathbf{r} \times (f(r)\hat{r}) \\ &= f(r) (\mathbf{r} \times \hat{r}). \end{aligned}$$

This implies that  $\frac{d\mathbf{L}}{dt} = 0$  under a central force.

**Example (Determinant).** Let  $\mathbf{M} = M_{ij}$  be square. We denote  $\mathbf{M}_i$  to be the vector denoting the  $i$ th-row. Then,

$$\begin{aligned} m &= |\mathbf{M}| \\ &= \mathbf{M}_1 \cdot (\mathbf{M}_2 \times \mathbf{M}_3) \\ &= \mathbf{M}_3 \cdot (\mathbf{M}_1 \times \mathbf{M}_2) \\ &= \mathbf{M}_2 \cdot (\mathbf{M}_3 \times \mathbf{M}_1). \end{aligned}$$

**Example (Trace).** Let  $\mathbf{M} = M_{ij}$  be a square matrix. We define  $\text{tr}(\mathbf{M}) = \sum_i M_{ii}$ . Equivalently,

$$\begin{aligned} \text{tr}(\mathbf{M}) &= \sum_{ij} M_{ij} \delta_{ij} \\ &= \sum_i M_{ii}. \end{aligned}$$

Note that

$$\begin{aligned} \text{tr}(\mathbf{I}_n) &= \sum_i \delta_{ii} \\ &= n. \end{aligned}$$

When we upgrade to 3 matrices, we take

$$\begin{aligned} \text{tr}(ABC) &= \sum_{i,j} \left( \sum_{k,\ell} A_{ik} B_{k\ell} C_{\ell j} \right) \delta_{ij} \\ &= \sum_{i,k,\ell} A_{ik} B_{k\ell} C_{\ell i} \\ &= \sum_{i,k,\ell} C_{\ell i} A_{ik} B_{k\ell} \\ &= \text{tr}(CAB). \end{aligned}$$

In other words, the trace is invariant under cyclic permutations.

**Example (Moment of Inertia Tensor).**

Recall that

$$\mathbf{L} = \mathbf{r} \times \mathbf{p},$$

$$= I\omega.$$

where  $\mathbf{p} = m\dot{\mathbf{x}}$ , and  $I$  denotes the moment of inertia. Note that  $I \sim mr^2$ . On a more fundamental level, it is the case that the first equation,  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ , is the “true” definition of  $\mathbf{L}$ .

Consider a small portion  $m_\alpha$  about some axis at radius  $\mathbf{r}_\alpha$  and momentum  $\mathbf{p}_\alpha$ . Then, we have

$$\begin{aligned} \mathbf{L}_\alpha &= \sum_\alpha \mathbf{r}_\alpha \times \mathbf{p}_\alpha \\ &= \sum_\alpha m_\alpha (\mathbf{r}_\alpha \times (\boldsymbol{\omega} \times \mathbf{r}_\alpha)). \end{aligned}$$

In the infinitesimal case (i.e., as  $\alpha \rightarrow 0$ ), we get

$$\mathbf{L} = \int \mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r}) \rho \, d\tau,$$

where  $\rho$  denotes volume density. Applying the identity  $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ , we find

$$\mathbf{L} = \int (\boldsymbol{\omega}(\mathbf{r} \cdot \mathbf{r}) - \mathbf{r}(\mathbf{r} \cdot \boldsymbol{\omega})) \rho \, d\tau.$$

Switching to index notation, we have

$$\begin{aligned} L_i &= \int \left( \omega_i r^2 - r_i \sum_j r_j \omega_j \right) \rho \, d\tau \\ &= \sum_j \int \omega_j \left( \delta_{ij} r^2 - r_i r_j \right) \rho \, d\tau \\ &= \sum_j \omega_j \underbrace{\left( \int \left( \delta_{ij} r^2 - r_i r_j \right) \rho \, d\tau \right)}_{\text{moment of inertia tensor}} \\ &= \sum_j I_{ij} \omega_j. \end{aligned}$$

## Binomial Theorem

The binomial theorem allows us to calculate the expansion

$$(x + y)^n = \sum_{m=0}^n \binom{n}{m} x^{n-m} y^m.$$

In the case of  $(x + y)^2 = x^2 y^0 + 2x^1 y^1 + x^0 y^2 = x^2 + 2xy + y^2$ . Recall that

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}.$$

Recall that  $0! = 1$ .

## Infinite Series

Let

$$S = \sum_{k=0}^{\infty} a_k$$

be an infinite series. We are often curious as to the convergence of this sum (for a variety of reasons). Formally, we have to invoke partial sums

$$S_N = \sum_{k=0}^N a_k,$$

and see if the sequence of partial sums is convergent. However, we will prefer to use series convergence tests.

**Example (Geometric Series).** Let

$$\begin{aligned} S &= \sum_{k=0}^{\infty} r^k \\ &= 1 + r + r^2 + \dots \end{aligned}$$

Then, we have

$$\begin{aligned} S_N &= \sum_{k=0}^N r^k \\ rS_N &= \sum_{k=0}^N r^{k+1}. \end{aligned}$$

Subtracting, we get

$$\begin{aligned} (1-r)S_N &= 1 - r^{N+1} \\ S_N &= \frac{1 - r^{N+1}}{1 - r}. \end{aligned}$$

In the limit, we expect that if  $r \rightarrow \infty$ , and  $r < 1$ , then  $r^{N+1} \rightarrow 0$ . In the infinite case, we have

$$\begin{aligned} S &= \sum_{k=0}^{\infty} r^k \\ &= \frac{1}{1-r}, \end{aligned}$$

if  $r < 1$ .

There are a few prerequisites for series convergence:

- there exists some  $K$  for which for all  $k \geq K$ ,  $a_{k+1} \leq a_k$ ;
- $\lim_{k \rightarrow \infty} a_k < \infty$ ;
- we need the series to reduce “quickly” enough.

**Example (Ratio Test).** A series  $S = \sum_k a_k$  converges if the ratio of consecutive terms is (eventually) less than 1:

$$r = \lim_{k \rightarrow \infty} \frac{a_{k+1}}{a_k} < 1.$$

**Example (Applying the Ratio Test).** Consider  $S = \sum_k \frac{1}{k!}$ . Then,

$$\begin{aligned} r &= \lim_{k \rightarrow \infty} \frac{\frac{1}{(k+1)!}}{\frac{1}{k!}} \\ &= \lim_{k \rightarrow \infty} \frac{1}{k+1} \\ &= 0 < 1 \end{aligned}$$

**Example** (Riemann Zeta Function). We write

$$\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}.$$

In order to evaluate the convergence of the Riemann zeta function. We have

$$\begin{aligned} r &= \lim_{k \rightarrow \infty} \frac{\frac{1}{(k+1)^s}}{\frac{1}{k^s}} \\ &= \lim_{k \rightarrow \infty} \left( \frac{k}{k+1} \right)^s \\ &= 1. \end{aligned}$$

Unfortunately, this means the ratio test is inconclusive.

For examples of evaluations of the zeta function, we have

$$\begin{aligned} \zeta(1) &= 1 + \frac{1}{2} + \frac{1}{3} + \cdots \\ \zeta(2) &= 1 + \frac{1}{4} + \frac{1}{9} + \cdots \\ &= \frac{\pi^2}{6}. \end{aligned}$$

**Example** (Absolute Convergence). In our original ratio test, we had assumed that  $a_k$  are real and positive. However, if the  $a_k \in \mathbb{C}$ , we have to look at the convergence in modulus:

$$r = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right|.$$

If  $\sum_k |a_k|$  converges, this is known as absolute convergence.

**Example** (Alternating Series Test). If the series

$$\sum_{k=0}^{\infty} (-1)^k a_k$$

has the following conditions:

- $a_{k+1} < a_k$  for  $k > K$ ;
- $\lim_{k \rightarrow \infty} a_k = 0$ ;

then  $\sum_k (-1)^k a_k$  converges.

For instance, the alternating harmonic series converges

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} = \ln 2.$$

## Power Series

Consider the function

$$S(x) = \sum_{k=0}^{\infty} a_k x^k.$$



This is a series both in  $a_k$  and in  $x$ . In order to determine convergence, we use the ratio test as follows:

$$\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}x^{k+1}}{a_k x^k} \right| = |x| \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| \\ \equiv |x| r.$$

In particular, for convergence, it must be the case that

$$|x| r < 1.$$

We define

$$R = \begin{cases} \frac{1}{r} & 0 < r < \infty \\ 0 & r = \infty \\ \infty & r = 0 \end{cases}.$$

In particular, this means

$$|x| < R.$$

**Definition** (Radius of Convergence). For a power series  $\sum_k a_k x^k$ , the series converges for  $|x| < R$ ,<sup>v</sup> where

$$r = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| \\ R = \begin{cases} \frac{1}{r} & 0 < r < \infty \\ 0 & r = \infty \\ \infty & r = 0 \end{cases}.$$

Note that convergence for  $|x| < R$  does not provide information regarding convergence at the boundary.

**Example** (Geometric Series). We have

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$$

has  $R = 1$ , meaning the power series converges for  $|x| < 1$ .

**Example** (Exponential Function). We have

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

with  $R = \infty$ .

**Example** (Natural Log). We have

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \dots$$

In particular, since  $R = 1$ , we know that the radius of convergence is  $|x| < 1$ . However, the series does converge on the boundary when  $x = 2$ , but not when  $x = 0$  (for obvious reasons).

---

<sup>v</sup>The definition is not the true radius of convergence; it is actually that  $r = \limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|}$ . It just happens to be the case that the ratio test and root test return the same value when they're regular limits (rather than limits superior).

**Example** (Why Radius of Convergence?). Consider two series

$$\frac{1}{1-x^2} = \sum_{k=0}^{\infty} x^{2k}$$

$$\frac{1}{1+x^2} = \sum_{k=0}^{\infty} (-1)^k x^{2k}.$$

We can see that the first series converges for  $|x| < 1$ . However, even though  $\frac{1}{1+x^2}$  has a domain across the entire real numbers, it is still the case that the *series* converges for  $|x| < 1$ .

The primary reason that the radius of convergence is defined as such is because, over the complex numbers, it is the case that  $x^2 + 1 = 0$  at  $x = \pm i$ , meaning  $\frac{1}{1+z^2}$  has singularities at those values of  $z$ .

The main reason power series are useful is that, when truncated, they are simply polynomials. In particular, with power series, we can reverse the order of sum and derivative.

### Taylor Series

**Definition.** The Taylor series of a function  $f(x)$  about  $x_0$  is defined by

$$f(x) = \sum_{n=0}^{\infty} \frac{(x-x_0)^n}{n!} \left( \left. \frac{d^n f}{dx^n} \right|_{x=x_0} \right).$$

**Remark:** The reason we write  $\frac{d^n f}{dx^n}$  is because  $\frac{d^n}{dx^n}$  is an operator in and of itself.

**Example** (The Most Important Taylor Series).

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$$

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^{n+1} x^{2n+1}}{(2n+1)!}$$

**Example** (Equilibrium Points). Let  $U(x)$  denote a potential over  $x$ . Then,  $F = -\nabla U$ . We have

$$U(x) = U(x_0) + (x-x_0) U'(x_0) + \frac{1}{2!} (x-x_0)^2 U''(x_0) + \frac{1}{3!} (x-x_0)^3 U'''(x_0) + \dots$$

When we analyze an equilibrium point, we disregard the  $U(x_0)$  term, and see that the derivative of  $U$  is zero; thus, we can truncate our series at the second derivative close to  $x = x_0$ :

$$U(x) \approx \frac{1}{2} U''(x_0) (x-x_0)^2$$

$$= \frac{1}{2} m \omega^2 (x-x_0)^2.$$

In other words, when we are very close to equilibrium, we have simple harmonic motion.

**Example** (Faster Taylor Series). Consider the function

$$\exp\left(\frac{x}{1-x}\right).$$

In order to create its Taylor series, we can create this Taylor series piecewise:

$$\exp\left(\frac{x}{1-x}\right) = 1 + \left(\frac{x}{1-x}\right) + \frac{1}{2!}\left(\frac{x}{1-x}\right)^2 + \frac{1}{3!}\left(\frac{x}{1-x}\right)^3.$$

Now, we expand the denominators as geometric series:

$$= 1 + x \left( \sum_{k=0}^{\infty} x^k \right) + \frac{x^2}{2!} \left( \sum_{k=0}^{\infty} x^k \right)^2 + \frac{x^3}{3!} \left( \sum_{k=0}^{\infty} x^k \right) + \dots$$

If we want to expand through  $x^3$ , we have to expand by keeping track of *every* term:

$$= 1 + x + \frac{3}{2}x^2 + \frac{13}{6}x^3 + O(x^4).$$

We say we have expanded the series through the third order; the lowest order correction, denoted  $O(x^n)$ , is the fourth order (in this case).

**Example** (Exponentiated Operator). Consider a (square) matrix  $M$ . Then, we define

$$e^M = \sum_{k=0}^{\infty} \frac{M^k}{k!},$$

where  $M^k = \prod_{i=1}^k M$ ; we define  $M^0 = I$ . Similarly,

$$e^{\frac{d}{dx}} = \sum_{k=0}^{\infty} \frac{d^k}{dx^k} \frac{1}{k!}.$$

In particular,  $e^{\frac{d}{dx}}$  is the Taylor series operator.

**Remark:** In quantum mechanics, the momentum operator is

$$P = -i\hbar \frac{d}{dx}.$$

**Example** (Binomial Expansion). For any  $\alpha \in \mathbb{C}$  and  $|x| < 1$ , we have

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \dots$$

Note that if  $\alpha \in \mathbb{Z}^+$ , then the series truncates (and we recover the binomial theorem again).

The main use of the binomial expansion is with very small quantities. For instance,

$$\begin{aligned} E &\sim \frac{1}{(x^2 + a^2)^{3/2}} \\ &= \frac{1}{x^3 \left(1 + \frac{a^2}{x^2}\right)^{3/2}} \\ &\approx \frac{1}{x^3} \left(1 - \frac{3}{2} \frac{a^2}{x^2}\right) \end{aligned} \quad \text{For } x \gg a$$

**Remark:** The binomial expansion only applies to the form  $(1+x)^\alpha$ . If we are dealing with an expression of the form  $(a+x)^\alpha$ , we need to factor out  $a$ , making the expression  $a^\alpha (1+x/a)^\alpha$ .

**Example** (Special Relativity with the Binomial Expansion). In the theory of special relativity, Einstein came up with the equations

$$E = \gamma mc^2$$

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

We can use the binomial expansion to find more information about  $\gamma$ .

$$\begin{aligned} E &= \left(1 - \frac{v^2}{c^2}\right)^{-1/2} mc^2 \\ &= \left(1 + \frac{1}{2} \frac{v^2}{c^2} + \frac{(-\frac{1}{2})(-\frac{3}{2})}{2!} \left(-\frac{v^2}{c^2}\right)^2 + \dots\right) mc^2 \\ &= mc^2 + \underbrace{\frac{1}{2} mv^2 \left(1 + \frac{3}{4} \frac{v^2}{c^2} + \frac{5}{8} \left(\frac{v^2}{c^2}\right)^2 + \dots\right)}_{\text{Kinetic Energy}} \end{aligned}$$

As we take  $v \ll c$ , we only need to keep the first order term in the expansion, meaning we have  $E = mc^2 + \frac{1}{2}mv^2$ .

Thus, we can find kinetic energy as  $KE = (\gamma - 1) mc^2$ . Notice that this means that *most* energy is internal energy emergent as mass.

## Ten Integration Techniques

While Mathematica may exist,<sup>vi</sup> it is still valuable to know how to take various integrals. More importantly, knowing how to take integrals provides valuable insights into *what* exactly integrals are.

### Integration by Parts

**Definition** (Integration by Parts). Using the product rule, we have

$$\begin{aligned} \int \frac{d}{dx} (uv) \, dx &= \int \frac{du}{dx} v - \frac{dv}{dx} u \, dx \\ &= \int \frac{du}{dx} v \, dx - \int \frac{dv}{dx} u \, dx. \end{aligned}$$

Thus, we get

$$\int u \, dv = uv - \int v \, du.$$

In the case where our integrals are definite, we have

$$\int_a^b u \, dv = uv \Big|_a^b - \int_a^b v \, du.$$

We say  $uv \Big|_a^b$  is the boundary term (or surface term).<sup>vii</sup>

<sup>vi</sup>Citation needed.

<sup>vii</sup>We can also use integration by parts to define the (weak) derivative, assuming the boundary term is zero.

**Example.**

$$\begin{aligned}\int x e^{ax} dx &= \frac{1}{a} x e^{ax} - \int \frac{1}{a} e^{ax} dx & u = x, dv = e^{ax} dx \\ &= \frac{1}{a} x e^{ax} - \frac{1}{a^2} e^{ax} \\ &= \frac{1}{a^2} e^{ax} (ax - 1).\end{aligned}$$

The +C is implicit.

**Example.**

$$\begin{aligned}\int \ln x dx &= x \ln x - \int x \left( \frac{1}{x} \right) dx & u = \ln x, dv = dx \\ &= x \ln x - x.\end{aligned}$$

### Change of Variables

**Definition** (u-Substitution). Let  $x = x(u)$ , meaning  $dx = \frac{dx}{du} du$ . Thus, we get

$$\int_{x_1}^{x_2} f(x) du = \int_{u(x_1)}^{u(x_2)} f(x(u)) \frac{dx}{du} du.$$

**Example.**

$$\begin{aligned}I_1 &= \int_0^\infty x e^{-ax^2} dx \\ &= \frac{1}{2} \int_0^\infty e^{-au} du & u = x^2 \\ &= \frac{1}{2a}\end{aligned}$$

**Example.**

$$\begin{aligned}\int_0^\pi \sin \theta d\theta &= \int_{-1}^1 du & u = \cos \theta \\ &= 2.\end{aligned}$$

More generally, we have, for  $f(\theta) = f(\cos \theta)$ ,

$$\int_0^\pi f(\theta) \sin \theta d\theta = \int_{-1}^1 f(u) du.$$

**Example** (Trig Substitution).

$$\begin{aligned}\int_0^a \frac{x}{x^2 + a^2} dx &= \int_0^{\pi/4} \frac{a^2 \tan \theta \sec^2 \theta}{a^2 (1 + \tan^2 \theta)} d\theta & x = a \tan \theta \\ &= \int_0^{\pi/4} \tan \theta d\theta \\ &= -\ln(\cos \theta) \Big|_0^{\pi/4} \\ &= \ln(\sqrt{2}) \\ &= \frac{1}{2} \ln(2).\end{aligned}$$

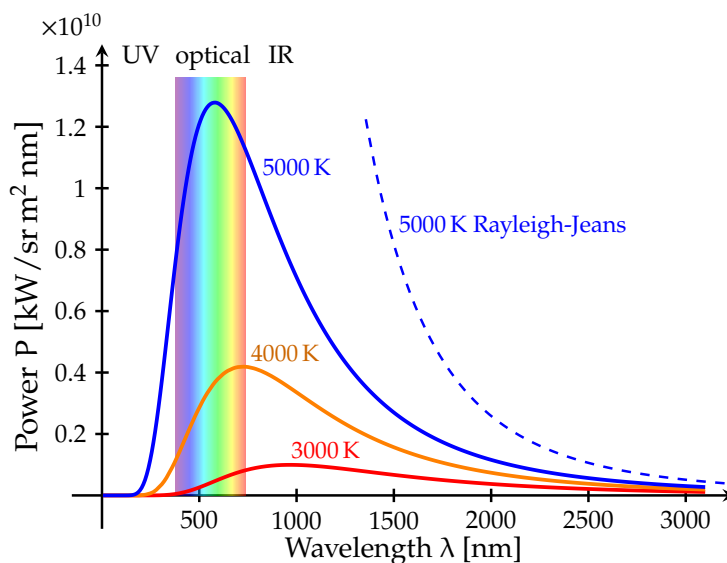
**Example** (Trig Substitution 2.0). For rational functions of  $\sin \theta$  and  $\cos \theta$ , we can use the half-angle trig substitution  $u = \tan(\theta/2)$ .<sup>viii</sup> This yields

$$\begin{aligned} d\theta &= \frac{2du}{1+u^2} \\ \sin \theta &= \frac{2u}{1+u^2} \\ \cos \theta &= \frac{1-u^2}{1+u^2}. \end{aligned}$$

For instance,

$$\begin{aligned} \int \frac{1}{1+\cos \theta} d\theta &= \int \frac{1}{1+\frac{1-u^2}{1+u^2}} \frac{2}{1+u^2} du \\ &= \int du \\ &= \tan(\theta/2) \\ &= \frac{\sin \theta}{1+\cos \theta}. \end{aligned}$$

**Example** (Dimensionless Integrals).



Anything that has a nonzero absolute temperature radiates some energy. In particular, we want to know how this radiation is distributed among various wavelengths.

For a box of photons in equilibrium at temperature  $T$ , the energy per volume per wavelength  $\lambda^{\text{ix}}$  is

$$u(\lambda) = \frac{8\pi hc}{\lambda^5 (e^{hc/\lambda kT} - 1)}.$$

Here,  $h$  denotes Planck's constant,  $c$  is the speed of light, and  $k$  is Boltzmann's constant.

<sup>viii</sup> $\tan(\theta/2) = \frac{\sin \theta}{1+\cos \theta}$

<sup>ix</sup>read as (energy per volume) per wavelength

In order to find the total energy density, we have to integrate  $u(\lambda)$  over all possible values of  $\lambda$ :

$$\begin{aligned} U &= \int_0^\infty u(\lambda) d\lambda \\ &= 8\pi hc \int_0^\infty \frac{1}{\lambda^5 (e^{hc/\lambda kT} - 1)} d\lambda \end{aligned}$$

This integral is, for lack of a better word, hard. However, if we remove the dimensions of  $\lambda$  by substituting  $x = \frac{hc}{\lambda kT}$ , we can verify that the value of  $U$  now becomes

$$U = 8\pi hc \left( \frac{kT}{hc} \right)^4 \underbrace{\int_0^\infty \frac{x^3}{e^x - 1} dx}_{\text{scalar}}.$$

Thus, all the physics<sup>x</sup> is captured as a coefficient on the integral; namely, this integral captures the Stefan-Boltzmann law, which has that energy density scales by  $T^4$ .

Using some fancy techniques we will learn later, we can evaluate

$$\int_0^\infty \frac{x^3}{e^x - 1} dx = \frac{\pi^4}{15}.$$

## Even/Odd

**Definition** (Even and Odd Functions). A function  $f(x)$  is

- even if  $f(-x) = f(x)$ ;
- odd if  $f(-x) = -f(x)$ .

Just as a matrix can be decomposed into a sum of a symmetric and antisymmetric matrix, we can decompose a function into a sum of an even function and an odd function.

Integrals over symmetric intervals on functions with definite parity are very simple:

$$\int_{-a}^a f(x) dx = \begin{cases} 2 \int_0^a f(x) dx & f \text{ odd} \\ 0 & f \text{ even} \end{cases}.$$

For the case of a function  $g(x) = g(|x|)$ , we have

$$\int_{-a}^b g(|x|) dx = \int_{-a}^0 g(-x) dx + \int_0^b g(x) dx.$$

## Products and Powers of Sines and Cosines

**Example.** If we have an integral

$$\begin{aligned} \int \sin(3x) \cos(2x) dx &= \frac{1}{2} \int \sin(5x) + \sin(x) dx \\ &= \frac{1}{2} \left( -\frac{1}{5} \cos(5x) - \cos(x) \right). \end{aligned}$$

---

<sup>x</sup>Who cares about that stuff?

**Example.** To evaluate

$$\int \sin^2(x) dx,$$

$$\int \cos^2(x) dx$$

we use the identity

$$\sin^2(x) = \frac{1}{2} (1 - \cos(2x))$$

$$\cos^2(x) = \frac{1}{2} (1 + \cos(2x)),$$

and take

$$\begin{aligned} \int \sin^2(x) dx &= \frac{1}{2} \int (1 - \cos(2x)) dx \\ &= \frac{x}{2} - \frac{1}{4} \sin(2x) \\ \int \cos^2(x) dx &= \frac{1}{2} \int (1 + \cos(2x)) dx \\ &= \frac{x}{2} + \frac{1}{4} \sin(2x). \end{aligned}$$

Thus, we can see that

$$\int_0^\pi \sin^2(x) dx = \frac{\pi}{2}$$

$$\int_0^\pi \cos^2(x) dx = \frac{\pi}{2}$$

### Axial and Spherical Symmetry

Consider a function of the form  $f(x, y) = x^2 + y^2$ . If we were to integrate with respect to  $dx dy$ , we would need a two dimensional integral. With polar coordinates, though, we would have  $dx dy = r dr d\phi$ . Since  $f$  is axially symmetric, we would have our  $dx dy = 2\pi r dr$ , which is a one-dimensional integral.

If we have something with spherical symmetry, then there is no dependence on either  $\theta$  or  $\phi$ , yielding a function  $f(\mathbf{r}) = f(r)$ , meaning

$$\begin{aligned} \int f(\mathbf{r}) d\tau &= \int f(r) r^2 \sin \theta dr d\theta d\phi \\ &= 4\pi \int f(r) r^2 dr. \end{aligned}$$

Note that  $\int \sin \theta d\theta d\phi$  over the sphere is  $4\pi$ .

**Example.** Consider a surface  $S$  with charge density  $\sigma(\mathbf{r})$ . Finding the total charge requires evaluating

$$Q = \int_S \sigma(\mathbf{r}) dA.$$

If  $S$  is hemispherical with  $z > 0$  with radius  $R$ , and  $\sigma = k \frac{x^2 + y^2}{R^2}$ , the integrand is axially symmetric.



Using spherical coordinates, we evaluate

$$\begin{aligned}
 Q &= \int_S \sigma(\mathbf{r}) \, dA \\
 &= \frac{k}{R^2} \int x^2 + y^2 \, dA \\
 &= \frac{k}{R^2} \int \left( R^2 \sin^2 \theta \cos^2 \phi + R^2 \sin^2 \theta \sin^2 \phi \right) R^2 \sin \theta \, d\theta d\phi \\
 &= kR^2 \int_S \sin^3 \theta \, d\theta d\phi \\
 &= 2\pi kR^2 \int_0^{\pi/2} \sin^3 \theta \, d\theta \\
 &= \frac{4\pi kR^2}{3}.
 \end{aligned}$$

**Example.** Let

$$\Phi(\mathbf{r}) = \int \frac{e^{-i\mathbf{k} \cdot \mathbf{r}}}{(2\pi)^3 \|\mathbf{k}\|^2} \, d^3k$$

where  $k$ -space is an abstract 3-dimensional Euclidean space. In Cartesian coordinates,  $d^3k = dk_x dk_y dk_z$ , which yields the integral

$$\Phi(\mathbf{r}) = \int \frac{e^{-ik_x x} e^{-ik_y y} e^{-ik_z z}}{(2\pi)^3 (k_x^2 + k_y^2 + k_z^2)} \, dk_x dk_y dk_z.$$

This integral is very hard to evaluate (over Cartesian coordinates, anyway),<sup>XI</sup> so we need to use some other methods.

In spherical coordinates, we have  $d^3k = k^2 dk d\Omega$ , yielding

$$\Phi(\mathbf{r}) = \frac{1}{(2\pi)^3} \int k^2 \frac{e^{-ikr \cos \theta}}{k^2} \, dk d(\cos \theta) d\phi.$$

Since we are summing away all our  $k$ -dependence, we can orient  $\mathbf{r}$  along the  $k_z$  axis. Thus, we can evaluate the integral as

$$\begin{aligned}
 \Phi(\mathbf{r}) &= \frac{1}{(2\pi)^3} \int k^2 \frac{e^{-ikr \cos \theta}}{k^2} \, dk d(\cos \theta) d\phi \\
 &= \frac{1}{(2\pi)^2} \int_{-1}^1 \int_0^\infty e^{-ikr \cos \theta} \, dk d(\cos \theta) \\
 &= \frac{1}{(2\pi)^2} \int \frac{1}{(-ikr)} \left( e^{-ikr} - e^{ikr} \right) \, dk \\
 &= \frac{1}{(2\pi)^2} \int_0^\infty \frac{2 \sin(kr)}{kr} \, dk \\
 &= \frac{1}{2\pi^2} \underbrace{\int_0^\infty \frac{\sin(kr)}{kr} \, dk}_{\text{sinc integral}}.
 \end{aligned}$$

In order to evaluate the sinc integral, we have to use some different techniques.

<sup>XI</sup>Citation needed.

## Differentiation with Respect to a Parameter

**Example.** We can evaluate

$$\begin{aligned}\int x e^{ax} dx &= \frac{\partial}{\partial a} \left( \int e^{ax} dx \right) \\ &= \frac{\partial}{\partial a} \left( \frac{1}{a} e^{ax} \right) \\ &= -\frac{1}{a^2} e^{ax} + \frac{1}{a} x e^{ax} \\ &= \frac{1}{a^2} e^{ax} (ax - 1)\end{aligned}$$

When differentiating with respect to a parameter, it is important to remember that we are often differentiating *with respect to the parameter*, not with respect to our main variable.

**Example (Introducing a Parameter).** We wish to solve the sinc integral,

$$\int_0^\infty \frac{\sin x}{x} dx.$$

In order to do this, we will introduce a parameter such that differentiation will cancel out the  $x$  in the denominator:

$$J(\alpha) = \int_0^\infty e^{-\alpha x} \frac{\sin x}{x} dx. \quad \alpha > 0$$

In particular,  $\alpha > 0$ . We calculate

$$\begin{aligned}\frac{dJ}{d\alpha} &= - \int_0^\infty e^{-\alpha x} \sin x dx \\ &= -\frac{1}{1 + \alpha^2}.\end{aligned}$$

Therefore,

$$\begin{aligned}J(\alpha) &= - \int \frac{1}{\alpha^2} d\alpha \\ &= -\arctan(\alpha) + C.\end{aligned}$$

In order to determine the value of  $C$ , we need to make sure  $J(\infty) = 0$ . Therefore,  $C = \frac{\pi}{2}$ . Therefore, we have

$$J(0) = \frac{\pi}{2}.$$

## Gaussian Integral

We cannot evaluate  $I_0 = \int_0^\infty e^{-ax^2} dx$  using elementary methods, because  $e^{-ax^2}$  is not an elementary function. The reason we care a lot about  $e^{-ax^2}$  is because it is very important in quantum mechanics and statistics.<sup>xii</sup>

It is clear that  $I_0$  converges. We can see that the dimension of  $a$  is  $x^{-2}$ , and since we are integrating with respect to  $dx$ , we can see that our integral is related to  $\frac{1}{\sqrt{a}}$ .

---

<sup>xii</sup>Who cares about that?

**Example.** We will not solve for  $I_0$ , but for  $I_0^2$ . Thus, we have

$$\begin{aligned}
 I_0^2 &= \left( \frac{1}{2} \int_{-\infty}^{\infty} e^{-ax^2} dx \right) \left( \frac{1}{2} \int_{-\infty}^{\infty} e^{-ay^2} dy \right) \\
 &= \frac{1}{4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-a(x^2+y^2)} dx dy \\
 &= \frac{1}{4} \int_0^{2\pi} \int_0^{\infty} r e^{-ar^2} dr d\phi \\
 &= \frac{\pi}{2} \int_0^{\infty} r e^{-ar^2} dr \\
 &= \frac{\pi}{2} \left( \frac{1}{2} \int_0^{\infty} e^{-au} du \right) \\
 &= \frac{\pi}{4a}.
 \end{aligned}$$

Therefore,  $I_0 = \frac{1}{2} \sqrt{\frac{\pi}{a}}$ .

**Definition** (Family of Gaussian Integrals).

$$I_n = \int_0^{\infty} x^n e^{-ax^2} dx.$$

It is important to note that there are different expressions for the Gaussian integral:

$$\begin{aligned}
 &\int e^{-ax^2} dx \\
 &\int e^{-a^2x^2} dx \\
 &\int e^{-a^2x^2/2} dx \\
 &\int e^{-x^2/a} dx \\
 &\int e^{-x^2/a^2} dx,
 \end{aligned}$$

meaning we have to be careful when evaluating these integrals.

**Example** (Error Function). Consider the integral

$$\int_0^{53} e^{-ax^2} dx.$$

Unfortunately, there is no way to do this integral analytically. It is only able to be calculated numerically.

We define

$$\text{erf}(u) = \int_0^u e^{-ax^2} dx$$

### Completing the Square

**Example.** Consider the integral

$$\int_{-\infty}^{\infty} e^{-ax^2-bx} dx.$$

This integral is Gaussian-esque, but it isn't fully Gaussian, yet.

To do this, we will complete the square:

$$\begin{aligned} ax^2 + bx &= a \left( x^2 + \frac{b}{a}x \right) \\ &= a \left( x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} - \frac{b^2}{4a^2} \right) \\ &= a \left( x + \frac{b}{2a} \right)^2 - \frac{b^2}{4a}. \end{aligned}$$

In particular, this turns the integral into

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-ax^2-bx} dx &= \int_{-\infty}^{\infty} e^{-a(x+b/2a)^2+b^2/4a} dx \\ &= e^{b^2/4a} \int_{-\infty}^{\infty} e^{-a(x+b/2a)} dx \\ &= e^{b^2/4a} \left( \sqrt{\frac{\pi}{a}} \right) \\ &= e^{b^2/4a} \sqrt{\frac{\pi}{a}}. \end{aligned}$$

### Series Expansion

Consider the integral

$$\int_0^{\infty} \frac{x^{s-1}}{e^x - 1} dx.$$

This is a very nasty integral,<sup>xiii</sup> but we will need to know this value because it is useful in statistical mechanics.<sup>xiv</sup> We want to ensure this converges.

Notice that for large  $x$ , the integrand looks like  $e^{-x}x^{s-1}$ .

**Example.** To resolve the integral we take

$$\int_0^{\infty} \frac{x^{s-1}}{e^x - 1} dx = \int_0^{\infty} \frac{e^{-x}x^{s-1}}{1 - e^{-x}} dx$$

We will use the geometric series expansion for the denominator:

$$\begin{aligned} &= \int_0^{\infty} e^{-x}x^{s-1} \sum_{k=0}^{\infty} e^{-kx} dx \\ &= \sum_{k=0}^{\infty} \int_0^{\infty} x^{s-1} e^{-(k+1)x} dx. \end{aligned}$$

We make the change of variables  $u = (n+1)x$ .

$$= \sum_{n=0}^{\infty} \frac{1}{(n+1)^s} \int_0^{\infty} u^{s-1} e^{-u} du$$

<sup>xiii</sup>Citation needed.

<sup>xiv</sup>Okay actually I do kinda care about this.

$$= \underbrace{\sum_{n=1}^{\infty} \frac{1}{n^s}}_{\zeta(s)} \underbrace{\int_0^{\infty} u^{s-1} e^{-u} du}_{\Gamma(s)}.$$

Thus, our integral resolves to

$$\int_0^{\infty} \frac{x^{s-1}}{e^x - 1} dx = \Gamma(s)\zeta(s).$$

## Partial Fractions

**Example** (A Partial Fraction Decomposition).

$$\begin{aligned} \frac{1}{1-x^2} &= \frac{\alpha}{1-x} + \frac{\beta}{1+x} \\ &= \frac{1/2}{1-x} + \frac{1/2}{1+x}. \end{aligned}$$

**Example** (Integrating using Partial Fractions). To evaluate

$$\int \frac{4-2x}{(x^2+1)(x-1)^2} dx,$$

we do the partial fraction decomposition to find

$$\int \frac{4-2x}{(x^2+1)(x-1)^2} dx = \int \frac{2x+1}{x^2+1} + \frac{-2}{x-1} + \frac{1}{(x-1)^2} dx.$$

**Example** (Mean Value Theorem). If we have a function  $f$  defined on  $[a, b]$ , then there is a point  $c \in (a, b)$  such that

$$f(c)(b-a) = \int_a^b f(x) dx.$$

More generally, the mean value theorem says there exists  $c \in (a, b)$  such that

$$\int_a^b f(x)g(x) dx = f(c) \int_a^b g(x) dx$$

## Delta Distribution

Consider a “function”  $\delta(x)$  such that

$$\int_{-\infty}^{\infty} f(x)\delta(x-a) dx = f(a).$$

This idea seems absurd on its face — after all, singletons have measure zero, so the idea of an integral collapsing into a single point doesn’t sound normal.

The structure of the delta distribution is

$$\delta(x-a) = \begin{cases} +\infty & x = a \\ 0 & \text{else} \end{cases}.$$

In particular, we also have to define

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

This is known as the Dirac delta function (or rather, distribution). The delta distribution “weights”  $f$  to infinity at  $x = a$  and zero everywhere else.

**Example** (Delta Distribution as a Limit). Imagine a Gaussian function with area under the curve 1. In particular,

$$f_n(x) = \frac{1}{\sqrt{\pi}} n e^{-n^2 x^2}.$$

In particular, we have

$$\delta(x) = \lim_{n \rightarrow \infty} \frac{1}{\sqrt{\pi}} n e^{-n^2 x^2}$$

**Example** (A Physical Example). Imagine a ball is kicked. The force is dependent on time,  $F(t)$ .

There isn't an easy way to find the force, but by Newton's second law, we have

$$\delta p = \int F(t) dt,$$

where

$$I \equiv \int F(t) dt$$

is the impulse.

If we want to model  $F(t)$ , where we don't care about a nonzero time over which the force is occurring, we can simply state  $F(t)$  as

$$F(t) = \Delta p \delta(t - t_0).$$

Taking this integral yields  $I$ .

**Example** (Fourier Integral Representation of Delta Distribution). A different representation of  $\delta(x)$  is

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk.$$

We are superimposing all the waves  $e^{ikx}$  — in particular, for all values of  $k \neq 0$ , both  $e^{ikx}$  and  $e^{-ikx}$  are “added” together, yielding absolute destructive interference.

The factor of  $\frac{1}{2\pi}$  is necessary to normalize the integral.

### Properties of the Delta Distribution

#### Normalization:

$$\begin{aligned} \int_{-\infty}^{\infty} \delta(x) dx &= 1 \\ \int_{x_1}^{x_2} \delta(x - a) dx &= \begin{cases} 1 & x_1 < a < x_2 \\ 0 & \text{else} \end{cases}. \end{aligned}$$

#### Sieve:

$$\int_{x_1}^{x_2} f(x) \delta(x - a) dx = \begin{cases} f(a) & x_1 < a < x_2 \\ 0 & \text{else} \end{cases}.$$

**Example** (Delta Distribution as a Limit of Rectangles). We define the family of functions

$$\phi_k(x) = \begin{cases} k/2 & |x| < 1/k \\ 0 & |x| > 1/k \end{cases}.$$

We can see that integrating  $\phi_k$  over  $\mathbb{R}$  yields 1 for each  $k$ .

We now need to evaluate if  $\lim_{k \rightarrow \infty} \phi_k(x) = \delta(x)$ . In order to see this, we take

$$\begin{aligned} \lim_{k \rightarrow \infty} \int_{-\infty}^{\infty} f(x) \phi_k(x) dx &= \lim_{k \rightarrow \infty} \frac{k}{2} \int_{-1/k}^{1/k} f(x) dx \\ &= \lim_{k \rightarrow \infty} f(c_k) \left( \frac{k}{2} \int_{-1/k}^{1/k} dx \right) \\ &= \lim_{k \rightarrow \infty} f(c_k), \end{aligned}$$

where we define  $c_k$  from the mean value theorem. In particular, since  $c_k \in (-1/k, 1/k)$ , it is the case that  $c_k \rightarrow 0$  as  $k \rightarrow \infty$ , so

$$\begin{aligned} \lim_{k \rightarrow \infty} \int_{-\infty}^{\infty} f(x) \phi_k(x) dx &= \lim_{k \rightarrow \infty} f(c_k) \\ &= f(0). \end{aligned}$$

Thus,  $\lim_{k \rightarrow \infty} \phi_k(x) = \delta(x)$ .

We can imagine the delta distribution to be the density distribution of a single point.

The units of  $\delta(x)$  are

$$[\delta(x)] = x^{-1}.$$

**Example** (Linear Argument for  $\delta$ ). Consider  $\delta(ax)$ . For instance, we want to evaluate

$$\int_{-\infty}^{\infty} f(x) \delta(ax) dx.$$

To do so, we use  $u$  substitution with  $u = ax$ :

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) \delta(ax) dx &= \frac{1}{a} \int_{-\infty}^{\infty} f(u/a) \delta(u) du \\ &= \frac{1}{a} f(0). \end{aligned}$$

It is important to note that the integration variable  $dx$  and the argument of  $\delta(x)$  must be equal.

In general, we have

$$\delta(ax) = \frac{1}{|a|} \delta(x).$$

**Example** (Function Argument for  $\delta$ ). We now want to evaluate

$$\int_{-\infty}^{\infty} f(x) \delta(g(x)) dx.$$

When we take the change of variables, we have

$$\int_{y_1}^{y_2} f(y) \delta(y) dy = \int_{x_1}^{x_2} f(g(x)) \delta(g(x)) \left| \frac{dg}{dx} \right| dx.$$

Therefore, we must have  $\delta(g(x)) = \frac{1}{|dg/dx|_{g(x)=0}} \delta(x)$ .

In the general case, we have

$$\delta(g(x)) = \frac{1}{|dg/dx|_{x_0}} \delta(x - x_0)$$

where  $g(x_0) = 0$ .

If, in the region of integration,  $g$  takes multiple zeros, we must take a sum:

$$\delta(g(x)) = \sum_i \frac{1}{|dg/dx|_{x_i}} \delta(x_i);$$

where we assume  $\left| \frac{dg}{dx} \right|_{x_i} \neq 0$ .

**Example** ( $x^2 - a^2$  Argument for  $\delta$ ). Consider the distribution

$$\delta(x^2 - a^2).$$

The derivative of  $g(x)$  is  $2x$ ; the two zeros of  $g$  are at  $x = \pm a$ . Therefore,

$$\begin{aligned} \delta(x^2 - a^2) &= \frac{1}{|2x|_a} \delta(x - a) + \frac{1}{|2x|_{-a}} \delta(x + a) \\ &= \frac{1}{2a} (\delta(x - a) + \delta(x + a)). \end{aligned}$$

For example, if we took

$$\begin{aligned} \int_{-\infty}^{\infty} x^3 \delta(x^2 - a^2) dx &= \frac{1}{2a} \int_{-\infty}^{\infty} x^3 (\delta(x - a) + \delta(x + a)) dx \\ &= \frac{1}{2a} (a^3 + (-a)^3) \\ &= 0. \end{aligned}$$

Now, evaluating

$$\begin{aligned} \int_0^{\infty} x^3 (\delta(x^2 - a^2)) dx &= \frac{1}{2a} \int_0^{\infty} x^3 (\delta(x - a) + \delta(x + a)) dx \\ &= \frac{1}{2a} (a^3) \\ &= \frac{1}{2} a^2. \end{aligned}$$

**Example** ((Weak) Derivative of  $\delta$ ). Obviously we cannot formally take  $\delta'(x)$ , but we can always place  $\delta(x)$  under the integral sign and treat  $\delta'(x)$  as the “derivative” via integration by parts:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) \delta'(x) dx &= f(x) \delta(x) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{df}{dx} \delta(x) dx \\ &= - \frac{df}{dx} \Big|_0 \end{aligned}$$



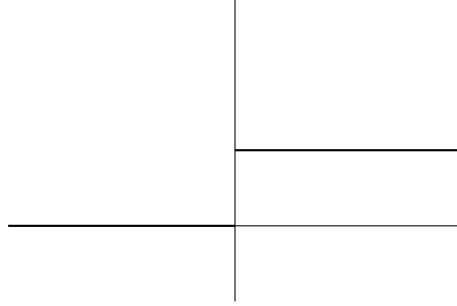
$$= -f'(0).$$

The “identity” for the delta function’s derivatives is

$$f(x)\delta'(x) = -f'(x)\delta(x).$$

**Example** (Heaviside Step Function). The Heaviside step function,  $\Theta(x)$ , is

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



**Example** (Higher Dimension Delta Distributions). In higher dimensions,

$$\int_{\text{all space}} \delta(\mathbf{r}) \, d\tau = 1,$$

and

$$\int_V f(\mathbf{r}) \delta(\mathbf{r} - \mathbf{a}) \, d\tau = \begin{cases} f(\mathbf{a}) & \mathbf{a} \in V \\ 0 & \text{otherwise} \end{cases}.$$

One of the common notations for higher dimensional delta functions is  $\delta^{(n)}(\mathbf{r})$ , where  $(n)$  denotes the dimension (not to be confused with  $n$ th derivative).

Instead, we can use  $\delta(\mathbf{r})$ , which lets us know that we are dealing in higher dimensions, and context is evident.

**Example** (Voltage under a Point Charge). The voltage of a point charge  $q$  at a position  $\mathbf{a}$  is given by Coulomb’s law

$$\Phi(\mathbf{r}) = \frac{q}{4\pi\epsilon_0} \frac{1}{|\mathbf{r} - \mathbf{a}|},$$

with  $\Phi = 0$  at  $\infty$ .

For a continuous point charge distribution  $\rho(\mathbf{r})$ , we consider each element of the volume  $d\tau$  centered at  $\mathbf{r}$  with charge  $dq = \rho(\mathbf{r}) \, d\tau$ .

$$\begin{aligned} d\Phi(\mathbf{r}) &= \frac{dq}{4\pi\epsilon_0} \frac{1}{|\mathbf{r} - \mathbf{a}|} \\ &= \frac{1}{4\pi\epsilon_0} \frac{\rho(\mathbf{r}) \, d\tau}{|\mathbf{r} - \mathbf{a}|}. \end{aligned}$$

In particular, for some  $\mathbf{r}$ , we need to add up over  $\mathbf{a}$ , yielding

$$\Phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int_V \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d\tau'$$

This expression should hold for every physically reasonable volume charge distribution  $\rho$ , what  $\rho(\mathbf{r})$  denotes a point charge?

In particular, if  $\rho(\mathbf{r})$  is a point charge, then  $\rho = q\delta(\mathbf{r} - \mathbf{a})$ .

**Example** (Using the Multi-Dimensional Delta Distribution). In Cartesian coordinates, we have

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

If we want to transform  $\delta(\mathbf{r} - \mathbf{r}_0)$  into a different coordinate system such as  $d\tau = du dv dw$ , we need the Jacobian. Thus,

$$\delta(\mathbf{r} - \mathbf{r}_0) = \frac{1}{|J|} \delta(u - u_0) \delta(v - v_0) \delta(w - w_0).$$

For instance, in spherical coordinates, we have

$$\begin{aligned} \delta(\mathbf{r} - \mathbf{r}_0) &= \frac{1}{r^2 \sin \theta} \delta(r - r_0) \delta(\theta - \theta_0) \delta(\phi - \phi_0) \\ &= \frac{1}{r^2} \delta(r - r_0) \delta(\cos \theta - \cos \theta_0) \delta(\phi - \phi_0). \end{aligned}$$

## Vector Calculus

**Question:** What is a vector?

**Answer:** A vector is an element of a vector space.

**Remark:** Yes, vectors as defined by “magnitude and direction” also are elements of vector spaces.

For the purposes of this unit, we will focus on vectors in the vector space  $\mathbb{R}^n$  over  $\mathbb{R}$ .

**Notation:** Vector-valued functions with vector-valued outputs will be denoted

$$\mathbf{F}(\mathbf{r}).$$

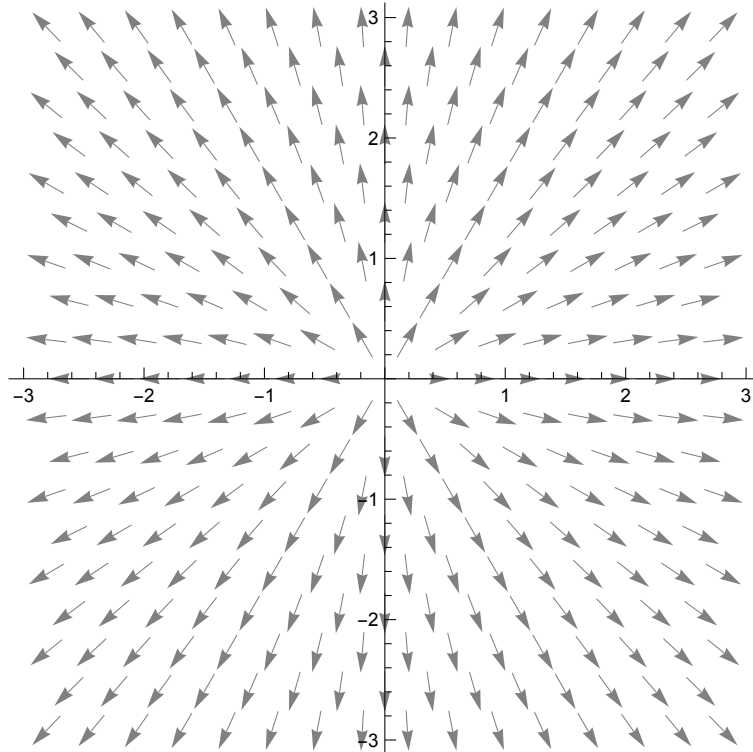
## Vector Fields

**Definition** (Vector Field). A vector-valued function  $\mathbf{F}(\mathbf{r})$  with vector-valued outputs is known as a vector field.

**Example.** The field

$$\mathbf{F}(x, y, z) = x\hat{i} + y\hat{j} + z\hat{k}$$

can be seen below.

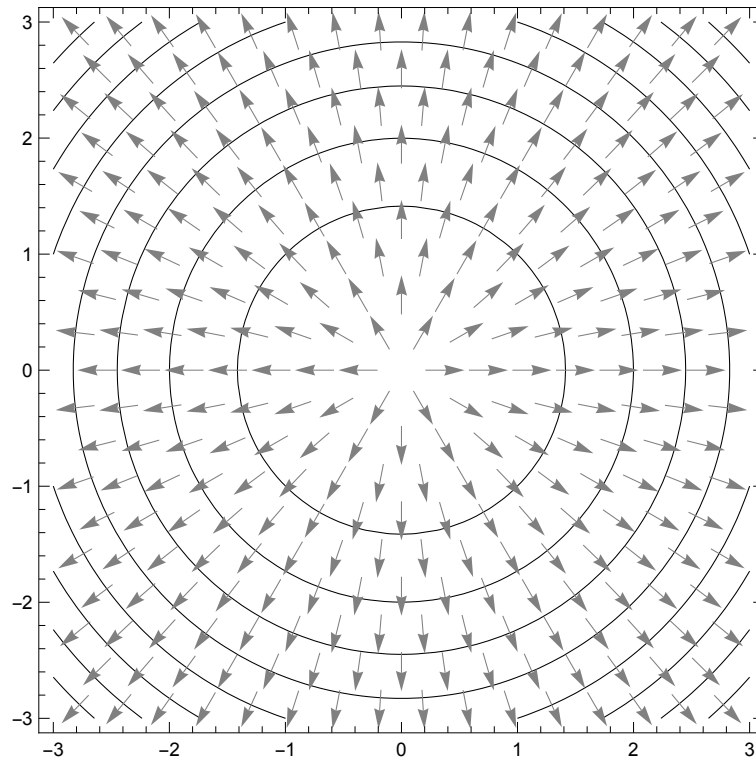


Notice that, in terms of spherical coordinates,  $\mathbf{F}(x, y, z) = r\hat{\mathbf{r}} = \mathbf{r}$ .

**Definition** (Incompressible Fluid). A fluid is incompressible if its density is constant.

In particular, incompressible fluids cannot have either sources or sinks, since sources imply a local reduction in density, while sinks imply a local increase in density.

**Example.** Consider a sprinkler with  $N$  streams. Since water is incompressible, the density of streamlines  $\sigma$  and the surface area of the spherical shells,  $A$  must be inversely proportional to each other.

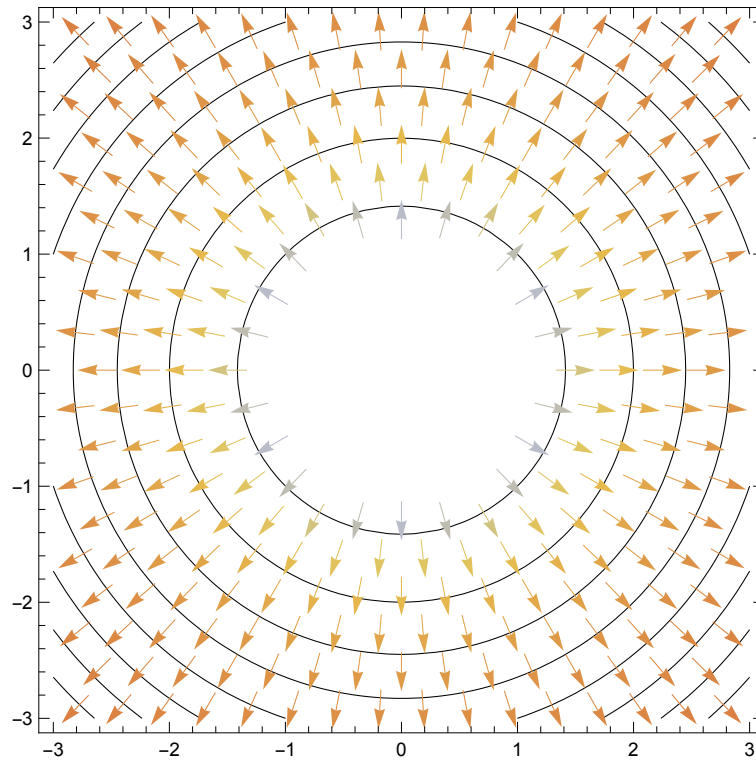


Thus, we have

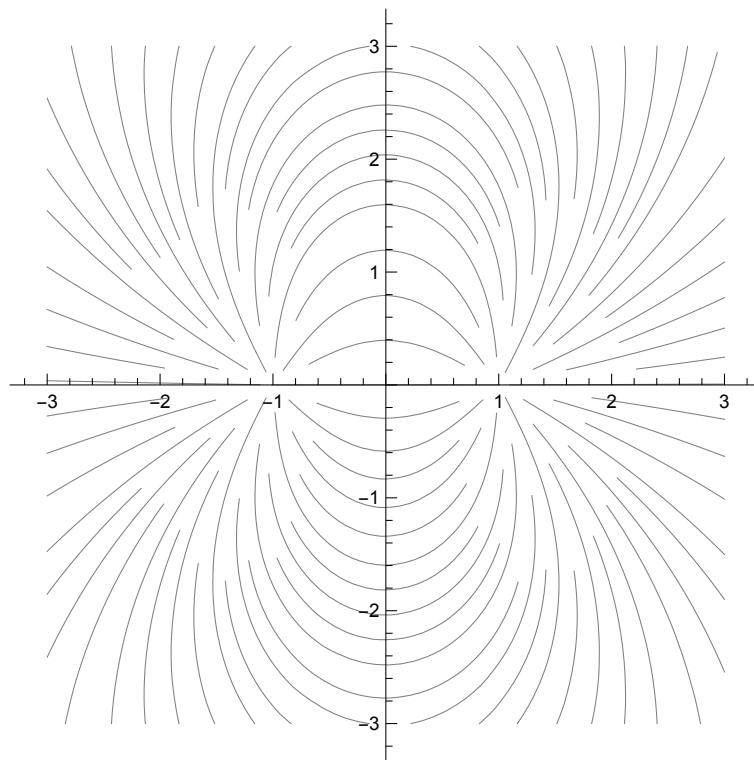
$$N = \sigma A,$$

meaning  $\sigma \sim \frac{1}{r^2}$  since  $A \sim r^2$ .

In particular, the strength of the vector field must diminish with the square of the distance.



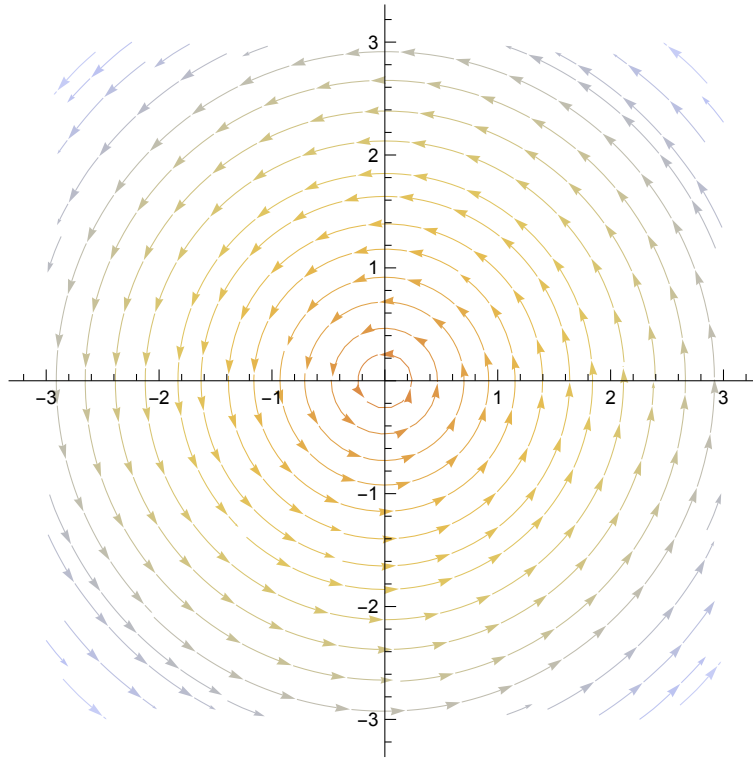
**Example.** Vector fields can be added.



**Example.** Consider the field

$$\mathbf{G}(x, y) = \frac{1}{\sqrt{x^2 + y^2}} (-y\hat{i} + x\hat{j})$$

As depicted, we can see that the vector field looks as follows.



In particular, we can see that  $\mathbf{G} = \hat{\phi}$ .

Notice that our vector fields are dependent on both the basis and the coordinate system.

In particular, we have reason to prefer a Cartesian basis over the polar or spherical basis, since the Cartesian basis is position-independent.

## Gradient, Divergence, and Curl

### The $\nabla$ Operator

Consider a scalar function  $f(\mathbf{r})$ . If we want to imagine how  $f$  changes as we move  $\mathbf{r}$  to  $\mathbf{r} + d\mathbf{r}$ , we use the chain rule.

$$\begin{aligned} df &= \left( \frac{\partial f}{\partial x} \right) dx + \left( \frac{\partial f}{\partial y} \right) dy + \left( \frac{\partial f}{\partial z} \right) dz \\ &= \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} \\ &= \nabla f \cdot d\mathbf{r}. \end{aligned}$$

In particular, we define

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

Notice that, since  $dx\hat{i} + dy\hat{j} + dz\hat{k}$  is a vector, and  $df$  is a scalar, we know that  $\nabla f$  *must* be a vector.

**Definition** (Gradient).

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

$$df = |\nabla f| |d\mathbf{r}| \cos \theta.$$

If  $\nabla f$  is in the direction of  $d\mathbf{r}$ , then  $\cos \theta = 1$ , meaning  $df$  is maximized. In particular,  $\nabla f$  points in the direction of maximum change in  $f$ .

In particular, this means that for every (differentiable) scalar field, there is a natural vector field associated with the direction of largest increase.

**Example.** The electric field

$$\mathbf{E} = -\nabla V.$$

Similarly, for any given potential  $U$ ,

$$\mathbf{F} = -\nabla U.$$

**Definition** (The  $\nabla$  Operator).

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}}_{\nabla} (f)$$

Thus, we get

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

**Example.**

(1) For some scalar field  $f(\mathbf{r})$ , we can take

$$\nabla(f) = \nabla f,$$

which yields the gradient field.

(2) For some vector field  $\mathbf{E}$ , we can take

$$\nabla \cdot \mathbf{E} = g$$

which yields a scalar field known as the divergence of  $\mathbf{E}$ .

In particular,

$$\nabla \cdot \mathbf{E} = \frac{\partial}{\partial x} (\mathbf{E} \cdot \hat{i}) + \frac{\partial}{\partial y} (\mathbf{E} \cdot \hat{j}) + \frac{\partial}{\partial z} (\mathbf{E} \cdot \hat{k}).$$

(3) For some vector field  $\mathbf{B}$ , we can take

$$\nabla \times \mathbf{B} = \mathbf{A},$$

which yields a vector field known as the curl of  $\mathbf{B}$ .

(4)

$$\begin{aligned}
\nabla \cdot (\nabla f) &= (\nabla \cdot \nabla) f \\
&= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \\
&= \nabla^2 f \\
&= \Delta f,
\end{aligned}$$

which yields an operator known as the Laplacian.

**Example.** Let  $\mathbf{v}_1 = xy\hat{i} + y^2\hat{j}$ . Then,

$$\begin{aligned}
\nabla \cdot \mathbf{v}_1 &= \frac{\partial}{\partial x}(xy) + \frac{\partial}{\partial y}(y^2) \\
&= y + 2y \\
&= 3y,
\end{aligned}$$

and

$$\begin{aligned}
\nabla \times \mathbf{v}_1 &= \left( \frac{\partial}{\partial x}(y^2) - \frac{\partial}{\partial y}(xy) \right) \hat{k} \\
&= -x\hat{k}.
\end{aligned}$$

**Example.** Let  $\mathbf{v}_2 = \frac{1}{x^2+y^2+z^2} (x\hat{i} + y\hat{j} + z\hat{k})$ . Then,

$$\begin{aligned}
\nabla \cdot \mathbf{v}_2 &= \frac{\partial}{\partial x} \left( \frac{x}{x^2+y^2+z^2} \right) + \frac{\partial}{\partial y} \left( \frac{y}{x^2+y^2+z^2} \right) + \frac{\partial}{\partial z} \left( \frac{z}{x^2+y^2+z^2} \right) \\
&= \frac{1}{x^2+y^2+z^2} \nabla \cdot \mathbf{v}_2 = 0
\end{aligned}$$

**Example.** Consider

$$\begin{aligned}
\mathbf{v} &= x^2\hat{i} + y^2\hat{j} + z^2\hat{k} \\
\mathbf{u} &= yz\hat{i} + zx\hat{j} + xy\hat{k}.
\end{aligned}$$

In particular, it is easily verified that  $\nabla \times \mathbf{v} = 0$  and  $\nabla \cdot \mathbf{u} = 0$ .

### Applying Vector Identities to the $\nabla$ Operator

We are aware that

$$\begin{aligned}
a\mathbf{V} &= \mathbf{V}a \\
\mathbf{A} \cdot \mathbf{B} &= \mathbf{B} \cdot \mathbf{A}.
\end{aligned}$$

However, when we deal with  $\nabla$ , we have to respect both the properties of the vectors *and* the properties of the operator. In particular,

$$f\nabla \neq \nabla f.$$

This is because  $f\nabla$  is a vector operator, while  $\nabla f$  is a vector field. Similarly,

$$\nabla \cdot \mathbf{E} \neq \mathbf{E} \cdot \nabla,$$

since  $\nabla \cdot \mathbf{E}$  is a scalar field, while  $\mathbf{E} \cdot \nabla$  is a scalar operator.



**Example (Curl of Curl).** Consider

$$\nabla \times (\nabla \times \mathbf{v}).$$

On first glance, we want to use the identity  $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ , yielding

$$\nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \mathbf{v}(\nabla \cdot \nabla).$$

However, notice that  $\mathbf{v}(\nabla \cdot \nabla)$  is a scalar operator, while  $\nabla \times (\nabla \times \mathbf{v})$  is a vector field. Thus, we have to modify the double cross product to  $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - (\mathbf{A} \times \mathbf{B})\mathbf{C}$

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

**Example (Curl of Gradient and Divergence of Curl).** Consider

$$\nabla \times (\nabla f).$$

In particular, we are tempted to take

$$\begin{aligned} \nabla \times (\nabla f) &= (\nabla \times \nabla) f \\ &= 0. \end{aligned}$$

This is allowed, since we do not affect the property of the operation.

The following identity is also true,

$$\nabla \cdot (\nabla \times \mathbf{v}) = 0,$$

but we cannot use a cheesy method to prove this.

**Remark:** Faraday's law is  $\nabla \times \mathbf{E} = 0$ , and Gauss's law for magnetism is  $\nabla \cdot \mathbf{B} = 0$ , where  $\mathbf{E}$  denotes the electric field and  $\mathbf{B}$  denotes the magnetic field.

If  $\nabla \times \mathbf{E} = 0$  in electrostatics, then  $\mathbf{E} = \nabla A$  for some scalar function  $A$ . In particular, we say  $\mathbf{E} = -\nabla V$ . We call  $V$  the scalar potential.

Similarly, if  $\nabla \cdot \mathbf{B} = 0$ , then  $\mathbf{B} = \nabla \times \mathbf{A}$  for some vector field  $\mathbf{A}$ . We call  $\mathbf{A}$  the vector potential.

**Example (Products).** Consider

$$\nabla(fg).$$

Using the product rule, we have

$$\nabla(fg) = (\nabla f)g + f(\nabla g).$$

However, when we look at

$$\nabla \cdot (f\mathbf{A}),$$

things get a little more complicated. Notice that  $\nabla \cdot (f\mathbf{A})$  is a scalar, meaning we apply the product rule using dot products to yield such a scalar.

$$\nabla \cdot (f\mathbf{A}) = \nabla f \cdot \mathbf{A} + f \nabla \cdot \mathbf{A}.$$

Similarly,

$$\begin{aligned} \nabla \times (f\mathbf{A}) &= (\nabla \times \mathbf{A})f + \nabla f \times \mathbf{A} \\ &= (\nabla \times \mathbf{A})f - \mathbf{A} \times \nabla f. \end{aligned}$$

## Changing Coordinates

Vector equations are fundamentally independent of their coordinate systems. Thus, the established identities in the previous subsection must be valid regardless of the coordinate system.

For instance, we should be able to calculate

$$\nabla \cdot \mathbf{E}$$

regardless of the coordinate system.

**Example** (Converting to Polar Coordinates). Consider

$$\mathbf{v}_1 = xy\hat{i} + y^2\hat{j}.$$

Conversion to polar coordinates yields

$$\mathbf{v}_1 = (r^2 \sin \phi) \hat{r}.$$

## Understanding $\nabla^2$ , $\nabla \cdot$ , and $\nabla \times$

**Example** (Understanding the Laplacian). One of the most important differential equations is the equation for simple harmonic motion:

$$\frac{d^2}{dt^2} f(t) = -\omega^2 f(t).$$

However, this equation does not need to be in time. We can also imagine this oscillation happening in space:

$$\frac{d^2}{dx^2} f(x) = -k^2 f(x).$$

This doesn't need to occur in one dimension, though. A reasonable assumption is that what occurs in  $x$  should occur in  $y$  and  $z$  symmetrically.

$$\underbrace{\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right)}_{\nabla^2} f(x, y, z) = - \underbrace{(k_x^2 + k_y^2 + k_z^2)}_{\|\mathbf{k}\|^2} f(x, y, z).$$

Essentially,  $\nabla^2$  is a measure of the concavity of the function in a particular direction.

In one dimension, if  $\frac{df}{dx}|_P = 0$  and  $\frac{d^2f}{dx^2}|_{x=P} < 0$ , we know that  $f(P)$  is lower than the average value "around" the point  $x = P$ .

**Example** (Understanding the Divergence). If we have a vector field  $\mathbf{F}$  in three dimensions, there are nine different ways to understand a rate of change,  $\partial_i F_j$ .

We start by choosing axes such that  $\mathbf{F}(P) = F_x(P)\hat{i}$ . Moving along the streamline by  $dx$ , a simple linear approximation gives,

$$F_x(P + dx\hat{i}) = F_x(P) + \frac{\partial F_x}{\partial x} \Big|_{x=P} dx.$$

When we move along the streamline by  $dx$ , if  $\frac{\partial F_x}{\partial x} \Big|_{x=P} > 0$ , we see that  $F_x(P + dx\hat{i})$  increases. Essentially, this derivative measures the "surge" from point  $P$ .

If we look at  $\mathbf{F}$  by moving along  $dy\hat{j}$ , we find

$$F_y(P + dy\hat{j}) = \left. \frac{\partial F_y}{\partial y} \right|_{y=P} dy,$$

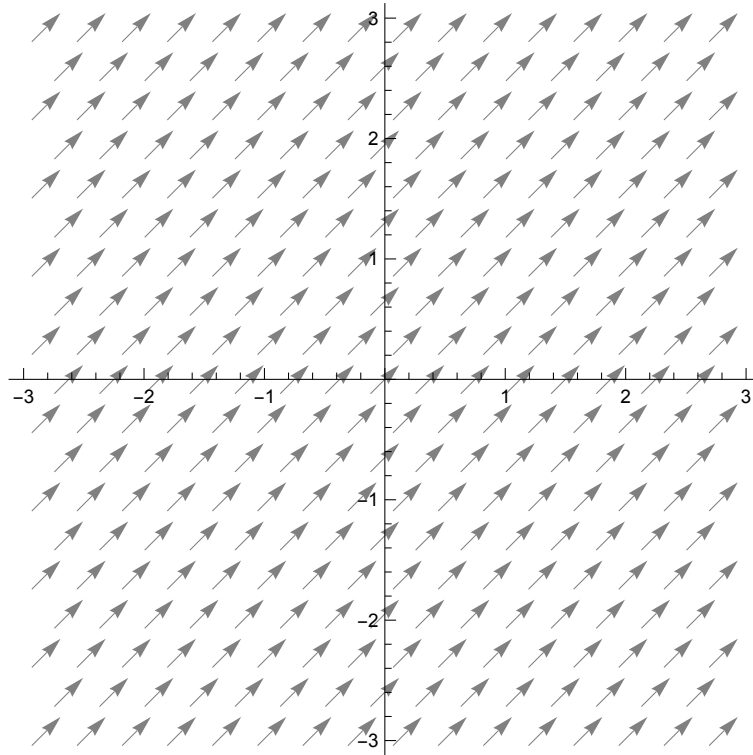
since  $F_y(P) = 0$ . If  $\frac{\partial F_y}{\partial y} > 0$ , the streamlines seem to “spread out” from each other.

These derivatives measure how a field “surges” out of a point and how it “spreads” out of a point.

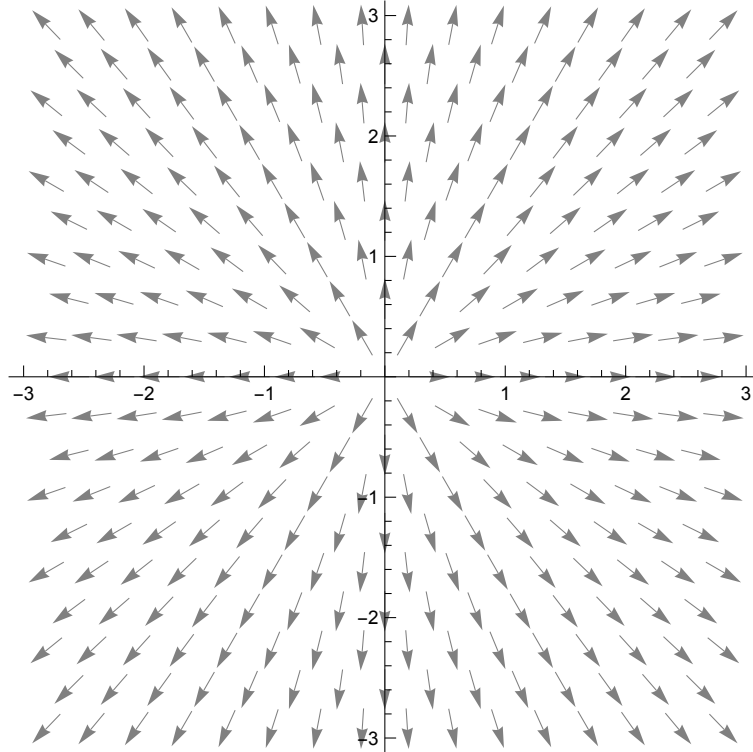
This measure of surge and/or spread has to be axis-independent. All of these have to coincide, meaning the full measure of surge and spread is

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

**Example** (Coordinate Dependence of Divergence). The following field represents  $\mathbf{F} = \hat{i} + \hat{j}$ , which has zero divergence.



The following field represents  $\mathbf{F} = \hat{r}$ , which has positive divergence (recall that  $\hat{r}$  is not position-independent).



**Definition** (Solenoidal Field). A field that has zero divergence is known as a solenoidal or divergence-free field.

Solenoidal fields are useful for modeling incompressible fluids, as incompressible fluids have constant density.

If  $\nabla \cdot \mathbf{F} = 0$ , then  $\mathbf{F} = \nabla \times \mathbf{A}$  for some other vector field  $\mathbf{A}$ . Additionally, since  $\mathbf{F} = \nabla \Phi$  for some scalar field  $\Phi$ , we recover Laplace's equation:

$$\begin{aligned}\nabla \cdot (\nabla \Phi) &= 0 \\ \nabla^2 \Phi &= 0.\end{aligned}$$

**Example** (Understanding the Curl). Of the nine combinations  $\partial_i F_j$ , we have used 3 of them via the divergence.

Now, we explore the curl, which is  $\nabla \times \mathbf{F}$ , which gives information about  $\partial_i F_j$  where  $i \neq j$ .

By the definition of the cross product, we have

$$(\nabla \times \mathbf{F})_k = \partial_i F_j - \partial_j F_i,$$

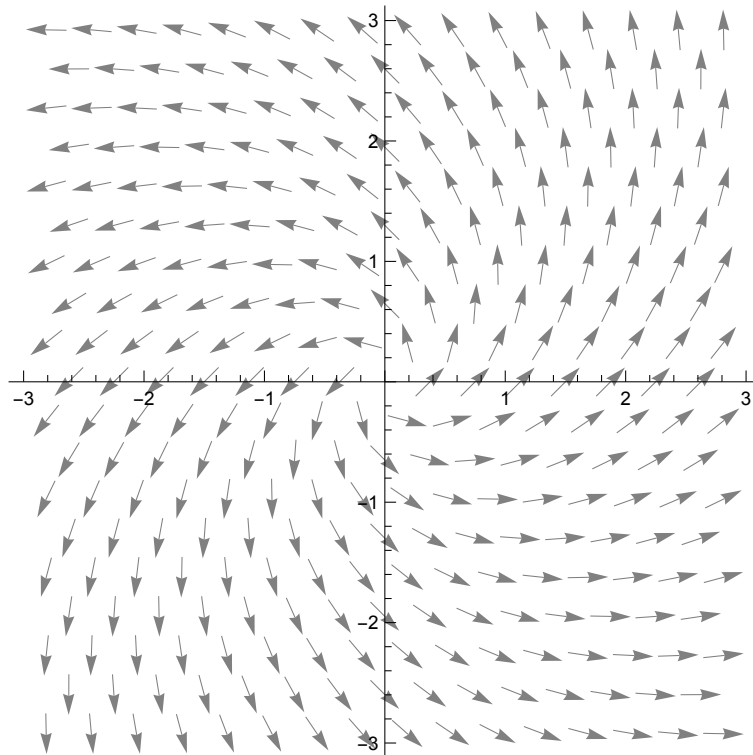
where  $i \neq j$ .

Considering water rotating with angular velocity  $\omega$ , we find  $\mathbf{v} = \omega \times \mathbf{r}$ . Now, taking

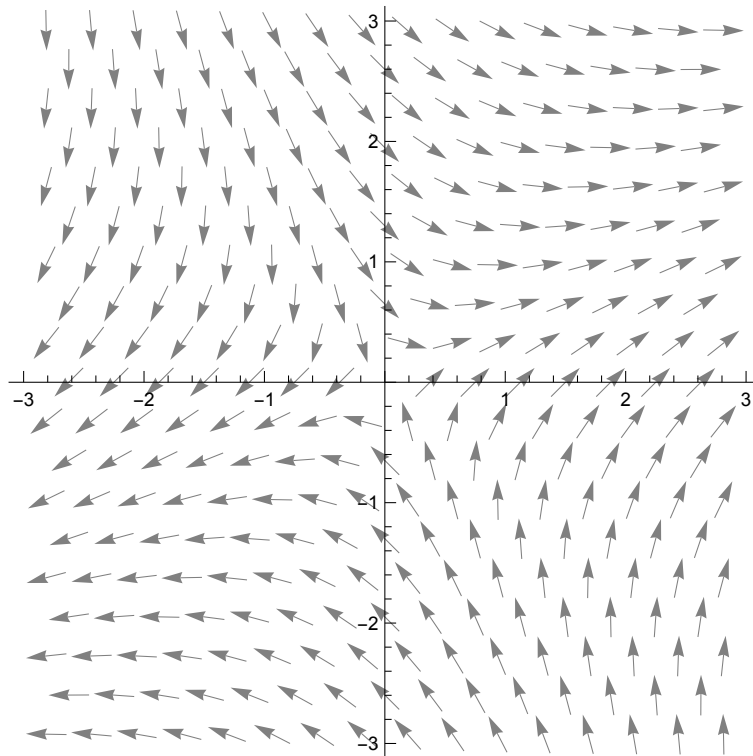
$$\begin{aligned}\nabla \times \mathbf{v} &= \nabla \times (\mathbf{r} \times \omega) \\ &= 2\omega.\end{aligned}$$

Therefore, curl measures some “swirl” of a given vector field.

**Example** (Fields with Differing Curl). The field  $\mathbf{F} = (x - y)\hat{i} + (x + y)\hat{j}$  has positive curl everywhere.



Meanwhile, the field  $\mathbf{F} = (x + y)\hat{i} + (x - y)\hat{j}$  has zero curl everywhere.



## Integrating Scalar and Vector Fields

Integration is summation,<sup>xv</sup> meaning the expression

$$\int_{x_1}^{x_2} f(x) dx$$

describes a sum along a unique interval defined by  $x_1$  and  $x_2$ .

When we go to higher dimensions, we first think of

$$\int_{P_1}^{P_2} f(x, y) dx,$$

which wants us to sum  $f(x, y)$  from  $P_1 = (x_1, y_1)$  to  $P_2 = (x_2, y_2)$ . However, this is not a fully specific expression — we need a *path* along which we integrate. We specify the path by  $y = g(x)$ . Now, the integral becomes

$$\int_{P_1}^{P_2} f(x, g(x)) dx.$$

Similarly, for a surface integral

$$\int_S f(x, y) dx dy,$$

we need the surface  $S$  along which we integrate, where we say  $z = g(x, y)$ .

However, while there are certain functions that are path-independent,<sup>xvi</sup> we must assume that every function is *path-dependent*.

### Line Integrals

A line integral is a sum over a curve  $C$ .

$$\int_C f(\mathbf{r}) d\ell.$$

Here,  $d\ell$  denotes the length element along  $C$ . For instance, if  $f$  denotes the charge density per unit length, and  $C$  is a wire, then  $\int_C f(\mathbf{r}) d\ell$ .

If we define  $C$  by  $y = g(x)$ , then the integral is

$$\int_C f(x, y) d\ell = \int_C f(x, g(x)) d\ell.$$

However, we need to figure out how to deal with  $d\ell$  — in particular, we need  $d\ell$  to be an expression only in  $dx$ . In particular,  $d\ell$  is given by the Pythagorean theorem:

$$(d\ell)^2 = (dx)^2 + (dy)^2.$$

**Notation:** Everyone (else) drops the parentheses:

$$d\ell^2 = dx^2 + dy^2.$$

---

<sup>xv</sup>kinda

<sup>xvi</sup>Holomorphic functions, for instance

Our integral now becomes

$$\begin{aligned}\int_C f(x, g(x)) \, d\ell &= \int_C f(x, g(x)) \sqrt{(dx)^2 + (dy)^2} \\ &= \int_{x_1}^{x_2} f(x, g(x)) \, dx \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\end{aligned}$$

**Example.** Let

$$\begin{aligned}f(\mathbf{r}) &= \sqrt{\frac{1}{4}x^2 + y^2} \\ y(x) &= \frac{1}{2}x^2,\end{aligned}$$

integrated from  $0 < x < 1$ .

Then,

$$\begin{aligned}\int_C f(\mathbf{r}) \, d\ell &= \int_0^1 \sqrt{\frac{1}{4}x^2 + \frac{1}{4}x^4} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \, dx \\ &= \frac{1}{2} \int_0^1 x \sqrt{1 + x^2} \sqrt{1 + x^2} \, dx \\ &= \frac{1}{2} \int_0^1 x (1 + x^2) \, dx \\ &= \frac{3}{8}.\end{aligned}$$

Note that we can also evaluate this integral as a function of  $y$ , and still get the same outcome.

**Example.** Let

$$\begin{aligned}f(\mathbf{r}) &= \sqrt{\frac{1}{4}x^2 + y^2} \\ y(x) &= x,\end{aligned}$$

integrated from  $0 < x < 1$ .

Then,

$$\begin{aligned}\int_C f(\mathbf{r}) \, d\ell &= \int_0^1 \sqrt{\frac{1}{4}x^2 + x^2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \, dx \\ &= \sqrt{\frac{5}{2}} \int_0^1 x \, dx \\ &= \sqrt{\frac{5}{8}}.\end{aligned}$$

We can also do this in polar coordinates. First, we have

$$(d\ell)^2 = (dr)^2 + r^2 (d\theta)^2.$$

Then,

$$\int_C f(\mathbf{r}) \, d\ell = \int_0^{\sqrt{2}} \sqrt{\frac{1}{4}r^2 \cos^2 \theta + r^2 \sin^2 \theta} \, dr$$

$$= \sqrt{\frac{5}{8}}.$$

When we integrate with respect to a parametrized curve in one dimension, we have

$$\int_{x_1}^{x_2} f(x) dx = \int_{t(x_1)}^{t(x_2)} f(x(t)) \frac{dx}{dt} dt.$$

In multiple dimensions for  $c(t) = (x(t), y(t))$ , we have

$$\begin{aligned} d\ell &= \sqrt{(dx)^2 + (dy)^2} \\ &= dt \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \\ &= \left\| \frac{dc}{dt} \right\| dt, \end{aligned}$$

yielding

$$\int_C f(\mathbf{r}) d\ell = \int_{t_1}^{t_2} f(x(t), y(t)) \left\| \frac{dc}{dt} \right\| dt.$$

### Surface Integrals

When we turn to a surface

$$\int_S f(\mathbf{r}) dA,$$

we parametrize  $z = g(x, y)$  to yield

$$\int_S f(x, y, z) dA = \int_D f(x, y, g(x, y)) \sqrt{1 + \|\nabla g\|^2} dx dy,$$

where  $D$  is the projection of  $S$  onto the  $x, y$ -plane.

### Circulation

So far, we have only considered scalar fields, When we look at

$$\int_C \mathbf{F}(\mathbf{r}) d\ell = \hat{i} \int_C F_x(\mathbf{r}) d\ell + \hat{j} \int_C F_y(\mathbf{r}) d\ell + \hat{k} \int_C F_z(\mathbf{r}) d\ell.$$

More commonly, though, we are interested in

$$\int_C \mathbf{F}(\mathbf{r}) \cdot d\vec{\ell},$$

which is a scalar quantity. When we add the dot product into the integral, we signal that we are most interested in the *parallel* component of  $\mathbf{F}$  to  $d\vec{\ell}$ .

In cartesian coordinates, this integral is equal to

$$\int_C \mathbf{F}(\mathbf{r}) \cdot d\vec{\ell} = \int_C F_x dx + F_y dy + F_z dz.$$

For a more physical example,<sup>xvii</sup> this integral is a measure of work if  $\mathbf{F}(\mathbf{r})$  denotes force. Similarly,

$$\Delta V = - \int_{P_1}^{P_2} \mathbf{E} \cdot d\vec{\ell}$$

---

<sup>xviii</sup>Who cares about that?



**Definition** (Circulation). When we integrate over a closed curve  $C$ , our line integral now becomes

$$\Psi = \oint_C \mathbf{F} \cdot d\vec{\ell}$$

**Remark:** It is not always the case that  $\Psi$  is zero.

Summing  $\mathbf{F}$  along a closed path  $C$  measures the net “swirl” of the vector field.

**Example** (Symmetry in Circulation). Ampère’s law says that

$$\oint_C \mathbf{B} \cdot d\vec{\ell} = \mu_0 I,$$

where  $\mu_0$  is the permeability of free space,  $I$  denotes current, and  $\mathbf{B}$  is the magnetic field.

Consider the case of a wire with constant radius  $a$  and uniform current  $I_0$ . From cylindrical symmetry, it must be the case that  $\mathbf{B}$  forms concentric circles about the wire.

Since Ampère’s law says we can *always* pick a curve that yields a constant, we can let  $C$  correspond to one of the circles with constant  $\|\mathbf{B}\|$ .

$$\begin{aligned} \oint_C \mathbf{B} \cdot d\vec{\ell} &= \oint_C B \, d\ell \\ &= B \oint_C d\ell \\ &= 2\pi r B, \end{aligned}$$

where  $r$  is the radius of the loop  $C$ .

Note that this value does not depend on the radius of  $C$ , but the value of  $\mu_0 I$  does depend on the radius of  $C$ .

In particular,

$$I(r < a) = I_0 \frac{r^2}{a^2},$$

yielding

$$\begin{aligned} \mathbf{B}(r < a) &= \frac{\mu_0 I_0}{2\pi} \frac{r}{a^2} \hat{\phi} \\ \mathbf{B}(r > a) &= \frac{\mu_0 I_0}{2\pi r} \hat{\phi}. \end{aligned}$$

Inside the wire, the magnitude of the  $B$  field grows linearly with respect to  $r$ , and outside, the magnitude of the  $B$  field falls off by a factor of  $r^{-1}$ .

However, in the general case, we need to parametrize our line integral.

$$\int_C \mathbf{F} \cdot d\vec{\ell} = \int_{t_1}^{t_2} \left( \mathbf{F} \cdot \frac{d\mathbf{c}}{dt} \right) dt$$

**Example** (Path-Dependence (or Path-Independence) of Work). Let  $\mathbf{B} = x^2 y \hat{i} - x y^2 \hat{j}$ ,  $P_1 = (0, 0)$ , and  $P_2 = (1, 1)$ . We will calculate the work done by  $\mathbf{B}$  as follows:

$$\int_C \mathbf{B} \cdot d\vec{\ell} = \int_C B_x \, dx + B_y \, dy.$$

Along the path  $C_1$ , which goes from  $(0, 0)$  to  $(1, 0)$ , then  $(1, 0)$  to  $(1, 1)$ , the work is given by

$$\begin{aligned}\int_C \mathbf{B} \cdot d\vec{\ell} &= \int_{C_1} x^2 y \, dx - xy^2 \, dy = \int_0^1 x^2 y|_{y=0} \, dx - \int_0^1 xy^2|_{x=1} \, dy \\ &= -\frac{1}{3}.\end{aligned}$$

Along the path  $C_2$ , which goes from  $(0, 0)$  to  $(1, 1)$  directly, with  $x = y$  and  $dy = dx$ , we have

$$\begin{aligned}\int_{C_2} \mathbf{B} \cdot d\vec{\ell} &= \int_{C_2} (x^2 y - xy^2)|_{y=x} \, dx \\ &= 0.\end{aligned}$$

Along the path  $C_3$ , defined by  $x = 1 - \cos \theta$  and  $y = \sin \theta$ , we have  $(dx, dy) = (\sin \theta, \cos \theta) d\theta$ , meaning

$$\begin{aligned}\int_{C_3} \mathbf{B} \cdot d\vec{\ell} &= \int_0^{\pi/2} ((1 - \cos \theta)^2 \sin \theta) \sin \theta \, d\theta - \int_0^{\pi/2} ((1 - \cos \theta) \sin^2 \theta) \cos \theta \, d\theta \\ &= \frac{3\pi}{8} - 1.\end{aligned}$$

Meanwhile, for  $\mathbf{E} = xy^2 \hat{i} + x^2 y \hat{j}$ , it is the case that

$$\begin{aligned}\int_{C_1} \mathbf{E} \cdot d\vec{\ell} &= \frac{1}{2} \\ \int_{C_2} \mathbf{E} \cdot d\vec{\ell} &= \frac{1}{2} \\ \int_{C_3} \mathbf{E} \cdot d\vec{\ell} &= \frac{1}{2}.\end{aligned}$$

**Definition** (Path-Independence). Let  $\mathbf{F}(\mathbf{r})$  be a vector field. If, for every closed path  $C$ , it is the case that

$$\oint_C \mathbf{F} \cdot d\vec{\ell} = 0,$$

then we say  $\mathbf{F}$  is path-independent.

If our vector field is path-independent, then we are allowed to pick the path that is easiest to calculate.

If  $\mathbf{E}$  is path-independent, then there has to exist some scalar field  $\Phi$  such that

$$\int_{t_1}^{t_2} \mathbf{E} \cdot d\vec{\ell} = \Phi(t_2) - \Phi(t_1)$$

**Definition** (Equivalent Conditions for Path-Independence). Let  $\mathbf{F}(\mathbf{r})$  be a vector field. Then, the following are equivalent:

- $\oint_C \mathbf{F} \cdot d\vec{\ell} = 0$
- $\mathbf{F} = \nabla \Phi$
- $\nabla \times \mathbf{F} = 0$

**Example** (Finding the Scalar Field). Consider  $\mathbf{E} = xy^2 \hat{i} + x^2 y \hat{j}$ . Since  $\mathbf{E}$  has curl 0, we know there must exist some  $\Phi$  such that

$$\frac{\partial \Phi}{\partial x} = xy^2$$

$$\begin{aligned}\Phi &= \frac{1}{2}x^2y^2 + f(y) \\ \frac{\partial \Phi}{\partial y} &= xy \\ \Phi &= \frac{1}{2}x^2y^2 + g(x).\end{aligned}$$

Such a  $\Phi$  exists if and only if we can choose  $f$  and  $g$  to be the same function. Since we can set  $f(y) = g(x) = c \in \mathbb{R}$ , such a  $\Phi$  exists — namely,  $\Phi(x, y) = \frac{1}{2}x^2y^2$ .

Consider  $\mathbf{B} = x^2y\hat{i} = xy^2\hat{j}$ . We know such a  $\Phi$  must not exist.

$$\begin{aligned}\frac{\partial \Phi}{\partial x} &= x^2y \\ \Phi &= \frac{1}{3}x^3y + f(y) \\ \frac{\partial \Phi}{\partial y} &= -xy^2 \\ \Phi &= -\frac{1}{3}xy^3 + g(x).\end{aligned}$$

There cannot exist such  $f$  and  $g$  that makes these antiderivatives equal to each other.

**Example (Gravity).** A mass  $m$  has a gravitational field

$$\mathbf{g} = -Gm \frac{\hat{r}}{r^2}.$$

Conventionally, the gravitational potential is the negative of the line integral of  $\mathbf{g}$ , since we want  $\Phi(\infty) = 0$ , yielding

$$\begin{aligned}\Phi(r) &= - \int_{\infty}^r \mathbf{g} \cdot d\vec{\ell} \\ &= Gm \int_{\infty}^r \frac{1}{s^2} ds \\ &= -\frac{Gm}{r}.\end{aligned}$$

Consider a planet. A planet is a bunch of point masses, which means we orient our focus away from  $\mathbf{g}$  towards  $d\mathbf{g}$ , and  $m$  to  $dm = \rho d\tau$ .

However, instead of trying to work out this integral in vector form, we want to find a scalar field  $\Phi$ , then take  $\mathbf{g} = -\nabla\Phi$ .

Let's construct a planet that is a spherical shell at radius  $a$  with mass  $m$ . Then,  $\sigma = \frac{m}{4\pi a^2}$  (constant), with the area  $dA$  has mass  $dm = \sigma dA$ . Thus, the mass at distance  $\mathbf{r}$  is

$$d\mathbf{g} = -G \frac{(\mathbf{r} - \mathbf{s})}{\|\mathbf{r} - \mathbf{s}\|^3} \sigma dA.$$

To find the total gravitational field, we find

$$\mathbf{g}(\mathbf{r}) = -G \int_S \frac{\mathbf{r} - \mathbf{s}}{\|\mathbf{r} - \mathbf{s}\|^3} \sigma dA,$$

where the integral with respect to  $\mathbf{s}$  is over the surface.

This vector-valued integral kind of sucks,<sup>xviii</sup> so instead, we want to integrate

$$\Phi(\mathbf{r}) = -G \int_S \frac{1}{\|\mathbf{r} - \mathbf{s}\|} \sigma \, dA,$$

which is much easier to do. Notice that

$$\begin{aligned} \|\mathbf{r} - \mathbf{s}\|^2 &= \langle \mathbf{r} - \mathbf{s}, \mathbf{r} - \mathbf{s} \rangle \\ &= \|\mathbf{r}\|^2 + \|\mathbf{s}\|^2 - 2\|\mathbf{r}\|\|\mathbf{s}\|\cos\theta, \end{aligned}$$

meaning, with  $\|\mathbf{r}\| = r$  and  $\|\mathbf{s}\| = a$ ,

$$\|\mathbf{r} - \mathbf{s}\| = \sqrt{r^2 + a^2 - 2ra\cos\theta}.$$

Using the area element,  $dA = a^2 d\Omega$ , or  $dA = 2\pi a^2 d(\cos\theta)$ , we get

$$\begin{aligned} \Phi(\mathbf{r}) &= -2\pi G\sigma a^2 \int_{-1}^1 \frac{1}{\sqrt{r^2 + a^2 - 2ra\cos\theta}} d(\cos\theta) \\ &= \frac{-2\pi G\sigma a}{r} ((r+a) - |r-a|). \end{aligned}$$

We need to evaluate  $\Phi$  for  $r < a$  and for  $r > a$ .

$$\begin{aligned} \Phi(r < a) &= -4\pi G\sigma a \\ &= -\frac{Gm}{a} \\ \Phi(r > a) &= -\frac{4\pi G\sigma a^2}{r} \\ &= -\frac{Gm}{r}. \end{aligned}$$

Notice that inside the shell, the gravitational potential is constant, while outside the shell, the gravitational potential falls off as if the shell has its mass concentrated at a point in the center.

When we convert the spherical shell into a ball (by integrating over  $a$ ), we can see that, outside the planet, it is still the case that gravitational potential depends solely on the distance to the center of the sphere.

## Flux

While circulation is the measure of the “swirl” of a vector field about a curve, we are also interested in the “surge” of a vector field about a surface. This is what flux is.

**Definition (Flux).** The flux is the product of the field’s magnitude and surface area  $A$  that it crosses.

$$\Phi \equiv EA.$$

A surface is identified with its orientation,  $\hat{n}$ , meaning

$$\Phi = \mathbf{E} \cdot \hat{n} A.$$

If  $\hat{n}$  is not constant along a curve’s surface, flux is calculated with an integral.

$$\Phi = \int_S \mathbf{E} \cdot \hat{n} \, da$$

---

<sup>xviii</sup>Citation needed.

$$= \int_S \mathbf{E} \cdot d\mathbf{a}.$$

If  $S$  is a closed surface,<sup>xix</sup> the flux is known as the net flux, and is denoted

$$\Phi_{\text{net}} = \oint_S \mathbf{E} \cdot d\mathbf{a}.$$

**Example** (Continuity Equation). The conservation of electric charge is expressed by continuity — the rate at which charge is lost in a surface,  $Q_{\text{in}}$  must be accounted for by the current flowing out of the surface,  $I_{\text{out}}$ .

$$I_{\text{out}} = -\frac{dQ_{\text{in}}}{dt}.$$

We define the charge density  $\rho$  as

$$Q_{\text{in}} = \int_V \rho \, d\tau,$$

and the movement of charge density with average drift velocity  $\mathbf{v}$  gives the current density,  $\mathbf{J} = \rho\mathbf{v}$ . Since  $\mathbf{J}$  has units of current per area, the usual current is measured by

$$I = \int_S \mathbf{J} \cdot d\mathbf{a}.$$

In particular, continuity is concerned with net flux, meaning

$$I_{\text{out}} = \oint_S \mathbf{J} \cdot d\mathbf{a},$$

so the continuity equation is

$$\oint_S \mathbf{J} \cdot d\mathbf{a} = -\frac{d}{dt} \int_V \rho \, d\tau,$$

where  $S$  bounds the volume  $V$ .

**Example** (Gauss's Law). Gauss's Law states that the net flux about an enclosed surface is proportional to the charge enclosed.

$$\oint_S \mathbf{E} \cdot d\mathbf{a} = \frac{q_{\text{encl}}}{\epsilon_0}.$$

We would find it weird to find  $\mathbf{E}$  using Gauss's law, unless the symmetry works heavily in our favor.

Consider a uniform ball with positive charge  $Q_0$  and radius  $a$ . Since the charge is spherically symmetric,  $\mathbf{E}$  must be radial, implying  $\mathbf{E} = E(r)\hat{r}$ .

With this in mind, we can choose a surface  $S$  such that  $\mathbf{E}$  is parallel to  $d\mathbf{a} = \hat{n}da$ ,<sup>xx</sup> and of constant magnitude on  $S$ .

For  $r > a$ , our enclosed charge is

$$\oint_S \mathbf{E} \cdot d\mathbf{a} = \oint_S E(r)\hat{r} \cdot \hat{r} \, da$$

---

<sup>xix</sup>A surface with no boundary.

<sup>xx</sup>For a sphere,  $\hat{n} = \hat{r}$

$$\begin{aligned}
&= E(r) \oint_S da \\
&= 4\pi r^2 E(r) \\
&= \sigma 4\pi a^2 \\
E &= \frac{Q_0}{4\pi\epsilon r^2} \\
\mathbf{E} &= \frac{Q_0}{4\pi\epsilon_0 r^2} \hat{r}.
\end{aligned}$$

If our Gaussian surface is inside the sphere, then with our charge density  $\rho = \frac{Q_0}{4/3\pi a^3}$ , implying

$$\begin{aligned}
q_{\text{encl}} &= \rho \frac{4}{3}\pi r^3 \\
&= Q_0 \frac{r^3}{a^3}.
\end{aligned}$$

Thus, inside the ball, we get

$$\mathbf{E} = \frac{Q_0}{4\pi\epsilon_0} \frac{r}{a^3} \hat{r}.$$

**Recall:** A vector field  $\mathbf{E}$  is path independent if

$$\begin{aligned}
\oint_C \mathbf{E} \cdot d\vec{\ell} &= 0 \\
\mathbf{E} &= \nabla \Phi \\
\nabla \times \mathbf{E} &= 0.
\end{aligned}$$

This is also known as an irrotational vector field.

**Example (Surface Independence).** Let  $\mathbf{E} = xy^2\hat{i} + x^2y\hat{j}$  and  $\mathbf{B} = x^2y\hat{i} - xy^2\hat{j}$ . We will examine the flux of these fields about different surfaces.

**Example (Flux of  $\mathbf{E}$  about  $S_1$ ).** We define  $S_1$  to be a quarter circle of radius  $a$  in the  $yz$ -plane, defined at  $x = b$ . To define  $d\mathbf{a}$ , we define the orientation to be in the  $\hat{i}$  direction, meaning  $d\mathbf{a} = \hat{i} dydz$ . Our integral becomes

$$\int_{S_1} \mathbf{E} \cdot d\mathbf{a} = \int_{S_1} xy^2 dydz.$$

We don't want to do this integral in cartesian coordinates, so we do a change of coordinates

$$\begin{aligned}
\int_{S_1} \mathbf{E} \cdot d\mathbf{a} &= \int_{S_1} xy^2 dydz \\
&= \int_{S_1} xy^2 r dr d\phi, \\
&= b \int_{S_1} y^2 dr d\phi \\
&= b \int_0^a r^3 dr \int_0^{\pi/2} \cos^2 \phi d\phi \\
&= \frac{\pi}{16} a^4 b.
\end{aligned}$$

**Example (Flux of  $\mathbf{E}$  about  $S_2$ ).** Let  $S_2$  be a quarter cylinder of radius  $a$  in the  $yz$ -plane, with  $b < x < 2b$ . We are integrating over four faces — the rectangle in the  $xz$ -plane, the rectangle in the  $xy$ -plane, the

quarter-circle at  $x = 2b$  in the  $yz$ -plane, and the radial face between  $x = b$  and  $x = 2b$ .

The rectangle in the  $xy$ -plane has its orientation in the  $-\hat{k}$  direction, meaning  $\mathbf{E} \cdot d\mathbf{a}_1 = 0$ , and similarly, the rectangle in the  $xz$ -plane has its orientation in the  $-\hat{j}$  direction, meaning  $\mathbf{E} \cdot d\mathbf{a}_2 = 0$ . The quarter circle cap at  $x = 2b$  has its orientation in the positive  $x$  direction — in cylindrical coordinates, this becomes  $d\mathbf{a}_3 = \hat{i} \rho d\rho d\phi$ . On the radial face, we have  $d\mathbf{a}_4 = \hat{\rho} a d\phi dx$ .

$$\begin{aligned} \int_{S_2} \mathbf{E} \cdot d\mathbf{a} &= \sum_i \int_{S_i} \mathbf{E} \cdot d\mathbf{a}_i \\ &= 0 + 0 + \int_{S_3} xy^2 \rho d\rho d\phi + \int_{S_4} \mathbf{E} \cdot \hat{\rho} a d\phi dx \\ &= \frac{\pi}{8} a^2 b \left( a^2 + \frac{14}{3} b^2 \right). \end{aligned}$$

Note that in our previous two examples, the surface affected the flux.

**Example** (Flux of  $\mathbf{B}$  about  $S_1$ ).

$$\int_{S_1} \mathbf{B} \cdot d\mathbf{a} = \frac{1}{3} a^3 b^2$$

**Example** (Flux of  $\mathbf{B}$  about  $S_2$ ).

$$\begin{aligned} \int_{S_2} \mathbf{B} \cdot d\mathbf{a} &= \int_{S_3} xy^2 \rho d\rho d\phi + \int_{S_4} \mathbf{B} \cdot \hat{\rho} a d\phi dx \\ &= \frac{1}{3} a^3 b^2. \end{aligned}$$

In this case, we can see that  $\mathbf{B}$  has the same flux integral. Later, we will see that  $\mathbf{B}$  is “surface independent,” in the sense that since both  $S_1$  and  $S_2$  have the same boundary, the flux integral of  $\mathbf{B}$  only depends on the boundary.

**Definition** (Surface Independence). Let  $\mathbf{B}$  be a vector field. If, for any closed surface  $S$ ,

$$\oint_S \mathbf{B} \cdot d\mathbf{a} = 0,$$

then we say  $\mathbf{B}$  is surface independent.

For a surface-independent field, the integral about an arbitrary surface must be the same through any surface bounded by the same closed curve, implying the existence of a vector potential  $\mathbf{A}$  such that

$$\int_S \mathbf{B} \cdot d\mathbf{a} = \oint_C \mathbf{A} \cdot d\vec{\ell},$$

where  $C$  is the boundary of  $S$  — note that this equivalence only holds with respect to a particular orientation. By convention, we use the positive orientation on  $C$ , which implies that  $\hat{n}$  has outward orientation on  $S$ .

Since  $\mathbf{A}$  is a vector potential, we can yield another field  $\mathbf{F} = \nabla \times \mathbf{A}$ . Thus, we get Stokes’s Theorem

$$\int_S \nabla \times \mathbf{A} \cdot d\mathbf{a} = \oint_C \mathbf{A} \cdot d\vec{\ell}.$$

Notice that the divergence of a curl is zero, so we can see that  $\nabla \cdot \mathbf{F} = 0$ .

**Definition** (Three Equivalent Criteria for Surface Independence). Let  $\mathbf{F}$  be a surface independent vector field. Then, the following are equivalent:

$$\begin{aligned}\oint_S \mathbf{F} \cdot d\mathbf{a} &= 0 \\ \mathbf{F} &= \nabla \times \mathbf{A} \\ \nabla \cdot \mathbf{F} &= 0.\end{aligned}$$

**Example** (Finding  $\mathbf{A}$ ). Let  $\mathbf{B}$  be such that  $\nabla \mathbf{A} = \mathbf{B}$ . Then,

$$\mathbf{B} = \begin{cases} x^2 y = \partial_y A_z - \partial_z A_y \\ -xy^2 = \partial_z A_x - \partial_x A_z \\ 0 = \partial_x A_y - \partial_y A_x \end{cases}.$$

If we let  $A_x = A_y = 0$ , then we solve the equation by taking

$$\begin{aligned}x^2 y &= \partial_y A_z \\ xy^2 &= \partial_x A_z,\end{aligned}$$

yielding

$$\begin{aligned}A_z &= \frac{1}{2}x^2 y^2 + \alpha(x, z) \\ A_z &= \frac{1}{2}x^2 y^2 + \beta(y, z).\end{aligned}$$

For consistency, it must be the case that  $\alpha(x, z) = \beta(y, z) = \gamma(z)$ . Thus, we get

$$\mathbf{A}_1 = \left( \frac{1}{2}x^2 y^2 + \gamma(z) \right) \hat{\mathbf{k}}.$$

We can also find

$$\mathbf{A}_2 = \left( -xy^2 z + \beta(x, y) \right) \hat{\mathbf{i}} + \left( -x^2 y z + \alpha(x, y) \right) \hat{\mathbf{j}},$$

whose curl is also  $\mathbf{B}$ .

Note that this means  $\mathbf{A}$  is not unique. In particular, for any  $\mathbf{A}$ , we can add any gradient field  $\nabla \chi$ , since  $\nabla \times \nabla \chi = 0$ .

## Calculating Flux Integrals with Boundary Geometry

Recall the fundamental theorem of calculus:

$$\int_a^b \frac{df}{dx} dx = f(b) - f(a).$$

In higher dimensions along a path  $C$  traveling from  $\mathbf{r}_1$  to  $\mathbf{r}_2$ , we have

$$\int_C \nabla f \cdot d\vec{\ell} = f(\mathbf{r}_2) - f(\mathbf{r}_1).$$

Notice that, via this process, we convert from a local property (derivatives) to a property defined by the boundary of a region (in this case, the endpoints of an interval and/or curve).



### Divergence Theorem

Consider an infinitesimal box with length  $dx$ , height  $dy$ , and depth  $dz$ . We can see that the flux of  $\mathbf{E}$  about these faces is

$$d\Phi_{\text{net}} = \sum_i \mathbf{E} \cdot d\mathbf{a}_i.$$

Evaluating  $\mathbf{E}$  at the faces on the  $yz$  plane parallel to the  $x$  axis, we get

$$\mathbf{E}(0, 0, 0) \cdot d\mathbf{a}_{\text{left}} + \mathbf{E}(dx, 0, 0) \cdot d\mathbf{a}_{\text{right}} = -E_x(0, 0, 0)dydz + E_x(dx, 0, 0)dydz.$$

The  $x$  component of  $d\Phi_{\text{net}}$  is, thus,

$$\begin{aligned} d\phi_1 &= (E_x(dx, 0, 0) - E_x(0, 0, 0))dydz \\ &= \left( \frac{\partial E_x}{\partial x} dx \right) dydz. \end{aligned}$$

Similarly,

$$\begin{aligned} d\phi_2 &= \left( \frac{\partial E_y}{\partial y} dy \right) dzdx \\ d\phi_3 &= \left( \frac{\partial E_z}{\partial z} dz \right) dxdy. \end{aligned}$$

Therefore,

$$\begin{aligned} d\Phi_{\text{net}} &= \left( \frac{\partial E_x}{\partial x} dx \right) dydz + \left( \frac{\partial E_y}{\partial y} dy \right) dzdx + \left( \frac{\partial E_z}{\partial z} dz \right) dxdy \\ &= (\nabla \cdot \mathbf{E}) d\tau, \end{aligned}$$

where  $d\tau = dxdydz$  is the volume element. Thus, we see that the infinitesimal flux is proportional to the divergence of  $\mathbf{E}$ .

To convert from infinitesimal boxes to a full (bounded) region, we cut up  $V$  into a bunch of small boxes. The interior box faces' respective fluxes cancel each other out since they have opposite-facing normals.

**Theorem:** Let  $V$  be a compact surface in  $\mathbb{R}^3$  with piecewise smooth boundary  $\partial V$ , and  $\mathbf{E}$  a continuously differentiable vector field defined on a neighborhood of  $V$ . Then, the following is true

$$\int_V \nabla \cdot \mathbf{E} d\tau = \oint_{\partial V} \mathbf{E} \cdot d\mathbf{a}.$$

**Example.** Let  $\mathbf{E} = r^\alpha \hat{r}$ , and  $V$  be a sphere of radius  $R$ . Since we have spherical symmetry, we can find the flux of  $\mathbf{E}$  about the sphere by taking

$$\begin{aligned} \oint_S \mathbf{E} \cdot d\mathbf{a} &= \oint_S r^\alpha R^2 d\Omega \\ &= 4\pi R^{\alpha+2}. \end{aligned}$$

Using the divergence theorem, we take  $\nabla \cdot \mathbf{E} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 E_r) = (\alpha + 2) r^{\alpha-1}$ , and take

$$\begin{aligned} \int_V \nabla \cdot \mathbf{E} d\tau &= (\alpha + 2) \int_0^R r^{\alpha-1} r^2 dr d\Omega \\ &= 4\pi R^{\alpha+2}. \end{aligned}$$

Notice that at  $\alpha = -2$ , this expression appears to fail — the divergence is zero, but the flux is clearly not zero.

**Example.** Let  $\mathbf{E} = r^\alpha \hat{\mathbf{r}}$ , and  $V$  be the region between a sphere of radius  $R_1$  and radius  $R_2$ , with  $R_1 < R_2$ .

To find the flux about  $\mathbf{E}$ , we find

$$\begin{aligned}\int_{\partial V} \mathbf{E} \cdot d\mathbf{a} &= \oint_{S_2} \mathbf{E} \cdot d\mathbf{a} + \oint_{S_1} \mathbf{E} \cdot d\mathbf{a} \\ &= 4\pi \left( R_2^{\alpha+2} - R_1^{\alpha+2} \right).\end{aligned}$$

Similarly, we get

$$\int_V \nabla \cdot \mathbf{E} d\tau = 4\pi \left( R_2^{\alpha+2} - R_1^{\alpha+2} \right).$$

Calculating for  $\alpha = -2$ , we know that

$$\begin{aligned}\int_V \nabla \cdot \frac{\hat{\mathbf{r}}}{r^2} d\tau &= \oint_S \frac{\hat{\mathbf{r}} \cdot \hat{\mathbf{n}}}{r^2} d\mathbf{a} \\ &= \oint_{S_1} \frac{\hat{\mathbf{r}} \cdot \hat{\mathbf{n}}_1}{r^2} d\mathbf{a} + \underbrace{\oint_{S_2} \frac{\hat{\mathbf{r}} \cdot \hat{\mathbf{n}}_2}{r^2} d\mathbf{a}}_{-\oint_{S_1} \frac{\hat{\mathbf{r}} \cdot \hat{\mathbf{n}}_1}{r^2} d\mathbf{a}} \\ &= 0.\end{aligned}$$

Thus,

$$\begin{aligned}\oint_{S_2} \frac{\hat{\mathbf{r}}}{r^2} d\mathbf{a} &= \oint_{S_1} \frac{\hat{\mathbf{r}}}{r^2} d\mathbf{a} \\ &= \oint_{S_1} \frac{R_1^2}{R_1^2} d\Omega \\ &= 4\pi.\end{aligned}$$

Thus, the integral over  $R_2$  has *no* dependence on  $R_2$ , and (as long as  $R_1 \neq 0$ ), we can change  $R_1$  however we like. If we let  $R_1 \rightarrow 0$ , we find

$$\int_V \left( \nabla \cdot \frac{\hat{\mathbf{r}}}{r^2} \right) d\tau = 4\pi.$$

Since the divergence is 0 everywhere except the origin, this means

$$\nabla \cdot \frac{\hat{\mathbf{r}}}{r^2} = 4\pi\delta(\mathbf{r})$$

Thus,

$$\begin{aligned}\Phi_{\text{net}} &= \oint_S \mathbf{E} \cdot d\mathbf{a} \\ &= \oint_S \frac{R^2}{R^2} d\Omega \\ &= 4\pi \\ &= \frac{1}{\epsilon_0} Q_{\text{encl}},\end{aligned}$$

meaning we get back Gauss's law:

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_0} \frac{\hat{\mathbf{r}}}{r^2}.$$

**Remark:** Gauss's law for magnetism has

$$\begin{aligned}\Phi_{\text{net}} &= \oint_S \mathbf{B} \cdot d\mathbf{a} \\ &= 0.\end{aligned}$$

This means that  $\nabla \cdot \mathbf{B} = 0$  at all points (rather than yielding the Dirac delta distribution), or that there are no magnetic monopoles. All magnetic fields must form closed loops.

### Stokes's Theorem

The divergence theorem allows us to move from flux (which is defined on the boundary of a region) and the divergence (defined on the interior). Similarly, we want to find some similar relation for circulation.

Consider a field  $\mathbf{B}$  and an infinitesimal rectangular curve with length  $dx$  and width  $dy$ . Then,

$$d\Psi = \sum_i \mathbf{B} \cdot d\vec{\ell}_i.$$

Looking at the bottom and top segments, we find

$$\mathbf{B}(0, 0, 0) \cdot d\vec{\ell}_{\text{bottom}} + \mathbf{B}(0, dy, 0) \cdot d\vec{\ell}_{\text{top}} = B_x(0, 0, 0) dx - B_x(0, dy, 0) dx,$$

where the minus sign appears from integrating right to left along the top segment. Thus, we get

$$\begin{aligned}d\psi_1 &= (-B_x(0, dy, 0) + B_x(0, 0, 0)) dx \\ &= -\left(\frac{\partial B_x}{\partial y} dy\right) dx.\end{aligned}$$

Similarly, for left and right segments, we get

$$\begin{aligned}d\psi_2 &= (B_y(dx, 0, 0) - B_y(0, 0, 0)) dy \\ &= \left(\frac{\partial B_y}{\partial x} dx\right) dy.\end{aligned}$$

Thus, we get

$$\begin{aligned}d\Psi &= \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right) dx dy \\ &= \left(\frac{B_y}{x} - \frac{\partial B_x}{\partial y}\right) dA,\end{aligned}$$

where  $dA = dx dy$  is the area element.

To build up to a simply connected region, split the region into rectangles. The interior infinitesimal segments cancel each other out, yielding Green's Theorem:

**Theorem:** Let  $C$  be a piecewise smooth simple closed curve in  $\mathbb{R}^2$ , and let  $S$  be the region in  $\mathbb{R}^2$  that is bounded by  $C$ . Then, for a continuously differentiable vector field  $\mathbf{B}$ ,

$$\oint_C \mathbf{B} \cdot d\vec{\ell} = \int_S \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right) dx dy,$$

Notice that

$$\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} = (\nabla \times \mathbf{B})_z.$$

With this in mind, we can move to three dimensions using the curl, yielding Stokes's Theorem.

**Theorem:** Let  $S$  be a smooth oriented surface with boundary  $C = \partial S$ . Then, for a continuously differentiable vector field  $\mathbf{B}$ ,

$$\oint_C \mathbf{B} \cdot d\vec{\ell} = \int_S \nabla \times \mathbf{B} \cdot d\mathbf{a}.$$

**Example.** Let  $\mathbf{B} = x^2y\hat{i} - xy^2\hat{j}$ .

We calculated the circulation from  $P_1 = (0,0)$  and  $P_2 = (1,1)$ , via  $C_1$ , defined by the straight line from  $(0,0)$  to  $(1,0)$  to  $(1,1)$ , to be

$$\int_{C_1} \mathbf{B} \cdot d\vec{\ell} = -\frac{1}{3},$$

and the circulation along  $C_2$ , which is a straight line from  $(0,0)$  to  $(1,1)$ , to be

$$\int_{C_2} \mathbf{B} \cdot d\vec{\ell} = 0.$$

Thus, the circulation along the path  $C_2 - C_1$  is  $\frac{1}{3}$ .

Verifying with Green's Theorem, we find  $\nabla \times \mathbf{B} = -(x^2 + y^2)\hat{k}$ , and take

$$\begin{aligned} \int_S \nabla \times \mathbf{B} \cdot d\mathbf{a} &= \int_S \left( -(x^2 + y^2)\hat{k} \right) \cdot (-\hat{k}) \, dx dy \\ &= \int_0^1 dx \int_0^x (x^2 + y^2) \, dy \\ &= \frac{4}{3} \int_0^1 x^3 \, dx \\ &= \frac{1}{3}. \end{aligned}$$

**Example.** Let  $\mathbf{B} = r^\alpha \hat{\phi}$ , and consider a circle of radius  $R$ .

Since our field is axially symmetric, we evaluate the circulation to find

$$\begin{aligned} \int_C \mathbf{B} \cdot d\vec{\ell} &= \oint_C r^\alpha R \, d\phi \\ &= R^{\alpha+1} \int_0^{2\pi} d\phi \\ &= 2\pi R^{\alpha+1}. \end{aligned}$$

Calculating the curl, since the field only has a  $\hat{\phi}$  component, we get

$$\begin{aligned} \nabla \times \mathbf{B} &= \frac{1}{r} \frac{\partial}{\partial r} (r B_\phi) \hat{z} \\ &= (\alpha + 1) r^{\alpha-1} \hat{z}, \end{aligned}$$

meaning over the circular disc bounded by  $C$ , we have

$$\begin{aligned} \int_S (\nabla \times \mathbf{B}) \cdot d\mathbf{a} &= (\alpha + 1) \int_S r^\alpha \, dr d\phi \\ &= 2\pi R^{\alpha+1}. \end{aligned}$$

Notice that we seem to be missing a curl here — if  $\alpha = -1$ , one side of Stokes's theorem says  $\int_S (\nabla \times \mathbf{B}) \cdot d\mathbf{a} = 0$ , but the circulation is not zero.

**Example.** Consider a field of  $\mathbf{B} = -\hat{k}$ , and we want to know the flux about a hemispherical surface of radius  $R$  bounded by a curve  $C$  in the  $xy$ -plane.

We can integrate about the hemisphere,

$$\begin{aligned}\int_{S_1} \mathbf{B} \cdot d\mathbf{a} &= \int_S (\hat{r} \cos \theta - \hat{\theta} \sin \theta) \cdot \hat{r} R^2 d\Omega \\ &= -2\pi R^2 \int_0^1 \cos \theta d(\cos \theta) \\ &= -\pi R^2.\end{aligned}$$

We can also use the boundary integral. In particular, since  $\nabla \times \mathbf{B} = 0$ , we can find  $\mathbf{A} = -y\hat{i}$ . Now, we can evaluate it as the circulation of  $\mathbf{A}$ .

Notice that  $d\vec{\ell} = \hat{\phi} R d\phi$ , meaning we only need  $A_\phi = -y \sin \phi$ ; since  $y = R \sin \theta \sin \phi$  and  $\theta = \pi/2$ , we have

$$\begin{aligned}\int_C \mathbf{A} \cdot d\vec{\ell} &= -\oint_C y \sin \phi R d\phi \\ &= -\pi R^2.\end{aligned}$$

However, the easiest way to do this integral is by selecting our surface to be the disk in the plane bounded by  $C$ . Thus,

$$\begin{aligned}\int_{S_2} \mathbf{B} \cdot d\mathbf{a} &= -\int \hat{k} \cdot \hat{k} r dr d\phi \\ &= -\pi R^2.\end{aligned}$$

## The Fundamental Theorem of Calculus, Revisited

The fundamental theorem of calculus says

$$f(P_2) - f(P_1) = \int_C \nabla f \cdot d\vec{\ell},$$

which expresses the accumulation of  $\nabla f \cdot d\vec{\ell}$  along the one-dimensional curve  $C$  via its expression over the (zero-dimensional) boundary. Similarly, the divergence theorem says

$$\oint_S \mathbf{E} \cdot d\mathbf{a} = \int_V \nabla \cdot \mathbf{E} d\tau,$$

which expresses the accumulation of  $\nabla \cdot \mathbf{E} d\tau$  over the three-dimensional volume via its expression over the two-dimensional boundary. Similarly,

$$\oint_C \mathbf{B} \cdot d\vec{\ell} = \int_S (\nabla \times \mathbf{B}) \cdot d\mathbf{a}$$

expresses the accumulation of  $(\nabla \times \mathbf{B}) \cdot d\mathbf{a}$  over the two-dimensional surface via its expression on the one-dimensional boundary  $C$ .

**Example (Integration By Parts).** Recall that integration by parts is the inverse of the product rule

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

meaning

$$\int_a^b f(x) \frac{dg}{dx} dx = f(x)g(x)|_a^b - \int_a^b \frac{df}{dx} g(x) dx.$$

However, we can do this for higher dimensions too, since the fundamental theorem of calculus still holds.

$$\nabla \cdot (f\mathbf{E}) = f(\nabla \cdot \mathbf{E}) + (\nabla f) \cdot \mathbf{E},$$

meaning

$$\int_V f(\nabla \cdot \mathbf{E}) \, d\tau = \oint_S f\mathbf{E} \cdot d\mathbf{a} - \int_V (\nabla f) \cdot \mathbf{E} \, d\tau,$$

which yields a boundary term.

Similarly,

$$\nabla \times (f\mathbf{B}) = f(\nabla \times \mathbf{B}) + (\nabla f) \times \mathbf{B},$$

meaning

$$\int_S f(\nabla \times \mathbf{B}) \cdot d\mathbf{a} = \oint_C f\mathbf{B} \cdot d\vec{\ell} - \int_S ((\nabla f) \times \mathbf{B}) \cdot d\mathbf{a}.$$

**Example** (Boundary of the Boundary). The fundamental theorem of calculus for an arbitrary field  $f$  in relation to its infinitesimal change  $df$  integrated over  $\Omega$  can be expressed by the generalized Stokes's theorem (which is the fundamental theorem of calculus):

$$\int_{\partial\Omega} f = \int_{\Omega} df.$$

When we take the integral of the second derivative, we get

$$\begin{aligned} \int_{\Omega} d^2f &= \int_{\partial\Omega} df \\ &= \int_{\partial^2\Omega} f. \end{aligned}$$

Since there is no boundary of a boundary, this means the integral is

$$= 0.$$

Consider a field  $\mathbf{E} = \nabla\Phi$ . Using generalized Stokes's theorem

$$\begin{aligned} \int_S (\nabla \times \mathbf{E}) \cdot d\mathbf{a} &= \int_S \nabla \times (\nabla\Phi) \cdot d\mathbf{a} \\ &= \oint_C \nabla\Phi \cdot d\vec{\ell} \\ &= \Phi(P_2) - \Phi(P_1) \\ &= 0. \end{aligned}$$

Similarly, for  $\mathbf{B} = \nabla \times \mathbf{A}$ , using the generalized Stokes theorem, we get

$$\begin{aligned} \int_V \nabla \times \mathbf{B} \, d\tau &= \int_V \nabla \times (\nabla \times \mathbf{A}) \, d\tau \\ &= \oint_S (\nabla \times \mathbf{A}) \cdot d\mathbf{a} \\ &= \oint_C \mathbf{A} \cdot d\vec{\ell} \\ &= 0. \end{aligned}$$

Since these hold for *any* arbitrary region in  $\mathbb{R}^3$ , we get the previously shown results that

$$\nabla \times \nabla\Phi = 0$$

$$\nabla \cdot (\nabla \times \mathbf{A}) = 0.$$

## Helmholtz Decomposition

In electrostatics, we have

$$\Phi(\mathbf{r}) = \int \frac{\rho(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} d^3x$$

$$\mathbf{E} = -\nabla\Phi.$$

However, this integral is hard to calculate,<sup>xxi</sup> so we focus on trying to find the behavior of the field.

The Helmholtz decomposition basically tells us that, if the field goes to zero “fast enough,” then we can reconstruct the field from the divergence and the curl.

We start with a generic field,

$$\mathbf{F}(\mathbf{r}) = \int_V \mathbf{F}(\mathbf{x}) \delta(\mathbf{r} - \mathbf{x}) d^3x,$$

where  $V$  includes  $\mathbf{r}$ . The identity

$$\nabla \cdot \left( \nabla \left( \frac{1}{|\mathbf{r} - \mathbf{x}|} \right) \right) = -4\pi\delta(\mathbf{r} - \mathbf{x})$$

allows us to substitute into the original expression of  $\mathbf{F}$ .

$$\begin{aligned} \mathbf{F}(\mathbf{r}) &= -\frac{1}{4\pi} \int_V \mathbf{F}(\mathbf{x}) \nabla^2 \left( \frac{1}{|\mathbf{r} - \mathbf{x}|} \right) d^3x \\ &= -\frac{1}{4\pi} \nabla^2 \int_V \frac{\mathbf{F}(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} d^3x, \end{aligned}$$

which we are allowed to do<sup>xxii</sup> since  $\nabla^2$  acts on  $\mathbf{x}$  rather than  $\mathbf{r}$ .

The integrand can be rewritten by taking  $\nabla^2 \mathbf{v} = \nabla(\nabla \cdot \mathbf{v}) - \nabla \times (\nabla \times \mathbf{v})$ , yielding

$$\begin{aligned} \mathbf{F}(\mathbf{r}) &= -\nabla \underbrace{\left( \frac{1}{4\pi} \nabla \cdot \int_V \frac{\mathbf{F}(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} d^3x \right)}_{\Phi} + \nabla \times \underbrace{\left( \frac{1}{4\pi} \nabla \times \int_V \frac{\mathbf{F}(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} d^3x \right)}_{\mathbf{A}}, \\ &= -\nabla\Phi + \nabla \times \mathbf{A}. \end{aligned}$$

Now that we have completed the decomposition, we can reconstruct our field. Starting with  $\Phi$ , we take

$$\begin{aligned} \Phi(\mathbf{r}) &= \frac{1}{4\pi} \nabla \cdot \int_V \frac{\mathbf{F}(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} d^3x \\ &= \frac{1}{4\pi} \int_V \mathbf{F}(\mathbf{x}) \cdot \nabla \frac{1}{|\mathbf{r} - \mathbf{x}|} d^3x \\ &= -\frac{1}{4\pi} \int_V \mathbf{F}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} \frac{1}{|\mathbf{r} - \mathbf{x}|} d^3x \end{aligned}$$

where we move the  $\mathbf{r}$  derivative in the integral, and use the technique that

$$\nabla \frac{1}{|\mathbf{r} - \mathbf{x}|} = -\nabla_{\mathbf{x}} \frac{1}{|\mathbf{r} - \mathbf{x}|},$$

<sup>xxi</sup>Citation needed.

<sup>xxii</sup>Well, we need a little more, but we’re physicists so everything is assumed to be nice.

to move the derivative to the  $\mathbf{x}$ . Now, we can use integration by parts to take

$$\Phi(\mathbf{r}) = -\frac{1}{4\pi} \oint_S \frac{\mathbf{F}(\mathbf{x}) \cdot d\mathbf{a}}{|\mathbf{r} - \mathbf{x}|} + \frac{1}{4\pi} \int_V \frac{\nabla_{\mathbf{x}} \cdot \mathbf{F}(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} d^3x.$$

Using a similar technique for  $\mathbf{A}$ , we get

$$\mathbf{A}(\mathbf{r}) = -\frac{1}{4\pi} \oint_S \frac{d\mathbf{a} \times \mathbf{F}(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} + \frac{1}{4\pi} \int_V \frac{\nabla_{\mathbf{x}} \times \mathbf{F}(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} d^3x.$$

Thus, assuming these integrals converge, these expressions allow us to *uniquely* determine  $\mathbf{F}$  from its curl and divergence.

## Maxwell's Equations

Maxwell's equations (see Table 17) help move beyond Coulomb's Law and the Biot-Savart law by showing that electricity and magnetism are connected to each other once time is involved.

### Integrating Maxwell

We saw that we can derive the expression for conservation of charge by the continuity equation:

$$\oint_S \mathbf{J} \cdot d\mathbf{a} = -\frac{d}{dt} \int_V \rho d\tau,$$

where  $\rho$  is the charge density and  $\mathbf{J}$  is the current density. The continuity equation states that the charge that disappears from a volume is equal to the flux of the charge.

Since Maxwell's equations lay claim to being able to describe everything with respect to electric and magnetic fields (as they include both a divergence and curl), we see if we can derive the continuity equation from Maxwell's equations.

Only one of Maxwell's equations uses current: Ampère's law. In a time-independent field, we have

$$\oint_C \mathbf{B} \cdot d\vec{\ell} = \mu_0 \int_S \mathbf{J} \cdot d\mathbf{a}.$$

Assuming that both  $\mathbf{B}$  and  $\mathbf{J}$  are continuous, we can shrink the bounding curve to a single point, meaning  $\mathbf{S}$  becomes a closed surface, yielding

$$\oint_S \mathbf{J} \cdot d\mathbf{a} = 0.$$

This is what we expected for time-independence.

However, when we deal with a time-dependent system, we need a displacement current,  $I_{\text{disp}} = \epsilon_0 \frac{d}{dt} \int_S \mathbf{E} \cdot d\mathbf{a}$ . Thus, our modified Ampère's law for a closed surface yields

$$\begin{aligned} 0 &= \oint_S \mathbf{B} \cdot \vec{\ell} \\ &= \mu_0 \oint_S \mathbf{J} \cdot d\mathbf{a} + \underbrace{\mu_0 \epsilon_0 \frac{d}{dt} \oint_S \mathbf{E} \cdot d\mathbf{a}}_{\frac{Q_{\text{encl}}}{\epsilon_0}} \\ \oint_S \mathbf{J} \cdot d\mathbf{a} &= -\frac{dQ}{dt}. \end{aligned}$$



## From Integrals to Derivatives

**Derivation** (Gauss's Law). Let's start with Gauss's Law:

$$\oint_S \mathbf{E} \cdot d\mathbf{a} = \frac{Q_{\text{encl}}}{\epsilon_0}.$$

Using the Divergence theorem, we rewrite the integral as

$$\oint_S \mathbf{E} \cdot d\mathbf{a} = \int_V \nabla \cdot \mathbf{E} d\tau$$

and, setting  $\rho$  to be the charge density,

$$\frac{Q_{\text{encl}}}{\epsilon_0} = \frac{1}{\epsilon_0} \int_V \rho d\tau.$$

Thus, we get

$$\int_V \nabla \cdot \mathbf{E} d\tau = \frac{1}{\epsilon_0} \int_V \rho d\tau.$$

Since this is true for arbitrary  $V$ , we can remove the integrals, and take

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}.$$

**Derivation** (Gauss's Law for Magnetism). Similarly, with Gauss's Law for Magnetism, we get

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

**Derivation** (Faraday's Law). Turning our attention to Faraday's Law, we take

$$\oint_C \mathbf{E} \cdot d\vec{\ell} = -\frac{d}{dt} \int_S \mathbf{B} \cdot d\mathbf{a}.$$

Using Stokes's theorem, we get

$$\int_S (\nabla \times \mathbf{E}) \cdot d\mathbf{a} = \int_S -\frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{a}.$$

This this holds for arbitrary  $S$ , we get

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

**Derivation** (Ampère's Law). Finally, in Ampère's law, we get

$$\oint_C \mathbf{B} \cdot d\vec{\ell} = \mu_0 I_{\text{encl}} + \mu_0 \epsilon_0 \frac{d}{dt} \int_S \mathbf{E} \cdot d\mathbf{a}.$$

Using the fact that  $I_{\text{encl}} = \int_S \mathbf{J} \cdot d\mathbf{a}$ , and Stokes's theorem, we get

$$\begin{aligned} \int_S (\nabla \times \mathbf{B}) \cdot d\mathbf{a} - \mu_0 \left( \int_S \mathbf{J} \cdot d\mathbf{a} + \epsilon_0 \int_S \frac{\partial \mathbf{E}}{\partial t} \cdot d\mathbf{a} \right) \\ = \mu_0 \int_S \left( \mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \cdot d\mathbf{a}. \end{aligned}$$

Since this holds for arbitrary  $S$ , we get

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

**Derivation** (Continuity Equation). If we take the divergence of Ampère's Law, we get

$$\begin{aligned}\nabla \cdot (\nabla \times \mathbf{B}) &= \mu_0 \nabla \cdot \mathbf{J} + \mu_0 \epsilon_0 \nabla \cdot \frac{\partial \mathbf{E}}{\partial t} \\ 0 &= \nabla \cdot \mathbf{J} + \epsilon_0 \frac{\partial}{\partial t} (\nabla \cdot \mathbf{E}) \\ \nabla \cdot \mathbf{J} &= -\frac{\partial \rho}{\partial t},\end{aligned}$$

which is exactly the continuity equation.

**Derivation** (Wave Equation). If we take the curl of Faraday's Law, we get

$$\begin{aligned}\nabla \times (\nabla \times \mathbf{E}) &= -\nabla \times \frac{\partial \mathbf{B}}{\partial t} \\ \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} &= -\frac{\partial}{\partial t} (\nabla \times \mathbf{B}) \\ &= -\frac{\partial}{\partial t} \left( \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right).\end{aligned}$$

We are now interested to see what happens in empty space — i.e., we want to see what happens in the homogeneous case (where there are no charges or currents).

In particular,  $\rho = 0$  and  $\mathbf{J} = 0$ .

$$\begin{aligned}-\nabla^2 \mathbf{E} &= -\frac{\partial}{\partial t} \left( \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \\ \nabla^2 \mathbf{E} &= \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}.\end{aligned}$$

This is the wave equation.

In particular,  $\frac{1}{\sqrt{\mu_0 \epsilon_0}} = c$ , so  $v = c$ . This is the connection between optics and electromagnetism.

### Substituting Potentials

In a static field, we get

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} \\ \nabla \times \mathbf{E} &= 0 \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{B} &= \mu_0 \mathbf{J}.\end{aligned}$$

In particular, this case means we can write  $\mathbf{E}$  as  $-\nabla\Phi$ , and  $\mathbf{B} = \nabla \times \mathbf{A}$  for some scalar potential  $\Phi$  and vector potential  $\mathbf{A}$ .

Rewriting Gauss's law in terms of  $\Phi$ , we get

$$\nabla^2 \Phi = -\frac{\rho}{\epsilon_0},$$

which is known as Poisson's equation.

Rewriting the time-independent Ampère's law in terms of  $\mathbf{A}$ , we get

$$\nabla^2 \mathbf{A} - \nabla (\nabla \cdot \mathbf{A}) = -\mu_0 \mathbf{J}.$$

When we add in time dependence, unfortunately we don't have  $\Phi$  anymore. However, we still have  $\mathbf{A}$ , so

$$\begin{aligned}\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ &= -\frac{\partial}{\partial t} (\nabla \times \mathbf{A}) \\ &= -\nabla \times \frac{\partial \mathbf{A}}{\partial t} \\ \nabla \times \left( \mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) &= 0.\end{aligned}$$

We redefine  $\mathbf{E}$  with this in mind to yield

$$\mathbf{E} = -\nabla \Phi - \frac{\partial \mathbf{A}}{\partial t},$$

meaning Gauss's law yields

$$\nabla^2 \Phi + \frac{\partial}{\partial t} (\nabla \cdot \mathbf{A}) = -\frac{\rho}{\epsilon_0}$$

and Ampère's Law yields

$$\nabla^2 \mathbf{A} - \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla \left( \nabla \cdot \mathbf{A} + \mu_0 \epsilon_0 \frac{\partial \Phi}{\partial t} \right) = -\mu_0 \mathbf{J}.$$

Note that for any particular  $\lambda(x, t)$ , we can make the transformation

$$\begin{aligned}\mathbf{A} &\mapsto \mathbf{A} + \nabla \lambda \\ \Phi &\mapsto \Phi - \frac{\partial \lambda}{\partial t}.\end{aligned}$$

In particular, we can choose a gauge such that  $\nabla \cdot \mathbf{A} = 0$ , which is known as the Coulomb gauge. We can also use the Lorenz gauge, which sets

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \Phi}{\partial t} = 0.$$

## Linear Algebra

### Superposition

Waves, unlike particles, are able to be superimposed — we can add waves together to get new waves.

Consider a charge with density  $\rho = \rho_1 + \rho_2$ . Then, each charge density yields  $\mathbf{E}_1$  and  $\mathbf{E}_2$ , where  $\nabla \cdot \mathbf{E}_1 + \nabla \cdot \mathbf{E}_2 = \nabla \cdot (\mathbf{E}_1 + \mathbf{E}_2) = \nabla \cdot \mathbf{E}$ .

Generally, for superposition to hold, we need an operator  $\mathcal{L}$  to be linear. For any inputs  $f$  and  $g$  and  $a, b \in \mathbb{F}$ , we must have

$$\mathcal{L}(af + bg) = a\mathcal{L}(f) + b\mathcal{L}(g).$$

**Example** (Some Linear and Nonlinear Operators). (1) Consider  $f(x) = mx + b$  with  $b \neq 0$ . Then,  $f$  is a linear function, but  $f$  is not a linear operator, since

$$\begin{aligned}f(ax_1 + bx_2) &= m(ax_1 + bx_2) + b \\ &\neq af(x_1) + bf(x_2).\end{aligned}$$

We say  $f$  is an affine transformation of  $\mathbb{R}$  (or  $\mathbb{C}$ ).

(2) In finite dimensions, a linear operator can always be expressed as a matrix.

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$y_i = \sum_j M_{ij} x_j.$$

(3) The linear operator

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

is a projection. In particular,  $P$  maps any vector  $v$  to its  $x$  and  $y$  components. The projection operator is idempotent —  $P^2 = P$ .

Note that “magnitude and direction” follows from the idea of a vector.

## Vector Spaces

### Essentials of Vector Spaces

A vector space is a free module generated over a field. The generating set is known as a basis.

Unfortunately, physicists don’t care about that stuff so we have to go the long way to define a vector space.

**Definition** (Properties of a Vector Space). Let  $\mathbf{A}$  and  $\mathbf{B}$  be vectors and  $a, b \in \mathbb{F}$ . Our vector space must be closed under addition. Then, we must have

- Commutativity:  $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ .
- Associativity:  $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$ .
- Additive Inverse:  $\mathbf{A} + (-\mathbf{A}) = 0$ .
- Distributivity of Scalar Multiplication:  $a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}$ .
- Associativity of Scalar Multiplication:  $(ab)\mathbf{A} = a(b\mathbf{A})$ .

**Example** (Understanding  $\mathbb{R}^3$ ). • Arrows in  $\mathbb{R}^3$  are a vector space.

- Vectors in  $\mathbb{R}^3$  with one vanishing component *are not* a vector space. However, vectors in  $\mathbb{R}^3$  with the same vanishing component *are* a vector space (also known as a subspace) — this vector space is isomorphic to  $\mathbb{R}^2$ .
- A plane that slices through the origin in  $\mathbb{R}^3$  is a subspace, and any line that slices through the origin in  $\mathbb{R}^3$  is also a subspace.

### Touching Base(s)

The primary circumstance under which  $a\mathbf{A} + b\mathbf{B} = 0$  is if  $\mathbf{B} = -\left(\frac{a}{b}\right)\mathbf{A}$ , meaning  $\mathbf{A}$  and  $\mathbf{B}$  are parallel.

If the only way  $a\mathbf{A} + b\mathbf{B} = 0$  is if  $a = b = 0$ , then  $\mathbf{A}$  and  $\mathbf{B}$  are independent of each other.

**Definition.** A collection  $\{\mathbf{A}_i\}_{i=1}^k$  of vectors is called linearly independent if

$$\sum_{i=1}^k c_i \mathbf{A}_i = 0$$

if and only if  $c_i = 0$  for all  $i = 1, \dots, k$ .

**Definition** (Spanning Set). A collection  $\{\mathbf{B}_j\}_{j=1}^m$  is said to be spanning if for any vector  $\mathbf{A}$ , there exist  $c_1, \dots, c_m$  such that

$$\mathbf{A} = \sum_{j=1}^m c_j \mathbf{B}_j.$$

**Definition** (Basis). A collection of vectors  $\{\mathbf{e}_i\}_{i=1}^n$  is a basis if it is both linearly independent and spanning for the vector space.

The dimension of a vector space is equal to the cardinality of the basis.

The basis  $\{\mathbf{e}_i\}_{i=1}^n$  “supports” any vector in a set. For any vector  $\mathbf{A}$ , then

$$\mathbf{A} = \sum_{j=1}^n c_j \mathbf{e}_j.$$

We say the  $c_j$  are the components of  $\mathbf{A}$  on the basis  $\{\mathbf{e}_i\}_{i=1}^n$ .

**Example** (Standard Basis). Consider the set of  $n \times 1$  arrays,

$$\mathbf{A} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

This set obeys the necessary rules for the column vectors. We define the standard basis for this set by

$$\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

where the 1 is at the  $i$ th position. We will write  $\mathbf{e}_i = \hat{\mathbf{e}}_i$ .

We can write

$$\mathbf{A} = \sum_{i=1}^n a_i \hat{\mathbf{e}}_i.$$

Note that the expansion coefficients are the elements of the column vector  $\mathbf{A}$ .

**Example** (Non-standard Basis). Let

$$\begin{aligned} \epsilon_1 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ \epsilon_2 &= \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \\ \epsilon_3 &= \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \end{aligned}$$

Then,  $\{\epsilon_1, \epsilon_2, \epsilon_3\}$  forms a basis.

We have

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \left( \frac{a_1 + a_2 + a_3}{3} \right) \epsilon_1 + \left( \frac{a_1 - 2a_2 + a_3}{6} \right) \epsilon_2 + \left( \frac{a_1 - a_3}{2} \right) \epsilon_3.$$

$$= \sum_{i=1}^3 c_i \epsilon_i.$$

**Example (Matrix Basis).** Consider  $\text{Mat}_2(\mathbb{C})$ , the set of all  $2 \times 2$  matrices over  $\mathbb{C}$ . Then, the set

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

is linearly independent. Also,

$$\mathcal{S} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

$$= \{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$$

forms a basis for  $\text{Mat}_2(\mathbb{C})$ . The set  $\sigma_1, \sigma_2, \sigma_3$  are known as the Pauli matrices.

**Example (Vector Spaces of Polynomials).** Consider the set of polynomials with degree  $n$ ,  $\mathcal{P}_n$ . An element of  $\mathcal{P}_n$  is written

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

This is *not* closed under addition, since we can subtract two polynomials to yield a polynomial with degree  $n - 1$ . However, the set of polynomials with degree *at most*  $n$  is closed under  $n$ .

A basis for  $\mathcal{P}_n$  is  $\mathcal{B} = \{1, x, x^2, \dots, x^n\}$ .

**Example (Vector Spaces of Periodic Functions).** Consider the set of functions  $f$  such that  $f(x + T) = f(x)$  for some fixed period  $T$ . This *is* a vector space, since we can add two functions with the same period to yield another function with that same period.

Finding a basis is much more difficult. However, we know that sine and cosine have a periodicity. In particular,  $\sin(2\pi n/T)$  and  $\cos(2\pi n/T)$  can form a Schauder basis for the set of functions with period  $T$ .

$$f(t) = \sum_{n=0}^{\infty} a_n \cos(2\pi n t/T) + b_n \sin(2\pi n t/T).$$

However, when we have infinitely many dimensions, we run into issues of convergence, which we will deal with in a future part of this course.<sup>xxiii</sup>

## Kets and Reps

Since vectors are not universally arrows (as vectors are elements of vector spaces), we want to find a universal way to represent vectors that is not  $\mathbf{v}$  or  $\vec{v}$ .

**Definition (Ket).** Let  $V$  be a vector space. Any element of  $V$  is written as  $|v\rangle$ .

<sup>xxiii</sup>Presumably very poorly, because this isn't a math class.

For a basis  $\{|e_i\rangle\}_{i \in I}$ , we abbreviate it as  $|i\rangle$ .

We express  $|v\rangle$  on a basis as follows.

$$\begin{aligned} |v\rangle &= \sum_{i=1}^n v_i |i\rangle \\ &= v_1 |1\rangle + v_2 |2\rangle + \cdots + v_n |n\rangle. \end{aligned}$$

We can express addition of  $|v\rangle$  and  $|w\rangle$  on a common basis as follows.

$$\begin{aligned} |v\rangle + |w\rangle &= \sum_{i=1}^n v_i |i\rangle + w_i |i\rangle \\ &= \sum_{i=1}^n (v_i + w_i) |i\rangle \\ &= |v + w\rangle. \end{aligned}$$

Scalar multiplication works as follows

$$\begin{aligned} c |v\rangle &= \sum_{i=1}^n c v_i |i\rangle \\ &= |c v\rangle \end{aligned}$$

Note that we use 0 to denote the zero vector in  $V$ , rather than  $|0\rangle$ , which we use to denote a concrete basis vector.

**Example (Representing Vectors).** Let  $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$  be a four-dimensional basis. Let

$$|P\rangle = 2|0\rangle - |1\rangle + 3|2\rangle + |3\rangle.$$

We can let  $|P\rangle$  be abstract, or create a concrete representation. For instance, if we take

$$|i\rangle = x^i$$

for each  $i \in \{0, 1, 2, 3\}$ , then

$$P(x) = x^3 + 3x^2 - x + 2.$$

However, if we take  $|i\rangle = \sigma_i$ , where  $\sigma_i$  are the Pauli matrices, then we get

$$|P\rangle = \begin{pmatrix} 3 & -1 - 3i \\ -1 + 3i & 1 \end{pmatrix}.$$

We can also take  $|i\rangle = e_{i+1}$ , where  $e_j$  represents the standard basis in  $\mathbb{C}^4$ . This representation yields

$$|P\rangle = \begin{pmatrix} 2 \\ -1 \\ 3 \\ 1 \end{pmatrix}.$$

## Inner Products

We have yet to define the “multiplication” of vectors. In order to do this, taking  $|A\rangle$  and  $|B\rangle$ , we let  $\langle A|B\rangle$  be the inner product on  $|A\rangle$  and  $|B\rangle$ . Note that  $\langle A|B\rangle$  is a scalar.

We will delay the formal definition of an inner product; we will start by discussing a more intuitive definition of inner products.

## Adjoins

First, let's try to figure out why  $\langle A|B \rangle \neq |A\rangle \cdot |B\rangle$  generally.

In  $\mathbb{R}^2$ , however,  $\langle A|B \rangle = |A\rangle \cdot |B\rangle$ . We represent

$$\mathbf{A} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

and

$$\begin{aligned} \langle A|B \rangle &= \mathbf{A}^T \mathbf{B} \\ &= (a_1 \quad a_2) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\ &= \sum_i a_i b_i. \end{aligned}$$

Note that we need to take the transpose of  $\mathbf{A}$ . This seems trivial in  $\mathbb{R}^n$ , since  $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$  over the real numbers.

However, over  $\mathbb{C}^n$ , things become slightly different. Note that the norm

$$\begin{aligned} \|\mathbf{A}\| &= \langle A|A \rangle \\ &\geq 0 \end{aligned}$$

for all  $\mathbf{A} \in \mathbb{C}^n$ . Using the traditional transpose on  $\mathbf{A}$ , we would get

$$\begin{aligned} \mathbf{A}^T \mathbf{A} &= (a_1 \quad \cdots \quad a_n) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \\ &= \sum_{i=1}^n a_i^2, \end{aligned}$$

which is not necessarily in  $[0, \infty)$  (let alone  $\mathbf{A}^T \mathbf{A}$  is a norm).

Therefore, we need to take the conjugate of  $\mathbf{A}^T$ ,  $\mathbf{A}^{*\text{xxiv}}$  in order to use the inner product.

$$\begin{aligned} \langle A|A \rangle &= \overline{(a_1 \quad \cdots \quad a_n)} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \\ &= \sum_{i \in I} |a_i|^2. \end{aligned}$$

In the general case, we take

$$\begin{aligned} \langle A|B \rangle &= \mathbf{A}^* \mathbf{B} \\ &= \sum_i \overline{a_i} b_i. \end{aligned}$$

---

<sup>xxiv</sup>The physicists use  $\dagger$  instead of  $*$  to denote the conjugate transpose, but they're wrong. I will also begin using  $\bar{z}$  to denote the complex conjugate now.



We let  $A^*$  be the Hermitian adjoint for  $A$ .

Note that, with this definition, we now have

$$\begin{aligned}\langle A|B\rangle &= \sum_i \overline{a_i} b_i \\ &= \overline{\left( \sum_i a_i \overline{b_i} \right)} \\ &= \overline{\langle B|A\rangle},\end{aligned}$$

meaning the inner product is also *conjugate symmetric*. Additionally,  $A^{**} = A$ .<sup>xxv</sup>

Note that since  $\langle A|B\rangle = \overline{\langle B|A\rangle}$  is an equality, this means the inner product is basis-independent.

In matrices, we have

$$M_{ij}^* = \overline{M_{ji}}.$$

**Example.** Let

$$\begin{aligned}v &= \begin{pmatrix} -1 \\ 2 \end{pmatrix} \\ w &= \begin{pmatrix} 1 \\ -i \end{pmatrix}.\end{aligned}$$

Then,

$$\begin{aligned}\langle v|w\rangle &= -1 - 2i \\ \langle w|v\rangle &= -1 + 2i.\end{aligned}$$

**Example.** Let

$$\begin{aligned}S &= \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \\ T &= \begin{pmatrix} 1 & 2i \\ -i & 1 \end{pmatrix}.\end{aligned}$$

Then,

$$\begin{aligned}S^* &= \begin{pmatrix} -i & 1 \\ -1 & i \end{pmatrix} \\ T^* &= \begin{pmatrix} 1 & i \\ -2i & 1 \end{pmatrix}.\end{aligned}$$

We can see that

$$\begin{aligned}(ST)^* &= \begin{pmatrix} -2i & 2 \\ -3 & -3i \end{pmatrix} \\ &= T^* S^*.\end{aligned}$$

**Definition** (Dirac Notation). Let  $V$  be a vector space.

We denote any vector in  $V$  by  $|v\rangle$ , and its adjoint,  $|v\rangle^* = \langle v|$  the unique linear functional such that

$$|v\rangle^* (|w\rangle) = \langle v|w\rangle.$$

Note that for any  $|v\rangle \in V$ ,  $\langle v| \in V^*$  necessarily exists by the Riesz representation theorem.

<sup>xxv</sup>We refer to the  $*$  operation as an involution.

**Example** (Vector Space of Kets and Bras). Note that we have

$$|v + w\rangle = |v\rangle + |w\rangle,$$

and similarly,

$$\langle v + w| = \langle v| + \langle w|.$$

However, unlike in kets, where we have

$$|cv\rangle = c|v\rangle,$$

for bras, we have

$$\begin{aligned}\langle cv| &= |cv\rangle^* \\ &= \bar{c}|v\rangle^* \\ &= \bar{c}\langle v|.\end{aligned}$$

We can generalize to application by any linear operator.

**Example** (Linear Operators on Kets and Bras). We can take  $\mathcal{L}$  to be some linear operator, giving

$$|w\rangle = \mathcal{L}|v\rangle.$$

When it comes to bras, though, we get

$$\begin{aligned}\langle w| &= |w\rangle^* \\ &= (\mathcal{L}|v\rangle)^* \\ &= |v\rangle^* \mathcal{L}^* \\ &= \langle v| \mathcal{L}^*.\end{aligned}$$

Similarly, we can define

$$\begin{aligned}|w\rangle &= |\mathcal{L}v\rangle \\ &= \mathcal{L}|v\rangle,\end{aligned}$$

and

$$\begin{aligned}\langle w| &= \langle \mathcal{L}v| \\ &= \langle v| \mathcal{L}^*.\end{aligned}$$

**Definition** (The Adjoint through Inner Products). Let  $\mathcal{L}$  be a linear operator. The adjoint,  $\mathcal{L}^*$ , is the unique linear operator such that

$$\langle v|\mathcal{L}w\rangle = \langle \mathcal{L}^*v|w\rangle$$

or, equivalently,

$$\overline{\langle v|\mathcal{L}|w\rangle} = \langle w|\mathcal{L}^*|v\rangle$$

**Definition** (Hermitian Operator). A linear operator  $\mathcal{L}$  is known as Hermitian (or self-adjoint) if  $\mathcal{L}^* = \mathcal{L}$ .

**Example.** Let

$$v = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

$$w = \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

$$S = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}.$$

Then,

$$\begin{aligned} \overline{\langle v | S | w \rangle} &= \langle v | S w \rangle \\ &= \overline{\begin{bmatrix} -1 & 2 \end{bmatrix} \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix}} \\ &= 4 + 2i \\ \langle w | S^* | v \rangle &= \langle w | S^* v \rangle \\ &= \overline{\begin{pmatrix} 1 & -i \end{pmatrix} \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} -1 \\ 2 \end{pmatrix}} \\ &= 4 + 2i. \end{aligned}$$

**Example** (A Hermitian Operator). The matrix

$$A = \begin{pmatrix} 1 & i \\ -i & 3 \end{pmatrix}$$

is a Hermitian operator.

### Cauchy-Schwarz Inequality

**Definition** (Inner Product Space). An inner product space is a vector space with an additional operation which takes two vectors,  $|v\rangle$  and  $|w\rangle$  such that the following are true.

- (1) Positive definite norm:  $\langle v | v \rangle = \|v\|^2$ .
- (2) Conjugate symmetry:  $\langle v | w \rangle = \overline{\langle w | v \rangle}$ .
- (3) Linearity:  $\langle u | (a |v\rangle + b |w\rangle) = a \langle u | v \rangle + b \langle u | w \rangle$ .

Consider  $|A\rangle, |B\rangle$  be vectors, and let

$$\langle A | B \rangle = e^{i\phi} |\langle A | B \rangle|.$$

Define

$$|C\rangle = s e^{i\phi} |A\rangle + |B\rangle$$

for some  $s \in \mathbb{R}$ . Then, we have

$$\langle C | = s e^{-i\phi} \langle A | + \langle B |,$$

meaning

$$\begin{aligned} \langle C | C \rangle &= s^2 \langle A | A \rangle + s e^{-i\phi} \langle A | B \rangle + s e^{i\phi} \langle B | A \rangle + \langle B | B \rangle \\ &= s^2 \langle A | A \rangle + 2s \operatorname{Re} \left( e^{-i\phi} \langle A | B \rangle \right) + \langle B | B \rangle \\ &= s^2 \langle A | A \rangle + 2s |\langle A | B \rangle| + \langle B | B \rangle. \end{aligned}$$

Since  $\langle C | C \rangle \geq 0$  necessarily, it is the case that

$$|\langle A | B \rangle|^2 \leq \langle A | A \rangle \langle B | B \rangle.$$

**Definition** (Cauchy–Schwarz Inequality). Let  $V$  be an inner product space, and let  $A, B \in V$ . Then,

$$|\langle A|B \rangle| \leq \|A\| \|B\|.$$

**Example.** In  $\mathbb{C}^n$ , the Cauchy–Schwarz inequality becomes

$$\left| \sum_{i=1}^n \overline{a_i} b_i \right|^2 \leq \left( \sum_{i=1}^n |a_i|^2 \right) \left( \sum_{i=1}^n |b_i|^2 \right).$$

From the Cauchy–Schwarz inequality, we can find a limit on the norm of  $\mathbf{A} + \mathbf{B}$ .

$$\begin{aligned} \|\mathbf{A} + \mathbf{B}\|^2 &= \langle \mathbf{A} + \mathbf{B} | \mathbf{A} + \mathbf{B} \rangle \\ &= \langle \mathbf{A} | \mathbf{A} \rangle + \langle \mathbf{B} | \mathbf{B} \rangle + \langle \mathbf{A} | \mathbf{B} \rangle + \langle \mathbf{B} | \mathbf{A} \rangle \\ &= \langle \mathbf{A} | \mathbf{A} \rangle + \langle \mathbf{B} | \mathbf{B} \rangle + 2 \operatorname{Re}(\langle \mathbf{A} | \mathbf{B} \rangle) \\ &\leq \langle \mathbf{A} | \mathbf{A} \rangle + \langle \mathbf{B} | \mathbf{B} \rangle + 2 |\langle \mathbf{A} | \mathbf{B} \rangle| \\ &\leq \langle \mathbf{A} | \mathbf{A} \rangle + \langle \mathbf{B} | \mathbf{B} \rangle + 2 \|\mathbf{A}\| \|\mathbf{B}\| \\ &= (\|\mathbf{A}\| + \|\mathbf{B}\|)^2. \end{aligned}$$

Thus, we recover the triangle inequality.

**Definition** (Triangle Inequality). Let  $\mathbf{A}, \mathbf{B} \in V$ , where  $V$  is any normed vector space. Then,

$$\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|.$$

We define  $\cos \theta = \frac{|\langle \mathbf{A} | \mathbf{B} \rangle|}{\|\mathbf{A}\| \|\mathbf{B}\|}$ . The geometric interpretation of  $\langle \mathbf{A} | \mathbf{B} \rangle$  is, then, the projection of  $|\mathbf{B}\rangle$  onto  $|\mathbf{A}\rangle$ .

Note that the inner product, norm, and the Cauchy–Schwarz inequality are completely basis-independent.

### Orthonormal Sets and Bases

Now that we have shown that the inner product is basis-independent, we can select a basis to make our lives a lot easier. Specifically, we can select  $\{|\hat{e}_i\rangle\}_{i \in I}$  to be a (Schauder) basis such that

$$\langle \hat{e}_i | \hat{e}_j \rangle = \delta_{ij}.$$

In the case of the inner product between two vectors, this basis collapses to yield the standard inner product on  $\mathbb{C}^n$ :

$$\begin{aligned} \langle \mathbf{A} | \mathbf{B} \rangle &= \sum_{i,j} \overline{a_i} b_j \langle \hat{e}_i | \hat{e}_j \rangle \\ &= \sum_i \overline{a_i} b_i. \end{aligned}$$

For instance, the standard basis on  $\mathbb{C}^2$  is

$$\begin{aligned} \hat{e}_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \hat{e}_2 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

**Example** (A Non-Standard Basis). Let

$$|e_{1,2}\rangle = \frac{1}{\sqrt{2}} (\pm |\hat{e}_1\rangle + |\hat{e}_2\rangle).$$

Then,

$$\begin{aligned}
 \langle \epsilon_{1,2} | \epsilon_{1,2} \rangle &= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} (\pm \langle \hat{e}_1 | + \langle \hat{e}_2 |) (|\hat{e}_1\rangle + |\hat{e}_2\rangle) \\
 &= \frac{1}{2} (\langle \hat{e}_1 | \hat{e}_1 \rangle + \langle \hat{e}_1 | \hat{e}_2 \rangle + \langle \hat{e}_2 | \hat{e}_1 \rangle + \langle \hat{e}_2 | \hat{e}_2 \rangle) \\
 &= 1. \\
 \langle \epsilon_1 | \epsilon_2 \rangle &= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} (\langle \hat{e}_1 | + \langle \hat{e}_2 |) (-|\hat{e}_1\rangle + |\hat{e}_2\rangle) \\
 &= \frac{1}{2} (-\langle \hat{e}_1 | \hat{e}_1 \rangle + \langle \hat{e}_1 | \hat{e}_2 \rangle - \langle \hat{e}_2 | \hat{e}_1 \rangle + \langle \hat{e}_2 | \hat{e}_2 \rangle) \\
 &= 0.
 \end{aligned}$$

**Example** (Using an Orthonormal Basis). With the previous example, we have

$$\begin{aligned}
 |\hat{e}_1\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
 |\hat{e}_2\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.
 \end{aligned}$$

Consider the vectors

$$\begin{aligned}
 |v\rangle &= \begin{pmatrix} -1 \\ 2 \end{pmatrix} \\
 |w\rangle &= \begin{pmatrix} 1 \\ -i \end{pmatrix}.
 \end{aligned}$$

We wish to expand these vectors along the  $\{|\epsilon_1\rangle, |\epsilon_2\rangle\}$  basis.

To do this, we can project using the inner product's projection properties:

$$\begin{aligned}
 v_1 &= \langle \hat{e}_1 | v \rangle \\
 &= \frac{1}{\sqrt{2}} \overline{\begin{pmatrix} 1 & 1 \end{pmatrix}} \begin{pmatrix} -1 \\ 2 \end{pmatrix} \\
 &= \frac{1}{\sqrt{2}} \\
 v_2 &= \langle \hat{e}_2 | v \rangle \\
 &= \frac{1}{\sqrt{2}} \overline{\begin{pmatrix} -1 & 1 \end{pmatrix}} \begin{pmatrix} -1 \\ 2 \end{pmatrix} \\
 &= \frac{3}{\sqrt{2}} \\
 v &= \frac{1}{\sqrt{2}} |\hat{e}_1\rangle + \frac{3}{\sqrt{2}} |\hat{e}_2\rangle
 \end{aligned}$$

$$\begin{aligned}
 w_1 &= \langle \hat{e}_1 | w \rangle \\
 &= \frac{1}{\sqrt{2}} \overline{\begin{pmatrix} 1 & 1 \end{pmatrix}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \\
 &= (1 - i) \frac{1}{\sqrt{2}} \\
 w_2 &= \langle \hat{e}_2 | w \rangle
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sqrt{2}} \overline{(-1 \ 1)} \begin{pmatrix} 1 \\ -i \end{pmatrix} \\
&= (-1 - i) \frac{1}{\sqrt{2}} \\
w &= \frac{1-i}{\sqrt{2}} |\hat{e}_I\rangle + \frac{-1-i}{\sqrt{2}} |\hat{e}_{II}\rangle.
\end{aligned}$$

**Example** (Using a Complex-Valued Basis). Consider the basis

$$\begin{aligned}
|\hat{e}_I\rangle &= \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2i \end{pmatrix} \\
|\hat{e}_{II}\rangle &= \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -i \end{pmatrix}.
\end{aligned}$$

This is an orthonormal basis.

$$\begin{aligned}
\langle \hat{e}_I | \hat{e}_{II} \rangle &= \frac{1}{\sqrt{5}} \frac{1}{\sqrt{5}} (1 \ -2i) \begin{pmatrix} 2 \\ -i \end{pmatrix} \\
&= 0.
\end{aligned}$$

The components of a vector

$$|A\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

on this basis are

$$\begin{aligned}
a_I &= \langle \hat{e}_I | A \rangle \\
&= \frac{1}{\sqrt{5}} \overline{(1 \ 2i)} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\
&= \frac{a_1 - 2ia_2}{\sqrt{5}} \\
a_{II} &= \langle \hat{e}_{II} | A \rangle \\
&= \frac{1}{\sqrt{5}} \overline{(2 \ -i)} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\
&= \frac{2a_1 + ia_2}{\sqrt{5}} \\
A &= \left( \frac{a_1 - 2ia_2}{\sqrt{5}} \right) |\hat{e}_I\rangle + \left( \frac{2a_1 + ia_2}{\sqrt{5}} \right) |\hat{e}_{II}\rangle.
\end{aligned}$$

We can see that

$$\|A\|^2 = |a_1|^2 + |a_2|^2.$$

It can also be shown<sup>xxvi</sup> that

$$\begin{aligned}
\|A\|^2 &= |a_I|^2 + |a_{II}|^2 \\
&= |a_1|^2 + |a_2|^2.
\end{aligned}$$

---

<sup>xxvi</sup>I will not show this, though, because I don't want to.

**Example.** Consider the vectors

$$\begin{aligned} |\epsilon_1\rangle &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ |\epsilon_2\rangle &= \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \\ |\epsilon_3\rangle &= \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \end{aligned}$$

This is an orthogonal basis, but not an orthonormal basis. To normalize these vectors, we divide out by the magnitude.

$$\begin{aligned} |\hat{\epsilon}_1\rangle &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ |\hat{\epsilon}_2\rangle &= \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \\ |\hat{\epsilon}_3\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \end{aligned}$$

We project the arbitrary vector

$$A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

onto each basis element to obtain a new representation.

$$\begin{aligned} \langle \hat{\epsilon}_1 | A \rangle &= \frac{1}{\sqrt{3}} (a_1 + a_2 + a_3) \\ \langle \hat{\epsilon}_2 | A \rangle &= \frac{1}{\sqrt{6}} (a_1 - 2a_2 + a_3) \\ \langle \hat{\epsilon}_3 | A \rangle &= \frac{1}{\sqrt{2}} (a_1 - a_3). \end{aligned}$$

Thus, we can write

$$A = \frac{a_1 + a_2 + a_3}{\sqrt{3}} |\hat{\epsilon}_1\rangle + \frac{a_1 - 2a_2 + a_3}{\sqrt{6}} |\hat{\epsilon}_2\rangle + \frac{a_1 - a_3}{\sqrt{2}} |\hat{\epsilon}_3\rangle.$$

We can actually write in terms of the non-normalized basis, by dividing each  $|\hat{\epsilon}_i\rangle$  by its corresponding normalization constant.

$$A = \frac{a_1 + a_2 + a_3}{3} |\epsilon_1\rangle + \frac{a_1 - 2a_2 + a_3}{6} |\epsilon_2\rangle + \frac{a_1 - a_3}{2} |\epsilon_3\rangle$$

**Example (Pauli Basis).** Consider the three Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We can construct an inner product on these matrices. Specifically, we see that the Frobenius inner product<sup>xxvii</sup>

$$\begin{aligned} \langle \sigma_i | \sigma_j \rangle &= \text{tr} (\sigma_i^* \sigma_j) \\ &= 2\delta_{ij}, \end{aligned}$$

meaning that we need to redefine our inner product. Specifically, we take

$$\langle \sigma_i | \sigma_j \rangle = \frac{1}{2} \text{tr} (\sigma_i^* \sigma_j).$$

Note that with the Pauli matrices,  $\sigma_i^* = \sigma_i$ . If we add

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

we now form a complete orthonormal basis in  $\text{Mat}_2(\mathbb{C})$ .

If  $A$  is any complex-valued matrix, then

$$\begin{aligned} A &= \sum_{\alpha=0}^3 a_{\alpha} \sigma_{\alpha} \\ &= \sum_{\alpha=0}^3 \langle \sigma_{\alpha} | A \rangle \sigma_{\alpha} \\ &= \sum_{\alpha=0}^3 \frac{1}{2} \text{tr} (\sigma_{\alpha}^* A) \sigma_{\alpha}. \\ &= \frac{1}{2} \text{tr}(A) I_2 + \frac{1}{2} \sum_{i=1}^3 \text{tr} (\sigma_i^* A) \sigma_i. \end{aligned}$$

One of the interesting applications of the Pauli basis is that, using the Pauli basis, we can represent any  $\mathbf{v} \in \mathbb{R}^3$  and represent with the Pauli matrices.

$$\begin{aligned} V &= \sum_{i=1}^3 v_i \sigma_i \\ &= \sum_{i=1}^3 \langle \sigma_i | \mathbf{v} \rangle \sigma_i \\ &= \begin{pmatrix} v_z & v_x - i v_y \\ v_x + i v_y & -v_z \end{pmatrix}. \end{aligned}$$

Note that  $V = V^*$  and  $\text{tr}(V) = 0$ . Thus, using the Pauli basis, we can express any 3-dimensional vector as a trace-free Hermitian matrix.

---

<sup>xxvii</sup> $\langle A | B \rangle = \text{tr} (A^* B)$



### Constructing an Orthonormal Basis: Gram–Schmidt Decomposition

The ultimate question behind linear algebra is what basis to choose.

However, even if we are given a basis, it may not be an orthonormal basis.

**Method.** Consider a set of  $n$  linearly independent vectors in  $\mathbb{C}^n$ ,  $\{|\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_n\rangle\}$ . From these  $n$  vectors, we want to create a set of  $n$  orthonormal vectors.

Starting with  $|\phi_1\rangle$ , we normalize to make

$$|\hat{e}_1\rangle = \frac{|\phi_1\rangle}{\sqrt{\langle\phi_1|\phi_1\rangle}}.$$

To get  $|\hat{e}_2\rangle$ , we take  $|\phi_2\rangle$  and remove the component of  $|\phi_2\rangle$  along the “direction” of  $|\hat{e}_1\rangle$ . This gives us

$$|e_2\rangle = |\phi_2\rangle - \langle\hat{e}_1|\phi_2\rangle |\hat{e}_1\rangle.$$

Now, we normalize to get

$$|\hat{e}_2\rangle = \frac{|e_2\rangle}{\sqrt{\langle e_2|e_2\rangle}}.$$

Note that

$$\begin{aligned}\langle\hat{e}_1|\hat{e}_2\rangle &= \langle\hat{e}_1|\phi_2\rangle - \langle\hat{e}_1|\phi_2\rangle \langle\hat{e}_1|\hat{e}_1\rangle \\ &= 0.\end{aligned}$$

For  $|\hat{e}_3\rangle$ , we start by taking  $|\phi_3\rangle$ , removing the components along the  $|\hat{e}_1\rangle$  and  $|\hat{e}_2\rangle$  directions, giving

$$|e_3\rangle = |\phi_3\rangle - \langle\hat{e}_1|\phi_3\rangle |\hat{e}_1\rangle - \langle\hat{e}_2|\phi_3\rangle |\hat{e}_2\rangle.$$

Normalizing, we get

$$|\hat{e}_3\rangle = \frac{|e_3\rangle}{\sqrt{\langle e_3|e_3\rangle}}.$$

In the general form, we have

$$\begin{aligned}|e_m\rangle &= |\phi_m\rangle - \sum_{\ell=1}^{m-1} \langle\hat{e}_\ell|\phi_m\rangle |\hat{e}_\ell\rangle \\ |\hat{e}_m\rangle &= \frac{|e_m\rangle}{\sqrt{\langle e_m|e_m\rangle}}.\end{aligned}$$

**Example.** Consider the vectors

$$\begin{aligned}|\phi_1\rangle &= \begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix} \\ |\phi_2\rangle &= \begin{pmatrix} i \\ 0 \\ i \end{pmatrix} \\ |\phi_3\rangle &= \begin{pmatrix} 1 \\ i \\ -1 \end{pmatrix}.\end{aligned}$$

This set is linearly independent. We start with

$$|\hat{e}_1\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix}.$$

Then,

$$\begin{aligned} |e_2\rangle &= |\phi_2\rangle - \langle \hat{e}_1 | \phi_2 \rangle |\hat{e}_1\rangle \\ &= \begin{pmatrix} i \\ 0 \\ i \end{pmatrix} - \left[ \overline{\left( \frac{1}{\sqrt{3}} \right)} \begin{pmatrix} 1 & -i & 1 \end{pmatrix} \begin{pmatrix} i \\ 0 \\ i \end{pmatrix} \right] \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} i \\ 0 \\ i \end{pmatrix} - \frac{2i}{3} \begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} i \\ -2 \\ i \end{pmatrix}. \end{aligned}$$

Normalizing, we get

$$|\hat{e}_2\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} i \\ -2 \\ i \end{pmatrix}.$$

Finally, to get  $|\hat{e}_3\rangle$ , we take

$$\begin{aligned} |e_3\rangle &= |\phi_3\rangle - \langle \hat{e}_1 | \phi_3 \rangle |\hat{e}_1\rangle - \langle \hat{e}_2 | \phi_3 \rangle |\hat{e}_2\rangle \\ &= \begin{pmatrix} 1 \\ i \\ -1 \end{pmatrix} - \left[ \frac{1}{\sqrt{3}} \overline{\begin{pmatrix} 1 & -i & 1 \end{pmatrix}} \begin{pmatrix} 1 \\ i \\ -1 \end{pmatrix} \right] \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} - \left[ \frac{1}{\sqrt{6}} \overline{\begin{pmatrix} i & -2 & i \end{pmatrix}} \begin{pmatrix} 1 \\ i \\ -1 \end{pmatrix} \right] \begin{pmatrix} i \\ 3 \\ i \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \end{aligned}$$

Thus,

$$|\hat{e}_3\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

The orthonormal basis is, as follows:

$$\begin{aligned} |\hat{e}_1\rangle &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix} \\ |\hat{e}_2\rangle &= \frac{1}{\sqrt{6}} \begin{pmatrix} i \\ -2 \\ i \end{pmatrix} \\ |\hat{e}_3\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \end{aligned}$$

**Example** (Gram–Schmidt with Matrices). Let

$$|M_1\rangle = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$|M_2\rangle = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

$$|M_3\rangle = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

$$|M_4\rangle = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

These matrices are linearly independent, but it is not the case that

$$\begin{aligned} \langle M_i | M_j \rangle &= \frac{1}{2} \text{tr} (M_i^* M_j) \\ &\neq \delta_{ij}. \end{aligned}$$

Let

$$\begin{aligned} |\hat{e}_1\rangle &= \frac{|M_1\rangle}{\| |M_1\rangle \|} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \frac{1}{\sqrt{\frac{1}{2} \text{tr} (M_1^* M_1)}} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \end{aligned}$$

Then, we have

$$\begin{aligned} |e_2\rangle &= |M_2\rangle - \langle \hat{e}_1 | M_2 \rangle |\hat{e}_1\rangle \\ &= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} - \frac{1}{2} \text{tr} (\hat{e}_1^* M_2) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} - \frac{i}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{i}{2} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}. \end{aligned}$$

Normalizing, we get

$$\begin{aligned} |\hat{e}_2\rangle &= \frac{|e_2\rangle}{\| |e_2\rangle \|} \\ &= \frac{i}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}. \end{aligned}$$

Continuing, we get

$$\begin{aligned} |\hat{e}_3\rangle &= \frac{1-i}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ |\hat{e}_4\rangle &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \end{aligned}$$

## Completeness

The most important fact of a basis is that it can support any vector in  $V$ .

$$|v\rangle = \sum_{i=1}^n \langle \hat{e}_i | v \rangle |\hat{e}_i\rangle.$$

This is redundant in finite dimensions, but in infinite dimensions, we actually need our basis to support any vector.

Consider

$$\begin{aligned} |A\rangle &= \sum_{\ell} a_{\ell} |\hat{e}_{\ell}\rangle \\ &= \sum_{\ell} \langle \hat{e}_{\ell} | A \rangle |\hat{e}_{\ell}\rangle. \end{aligned}$$

Since  $\langle \hat{e}_{\ell} | A \rangle$  is a number, we can switch the order, yielding

$$\begin{aligned} &= \sum_{\ell} |\hat{e}_{\ell}\rangle \langle \hat{e}_{\ell} | A \rangle \\ &= \left( \sum_{\ell} |\hat{e}_{\ell}\rangle \langle \hat{e}_{\ell}| \right) |A\rangle \\ &= I_n |A\rangle. \end{aligned}$$

**Definition** (Completeness Relation). The completeness relation states

$$\sum_{\ell} |\hat{e}_{\ell}\rangle \langle \hat{e}_{\ell}| = \text{id}_V.$$

We let  $\mathcal{P}_{\ell} = |\hat{e}_{\ell}\rangle \langle \hat{e}_{\ell}|$  be the projection onto the subspace  $\text{span}(|\hat{e}_{\ell}\rangle)$ . Note that

$$\mathcal{P} = \sum_{\ell} \mathcal{P}_{\ell}$$

satisfies the definition of the projection operator,  $\mathcal{P}^2 = \mathcal{P}$ :

$$\begin{aligned} \mathcal{P}^2 &= \left( \sum_{\ell} |\hat{e}_{\ell}\rangle \langle \hat{e}_{\ell}| \right) \left( \sum_{m} |\hat{e}_m\rangle \langle \hat{e}_m| \right) \\ &= \sum_{\ell, m} |\hat{e}_{\ell}\rangle \langle \hat{e}_{\ell} | \hat{e}_m \rangle \langle \hat{e}_m| \\ &= \sum_{\ell, m} |\hat{e}_{\ell}\rangle \delta_{\ell m} \langle \hat{e}_m| \\ &= \sum_{\ell} |\hat{e}_{\ell}\rangle \langle \hat{e}_{\ell}|. \end{aligned}$$

We sometimes write  $|A\rangle \langle B|$  to be the outer product, which is the unique operator such that, for any  $C \in V$ , we have

$$(|A\rangle \langle B|) |C\rangle = |A\rangle \langle B | C \rangle.$$

**Example** (Completeness Relation in  $\mathbb{R}^3$ ). Consider

$$|\hat{e}_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$|\hat{e}_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$|\hat{e}_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Note that we have

$$|\hat{e}_1\rangle \langle \hat{e}_1| = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$|\hat{e}_2\rangle \langle \hat{e}_2| = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$|\hat{e}_3\rangle \langle \hat{e}_3| = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so the completeness relation is satisfied.

**Definition** (Complete Orthonormal Basis). A set  $\{|\hat{e}_i\rangle\}_{i \in I}$  of normalized vectors is said to be a complete orthonormal basis if

$$\langle \hat{e}_i | \hat{e}_j \rangle = \delta_{ij}$$

$$\sum_{i \in I} |\hat{e}_i\rangle \langle \hat{e}_i| = \text{id}.$$

Note that we get

$$|A\rangle = \text{id}_V |A\rangle$$

$$= \left( \sum_{\ell} |\hat{e}_{\ell}\rangle \langle \hat{e}_{\ell}| \right) |A\rangle$$

$$= \sum_{\ell} \langle \hat{e}_{\ell} | A \rangle |\hat{e}_{\ell}\rangle$$

$$= \sum_{\ell} a_{\ell} |\hat{e}_{\ell}\rangle.$$

Additionally, this allows us to decompose along any complete basis:

$$|A\rangle = \sum_{\ell} a_{\ell} |\hat{e}_{\ell}\rangle$$

$$= \sum_{\ell} a_{\ell} \text{id}_V |\hat{e}_{\ell}\rangle$$

$$= \sum_{\ell} a_{\ell} \left( \sum_{\mathbf{k}} |\hat{e}_{\mathbf{k}}\rangle \langle \hat{e}_{\mathbf{k}}| \right) |\hat{e}_{\ell}\rangle$$

$$= \sum_{\mathbf{k}, \ell} a_{\ell} \langle \hat{e}_{\mathbf{k}} | \hat{e}_{\ell} \rangle |\hat{e}_{\mathbf{k}}\rangle$$

$$= \sum_{\mathbf{k}} \underbrace{\left( \sum_{\ell} a_{\ell} \langle \hat{e}_{\mathbf{k}} | \hat{e}_{\ell} \rangle \right)}_{a'_{\mathbf{k}}} |\hat{e}_{\mathbf{k}}\rangle$$

$$= \sum_k a'_k |\hat{e}_k\rangle.$$

**Example (Quantum Mechanics).** Fundamental particles like protons, electrons, etc. have an intrinsic angular momentum, known as spin. This angular momentum has a value  $\frac{1}{2}\hbar$ , where  $\hbar = \frac{h}{2\pi}$  is the reduced Planck constant.

We represent

$$|\psi\rangle = a |\hat{e}_\uparrow\rangle + b |\hat{e}_\downarrow\rangle,$$

subject to the fact that

$$\begin{aligned} E_\psi(\uparrow) &= |\langle \hat{e}_\uparrow | \psi \rangle|^2 \\ E_\psi(\downarrow) &= |\langle \hat{e}_\downarrow | \psi \rangle|^2. \end{aligned}$$

However, since we have to measure *some* probability, we must have

$$\underbrace{|\langle \hat{e}_\uparrow | \psi \rangle|^2}_{|a|^2} + \underbrace{|\langle \hat{e}_\downarrow | \psi \rangle|^2}_{|b|^2} = \langle \psi | \psi \rangle = 1$$

On the  $z$  basis, we have

$$\begin{aligned} |\hat{e}_\uparrow\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |\hat{e}_\downarrow\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

However, on the  $x$  basis, we have

$$\begin{aligned} |\hat{e}_+\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ |\hat{e}_-\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned}$$

However, despite this basis change, since these are a complete orthonormal basis for  $\mathbb{C}^2$ , we must have

$$|\langle \hat{e}_\uparrow | \psi \rangle|^2 + |\langle \hat{e}_\downarrow | \psi \rangle|^2 = |\langle \hat{e}_+ | \psi \rangle|^2 + |\langle \hat{e}_- | \psi \rangle|^2.$$

### Matrix Representation of Operators

Consider a linear operator  $|W\rangle = \mathcal{M} |V\rangle$  with vectors  $|V\rangle$  and  $|W\rangle$ . We can project this onto a basis  $\{\hat{e}_i\}$  to find the relationship between the components.

$$\begin{aligned} w'_i &= \langle \hat{e}_i | W \rangle \\ &= \langle \hat{e}_i | \mathcal{M} | V \rangle \\ &= \sum_j \langle \hat{e}_i | \mathcal{M} | \hat{e}_j \rangle \langle \hat{e}_j | V \rangle \\ &= \sum_j M_{ij} v_j. \end{aligned}$$

This yields

$$\begin{pmatrix} w'_1 \\ w'_2 \\ \vdots \\ w'_n \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Note that the element  $M_{ij}$  requires *two* projections.

$$M_{ij} = \langle \hat{e}_i | \mathcal{M} | \hat{e}_j \rangle$$

However, consider a different basis, with

$$\begin{aligned} \mathcal{M} &= \mathcal{M}'_{ab} \\ &= \langle \hat{e}_a | \mathcal{M} | \hat{e}_b \rangle \\ &= \sum_{ij} \langle \hat{e}_a | \hat{e}_i \rangle \langle \hat{e}_i | \mathcal{M} | \hat{e}_j \rangle \langle \hat{e}_j | \hat{e}_b \rangle \\ &= \sum_{ij} \bar{a}_i M_{ij} b_j, \end{aligned}$$

where  $a_i = \langle \hat{e}_i | \hat{e}_a \rangle$  and  $b_i = \langle \hat{e}_i | \hat{e}_b \rangle$  are the projections of  $|\hat{e}_a\rangle$  and  $|\hat{e}_b\rangle$  onto the  $|\hat{e}_i\rangle$  basis respectively.

**Example.** Let  $\mathcal{M}$  be defined by

$$\begin{aligned} \langle \hat{e}_1 | \mathcal{M} | \hat{e}_1 \rangle &= 1 \\ \langle \hat{e}_2 | \mathcal{M} | \hat{e}_2 \rangle &= 1 \\ \langle \hat{e}_1 | \mathcal{M} | \hat{e}_2 \rangle &= 2 \\ \langle \hat{e}_2 | \mathcal{M} | \hat{e}_1 \rangle &= 2. \end{aligned}$$

In the standard basis, we get

$$\mathcal{M} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

On the  $|\hat{e}\rangle$  basis, we have

$$\begin{aligned} |\hat{e}_1\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ |\hat{e}_2\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned}$$

$\mathcal{M}$  has the matrix elements

$$\begin{aligned} M'_{11} &= \langle \hat{e}_1 | \mathcal{M} | \hat{e}_1 \rangle \\ &= \frac{1}{\sqrt{2}} \overline{\begin{pmatrix} 1 & 1 \end{pmatrix}} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= 3 \end{aligned}$$

$$\begin{aligned} M'_{22} &= \langle \hat{e}_1 | \mathcal{M} | \hat{e}_1 \rangle \\ &= \frac{1}{\sqrt{2}} \overline{\begin{pmatrix} 1 & -1 \end{pmatrix}} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= -1, \end{aligned}$$

and similarly,  $M'_{12} = 0$  and  $M'_{21} = 0$ . On the  $|\hat{e}\rangle$  basis, we have

$$M' = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}.$$

Similarly, if we have the basis

$$|\hat{e}_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$$

$$|\hat{e}_2\rangle = \frac{1}{\sqrt{2}} (1/-i),$$

we calculate

$$M''_{11} = \langle \hat{e}_1 | M | \hat{e}_1 \rangle$$

$$= \frac{1}{\sqrt{2}} \overline{\begin{pmatrix} 1 & i \end{pmatrix}} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$$

$$= 1$$

$$\vdots$$

$$M'' = \begin{pmatrix} 1 & -2i \\ 2i & 1 \end{pmatrix}.$$

Note that the following are true:

$$M = M^*$$

$$M' = (M')^*$$

$$M'' = (M'')^*,$$

and  $\text{tr}(M) = \text{tr}(M') = \text{tr}(M'')$ .

This is an example of the fact that hermiticity, trace, determinant, etc. are all basis-independent.

Some bases are better than others — specifically, for normal operators, we will be able to choose a basis of orthonormal eigenvectors, which will allow the operator to be realized as a diagonal matrix.

## Rotations

Rotations are important in physics.<sup>xxviii</sup> Therefore, we must study the groups of (continuous) rotations — and, specifically, their representations as matrices.

### Introduction to Rotations

Consider a two dimensional vector  $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ . Here, we define

$$v_1 = v \cos \alpha$$

$$v_2 = v \sin \alpha$$

$$v'_1 = v \cos \beta$$

$$v'_2 = v \sin \beta,$$

---

<sup>xxviii</sup>Citation needed.



subject to  $\beta = \alpha + \phi$ .

$$\begin{aligned} v'_1 &= v \cos(\alpha + \phi) \\ &= v \cos \alpha \cos \phi - v \sin \alpha \sin \phi \\ &= v_1 \cos \phi - v_2 \sin \phi \\ v'_2 &= v \sin(\alpha + \phi) \\ &= v \sin \alpha \cos \phi + v \cos \alpha \sin \phi \\ &= v_2 \cos \phi + v_1 \sin \phi. \end{aligned}$$

Thus, we can write

$$\begin{pmatrix} v'_1 \\ v'_2 \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Here, we denote

$$R(\phi) = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$$

to be the rotation matrix. Note that  $R^{-1}(\phi) = R(-\phi) = R^T(\phi)$ . The definition that  $R^{-1} = R^T$  is the crucial, defining property of rotations in  $\mathbb{R}^n$ .

Note that we rotated the vector  $\mathbf{v}$  to make  $\mathbf{v}'$ . However, we can also rotate the *axes*.

Rotating the vector is known as an “active” rotation, while rotating the axes is known as a “passive” rotation.

To execute a passive rotation, we need to use the rotation matrix on the basis vectors. This yields

$$\begin{pmatrix} |\hat{e}'_1\rangle \\ |\hat{e}'_2\rangle \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} |\hat{e}_1\rangle \\ |\hat{e}_2\rangle \end{pmatrix}.$$

Note that we are having the rotation matrix act on the “vector” of basis vectors to yield a vector of basis vectors.

### What is a Rotation?

Consider successive rotations,  $\phi_1 + \phi_2$ . Then,

$$R(\phi_1 + \phi_2) = \begin{pmatrix} \cos(\phi_1 + \phi_2) & -\sin(\phi_1 + \phi_2) \\ \sin(\phi_1 + \phi_2) & \cos(\phi_1 + \phi_2) \end{pmatrix}.$$

Intuitively, we must have

$$\begin{aligned} R(\phi_1 + \phi_2) &= R(\phi_2) R(\phi_1) \\ &= \begin{pmatrix} \cos \phi_2 & -\sin \phi_2 \\ \sin \phi_2 & \cos \phi_2 \end{pmatrix} \begin{pmatrix} \cos \phi_1 & -\sin \phi_1 \\ \sin \phi_1 & \cos \phi_1 \end{pmatrix} \\ &= \begin{pmatrix} \cos \phi_2 \cos \phi_1 - \sin \phi_2 \sin \phi_1 & -\sin \phi_1 \cos \phi_2 - \sin \phi_2 \cos \phi_1 \\ \sin \phi_1 \cos \phi_2 + \cos \phi_1 \sin \phi_2 & \cos \phi_2 \cos \phi_1 - \sin \phi_1 \sin \phi_2 \end{pmatrix} \end{aligned}$$

Additionally, the rotation of doing nothing is

$$R(\phi = 0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= I_2.$$

Additionally, we must have

$$R^{-1}(\phi) = R(-\phi).$$

Thus, we have established that rotations indeed form a group.

**Example** (Rotations in Real Space). By the definition of a rotation, we must have

$$\|\mathbf{v}\|^2 = \sum_i v_i^2.$$

remain constant. Thus, we have

$$\begin{aligned} \sum_i v_i v_i &= \sum_i v'_i v'_i \\ &= \sum_i \left( \sum_k R_{ik} v_k \right) \left( \sum_\ell R_{i\ell} v_\ell \right) \\ &= \sum_{i,k,\ell} R_{ik} R_{i\ell} v_k v_\ell \\ &= \sum_{k,\ell} \left( \sum_i R_{ki}^T R_{i\ell} \right) v_k v_\ell \end{aligned}$$

Thus, we must have

$$\sum_i R_{ki}^T R_{i\ell} = \delta_{k\ell},$$

so

$$R^T R = I_2,$$

meaning

$$R^T = R^{-1}.$$

These are known as the orthogonal matrices.

The orthogonal matrices form a group,  $O(n)$ . In two dimensions, the rotations are (part of) the group  $O(2)$ .

**Example** (Rotations in Complex Space). Using a similar derivation to the case of real space, we are able to determine the criterion for a rotation in complex space, yielding a unitary matrix

$$\begin{aligned} U^* U &= U U^* \\ &= I. \end{aligned}$$

where  $U^*$  is the adjoint of  $U$ .

We can also obtain this result by taking

$$\begin{aligned} \langle v' | w' \rangle &= \langle Uv | Uw \rangle \\ &= \langle v | U^* U | w \rangle. \\ &= \langle v | w \rangle \end{aligned}$$

Thus, we must have  $U^* U = I$ .

**Example** (Rotations in  $\mathbb{R}^2$  and Rotations in  $\mathbb{C}$ ). We know that  $\mathbb{C}$  is isomorphic to  $\mathbb{R}^2$  as  $\mathbb{R}$ -vector spaces. Additionally, for

$$\begin{aligned} z_1 &= r_1 e^{i\phi_1} \\ z_2 &= r_2 e^{i\phi_2}, \end{aligned}$$

we have

$$z_1 z_2 = r_1 r_2 e^{i(\phi_1 + \phi_2)}.$$

Complex multiplication is, essentially, a rotation.

If we take

$$z = a + ib,$$

we can represent it as

$$= \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

We need to verify that this is a faithful representation.

$$\begin{aligned} z_1 + z_2 &= \begin{pmatrix} a_1 + a_2 & -b_1 - b_2 \\ b_1 + b_2 & a_1 + a_2 \end{pmatrix} \\ z_1 z_2 &= \begin{pmatrix} a_1 a_2 - b_1 b_2 & -a_1 b_2 - a_2 b_1 \\ a_1 b_2 + a_2 b_1 & a_1 a_2 - b_1 b_2 \end{pmatrix}. \end{aligned}$$

Since these are faithful representations, we can see that, for the circle group  $\mathbb{T}$  consisting of all complex numbers with modulus 1, we have

$$\mathbb{T} \cong \text{SO}(2).$$

In other words, complex numbers consisting of phase correspond one to one to rotations.

**Example** (Rotations in  $\mathbb{C}^2$ ). Consider the matrix

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

This matrix has 8 free parameters. However, by definition of unitary, we must have

$$\begin{aligned} U^* U &= I \\ &= \begin{cases} |a|^2 + |c|^2 = 1 \\ |b|^2 + |d|^2 = 1 \\ \bar{a}b + \bar{c}d = 0 \end{cases}. \end{aligned}$$

Applying these constraints, now we have 4 constraints. We have to add one more constraint — since  $\det U = e^{i\alpha}$  for any phase  $\alpha$ , and as  $U$  is a rotation, we must have  $\det U = ad - bc = 1$ .

We can summarize the necessary constraints as follows:

$$U = \begin{pmatrix} a & -b \\ \bar{b} & \bar{a} \end{pmatrix}, \quad |a|^2 + |b|^2 = 1$$

**Example** (Using the Complex Rotations). Let

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

Then,

$$\begin{aligned} \xi' &= U\xi \\ &= \begin{pmatrix} a & -b \\ \bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \\ &= \begin{pmatrix} a\xi_1 - b\xi_2 \\ \bar{b}\xi_1 + \bar{a}\xi_2 \end{pmatrix}. \end{aligned}$$

To ensure that we did this right, we require  $\langle \xi | \xi \rangle = \langle \xi' | \xi' \rangle$ .

**Example** (Converting from Rotations in  $\mathbb{R}^3$  to rotations in  $\mathbb{C}$ ). Note that for any complex matrix  $U$ , we can write

$$U = n_0 I + i \sum_{j=1}^3 n_j \sigma_j$$

where  $\sigma_j$  are the Pauli matrices,

$$\begin{aligned} \sigma_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \sigma_3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

As a matrix, we have

$$U = \begin{pmatrix} n_0 + in_3 & i(n_1 - in_2) \\ i(n_1 + in_2) & n_0 - in_3 \end{pmatrix}.$$

Note that, for  $n_i \in \mathbb{R}$ , we have  $U$  is Hermitian.

We call this the quaternion representation of the matrix.

Note that

$$UU^* = \begin{pmatrix} n_0^2 + n_1^2 + n_2^2 + n_3^2 & 0 \\ 0 & n_0^2 + n_1^2 + n_2^2 + n_3^2 \end{pmatrix},$$

meaning for  $U$  to be unitary, we must have  $\det U = n_0^2 + n_1^2 + n_2^2 + n_3^2 = 1$ .

Using the  $n_i$  is annoying, though, so we introduce angles. Define  $\varphi$  to be such that  $n_0 = \cos(\varphi/2)$ , and  $\vec{n} = \hat{n} \sin(\varphi/2)$ . Note that this satisfies

$$\begin{aligned} n_0^2 + \vec{n} \cdot \vec{n} &= 1 \\ &= n_0^2 + n_1^2 + n_2^2 + n_3^2. \end{aligned}$$

We now have

$$U = \cos(\varphi/2) I + i \hat{n} \cdot \sigma \sin(\varphi/2).$$

Incidentally, this looks a lot like  $e^{i\phi} = \cos \phi + i \sin \phi$ . In particular, we now have

$$\begin{aligned} U^* &= \cos(\phi/2) I + i \hat{n} \cdot \sigma^* \sin(\phi/2) \\ &= \cos(\phi/2) I - i \hat{n} \cdot \sigma \sin(\phi/2). \end{aligned}$$

Note that this means

$$\begin{aligned} U(\phi) &= e^{i\phi \hat{n} \cdot \sigma / 2} \\ U^* &= e^{-i\phi \hat{n} \cdot \sigma / 2} \\ U U^* &= I. \end{aligned}$$

We can also use the determinant/trace identity

$$\det(e^M) = e^{\text{tr}(M)}$$

to show that

$$\begin{aligned} \det(e^{i\phi \hat{n} \cdot \sigma / 2}) &= e^{i(\phi/2) \text{tr}(\hat{n} \cdot \sigma)} \\ &= 1, \end{aligned}$$

since  $\text{tr}(\sigma_i) = 0$ .

### Improper Rotations

Note that, by the definition of determinant, we have

$$\begin{aligned} 1 &= \det(R^T R) \\ &= \det(R^T) \det(R) \\ &= (\det(R))^2. \end{aligned}$$

We can have  $\det(R) = \pm 1$ . However, only  $\det(R) = 1$  yields a proper rotation.

When it comes to improper rotations, the “closest” to the identity is the pure reflection:

$$B_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Note that this yields

$$\begin{aligned} B_0 \mathbf{v} &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix} \\ &= \begin{pmatrix} v_x \\ -v_y \end{pmatrix}. \end{aligned}$$

## List of Tables

1	Coordinate Conversions . . . . .	95
2	Line, Area, and Volume Elements in Different Coordinate Systems . . . . .	95
3	Complex Number Identities . . . . .	95
4	$\delta_{ij}$ and $\epsilon_{ijk}$ . . . . .	95
5	Values of $\epsilon_{ijk}$ . . . . .	96
6	Vector Identities, $\epsilon_{ijk}$ , and $\delta_{ij}$ . . . . .	96
7	Important Taylor Series . . . . .	96
8	Useful Trig Identities . . . . .	97
9	Integrals of Powers and Products of Sine and Cosine . . . . .	97
10	Gaussian Integrals . . . . .	97
11	Gamma and Zeta Functions . . . . .	97
12	Gradient, Divergence, and Curl . . . . .	97
13	Gradients in Coordinate Systems . . . . .	98
14	Integrating Scalar Fields . . . . .	98
15	Integrating Vector Fields . . . . .	98
16	Generalized Stokes's Theorem . . . . .	99
17	Maxwell's Equations . . . . .	99

Polar	Cylindrical	Spherical
$\mathbf{s} = s(\rho, \phi)$	$\mathbf{s} = s(\rho, \phi, z)$	$\mathbf{s} = s(r, \phi, \theta)$
$\mathbf{s} = \rho \cos \phi \hat{\mathbf{i}} + \rho \sin \phi \hat{\mathbf{j}}$	$\mathbf{s} = \rho \cos \phi \hat{\mathbf{i}} + \rho \sin \phi \hat{\mathbf{j}} + z \hat{\mathbf{k}}$	$\mathbf{s} = r \cos \phi \sin \theta \hat{\mathbf{i}} + r \sin \phi \sin \theta \hat{\mathbf{j}} + r \cos \theta \hat{\mathbf{k}}$

Table 1: Coordinate Conversions

Coordinate System	Line Element	Area Element	Volume Element
Polar	$d\mathbf{s} = \hat{\rho} d\rho + \rho \hat{\phi} d\phi$	$d\mathbf{a} = r dr d\phi$	—
Cylindrical	$d\mathbf{s} = \hat{\rho} d\rho + \rho \hat{\phi} d\phi + \hat{\mathbf{k}} dz$	$d\mathbf{a} = \hat{\rho} \rho d\phi dz$	$d\tau = r dr d\phi dz$
Spherical	$d\mathbf{s} = \hat{\mathbf{r}} dr + r \sin \theta \hat{\phi} d\phi + r \hat{\theta} d\theta$	$d\mathbf{a} = \hat{\mathbf{r}} r^2 \sin \theta d\phi d\theta$	$d\tau = r^2 \sin \theta dr d\phi d\theta$

Table 2: Line, Area, and Volume Elements in Different Coordinate Systems

Quantity	Expression and/or Criterion
Cartesian form	$z = a + bi$
Polar form	$z = re^{i\phi}$
$r$	$\sqrt{a^2 + b^2}$
$\phi$	$\arg z = \arctan\left(\frac{b}{a}\right)$
Cartesian $\bar{z}$	$\bar{z} = a - bi$
Polar $\bar{z}$	$\bar{z} = re^{-i\phi}$
$ z $	$\sqrt{z\bar{z}}$
$\text{Re}(z)$	$\text{Re}(z) = \frac{z+\bar{z}}{2}$
$\text{Im}(z)$	$\text{Im}(z) = \frac{z-\bar{z}}{2i}$
$\cos \phi$	$\frac{e^{i\phi} + e^{-i\phi}}{2}$
$\sin \phi$	$\frac{e^{i\phi} - e^{-i\phi}}{2i}$
$e^{i\phi}$	$\cos \phi + i \sin \phi$
$e^{in\phi}$	$\cos(n\phi) + i \sin(n\phi)$

Table 3: Complex Number Identities

Name	Notation	Definition
Kronecker Delta	$\delta_{ij}$	$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$
Levi-Civita Symbol	$\epsilon_{ijk}$	$\epsilon_{ijk} = \begin{cases} 1 & (i, j, k) = (1, 2, 3) \text{ cyclically} \\ -1 & (i, j, k) = (2, 1, 3) \text{ cyclically} \\ 0 & \text{else} \end{cases}$

Table 4:  $\delta_{ij}$  and  $\epsilon_{ijk}$

Order of (i, j, k)	Value of $\epsilon_{ijk}$
1, 2, 3	1
3, 1, 2	1
2, 3, 1	1
1, 3, 2	-1
2, 1, 3	-1
3, 2, 1	-1
else	0

Table 5: Values of  $\epsilon_{ijk}$ 

Value	Index Notation
$\mathbf{A} \times \mathbf{B}$	$\sum_{i,j,k} \epsilon_{ijk} A_i B_j \hat{e}_k$
$(\mathbf{A} \times \mathbf{B})_\ell$	$\sum_{i,j} \epsilon_{ij\ell} A_i B_j$
$(\hat{e}_i \times \hat{e}_j) \cdot \hat{e}_k$	$\epsilon_{ijk}$
$B_i$	$\sum_{\alpha} B_{\alpha} \delta_{\alpha i}$
$\mathbf{A} \cdot \mathbf{B}$	$\sum_{i,j} A_i B_j \delta_{ij}$
$\sum_{j,k} \epsilon_{mjk} \epsilon_{nj k}$	$2\delta_{mn}$
$\sum_{\ell} \epsilon_{mnl} \epsilon_{ij \ell}$	$\delta_{mi} \delta_{nj} - \delta_{mj} \delta_{ni}$

Table 6: Vector Identities,  $\epsilon_{ijk}$ , and  $\delta_{ij}$ 

Function	Taylor Series
$f(x)$	$\sum_{k=0}^{\infty} \frac{(x-x_0)^k}{k!} \left( \left. \frac{d^k f}{dx^k} \right _{x=x_0} \right)$
$e^x$	$\sum_{n=0}^{\infty} \frac{x^n}{n!}$
$\cos x$	$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$
$\sin x$	$\sum_{n=0}^{\infty} \frac{(-1)^{n+1} x^{2n+1}}{(2n+1)!}$
$(1+x)^\alpha$	$\sum_{n=0}^{\infty} \frac{\prod_{k=0}^{n-1} (\alpha-k)}{n!} x^n$

Table 7: Important Taylor Series



Value	Expression
$\sin(\alpha \pm \beta)$	$\sin \alpha \cos \beta \pm \sin \beta \cos \alpha$
$\cos(\alpha \pm \beta)$	$\cos \alpha \cos \beta \mp \sin \alpha \sin \beta$
$\sin \alpha \cos \beta$	$\frac{1}{2} (\sin(\alpha + \beta) + \sin(\alpha - \beta))$
$\cos \alpha \cos \beta$	$\frac{1}{2} (\cos(\alpha - \beta) + \cos(\alpha + \beta))$
$\sin \alpha \sin \beta$	$\frac{1}{2} (\cos(\alpha - \beta) - \cos(\alpha + \beta))$

Table 8: Useful Trig Identities

Integral	Shortcut
$\int \sin^m(x) \cos^{2k+1}(x) dx$	$\int u^m (1 - u^2)^k du$
$\int \sin^{2k+1}(x) \cos^n(x) dx$	$-\int (1 - u^2)^k u^n du$
$\int \sin^2(x) dx$	$\frac{x}{2} - \frac{1}{4} \sin(2x)$
$\int \cos^2(x) dx$	$\frac{x}{2} + \frac{1}{4} \sin(2x)$

Table 9: Integrals of Powers and Products of Sine and Cosine

Expression	Value
$I_0$	$\frac{1}{2} \sqrt{\frac{\pi}{a}}$
$I_1$	$\frac{1}{2a}$
$I_{2n}$	$(-1)^n \frac{d^n}{da^n} I_0$
$I_{2n+1}$	$(-1)^n \frac{d^n}{da^n} I_1$

Table 10: Gaussian Integrals

Function	Expression
$\Gamma(s)$	$\int_0^\infty x^{s-1} e^{-x} dx$
$\zeta(s)$	$\sum_{k=1}^\infty \frac{1}{k^s}$
$\Gamma(s+1)$	$s\Gamma(s)$

Table 11: Gamma and Zeta Functions

Value	Expression In Terms of $\nabla$	Expression In Terms of $\partial$
Gradient	$\nabla f$	$\frac{\partial f}{\partial x} \hat{i} + \frac{\partial f}{\partial y} \hat{j} + \frac{\partial f}{\partial z} \hat{k}$
Divergence	$\nabla \cdot \mathbf{E}$	$\sum_i \partial_i E_i$
Curl	$\nabla \times \mathbf{B}$	$\sum_{i,j,k} \epsilon_{ijk} \partial_i B_j \hat{e}_k$
Laplacian of a scalar field	$\nabla^2 f$	$\sum_i \frac{\partial^2}{\partial i^2} f$
Laplacian of a vector field	$\nabla^2 \mathbf{v}$	$\sum_i \frac{\partial^2}{\partial i^2} v_i \hat{e}_i$

Table 12: Gradient, Divergence, and Curl

Coordinate System	Value	Expression
Cartesian	Gradient	$\nabla f = \frac{\partial f}{\partial x} \hat{i} + \frac{\partial f}{\partial y} \hat{j} + \frac{\partial f}{\partial z} \hat{k}$
	Laplacian	$\nabla \cdot \nabla f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$
	Divergence	$\nabla \cdot \mathbf{A} = \frac{\partial}{\partial x} A_x + \frac{\partial}{\partial y} A_y + \frac{\partial}{\partial z} A_z$
	Curl	$\nabla \times \mathbf{A} = \left( \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{i} + \left( \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{j} + \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{k}$
Cylindrical	Gradient	$\nabla f = \frac{\partial f}{\partial \rho} \hat{\rho} + \frac{1}{\rho} \frac{\partial f}{\partial \phi} \hat{\phi} + \frac{\partial f}{\partial z} \hat{z}$
	Laplacian	$\nabla \cdot \nabla f = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial f}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}$
	Divergence	$\nabla \cdot \mathbf{A} = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial A_\rho}{\partial \rho} \right) + \frac{1}{\rho} \frac{\partial A_\phi}{\partial \phi} + \frac{\partial A_z}{\partial z}$
	Curl	$\left( \frac{1}{\rho} \frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z} \right) \hat{\rho} + \left( \frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial \rho} \right) \hat{\phi} + \frac{1}{\rho} \left( \frac{\partial}{\partial \rho} (\rho A_\phi) - \frac{\partial A_z}{\partial \phi} \right) \hat{z}$
Spherical	Gradient	$\nabla f = \frac{\partial f}{\partial r} \hat{r} + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi} \hat{\phi} + \frac{1}{r} \frac{\partial f}{\partial \theta} \hat{\theta}$
	Laplacian	$\nabla \cdot \nabla f = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial f}{\partial \theta} \right)$
	Divergence	$\nabla \cdot \mathbf{A} = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial A_r}{\partial r} \right) + \frac{1}{r \sin \theta} \frac{\partial A_\phi}{\partial \phi} + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta A_\theta)$
	Curl	$\nabla \times \mathbf{A} = \frac{1}{r \sin \theta} \left( \frac{\partial}{\partial \theta} (\sin \theta A_\phi) - \frac{\partial A_\theta}{\partial \phi} \right) \hat{r} + \frac{1}{r} \left( \frac{\partial}{\partial r} (r A_\theta) - \frac{\partial A_r}{\partial \theta} \right) \hat{\phi} + \frac{1}{r} \left( \frac{1}{\sin \theta} \frac{\partial A_r}{\partial \phi} - \frac{\partial}{\partial r} (r A_\phi) \right) \hat{\theta}$

Table 13: Gradients in Coordinate Systems

Type of Integral	Expression
Line Integral, $C = y(x)$	$\int_{x_1}^{x_2} f(x, g(x)) \sqrt{1 + \left( \frac{dy}{dx} \right)^2} dx$
Line Integral, $\mathbf{c}(t) = (x(t), y(t))$	$\int_{t_1}^{t_2} f(x(t), y(t)) \left\  \frac{d\mathbf{c}}{dt} \right\  dt$
Surface Integral, $z = g(x, y)$	$\int_D f(x, y, g(x, y)) \sqrt{1 + \ \nabla g\ ^2} dx dy$
Surface Integral, $\mathbf{r} = \mathbf{x}(s, t)$	$\int_{t_1}^{t_2} \int_{s_1}^{s_2} f(\mathbf{r}) \left\  \frac{\partial \mathbf{x}}{\partial s} \times \frac{\partial \mathbf{x}}{\partial t} \right\  ds dt.$

Table 14: Integrating Scalar Fields

Type of Integral	Expression
$\int_C \mathbf{B} \cdot d\vec{\ell}, \mathbf{B} = B_x \hat{i} + B_y \hat{j} + B_z \hat{k}$	$\int_C B_x dx + B_y dy + B_z dz$
$\int_C \mathbf{B} \cdot d\vec{\ell}, \mathbf{B} = B_r \hat{r} + B_\phi \hat{\phi} + B_z \hat{k}$	$\int_C B_r dr + r B_\phi d\phi + B_z dz$
$\int_C \mathbf{B} \cdot d\vec{\ell}, \mathbf{B} = B_r \hat{r} + B_\phi \hat{\phi} + B_\theta \hat{\theta}$	$\int_C B_r dr + r \sin \theta B_\phi d\phi + B_\theta d\theta$
$\int_C \mathbf{E} \cdot d\vec{\ell}, \mathbf{E} = \nabla \Phi$	$\Phi(t_2) - \Phi(t_1)$
$\Phi$	$\int_S \mathbf{E} \cdot d\mathbf{a}$
$\Phi_{\text{net}}$	$\oint_S \mathbf{E} \cdot d\mathbf{a}$

Table 15: Integrating Vector Fields

Theorem	Expression
Divergence Theorem	$\int_V \nabla \cdot \mathbf{E} \, d\tau = \oint_{\partial V} \mathbf{E} \cdot d\mathbf{a}$
Divergence Theorem, Gradients	$\int_V \nabla \Phi \, d\tau = \oint_{\partial V} \Phi d\mathbf{a}$
Divergence Theorem, Curls	$\int_V \nabla \times \mathbf{E} \, d\tau = \oint_{\partial V} d\mathbf{a} \times \mathbf{E}$
Green's Theorem	$\int_C \mathbf{B} \cdot d\vec{\ell} = \int_S \left( \frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right) dx dy$
Stokes's Theorem	$\int_S \nabla \times \mathbf{B} \, d\mathbf{a} = \int_{\partial S} \mathbf{B} \cdot d\vec{\ell}$
Stokes's Theorem, Gradients	$\oint_C \Phi d\vec{\ell} = \int_S d\mathbf{a} \times \nabla \Phi$
Stokes's Theorem, Curls	$\oint_C d\vec{\ell} \times \mathbf{E} = \int_S (d\mathbf{a} \times \nabla) \times \mathbf{E}$

Table 16: Generalized Stokes's Theorem

Equation	Integral form	Differential Form
Gauss's Law	$\oint_S \mathbf{E} \cdot d\mathbf{a} = \frac{Q_{\text{encl}}}{\epsilon_0}$	$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$
Gauss's Law for Magnetism	$\oint_S \mathbf{B} \cdot d\mathbf{a} = 0$	$\nabla \cdot \mathbf{B} = 0$
Faraday's Law	$\oint_C \mathbf{E} \cdot d\vec{\ell} = -\frac{d}{dt} \int_S \mathbf{B} \cdot d\mathbf{a}$	$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$
Ampère's Law	$\oint_C \mathbf{B} \cdot d\vec{\ell} = \mu_0 I_{\text{encl}} + \mu_0 \epsilon_0 \int_S \mathbf{E} \cdot d\mathbf{a}$	$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$

Table 17: Maxwell's Equations