

Topics on Fundamental Science

D. T. McGuiness, Ph.D

**M.Sc  
Electrodynamics  
Lecture Book**

WS.2025







WS.2025

# Electrodynamics

Lecture Book

D. T. McGuiness, Ph.D

MCI

(2025, D. T. McGuines, Ph.D)

Current version is WS.2025.

This document includes the contents of Electrodynamics, official name being *Elektrodynamik*, taught at MCI in the Mechatronik Smart Technologies. This document is the part of the module MECH-M-1-EDY-EDY-VO taught in the M.Sc degree.

All relevant code of the document is done using *SageMath* where stated and Python v3.13.7.

This document was compiled with Lua $\text{\TeX}$  v1.22.0, and all editing were done using GNU Emacs v30.1 using AUCT $\text{\TeX}$  and org-mode package.

This document is based on the following books and resources shown in no particular order:

*A modern introduction to classical electrodynamics* by Maggiore Michele , Oxford University Press (2023) *Introduction to Electrodynamics (4th Edition)* by Griffiths David J. , Cambridge University Press (2023) *Field and Wave Electromagnetics* by Cheng David Keun. , Pearson Education India (1989). by Electromagnetic Waves and Radiation , by Schaums Electromagnetics ,

The document is designed with no intention of publication and has only been designed for education purposes.

The current maintainer of this work along with the primary lecturer  
is D. T. McGuiness, Ph.D. (dtm@mci4me.at).



# Table of Contents

## Part I Prologue

---

<b>Chapter 1</b>	<b>The Purpose of Electromagnetism</b>	<b>3</b>
1.1	Mechanics in Four Different Views . . . . .	3
	Four Kinds of Forces · Unifying Physical Theories · The Fields of Electrodynamics · Electric Charge	
<b>Chapter 2</b>	<b>Vector Calculus</b>	<b>9</b>
2.1	Vector Algebra . . . . .	9
	Vector Operations · Vector Component Forms · Triple Products · Position, Displacement, and Separation Vectors	
2.2	Differential Calculus . . . . .	15
	Ordinary Derivatives · Gradient · The Del Operator · Product Rules · Second Derivatives	
2.3	Integral Calculus . . . . .	24
	Line, Surface, and Volume Integrals · The Fundamental Theorems of Vector Calculus · The Fundamental Theorem for Curls	
2.4	Curvilinear Coordinates . . . . .	33
	Spherical Coordinate System · Cylindrical Coordinates	
2.5	Dirac Delta Function . . . . .	36
	A Mathematical Anomaly · The 1D Dirac Delta Function · The 3D Dirac Delta Function	
2.6	Vector Field Theory . . . . .	40
	Helmholtz Theorem · Potentials	

## Part II Electric Fields

---

<b>Chapter 3</b>	<b>Electrostatics</b>	<b>45</b>
3.1	The Electric Field . . . . .	45
	Coulomb's Law · The Electric Field · Continuous Charge Distributions	
3.2	Divergence and Curl of Electrostatic Fields . . . . .	49
	Quantification · Divergence of E · Applications of Gauss's Law · Curl of E	
3.3	Electric Potential . . . . .	55
	Defining the Potential · Some Digressions on Potential · Poisson's Equation and Laplace's Equation · The Potential of a Localised Charge Distribution · Boundary Conditions	
3.4	Work and Energy . . . . .	63
	To Move a Charge · The Energy of a Point Charge Distribution · The Energy of a Continuous Charge Distribution · Comments on Electrostatic Energy	
3.5	Conductors . . . . .	68
	Properties · Induced Charges · Surface Charge and the Force on a Conductor · Capacitors	
<b>Chapter 4</b>	<b>Advanced Mathematical Methods</b>	<b>75</b>
4.1	Solving Electrostatics . . . . .	75
	Laplace's Equation in One Dimension · Boundary Conditions and Uniqueness Theorems · Conductors and the Second Uniqueness Theorem	
4.2	Method of Images . . . . .	83
	A Classic Problem · Induced Surface Charge · Force and Energy	
4.3	Separation of Variables . . . . .	87
	Case Study: 2D Grounded Plates · Case Study: 3D Rectangular Pipe · Spherical Coordinates	
4.4	Multipole Expansion . . . . .	101
	Approximating Potentials at Large Distances · Monopole and Dipole Terms · Origin of Coordinates in Multipole Expansions · The Electric Field of a Dipole	

## Part III Magnetic Fields

---

## **Part IV** Electromagnetic Fields

---

**Glossary**

**111**



# Supplemental List of Items

## List of Figures

1.1	The Standard Model of particle physics is the theory describing three of the four known fundamental forces (electromagnetic, weak and strong interactions - excluding gravity) in the universe and classifying all known elementary particles. It was developed in stages throughout the latter half of the 20th century, through the work of many scientists worldwide, with the current formulation being finalized in the mid-1970s upon experimental confirmation of the existence of quarks. . . . .	7
2.1	A mnemonic, used to define the orientation of axes in three-dimensional space and to determine the direction of the cross product of two vectors, as well as to establish the direction of the force on a current-carrying conductor in a magnetic field. . . . .	11
2.2	The decomposition of the vector $\mathbf{A}$ into its components. . . . .	13
2.3	An example of a gradient field. Here the field itself is plotted by using arrows to designate the direction of the gradient. To explain the magnitude of the gradient it is either shown by the length of the arrow or by imposing colour on to the plot, which the latter has been used here. Extending this plot to our example, we can see that most of the high temperature resides on the edges of the room whereas the centre remains cool. . . . .	15
2.4	Visual description of the divergence operation. . . . .	18
2.5	Different behaviours of the curl operation. . . . .	19
2.6	The method in which line integral is calculated. At each point the dot product of the vector ( $\mathbf{v}$ ) is taken with the length vector ( $d\mathbf{l}$ ) which is always tangential to the point in which the integration is taken. . . . .	24
2.7	An visual description of the surface integral. . . . .	25
2.8	To measure the height of a mountain, it doesn't matter what way we take, as long as we know the base and the top, we will know the height. . . . .	27
2.9	The paraboloid of "Surface Area of an Implicit Surface". . . . .	31
2.10	The two types of coordinate systems in question. (a) Spherical coordinate system (b) Spherical coordinate system. . . . .	33
2.11	The vector plot of the "divergence problem". . . . .	36
2.12	A visual representation of a 1D Dirac Delta Function. Think of it as a distribution function being squeezed to an infinitely small width. . . . .	37

3.1	An example of a singular charge in space. As can be seen a charge with a plus value will generate a repulsive force. . . . .	46
3.2	An example of imagining electric field as a collection of lines passing through. Here we can see the flux lines (i.e., the arrows) passing through an object. . . . .	49
3.3	An example interaction of tow charges interacting with one another with their field lines visible. Please observe that the field line spread away from the plus charge and are collected by the minus charge. . . . .	50
3.4	A visual representation of the gauss' law. . . . .	51
3.5	. . . . .	53
3.6	The relation between three (3) major variables used in the study of electrostatics.	61
3.7	A thought experiment on how to figure out the discrepancy when a boundary is changed. . . . .	61
3.8	Some may say that superconductors can be considered as "perfect conductors", however, they exhibit quantum effects such as the Meissner effect and quantization of magnetic flux. In perfect conductors, the interior magnetic field must remain fixed but can have a zero or nonzero value [ <b>henyey1982distinction</b> ]. In real superconductors, all magnetic flux is expelled during the phase transition to superconductivity (the Meissner effect), and the magnetic field is always zero within the bulk of the superconductor [ <b>ScienceSup2023</b> ]. . . . .	68
3.9	A visual representation of treating a non-standard body and the generated electric field. Please observe that all lines flowing from the body is perpendicular. . . . .	69
3.10	A visual representation of the electric field inside a conductor. . . . .	69
3.11	Charges induced by a point charge in a conductor (a) When a charge approaches a conductor the charges will start to move where opposite charges will move towards the point charge and same charges will move away. While there is an imbalance now, the overall charge of the conductor remains the same (b) When a charge is given inside the cavity, again similar behaviour is to be expected inside the conductor.	70
3.12	Visual description of an electric field inside a cavity. As can be seen all the lines generated by the positive side of the edge is cancelled by the negative side with it absorbs making the inside of the cavity have 0 field. . . . .	70
3.13	A parallel-plate capacitor, as shown in part A, consists of two flat conducting plates. Here one plate is positively charged and the other one is charged negatively. At a distance of $d$ apart, an electric field ( $E$ ) is generated between the two (2) plates.	72
4.1	A particular solution to a 1D Laplace. . . . .	76
4.2	A visual aid on how a Laplace equation in 2D would be tackled. . . . .	77
4.3	. . . . .	79
4.4	A visual explanation of the second uniqueness theorem. Here the dashed lines represent the surface of integration and solid line describe the actual surface ( $S$ ) of the charge. . . . .	81
4.5	. . . . .	82
4.6	. . . . .	83

4.7	.....	87
4.8	The solution to the 2D-Grounded plate. Please observe that the solution is generated by summing up consecutive parts of a Fourier series to achieve a more accurate result.	91
4.9	.....	100
4.10	The setup in question. Here we have two (2) charges and we are trying to estimate its value from far away.	101
4.11	A visual comparison between (a) physical description of a dipole and (b) a <b>pure</b> dipole.	106

## List of Tables

1.1	The four fundamental forces and their respective properties.	5
2.1	Product rules of ordinary derivatives	20
2.2	Product rules of vector operations	21
4.1	Legendre Polynomials	95

## List of Examples

2.1	Finding Vector Components - I	16
2.2	Finding Vector Components - II	17
2.3	An Example of a Curl	19
2.4	The Laplacian of a Vector	22
2.5	Fluid Flow	24
2.6	Field Circulation	25
2.7	Double Integrals	26
2.8	An Example of Divergence Theorem	29
2.9	Divergence Theorem of an Octant of a Sphere	29
2.10	Surface Area of an Implicit Surface	31
2.11	Stokes' Theorem Over a Hemisphere	32
2.12	The Volume of a Sphere	34
2.13	A Simple Dirac Integral	38
2.14	1D Dirac Delta Function	38
3.1	Electric Field at a Distance	47
3.2	A Gaussian Sphere	52
4.1	Spherical Separation of Variables	91
4.2	Hollow Sphere Spherical Coordinates	96
4.3	Charge Density	96
4.4	Uncharged Metal Sphere	97
4.5	A Long Rectangular Tube	99

## List of Theorems

2.1	Calculus Theorem . . . . .	27
2.2	Gradient Theorem . . . . .	27
2.3	Line Independence of Gradient . . . . .	28
2.4	Divergence Theorem . . . . .	28
2.5	Stokes' Theorem . . . . .	30
2.6	Zero Curl Fields . . . . .	41
2.7	Zero Divergence Fields . . . . .	41
4.1	First Uniqueness . . . . .	79
4.2	Second Uniqueness . . . . .	80



## Part I

# Prologue

## Part Contents

---

Chapter 1	The Purpose of Electromagnetism . . . . .	3
Chapter 2	Vector Calculus . . . . .	9

But still try, for who knows what is possible . . .

---

*(Michael Faraday, in *The Life and Letters of Faraday* (1870) Vol. II, edited by Henry Bence Jones, p. 483)*



# Chapter 1

## The Purpose of Electromagnetism

### Table of Contents

1.1 Mechanics in Four Different Views . . . . .	3
---	---

### 1.1 Mechanics in Four Different Views

Newtonian mechanics is usually enough for most **everyday** applications ranging from calculating the motion of an object falling from a building to predicting the motion of planets. When objects start to move at high speeds<sup>1</sup>, however, the Newtonian model fails to predict accurately and was replaced by special relativity, which was introduced by Einstein in 1905 [**einstein1905electrodynamics**]. For objects which are **extremely small**, sizes comparable to that of atoms, then the theory of relativity fails for different reasons, and is superseded by quantum mechanics.<sup>2</sup>

<sup>1</sup>as in speeds comparable to the speed of light.

Finally, for objects which are both very fast and very small<sup>3</sup>, a mechanics which combines both relativity and quantum principles is needed:

<sup>2</sup>Developed by Bohr, Schrodinger, Heisenberg, and many others, in the 1920's, mostly.

This relativistic quantum mechanics is known as quantum field theory.

Even today this new model cannot claim to be a completely satisfactory system.

Luckily for us, in this **Electrodynamics Lecture**, we will **NOT** look at these models in detail but look at a specific subset of it called **electrodynamics** which is generally in the domain of classical mechanics.<sup>4</sup>

<sup>4</sup>which can be extended to other domains if needed.

**Information:** Electromagnetism and Relativity

Interestingly, electromagnetism was one of the main catalyst for developing general relativity.

Historically, the *Maxwell equations* were discovered before the theory of special relativity. It was thought light waves we derived above must be oscillations of some substance which fills all of space. This was dubbed the aether.<sup>5</sup> The idea was that Maxwell's equations only hold in the frame in which the aether is at rest; light should then travel at speed  $c$  relative to the aether.

We now know that the concept of the aether is unnecessary baggage. Instead, Maxwell's equations hold in all inertial frames and are the first equations of physics which are consistent with the laws of special relativity. Ultimately, it was by studying the Maxwell equations that Lorentz was able to determine the form of the Lorentz transformations which subsequently laid the foundation for Einstein's vision of space and time [DTong2024].

<sup>5</sup>the postulated medium for the propagation of light [brewer1852guide]. It was originally used to explain the ability of the apparently wave-based light to propagate through empty space, something that waves should not be able to do. The assumption of a spatial plenum of luminiferous aether, rather than a spatial vacuum, provided the theoretical medium that was required by wave theories of light.

### 1.1.1 Four Kinds of Forces

Classical Mechanics tells us how behaviour is when subjected to a given force. There are four (4) basic forces known to physics. Listing them in the order of decreasing strength:

#### Strong

Holds protons and neutrons together in the atomic nucleus, have extremely short range, so we do not "feel" them, in spite of the fact that they are a hundred times more powerful than electrical forces.

#### Weak

Accounts for certain kinds of radioactive decay, are also of short range, and they are far weaker than electromagnetic forces.

#### Gravitational

it is very weak<sup>6</sup> that it is only in scale of huge mass concentrations, for example the earth and the sun, that we ever notice it at all.

<sup>6</sup>compared to all of the others

#### Electromagnetic

It is the interaction that occurs between particles with electric charge via electromagnetic fields.

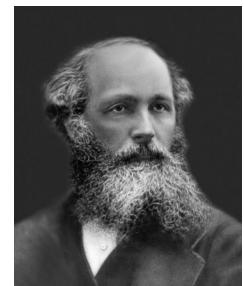
For reference, the electrical repulsion between two (2) electrons is  $1 \times 10^{42}$  times as large as their gravitational attraction.

Not only are electromagnetic forces overwhelmingly dominant in everyday life, they are also, the only ones which are completely understood.

There is, of course, a classical theory of gravity and a relativistic one, but no entirely satisfactory quantum mechanical theory of gravity has been constructed. At the present time there is a

Table 1.1: The four fundamental forces and their respective properties.

Traits	Strong Force	Electromagnetism	Weak Force	Gravity
Affected Particles	quarks <sup>1</sup> and gluons <sup>2</sup>	electrically charged	quarks and leptons <sup>3</sup>	all particles with mass
Force Carrying Particle	gluon ( $g$ )	photon <sup>4</sup> ( $\gamma$ )	W and Z bosons ( $W^+, W^-, Z^0$ )	graviton (unobserved)
Acting Range	short ( $\sim 1 \text{ fm}$ ) <sup>5</sup>	$\infty$	short	$\infty$
Strength	1	$1/137^6$	$1 \times 10^{-6}$	$6 \times 10^{-39}$

<sup>1</sup> Elementary particle responsible for making protons and neutron.<sup>2</sup> Elementary particle acting as exchange particle for the strong force between quarks.<sup>3</sup> Elementary particle affected by the weak force but not by the strong force.<sup>4</sup> Elementary particle that is a quantum of the electromagnetic field, such as light and radio waves.<sup>5</sup> 1 femtometre ( $1 \text{ fm} = 1 \times 10^{-15} \text{ m}$ ). i.e., the gold nucleus radius is approx. 8.45 fm.<sup>6</sup> Known as the fine structure constant.

<sup>7</sup> James Clerk Maxwell  
FRS FRSE  
(1831 - 1879)

A Scottish physicist and mathematician who was responsible for the classical theory of electromagnetic radiation, which was the first theory to describe electricity, magnetism and light as different manifestations of the same phenomenon. Maxwell's equations for electromagnetism achieved the second great unification in physics, where the first one had been realised by Isaac Newton. Maxwell was also key in the creation of statistical mechanics.

With the publication of *A Dynamical Theory of the Electromagnetic Field* in 1865, Maxwell demonstrated electric and magnetic fields travel through space as waves moving at the speed of light. He proposed light is an undulation in the same medium that is the cause of electric and magnetic phenomena. The unification of light and electrical phenomena led to his prediction of radio waves. As a result of his equations, and other contributions such as introducing an effective method to deal with network problems and linear conductors, he is regarded as a founder of the modern field of electrical engineering.

very successful theory for the weak interactions, and a strikingly attractive candidate, called chromodynamics. for the strong interactions.

All these theories draw their inspiration from electrodynamics. None can claim conclusive experimental verification at this stage. So electrodynamics, a complete and successful theory, has become a pedestal for physicists:

An ideal model that other theories emulate.

The laws of classical electrodynamics were discovered in bits and pieces by Franklin, Coulomb, Ampere, Faraday, and others, but the person who completed the job, and packaged it all in the compact and consistent form it has today, was James Clerk Maxwell.<sup>7</sup>

The theory is now about 150 years old.

## 1.1.2 Unifying Physical Theories

In the beginning, electricity and magnetism were entirely separate subjects. The one dealt with glass rods and cat's fur, batteries, currents, electrolysis, and lightning, whereas the other with bar magnets, iron filings, compass needles, and the North Pole. However, in 1820 Øersted noticed that an electric current could deflect a magnetic compass needle. Soon afterwards, Ampère correctly postulated all magnetic phenomena are due to electric charges in motion. Then, in 1831, Faraday discovered that a moving magnet generates an electric current. By the time Maxwell and Lorentz put the finishing touches on the theory, electricity and magnetism were inextricably intertwined.

They could no longer be regarded as separate subjects, but rather as two aspects of a single subject: electromagnetism.

Faraday speculated that light, too, is electrical in nature. Maxwell's theory provided justification for this hypothesis, and soon optics, which is the study of lenses, mirrors, prisms, interference, and diffraction, was incorporated into electromagnetism. Hertz, who presented the decisive experimental confirmation for Maxwell's theory in 1888, said:

The connection between light and electricity is now established . . . In every flame, in every luminous particle, we see an electrical process . . . Thus, the domain of electricity extends over the whole of nature. It even affects ourselves intimately: we perceive that we possess . . . an electrical organ—the eye.

By 1900, three (3) great branches of physics: **electricity**, **magnetism**, and **optics** had merged into a single unified theory.

It became apparent that visible light represents only a tiny “window” in the vast spectrum of electromagnetic radiation, from radio through microwaves, infrared and ultraviolet . . .

Einstein worked on a further unification, which would combine gravity and electrodynamics, in much the same way as electricity and magnetism had been combined a century earlier. His unified field theory was not particularly successful, but in recent years the same impulse has spawned a hierarchy of increasingly ambitious unification schemes, beginning in the 1960s with the electroweak theory of Glashow, Weinberg, and Salam, which joins the weak and electromagnetic forces, and culminating in the 1980s with the superstring theory.<sup>8</sup>

At each step in this hierarchy, the mathematical difficulties mount, and the gap between inspired conjecture and experimental test widens; nevertheless, it is clear the unification of forces initiated by electrodynamics has become a major theme in the progress of physics.

<sup>8</sup>an attempt to explain all of the particles and fundamental forces of nature in one theory by modeling them as vibrations of tiny supersymmetric strings.

### 1.1.3 The Fields of Electrodynamics

The essential problem the theory of electromagnetism hopes to solve is:

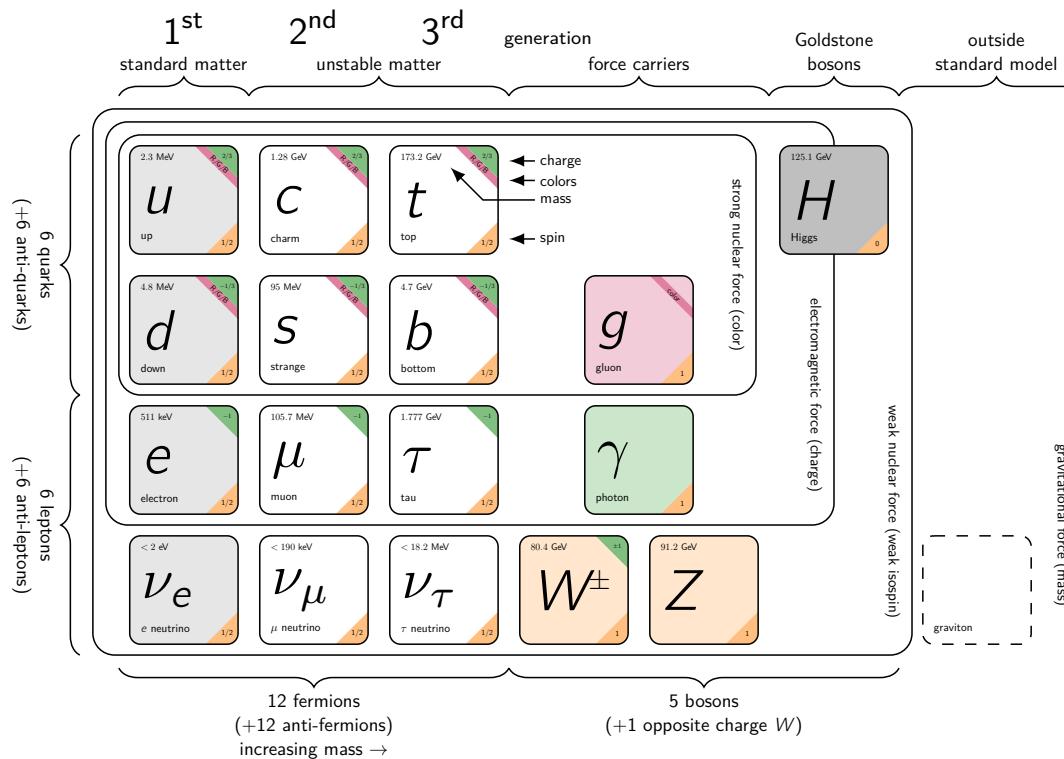
If there exists a bunch of electric charges here, what happens to some other charge, over another place?

The classical solution takes the form of a **field theory**:

We say that the space around an electric charge is permeated by **electric** and **magnetic** fields. A second charge, in the presence of these fields, experiences a force. The fields, then, transmit the influence from one charge to the other—they “mediate” the interaction.

When a charge undergoes acceleration, a portion of the field **detaches** itself, in sense, and travels off at the speed of light, carrying with it energy, momentum, and angular momentum. We call this **electromagnetic radiation**. Its existence invites<sup>9</sup> us to regard the fields as independent dynamical entities in their own right, every bit as **real** as atoms or baseballs.

<sup>9</sup>if not compels



**Figure 1.1:** The Standard Model of particle physics is the theory describing three of the four known fundamental forces (electromagnetic, weak and strong interactions - excluding gravity) in the universe and classifying all known elementary particles. It was developed in stages throughout the latter half of the 20th century, through the work of many scientists worldwide, with the current formulation being finalized in the mid-1970s upon experimental confirmation of the existence of quarks.

Our interest accordingly shifts from the study of forces between charges to the theory of the fields themselves. But it takes a charge to produce an electromagnetic field, and it takes another charge to detect one, so we had best begin by reviewing the essential properties of electric charge.

### 1.1.4 Electric Charge

#### Charge comes in two varieties

We call **plus** and **minus**, as their effects tend to cancel.<sup>10</sup> The interesting point here is that plus and minus charges occur in exactly equal amounts, to high precision, in bulk matter, so that their effects are almost completely neutralised.

<sup>10</sup>If we have  $+q$  and  $-q$  at the same point, electrically it is the same as having no charge there at all.

#### Charge is conserved

Charge cannot be created or destroyed [purcell2013electricity]. What there is now has always been. A plus charge can **annihilate** an equal minus charge, but a plus charge cannot simply disappear by itself, something must pick up that electric charge. So the total charge of the universe is fixed for all time.

This is called **global conservation of charge**. Charges also cannot disappear and then reappear

in some other place, the charge, if moved must follow a continuous path. This is called **local conservation of charge**.

### Charge is quantised

While there is no restrictions in classical electrodynamics which require it to be, electric charge comes only in **integer multiples** of the basic unit of charge. If we call the charge on the proton  $+e$ , then the electron carries charge  $e$ .<sup>11</sup>

<sup>11</sup>It is never  $7.392e$ , or  $1/2e$ .

#### Information: Interpreting Conservation of Charge

Charge conservation does **NOT** mean that individual positive and negative charges cannot be created or destroyed. Electric charge is carried by subatomic particles such as **electrons** and **protons**. Charged particles can be created and destroyed in elementary particle reactions.

In particle physics, charge conservation means, in reactions that create charged particles, equal numbers of positive and negative particles are always created, keeping the net amount of charge unchanged. Similarly, when particles are destroyed, equal numbers of positive and negative charges are destroyed.

This property is supported without exception by all empirical observations so far [[purcell2013electricity](#)].

Even though the conservation of charge requires that the total quantity of charge in the universe is constant, it leaves open the question of what that quantity is. Most evidence indicates that the net charge in the universe is zero [[orito1985can](#)]. That is, there are equal quantities of positive and negative charge.

# Chapter 2

## Vector Calculus

### Table of Contents

2.1	Vector Algebra . . . . .	9
2.2	Differential Calculus . . . . .	15
2.3	Integral Calculus . . . . .	24
2.4	Curvilinear Coordinates . . . . .	33
2.5	Dirac Delta Function . . . . .	36
2.6	Vector Field Theory . . . . .	40

### 2.1 Vector Algebra

#### 2.1.1 Vector Operations

Walking 5 km north and then 12 km east, we will have gone a total of 17 km, but if we carefully observe our current position, we will realise that we're not 17 km from where we set out, but only 13. To describe these quantities, We need a set of mathematics principles, which evidently do **NOT** add in the ordinary way.

The reason they don't, is **displacements** have *direction* as well as *magnitude*, and it is essential to take both into account when we combine them. Such objects are called **vectors**.

Examples include: velocity, acceleration, force, momentum . . .

By contrast, quantities that have magnitude but no direction are called **scalars**.

Examples include: mass, charge, density, temperature, . . .



We shall use **boldface** ( $\mathbf{A}$ ,  $\mathbf{B}$ , and so on) for vectors and **normal-font** for scalars. The magnitude of a vector  $\mathbf{A}$  is written  $|\mathbf{A}|$  or, more simply,  $A$ . In diagrams, vectors are denoted by arrows: the length of the arrow is proportional to the magnitude of the vector, and the arrow indicates its direction.

**Minus  $\mathbf{A}$**  ( $-\mathbf{A}$ ) is a vector with the same magnitude as  $\mathbf{A}$  but of opposite direction.

Vectors have magnitude and direction but *not location*.

Here we will define four (4) vector operations: addition and three kinds of multiplication.

**Addition of Two Vectors** Place the tail of  $\mathbf{B}$  at the head of  $\mathbf{A}$ . Their sum,  $\mathbf{A} + \mathbf{B}$ , is the vector from the tail of  $\mathbf{A}$  to the head of  $\mathbf{B}$ . Addition is *commutative*:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A},$$

5 kilometers east followed by 12 kilometers north gets us to the same place as 12 kilometers north followed by 5 kilometers east. Addition is also *associative*:

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}).$$

To subtract a vector, add its opposite

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B}).$$

**Multiplication by a Scalar Value** Multiplication of a vector by a positive scalar  $a$  multiplies the *magnitude* but leaves the direction **unchanged**. This means if  $a$  is negative, the direction is reversed. Scalar multiplication is *distributive*:

$$a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}.$$

**Dot Product of Two Vectors** The dot product of two vectors is defined by:

$$\mathbf{A} \cdot \mathbf{B} \equiv AB \cos \theta \quad (2.1)$$

where  $\theta$  is the angle they form when placed tail-to-tail.

$\mathbf{A} \cdot \mathbf{B}$  is itself a scalar.<sup>1</sup>

---

<sup>1</sup>which is why its alternative name is **scalar product**.

The dot product is *commutative*,

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A},$$

and *distributive*,

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}, \quad (2.2)$$

Geometrically,  $\mathbf{A} \cdot \mathbf{B}$  is the product of  $A$  times the projection of  $B$  along  $\mathbf{A}$ .<sup>2</sup> If the two vectors are parallel, then  $\mathbf{A} \cdot \mathbf{B} = AB$ . In particular, for any vector  $\mathbf{A}$ ,

$$\mathbf{A} \cdot \mathbf{A} = A^2 \quad (2.3)$$

If  $\mathbf{A}$  and  $\mathbf{B}$  are perpendicular, then  $\mathbf{A} \cdot \mathbf{B} = 0$ .

<sup>2</sup>Or the product of  $B$  times the projection of  $\mathbf{A}$  along  $B$ .

**Cross Product of Two Vectors** The cross product of two (2) vectors is defined by:

$$\mathbf{A} \times \mathbf{B} \equiv AB \sin \theta \hat{\mathbf{n}} \quad (2.4)$$

where  $\hat{\mathbf{n}}$  is a **unit vector** (vector of magnitude 1) pointing perpendicular to the plane of  $\mathbf{A}$  and  $\mathbf{B}$ . Of course, there are two directions perpendicular to any plane: **in** and **out**.

The ambiguity is resolved by the **right-hand rule**: let our fingers point in the direction of the first vector and curl around (via the smaller angle) toward the second; then our thumb indicates the direction of  $\hat{\mathbf{n}}$ .

**$\mathbf{A} \times \mathbf{B}$  is itself a vector and it is also known as vector product.**

The cross product is *distributive*,

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) + (\mathbf{A} \times \mathbf{C}) \quad (2.5)$$

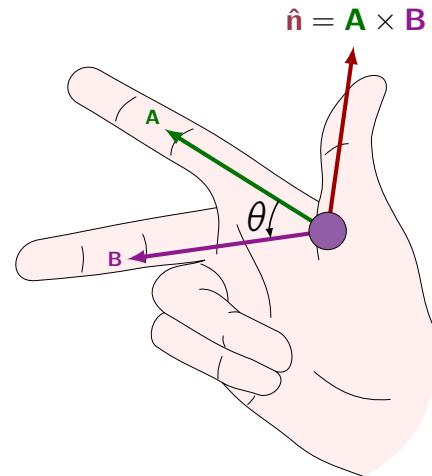
but **NOT** commutative:

$$(\mathbf{B} \times \mathbf{A}) = -(\mathbf{A} \times \mathbf{B}) \quad (2.6)$$

If two (2) vectors are parallel, their cross product is zero. In particular,

$$\mathbf{A} \times \mathbf{A} = 0,$$

for any vector  $\mathbf{A}$ .



**Figure 2.1:** A mnemonic, used to define the orientation of axes in three-dimensional space and to determine the direction of the cross product of two vectors, as well as to establish the direction of the force on a current-carrying conductor in a magnetic field.

### 2.1.2 Vector Component Forms

In the previous section, we defined the four (4) vector operations in abstract form, without reference to any particular coordinate system.

In practice, it is often easier to set up Cartesian coordinates ( $x, y, z$ ) and work with vector **components**. Let  $\hat{x}, \hat{y}, \hat{z}$  be unit vectors parallel to the  $x, y$ , and  $z$  axes, respectively. An arbitrary vector  $\mathbf{A}$  can be expanded in terms of these **basis vectors**:

$$\mathbf{A} = A_x \hat{x} + A_y \hat{y} + A_z \hat{z}$$

The symbols  $A_x$ ,  $A_y$ ,  $A_z$ , are the components of  $\mathbf{A}$ . In geometrical terms they are the **projections** of  $\mathbf{A}$  along the three (3) coordinate axes (i.e.,  $A_x = \mathbf{A} \cdot \hat{\mathbf{x}}$ ,  $A_y = \mathbf{A} \cdot \hat{\mathbf{y}}$ ,  $A_z = \mathbf{A} \cdot \hat{\mathbf{z}}$ ).

We can now reformulate each of the four (4) vector operations as a rule for manipulating components:

$$\begin{aligned}\mathbf{A} + \mathbf{B} &= (A_x \hat{\mathbf{x}} + A_y \hat{\mathbf{y}} + A_z \hat{\mathbf{z}}) + (B_x \hat{\mathbf{x}} + B_y \hat{\mathbf{y}} + B_z \hat{\mathbf{z}}) \\ &= (A_x + B_x) \hat{\mathbf{x}} + (A_y + B_y) \hat{\mathbf{y}} + (A_z + B_z) \hat{\mathbf{z}}\end{aligned}$$

The operation rules are as follows:

- i. To add vectors, add like components.

$$a\mathbf{A} = (aA_x) \hat{\mathbf{x}} + (aA_y) \hat{\mathbf{y}} + (aA_z) \hat{\mathbf{z}}$$

- ii. To multiply by a scalar, multiply each component. As  $\hat{\mathbf{x}}$ ,  $\hat{\mathbf{y}}$ ,  $\hat{\mathbf{z}}$  are mutually perpendicular unit vectors, the following properties are valid:

$$\hat{\mathbf{x}} \cdot \hat{\mathbf{x}} = \hat{\mathbf{y}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{z}} \cdot \hat{\mathbf{z}} = 1 \quad \hat{\mathbf{x}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{x}} \cdot \hat{\mathbf{z}} = \hat{\mathbf{y}} \cdot \hat{\mathbf{z}} = 0$$

Accordingly,

$$\mathbf{A} \cdot \mathbf{B} = (A_x \hat{\mathbf{x}} + A_y \hat{\mathbf{y}} + A_z \hat{\mathbf{z}}) \cdot (B_x \hat{\mathbf{x}} + B_y \hat{\mathbf{y}} + B_z \hat{\mathbf{z}})$$

- iii. To calculate the dot product, multiply like components, and add. In particular,

$$\mathbf{A} \cdot \mathbf{A} = A_x^2 + A_y^2 + A_z^2, \quad \text{so} \quad A = \sqrt{A_x^2 + A_y^2 + A_z^2}$$

Similarly the following relations can be derived:

$$\hat{\mathbf{x}} \times \hat{\mathbf{x}} = \hat{\mathbf{y}} \times \hat{\mathbf{y}} = \hat{\mathbf{z}} \times \hat{\mathbf{z}} = 0,$$

$$\hat{\mathbf{x}} \times \hat{\mathbf{y}} = -\hat{\mathbf{y}} \times \hat{\mathbf{x}} = \hat{\mathbf{z}},$$

$$\hat{\mathbf{y}} \times \hat{\mathbf{z}} = -\hat{\mathbf{z}} \times \hat{\mathbf{y}} = \hat{\mathbf{x}},$$

$$\hat{\mathbf{z}} \times \hat{\mathbf{x}} = -\hat{\mathbf{x}} \times \hat{\mathbf{z}} = \hat{\mathbf{y}}.$$

Therefore,

$$\begin{aligned}\mathbf{A} \times \mathbf{B} &= (A_x \hat{\mathbf{x}} + A_y \hat{\mathbf{y}} + A_z \hat{\mathbf{z}}) \times (B_x \hat{\mathbf{x}} + B_y \hat{\mathbf{y}} + B_z \hat{\mathbf{z}}) \\ &= (A_y B_z - A_z B_y) \hat{\mathbf{x}} + (A_z B_x - A_x B_z) \hat{\mathbf{y}} + (A_x B_y - A_y B_x) \hat{\mathbf{z}}.\end{aligned}$$

This expression look unruly but we can tidy it up and write it more neatly as a **determinant**:

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \quad (2.7)$$

- iv. To calculate the cross product, form the determinant whose first row is  $\hat{\mathbf{x}}$ ,  $\hat{\mathbf{y}}$ ,  $\hat{\mathbf{z}}$ , whose second row is  $\mathbf{A}$  (in component form), and whose third row is  $\mathbf{B}$ .

### 2.1.3 Triple Products

As the cross product of two (2) vectors is itself a vector, it can be dotted or crossed with a 3<sup>rd</sup> vector to form a **triple** product.

**Scalar Triple Product** Evidently,

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}),$$

for they all correspond to the same value.

The **alphabetical** order is preserved.

Alternatively, the **non-alphabetical** triple products,

$$\mathbf{A} \cdot (\mathbf{C} \times \mathbf{B}) = \mathbf{B} \cdot (\mathbf{A} \times \mathbf{C}) = \mathbf{C} \cdot (\mathbf{B} \times \mathbf{A}),$$

have the **opposite** sign. In component form,

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}.$$

A final point worth mentioning is the dot and cross can be interchanged:

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C},$$

however, it is important to stress it out, the placement of the parentheses is **critical**:

**Information:** Cross Product of Two Scalars

$(\mathbf{A} \cdot \mathbf{B}) \times \mathbf{C}$  is a meaningless expression. We can't make a cross product from a scalar and a vector.

**Vector Triple Product** The vector triple product can be simplified by the **BAC-CAB** rule:

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}).$$

Please observe that the following triple product

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = -\mathbf{C} \times (\mathbf{A} \times \mathbf{B}) = -\mathbf{A}(\mathbf{B} \cdot \mathbf{C}) + \mathbf{B}(\mathbf{A} \cdot \mathbf{C})$$

is an entirely different vector.<sup>3</sup> All *higher* vector products can be similarly reduced, often by repeated application, so it is never necessary for an expression to contain more than one cross product in any term. For instance,

$$\begin{aligned} (\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) &= (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}), \\ \mathbf{A} \times [\mathbf{B} \times (\mathbf{C} \times \mathbf{D})] &= \mathbf{B}[\mathbf{A} \cdot (\mathbf{C} \times \mathbf{D})] - (\mathbf{A} \cdot \mathbf{B})(\mathbf{C} \times \mathbf{D}). \end{aligned}$$

<sup>3</sup>To reiterate cross-products are not associative.

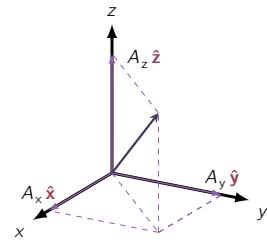


Figure 2.2: The decomposition of the vector  $\mathbf{A}$  into its components.

## 2.1.4 Position, Displacement, and Separation Vectors

The location of a point in three dimensions can be described by listing its Cartesian coordinates  $(x, y, z)$ . The vector to that point from the origin  $(\mathcal{O})$  is called the **position vector**:

$$\mathbf{r} \equiv x \hat{\mathbf{x}} + y \hat{\mathbf{y}} + z \hat{\mathbf{z}}$$

Throughout this course,  $\mathbf{r}$  will be used to measure **distance**. Its magnitude:

$$r = \sqrt{x^2 + y^2 + z^2}$$

is the distance from the origin, and

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{r} = \frac{(x) \hat{\mathbf{x}} + (y) \hat{\mathbf{y}} + (z) \hat{\mathbf{z}}}{\sqrt{x^2 + y^2 + z^2}}$$

is a unit vector pointing **radially outward**. The **infinitesimal displacement vector**, from  $(x, y, z)$  to  $(x + dx, y + dy, z + dz)$ , is

$$d\mathbf{l} = (dx) \hat{\mathbf{x}} + (dy) \hat{\mathbf{y}} + (dz) \hat{\mathbf{z}}.$$

In electrodynamics, one frequently encounters problems involving two (2) points:

**source point** ( $\mathbf{r}'$ ), where an electric charge is located

**field point** ( $\mathbf{r}$ ), at which we are calculating the electric or magnetic field

To make these redundant calculations easier to handle, let's use the following short-hand notation:

$$\mathbf{z} \equiv \mathbf{r} - \mathbf{r}'$$

Its magnitude is:

$$z = |\mathbf{r} - \mathbf{r}'|$$

and a unit vector in the direction from  $\mathbf{r}'$  to  $\mathbf{r}$  is

$$\hat{\mathbf{z}} = \frac{\mathbf{z}}{z} = \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}$$

In Cartesian coordinates our new short-notations would be as following:

$$\begin{aligned}\mathbf{z} &= (x - x') \hat{\mathbf{x}} + (y - y') \hat{\mathbf{y}} + (z - z') \hat{\mathbf{z}} \\ z &= \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2} \\ \hat{\mathbf{z}} &= \frac{(x - x') \hat{\mathbf{x}} + (y - y') \hat{\mathbf{y}} + (z - z') \hat{\mathbf{z}}}{\sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}}\end{aligned}$$

## 2.2 Differential Calculus

### 2.2.1 Ordinary Derivatives

Assume a function of just one (1) variable:  $f(x)$ . Therefore, the question we need to answer is what does the derivative,  $df/dx$ , do and convey?

It tells us how rapidly the function  $f(x)$  **varies** when we change the argument  $x$  by an infinitesimal amount,  $dx$ :

$$df = \left( \frac{df}{dt} \right) dx$$

If we increment  $x$  by an infinitesimal amount  $dx$ , then  $f$  changes by an amount  $df$ .<sup>4</sup>

Geometrically, the derivative  $df/dt$  is the *slope* of the graph of  $f$  versus  $x$ .

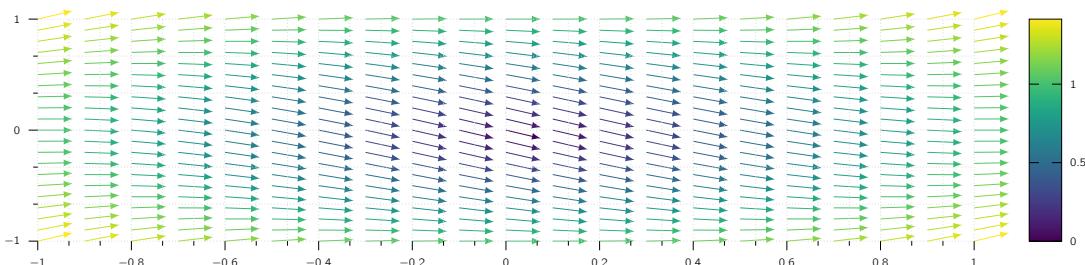
<sup>4</sup>Here we can think the derivative as the proportionality factor.

### 2.2.2 Gradient

Assume a function which accepts three (3) variables. As an example, lets take the temperature  $T(x, y, z)$  in the lecture room. Start out in one corner, and set up a system of cardinal directions. Then for each point  $(x, y, z)$  in the room,  $T$  gives the temperature at that spot. We want to generalise the notion of **derivative** to functions like  $T$ , which depend **NOT** on one but on three variables.

A derivative tells us **how fast the function varies**, if we move a little distance. But this time the situation is more complicated, because it depends on what **direction** we move:

- Going straight up, the temperature will probably increase fairly rapidly,
- Moving horizontally, it may not change much at all.



**Figure 2.3:** An example of a gradient field. Here the field itself is plotted by using arrows to designate the direction of the gradient. To explain the magnitude of the gradient it is either shown by the length of the arrow or by imposing colour on to the plot, which the latter has been used here. Extending this plot to our example, we can see that most of the high temperature resides on the edges of the room whereas the centre remains cool.

In fact, the question of "How fast does  $T$  vary?" can have an infinite number of answers, one for each direction we might choose to explore. Fortunately, the problem is not as bad as it looks. A theorem on partial derivatives states:

$$dT = \left( \frac{\partial T}{\partial x} \right) dx + \left( \frac{\partial T}{\partial y} \right) dy + \left( \frac{\partial T}{\partial z} \right) dz$$

This tells us how  $T$  changes when we alter all three (3) variables by the infinitesimal amounts of  $dx$ ,  $dy$ ,  $dz$ . We can write the aforementioned equation as a dot product:

$$dT = \left( \frac{dT}{dx} \hat{x} + \frac{dT}{dy} \hat{y} + \frac{dT}{dz} \hat{z} \right) \cdot ((dx) \hat{x} + (dy) \hat{y} + (dz) \hat{z}) = (\nabla T) \cdot (dI),$$

where

$$\nabla T \equiv \frac{dT}{dx} \hat{x} + \frac{dT}{dy} \hat{y} + \frac{dT}{dz} \hat{z}$$

is the **gradient** of  $T$ . Note that  $\nabla T$  is a **vector quantity**, with three (3) components.<sup>5</sup>

<sup>5</sup>This is the generalised derivative we have been looking for.

### Exercise 2.1 Finding Vector Components - I

Find the components of the vector  $v$  with given initial point  $P$  and terminal point  $Q$ . Find  $|v|$  and unit vector  $\hat{v}$ .

$$\begin{array}{lll} P(3, 2, 0), & Q(5, -2, 2), & P(1, 1, 1), \\ P(1, 0, 1.2), & Q(0, 0, 6.2), & P(2, -2, 0), \\ P(4, 3, 2), & Q(-4, -3, 2), & P(0, 0, 0), \end{array} \quad \begin{array}{lll} Q(-4, -4, -4) \\ Q(0, 4, 6) \\ Q(6, 8, 10) \end{array}$$

**SOLUTION** The solution is as follows:

$$v = (5 - 3) \hat{x} + (-2 - 2) \hat{y} + (2 - 0) \hat{z} = (2) \hat{x} + (-4) \hat{y} + (2) \hat{z},$$

$$|v| = \sqrt{(2)^2 + (-4)^2 + (2)^2} = 2\sqrt{6}.$$

$$\hat{v} = \frac{v}{|v|} = \frac{(2) \hat{x} + (-4) \hat{y} + (2) \hat{z}}{2\sqrt{6}} = \left( \frac{1}{\sqrt{6}} \right) \hat{x} + \left( -\frac{2}{\sqrt{6}} \right) \hat{y} + \left( \frac{1}{\sqrt{6}} \right) \hat{z} \blacksquare$$

$$v = (-4 - 1) \hat{x} + (-1 - 0) \hat{y} + (-1 - 1) \hat{z} = (-5) \hat{x} + (-5) \hat{y} + (-5) \hat{z},$$

$$|v| = \sqrt{(-5)^2 + (-5)^2 + (-5)^2} = 5\sqrt{3}.$$

$$\hat{v} = \frac{v}{|v|} = \frac{(-5) \hat{x} + (-5) \hat{y} + (-5) \hat{z}}{5\sqrt{3}} = \left( -\frac{1}{\sqrt{3}} \right) \hat{x} + \left( -\frac{1}{\sqrt{3}} \right) \hat{y} + \left( -\frac{1}{\sqrt{3}} \right) \hat{z} \blacksquare$$

$$v = (0 - 1) \hat{x} + (0 - 0) \hat{y} + (6.2 - 1.2) \hat{z} = (-1) \hat{x} + (0) \hat{y} + (5) \hat{z},$$

$$|v| = \sqrt{(-1)^2 + (0)^2 + (5)^2} = \sqrt{26}.$$

$$\hat{v} = \frac{v}{|v|} = \frac{(-1) \hat{x} + (0) \hat{y} + (5) \hat{z}}{\sqrt{26}} = \left( -\frac{1}{\sqrt{26}} \right) \hat{x} + (0) \hat{y} + \left( -\frac{5}{\sqrt{26}} \right) \hat{z} \blacksquare$$

$$v = (0 - 2) \hat{x} + (4 - (-2)) \hat{y} + (6 - 0) \hat{z} = (-2) \hat{x} + (6) \hat{y} + (6) \hat{z},$$

$$|v| = \sqrt{(-2)^2 + (6)^2 + (6)^2} = 2\sqrt{19}.$$

$$\hat{v} = \frac{v}{|v|} = \frac{(-2) \hat{x} + (6) \hat{y} + (6) \hat{z}}{2\sqrt{19}} = \left( -\frac{1}{\sqrt{19}} \right) \hat{x} + \left( \frac{3}{\sqrt{19}} \right) \hat{y} + \left( \frac{3}{\sqrt{19}} \right) \hat{z} \blacksquare$$

$$v = (-4 - 4) \hat{x} + (-3 - 3) \hat{y} + (2 - 2) \hat{z} = (-8) \hat{x} + (-6) \hat{y} + (0) \hat{z},$$

$$|v| = \sqrt{(-8)^2 + (-6)^2 + (0)^2} = 10.$$

$$\hat{v} = \frac{v}{|v|} = \frac{(-8) \hat{x} + (-6) \hat{y} + (0) \hat{z}}{10} = \left( -\frac{3}{5} \right) \hat{x} + \left( -\frac{4}{5} \right) \hat{y} + (0) \hat{z} \blacksquare$$

$$v = (6 - 0) \hat{x} + (8 - 0) \hat{y} + (10 - 0) \hat{z} = (6) \hat{x} + (8) \hat{y} + (10) \hat{z},$$

$$|\mathbf{v}| = \sqrt{(6)^2 + (8)^2 + (10)^2} = 10\sqrt{2}.$$

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{(6) \hat{\mathbf{x}} + (8) \hat{\mathbf{y}} + (0) \hat{\mathbf{z}}}{10\sqrt{2}} = \left( \frac{3}{5\sqrt{2}} \right) \hat{\mathbf{x}} + \left( \frac{4}{5\sqrt{2}} \right) \hat{\mathbf{y}} + \left( \frac{1}{\sqrt{2}} \right) \hat{\mathbf{z}} \quad \blacksquare$$

$$\nabla = i \frac{d}{dx} + j \frac{d}{dy} + k \frac{d}{dz}$$

<sup>6</sup>The History of Nabla

### Exercise 2.2 Finding Vector Components - II

Given the components of a vector  $\mathbf{v} = [v_x, v_y, v_z]$  and a particular initial point  $P$ , find the corresponding terminal point  $Q$  and the length of  $\mathbf{v}$  (i.e.,  $|\mathbf{v}|$ ).

$$\begin{array}{lll} \mathbf{v} = [3, -1, 0]; & P(4, 6, 0), & \mathbf{v} = [8, 4, 2]; & P(-8, -4, -2), \\ \mathbf{v} = [0.25, 2, 0.75]; & P(0, -0.5, 0), & \mathbf{v} = [3, 2, 6]; & P(4, 6, 0), \\ \mathbf{v} = [4, 2, -2]; & P(4, 6, 0), & \mathbf{v} = [3, -3, 3]; & P(4, 6, 0), \end{array}$$

**SOLUTION** Previously we have defined  $\mathbf{v} = Q - P$ . Here we have  $\mathbf{v}$  and  $P$ . To calculate  $Q$  we only need to add individual components of the vector with the initial point  $P$ .

$$\begin{aligned} Q &= \mathbf{v} + P = (3+4) \hat{\mathbf{x}} + (-1+6) \hat{\mathbf{y}} + (0+0) \hat{\mathbf{z}} = (7) \hat{\mathbf{x}} + (5) \hat{\mathbf{y}} + (0) \hat{\mathbf{z}}, \\ |\mathbf{v}| &= \sqrt{(3)^2 + (-1)^2 + (0)^2} = \sqrt{10} \quad \blacksquare \\ Q &= \mathbf{v} + P = (8+(-8)) \hat{\mathbf{x}} + (4+(-4)) \hat{\mathbf{y}} + (-2+2) \hat{\mathbf{z}} = (0) \hat{\mathbf{x}} + (0) \hat{\mathbf{y}} + (0) \hat{\mathbf{z}}, \\ |\mathbf{v}| &= \sqrt{(8)^2 + (4)^2 + (2)^2} = 2\sqrt{21} \quad \blacksquare \\ Q &= \mathbf{v} + P = (0.25+0) \hat{\mathbf{x}} + (2+(-0.5)) \hat{\mathbf{y}} + (0.75+0) \hat{\mathbf{z}} = (0.25) \hat{\mathbf{x}} + (1.5) \hat{\mathbf{y}} + (0.75) \hat{\mathbf{z}}, \\ |\mathbf{v}| &= \sqrt{(0.25)^2 + (1.5)^2 + (0.75)^2} = \sqrt{74}/4 \quad \blacksquare \\ Q &= \mathbf{v} + P = (3+4) \hat{\mathbf{x}} + (2+6) \hat{\mathbf{y}} + (6+0) \hat{\mathbf{z}} = (7) \hat{\mathbf{x}} + (8) \hat{\mathbf{y}} + (6) \hat{\mathbf{z}}, \\ |\mathbf{v}| &= \sqrt{(7)^2 + (8)^2 + (6)^2} = \sqrt{149} \quad \blacksquare \\ Q &= \mathbf{v} + P = (4+4) \hat{\mathbf{x}} + (2+6) \hat{\mathbf{y}} + (-2+0) \hat{\mathbf{z}} = (8) \hat{\mathbf{x}} + (8) \hat{\mathbf{y}} + (-2) \hat{\mathbf{z}}, \\ |\mathbf{v}| &= \sqrt{(8)^2 + (8)^2 + (-2)^2} = 2\sqrt{33} \quad \blacksquare \\ Q &= \mathbf{v} + P = (3+4) \hat{\mathbf{x}} + (-3+6) \hat{\mathbf{y}} + (3+0) \hat{\mathbf{z}} = (7) \hat{\mathbf{x}} + (3) \hat{\mathbf{y}} + (3) \hat{\mathbf{z}}, \\ |\mathbf{v}| &= \sqrt{(7)^2 + (3)^2 + (3)^2} = 2\sqrt{67} \quad \blacksquare \end{aligned}$$

The operator ( $\nabla$ ), was originally introduced by William Rowan Hamilton (1805-1865). Hamilton wrote the operator as a rotated nabla and it was P. G. Tait who established An upside-down delta as the conventional symbol in *An Elementary Treatise on Quaternions* (1867). Tait was also responsible for establishing the term *nabla*.

David Wilkins suggests Hamilton may have used the nabla as a general purpose symbol or abbreviation for whatever operator he wanted to introduce at any time. In 1837 Hamilton used the nabla, in its modern orientation, as a symbol for any arbitrary function in *Trans. R. Irish Acad. XVII. 236.* (OED.) He used the nabla to signify a permutation operator in "On the Argument of Abel, respecting the Impossibility of expressing a Root of any General Equation above the Fourth Degree, by any finite Combination of Radicals and Rational Functions, (1839).

Hamilton used the rotated nabla ( $\nabla$ ), for the vector differential operator in the *Proceedings of the Royal Irish Academy* (1846). Hamilton also used the rotated nabla as the vector differential operator in *On Quaternions; or on a new System of Imaginaries in Algebra*. For more information, please visit [here](#).

### 2.2.3 The Del Operator

The gradient has the formal appearance of a vector, ( $\nabla$ )<sup>6</sup> multiplying a scalar  $T$ :

$$\nabla T = \left( \hat{\mathbf{x}} \frac{\partial}{\partial x} + \hat{\mathbf{y}} \frac{\partial}{\partial y} + \hat{\mathbf{z}} \frac{\partial}{\partial z} \right) T$$

The term in parentheses is called **del** operator:

$$\nabla = \hat{\mathbf{x}} \frac{\partial}{\partial x} + \hat{\mathbf{y}} \frac{\partial}{\partial y} + \hat{\mathbf{z}} \frac{\partial}{\partial z} \quad (2.8)$$

Del is **NOT** a vector, in the usual sense. It doesn't mean much until we provide it with a function to act upon. Furthermore, it does **NOT** "multiply"  $T$ . Rather, it is an instruction to **differentiate**

what follows. To be precise, then, we say that  $\nabla$  is a vector operator which *acts upon*  $T$ , not a vector that multiplies  $T$ .

With this qualification, though,  $\nabla$  mimics the behaviour of an ordinary vector in virtually every way. Almost anything that can be done with other vectors can also be done with  $\nabla$ . Now, an ordinary vector  $\mathbf{A}$  can multiply in three (3) ways:

1. By a scalar  $a$ :  $\mathbf{A}a$ ;
2. By a vector  $\mathbf{B}$ , via the dot product:  $\mathbf{A} \cdot \mathbf{B}$ ;
3. By a vector  $\mathbf{B}$  via the cross product:  $\mathbf{A} \times \mathbf{B}$ .

Correspondingly, there are three ways the operator  $\nabla$  can act:

1. On a scalar function  $T$ :  $\nabla T$  (the gradient);
2. On a vector function  $\mathbf{v}$ , via the dot product:  $\nabla \cdot \mathbf{v}$  (divergence)
3. On a vector function  $\mathbf{v}$ , via the cross product:  $\nabla \times \mathbf{v}$  (curl).

It is time to examine the other two (2) vector derivatives: divergence and curl.

**Divergence** From the definition of  $\nabla$  we construct the divergence as follows:

$$\begin{aligned}\nabla \cdot \mathbf{v} &= \left( \left( \frac{\partial}{\partial x} \right) \hat{x} + \left( \frac{\partial}{\partial y} \right) \hat{y} + \left( \frac{\partial}{\partial z} \right) \hat{z} \right) \cdot \left( (v_x) \hat{x} + (v_y) \hat{y} + (v_z) \hat{z} \right) \\ &= \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}.\end{aligned}$$

Observe that the divergence of a vector function  $\mathbf{v}$  is itself a scalar  $\nabla \cdot \mathbf{v}$ .

Let's try to visualise this concept. The name divergence is well chosen, for  $\nabla \cdot \mathbf{v}$  is a measure of how much  $\mathbf{v}$  spreads out<sup>7</sup> from the initial point.

<sup>7</sup>i.e., diverges

For example, a vector function in **Fig. 2.4** has a large (positive) divergence,<sup>8</sup> two functions in **Fig. 2.4**

<sup>8</sup>if the arrows pointed in, it would be a negative divergence

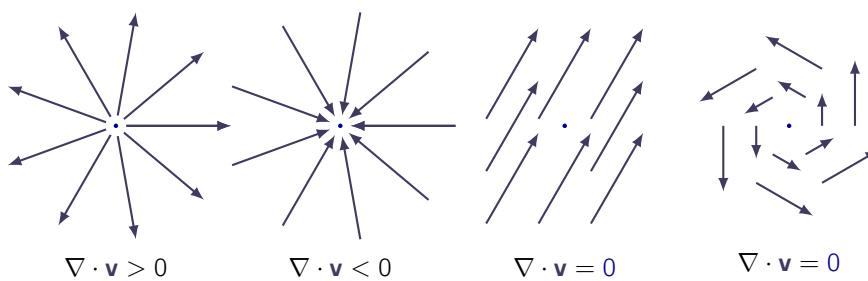


Figure 2.4: Visual description of the divergence operation.

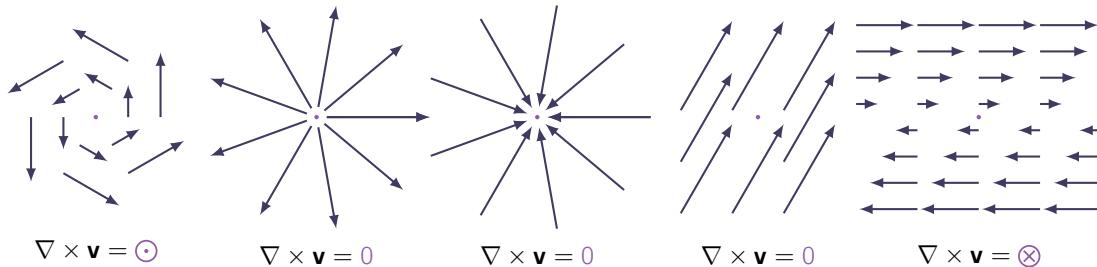


Figure 2.5: Different behaviours of the curl operation.

has zero divergence, and one function has a negative divergence.

As an example, imagine standing at the edge of a pond. Sprinkle some sawdust on the surface.

1. If the material spreads out, then we dropped it at a point of positive divergence;
2. If it collects together, we dropped it at a point of negative divergence.

**Curl** From the definition of  $\nabla$  we construct the curl:

$$\begin{aligned}\mathbf{v} \times \mathbf{v} &= \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ v_x & v_y & v_z \end{vmatrix} \\ &= \hat{\mathbf{x}} \left( \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} \right) + \hat{\mathbf{y}} \left( \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) + \hat{\mathbf{z}} \left( \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right).\end{aligned}$$

As with divergence, the name curl is also well chosen, for  $\nabla \times \mathbf{v}$  is a measure of how much  $\mathbf{v}$  swirls around the point in question. Therefore the functions in Fig. 1.18 all have zero curl, whereas the functions in **Fig. 2.5** all have a substantial curl sans one, pointing in the  $\hat{\mathbf{z}}$  direction, as the natural right-hand rule would suggest.

To finish these two (2) important operations, let's again imagine we are standing at the edge of a pond. Float a small flower, if it starts to rotate, then we placed it at a point of nonzero curl. A whirlpool would be a region of large curl.

### Exercise 2.3 An Example of a Curl

Find the curl ( $\nabla \times$ ) of the following functions.

$$\begin{aligned}\mathbf{v} &= (y) \hat{\mathbf{x}} + (2x^2) \hat{\mathbf{y}} + (0) \hat{\mathbf{z}}, & \mathbf{v} &= (y^n) \hat{\mathbf{x}} + (z^n) \hat{\mathbf{y}} + (x^n) \hat{\mathbf{z}}, \\ \mathbf{v} &= (\sin y) \hat{\mathbf{x}} + (\cos z) \hat{\mathbf{y}} + (-\tan x) \hat{\mathbf{z}}, & \mathbf{v} &= (x^2 - z) \hat{\mathbf{x}} + (xe^z) \hat{\mathbf{y}} + (xy) \hat{\mathbf{z}}.\end{aligned}$$

**SOLUTION** The curl ( $\nabla \times$ ) of the functions are as follows:

$$\begin{aligned}f(x, y, z) &= (y) \hat{\mathbf{x}} + (2x^2) \hat{\mathbf{y}} + (0) \hat{\mathbf{z}}, \\ \nabla \times f &= (0) \hat{\mathbf{x}} + (0) \hat{\mathbf{y}} + (-1 + 4x) \hat{\mathbf{z}}.\end{aligned}$$

$$\begin{aligned}
 f(x, y, z) &= (y^n) \hat{x} + (z^n) \hat{y} + (x^n) \hat{z}, \\
 \nabla \times f &= (-nz^{n-1}) \hat{x} + (-nx^{n-1}) \hat{y} + (-ny^{n-1}) \hat{z}. \\
 f(x, y, z) &= (\sin y) \hat{x} + (\cos z) \hat{y} + (-\tan x) \hat{z}, \\
 \nabla \times f &= (\sin z) \hat{x} + (\sec^2 x) \hat{y} + (-\cos y) \hat{z}. \\
 f(x, y, z) &= (x^2 - z) \hat{x} + (xe^z) \hat{y} + (xy) \hat{z}, \\
 \nabla \times f &= (x - e^z x) \hat{x} + (-1 - y) \hat{y} + (e^z) \hat{z}.
 \end{aligned}$$

## 2.2.4 Product Rules

The calculation of ordinary derivatives is facilitated by a number of rules which are as follows:

Table 2.1: Product rules of ordinary derivatives

Sum Rule	$\frac{d}{dx}(f + g) = \frac{df}{dx} + \frac{dg}{dx}$
Constant Multiplication	$\frac{d}{dx}(kf) = k \frac{df}{dx}$
Product Rule	$\frac{d}{dx}(fg) = f \frac{dg}{dx} + g \frac{df}{dx}$
Quotient Rule	$\frac{d}{dx}\left(\frac{f}{g}\right) = \frac{g \frac{df}{dx} - f \frac{dg}{dx}}{g^2}$

Similar relations hold for the vector derivatives. Thus,

$$\begin{aligned}
 \nabla(f + g) &= \nabla f + \nabla g, \\
 \nabla \cdot (\mathbf{A} + \mathbf{B}) &= (\nabla \cdot \mathbf{A}) + (\nabla \cdot \mathbf{B}), \\
 \nabla \times (\mathbf{A} + \mathbf{B}) &= (\nabla \times \mathbf{A}) + (\nabla \times \mathbf{B}),
 \end{aligned}$$

and

$$\nabla(kf) = k\nabla f, \quad \nabla \cdot (k\mathbf{A}) = k(\nabla \cdot \mathbf{A}), \quad \nabla \times (k\mathbf{A}) = k(\nabla \times \mathbf{A}), \quad (2.9)$$

The product rules, on the other hand, are not quite so simple. There are two (2) ways to construct a scalar as the product of two functions:

$$\begin{aligned}
 fg &\quad (\text{product of two scalar functions}), \\
 \mathbf{A} \cdot \mathbf{B} &\quad (\text{dot product of two vector functions}),
 \end{aligned}$$

and two ways to make a vector:

$$\begin{aligned}
 f\mathbf{A} &\quad (\text{scalar times vector}), \\
 \mathbf{A} \times \mathbf{B} &\quad (\text{cross product of two vectors}).
 \end{aligned}$$

Accordingly, there are six (6) product rules, which are given in **Tbl.** 2.2.

Table 2.2: Product rules of vector operations

Gradient I	$\nabla(fg) = f\nabla g + g\nabla f$
Gradient II	$\nabla(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A}$
Divergence I	$\nabla \cdot (f\mathbf{A}) = f(\nabla \cdot \mathbf{A}) + \mathbf{A} \cdot (\nabla f)$
Divergence II	$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B})$
Curl I	$\nabla \times (f\mathbf{A}) = f(\nabla \times \mathbf{A}) - \mathbf{A} \times (\nabla f),$
Curl II	$\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} + \mathbf{A}(\nabla \cdot \mathbf{B}) - \mathbf{B}(\nabla \cdot \mathbf{A})$

The proofs for rules given in **Tbl.** 2.2 come straight from the product rule for ordinary derivatives. As an example:

$$\begin{aligned}\nabla \cdot (f\mathbf{A}) &= \frac{\partial}{\partial x}(fA_x) + \frac{\partial}{\partial y}(fA_y) + \frac{\partial}{\partial z}(fA_z) \\ &= \left( \frac{\partial f}{\partial x}A_x + f\frac{\partial A_x}{\partial x} \right) + \left( \frac{\partial f}{\partial y}A_y + f\frac{\partial A_y}{\partial y} \right) + \left( \frac{\partial f}{\partial z}A_z + f\frac{\partial A_z}{\partial z} \right) \\ &= (\nabla f) \cdot \mathbf{A} + f(\nabla \cdot \mathbf{A}).\end{aligned}$$

It is also possible to formulate three quotient rules:

$$\begin{aligned}\nabla \left( \frac{f}{g} \right) &= \frac{g\nabla f - f\nabla g}{g^2}, \\ \nabla \cdot \left( \frac{\mathbf{A}}{g} \right) &= \frac{g(\nabla \cdot \mathbf{A}) - \mathbf{A} \cdot (\nabla g)}{g^2}, \\ \nabla \times \left( \frac{\mathbf{A}}{g} \right) &= \frac{g(\nabla \times \mathbf{A}) + \mathbf{A} \times (\nabla g)}{g^2}.\end{aligned}$$

However, given these can be obtained quickly from the previously mentioned product rules, there is no point in listing them separately and is left for the reader as exercise.

## 2.2.5 Second Derivatives

The gradient, the divergence, and the curl are the only first derivatives we can make with  $\nabla \cdot \mathbf{v}$  by applying  $\nabla$  twice, we can construct five (5) types of 2<sup>nd</sup> derivatives.

The gradient  $\nabla T$  is a vector, so we can take the divergence and curl of it:

1. Divergence of gradient:  $\nabla \cdot (\nabla T)$ .
2. Curl of gradient:  $\nabla \times (\nabla T)$ .

The divergence  $\nabla \cdot \mathbf{v}$  is a scalar, therefore all we can do is take its gradient:

3. Gradient of divergence:  $\nabla(\nabla \cdot \mathbf{v})$ .

The curl  $\nabla \times \mathbf{v}$  is a vector, so we can take its divergence and curl:

4. Divergence of curl:  $\nabla \cdot (\nabla \times \mathbf{v})$ .5. Curl of curl:  $\nabla \times (\nabla \times \mathbf{v})$ .

This exhausts the possible combinations, and in fact **NOT** all of them give anything new. Let's consider them one at a time:

**Divergence of a Gradient**

$$\begin{aligned}\nabla \cdot (\nabla T) &= \left( \hat{\mathbf{x}} \frac{\partial}{\partial x} + \hat{\mathbf{y}} \frac{\partial}{\partial y} + \hat{\mathbf{z}} \frac{\partial}{\partial z} \right) \cdot \left( \frac{\partial T}{\partial x} \hat{\mathbf{x}} + \frac{\partial T}{\partial y} \hat{\mathbf{y}} + \frac{\partial T}{\partial z} \hat{\mathbf{z}} \right) \\ &= \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}.\end{aligned}$$

This object, which we write as  $\nabla^2 T$  for short, is called the Laplacian<sup>9</sup> of  $T$ , which will be our focus later.

The Laplacian of a scalar  $T$  is a scalar value.

Occasionally, we will use the Laplacian of a vector,  $\nabla^2 \mathbf{v}$ . By this we mean a **vector** quantity whose  $x$ -component is the Laplacian of  $v_x$ , and so on.

$$\nabla^2 \mathbf{v} \equiv \left( \nabla^2 v_x \right) \hat{\mathbf{x}} + \left( \nabla^2 v_y \right) \hat{\mathbf{y}} + \left( \nabla^2 v_z \right) \hat{\mathbf{z}}$$

This is nothing more than a convenient extension of the meaning of  $\nabla^2$ .

**Exercise 2.4 | The Laplacian of a Vector**

Calculate the Laplacian of the following functions:

$$\begin{aligned}(i) \quad T_a &= x_2 + 3xy + 3z + 4, & (ii) \quad T_b &= \sin x \sin y \sin z, \\ (iii) \quad T_c &= e^{-5x} \sin 4y \cos 3z, & (iv) \quad \mathbf{v} &= (x^2) \hat{\mathbf{x}} + (3xz^2) \hat{\mathbf{y}} + (-2xz) \hat{\mathbf{z}}.\end{aligned}$$

**SOLUTION** The solution to the Laplacian of the functions are as follows:

$$\begin{aligned}(i) \quad \frac{\partial^2 T_a}{\partial x^2} &= 2; \quad \frac{\partial^2 T_a}{\partial y^2} = 0; \quad \frac{\partial^2 T_a}{\partial z^2} = 0 \quad \rightarrow \quad \nabla^2 T_a = 2 \quad \blacksquare \\ (ii) \quad \frac{\partial^2 T_b}{\partial x^2} &= \frac{\partial^2 T_b}{\partial y^2} = \frac{\partial^2 T_b}{\partial z^2} = -3T_b \quad \rightarrow \quad \nabla^2 T_b = -3T_b = 3 \sin x \sin y \sin z \quad \blacksquare \\ (iii) \quad \frac{\partial^2 T_c}{\partial x^2} &= 25T_c; \\ \frac{\partial^2 T_c}{\partial y^2} &= -16T_c; \quad \frac{\partial^2 T_c}{\partial z^2} = -9T_c \quad \rightarrow \quad \nabla^2 T_c = 0 \quad \blacksquare \\ (iv) \quad \frac{\partial^2 v_x}{\partial x^2} &= 2; \quad \frac{\partial^2 v_x}{\partial y^2} = 0; \quad \frac{\partial^2 v_x}{\partial z^2} = 0 \quad \rightarrow \quad \nabla^2 v_x = 0, \\ \frac{\partial^2 v_y}{\partial x^2} &= 0; \quad \frac{\partial^2 v_y}{\partial y^2} = 0; \quad \frac{\partial^2 v_y}{\partial z^2} = 6 \quad \rightarrow \quad \nabla^2 v_y = 6x,\end{aligned}$$

<sup>9</sup>A differential operator given by the divergence of the gradient of a scalar function on Euclidean space. It is usually denoted by the symbols  $\nabla \cdot \nabla$ ,  $\nabla^2$  or  $\Delta$ . In a Cartesian coordinate system, the Laplacian is given by the sum of second partial derivatives of the function with respect to each independent variable.

The Laplace operator is named after the French mathematician Pierre-Simon de Laplace (1749-1827), who first applied the operator to the study of celestial mechanics: the Laplacian of the gravitational potential due to a given mass density distribution is a constant multiple of that density distribution. Solutions of Laplace's equation  $\Delta f = 0$  are called harmonic functions and represent the possible gravitational potentials in regions of vacuum.

The Laplacian occurs in many differential equations describing physical phenomena. Poisson's equation describes electric and gravitational potentials; the diffusion equation describes heat and fluid flow; the wave equation describes wave propagation; and the Schrödinger equation describes the wave function in quantum mechanics. In image processing and computer vision, the Laplacian operator has been used for various tasks, such as blob and edge detection.

$$\frac{\partial^2 v_z}{\partial x^2} = 0; \frac{\partial^2 v_z}{\partial y^2} = 0; \frac{\partial^2 v_z}{\partial z^2} = 0 \quad \rightarrow \quad \nabla^2 v_z = 0,$$

$\nabla^2 v = 2\hat{x} + 6x\hat{y}$  ■

---

### Curl of a Gradient

The curl of a gradient is **always** zero:

$$\nabla \times (\nabla T)$$

This is an **important fact**, which will be used repeatedly. Without going into too much detail into the proof, it relies on the following relation:

$$\frac{\partial}{\partial x} \left( \frac{\partial T}{\partial y} \right) = \frac{\partial}{\partial y} \left( \frac{\partial T}{\partial x} \right)$$

### Gradient of Divergence

This operation rarely occurs in physical applications, and it has not been given any special name of its own. Notice that  $\nabla(\nabla \cdot \mathbf{v})$  is **NOT** the same as the Laplacian of a vector:

$$\nabla^2 = (\nabla \cdot \nabla) \neq \nabla(\nabla \cdot \mathbf{v})$$

### Divergence of a Curl

Similar to the curl of a gradient, it is always zero:

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

### Curl of a Curl

As we can check from the definition of  $\nabla$ :

$$\nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \nabla^2 \mathbf{v}. \quad (2.10)$$

So curl-of-curl gives nothing new as the first term is just the divergence of a curl, and the second is the Laplacian. To put it short, then, there are just two kinds of second derivatives:<sup>10</sup>

1. Laplacian,
2. Gradient of divergence.

<sup>10</sup>It is possible to work out 3<sup>rd</sup> derivatives, but fortunately second derivatives suffice for practically all physical applications.

## 2.3 Integral Calculus

### 2.3.1 Line, Surface, and Volume Integrals

In electrodynamics, we encounter several different kinds of integrals, among which the most important are **line** (or **path**) **integrals**, **surface integrals** (or **flux**), and **volume integrals**, which will be the focus of this section.

**Line Integrals** Has an expression of the form:

$$\int_a^b \mathbf{v} \cdot d\mathbf{l}$$

where  $\mathbf{v}$  is a vector function,  $d\mathbf{l}$  is the infinitesimal displacement vector, and the integral is to be carried out along a prescribed path  $\mathcal{P}$  from point  $a$  to point  $b$ . If the path forms a closed loop,<sup>11</sup>

i.e., if  $b = a$ .

We put a circle on the integral sign:

$$\oint \mathbf{v} \cdot d\mathbf{l}$$

At each point on the path, we take the dot product of  $\mathbf{v}$  (evaluated at that point) with the displacement  $d\mathbf{l}$  to the next point on the path.

A good example of a line integral is the work done by a force  $\mathbf{F}$ :

$$W = \int \mathbf{F} \cdot d\mathbf{l}$$

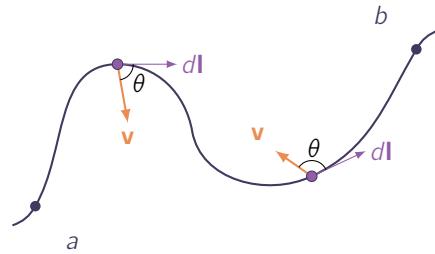


Figure 2.6: The method in which line integral is calculated. At each point the dot product of the vector ( $\mathbf{v}$ ) is taken with the length vector ( $d\mathbf{l}$ ) which is always tangential to the point in which the integration is taken.

Ordinarily, the value of a line integral depends critically on the path taken from  $a$  to  $b$ , but there is an important special class of vector functions for which the line integral is independent of path and is determined entirely by the end points.

It will be our business in due course to characterise this special class of vectors.

A force which has this property is called **conservative**.

### Exercise 2.5 Fluid Flow

A fluid's velocity field is  $\mathbf{F} = (x) \hat{\mathbf{x}} + (z) \hat{\mathbf{y}} + (y) \hat{\mathbf{z}}$ . Find the flow along the helix  $\mathbf{l}(t) = (\cos t) \hat{\mathbf{x}} + (\sin t) \hat{\mathbf{y}} + (t) \hat{\mathbf{z}}$  with a range of  $0 \leq t \leq \pi/2$ .

**SOLUTION** We first evaluate  $\mathbf{F}$  on the curve:

$$\mathbf{F} = (x) \hat{\mathbf{x}} + (z) \hat{\mathbf{y}} + (y) \hat{\mathbf{z}} = (\cos t) \hat{\mathbf{x}} + (t) \hat{\mathbf{y}} + (\sin t) \hat{\mathbf{z}} \quad \text{Substitute } x = \cos t, z = t, y = \sin t.$$

and then find  $d\ell/dt$ :

$$\frac{d\ell}{dt} = (-\sin t) \hat{x} + (\cos t) \hat{y} + (0) \hat{z}.$$

Then we integrate  $\mathbf{F} \cdot (d\ell/dt)$  from  $t = 0$  to  $t = \pi/2$ :

$$\begin{aligned}\mathbf{F} \cdot \frac{d\ell}{dt} &= (\cos t) (-\sin t) + (t) (\cos t) + (\sin t) (1), \\ &= -\sin t \cos t + t \cos t + \sin t.\end{aligned}$$

Which makes,

$$\begin{aligned}\text{Flow} &= \int_{t=a}^{t=b} \mathbf{F} \cdot \frac{d\ell}{dt} dt = \int_0^{\pi/2} (-\sin t \cos t + t \cos t + \sin t) dt, \\ &= \left[ \frac{\cos^2 t}{2} + t \sin t \right] \Big|_0^{\pi/2} = \left( 0 + \frac{\pi}{2} \right) - \left( \frac{1}{2} + 0 \right) = \frac{\pi}{2} - \frac{1}{2} \quad \blacksquare\end{aligned}$$

### Exercise 2.6 Field Circulation

Find the circulation of the field  $\mathbf{F} = (x - y) \hat{x} + x \hat{y}$  around the circle  $\ell(t) = (\cos t) \hat{x} + (\sin t) \hat{y} + (0) \hat{z}$  with a range of  $0 \leq t \leq 2\pi$ .

**SOLUTION** On the circle,  $\mathbf{F} = (x - y) \hat{x} + (x) \hat{y} + (0) \hat{z} = (\cos t - \sin t) \hat{x} + (\cos t) \hat{y} + (0) \hat{z}$  and  $\frac{d\ell}{dt} = (-\sin t) \hat{x} + (\cos t) \hat{y} + (0) \hat{z}$ .

Then

$$\mathbf{F} \cdot \frac{d\ell}{dt} = -\sin t \cos t + \underbrace{\sin^2 t + \cos^2 t}_1,$$

Gives.

$$\begin{aligned}\text{Circulation} &= \int_0^{2\pi} \mathbf{F} \cdot \frac{d\ell}{dt} dt = \int_0^{2\pi} (1 - \sin t \cos t) dt \\ &= \left[ t - \frac{\sin^2 t}{2} \right] \Big|_0^{2\pi} = 2\pi \quad \blacksquare\end{aligned}$$

**Surface Integrals** A surface integral is an expression of the form:

$$\int_S \mathbf{v} \cdot d\mathbf{a}$$

where  $\mathbf{v}$  is a vector function, and the integral is over a specified surface  $S$ . Here  $d\mathbf{a}$  is an infinitesimal patch of area, with direction **perpendicular to the surface**. There are, two (2) directions perpendicular to any surface, so the **sign** of a surface integral is intrinsically ambiguous.

If the surface is closed,<sup>12</sup> we put a circle on the integral sign:

$$\oint \mathbf{v} \cdot d\mathbf{a}$$

Tradition dictates that **outward** is positive, but for open surfaces it's arbitrary.

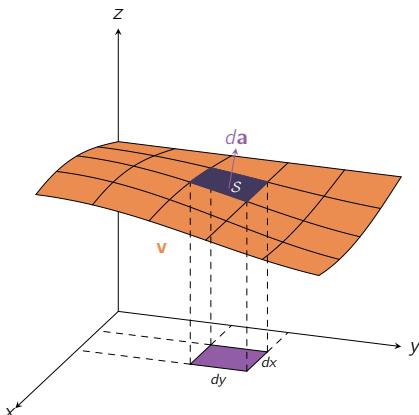


Figure 2.7: A visual description of the surface integral.

<sup>12</sup>imagine it forming a balloon.

As an example, if  $\mathbf{v}$  describes the flow of a fluid,<sup>13</sup> then  $\int \mathbf{v} \cdot d\mathbf{a}$  represents the total mass per unit time passing through the surface.

<sup>13</sup>Measured in mass per unit area per unit time.

Ordinarily, the value of a surface integral depends on the particular surface chosen, but there is a special class of vector functions for which it is **independent** of the surface and is determined entirely by the boundary line. An important task will be to characterise this special class of functions.

**Volume Integrals** A volume integral is an expression of the form:

$$\int_V T \, d\tau$$

where  $T$  is a scalar function and  $d\tau$  is an infinitesimal volume element. In Cartesian coordinates,

$$d\tau = dx \, dy \, dz$$

As an example, if  $T$  is the density of a substance,<sup>14</sup> then the volume integral would give the total mass.

<sup>14</sup>This might vary from point to point.

Occasionally we shall encounter volume integrals of vector functions:

$$\int \mathbf{v} \, d\tau = \int (v_x \hat{x} + v_y \hat{y} + v_z \hat{z}) \, d\tau = \hat{x} \int v_x \, d\tau + \hat{y} \int v_y \, d\tau + \hat{z} \int v_z \, d\tau.$$

As the unit vectors ( $\hat{x}$ ,  $\hat{y}$ ,  $\hat{z}$ ) are constants, they can be taken outside the integral.

### Exercise 2.7 Double Integrals

Find the following double integrals:

$$\begin{aligned} & \int_0^1 \int_x^{2x} (x+y)^2 \, dy \, dx, & & \int_0^1 \int_y^{\sqrt{y}} (1-2xy) \, dx \, dy, \\ & \int_0^3 \int_x^3 \cosh(x+y) \, dy \, dx, & & \int_0^1 \int_0^{y^3} \exp y^4 \, dx \, dy. \end{aligned}$$

#### SOLUTION

The solution to integrations are as follows:

$$\begin{aligned} \int_0^1 \int_x^{2x} (x+y)^2 \, dy \, dx &= \int_0^1 \int_x^{2x} x^2 + 2xy + y^2 \, dy \, dx = \int_0^1 \left[ yx^2 + xy^2 + \frac{y^3}{3} \right]_x^{2x} \, dx, \\ &= \int_0^1 \left( 4x^3 + \frac{7x^3}{3} \right) \, dx = \left[ 4x^3 + \frac{7x^4}{12} \right]_0^1 = \frac{19}{12} \quad \blacksquare \\ \int_0^1 \int_y^{\sqrt{y}} (1-2xy) \, dx \, dy &= \int_0^1 \left[ x - x^2 y \right]_y^{\sqrt{y}} \, dy, \\ &= \int_0^1 \left[ (\sqrt{y} - y^2) - (y - y^3) \right] \, dy = \int_0^1 \left[ y^3 + \sqrt{y} - y^2 - y \right] \, dy, \\ &= \left[ \frac{y^4}{4} + \frac{2}{3}y^{3/2} - \frac{y^3}{3} - \frac{y^2}{2} \right]_0^1 = \left( \frac{1}{4} + \frac{2}{3} - \frac{1}{3} - \frac{1}{2} \right) - (0) = \frac{1}{12} \quad \blacksquare \\ \int_0^3 \int_x^3 \cosh(x+y) \, dy \, dx &= \int_0^1 \left[ \sinh(x+y) \right]_x^3 \, dx = \int_0^1 \left[ \sinh(3+x) - \sinh(2x) \right] \, dx \end{aligned}$$



**Figure 2.8:** To measure the height of a mountain, it doesn't matter what way we take, as long as we know the base and the top, we will know the height.

### 2.3.2 The Fundamental Theorems of Vector Calculus

Assume  $f(x)$  is a function of one (1) variable. Based on this, the fundamental theorem of calculus says:

#### Theory 2.1: Calculus Theorem

The **integral** of a **derivative** over some **region** is given by the **value of the function** at the end points (**boundaries**)

$$\int_a^b \left( \frac{df}{dx} \right) dx = f(x) - f(a) \quad \text{or} \quad \int_a^b F(x) dx = f(x) - f(a)$$

In vector calculus there are three (3) species of derivative<sup>15</sup> and each has its own “fundamental theorem” with essentially the same format. Our purpose here is to not prove these theorems here, but rather, understand what they mean.

<sup>15</sup>These are gradient, divergence, and curl

**The Fundamental Theorem for Gradients** Suppose we have a scalar function of three (3) variables  $T(x, y, z)$ . Starting at point  $\mathbf{a}$ , move a small distance  $d\mathbf{l}_1$ . The function  $T$  will change by an amount:

$$dT = (\nabla T) \cdot d\mathbf{l}_1$$

Now let's move an additional small displacement  $d\mathbf{l}_2$ . The incremental change in  $T$  will be now:

$$dT = (\nabla T) \cdot d\mathbf{l}_2$$

In this manner, proceeding by infinitesimal steps, we make the journey to point  $\mathbf{b}$ . At each step we compute the gradient of  $T$ , at that point, and dot it into the displacement  $d\mathbf{l}$ ...this gives us the change in  $T$ .

#### Theory 2.2: Gradient Theorem

The total change in  $T$  in going from  $\mathbf{a}$  to  $\mathbf{b}$  (along the path selected) is:

$$\int_{\mathbf{a}}^{\mathbf{b}} (\nabla T) \cdot d\mathbf{l} = T(\mathbf{b}) - T(\mathbf{a})$$

Similar to “ordinary” fundamental theorem, this theorem says the integral<sup>16</sup> of a derivative, which here the gradient, is given by the value of the function at the boundaries which are  $\mathbf{a}$  and  $\mathbf{b}$  respectively.

<sup>16</sup>In this case it is a line integral

As an example, assume we want to measure the height of GrossGlockner. We could climb the mountain from base, or take the high alpine road, or take a helicopter ride all the way up to top. Regardless of the options we take, we should get the same answer either way.<sup>17</sup>

<sup>17</sup>This is the essence of the fundamental theorem

### Theory 2.3: Line Independence of Gradient

**Gradients** have the special property that their line integrals are path independent:

- $\int_a^b (\nabla T) \cdot dI$  is independent of the path taken from  $a$  to  $b$ .
- $\oint (\nabla T) \cdot dI = 0$ , since the beginning and end points are identical, and hence  $T(b) - T(a) = 0$ .

### The Fundamental Theorem for Divergences

This theorem has at least three (3) special names:

1. Gauss's theorem,
2. Green's theorem,
3. Divergence theorem.

The fundamental theorem for divergences states:

### Theory 2.4: Divergence Theorem

the **integral** of a **derivative** (the **divergence**) over a **region** (in this case the **volume**,  $\mathcal{V}$ ) is equal to the value of the function at the **boundary** (in this case the **surface**  $\mathcal{S}$  bounding the volume).

$$\int_{\mathcal{V}} (\nabla \cdot \mathbf{v}) \, d\tau = \oint_{\mathcal{S}} \mathbf{v} \cdot d\mathbf{a}.$$

The boundary term is itself an integral, more specifically, a surface integral. This is reasonable: the "boundary" of a line is just two end points, but the boundary of a volume is a closed surface. To create an analogy, if  $\mathbf{v}$  represents the flow of an incompressible fluid, then the flux  $\mathbf{v}$  is the total amount of fluid passing out through the surface, per unit time. Now, the divergence measures the *spreading out* of the vectors from a point, a place of high divergence is like a tap, pouring out liquid. If we have a bunch of tap in a region filled with incompressible fluid, an equal amount of liquid will be forced out through the boundaries of the region. In fact, there are two (2) ways we could determine how much is being produced:

- i. we could count up all the faucets, recording how much each puts out
- ii. we could go around the boundary, measuring the flow at each point, and add it all up

We get the same answer either way:

$$\int (\text{faucets within the volume}) = \oint (\text{flow out through the surface})$$

**Exercise 2.8** An Example of Divergence Theorem

Evaluate both sides of the Divergence theorem for the expanding vector field  $\mathbf{F} = (x) \hat{\mathbf{x}} + (y) \hat{\mathbf{y}} + (z) \hat{\mathbf{z}}$  over the sphere  $x^2 + y^2 + z^2 = a^2$

**SOLUTION** The outer unit normal to  $S$ , calculated from the gradient of  $f(x, y, z) = x^2 + y^2 + z^2 - a^2$ , is:

$$\hat{\mathbf{n}} = \frac{\nabla S}{|\nabla S|} = \frac{(2x) \hat{\mathbf{x}} + (2y) \hat{\mathbf{y}} + (2z) \hat{\mathbf{z}}}{\sqrt{4x^2 + 4y^2 + 4z^2}} = \frac{(x) \hat{\mathbf{x}} + (y) \hat{\mathbf{y}} + (z) \hat{\mathbf{z}}}{a}. \quad x^2 + y^2 + z^2 = a^2 \text{ on } S$$

Therefore:

$$(\mathbf{F} \cdot \hat{\mathbf{n}}) da = \frac{x^2 + y^2 + z^2}{a} da = \frac{a^2}{a} da = a da.$$

This in turn gives us:

$$\iint_S (\mathbf{F} \cdot \hat{\mathbf{n}}) da = \iint_S a da = a \iint_S da = a (4\pi a^2) = 4\pi a^3. \quad \text{Area of } S \text{ is } 4\pi a^2$$

The divergence of  $\mathbf{F}$  is:

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x} (x) + \frac{\partial}{\partial y} (y) + \frac{\partial}{\partial z} (z) = 3,$$

So,

$$\iiint_V (\nabla \cdot \mathbf{v}) dV = \iiint_V 3 dV = 3 \left( \frac{4}{3} \pi a^3 \right) = 4\pi a^3 \blacksquare$$

**Exercise 2.9** Divergence Theorem of an Octant of a Sphere

Check the divergence theorem for the function:

$$\mathbf{v} = (r^2 \cos \theta) \hat{\mathbf{r}} + (r^2 \cos \phi) \hat{\mathbf{\theta}} + (-r^2 \cos \theta \sin \phi) \hat{\mathbf{\phi}},$$

using as your volume one octant of the sphere of radius  $R$ .

**SOLUTION** It is always useful to write the theorem we are going to work on:

$$\begin{array}{c} \iiint_V (\nabla \cdot \mathbf{v}) dV = \iint_S \mathbf{v} \cdot \mathbf{n} da. \\ \text{Divergence integral} \qquad \text{Outward flux} \end{array}$$

First solve the left hand side of the equation:

$$\begin{aligned} \nabla \cdot \mathbf{v} &= \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 r^2 \cos \theta) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta r^2 \cos \phi) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} (-r^2 \cos \theta \sin \phi), \\ &= \frac{1}{r^2} 4r^3 \cos \theta + \frac{1}{r \sin \theta} \cos \theta r^2 \cos \phi + \frac{1}{r \sin \theta} (-r^2 \cos \theta \cos \phi), \\ &= \frac{r \cos \theta}{\sin \theta} [4 \sin \theta + \cos \phi - \cos \phi] = 4r \cos \theta. \end{aligned}$$

$$\begin{aligned} \int (\nabla \cdot \mathbf{v}) dV &= \int (4r \cos \theta) r^2 \sin \theta dr d\theta d\phi = 4 \int_0^R r^3 dr \int_0^{\pi/2} \cos \theta \sin \theta d\theta \int_0^{\pi/2} d\phi, \\ &= (R^4) \left( \frac{1}{2} \right) \left( \frac{\pi}{2} \right) = \frac{\pi R^4}{4} \blacksquare \end{aligned}$$

Now it is time to solve the right hand side of the question. As we are aware from the shape, an octant of the sphere has 4 sides to it: the curved surface  $xyz \rightarrow a_1$ , and  $xz \rightarrow a_2$ ,  $yz \rightarrow a_3$  and  $xy \rightarrow a_4$ . These are

$$\begin{aligned} da_1 &= \hat{\mathbf{r}} dl_\theta dl_\phi = \hat{\mathbf{r}} R^2 \sin \theta d\phi d\theta, & da_2 &= dl_r dl_\theta = -\hat{\mathbf{\theta}} r dr d\theta, \\ da_3 &= \hat{\mathbf{\phi}} dl_r dl_\theta = \hat{\mathbf{\phi}} r dr d\theta, & da_4 &= dl_r dl_\phi = \hat{\mathbf{\theta}} r dr d\theta. \quad (\theta = \pi/2) \end{aligned}$$

$$\begin{aligned} \iint_S \mathbf{v} \cdot da &= \iint_{S_1} \mathbf{v} \cdot da + \iint_{S_2} \mathbf{v} \cdot da + \iint_{S_3} \mathbf{v} \cdot da + \iint_{S_4} \mathbf{v} \cdot da, \\ &= \int_0^{\pi/2} \int_0^{\pi/2} [r^2 \cos \theta \hat{\mathbf{r}} + r^2 \cos \phi \hat{\mathbf{\theta}} - r^2 \cos \theta \sin \phi \hat{\mathbf{\phi}}] \Big|_{r=R} \cdot (\hat{\mathbf{r}} R^2 \sin \theta d\phi d\theta) \\ &\quad + \int_0^{\pi/2} \int_0^R [r^2 \cos \theta \hat{\mathbf{r}} + r^2 \cos \phi \hat{\mathbf{\theta}} - r^2 \cos \theta \sin \phi \hat{\mathbf{\phi}}] \Big|_{\phi=0} \cdot (-\hat{\mathbf{\theta}} r dr d\theta) \end{aligned}$$

$$\begin{aligned}
 & + \int_0^{\pi/2} \int_0^R [r^2 \cos \theta \hat{r} + r^2 \cos \phi \hat{\theta} - r^2 \cos \theta \sin \phi \hat{\phi}] \Big|_{\phi=\pi/2} \cdot (\hat{\phi} r dr d\theta) \\
 & + \int_0^{\pi/2} \int_0^R [r^2 \cos \theta \hat{r} + r^2 \cos \phi \hat{\theta} - r^2 \cos \theta \sin \phi \hat{\phi}] \Big|_{\theta=\pi/2} \cdot (\hat{\theta} r dr d\theta) ,
 \end{aligned}$$

Time to do some integration.

$$\begin{aligned}
 \iint_S \mathbf{v} \cdot d\mathbf{a} = & \int_0^{\pi/2} \int_0^{\pi/2} [R^2 \cos \theta \hat{r} + R^2 \cos \phi \hat{\theta} - R^2 \cos \theta \sin \phi \hat{\phi}] \cdot (\hat{r} R^2 \sin \theta d\phi d\theta) \\
 & + \int_0^{\pi/2} \int_0^R [r^2 \cos \theta \hat{r} + r^2(1) \hat{\theta} - (0) \sin \phi \hat{\phi}] \cdot (-\hat{\phi} r dr d\theta) \\
 & + \int_0^{\pi/2} \int_0^R [r^2 \cos \theta \hat{r} + (0)\phi \hat{\theta} - r^2 \cos \theta (1) \hat{\phi}] \cdot (\hat{\phi} r dr d\theta) \\
 & + \int_0^{\pi/2} \int_0^R [(0) \hat{r} + r^2 \cos \phi \hat{\theta} - (0) \hat{\phi}] \cdot (\hat{\theta} r dr d\theta) .
 \end{aligned}$$

Final touches and cleaning up,

$$\begin{aligned}
 \iint_S \mathbf{v} \cdot d\mathbf{a} = & \int_0^{\pi/2} \int_0^{\pi/2} R^4 \sin \theta \cos \theta d\phi d\theta + 0 - \overbrace{\int_0^{\pi/2} \int_0^R r^3 \cos \theta dr d\theta + \int_0^{\pi/2} \int_0^R r^3 \cos \theta dr d\phi}^{=0} , \\
 = & R^4 \left( \int_0^{\pi/2} d\phi \right) \left( \int_0^{\pi/2} \sin \theta \cos \theta d\theta \right) , \\
 = & R^4 \left( \frac{\pi}{2} \right) \left( \frac{\pi}{2} \right) , \\
 = & \frac{\pi R^4}{4} \quad \blacksquare
 \end{aligned}$$



<sup>18</sup>Sir George Gabriel Stokes, 1st Baronet, (1819 - 1903)

### 2.3.3 The Fundamental Theorem for Curls

The fundamental theorem for curls, also known as **Stokes' theorem**, states:<sup>18</sup>

#### Theory 2.5: Stokes' Theorem

**Integral** of a **derivative** over a **region** ( $\mathcal{S}$ ) is equal to the value of the function at **boundary** ( $\mathcal{P}$ ).

$$\int_{\mathcal{S}} (\nabla \times \mathbf{v}) \cdot d\mathbf{a} = \oint_{\mathcal{P}} \mathbf{v} \cdot d\mathbf{l}.$$

Similar to the divergence theorem, the boundary term is itself an integral. Specifically, a *closed line integral*.

Remember the curl measures the *twist* of the vectors  $\mathbf{v}$ . Think of a region of high curl as a whirlpool, where if we put a wheel there, it will rotate. Now, the integral of the curl over some surface (or, more precisely, the *flux* of the curl through the surface) represents the *total amount of swirl*, and we can determine that just as well by going around the edge and finding how much the flow is following the boundary.

$\oint \mathbf{v} \cdot d\mathbf{l}$  is sometimes called the **circulation** of  $\mathbf{v}$ .

was an Irish mathematician and physicist. Born in County Sligo, Ireland, Stokes spent his entire career at the University of Cambridge, where he served as the Lucasian Professor of Mathematics for 54 years, from 1849 until his death in 1903, the longest tenure held by the Lucasian Professor. As a physicist, Stokes made seminal contributions to fluid mechanics, including the Navier-Stokes equations; and to physical optics, with notable works on polarisation and fluorescence. As a mathematician, he popularised "Stokes' theorem" in vector calculus and contributed to the theory of asymptotic expansions. Stokes, along with Felix Hoppe-Seyler, first demonstrated the oxygen transport function of haemoglobin, and showed colour changes produced by the aeration of haemoglobin solutions.

There seems to be an ambiguity in Stokes' theorem: concerning the boundary line integral:

Which way are we supposed to go around?<sup>19</sup>

<sup>19</sup>clockwise or  
counterclockwise.

The answer is that it doesn't matter which way we go **as long as we are consistent**, for there is an additional sign ambiguity in the surface integral:

Which way does  $d\mathbf{a}$  point?

For a closed surface,<sup>20</sup>  $d\mathbf{a}$  points in the direction of the outward normal. But for an open surface, which way would be defined as out? Consistency in Stokes' theorem is given by the right-hand rule. If our rings point in the direction of the line integral, then our thumb fixes the direction of  $d\mathbf{a}$ .

<sup>20</sup>i.e., the divergence theorem.

Ordinary, a flux integral depends critically on what surface we integrate over, but this is **NOT** the case with curls. For Stokes' theorem says that  $\int(\nabla \times \mathbf{v}) \cdot d\mathbf{a}$  is equal to the line integral of  $\mathbf{v}$  around the boundary, and the latter makes no reference to the specific surface we choose.

**Preposition I**  $\int(\nabla \times \mathbf{v}) \cdot d\mathbf{a}$  depends only on the boundary line, not on the particular surface used.

**Preposition II**  $\oint(\nabla \times \mathbf{v}) \cdot d\mathbf{a} = 0$  for any closed surface.

### Exercise 2.10 Surface Area of an Implicit Surface

Find the area of the surface cut from the bottom of the paraboloid  $x^2 + y^2 - z = 0$  by the plane  $z = 4$ .

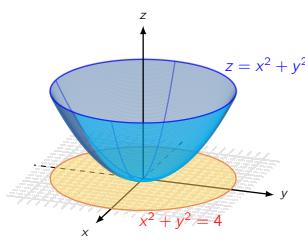


Figure 2.9: The paraboloid of "Surface Area of an Implicit Surface".

**SOLUTION** We sketch the surface  $S$  and the region  $R$  below it in the  $xy$ -plane (Fig. 2.9). The surface  $S$  is part of the level surface  $F(x, y, z) = x^2 + y^2 - z = 0$ , and  $R$  is the disk  $x^2 + y^2 \leq 4$  in the  $xy$ -plane.

To get a unit vector normal (i.e.,  $\hat{\mathbf{n}}$ ) to the plane  $R$ , we can take  $\hat{\mathbf{n}} = \hat{\mathbf{z}}$ . At any point  $(x, y, z)$  on the surface, we have:

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

$$\nabla F = (2x) \hat{\mathbf{x}} + (2y) \hat{\mathbf{y}} + (-1) \hat{\mathbf{z}}$$

$$\begin{aligned} |\nabla F| &= \sqrt{(2x)^2 + (2y)^2 + (-1)^2} \\ &= \sqrt{4x^2 + 4y^2 + 1} \\ |\nabla F \cdot \hat{\mathbf{n}}| &= |\nabla F \cdot \hat{\mathbf{z}}| = |-1| = 1. \end{aligned}$$

In the region  $R$ , the area is defined to be  $dA = dx dy$ . Therefore:

$$\begin{aligned} \text{Surface Area} &= \iint_R \frac{|\nabla F|}{|\nabla F \cdot \hat{\mathbf{n}}|} dA \\ &= \iint_{x^2+y^2 \leq 4} \sqrt{4x^2 + 4y^2 + 1} dx dy \\ &= \int_0^{2\pi} \int_0^2 \sqrt{4r^2 + 1} r dr d\theta \\ &= \int_0^{2\pi} \frac{1}{12} (4r^2 + 1)^{3/2} \Big|_0^2 d\theta \\ &= \int_0^{2\pi} \frac{1}{12} (17^{3/2} - 1) d\theta \\ &= \frac{\pi}{6} (17\sqrt{17} - 1) \blacksquare \end{aligned}$$

**Exercise 2.11** Stokes' Theorem Over a Hemisphere

Evaluate Stokes's theorem for the hemisphere  $S : x^2 + y^2 + z^2 = 9, z \geq 0$ , its bounding circle  $C : x^2 + y^2 = 9, z = 0$  and the field  $\mathbf{F} = (y) \hat{\mathbf{x}} + (-x) \hat{\mathbf{y}} + (0) \hat{\mathbf{z}}$ .

**Tip:** Parametrisation of a circle is:  $x = r \cos \theta, y = r \sin \theta$  and  $da = \frac{3}{z} dA$

**SOLUTION** We start by calculating the counter-clockwise circulation around  $C$  using the following parametrisation:

$$\mathbf{C}(\theta) = (3 \cos \theta) \hat{\mathbf{x}} + (3 \sin \theta) \hat{\mathbf{y}} + (0) \hat{\mathbf{z}},$$

where  $0 \leq \theta \leq 2\pi$ .

Using this we can calculate the counter-clockwise circulation.

$$d\mathbf{C} = (-3 \sin \theta d\theta) \hat{\mathbf{x}} + (3 \cos \theta d\theta) \hat{\mathbf{y}} + (0) \hat{\mathbf{z}},$$

$$\mathbf{F} = (y) \hat{\mathbf{x}} + (-x) \hat{\mathbf{y}} + (0) \hat{\mathbf{z}}$$

$$= (3 \sin \theta) \hat{\mathbf{x}} + (-3 \cos \theta) \hat{\mathbf{y}} + (0) \hat{\mathbf{z}},$$

$$\mathbf{F} \cdot d\mathbf{C} = -9 \sin^2 \theta d\theta - 9 \cos^2 \theta d\theta = -9 d\theta,$$

$$\oint_C \mathbf{F} \cdot d\mathbf{C} = \int_0^{2\pi} -9 d\theta = -18\pi.$$

For the curl of integral we have:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ F_x & F_y & F_z \end{vmatrix}$$

$$\begin{aligned} &= (0 - 0) \hat{\mathbf{x}} + (0 - 0) \hat{\mathbf{y}} + (-1 - 1) \hat{\mathbf{z}} = -2 \hat{\mathbf{z}} \\ \hat{\mathbf{n}} &= \frac{\nabla S}{|\nabla S|} = \frac{(x) \hat{\mathbf{x}} + (y) \hat{\mathbf{y}} + (z) \hat{\mathbf{z}}}{\sqrt{x^2 + y^2 + z^2}} \\ &= \frac{(x) \hat{\mathbf{x}} + (y) \hat{\mathbf{y}} + (z) \hat{\mathbf{z}}}{3} \quad \text{Unit normal} \end{aligned}$$

Now it is time to define the area of integration ( $da$ ):

$$\begin{aligned} da &= \frac{|\nabla S|}{|\nabla S \cdot \hat{\mathbf{z}}|} dA \\ &= \frac{|(2x) \hat{\mathbf{x}} + (2y) \hat{\mathbf{y}} + (2z) \hat{\mathbf{z}}|}{2z} dA \\ &= \frac{2 \sqrt{x^2 + y^2 + z^2}}{2z} dA \\ &= \frac{3}{z} dA, \\ \nabla \times \mathbf{F} \cdot \mathbf{n} da &= -\frac{2z}{3} \frac{3}{z} dA = -2 dA \end{aligned}$$

The cardinal direction  $\hat{\mathbf{z}}$  comes from being the direction perpendicular to the surface ( $S$ ).

$$\iint_S \nabla \times \mathbf{F} \cdot \mathbf{n} da = \iint_{x^2+y^2 \leq 9} -2 dA = -18\pi$$

The circulation around the circle equals the integral of the curl over the hemisphere ■

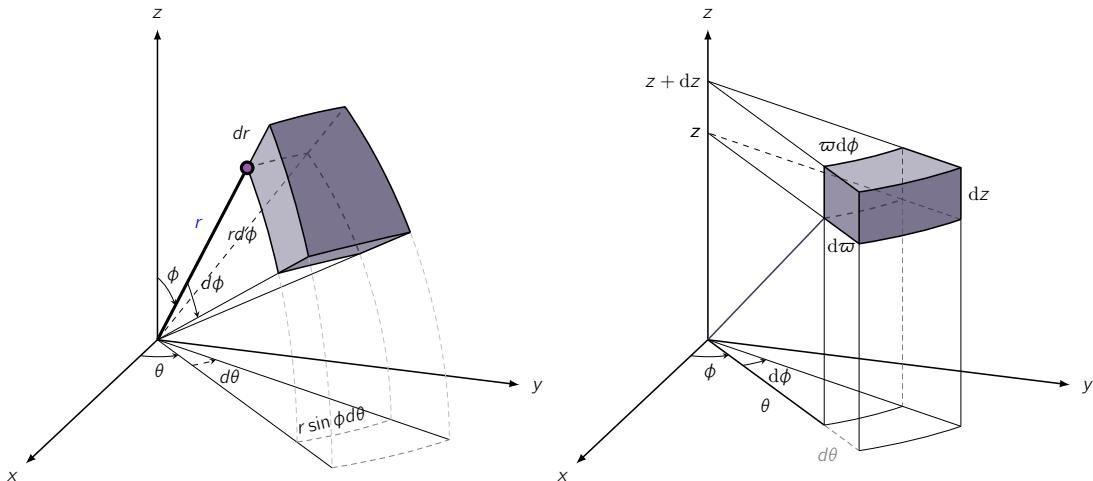


Figure 2.10: The two types of coordinate systems in question. (a) Spherical coordinate system (b) Spherical coordinate system.

## 2.4 Curvilinear Coordinates

### 2.4.1 Spherical Coordinate System

It is possible to label a point  $P$  in Cartesian coordinates  $(x, y, z)$ , but sometimes it is more convenient to use **spherical** coordinates  $(r, \theta, \phi)$ ;  $r$  is the distance from the origin,  $\theta$  is called the **polar angle**, and  $\phi$  is the **azimuthal angle**. Their relation to Cartesian coordinates can be read from **Fig. 2.10a**.

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

**Fig. 2.10a** also shows three unit vectors,  $\hat{r}$ ,  $\hat{\theta}$ ,  $\hat{\phi}$ , pointing in the direction of increase of the corresponding coordinates.

They constitute an **orthogonal** basis set,<sup>21</sup> similar to  $\hat{x}$ ,  $\hat{y}$ ,  $\hat{z}$ , and any vector  $\mathbf{A}$  can be expressed in terms of them, in the usual way:

$$\mathbf{A} = (A_r) \hat{r} + (A_\theta) \hat{\theta} + (A_\phi) \hat{\phi}$$

Here,  $A_r$ ,  $A_\theta$ ,  $A_\phi$  are the radial, polar, and azimuthal components of vector  $\mathbf{A}$ . In terms of the Cartesian unit vectors:

$$\begin{aligned}\hat{r} &= (\sin \theta \cos \phi) \hat{x} + (\sin \theta \sin \phi) \hat{y} + (\cos \theta) \hat{z}, \\ \hat{\theta} &= (\cos \theta \cos \phi) \hat{x} + (\cos \theta \sin \phi) \hat{y} + (-\sin \theta) \hat{z}, \\ \hat{\phi} &= (-\sin \phi) \hat{x} + (\cos \phi) \hat{y} + (0) \hat{z}.\end{aligned}$$

An infinitesimal displacement in the  $\hat{r}$  direction is simply  $dr$ , just as an infinitesimal element of length in the  $\hat{x}$  direction is  $dx$ :

$$dl_r = dr$$

<sup>21</sup>meaning mutually perpendicular.



On the other hand, an infinitesimal element of length in the  $\hat{\theta}$  direction is **NOT** just  $d\theta$  but rather,

$$dl_\theta = r d\theta$$

Similarly, an infinitesimal element of length in the  $\hat{\phi}$  direction is

$$dl_\phi = r \sin \theta d\phi$$

Thus the general infinitesimal displacement  $dl$  is:

$$dl = (dr) \hat{r} + (r d\theta) \hat{\theta} + (r \sin \theta) \hat{\phi}$$

This plays the role  $dl = (dx) \hat{x} + (dy) \hat{y} + (dz) \hat{z}$  plays in Cartesian coordinates. The infinitesimal volume element  $d\tau$ , in spherical coordinates, is the product of the three (3) infinitesimal displacements:

$$d\tau = dl_r dl_\theta dl_\phi = r^2 \sin \theta dr d\theta d\phi.$$

It is not possible to give a general expression for **surface** elements  $da$ , since these depend on the orientation of the surface. We simply have to analyse the geometry for any given case, which goes for Cartesian and curvilinear coordinates.

Integrating over the surface of a sphere, for instance, makes  $r$  constant, whereas  $\theta$  and  $\phi$  change:

$$da_1 = dl_\theta dl_\phi \hat{r} = r^2 \sin \theta d\theta d\phi \hat{r}$$

On the other hand, if the surface lies in the  $xy$  plane, making  $\theta$  is constant, while  $r$  and  $\phi$  vary:

$$da_2 = dl_r dl_\phi \hat{\theta} = r dr d\phi \hat{\theta}$$

Finally:  $r$  ranges from 0 to  $\infty$ ,  $\phi$  from 0 to  $2\pi$ , and  $\theta$  from 0 to  $\pi$ .

### Exercise 2.12 | The Volume of a Sphere

Find the volume of a sphere of radius  $R$ .

**SOLUTION** The derivation is as follows:

$$V = \int d\tau$$

$$\begin{aligned} & \int_{r=0}^R \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} r^2 \sin \theta dr d\theta d\phi, \\ &= \left( \int_0^R r^2 dr \right) \left( \int_0^{\pi} \sin \theta d\theta \right) \left( \int_0^{2\pi} d\phi \right) \\ &= \left( \frac{R^3}{3} \right) (2) (2\pi) = \frac{4}{3} \pi R^3 \blacksquare \end{aligned}$$

### 2.4.2 | Cylindrical Coordinates

The cylindrical coordinates  $(s, \phi, z)$  of a point  $P$  are defined in **Fig. 2.10b**. Observe that  $\phi$  has the same meaning as in spherical coordinates, and  $z$  is the same as Cartesian;  $s$  is the distance to  $P$  from the  $z$  axis, whereas the spherical coordinate  $r$  is the distance from the origin. The relation to Cartesian coordinates is:

$$x = s \cos \phi \quad y = s \sin \phi \quad z = z.$$

The unit vectors are:

$$\begin{aligned}\hat{s} &= \cos \phi \hat{x} + \sin \phi \hat{y} \\ \hat{\phi} &= -\sin \phi \hat{x} + \cos \phi \hat{y} \\ \hat{z} &= \hat{z}\end{aligned}$$

The infinitesimal displacements are

$$dl_s = ds \quad dl_\phi = s d\phi, \quad dl_z = dz$$

which makes:

$$dl = ds \hat{s} + s d\phi \hat{\phi} + dz \hat{z}.$$

and the volume element is

$$d\tau' = s ds d\phi dz$$

The range of  $s$  is  $(0, \infty)$ ,  $\phi$  is from 0 to  $2\pi$  and  $z$  is from  $-\infty$  to  $+\infty$

## 2.5 Dirac Delta Function

### 2.5.1 A Mathematical Anomaly

Consider the following vector function:

$$\mathbf{v} = \frac{1}{r^2} \hat{\mathbf{r}}$$

At every location,  $\mathbf{v}$  is directed **radially outward** which we can see in **Fig. 2.11**. Let's calculate its divergence:

$$\nabla \cdot \mathbf{v} = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{1}{r^2} \right) = \frac{1}{r^2} \frac{\partial}{\partial r} (1) = 0$$

This is interesting as this calculation gives us an unforeseen solution. Let's look at it closer. Suppose we integrate over a sphere of radius  $R$ , centred at the origin. The surface integral therefore is

$$\oint \mathbf{v} \cdot d\mathbf{a} = \int \left( \frac{1}{R^2} \hat{\mathbf{r}} \right) \cdot \left( R^2 \sin \theta d\theta d\phi \hat{\mathbf{r}} \right) = \left( \int_0^\pi \sin \theta d\theta \right) \left( \int_0^{2\pi} d\phi \right) = 4\pi$$

But the volume integral ( $\int \nabla \cdot \mathbf{v} d\tau$ ), should be zero (0) if we assume the aforementioned calculation to be true.

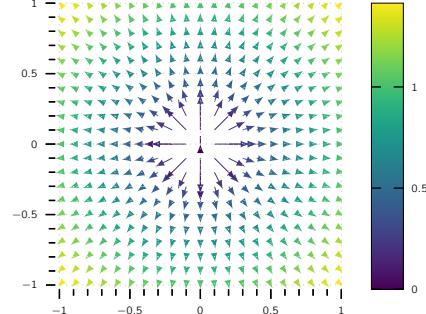
Does this mean that the divergence theorem is false?

The source of the problem lies at the point  $r = 0$ , where  **$\mathbf{v}$  blows up**. It is true that  $\nabla \cdot \mathbf{v} = 0$  everywhere **except** the origin, but right at the origin is the situation gets a little bit complicated.

Observe, the surface integral is **independent** of  $R$ . If the divergence theorem is to be true, we should expect

$$\int \nabla \cdot \mathbf{v} d\tau = 4\pi,$$

for any non-zero vector and the origin.



**Figure 2.11:** The vector plot of the “divergence problem”.

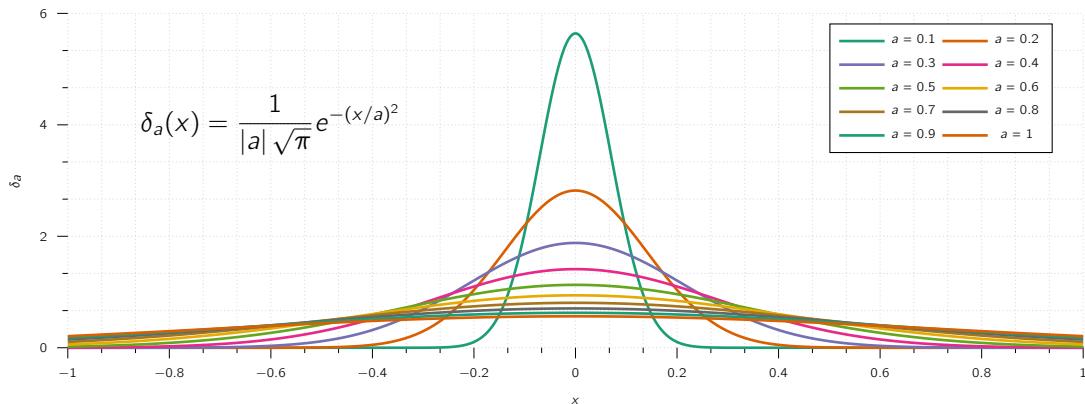
This means the value of  $4\pi$  must be coming from the point  $r = 0$ . Therefore,  $\nabla \cdot \mathbf{v}$  has the unique property that it vanishes everywhere except at one point, and yet its **integral** is  $4\pi$ .

No normal function behaves like that.

#### Information: An Analogy with Density

To wrap our heads around this property think of **density**.

The density of a point particle is zero except at the exact location of the particle, and yet its **integral** is finite. Namely, the mass of the particle.



**Figure 2.12:** A visual representation of a 1D Dirac Delta Function. Think of it as a distribution function being squeezed to an infinitely small width.

What we have stumbled upon is called the **Dirac delta function**. It arises in numerous branches of theoretical physics and plays a central role in the theory of electrodynamics.

### 2.5.2 The 1D Dirac Delta Function

The one-dimensional Dirac delta function ( $\delta(x)$ ), can be pictured as an infinitely high, infinitesimally narrow “spike”, with area 1.

That is to say:

$$\delta(x) = \begin{cases} 0, & \text{if } x \neq 0 \\ \infty, & \text{if } x = 0 \end{cases}$$

and,

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

In a strict sense of definition,  $\delta(x)$  is **NOT** a function at all, as its value is not finite at  $x = 0$ . In literature it is known as a generalised function.<sup>22</sup>

If  $f(x)$  is some “ordinary” function, then the product ( $f(x)\delta(x)$ ) is zero everywhere except at  $x = 0$ . It follows that:

$$f(x)\delta(x) = f(0)\delta(x). \quad (2.11)$$

The product is zero anyway except at  $x = 0$ . Based on this property, we may as well replace  $f(x)$  by the value it assumes at the origin.

<sup>22</sup>Objects extending the notion of functions on real or complex numbers. There is more than one recognised theory, for example the theory of distributions. Generalised functions are especially useful for treating discontinuous functions more like smooth functions, and describing discrete physical phenomena such as point charges.

Particularly

$$\int_{-\infty}^{\infty} f(x)\delta(x) dx = f(0) \int_{-\infty}^{\infty} \delta(x) dx = f(0). \quad (2.12)$$

Under an integral, then, the delta function “picks out” the value of  $f(x)$  at  $x = 0$ .<sup>23</sup> Of course, we can shift the spike from  $x = 0$  to some other point,  $x = a$ :

$$\delta(x - a) = \begin{cases} 0, & \text{if } x \neq a \\ \infty, & \text{if } x = a \end{cases} \quad \text{with} \quad \int_{-\infty}^{\infty} \delta(x - a) dx = 1. \quad (2.13)$$

<sup>23</sup>Here and below, the integral need not run from  $-\infty$  to  $+\infty$ ; it is sufficient that the domain extend across the delta function, and  $-\epsilon$  to  $+\epsilon$  would do as well.

which turns Eq. (2.11) into:

$$f(x)\delta(x - a) = f(a)\delta(x - a), \quad (2.14)$$

and Eq. (2.12) generalises to:

$$\int_{-\infty}^{\infty} f(x)\delta(x - a) dx = f(a). \quad (2.15)$$

While  $\delta$  itself is not a proper function, integrals over  $\delta$  are perfectly acceptable. In fact, it's best to think of the delta function as something that is always intended for use under an integral sign. In particular, two expressions involving delta functions, say  $D_1(x)$  and  $D_2(x)$ , are considered equal if:

$$\int_{-\infty}^{\infty} f(x)D_1(x) dx = \int_{-\infty}^{\infty} f(x)D_2(x) dx, \quad (2.16)$$

for all **ordinary** functions  $f(x)$ .

### Exercise 2.13 A Simple Dirac Integral

Evaluate the following integral:

$$\int_0^3 x^3 \delta(x - 2) dx$$

**SOLUTION** The delta function picks out the value of  $x^3$  at the point  $x = 2$ , so the integral is  $2^3 = 8$ . Notice, however, that if the upper limit had been 1 (instead of 3), the answer would be 0, because the spike would then be outside the domain of integration.

### Exercise 2.14 1D Dirac Delta Function

Evaluate the following integrals with Dirac delta functions:

$$\int_2^6 (3x^2 - 2x - 1) \delta(x - 3) dx, \quad (\text{xvii})$$

$$\int_0^5 \cos x \delta(x - \pi) dx, \quad (\text{xviii})$$

$$\int_0^3 x^3 \delta(x + 1) dx, \quad (\text{xix})$$

$$\int_{-\infty}^{+\infty} \ln(x + 3) \delta(x + 2) dx. \quad (\text{xx})$$

**SOLUTION** The solution are as follows:

$$(a) 3(3^2) - 2(3) - 1 = 27 - 6 - 1 = 20 \blacksquare$$

$$(b) \cos \pi = -1 \blacksquare$$

$$(c) 0 \blacksquare$$

$$(d) \ln(-2 + 3) = \ln 1 = 0 \blacksquare$$

### 2.5.3 The 3D Dirac Delta Function

Once we have defined the 1D Dirac, it is simple to generalise it to 3D with the following:

$$\delta^3(\mathbf{r}) = \delta(x)\delta(y)\delta(z), \quad (2.21)$$

As we can see, it is similar to 1D, where 3D Dirac is zero everywhere except at  $(0, 0, 0)$ , where it blows up. Its volume integral is 1:

$$\int_{\text{all space}} \delta^3(\mathbf{r}) d\tau = \iiint_{-\infty}^{\infty} \delta(x)\delta(y)\delta(z) dx dy dz = 1$$

And, the general form is:

$$\int_{\text{all space}} f(\mathbf{r})\delta^3(\mathbf{r} - \mathbf{a}) d\tau = f(\mathbf{a}). \quad (2.22)$$

As in the 1D case, integration with  $\delta$  picks out the value of the function  $f$  at the location of the spike.

We can fix the paradox introduced in Section 2.5. Remember, the divergence of  $\hat{\mathbf{r}}/r^2$  is zero everywhere except at the origin, however, its integral over any volume containing the origin is a constant.

These are precisely the defining conditions for the Dirac delta function; evidently

$$\nabla \cdot \left( \frac{\hat{\mathbf{r}}}{r^2} \right) = 4\pi\delta^3(\mathbf{r})$$

Or in a more general fashion:

$$\nabla \cdot \left( \frac{\hat{\mathbf{r}}'}{r'^2} \right) = 4\pi\delta^3(\mathbf{r}') \quad (2.23)$$

Differentiation here is with respect to  $\mathbf{r}$ , while  $\mathbf{r}'$  is held constant.

## 2.6 Vector Field Theory

### 2.6.1 Helmholtz Theorem

Electricity and magnetism are generally expressed as **electric and magnetic fields**,  $\mathbf{E}$  and  $\mathbf{B}$  and like many physical laws, these are most compactly expressed as **differential equations**.

As  $\mathbf{E}$  and  $\mathbf{B}$  are **vectors**, the differential equations naturally involve vector derivatives: *divergence* and *curl*. Maxwell reduced the entire theory to four (4) fundamental equations, specifying respectively the divergence and the curl of  $\mathbf{E}$  and  $\mathbf{B}$ .

This formulation raises an interesting question:

To what extent is a vector function determined by its divergence and curl?

To study this case lets assume a vector of  $\mathbf{F}$ . If the divergence of  $\mathbf{F}$  is a specified scalar function  $D$ ,

$$\nabla \cdot \mathbf{F} = D,$$

and the curl of  $\mathbf{F}$  is a specified vector function  $\mathbf{C}$ ,

$$\nabla \times \mathbf{F} = \mathbf{C},$$

and for consistency, we assume  $\mathbf{C}$  to have **NO** divergence,

$$\nabla \cdot \mathbf{C} = 0,$$

Remember, the divergence of a curl is **ALWAYS** zero.

Using this knowledge, is it possible to determine the function  $\mathbf{F}$ ?

Without knowing more information, it is not really possible. There are many functions whose divergence and curl are both zero everywhere.

Some examples are:

$$\mathbf{F} = 0,$$

$$\mathbf{F} = (y) \hat{\mathbf{x}} + (zx) \hat{\mathbf{y}} + (xy) \hat{\mathbf{z}},$$

$$\mathbf{F} = (\sin x \cosh y) \hat{\mathbf{x}} + (-\cos x \sinh y) \hat{\mathbf{y}} + (.) \hat{\mathbf{z}}$$

If we recall **Higher Mathematics I**, to solve a differential equation with a particular solution, we must also be supplied with appropriate **boundary conditions**.

In electrodynamics we typically require the fields go to zero at infinity. With that extra information, the **Helmholtz theorem**<sup>24</sup> guarantees the field is uniquely determined by its divergence and curl.



<sup>24</sup>Hermann Ludwig Ferdinand von Helmholtz (1821 - 1894)

was a German physicist and physician who made significant contributions in several scientific fields, particularly hydrodynamic stability. The Helmholtz Association, the largest German association of research institutions, was named in his honour.

In physics, he is known for his theories on the conservation of energy and on the electrical double layer, work in electrodynamics, chemical thermodynamics, and on a mechanical foundation of thermodynamics. Although credit is shared with Julius von Mayer, James Joule, and Daniel Bernoulli among others-for the energy conservation principles that eventually led to the first law of thermodynamics, he is credited with the first formulation of the energy conservation principle in its maximally general form.

## 2.6.2 Potentials

If the curl of a vector field ( $\mathbf{F}$ ) vanishes everywhere, then  $\mathbf{F}$  can be written as the **gradient of a scalar potential** ( $V$ ):

$$\nabla \times \mathbf{F} = 0 \iff \mathbf{F} = -\nabla V$$

The minus sign is purely conventional.

That's the essential idea of the following theorem:

### Theory 2.6: Zero Curl Fields

The following conditions are **equivalent**:

- i.  $\nabla \times \mathbf{F} = 0$  everywhere,
- ii.  $\int_a^b \mathbf{F} \cdot d\mathbf{l}$  is independent of path, for any given end points,
- iii.  $\oint \mathbf{F} \cdot d\mathbf{l} = 0$  for any closed loop,
- iv.  $\mathbf{F}$  is the gradient of some scalar function:  $\mathbf{F} = -\nabla V$ .

The potential is **NOT** unique as any constant can be added to  $V$ , since this will not affect its gradient.

If the divergence of a vector field ( $\mathbf{F}$ ) vanishes everywhere, then  $\mathbf{F}$  can be expressed as the curl of a **vector potential** ( $\mathbf{A}$ ):

$$\nabla \cdot \mathbf{F} = 0 \iff \mathbf{F} = \nabla \times \mathbf{A}$$

That's the main conclusion of the following theorem:

### Theory 2.7: Zero Divergence Fields

The following conditions are **equivalent**:

- i.  $\nabla \cdot \mathbf{F} = 0$  everywhere.
- ii.  $\oint f \mathbf{F} \cdot d\mathbf{a}$  is independent of surface, for any given boundary line.
- iii.  $\oint f \mathbf{F} \cdot d\mathbf{a} = 0$  for any closed surface.
- iv.  $\mathbf{F}$  is the curl of some vector function:  $\mathbf{F} = \nabla \times \mathbf{A}$ .

The vector potential is **NOT** unique as the gradient of any scalar function can be added to  $\mathbf{A}$  without affecting the curl, given the curl of a gradient is zero.

Incidentally, in all cases, a vector field  $\mathbf{F}$  can be written as the gradient of a scalar plus the curl of a vector.

$$\mathbf{F} = -\nabla V + \nabla \times \mathbf{A}$$





## Part II

# Electric Fields

### Part Contents

---

Chapter 3	Electrostatics	45
Chapter 4	Advanced Mathematical Methods	75

Wahrlich es ist nicht das Wissen, sondern das Lernen, nicht das Besitzen sondern das Erwerben, nicht das Da-Seyn, sondern das Hinkommen, was den grössten Genuss gewährt.

*It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment.*

---

*(Carl Friedrich Gauss in Letter to Farkas Bolyai (2 September 1808))*



# Chapter 3

## Electrostatics

### Table of Contents

3.1	The Electric Field . . . . .	45
3.2	Divergence and Curl of Electrostatic Fields . . . . .	49
3.3	Electric Potential . . . . .	55
3.4	Work and Energy . . . . .	63
3.5	Conductors . . . . .	68

### 3.1 The Electric Field

In essence, problem of electrodynamics is to solve:

We have some electric charges,  $q_1, q_2, q_3, \dots$ , let's call them **source charges**. What force do they exert on another charge (i.e., **test charge**),  $Q$ ?

The positions of the source charges are given<sup>1</sup> and the trajectory of the test particle is to be calculated. For most realistic cases, both the **source charges** and the **test charge** are in motion.

<sup>1</sup>usually as functions of time

The solution to this problem is solved with the **principle of superposition**, which states:

The interaction between any two (2) charges is **unaffected** by the presence of others.

This means to determine the force on  $Q$ , we can first calculate the force  $\mathbf{F}_1$ , caused by  $q_1$  alone, then we calculate the force  $\mathbf{F}_2$ , and then  $q_2$  alone and so on.

Finally, we take the vector sum of all these individual forces:

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3 + \dots$$

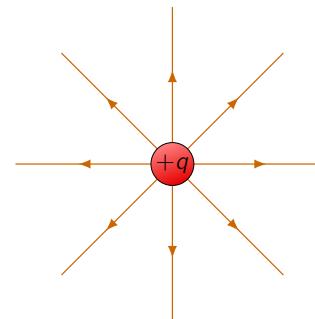
Therefore, if we can find the force on  $Q$  caused by a single source charge  $q$ , we are, in principle, done.<sup>2</sup>

It seems easy, right? Why don't we just write down the formula for the force on  $Q$  directed **towards**  $q$ , and calculate it?

We could, but some technicalities limit us, for not only does the force on  $Q$  depend on the separation distance  $\mathbf{r}$  between the charges, but also depends on *both* their velocities and on the *acceleration* of  $q$ .

Moreover, it is not the position, velocity, and acceleration of  $q$  **right now** that matter:

electromagnetic particles/waves travels at the speed of light, so what concerns  $Q$  is the position, velocity, and acceleration  $q$  *had* at some earlier time, when the message left.



**Figure 3.1:** An example of a singular charge in space. As can be seen a charge with a plus value will generate a repulsive force.

Therefore, in spite of the fact that the basic question is easy to state<sup>3</sup>, therefore we shall not answer this with a **direct** answer now but go at it by stages.

To simplify this problem, we assume time to be **irrelevant**, making the special case of **electrostatics** in which all the *source charges are stationary* whereas the test charge may be in motion.

<sup>2</sup>The principle of superposition may seem "obvious" to us, but it did **NOT** have to be simple. For example, if the electromagnetic force were proportional to the square of the total source charge, the principle of superposition would not hold, as  $(q_1 + q_2)^2 \neq q_1^2 + q_2^2$ , as there would be "cross terms" to consider. Superposition is **NOT** a logical necessity, but an experimental fact.

<sup>3</sup>Most complicated problems are funny enough relatively easy to state (see Collatz conjecture)

### 3.1.1 Coulomb's Law

What is the force on a test charge  $Q$  due to a single point charge  $q$ , that is at *rest* a distance  $\mathbf{r}$  away? The answer is given by Coulomb's law<sup>4</sup>:

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0} \frac{qQ}{\mathbf{r}^2} \mathbf{r}^2 \quad (3.1)$$

The constant  $\epsilon_0$  is the **permittivity of free space**. In SI units, where force is in newtons (N) distance in meters (m), and charge in coulombs (C).

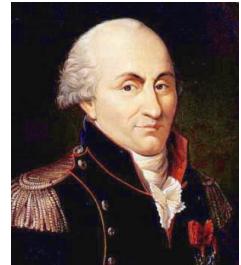
$$\epsilon_0 = 8.85 \times 10^{-12} \text{ C}^2 \cdot \text{N}^{-1} \cdot \text{m}^2$$

In other words,

the force is proportional to the product of the charges and inversely proportional to the square of the separation distance.

Remember,  $\mathbf{r}$  is the separation vector from  $\mathbf{r}_q$  (the location of  $q$ ) to  $\mathbf{r}_Q$  (the location of  $Q$ ):

$$\mathbf{r} = \mathbf{r}_Q - \mathbf{r}_q$$



<sup>4</sup>Charles-Augustin de Coulomb (1736 – 1806) was a French officer, engineer, physicist and the discoverer of what is now called Coulomb's law, the description of the electrostatic force of attraction and repulsion. He also did important work on friction, and his work on earth pressure formed the basis for the later development of much of the science of soil mechanics. The SI unit of electric charge, the **coulomb**, was named in his honor in 1880.

$\mathbf{F}$  is its magnitude, and  $\hat{\mathbf{r}}$  is its direction. The force points **along the line** from  $q$  to  $Q$ :

- it is repulsive if  $q$  and  $Q$  have the same sign,
- it is attractive if their signs are opposite

Coulomb's law and the principle of superposition form the physical foundation for electrostatics with the rest, except for some special properties of matter, is **mathematical derivation** of these fundamental rules.

### 3.1.2 The Electric Field

If we have *several* point charges  $q_1, q_2, \dots, q_n$ , at distances  $r_1, r_2, \dots, r_n$  from  $Q$ , the total force on  $Q$  is evidently

$$\begin{aligned} \mathbf{F} &= \mathbf{F}_1 + \mathbf{F}_2 + \dots = \frac{1}{4\pi\epsilon_0} \left( \frac{q_1 Q}{r_1^2} \hat{\mathbf{r}}_1 + \frac{q_2 Q}{r_2^2} \hat{\mathbf{r}}_2 + \dots \right) \\ &= Q \frac{1}{4\pi\epsilon_0} \left( \frac{q_1}{r_1^2} \hat{\mathbf{r}}_1 + \frac{q_2}{r_2^2} \hat{\mathbf{r}}_2 + \dots \right) \end{aligned}$$

or in a more cleaner way:

$$\mathbf{F} = Q \mathbf{E} \quad (3.2)$$

where

$$\mathbf{E}(\mathbf{r}) \equiv \frac{1}{4\pi\epsilon_0} \sum_{i=1}^n \frac{q_i}{r_i^2} \hat{\mathbf{r}}_i \quad (3.3)$$

$\mathbf{E}$  is called the **electric field** of the source charges. It is important to mention, it is a function of the position ( $\mathbf{r}$ ), as the separation vectors  $\hat{\mathbf{r}}_i$  depend on the location of the **field point**  $P$ .

But it makes no reference to the test charge  $Q$ .

The electric field is a **vector quantity** that varies from point to point and is determined by the configuration of source charges. Physically,  $\mathbf{E}(\mathbf{r})$  is the force per unit charge that would be exerted on a test charge, if you were to place one at  $P$ .

#### Exercise 3.1 Electric Field at a Distance

Find the  $\mathbf{E}$  field a distance  $z$  above the midpoint between two equal charges ( $q$ ), a distance  $d$  apart:

**SOLUTION** Let  $\mathbf{E}_1$  be the field of the left charge, and  $\mathbf{E}_2$  that of the right charge alone. Adding them, the horizontal components cancel and the vertical components add to:

$$E_z = 2 \frac{1}{4\pi\epsilon_0} \frac{q}{z^2} \cos\theta.$$

Here  $\mathbf{r} = \sqrt{z^2 + (d/2)^2}$  and  $\cos\theta = z/\mathbf{r}$ , therefore:

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \frac{2qz}{[z^2 + (d/2)^2]^{3/2}} \hat{\mathbf{z}} \quad \blacksquare$$

### 3.1.3 Continuous Charge Distributions

Our definition of the electric field Eq. (3.3) assumes the field source is a **collection** of discrete point charges  $q_i$ . If, the charge is distributed continuously over some region, the sum becomes an integral:

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{1}{r^2} \hat{\mathbf{r}} dq$$

1. If charge is spread out along a **line**, with charge-per-unit-length  $\lambda$ , then  $dq = \lambda dl'$ , where  $dl'$  is an element of length along the line.
2. If charge is smeared out over a **surface**, with charge-per-unit-area  $\sigma$ , then  $dq = \sigma da'$ , where  $da'$  is an element of area on the surface.
3. If charge fills a **volume** with charge-per-unit-volume  $\rho$ , then  $dq = \rho d\tau'$ , where  $d\tau'$  is an element of volume.

These aforementioned volumetric definitions can be rewritten as:

$$dq \rightarrow \lambda dl' \sim \sigma da' \sim \rho d\tau'. \quad (3.4)$$

Therefore the electric field of a line charge is,

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\lambda(r')}{r^2} \hat{\mathbf{r}} dl',$$

for a surface charge,

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\sigma(r')}{r^2} \hat{\mathbf{r}} da',$$

and for a volume charge,

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(r')}{r^2} \hat{\mathbf{r}} d\tau'. \quad (3.5)$$

Eq. (3.5) is often referred to as **Coulomb's law**, as it is such a short step from the original shown in Eq. (3.1), and because a volume charge is in a sense the most general and realistic case.



## 3.2 Divergence and Curl of Electrostatic Fields

As we have seen by now, the integrals involved in computing  $\mathbf{E}$  can be complicated and often laborious, even for reasonably simple charge distributions.

Therefore, rest of electrostatics from this point onwards is devoted to assembling methods for **avoiding these integrals**.

It all begins with the **divergence** and **curl** of  $\mathbf{E}$ . We shall calculate the divergence of  $\mathbf{E}$  directly from Eq. (3.5), but first we should develop an intuitive view of  $\mathbf{E}$ .

Let's begin with the simplest possible case:

A singular point charge ( $q$ ), situated at the **origin**:

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \hat{\mathbf{r}}$$

As the equation tells, the field falls off with a factor of  $1/r^2$ , the vectors get shorter as you go farther away from the origin and **they always point radially outward**. But there is a nicer way to represent this field, and that's to connect up the arrows, to form **field lines**. The magnitude of the field is indicated by the **density** of the field lines:

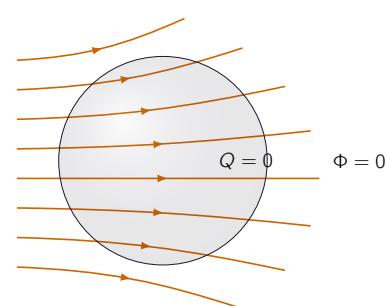
It's strong near the center where the field lines are close together, and weak farther out, where they are relatively far apart.

In truth, the field-line diagram can be deceptive, as when we draw it on a two-dimensional surface, for the density of lines passing through a circle of radius  $r$  is the total number divided by the circumference ( $n/2\pi r$ ), which goes at a rate of  $(1/r)$ , and not  $1/r^2$ .

But if were to imagine the model in three dimensions then the density of lines is the total number divided by the area of the sphere ( $n/4\pi r^2$ ), which **does** change with a factor of  $1/r^2$ .

Such diagrams are also convenient for representing more complicated fields. Field lines begin on positive charges and end on negative ones and they cannot simply terminate in midair, though they may extend out to infinity.

If they were to terminate, the divergence of  $\mathbf{E}$  would not be zero, and that cannot happen in empty space. In addition, field lines can never cross at the intersection,



**Figure 3.2**  
An example of imagining electric field as a collection of lines passing through. Here we can see the flux lines (i.e., the arrows) passing through an object.

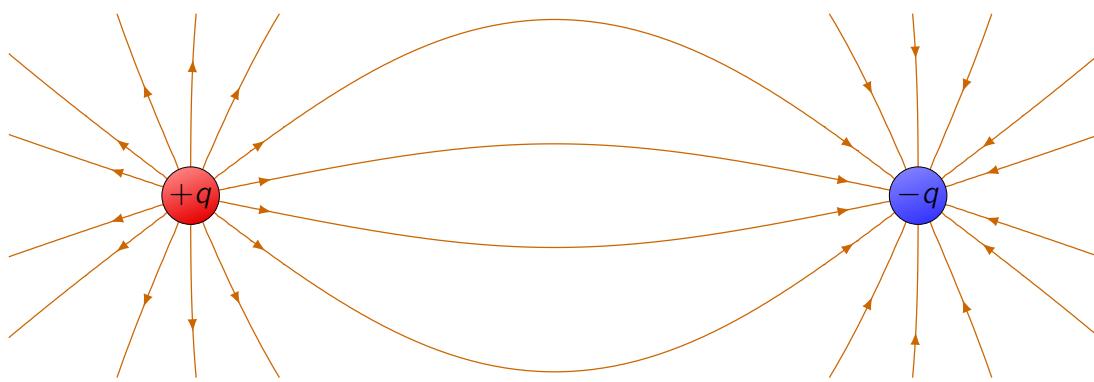


Figure 3.3

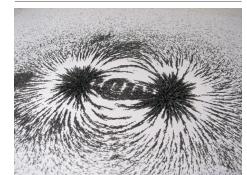
An example interaction of two charges interacting with one another with their field lines visible. Please observe that the field line spread away from the plus charge and are collected by the minus charge.

With all this in mind, it is easy to sketch the field of any simple configuration of point charges.

In this model, the **flux** of  $\mathbf{E}$  through a surface  $\mathcal{S}$ ,

$$\Phi_E = \int_{\mathcal{S}} \mathbf{E} \cdot d\mathbf{a}. \quad (3.6)$$

is a measure of **the number of field lines** passing through  $\mathcal{S}$ . It is worth mentioning here that it is only a representation of the field lines,<sup>5</sup> as the **total** number would be infinite. But **for a given sampling rate** the flux is *proportional* to the number of lines drawn, because the field strength, is proportional to the density of field lines, and therefore  $\mathbf{E} \cdot d\mathbf{a}$ , is proportional to the number of lines passing through the infinitesimal area  $d\mathbf{a}$ .



<sup>5</sup>In the history of physics, a line of force in Michael Faraday's extended sense is synonymous with James Clerk Maxwell's line of induction. According to J.J. Thomson, Faraday usually discusses lines of force as chains of polarized particles in a dielectric, yet sometimes Faraday discusses them as having an existence all their own as in stretching across a vacuum. From the 20th century perspective, lines of force are energy linkages embedded in a 19th-century field theory that led to more mathematically and experimentally sophisticated concepts and theories, including Maxwell's equations and Albert Einstein's theory of relativity.

This is the basic idea behind Gauss' law.

### 3.2.1 Quantification

In the case of a point charge  $q$  at the origin, the flux of  $\mathbf{E}$  through a spherical surface of radius ( $r$ ) is:

$$\oint \mathbf{E} \cdot d\mathbf{a} = \int \frac{1}{4\pi\epsilon_0} \left( \frac{q}{r^2} \hat{\mathbf{r}} \right) \cdot \left( r^2 \sin\theta d\theta d\phi \hat{\mathbf{r}} \right) = \frac{1}{\epsilon_0} q$$

Observe that the radius of the sphere **cancels out**, as the surface area goes **up** as  $r^2$ , the field goes **down** as  $1/r^2$ , so the product is **constant**. In terms of the field-line picture, this makes sense, given the same number of field lines pass through any sphere centered at the origin, regardless of its size.

In fact, it doesn't have to be a sphere. **Any** closed surface, whatever is shape, would be pierced by the same number of field lines. Evidently the flux through any surface enclosing the charge is  $q/\epsilon_0$ .

Let's extend this idea a bit further and assume instead of a single charge at the origin, we have a bunch of charges scattered about.

According to the principle of superposition, the total field is the vector sum of all the individual fields:

$$\mathbf{E} = \sum_{i=1}^n \mathbf{E}_i$$

The flux through a surface that encloses them all is:

$$\oint \mathbf{E} \cdot d\mathbf{a} = \sum_{i=1}^n \left( \oint \mathbf{E}_i \cdot d\mathbf{a} \right) = \sum_{i=1}^n \left( \frac{1}{\epsilon_0} q_i \right)$$

For any closed surface, then,

$$\oint \mathbf{E} \cdot d\mathbf{a} = \frac{1}{\epsilon_0} Q_{\text{enc}} \quad (3.7)$$

where  $Q_{\text{enc}}$  is the total charge enclosed within the surface. This is the quantitative statement of Gauss's law. As it stands, Gauss's law is an *integral* equation, but we can easily turn it into a *differential* one, by applying the divergence theorem:

$$\oint_S \mathbf{E} \cdot d\mathbf{a} = \int_V (\nabla \cdot \mathbf{E}) d\tau$$

Rewriting  $Q_{\text{enc}}$  in terms of the charge density  $\rho$ , we have

$$Q_{\text{enc}} = \int_V \rho d\tau$$

So Gauss's law becomes

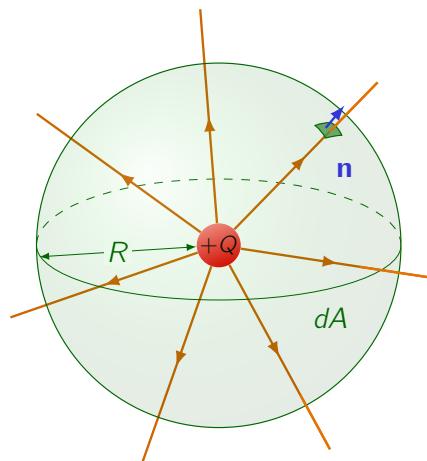


Figure 3.4: A visual representation of the gauss' law.

$$\int_V (\nabla \cdot \mathbf{E}) d\tau = \int_V \left( \frac{\rho}{\epsilon_0} \right) d\tau$$

And since this holds for *any* volume, the integrands must be equal:

$$\nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0} \rho \quad (3.8)$$

Eq. (3.8) carries the same message as Eq. (3.7) as it is **Gauss's law in differential form**. The differential version is idler, but the integral form has the advantage in that it accommodates point, line, and surface charges more naturally.

### 3.2.2 Divergence of $\mathbf{E}$

Let's go back, now, and calculate the divergence of  $\mathbf{E}$  directly from Eq. (3.5):

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int_{\text{all space}} \frac{\hat{\mathbf{z}}}{\mathbf{r}^2} \rho(\mathbf{r}') d\tau'$$

While the iteration is originally defined over the volume occupied by the charge, it is much easier to encompass all the space as  $\rho = 0$  outside the volume regardless.

Noting that the  $\mathbf{r}$ -dependence of the aforementioned equation is contained in the following:

$$\hat{\mathbf{z}} = \mathbf{r} - \mathbf{r}', \quad \text{where we have} \quad \nabla \cdot \mathbf{E} = \frac{1}{4\pi\epsilon_0} \int \nabla \cdot \left( \frac{\hat{\mathbf{z}}}{\mathbf{r}^2} \right) \rho(\mathbf{r}') d\tau'$$

This is precisely the divergence we calculated in the previous chapter regarding the divergence anomaly.

$$\left( \frac{\hat{\mathbf{z}}}{\mathbf{r}^2} \right) = 4\pi\delta^3(\hat{\mathbf{z}})$$

Using this relation we can arrive at the following:

$$\nabla \cdot \mathbf{E} = \frac{1}{4\pi\epsilon_0} \int 4\pi\delta^3(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') d\tau' = \frac{1}{\epsilon_0} \rho(\mathbf{r})$$

which is Gauss's law in differential form given in Eq. (3.8). To recover the integral form in Eq. (3.7), we run the previous argument **in reverse**, integrate over a volume and apply the divergence theorem:

$$\int_V \nabla \cdot \mathbf{E} d\tau = \oint_S \mathbf{E} \cdot d\mathbf{a} = \frac{1}{\epsilon_0} \int_V \rho d\tau = \frac{1}{\epsilon_0} Q_{\text{enc}}. \quad (3.9)$$

### 3.2.3 Applications of Gauss's Law

Let us look at some of the use cases of Gauss's law, in integral form. When the problem allows symmetry, it allows us the **quickest** and **easiest** way of computing electric fields. Let's understand this statement with a series of examples.

#### Exercise 3.2 A Gaussian Sphere

Find the field outside a uniformly charged solid sphere of radius  $R$  and total charge  $q$ .

**SOLUTION** Imagine a spherical surface at radius  $r > R$ . this is called a Gaussian surface. Gauss's law

says that:

$$\oint_S \mathbf{E} \cdot d\mathbf{a} = \frac{1}{\epsilon_0} Q_{\text{enc}},$$

For this case it is  $Q_{\text{enc}} = q$ . At first glance this doesn't seem to get us very far, as the quantity we want ( $\mathbf{E}$ )

is buried inside the surface integral. Luckily, symmetry allows us to extract  $\mathbf{E}$  from under the integral sign:  $\mathbf{E}$  certainly points radially outward,<sup>5</sup> as does  $d\mathbf{a}$ , so we can drop the dot product,

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

and the magnitude of  $\mathbf{E}$  is constant over the Gaussian

surface, so it comes outside the integral

$$\int_S |\mathbf{E}| d\mathbf{a} = |\mathbf{E}| \int_S d\mathbf{a} = |\mathbf{E}| 4\pi r^2$$

Therefore:

$$|\mathbf{E}| 4\pi r^2 = \frac{1}{\epsilon_0} q \quad \text{or} \quad \mathbf{E} = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \hat{\mathbf{r}} \blacksquare$$

Gauss's law is always **true**, however it may **NOT** always be useful. If  $\rho$  had not been uniform,<sup>6</sup> or if we had chosen some other shape for our Gaussian surface, it would still have been true that the flux of  $\mathbf{E}$  is  $q/\epsilon_0$ , but  $\mathbf{E}$  would not have pointed in the same direction as  $d\mathbf{a}$ , and its magnitude would not have been constant over the surface, and without that we cannot get  $|\mathbf{E}|$  outside of the integral.

Symmetry is crucial to this application of Gauss's law.

There are three (3) kinds of symmetry that work:

**Spherical** it is a concentric sphere,

**Cylindrical** it is a coaxial cylinder,

**Plane** A pillbox which straddles the surface.

While cylindrical and plane symmetries technically require infinitely long cylinders, and planes are assumed to extend to infinity, we shall often use them to get approximate answers for "long" cylinders or "large" planes, at points far from the edges.

Although the direct use of Gauss's law to compute electric fields is limited to cases of spherical, cylindrical, and planar symmetry, we can put together **combinations** of objects possessing such symmetry, even though the arrangement as a whole is not symmetrical. For example, invoking the principle of superposition, we could find the field in the vicinity of two uniformly charged parallel cylinders, or a sphere near an infinite charged plane.

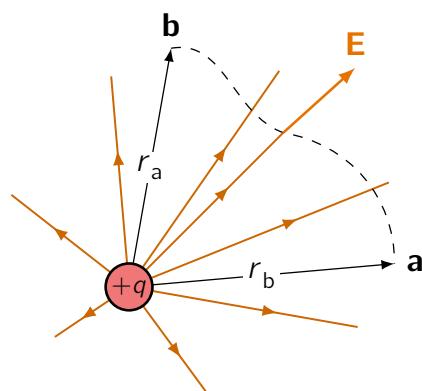


Figure 3.5

It is time to calculate the curl of  $\mathbf{E}$ , as we did the divergence. As usual, to derive this behaviour, we shall use the simplest case:

a point charge at the origin.

In this case we write down:

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \hat{\mathbf{r}}.$$

To get a feel of the curve of  $\mathbf{E}$ , let us calculate the line integral of this field from some point  $\mathbf{a}$  to some other point  $\mathbf{b}$ :

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{E} \cdot d\mathbf{l}$$

As we are working in spherical coordinates,  $(dr) \hat{\mathbf{r}} + (r d\theta) \hat{\mathbf{\theta}} + (r \sin \theta d\phi) \hat{\mathbf{\phi}}$ , so

$$\mathbf{E} \cdot d\mathbf{l} = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} dr$$

Therefore

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{E} \cdot d\mathbf{l} = \frac{1}{4\pi\epsilon_0} \int_{r_a}^{r_b} \frac{q}{r^2} dr = -\frac{1}{4\pi\epsilon_0} \frac{q}{r} \Big|_{r_a}^{r_b} = \frac{1}{4\pi\epsilon_0} \left( \frac{q}{r_a} - \frac{q}{r_b} \right)$$

where  $r_a$  is the distance from the origin to the point  $\mathbf{a}$  (i.e.,  $r_a = |\mathbf{a} - \mathcal{O}|$ ) and  $r_b$  is the distance to  $\mathbf{b}$  (i.e.,  $r_b = |\mathbf{b} - \mathcal{O}|$ ). The integral around a closed path is evidently zero (for then  $r_a = r_b$ ):

$$\oint \mathbf{E} \cdot d\mathbf{l} = 0 \quad (3.10)$$

And through the use of **Stokes' Theorem** one could arrive its differential form:

$$\nabla \times \mathbf{E} = 0 \quad (3.11)$$

Now, we proved Eq. (3.10) and Eq. (3.11) only for the field of a single point charge at the origin, but these results make no reference to what is, a perfectly arbitrary choice of coordinates.

They hold no matter where the charge is located. Moreover, if we have many charges, the principle of superposition states that the total field is a vector sum of their individual fields:

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \dots$$

which means

$$\nabla \times \mathbf{E} = \nabla \times (\mathbf{E}_1 + \mathbf{E}_2 + \dots)$$

Therefore, Eq. (3.10) and Eq. (3.11) hold for any static charge distribution.

## 3.3 Electric Potential

Now we look at another important concept in electric fields which is generally used as a mathematical concept but has significant use nonetheless.

### 3.3.1 Defining the Potential

The electric field  $\mathbf{E}$  is a **special kind** of vector function, one whose curl is zero.

For example,  $\mathbf{E} = y \hat{x}$ , could **NOT** possibly be an electrostatic field as **NO** set of charges, regardless of their sizes and positions, could ever produce such a field.

We're going to exploit this special property of electric fields to reduce a **vector** problem (finding  $\mathbf{E}$ ) to a **scalar** problem. We know from the previous chapter that any vector whose curl is zero is equal to the gradient of some scalar. What we are going to do now amounts to a proof in the context of electrostatics.

As  $\nabla \times \mathbf{E} = 0$ , the line integral of  $\mathbf{E}$  around any closed loop is zero.<sup>7</sup> As  $\oint \mathbf{E} \cdot d\mathbf{l} = 0$ , the line integral of  $\mathbf{E}$  from point  $\mathbf{a}$  to point  $\mathbf{b}$  is the same for all paths.

<sup>7</sup>This is a consequence of the Stokes' Theorem.

Because the line integral is independent of path, we can define a function

$$V(\mathbf{r}) \equiv - \int_{\mathcal{O}}^{\mathbf{r}} \mathbf{E} \cdot d\mathbf{l} \quad (3.12)$$

Here  $\mathcal{O}$  is some standard reference point on which we have agreed beforehand.  $V$  then depends only on the point  $\mathbf{r}$ . We will call this the **electric potential**.

The potential difference between two (2) points  $\mathbf{a}$  and  $\mathbf{b}$  is

$$V(\mathbf{b}) - V(\mathbf{a}) = - \int_{\mathcal{O}}^{\mathbf{b}} \mathbf{E} \cdot d\mathbf{l} + \int_{\mathcal{O}}^{\mathbf{a}} \mathbf{E} \cdot d\mathbf{l} \quad (3.13)$$

$$= - \int_{\mathcal{O}}^{\mathbf{b}} \mathbf{E} \cdot d\mathbf{l} - \int_{\mathbf{a}}^{\mathcal{O}} \mathbf{E} \cdot d\mathbf{l} = - \int_{\mathbf{a}}^{\mathbf{b}} \mathbf{E} \cdot d\mathbf{l}. \quad (3.14)$$

Now, the fundamental theorem for gradients states the following:

$$V(\mathbf{b}) - V(\mathbf{a}) = \int_{\mathbf{a}}^{\mathbf{b}} (\nabla V) \cdot d\mathbf{l}, \quad \text{and} \quad \int_{\mathbf{a}}^{\mathbf{b}} (\nabla V) \cdot d\mathbf{l} = - \int_{\mathbf{a}}^{\mathbf{b}} \mathbf{E} \cdot d\mathbf{l}.$$

Since, finally, this is true for any points  $\mathbf{a}$  and  $\mathbf{b}$ , the integrands must be equal:

$$\mathbf{E} = -\nabla V \quad (3.15)$$



Eq. (3.15) is the differential version of Eq. (3.12) which says that the electric field is the gradient of a scalar potential, which is what we set out to prove.

Notice the subtle but crucial role played by path independence<sup>8</sup> in this argument. If the line integral of  $\mathbf{E}$  depended on the path taken, then the “definition” of  $V$ , Eq. (3.12), would **NOT** be useful as it simply would **NOT** define a function, since changing the path would alter the value of  $V(\mathbf{r})$ .

<sup>8</sup>or, equivalently, the fact that  $\nabla \times \mathbf{E} = 0$

The minus sign in  $\mathbf{E} = -\nabla V$  is there due to historical conventions.

### Information: An Important Distinction

It is easy to be confused about the term **electrical potential** and a **voltage source** we encounter when we analyse a normal circuit. However there is a significant difference between them.

**Electrical potential** is the energy per unit charge gained or lost when a charge is moved from some reference point at which the potential is defined to be zero. **Voltage**, on the other hand, is the difference in potential between two (2) arbitrary points at which the potential is **NOT** necessarily zero.

For example, the potential at point **A** relative to a defined reference point<sup>9</sup> might be 100 V, and the potential at point **B** might be 10 V volts. Then the voltage between **B** and **A** is 90 V volts.

<sup>9</sup>sometimes an infinite distance away

## 3.3.2 Some Digressions on Potential

### Naming

The naming of  $V$  as **potential** is unfortunate as it inevitably reminds us of potential **energy**. This is particularly problematic, as there is a connection between **potential** and **potential energy**. Incidentally, a surface over which the potential is constant is called an **equipotential**.

### Advantages of the Potential

If we know  $V$ , we can easily get  $\mathbf{E}$ . It's simple, just take the gradient:  $\mathbf{E} = -\nabla V$ . This is due to the three (3) components of  $\mathbf{E}$  not being independent. In fact, they are **explicitly** interrelated by the very condition we started with,  $\nabla \times \mathbf{E} = 0$ .

In terms of components,

$$\frac{\partial E_x}{\partial y} = \frac{\partial E_y}{\partial x}, \quad \frac{\partial E_z}{\partial y} = \frac{\partial E_y}{\partial z}, \quad \frac{\partial E_x}{\partial z} = \frac{\partial E_z}{\partial x}.$$

### The Reference Point

There is a **necessary** ambiguity in the definition of potential, as the choice of reference point  $\mathcal{O}$  was arbitrary. Changing reference points amounts to adding a constant  $K$  to the potential:

$$V'(\mathbf{r}) = - \int_{\mathcal{O}}^{\mathbf{r}} \mathbf{E} \cdot d\mathbf{l} = - \int_{\mathcal{O}'}^{\mathcal{O}} \mathbf{E} \cdot d\mathbf{l} - \int_{\mathcal{O}'}^{\mathbf{r}} \mathbf{E} \cdot d\mathbf{l} = K + V(\mathbf{r})$$

where  $K$  is the line integral of  $\mathbf{E}$  from the old reference point  $\mathcal{O}$  to the new one  $\mathcal{O}'$ . Of course, adding a constant to  $V$  will **NOT** affect the potential **difference** between two (2) points:

$$V'(\mathbf{b}) - V'(\mathbf{a}) = V(\mathbf{b}) - V(\mathbf{a}),$$

since the  $K$ 's cancel out.<sup>10</sup> Nor does the ambiguity affect the gradient of  $V$ :

$$\nabla V' = \nabla V,$$

since the derivative of a constant is zero (0). That's why all such  $V$ 's, differing only in their choice of reference point, correspond to the same field  $\mathbf{E}$ .

<sup>10</sup>Actually, it was already clear from Eq. (3.13) that the potential difference is independent of  $\mathcal{O}$ , as it can be written as the line integral of  $\mathbf{E}$  from  $a$  to  $a$ , with no reference to  $\mathcal{O}$ .

Potential as such carries no real physical significance, for at any given point we can adjust its value at will by a suitable relocation of  $\mathcal{O}$ .

However, therefore  $V$ 's a "reference level to use for  $\mathcal{O}$  in electrostatics –almost-one-like equations to sea level for altitude—and that is a point infinitely far from the charge. Ordinarily, then, we "set the zero of potential at infinity." (Since  $V(\mathcal{O}) = 0$ , choosing a reference point is equivalent to selecting a place where  $V$  is to be zero.) But I must warn you that there is one special circumstance in which this convention fails: when the charge distribution itself extends to infinity. The symptom of trouble, in such cases, is that the potential blows up. For instance, the field of a uniformly charged plane is  $(\sigma/2\epsilon_0)$

$$V(z) = - \int_{\infty}^z \frac{1}{2\epsilon_0} \sigma dz = - \frac{1}{2\epsilon_0} \sigma(z - \infty).$$

The remedy is simply to choose some other reference point (in this example you might use a point on the plane). Notice that the difficulty occurs only in textbook problems; in "real life" there is no such thing as a charge distribution that goes on forever, and we can always use infinity as our reference point.

## Superposition Principle

The original superposition principle pertains to the force on a test charge ( $Q$ ). It says that the total force on  $Q$  is the **vector sum** of the forces attributable to the source charges individually:

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \dots$$

Dividing through by  $Q$ , we see that the electric field, too, obeys the superposition principle:

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \dots$$

Integrating from the common reference point to  $\mathbf{r}$ , it follows that the potential also satisfies such a principle:

$$V = V_1 + V_2 + \dots$$

That is, the potential at any given point is the sum of the potentials due to all the source charges separately. Only this time it is an ordinary sum, not a vector sum, which makes it a lot easier to work with.

### Units of Potential

For our system of units, SI, force is measured in newtons (N) and charge in coulombs (C), so electric fields are in newtons per coulomb ( $N \cdot C^{-1}$ ).

Accordingly, potential is known-meters per coulomb ( $m \cdot C^{-1}$ ), or joules per coulomb ( $J \cdot C^{-1}$ ).

A joule per coulomb is a volt.

### 3.3.3 Poisson's Equation and Laplace's Equation

We now know the electric field ( $\mathbf{E}$ ) can be written as the **gradient of a scalar potential**:

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

The question arises:

*How are divergence and curl of  $\mathbf{E}$  related to our new definition of electric potential,  $V$ ?*

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad \text{and} \quad \nabla \times \mathbf{E} = 0$$

Well, we can write the aforementioned equations in the following form:

$$\nabla \cdot \mathbf{E} = \nabla \cdot -\nabla V = -\nabla^2 V$$

so, apart from that persistent minus sign, the divergence of  $\mathbf{E}$  is the **Laplacian** of  $V$ . Gauss's law, then, says:

$$\nabla^2 V = -\frac{\rho}{\epsilon_0} \tag{3.16}$$

This is known as **Poisson's equation**. In regions where there is no charge, so  $\rho = 0$ , Poisson's equation reduces to **Laplace's equation**:

$$\nabla^2 V = 0$$

We'll explore this in Chapter 3 when we work with on **Advanced Mathematical Methods**.

What about the curl? Let's do the math and write the following form:

$$\nabla \times \mathbf{E} = \nabla \times (-\nabla V) = 0$$

But that's no condition on  $V$  as curl of gradient is always zero.

Of course, here, we used curl to show  $\mathbf{E}$  could be expressed as the gradient of a scalar, so it's not really surprising that this works out. It takes only one (1) Partial Differential Equation (PDE) (Poisson's) to determine  $V$ , as  $V$  is a scalar; for  $\mathbf{E}$  we needed two (2), the divergence and the curl.<sup>11</sup>

<sup>11</sup>This comes from Helmholtz theorem, which we discussed previously.

### 3.3.4 The Potential of a Localised Charge Distribution

Previously, we have defined  $V$  in terms of  $\mathbf{E}$  as shown in Eq. (3.12).

Ordinarily, though, it's  $E$  that we're looking for.<sup>12</sup> The idea is that it might be easier to get  $V$  first, and then calculate  $\mathbf{E}$  by taking the gradient. Typically, then, we know where the charge is,<sup>13</sup> and we want to find  $V$ .

Now, Poisson's equation relates  $V$  and  $\rho$ , but unfortunately it's "the wrong way around":

it would give us  $\rho$ , if we knew  $V$ , whereas we want  $V$ , knowing  $\rho$ .

What we need to do is invert Poisson's equation. That's the idea for this section, although we shall do it by roundabout means, beginning, as always, with a point charge at the origin.

The electric field is:

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \left( \frac{1}{r^2} \right) \hat{\mathbf{r}},$$

so

$$\mathbf{E} \cdot d\mathbf{l} = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} dr \quad \text{where} \quad d\mathbf{l} = dr \hat{\mathbf{r}} + r d\theta \hat{\theta} + r \sin \theta d\phi \hat{\phi}.$$

Setting the reference point at a place in infinity,<sup>14</sup> the potential of a point charge  $q$  at the origin is

$$V(r) = - \int_{\infty}^r \mathbf{E} \cdot d\mathbf{l} = \frac{1}{4\pi\epsilon_0} \int_{\infty}^r \frac{q}{r'^2} dr' = \frac{1}{4\pi\epsilon_0} \frac{q}{r'} \Big|_{\infty}^r = \frac{1}{4\pi\epsilon_0} \frac{q}{r}.$$

<sup>12</sup>if we already knew  $\mathbf{E}$ , there wouldn't be much point in calculating  $V$

<sup>13</sup>that is, we know where the charge density ( $\rho$ ) is.

Please observe the sign of  $V$ ; presumably the conventional minus sign in the definition Eq. (3.12) was chosen to make the potential of a positive charge come out positive.

#### Information: Think Topographically

It is useful to remember that regions of positive charge are potential "hills", regions of negative charge are potential "valleys", and the electric field points "downhill", from plus toward minus.

In general, the potential of a point charge  $q$  is

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{q}{\mathbf{r}}, \quad (3.17)$$

where  $\mathbf{r}$ , as always, is the distance from  $q$  to  $\mathbf{r}$ . Invoking the superposition principle, then, the potential of a collection of charges is

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^n \frac{q_i}{\mathbf{r}_i}, \quad (3.18)$$

or, for a continuous distribution,

$$\int dV[\mathbf{r}] = \frac{1}{4\pi\epsilon_0} \int \frac{1}{\mathbf{r}} dq. \quad (3.19)$$

In particular, for a volume charge, it's

$$\int dV[\mathbf{r}] = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')}{\mathbf{r}} d\tau'. \quad (3.20)$$

This is the equation we were looking for, telling us how to compute  $V$  when we know  $\rho$ . It is useful to think of this equation as the "solution" to Poisson's equation, for a localized charge distribution.

Let's compare Eq. (3.20) with the corresponding formula for the electric field in terms of  $\rho$  Eq. (3.5):

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')}{\mathbf{r}^2} \hat{\mathbf{r}} d\tau'$$

The main point to notice is that the unit vector ( $\hat{\mathbf{r}}$ ) is gone, so there is no need to fuss with components. The potentials of line and surface charges are

$$V = \frac{1}{4\pi\epsilon_0} \int \frac{\lambda(r')}{\mathbf{r}} dl' \quad \text{and} \quad V = \frac{1}{4\pi\epsilon_0} \int \frac{\sigma(r')}{\mathbf{r}} da' \quad \text{and} \quad (3.21)$$

It is worth stressing that everything in this section is predicated on the assumption that the reference point is at [infinity](#). This is hardly apparent in Eq. (3.20), but remember that we got that equation from the potential of a point charge at the origin,  $(1/4\pi\epsilon_0)(q/r)$ , which is valid only when  $\mathcal{O} = \infty$ .

If we try to apply these formulas to one of those artificial problems in which the charge itself extends to infinity, the integral will diverge.

### 3.3.5 Boundary Conditions

In the typical electrostatic problem we are given a source charge distribution ( $\rho$ ), and we want to find the electric field ( $\mathbf{E}$ ) it produces.

Unless the symmetry of the problem allows a solution by Gauss's law, it is generally to our advantage to calculate the potential first, as an intermediate step.

These are the three fundamental quantities of electrostatics:

$\rho$ ,  $E$ , and  $V$ .

We have, in the course of our discussion, derived all six (6) formulas interrelating them. These equations are neatly summarized in **Fig. 3.6**. We began with just two (2) experimental observations:

1. The principle of superposition-a broad general rule applying to **all** electromagnetic forces, and
2. Coulomb's law-the fundamental law of electrostatics.

From these, all else are derived. You may have noticed that the electric field always undergoes a discontinuity when you cross a surface charge  $\sigma$ . In fact, it is a simple matter to find the **amount** by which  $E$  changes at such a boundary. Suppose we draw a wafer-thin Gaussian pillbox, extending just barely over the edge in each direction as seen in **Fig. 3.7**. Gauss's law says that

$$\oint \mathbf{E} \cdot d\mathbf{a} = \frac{1}{\epsilon_0} Q_{\text{enc}} = \frac{1}{\epsilon_0} \sigma A,$$

where  $A$  is the area of the pillbox lid ( $\text{m}^2$ ).<sup>15</sup> Now, the sides of the pillbox contribute **nothing** to the flux, in the limit as the thickness  $\epsilon$  goes to zero, so we are left with:

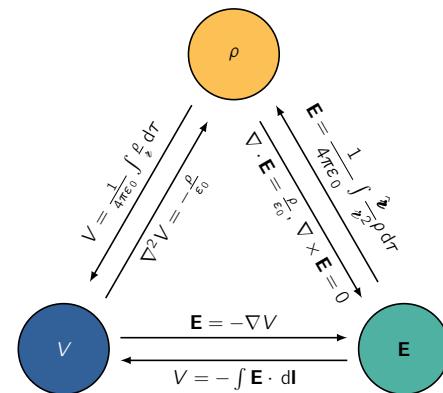
$$E_{\text{above}}^\perp - E_{\text{below}}^\perp = \frac{1}{\epsilon_0} \sigma, \quad (3.22)$$

where  $E_{\text{above}}^\perp$  denotes the component of  $E$  **perpendicular** to the surface immediately above, and  $E_{\text{below}}^\perp$  is the same, only just below the surface. For consistency, we let "upward" be the positive direction for both.

Based on this model we can say normal component of  $E$  (for clarity it is,  $E \cdot \hat{n}$ ) is discontinuous by an amount  $\sigma/\epsilon_0$  at any boundary. In particular, where there is **no** surface charge,  $E^\perp$  is continuous, as for instance at the surface of a uniformly charged solid sphere. The tangential component of  $vE$ , by contrast, is always continuous.

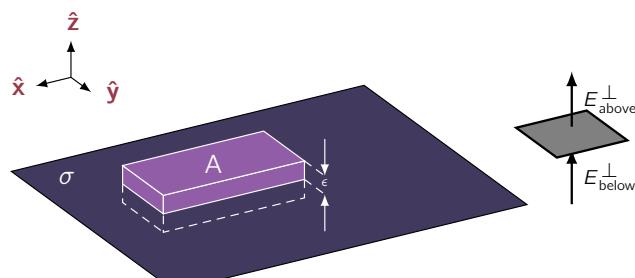
For if we apply:

$$\oint \mathbf{E} \cdot d\mathbf{l} = 0$$



**Figure 3.6**  
The relation between three (3) major variables used in the study of electrostatics.

<sup>15</sup>If  $\sigma$  varies from point to point or the surface is curved, we need to make sure  $A$  to be extremely small.



**Figure 3.7**  
A thought experiment on how to figure out the discrepancy when a boundary is changed.

to the thin rectangular loop of Fig. 2.37, the ends give nothing,<sup>16</sup> and the sides give  $(E_{\text{above}}^{\parallel} - E_{\text{below}}^{\parallel})$ ,<sup>16</sup> as  $\epsilon \rightarrow 0$ . so

$$E_{\text{above}}^{\parallel} = E_{\text{below}}^{\parallel} \quad (3.23)$$

where  $E^{\parallel}$  stands for the components of  $E$  parallel to the surface. The boundary conditions on  $E$  (Eqs. 2.31 and 2.32) can be combined into a single formula:

$$E_{\text{above}} - E_{\text{below}} = \frac{\sigma}{\epsilon_0} \hat{n}, \quad (3.24)$$

where  $\hat{n}$  is a unit vector perpendicular to the surface, pointing from "below" to "above."<sup>8</sup> The potential, meanwhile, is continuous across any boundary (Fig. 2.38), since  $V_{\text{above}} - V_{\text{below}} = - \int_0^b \mathbf{E} \cdot d\mathbf{l}$ ; as the path length shrinks to zero, so too does the integral:

$$V_{\text{above}} = V_{\text{below}}. \quad (3.25)$$

However, the **gradient** of  $V$  inherits the discontinuity in  $E$  given and following this  $E = -\nabla V$ , which Eq. (3.24) implies that:

$$\nabla V_{\text{above}} - \nabla V_{\text{below}} = -\frac{1}{\epsilon_0} \sigma \hat{n}, \quad (3.26)$$

or, more conveniently,

$$\frac{\partial V_{\text{above}}}{\partial n} - \frac{\partial V_{\text{below}}}{\partial n} = -\frac{1}{\epsilon_0} \sigma, \quad (3.27)$$

where

$$\frac{\partial V}{\partial n} = \nabla V \cdot \hat{n} \quad (3.28)$$

<sup>17</sup>That is, the rate of change in the direction perpendicular to the surface

denotes the **normal derivative** of  $V$ .<sup>17</sup> Please note that these boundary conditions relate the fields and potentials just above and **just** below the surface. For example, the derivatives in Eq. (3.27) are the **limiting** values as we approach the surface from either side.<sup>18</sup>

<sup>18</sup>Observe that it doesn't matter which side you call **above** and which **below**, given the reversal would switch the direction of  $\hat{n}$ .

Incidentally, if we're only interested in the field due to the essentially flat local patch of surface charge itself, the answer is  $(\sigma/2\epsilon_0) \hat{n}$  immediately above the surface, and  $-(\sigma/2\epsilon_0) \hat{n}$  immediately below. Evidently the entire discontinuity in  $E$  is attributable to this local patch of surface charge.

## 3.4 Work and Energy

### 3.4.1 To Move a Charge

To understand how to define work, Let us assume that we have a stationary configuration of source charges, and we want to move a test charge ( $Q$ ) from point  $a$  to point  $b$ . Here we need to ask the following question:

How much work will we have to do?

At any point along the path, the electric force on  $Q$  is  $\mathbf{F} = Q\mathbf{E}$ ; the force we must exert, in opposition to this electrical force, is  $-Q\mathbf{E}$ . The work we do is therefore:

$$W = \int_0^b \mathbf{F} \cdot d\mathbf{l} = -Q \int_0^b \mathbf{E} \cdot d\mathbf{l} = Q [V(b) - V(a)].$$

Please observe the answer is **independent of the path** we take from  $a$  to  $b$ ;

In mechanics, we call the electrostatic force “conservative”.<sup>19</sup>

Dividing through by  $Q$ , we have:

$$V(b) - V(a) = \frac{W}{Q}. \quad (3.29)$$

In other words, the potential difference between points  $a$  and  $b$  is equal to the work per unit charge required to carry a particle from  $a$  to  $b$ . In particular, if we want to bring  $Q$  in from far away and stick it at point  $r$ , the work we must do is

$$W = Q [V(r) - V(\infty)],$$

so, if we have set the reference point at infinity,

$$W = QV(r) \quad (3.30)$$

Here, **potential** is potential **energy**, the work it takes to create the system, per unit charge, just as the field is the force per unit charge.

<sup>19</sup>A force with the property which the total work done by the force in moving a particle between two (2) points is independent of the path taken. Equivalently, if a particle travels in a closed loop, the total work done by a conservative force is zero.

### 3.4.2 The Energy of a Point Charge Distribution

How much work would it take to assemble an entire **collection** of point charges? Imagine bringing in the charges, one by one, from far away (Fig. 2.40). The first charge,  $q_1$ , takes no work, since there is no field yet to fight against. Now bring in  $q_2$ . According to Eq. (3.30), this will cost you  $q_2V_1(\mathbf{r}_2)$ , where  $V_1$  is the potential due to  $q_1$ , and  $\mathbf{r}_2$  is the place we’re putting  $q_2$ :

$$W = \frac{1}{4\pi\epsilon_0} q_2 \left( \frac{q_1}{q_1^2} \right)$$

$(\lambda_{12})$  is the distance between  $q_1$  and  $q_2$  once they are in position). As you bring in each charge, nail it down in its final location, so it doesn't move when you bring in the next charge. Now bring in  $q_3$ ; this requires work  $q_3 V_{1,2}(\mathbf{r}_3)$ , where  $V_{1,2}$  is the potential due to charges  $q_1$  and  $q_2$ , namely,  $(1/4\pi\epsilon_0)(q_1/\nu_{13} + q_2/\nu_{23})$ . Thus

$$W_3 = \frac{1}{4\pi\epsilon_0} q_3 \left( \frac{q_1}{q_{13}} + \frac{q_2}{q_{23}} \right).$$

Similarly, the extra work to bring in  $q_4$  will be

$$W_4 = \frac{1}{4\pi\epsilon_0} q_4 \left( \frac{q_1}{\lambda_{14}} + \frac{q_2}{\lambda_{24}} + \frac{q_3}{\lambda_{34}} \right).$$

The total work necessary to assemble the first four charges, then, is

$$W = \frac{1}{4\pi\epsilon_0} \left( \frac{q_1 q_2}{\lambda_{12}} + \frac{q_1 q_3}{\lambda_{13}} + \frac{q_1 q_4}{\lambda_{14}} + \frac{q_2 q_3}{\lambda_{23}} + \frac{q_2 q_4}{\lambda_{24}} + \frac{q_3 q_4}{\lambda_{34}} \right).$$

You see the general rule: Take the product of each pair of charges, divide by their separation distance, and add it all up:

$$W = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^n \sum_{j>i}^n \frac{q_i q_j}{n_{ij}}. \quad (3.31)$$

The stipulation  $j > i$  is to remind you not to count the same pair twice. A nicer way to accomplish this is intentionally to count each pair twice, and then divide by 2:

$$W = \frac{1}{8\pi\epsilon_0} \sum_{i=1}^n \sum_{j=i}^n \frac{q_i q_j}{n_{ij}} \quad (3.32)$$

(we must still avoid  $i = j$ , of course). Notice that in this form the answer plainly does not depend on the order in which you assemble the charges, since every pair occurs in the sum. Finally, let's pull out the factor  $q_i$ :

$$W = \frac{1}{2} \sum_{i=1}^n q_i \left( \sum_{j \neq i}^n \frac{1}{4\pi\epsilon_0} \frac{q_j}{\nu_{ij}} \right).$$

The term in parentheses is the potential at point  $\mathbf{r}_i$  (the position of  $q_i$ ) due to all the other charges—all of them, now, not just the ones that were present at some stage during the assembly. Thus,

$$W = \frac{1}{2} \sum_{i=1}^n q_i V(\mathbf{r}_i). \quad (3.33)$$

That's how much work it takes to assemble a configuration of point charges; it's also the amount of work you'd get back if you dismantled the system. In the meantime, it represents energy stored in the configuration ("potential" energy, if you insist, though for obvious reasons I prefer to avoid that word in this context).

### 3.4.3 The Energy of a Continuous Charge Distribution

For a volume charge density ( $\rho$ ), Eq. (3.33) becomes:<sup>20</sup>

$$W = \frac{1}{2} \int \rho V d\tau \quad (3.34)$$

<sup>20</sup>The corresponding integrals for line and surface charges would be  $\int \lambda V dl$  and  $\int \sigma V da$ .

There is a tidy way to rewrite this result, in which  $\rho$  and  $V$  are eliminated in favor of  $\mathbf{E}$ . First use Gauss's law to express  $\rho$  in terms of  $\mathbf{E}$ :

$$\rho = \epsilon_0 \nabla \cdot \mathbf{E} \quad \text{so} \quad W = \frac{\epsilon_0}{2} \int (\nabla \cdot \mathbf{E}) V d\tau$$

Now use integration by parts, to transfer the derivative from  $\mathbf{E}$  to  $V$ :

$$W = \frac{\epsilon_0}{2} \left[ - \int \mathbf{E} \cdot (\nabla V) d\tau + \oint_S V \mathbf{E} \cdot da \right].$$

But we know  $\nabla V = -\mathbf{E}$ , so

$$W = \frac{\epsilon_0}{2} \left( \int_V E^2 d\tau + \oint_S V \mathbf{E} \cdot da \right) \quad (3.35)$$

But what volume is this we're integrating over? Let's go back to the formula we started with, Eq. (3.34). From its derivation, it is clear that we should integrate over the region where the charge is located. But actually, any larger volume would do just as well:

The “extra” territory we throw in will contribute nothing to the integral, since  $\rho = 0$  out there.

With this in mind, we return to Eq. (3.35). What happens here, as we enlarge the volume beyond the minimum necessary to trap all the charge? Well, the integral of  $E^2$  can only increase;<sup>21</sup> evidently the surface integral must decrease correspondingly to leave the sum intact.<sup>22</sup>

It is now worth stressing out that Eq. (3.35) gives us the correct energy  $W$ , whatever volume we use;<sup>23</sup> but the contribution from the volume integral goes up, and that of the surface integral goes down, as you take larger and larger volumes.

In particular, why not integrate over all space?

Then the surface integral goes to zero, and we are left with:

$$W = \frac{\epsilon_0}{2} \int E^2 d\tau \quad (3.36)$$

<sup>21</sup>the integrand being positive.

<sup>22</sup>In fact, at large distances from the charge,  $E$  goes  $1/r^2$  and  $V$   $1/r$ , while the surface area grows like  $r^2$ ; roughly speaking, then, the surface integral goes down like  $1/r$ .

<sup>23</sup>as long as it encloses all the charge

### 3.4.4 Comments on Electrostatic Energy

#### An Interesting Inconsistency

We can see Eq. (3.36) implies the energy of a stationary charge distribution is always **positive**. On the other hand, Eq. (3.33) which we derived Eq. (3.36), can be either **positive** or **negative**.

For instance, according to Eq. (3.33), the energy of two (2) equal but **opposite** charges a distance  $z$  apart is  $-(1/4\pi\epsilon_0)(q^2/z)$ .

What's gone wrong? Which equation is correct?

The answer is that **both** are correct, but they speak to slightly different questions. Eq. (3.33) does **NOT** take into account the work necessary to make the point charges in the first place; we **started** with point charges and simply found the work required to bring them together. This is wise strategy, given Eq. (3.36) indicates that the energy of a point charge is in fact **infinite**:

$$W = \frac{\epsilon_0}{2(4\pi\epsilon_0)^2} \int \left( \frac{q^2}{r^4} \right) (r^2 \sin \theta dr d\theta d\phi) = \frac{q^2}{8\pi\epsilon_0} \int_0^\infty \frac{1}{r^2} dr = \infty.$$

Eq. (3.36) is more **complete**, in the sense that it tells you the **total** energy stored in a charge configuration, but Eq. (3.33) is more appropriate when we're dealing with point charges, because we prefer to leave out that portion of the total energy that is attributable to the fabrication of the point charges themselves.

In practice, after all, the point charges are given to us ready-made; all we do is move them around. Since we did not put them together, and we cannot take them apart, it is immaterial how much work the process would involve.<sup>24</sup>

Now, it is a valid question to ask when the inconsistency crept into an apparently watertight derivation. The "flaw" lies between Eq. (3.33) and Eq. (3.34):

In the former,  $V(\mathbf{r}_i)$  represents the potential due to all the **other** charges **but not**  $q_i$ , whereas in the latter,  $V(\mathbf{r})$  is the **full** potential. For a **continuous** distribution, there is no distinction, given the amount of charge **right at the point**  $\mathbf{r}$  is vanishingly small, and its contribution to the potential is zero.

In the presence of point charges we should better stick with Eq. (3.33).

<sup>24</sup>Still, the infinite energy of a point charge is a recurring source of embarrassment for electromagnetic theory, afflicting the quantum version as well as the classical.

### The Place where Energy is Stored

Eq. (3.34) and Eq. (3.36) offer two (2) different ways of calculating the same thing.

1. The first is an integral over the charge distribution;
2. The second is an integral over the field.

These can involve completely different regions. For instance, in the case of the spherical shell, the charge is confined to the surface, whereas the electric field is everywhere **outside** this surface.

Where is the energy, then?

Is it stored in the field, as Eq. (3.36) seems to suggest, or is it stored in the charge, as Eq. (3.34) implies? At the present stage this is simply an unanswerable question: I can tell you what the total energy is, and I can provide you with several different ways to compute it, but it is impertinent to worry about where the energy is located. In the context of radiation theory it is useful to regard the energy as stored in the field, with a density:

$$\frac{\epsilon_0}{2} E^2 = \text{energy per unit volume.}$$

But in electrostatics one could just as well say it is stored in the charge, with a density  $\frac{1}{2}\rho V$ .

The difference is purely a matter of bookkeeping.

### Principle of Superposition

Because electrostatic energy is quadratic in the fields, it does **NOT** obey a superposition principle. The energy of a compound system is **NOT** the sum of the energies of its parts considered separately as there are also “cross terms” to take into account:

$$W_{\text{tot}} = \frac{\epsilon_0}{2} \int E^2 d\tau' = \frac{\epsilon_0}{2} \int (\mathbf{E}_1 + \mathbf{E}_2)^2 d\tau' \quad (3.37)$$

$$= \frac{\epsilon_0}{2} \int (E_1^2 + E_2^2 + 2\mathbf{E}_1 \cdot \mathbf{E}_2) d\tau \quad (3.38)$$

$$= W_1 + W_2 + \epsilon_0 \int \mathbf{E}_1 \cdot \mathbf{E}_2 d\tau. \quad (3.39)$$

For example, if you double the charge everywhere, you **quadruple** the total energy.

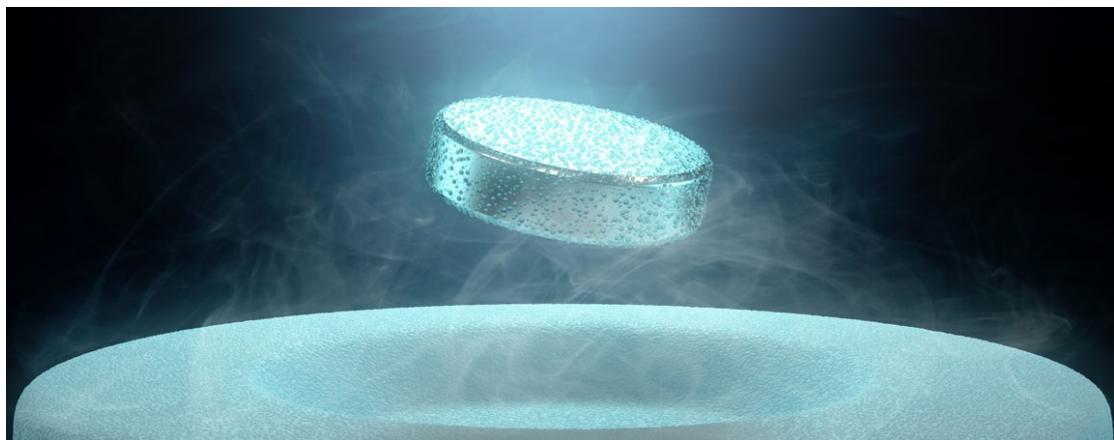


Figure 3.8

Some may say that superconductors can be considered as "perfect conductors", however, they exhibit quantum effects such as the Meissner effect and quantization of magnetic flux. In perfect conductors, the interior magnetic field must remain fixed but can have a zero or nonzero value [henyey1982distinction]. In real superconductors, all magnetic flux is expelled during the phase transition to superconductivity (the Meissner effect), and the magnetic field is always zero within the bulk of the superconductor [ScienceSup2023].

## 3.5 Conductors

### 3.5.1 Properties

In an insulator, for example glass or rubber, each electron is on a short leash, as it is attached to a particular atom. In a metallic **conductor**, by contrast, one or more electrons per atom are free to roam.<sup>25</sup> A **perfect** conductor would contain an **unlimited** supply of free charges.

In real life there are no such things perfect conductors, but metals come pretty close, for most purposes and therefore we can assume them to be perfect.

<sup>25</sup>In liquid conductors such as salt water, it is ions that do the moving. However, for now we shall only focus on solid materials.

From this aforementioned definition, the basic electrostatic properties of ideal conductors immediately follow:

#### i. No Electric Field Inside a Conductor

Let's have a think about this statement. If there **were** any field, those free charges would **move**, and it wouldn't be electrostatics any more. To give a more concrete explanation, we better examine what happens when we put a conductor into an external electric field  $E_0$ , seen in **Fig. 3.10**.

Initially, the field will drive any free positive charges to the right, and negative ones to the left.<sup>26</sup> When they come to the edge of the material, the charges pile up: plus on the right side, minus on the left. Now, these induced charges produce a field of their own ( $E_1$ ), which, as we can see from the figure, is in the **opposite direction** to  $E_0$ .

That's the crucial point, for it means the field of the induced charges tends to cancel the original field. Charge will continue to flow until this cancellation is complete, and the resultant field inside the conductor is precisely zero.<sup>27</sup> The whole process is practically instantaneous.

<sup>26</sup>In practice, it's the electrons that do the moving, but when they depart, the right side is left with a net positive charge, so it doesn't really matter which charges move; the effect is the same.

<sup>27</sup>Outside the conductor the field is not zero, for here  $E_0$  and  $E_1$  do NOT tend to cancel.



## ii. Charge Density is Zero inside a Conductor

This follows from Gauss's law:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

If  $\mathbf{E}$  is zero, then by extension, so is  $\rho$ . There is still charge around, but exactly as much plus as minus, so the net charge density in the interior is zero.

## iii. Any Net Charge is on the Surface

That's the only place left.

## iv. A Conductor is an Equipotential

For if  $a$  and  $b$  are any two (2) points within (or at the surface of) a given conductor,

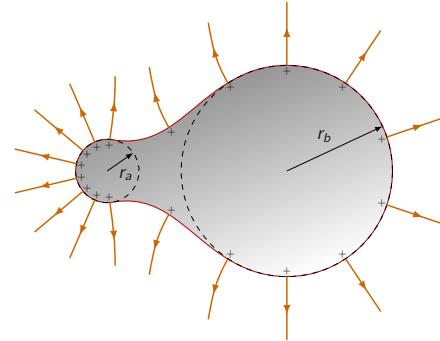
$$V(b) - V(a) = - \int_a^b \mathbf{E} \cdot d\mathbf{l} = 0, \quad \text{therefore} \quad V(b) = V(a).$$

## v. $\mathbf{E}$ is Perpendicular to the Surface

Otherwise, charge will immediately flow around the surface until it kills off the tangential component which are shown in Fig. 3.9.<sup>28</sup>

<sup>28</sup>Perpendicular to the surface, charge cannot flow, of course, since it is confined to the conducting object.

It is an astonishing fact the charge on a conductor flows to the surface. Given their mutual repulsion, the charges naturally spread out as much as possible, but for all of them to go to the surface seems like a waste of the interior space. Surely we could do better, from the point of view of making each charge as far as possible from its neighbors, to sprinkle some of them throughout the volume. Well, it simply is not so. You do best to put all the charge on the surface, and this is true regardless of the size or shape of the conductor.



**Figure 3.9**  
A visual representation of treating a non-standard body and the generated electric field. Please observe that all lines flowing from the body is perpendicular.

The problem can also be phrased in terms of energy.

Like any other free dynamical system, the charge on a conductor will seek the configuration that minimizes its potential energy. What property (iii) asserts is that the electrostatic energy of a solid object<sup>29</sup> is a minimum when charge is spread over the surface.

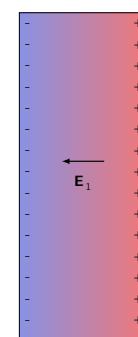
For instance, the energy of a sphere is

$$(1/8\pi\epsilon_0)(q^2/R)$$

if the charge is uniformly distributed over the surface, but it is,

$$(3/20\pi\epsilon_0)(q^2/R)$$

if the charge is uniformly distributed throughout the volume



**Figure 3.10**  
A visual representation of the electric field inside a conductor.

<sup>29</sup>with specified shape and total charge

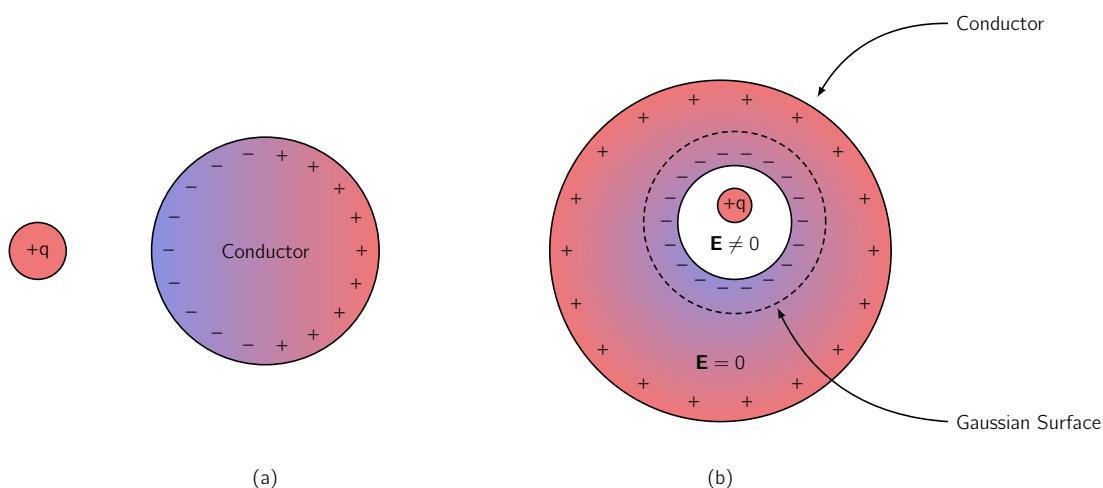


Figure 3.11

Charges induced by a point charge in a conductor (a) When a charge approaches a conductor the charges will start to move where opposite charges will move towards the point charge and same charges will move away. While there is an imbalance now, the overall charge of the conductor remains the same (b) When a charge is given inside the cavity, again similar behaviour is to be expected inside the conductor.

### 3.5.2 Induced Charges

If we were to hold a charge ( $+q$ ) near an **uncharged** conductor seen in **Fig. 3.11a**, the two (2) will attract one another. The reason for this is that  $+q$  will pull minus charges ( $-q$ ) over to the near side and repel plus charges to the far side.<sup>30</sup> Since the negative induced charge is closer to  $q$ , there is a net force of attraction.

When we speak of the field, charge, or potential “inside” a conductor, we should think of the “meat” of the conductor; if there is some hollow **cavity** in the conductor, and within that cavity we put some charge, then the field in the cavity will **NOT** be zero. But in a remarkable way the cavity and its contents are electrically isolated from the outside world by the surrounding conductor, shown in **Fig. 3.11b**.

<sup>30</sup>Another way to think of it is that the charge moves around in such a way as to kill off the field of  $q$  for points inside the conductor, where the total field must be zero.

No external fields penetrate the conductor; they are cancelled at the outer surface by the induced charge there. Similarly, the field due to charges within the cavity is cancelled, for all exterior points, by the induced charge on the inner surface. However, the compensating charge left over on the outer surface of the conductor effectively “communicates” the presence of  $q$  to the outside world.

The total charge induced on the cavity wall is equal and opposite to the charge inside, for if we surround the cavity with a Gaussian surface, all points of which are in the conductor (**Fig. 3.11b**),  $\oint \mathbf{E} \cdot d\mathbf{a} = 0$ , and therefore with Gauss’s law, the net enclosed charge must be zero (0). But  $Q_{\text{enc}} = q + q_{\text{induced}}$ , so  $q_{\text{induced}} = -q$ . Then if the conductor as a whole is electrically neutral, there must be a charge  $+q$  on its outer surface.

If a cavity surrounded by conducting material is itself empty of charge, then the field within the cavity is zero. For any field line would have to begin and end on the cavity wall, going from a plus charge to a minus charge, shown in **Fig. 3.12**.

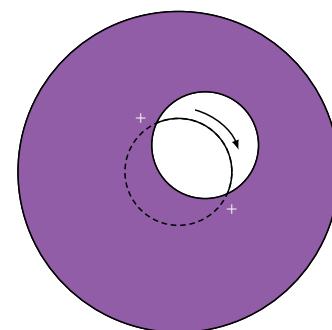


Figure 3.12

Visual description of an electric field inside a cavity. As can be seen all the lines generated by the positive side of the edge is cancelled by the negative side with it absorbs making the inside of the cavity have 0 field.

Letting that field line be part of a closed loop, the rest of which is entirely inside the conductor (where  $E = 0$ ), the integral  $\oint \mathbf{E} \cdot d\mathbf{l}$  is distinctly **positive**, in violation of  $\oint \mathbf{E} \cdot d\mathbf{l} = 0$ . It follows that  $E = 0$  within an **empty** cavity, and there is in fact **no** charge on the surface of the cavity.<sup>31</sup>



### 3.5.3 Surface Charge and the Force on a Conductor

Because the field inside a conductor is zero (0), boundary condition 2.33 requires the field immediately **outside** is:

$$E = \frac{\sigma}{\epsilon_0} \hat{n}, \quad (3.40)$$

consistent with our earlier conclusion that the field is **normal to the surface**. In terms of potential, Eq. (3.27) gives us:

$$\sigma = -\epsilon_0 \frac{\partial V}{\partial n} \quad (3.41)$$

These equations enable us to calculate the surface charge on a conductor, if we can determine  $E$  or  $V$ .<sup>32</sup> In the presence of an electric field, a surface charge will experience a force; the force per unit area ( $f$ ), is  $\sigma E$ .

But there's a problem here:

The electric field is **discontinuous** at a surface charge.

So what are we supposed to use:  $E_{\text{above}}$ ,  $E_{\text{below}}$ , or something in between? The answer is that we should use the average of the two:

$$f = \sigma E_{\text{average}} = \frac{1}{2}\sigma (E_{\text{above}} + E_{\text{below}}). \quad (3.42)$$

You may ask:, why the average?

The reason is very simple, though the telling makes it sound complicated: Let's focus our attention on a tiny patch of surface surrounding the point in question (Fig. 2.50).<sup>33</sup> The total field consists of two parts—that attributable to the patch itself, and that due to everything else (other regions of the surface, as well as any external sources that may be present):  $E = E_{\text{patch}} + E_{\text{other}}$ . Now, the patch cannot exert a force on itself, any more than you can lift yourself by standing in a basket and pulling up on the handles. The force on the patch, then, is due exclusively to  $E_{\text{other}}$ , and this suffers no discontinuity (if we removed the patch, the field in the "hole" would be perfectly smooth). The discontinuity is due entirely to the charge on the patch, which puts out a field ( $\sigma/2\epsilon_0$ ) on either side, pointing away from the surface. Thus,

$$\begin{aligned} E_{\text{above}} &= E_{\text{other}} + \frac{\sigma}{2\epsilon_0} \hat{n}, \\ E_{\text{below}} &= E_{\text{other}} - \frac{\sigma}{2\epsilon_0} \hat{n}. \end{aligned} \quad \text{therefore,} \quad E_{\text{other}} = \frac{1}{2}(E_{\text{above}} + E_{\text{below}}) = E_{\text{average}}.$$

Averaging is really just a device for removing the contribution of the patch itself. That argument applies to **any** surface charge; in the particular case of a conductor, the field is zero inside and

<sup>31</sup>This is why we are **relatively safe** inside a metal car during a thunderstorm—you may get cooked, if lightning strikes, but we will **NOT** be **electrocuted**. The same principle applies to the placement of sensitive apparatus inside a grounded Faraday cage, to shield out stray electric fields. In practice, the enclosure doesn't even have to be solid—conductor-chicken wire will often suffice [ck122025].

<sup>32</sup>we shall use them frequently in the next chapter.

<sup>33</sup>Make it small enough so it is essentially flat and the surface charge on it is essentially constant.

$(\sigma/\epsilon_0)$  m outside as shown in Eq. (3.40), so the average is  $(\sigma/2\epsilon_0) \hat{n}$ , and the force per unit area is

$$\mathbf{f} = \frac{1}{2\epsilon_0} \sigma^2 \hat{n}. \quad (3.43)$$

This amounts to an outward **electrostatic pressure** on the surface, tending to draw the conductor into the field, regardless of the sign of  $\sigma$ . Expressing the pressure in terms of the field just outside the surface,

$$P = \frac{\epsilon_0}{2} E^2. \quad (3.44)$$

### 3.5.4 Capacitors

Suppose we have two (2) conductors, and we put charge  $+Q$  on one and  $-Q$  on the other as shown in Fig. 3.13. Since  $V$  is **constant** over a conductor, we can speak unambiguously of the potential difference between them:

$$V = V_+ - V_- = - \int_{-}^{+} \mathbf{E} \cdot d\mathbf{l}.$$

We don't know how the charge distributes itself over the two conductors, and calculating the field would be a nightmare, if their shapes are complicated, but this much we do know:

$E$  is proportional to  $Q$ .

For  $E$  is given by Coulomb's law:

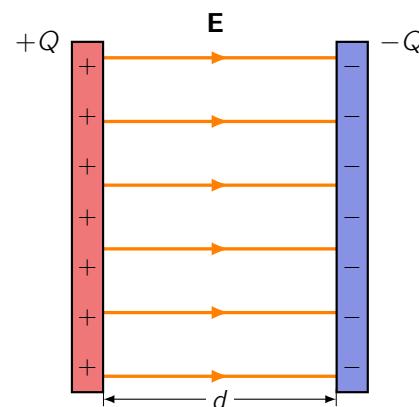
$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \int \frac{\rho}{\lambda^2} \hat{z} d\tau, \quad [\text{V} \cdot \text{m}^{-1}]$$

so if we double  $\rho$ , we double  $E$ .<sup>34</sup> Since  $E$  is proportional to  $Q$ , so also is  $V$ . The constant of proportionality is called the **capacitance** of the arrangement:

$$C \equiv \frac{Q}{V} \quad [\text{V} \cdot \text{C}^{-1} \text{ or F}] \quad (3.45)$$

Capacitance is a **purely geometrical quantity**, determined by the sizes, shapes, and separation of the two (2) conductors. In SI units,  $C$  is measured in farad (F); a farad is a coulomb-per-volt. This turns out to be inconveniently large; more practical units are the microfarad ( $1 \times 10^{-6}$  F) and the picofarad ( $1 \times 10^{-12}$  F).

Notice  $V$  is, by definition, the potential of the **positive** conductor less than that of the negative one; likewise,  $Q$  is the charge of the **positive** conductor. Accordingly, capacitance is an intrinsically positive quantity.<sup>35</sup>



**Figure 3.13**  
A parallel-plate capacitor, as shown in part A, consists of two flat conducting plates. Here one plate is positively charged and the other one is charged negatively. At a distance of  $d$  apart, an electric field ( $E$ ) is generated between the two (2) plates.

<sup>34</sup>Doubling  $Q$  doubles  $\rho$  everywhere, it doesn't shift the charge around.

<sup>35</sup>We may occasionally hear someone speak of the capacitance of a single conductor. In this case the "second conductor", with the negative charge, is an imaginary spherical shell of infinite radius surrounding the one conductor. It contributes **nothing** to the field, so the capacitance is given by Eq. (3.45), where  $V$  is the potential with infinity as the reference point.

### Charging a Capacitor

To “charge up” a capacitor, so to say, we have to remove electrons from the positive plate and carry them to the negative plate. By doing so, we fight against the electric field, which is pulling them back toward the positive conductor and pushing them away from the negative one.

*How much work does it take, then, to charge the capacitor up to a final amount  $Q$ ?*

Suppose that at some intermediate stage in the process the charge on the positive plate is  $q$ , so the potential difference is  $q/C$ . According to Eq. (3.29), the work we need do to transport the next piece of charge,  $dq$ , is:

$$dW = \left( \frac{q}{C} \right) dq. \quad [J]$$

The total work necessary, then, to go from  $q = 0$  to  $q = Q$ , is

$$W = \int_0^Q \left( \frac{q}{C} \right) dq = \frac{1}{2} \frac{Q^2}{C}, \quad [J]$$

or, since  $Q = CV$ ,

$$W = \frac{1}{2} CV^2 \quad [J] \quad (3.46)$$

where  $V$  is the final potential of the capacitor.



# Chapter 4

## Advanced Mathematical Methods

### Table of Contents

4.1 Solving Electrostatics . . . . .	75
4.2 Method of Images . . . . .	83
4.3 Separation of Variables . . . . .	87
4.4 Multipole Expansion . . . . .	101

### 4.1 Solving Electrostatics

The principle of electrostatics is to find the electric field ( $\mathbf{E}$ ) of a given stationary charge distribution. In terms of mathematical expression, this is accomplished by *Coulomb's law*, in the following form:

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\hat{\mathbf{z}}}{\mathbf{r}^2} \rho(\mathbf{r}') d\tau'.$$

Unfortunately, these types integrals can be difficult to calculate. Occasionally we can get around this by **exploiting symmetry** and using *Gauss's law*, but ordinarily the best strategy is first to calculate the electric potential ( $V$ ) which is given by the slightly less complicated expression:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{1}{\mathbf{r}} \rho(\mathbf{r}') d\tau'. \quad (4.1)$$

Even this integral is often too tough to handle analytically. Moreover, in problems involving conductors, the charge density ( $\rho$ ) itself may **NOT** be known in advance as charge is free to move around, the only thing we control directly is the total charge<sup>1</sup> of each conductor. For these cases, it is fruitful to recast the problem in differential form, using *Poisson's equation*:

$$\nabla^2 V = -\frac{1}{\epsilon_0} \rho,$$

<sup>1</sup>or perhaps the potential



which, together with appropriate **boundary conditions**, is equivalent to Eq. (4.1). Often we are interested in finding the potential in a region where  $\rho = 0$ <sup>2</sup>.

In this case, Poisson's equation reduces to **Laplace's equation**:

$$\nabla^2 V = 0, \quad \text{in Cartesian coordinates} \quad \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0.$$

<sup>2</sup>If  $\rho = 0$  everywhere, then  $V = 0$ , and there is nothing further to say. This NOT the intent. There can be plenty of charge elsewhere, but we're confining our attention to places where there is no charge.

The importance of this equation can't be downplayed as the field of electrostatics is the study of Laplace's equation. This equation very common in physics as it appears also in gravitation and magnetism, the theory of heat, and the study of bubbles.

To get an understanding of Laplace's equation and its solutions<sup>3</sup>, let's begin with the 1D and 2D versions, which are easier to picture, and illustrate all the essential properties of the 3D case.

### 4.1.1 Laplace's Equation in One Dimension

Let us assume  $V$  depends on only one (1) variable, say  $x$ . Then Laplace's equation becomes:

$$\frac{d^2 V}{dx^2} = 0 \quad \text{or} \quad V'' = 0.$$

The general solution becomes:

$$V(x) = mx + b, \quad (4.2)$$

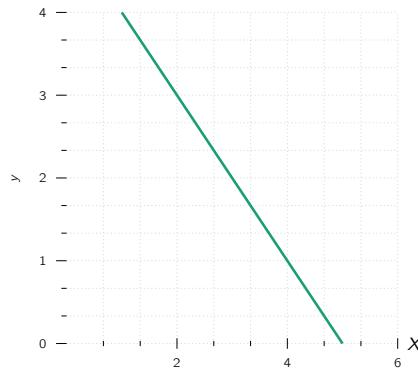


Figure 4.1: A particular solution to a 1D Laplace.

as we can see, the equation geometrically represents a **straight line**. It contains two (2) undetermined constants ( $m, b$ ), as is appropriate for a 2<sup>nd</sup>-order Ordinary Differential Equation (ODE). They are fixed, in any particular case, by the boundary conditions of that problem (or initial conditions if we recall **M.Sc Higher Mathematics I.**)

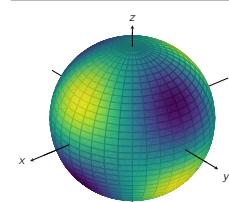
As an example, it might be specified that  $V = 4$  at  $x = 1$ , and  $V = 0$  at  $x = 5$ . In that case,  $m = -1$  and  $b = 5$ , so  $V = -x + 5$ , shown in **Fig. 4.1**.

Let's try to formulate the two (2) features of this result, while obvious in 1D, where we can write down the general solution **explicitly**, but the analogs in 2D and 3D dimensions are **powerful** and by no means obvious:

1.  $V(x)$  is the **average** of  $V(x+a)$  and  $V(x-a)$ , for any  $a$ :

$$V(x) = \frac{1}{2} [V(x+a) + V(x-a)]$$

Laplace's equation can be taught of as an **averaging instruction** as it tells to assign to the point  $x$  the average of the values to the left and to the right of  $x$ .



<sup>3</sup>known as **harmonic functions**

A harmonic function is a twice-continuously differentiable function that satisfies Laplace's equation, meaning the sum of its second partial derivatives is zero ( $u_{xx} + u_{yy} = 0$  for a two-variable function). This mathematical property is equivalent to the function's value at any point being the average of its values over any circle centered at that point. In physics and engineering, harmonic functions are known as potential functions and represent equilibrium conditions, such as steady-state temperature or electrostatic fields.



2. Laplace's equation **tolerates no local maxima or minima**.<sup>4</sup> Extreme values of  $V$  **must occur at the end points**. Actually, this is a consequence of the previous statement, for if there were a local maximum,  $V$  would be greater at that point than on either side, and therefore could not be the average.

<sup>4</sup>This is because the second derivative must be greater than or less than zero for maxima or minima but it is equal to zero as per laplace equation.

### Solving in Two Dimensions

When  $V$  depends on two (2) variables, Laplace's equation becomes:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0 \quad \text{or} \quad V_{xx} + V_{yy} = 0. \quad (4.3)$$

While the complexity does **NOT** rise in an exponential manner, nevertheless, we can no longer say we are working with an ODE.<sup>5</sup>. It is a PDE and therefore, some of the simple rules we may be familiar with will not apply.

<sup>5</sup>that is, one involving ordinary derivatives only

For instance, the general solution to the aforementioned equation doesn't contain just two (2) arbitrary constants.<sup>6</sup> Indeed, one cannot write down a **general solution**.<sup>7</sup>

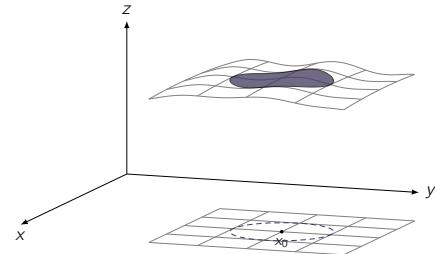
<sup>6</sup>or, for that matter, any finite number, despite being a second-order equation.

Nevertheless, it is possible to deduce certain properties **common to all solutions**.

<sup>7</sup>at least, NOT in a closed form we are familiar with.

It may help to have a physical example in mind. Picture a thin rubber sheet stretched over some support. For definiteness, suppose you take a cardboard box, cut a way line all the way around, and remove the top part, such as the one shown in **Fig. 4.2**.

Now glue a tightly stretched rubber membrane over the box, so it fits like a drum head.<sup>8</sup> Now, if you lay out coordinates  $(x, y)$  on the bottom of the box, the height  $V(x, y)$  of the sheet above the point  $(x, y)$  will satisfy Laplace's equation.



**Figure 4.2**  
A visual aid on how a Laplace equation in 2D would be tackled.

Harmonic functions in 2D have the same properties we noted in one dimension:

- The value of  $V$  at a point  $(x, y)$  is the **average** of those around the point. If we were to draw a circle of any radius  $R$  about the point  $(x, y)$ , the average value of  $V$  on the circle is **equal to the value at the centre**:

$$V(x, y) = \frac{1}{2\pi R} \oint_{\text{circle}} V \, dl.$$

This suggests the method of relaxation<sup>9</sup>, on which computer solutions to Laplace's equation are based. Starting with specified values for  $V$  at the boundary, and reasonable guesses for  $V$  on a grid of interior points, the first pass reassigns to each point the

<sup>9</sup>In numerical mathematics, relaxation methods are iterative methods for solving systems of equations, including nonlinear systems.



average of its nearest neighbors. The second pass repeats the process, using the corrected values, and so on. After a few iterations, the numbers begin to settle down, so that subsequent passes produce negligible changes, and a numerical solution to Laplace's equation, with the given boundary values, has been achieved.

2.  $V$  has **no local maxima or minima**; all extrema occur at the boundaries. Again, Laplace's equation picks the most featureless function possible, consistent with the boundary conditions: no hills, no valleys, just the smoothest conceivable surface. From a geometrical point of view, just as a straight line is the shortest distance between two points, so a harmonic function in two dimensions minimizes the surface area spanning the given boundary line.

### Solving in Three Dimensions

In 3D, unfortunately, we can neither derive an explicit solution nor describe the problem using a physical example as an analogy. However, even with all the additional complexities the third dimension brings, the same two properties remain true:

1. The value of  $V$  at some point  $r$  is the **average value** of  $V$  over a spherical surface of radius  $R$  centered at  $r$ :

$$V(r) = \frac{1}{4\pi R^2} \oint_{\text{sphere}} V \, da.$$

2. As a consequence of having more than one dimension,  $V$  can have no local maxima or minima. The extreme values of  $V$  must occur at the boundaries.<sup>10</sup>

<sup>10</sup>For if  $V$  had a local maximum at  $r$ , then by the very nature of maximum I could draw a sphere around  $r$  over which all values of  $V$ .

#### 4.1.2 Boundary Conditions and Uniqueness Theorems

Laplace's equation does **not by itself** determine  $V$ :

suitable boundary conditions must be supplied.

This raises an important question.

*What are considered appropriate conditions to determine the answer?*

For the case of 1D, it is easy, for here the general solution:

$$V = mx + b$$

contains two (2) arbitrary constants, and we require two (2) boundary conditions. We could:

- specify the value of the function at each end, [Dirichlet]
- give the value of the function and its derivative at one end, [Neumann]

■ give the value at one end and the derivative at the other,

[Robin]

and so on. But we cannot get away with just the value or just the derivative at one end. This is **insufficient** information or specifying the derivatives at both ends as this would either be redundant if the two are equal or inconsistent if they are **NOT**.

In 2D or 3D dimensions we are confronted by a PDE, and it is not so obvious **what would constitute acceptable boundary conditions**.

*Is the shape of a rubber membrane uniquely determined by the frame over which it is stretched, or, like a scanning jar lid, can it snap from one stable configuration to another?*

The answer, is that  $V$  is **uniquely determined by its value at the boundary**. However, other boundary conditions can also be used. The proof that a proposed set of boundary conditions will suffice is usually presented in the form of a **uniqueness theorem**. There are many such theorems for electrostatics, all sharing the same basic format. Let's look at the two (2) most useful ones.

#### Theory 4.8: First Uniqueness

The solution to Laplace's equation in some volume  $\mathcal{V}$  is **uniquely determined** if  $V$  is specified on the boundary surface  $\mathcal{S}$ .

##### Proof

In **Fig. 4.3** we can see such a region and its boundary.<sup>11</sup> Assume there were two (2) solutions to Laplace's equation:

$$\nabla^2 V_1 = 0, \quad \text{and} \quad \nabla^2 V_2 = 0,$$

both expression assume the specified value on the surface. Let's prove they must be equal. The main idea is look at their **difference**:

$$V_3 \equiv V_1 - V_2.$$

This obeys Laplace's equation,

$$\nabla^2 V_3 = \nabla^2 V_1 - \nabla^2 V_2 = 0,$$

and it takes the value **zero** on all boundaries.<sup>12</sup> But Laplace's equation allows no local maxima or minima as all extrema occur on the boundaries. Therefore, the maximum and minimum of  $V_3$  are both zero. This implies  $V_3$  must be zero everywhere, which means:

$$V_1 = V_2 \blacksquare$$

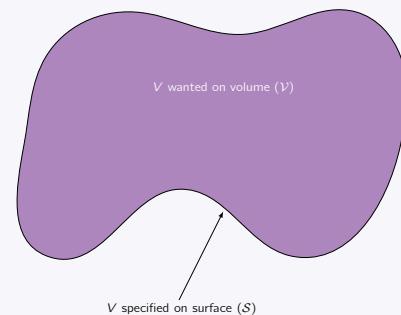


Figure 4.3

<sup>11</sup> There could also be "islands" inside, so long as  $V$  is given on all their surfaces. In addition, the outer boundary could be at infinity, where  $V$  is ordinarily taken to be zero.

<sup>12</sup> Given  $V_1$  and  $V_2$  are equal there.

The uniqueness theorem allows unique approaches as it doesn't matter how we come to a solution. If it satisfies Laplace's equation and it has the correct value on the boundaries, then

it's **right**.

We'll see the power of this argument when we come to the method of images. Incidentally, it is easy to improve on the first uniqueness theorem: I assumed there was no charge inside the region in question, so the potential obeyed Laplace's equation, but we may as well throw in some charge.<sup>13</sup> The argument is the same, only this time.

$$\nabla^2 V_1 = -\frac{1}{\epsilon_0} \rho, \quad \nabla^2 V_2 = -\frac{1}{\epsilon_0} \rho,$$

which makes:

$$\nabla^2 V_3 = \nabla^2 V_1 - \nabla^2 V_2 = -\frac{1}{\epsilon_0} \rho + \frac{1}{\epsilon_0} \rho = 0.$$

Once again the *difference* ( $V_3 \equiv V_1 - V_2$ ) satisfies Laplace's equation and has the value zero on all boundaries, so  $V_3 = 0$  and hence  $V_1 = V_2$ .

The potential in a volume  $\mathcal{V}$  is uniquely determined if (a) the charge density throughout the region, and (b) the value of  $V$  on all boundaries, are specified.

### 4.1.3 Conductors and the Second Uniqueness Theorem

The **easiest** way to set the boundary conditions for an electrostatic problem is to specify the value of  $V$  on all surfaces surrounding the region of interest. And this situation often occurs in practice:

In the laboratory, we have conductors connected to batteries, which maintain a given potential, or to **ground**, which is the experimental work for  $V = 0$ .

However, there are other circumstances in which we do not know the potential at the boundary, but rather the charges on various conducting surfaces. Suppose I put charge  $Q_a$  on the first conductor,  $Q_b$  on the second, and so on—I'm not telling you how the charge distributes itself over each conducting surface, because as soon as I put it on, it moves around in a way I do not control. And for good measure, let's say there is some specified charge density in the region between the conductors. Is the electric field now uniquely determined? Or are there perhaps a number of different ways the charges could arrange themselves on their respective conductors, each leading to a different field?

#### Theory 4.9: Second Uniqueness

In a volume  $\mathcal{V}$  surrounded by conductors and containing a specified charge density ( $\rho$ ), the electric field is uniquely determined if the *total charge* on each conductor is given,<sup>14</sup> which can be seen in **Fig. 4.4**.

<sup>14</sup>The region as a whole can be bounded by another conductor, or else unbounded.

#### Proof

Suppose there are two (2) fields satisfying the conditions of the given problem. They both obey



Gauss's law in differential form in the space between the conductors:

$$\nabla \cdot \mathbf{E}_1 = \frac{1}{\epsilon_0} \rho, \quad \nabla \cdot \mathbf{E}_2 = \frac{1}{\epsilon_0} \rho.$$

And they both obey Gauss's law in integral form for a Gaussian surface enclosing each conductor:

$$\oint_{i\text{th conducting surface}} \mathbf{E}_1 \cdot d\mathbf{a} = \frac{1}{\epsilon_0} Q_i, \quad \oint_{i\text{th conducting surface}} \mathbf{E}_2 \cdot d\mathbf{a} = \frac{1}{\epsilon_0} Q_i.$$

Similarly, for the outer boundary,<sup>15</sup>

$$\oint_{\text{outer boundary}} \mathbf{E}_1 \cdot d\mathbf{a} = \frac{1}{\epsilon_0} Q_{\text{tot}}, \quad \oint_{\text{outer boundary}} \mathbf{E}_2 \cdot d\mathbf{a} = \frac{1}{\epsilon_0} Q_{\text{tot}}.$$

As before, we examine the difference

$$\mathbf{E}_3 \equiv \mathbf{E}_1 - \mathbf{E}_2, \quad \text{which obeys} \quad \nabla \cdot \mathbf{E}_3 = 0, \quad (4.4)$$

in the region **between** the conductors, and,

$$\oint \mathbf{E}_3 \cdot d\mathbf{a} = 0 \quad (4.5)$$

over each boundary surface. Now there is one final piece of information we must use. While we don't know how the charge  $Q_i$  distributes itself over the  $i^{\text{th}}$  conductor, we **do** know that each conductor is an equipotential, and hence  $V_3$  is a **constant**<sup>16</sup> over each conducting surface.<sup>17</sup> Next comes a **trick**. Invoking product rule number 5, we find that:

$$\nabla \cdot (V_3 \mathbf{E}_3) = V_3 (\nabla \cdot \mathbf{E}_3) + \mathbf{E}_3 \cdot (\nabla V_3) = -(E_3)^2.$$

Here we have use Eq. (4.4), and  $\mathbf{E}_3 = -\nabla V_3$ . Integrating this over  $\mathcal{V}$ , and applying the **divergence theorem** to the Left Hand Side (LHS):

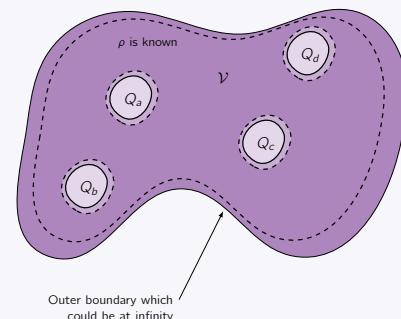
$$\int_{\mathcal{V}} \nabla \cdot (V_3 \mathbf{E}_3) d\tau = \oint_{\mathcal{S}} V_3 \mathbf{E}_3 \cdot d\mathbf{a} = - \int_{\mathcal{V}} (E_3)^2 d\tau.$$

The surface integral covers all boundaries of the region in question which is the conductors and outer boundary. Now  $V_3$  is a constant over each surface,<sup>18</sup> so it comes outside each integral, and what remains is zero, according to Eq. (4.5). Therefore,

$$\int_{\mathcal{V}} (E_3)^2 d\tau = 0.$$

But this integrand is **never negative**. The only way the integral can vanish is if  $E_3 = 0$  everywhere. Consequently,  $\mathbf{E}_1 = \mathbf{E}_2$ , and the theorem is proved ■

<sup>15</sup>whether this is just inside an enclosing conductor or at infinity.



<sup>16</sup>This does not necessarily needs to be the same constant.

Figure 4.4: A visual explanation of the second uniqueness theorem. Here the dashed lines represent the surface of integration and solid line describe the actual surface ( $\mathcal{S}$ ) of the charge.

<sup>17</sup>It need not be zero, for the potentials  $V_1$  and  $V_2$  may not be equal as all we know for sure is that **both** are **constant** over any given conductor.

<sup>18</sup>If the outer boundary is infinity,  $V_3 = 0$  there.

This proof was not easy, and there is a real danger that the theorem itself will seem more plausible to you than the proof. To give a better understanding of the second uniqueness theorem, consider

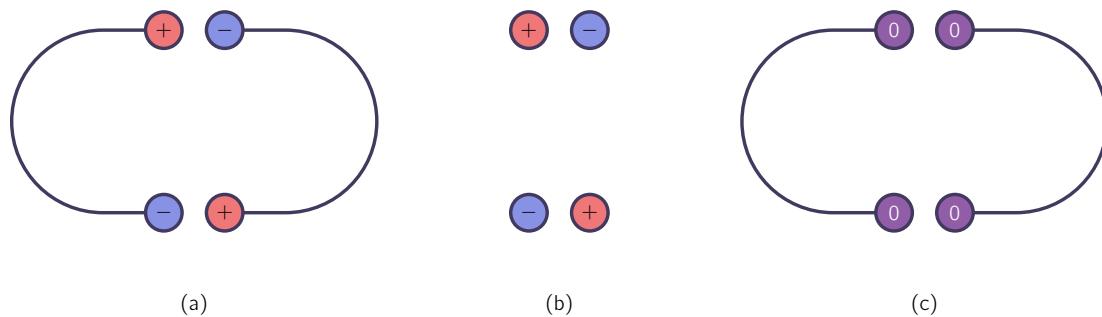


Figure 4.5

this example of Purcell's:

**Fig.** 4.5b shows a simple electrostatic configuration, consisting of four (4) conductors with charges  $Q$ , situated so that the pluses are near the minuses.

Up to now, it all looks very plausible. Now, what happens if we join them in pairs, by tiny wires, as indicated in **Fig.** 4.5a?

Since the positive charges are very near negative charges<sup>19</sup> we might well guess that **nothing** will happen. But it's wrong. The configuration in **Fig.** 4.5b is **impossible**.

<sup>19</sup>which is where we think they would be

For there are now effectively two (2) conductors, and the total charge on each is zero. **One** possible way to distribute zero charge over these conductors is to have no accumulation of charge anywhere, and hence zero field everywhere **Fig.** 4.5c.

By the second uniqueness theorem, this must be the solution. The charge will flow down the tiny wires, canceling itself off.

## 4.2 Method of Images

The method of images<sup>20</sup> is a mathematical tool for solving differential equations, in which boundary conditions are satisfied by combining a solution not restricted by the boundary conditions with its possibly weighted mirror image. This method has numerous applications in the field of electrical engineering which electrostatics is one of them.

<sup>20</sup>or method of mirror images.

### 4.2.1 A Classic Problem

To start let's discuss a simple problem. Let us assume a point charge ( $q$ ) is held a distance ( $d$ ) above an infinite grounded conducting plane (Fig. 3.10).

Now, we ask the Question:

*What is the potential in the region above the plane?*

Unfortunately we can't say it is  $(1/4\pi\epsilon_0) q/d$ , as  $q$  will induce a certain amount of negative charge on the nearby surface of the conductor; the total potential is due in part to  $q$  directly, and in part to this induced charge.

*But how can we possibly calculate the potential ( $V$ ), when we don't know how much charge is induced or how it is different?*

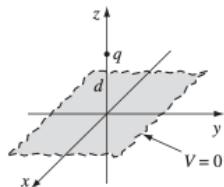


FIGURE 3.10

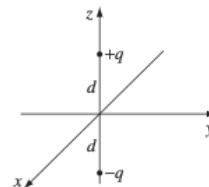


FIGURE 3.11

Figure 4.6

From a mathematical point of view, our problem is to solve Poisson's equation in the region  $z > 0$ , with a single point charge  $q$  at  $(0, 0, d)$ , subject to the boundary conditions:

1.  $V = 0$  when  $z = 0$ <sup>21</sup> and

$$[V(x, y, 0) = 0]$$

<sup>21</sup>since the conducting plane is grounded.

2.  $V \rightarrow 0$  far from the charge (that is, for  $x^2 + y^2 + z^2 \gg d^2$ ).

$$[\lim_{\infty} V = 0]$$

The first uniqueness theorem we looked previously guarantees that there is only one (1) function which meets these requirements.

The trick is to forget about the actual problem and study a *completely different* situation.

Let's rethink our problem and set a new configuration. This new configuration consists of two (2) point charges:

$+q$  at  $(0, 0, d)$ , and  $-q$  at  $(0, 0, -d)$ ,

and we will have no conducting plane in this new setup. For this configuration, we can easily write down the potential:

$$V(x, y, z) = \frac{1}{4\pi\epsilon_0} \left[ \frac{q}{\sqrt{x^2 + y^2 + (z-d)^2}} - \frac{q}{\sqrt{x^2 + y^2 + (z+d)^2}} \right]. \quad (4.6)$$

The denominators<sup>22</sup> are the distances from  $(x, y, z)$  to the charges  $+q$  and  $-q$ , respectively.

<sup>22</sup>Here, denominator means the value under the fraction sign.

It follows that:

$$1. V = 0 \text{ when } z = 0, \quad [V(x, y, 0) = 0]$$

$$2. V \rightarrow 0 \text{ for } x^2 + y^2 + z^2 \gg d^2, \quad [\lim_{\infty} V = 0]$$

which were defined previously, and the only charge in the region  $z > 0$  is the charge  $+q$  at  $(0, 0, d)$ .

These are precisely the conditions of the original problem.

Evidently the second configuration happens to produce **exactly the same potential** as the first configuration, in the “upper” region  $z \geq 0$ .<sup>23</sup>

Notice the crucial role played by the **uniqueness theorem** in this argument: without it, no one would believe this solution, since it was obtained for a completely different charge distribution. But the uniqueness theorem certifies it:

<sup>23</sup>The “lower” region,  $z > 0$ , is completely different, but as we don’t care about the lower region we can disregard that part.

If it satisfies Poisson’s equation in the region of interest, and assumes the correct value at the boundaries, then it must be right.

### 4.2.2 Induced Surface Charge

Now that we know the potential, it is a straightforward matter to compute the surface charge  $\sigma$  induced on the conductor. Looking back at the previous chapter:

$$\sigma = -\epsilon_0 \frac{\partial V}{\partial n}$$

where  $\partial V / \partial n$  is the normal derivative of  $V$  at the surface. In this case the normal direction is the  $z$  direction, so

$$\sigma = -\epsilon_0 \frac{\partial V}{\partial z} \Big|_{z=0}$$

And using our previously defined “mirror image” vector potential, in Eq. (4.6), we can write:

$$\frac{\partial V}{\partial z} = \frac{1}{4\pi\epsilon_0} \left[ \frac{-q(z-d)}{[x^2 + y^2 + (z-d)^2]^{3/2}} + \frac{q(z+d)}{[x^2 + y^2 + (z+d)^2]^{3/2}} \right],$$

so:

$$\sigma(x, y) = \frac{-qd}{2\pi(x^2 + y^2 + d^2)^{3/2}}.$$

As expected, the induced charge is negative<sup>24</sup> and greatest at  $x = y = 0$ . While we’re at it, let’s compute the **total** induced charge:

$$Q = \int \sigma da.$$

This integral, over the  $xy$  plane, could be done in Cartesian coordinates, with  $da = dx dy$ , but it’s a little easier to use polar coordinates  $(r, \phi)$ , with  $r^2 = x^2 + y^2$  and  $da = r dr d\phi$ . Then:

$$\sigma(r) = \frac{-qd}{2\pi(r^2 + d^2)^{3/2}},$$

and

$$Q = \int_0^{2\pi} \int_0^\infty \frac{-qd}{2\pi(r^2 + d^2)^{3/2}} r dr d\phi = \frac{qd}{\sqrt{r^2 + d^2}} \Big|_0^\infty = -q.$$

The total charge induced on the plane is  $-q$  ■

### 4.2.3 Force and Energy

The charge  $q$  is attracted toward the plane, because of the negative induced charge. Let’s calculate the force of attraction. Since the potential in the vicinity of  $q$  is the same as in the analogue problem,<sup>25</sup> so also is the field and, therefore, the force:

<sup>25</sup>the one with  $+q$  and  $-q$  but no conductor.

$$\mathbf{F} = -\frac{1}{4\pi\epsilon_0} \frac{q^2}{(2d)^2} \hat{\mathbf{z}}. \quad (4.7)$$

With the two (2) point charges and no conductor:

$$W = -\frac{1}{4\pi\epsilon_0} \frac{q^2}{2d}.$$

But for a single charge and conducting plane, the energy is *half* of this:

$$W = -\frac{1}{4\pi\epsilon_0} \frac{q^2}{4d}.$$

Why half? Think of the energy stored in the fields:

$$W = \frac{\epsilon_0}{2} \int E^2 d\tau.$$

In the first case, both the upper region ( $z > 0$ ) and the lower region ( $z < 0$ ) contribute equally. But in the second case, only the upper region contains a nonzero field, and hence the energy is half as great.

Of course, one could also determine the energy by calculating the work required to bring  $q$  in from infinity. The force required (to oppose the electrical force in Eq. (4.7)) is

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0} \left( q^2 / 4z^2 \right) \hat{\mathbf{z}}$$

so:

$$W = \int_{\infty}^d \mathbf{F} \cdot d\mathbf{l} = \frac{1}{4\pi\epsilon_0} \int_{\infty}^d \frac{q^2}{4z^2} dz = \frac{1}{4\pi\epsilon_0} \left( -\frac{q^2}{4z} \right) \Big|_{\infty}^d = -\frac{1}{4\pi\epsilon_0} \frac{q^2}{4d} \quad \blacksquare$$

As we move  $q$  towards the conductor, we do work **only on  $q$** . It is true that induced charge is moving in over the conductor, but this costs us nothing, since the whole conductor is at potential zero. By contrast, if I simultaneously bring in *two* point charges without any conductor present, we do work on **both** of them, and the total is twice as much.

## 4.3 Separation of Variables

In separation of variables<sup>26</sup> we will attack Laplace's equation **directly**. The method is applicable in situations where the potential ( $V$ ) or the charge density ( $\sigma$ ) is **specified** on the boundaries of some region, and we are tasked with finding the potential in the **interior**.

<sup>26</sup>This is the physicist's and engineer's favorite tool for solving PDEs.

We will have a look at this method in more detail in **M.Sc Higher Mathematics II** but for now let's use this opportunity to get a sneak-peek of this method.

The basic strategy is very simple:

We look for solutions that are products of functions, each of which depends on only one (1) of the coordinates.

The algebraic details, however, can be formidable, therefore let's try to understand this through exercises. We'll start with Cartesian coordinates and then do spherical coordinates.

### 4.3.1 Case Study: 2D Grounded Plates

To begin with let's assume the following problem

Two infinitely long grounded metal plates (■) at  $y = 0$  and  $y = a$  are connected at  $x = \pm b$  by metal strips maintained at a constant potential  $V_0$  (■) as shown. (a thin layer of insulation at each corner prevents the metal plates from shorting out just for pedantry). Find the potential inside the resulting rectangular pipe.

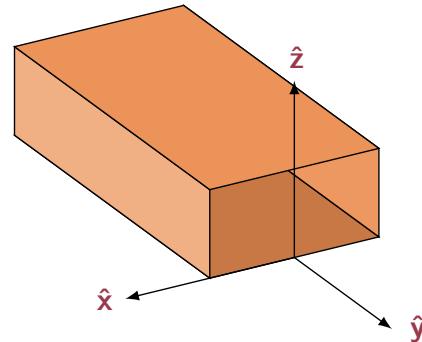


Figure 4.7

As can be seen, the configuration is **independent** from  $z$ -direction. Therefore, our goal here is to solve a 2D Laplace's equation ( $\nabla^2 f = 0$ ) is as follows:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0,$$

which are subject to the following boundary conditions:

- |                                       |                                       |
|---------------------------------------|---------------------------------------|
| <b>(i)</b> $V = 0$ when $y = 0$ ,     | <b>(ii)</b> $V = 0$ when $y = a$ ,    |
| <b>(iii)</b> $V = V_0$ when $x = b$ , | <b>(iv)</b> $V = V_0$ when $x = -b$ . |

To solve this problem we need to use **Separation of Variables** method which is as follows. We will assume a solution with the following description:

$$V(x, y) = X(x)Y(y) \quad \text{putting to Laplace, we get} \quad Y(y) \frac{d^2 X}{dx^2} + X(x) \frac{d^2 Y}{dy^2} = 0.$$

It is time to **isolate** the variables,

$$\frac{1}{X} \frac{d^2X}{dx^2} + \frac{1}{Y} \frac{d^2Y}{dy^2} = 0.$$

First term depends on  $x$ , and the second term depends on  $y$ . From this we can write:

$$f(x) + g(y) = 0.$$

For this to be true both of these functions needs to be constants.

$$\frac{d^2X}{dx^2} = k^2 X, \quad \frac{d^2Y}{dy^2} = -k^2 Y.$$

### Deriving the ODE Solutions

It is time to solve these equations, lets have a look at the first one:

$$\frac{d^2X}{dx^2} = k^2 X.$$

As we should be aware by now, this is classified as an homogeneous 2<sup>nd</sup>-order linear differential equation with the generic form of:<sup>27</sup>

$$a \frac{d^2y}{dx^2} + b \frac{dy}{dx} + c y(x) = 0.$$

where  $a, b, c$  are arbitrary constants to the equation. As this is a homogeneous equation, we can use the **characteristic** equation form to determine the quadratic form. The general solution for these kind of equation is along with the root formula:

$$y(x) = A e^{-\lambda_1 x} + B e^{-\lambda_2 x} \quad \text{where} \quad \lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Based on this form, the solution to the first ODE is;

$$X(x) = A \exp(kx) + B \exp(-kx), \quad \text{where} \quad \lambda_1 = -k \quad \text{and} \quad \lambda_2 = +k.$$

If we were to apply the second ODE to the previously mentioned form, we arrive at the following.

$$Y(y) = C \exp(-\mathbf{j} ky) + D \exp(\mathbf{j} ky)$$

While the imaginary solution is perfectly valid, we want to use the real only solution as the equation we are trying to solve is real as well. To convert this let's use Euler's identity.

$$\exp(\mathbf{j} \theta) = \cos \theta + \mathbf{j} \sin \theta \quad \text{which gives} \quad Y(y) = C \sin ky + D \cos ky.$$

The solution to this problem is as follows the same path we have looked at the lecture up to the following part:

$$V(x, y) = (A \exp(kx) + B \exp(-kx)) (C \sin ky + D \cos ky)$$

<sup>27</sup>If you need a light refreshment on this topic, please have a look at the beginning chapters of M.Sc Higher Mathematics I LectureBook.

## Applying the Boundary Conditions

Now we have the equation let's restrict the equation so we can get a singular solution.

To begin, we have to say, we cannot set  $A = 0$  as the region in question does not extend to  $x = \infty$ , so  $\exp kx$  is perfectly acceptable, whereas, the situation is symmetric with respect to  $x$ , so  $V(-x, y) = V(x, y)$  and it follows  $A = B$ . Using the hyperbolic cosine function identity:

$$A \exp kx + A \exp -kx = 2A \cosh kx,$$

distributing  $A$  to  $C$  and  $D$ , we have:<sup>28</sup>

$$V(x, y) = \cosh kx (2AC \sin ky + 2AD \cos ky)$$

Boundary conditions **(i)** and **(ii)** require  $D$  to be equal to zero and  $k = n\pi/a$  so:

$$V(x, y) = 2AC \cosh(n\pi x/a) \sin(n\pi y/a),$$

Because  $V(x, y)$  is even in  $x$ , it will automatically meet condition **(iv)** if it fits **(iii)**. It remains, therefore, to construct the linear combination,

$$V(x, y) = \sum_{n=1}^{\infty} G_n \cosh(n\pi x/a) \sin(n\pi y/a), \quad \text{where } G = 2AC.$$

and pick coefficients  $G_n$  such that it satisfies the condition on **(iii)**:

$$V(b, y) = \sum_{n=1}^{\infty} G_n \cosh(n\pi b/a) \sin(n\pi y/a) = V_0,$$

<sup>28</sup>We do this distribution as we don't care about individual coefficients, so only their end product. So we can state, here  $A \times C \rightarrow C$ .

## Simplifying the Fourier Expression

This is the same problem in *Fourier analysis* we worked in the exercise in part 3.

To solve this equation, we first need to multiply both sides with  $\sin \frac{m\pi y}{a}$ .

$$\sin \frac{m\pi y}{a} V_0(0) = \sum_{n=1}^{\infty} G_n \cosh(n\pi b/a) \sin \frac{n\pi y}{a} \sin \frac{m\pi y}{a},$$

Time to integrate both sides over  $y$ , from zero to  $a$ :

$$\int_0^a \sin \frac{m\pi y}{a} V_0(0) dy = \sum_{n=1}^{\infty} G_n \cosh(n\pi b/a) \int_0^a \sin \frac{n\pi y}{a} \sin \frac{m\pi y}{a} dy,$$

This is basically **Fourier Transform** on both sides. It is time to focus on the integration on the right hand side.

$$\begin{aligned} \int_0^a \sin \frac{n\pi y}{a} \sin \frac{m\pi y}{a} dy &= \left[ -\frac{a}{n\pi} \sin \frac{mxy}{a} \cos \frac{n\pi y}{a} \right]_0^a + \frac{m}{n} \int_0^a \cos \frac{m\pi y}{a} \cos \frac{n\pi y}{a} dy \\ &= \frac{m}{n} \left[ \frac{a}{n\pi} \cos \frac{m\pi y}{a} \sin \frac{n\pi y}{a} \right]_0^a + \frac{m}{n} \int_0^a \sin \frac{m\pi y}{a} \sin \frac{n\pi y}{a} dy \end{aligned}$$

Simplification of this integral yields,

$$\left(1 - \frac{m^2}{n^2}\right) \int_0^a \sin \frac{m\pi y}{a} \sin \frac{n\pi y}{a} dy = 0.$$

There are two possibilities here: either  $m = n$  or  $m \neq n$ . If the latter is the case, then:

$$\int_0^a \sin \frac{m\pi y}{a} \sin \frac{n\pi y}{a} dy = 0 \quad (m \neq n)$$

If on the other hand  $m = n$ , then:

$$\begin{aligned} \int_0^a \sin^2 \frac{n\pi y}{a} dy &= -\frac{a}{n\pi} \sin \frac{n\pi y}{a} \cos \frac{n\pi y}{a} \Big|_0^a + \int_0^a \cos^2 \frac{n\pi y}{a} dy \\ 2 \int_0^a \sin^2 \frac{n\pi y}{a} dy &= \int_0^a \sin^2 \frac{n\pi y}{a} dy + \int_0^a \cos^2 \frac{n\pi y}{a} dy = \int_0^a dy = a. \end{aligned}$$

Therefore:

$$\int_0^a \sin \frac{n\pi y}{a} \sin \frac{m\pi y}{a} dy = \begin{cases} 0, & \text{if } n \neq m \\ \frac{a}{2}, & \text{if } n = m \end{cases}$$

Based on this, we only care about the solutions where  $m = n$ . If we take our focus to the rest of the equation,

$$\int_0^a \sin \frac{n\pi y}{a} V_0 dy = \sum_1^\infty G_n \cosh(n\pi b/a) \overbrace{\int_0^a \sin \frac{n\pi y}{a} \sin \frac{n\pi y}{a} dy}^{=a/2},$$

If we isolate the  $G_n$  on one side:

$$G_n = \frac{2}{a \cosh(n\pi b/a)} \int_0^a \sin \frac{n\pi y}{a} V_0 dy$$

Doing the integral on the right hand side yields the following:

$$G_n = \frac{2}{a \cosh(n\pi b/a)} \int_0^a \sin \frac{n\pi y}{a} V_0 dy = \frac{1}{\cosh(n\pi b/a)} \frac{2V_0}{n\pi} (1 - \cos n\pi).$$

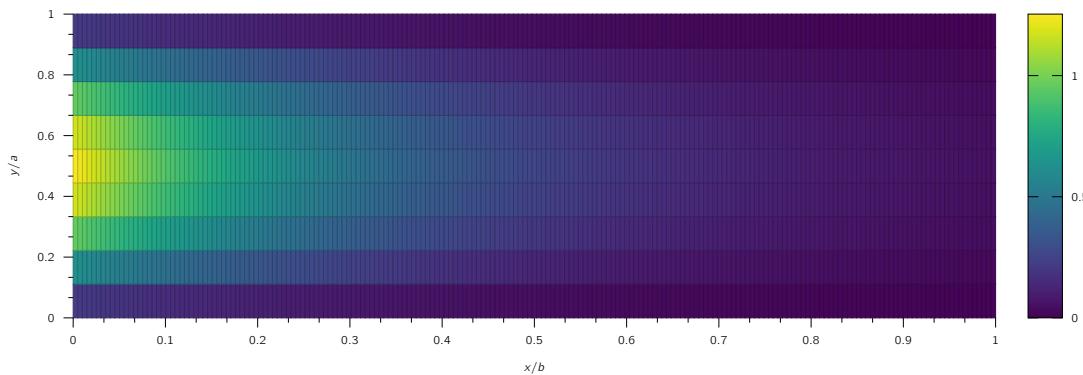
The results are:

$$C_n \cosh(n\pi b/a) = \begin{cases} 0, & \text{if } n \text{ is even,} \\ \frac{4V_0}{n\pi}, & \text{if } n \text{ is odd.} \end{cases}$$

The potential in this case is given by:

$$V(x, y) = \frac{4V_0}{\pi} \sum_{n=1, 3, 5, \dots} \frac{1}{n} \frac{\cosh(n\pi x/a)}{\cosh(n\pi b/a)} \sin(n\pi y/a) \quad \nabla \blacksquare$$

The success of this method relies on two (2) important properties of the separable solutions (Eqs. 3.28 and 3.29): **completeness** and **orthogonality**.



**Figure 4.8**  
The solution to the 2D-Grounded plate. Please observe that the solution is generated by summing up consecutive parts of a Fourier series to achieve a more accurate result.

A set of functions  $f_n(y)$  is said to be complete if any other function  $f(y)$  can be expressed as a linear combination of them:

$$f(y) = \sum_{n=1}^{\infty} C_n f_n(y)$$

The functions  $\sin(n\pi y/a)$  are complete on the interval  $0 \leq y \leq a$ . It was this fact, guaranteed by Dirichlet's theorem, which assured us Eq. 3.31 could be satisfied, given the proper choice of the coefficients  $C_n$ .<sup>29</sup>

A set of functions is **orthogonal** if the integral of the product of any two (2) different members of the set is zero:

$$\int_0^a f_n(y) f_{n'}(y) dy = 0 \quad \text{for} \quad n' \neq n.$$

<sup>29</sup>The proof of completeness, for a particular set of functions, is an extremely difficult business, and unfortunately this margin is too narrow to write.

The sine functions are orthogonal (Eq. 3.33). This is the property on which Fourier's trick is based, allowing us to kill off all terms but one in the infinite series and thereby solve for the coefficients  $C_n$ .

### Exercise 4.1: Spherical Separation of Variables

The potential  $V_0(\theta)$  is again specified on the surface of a sphere of radius  $R$ , but this time we are asked to find the potential **outside**, assuming there is no charge there.

#### Solution

In this case it's the  $A_l$ 's that must be zero (or else  $V$  would not go to zero at  $\infty$ ), so

$$V(r, \theta) = \sum_{l=0}^{\infty} \frac{B_l}{r^{l+1}} P_l(\cos \theta).$$

At the surface of the sphere, we require that

$$V(R, \theta) = \sum_{l=0}^{\infty} \frac{B_l}{R^{l+1}} P_l(\cos \theta) = V_0(\theta).$$

Multiplying by  $P_{l'}(\cos \theta) \sin \theta$  and integrating—exploiting, again, the orthogonality relation 3.68—we have

$$\frac{B_{l'}}{R^{l'+1}} \frac{2}{2l'+1} = \int_0^{\pi} \pi V_0(\theta) P_{l'}(\cos \theta) \sin \theta d\theta,$$

or

$$B_{l'} = \frac{2l'+1}{2} R^{l+1} \int_0^{\pi} \pi V_0(\theta) P_l(\cos \theta) \sin \theta d\theta.$$

Equation 3.72, with the coefficients given by Eq. 3.73, is the solution to our problem. ■

### 4.3.2 Case Study: 3D Rectangular Pipe

Let's turn the problem up a bit and study a 3D case. For this, imagine, an infinitely long rectangular metal pipe (sides  $a$  and  $b$ ) is grounded, but one end, at  $x = 0$ , is maintained at a specified potential  $V_0(y, z)$ , as indicated in Fig. 3.22.

We will try to find the potential inside the pipe. To start, we can say that this is a 3D problem with the following Laplace equation:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0,$$

which are subject to the following boundary conditions:

#### Deriving the ODEs

As usual, we are looking for solutions that are products:

$$V(x, y, z) = X(x) Y(y) Z(z).$$

Putting this into the Laplace equation presents us:

$$\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} + \frac{1}{Z} \frac{d^2 Z}{dz^2} = 0.$$

It follows that:

$$\frac{1}{X} \frac{d^2 X}{dx^2} = C_1, \quad \frac{1}{Y} \frac{d^2 Y}{dy^2} = C_2, \quad \frac{1}{Z} \frac{d^2 Z}{dz^2} = C_3, \quad \text{where } C_1 + C_2 + C_3 = 0.$$

For simplicity, we will assume  $C_1$  must be positive and  $C_2, C_3$  are negative. Setting  $C_2 = -k^2$  and  $C_3 = -l^2$ , we have  $C_1 = k^2 + l^2$ , and therefore:

$$\frac{d^2 X}{dx^2} = (k^2 + l^2) X, \quad \frac{d^2 Y}{dy^2} = -k^2 Y, \quad \frac{d^2 Z}{dz^2} = -l^2 Z$$

We have yet again turned a PDE into a series of ODE. The solutions are:

$$X(x) = A \exp(\sqrt{k^2 + l^2} x) + B \exp(-\sqrt{k^2 + l^2} x),$$

$$Y(y) = C \sin ky + D \cos ky,$$

$$Z(z) = E \sin lz + F \cos lz.$$

#### Applying the Boundary Conditions

Boundary condition **(v)** implies  $A = 0$ , **(i)** gives  $D = 0$  and **(iii)** gives  $F = 0$ , whereas **(ii)** and **(iv)** require  $k = n\pi/a$  and  $l = m\pi/b$ , where  $n$  and  $m$  are positive integers.

Combining the remaining constants, we are presented with:

$$V(x, y, z) = C \exp\left(-\pi\sqrt{(n/a)^2 + (m/b)^2} x\right) \sin(n\pi y/a) \sin(m\pi z/b).$$

This solution meets all the boundary condition except **(vi)**. It contains two (2) unspecified integers ( $m$  and  $n$ ) and the most general linear combination is a **double sum**:

$$V(x, y, z) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} C_{n,m} \exp\left(-\pi\sqrt{(n/a)^2 + (m/b)^2}x\right) \sin(n\pi y/a) \sin(m\pi z/b).$$

### Deriving the Coefficients

We hope to fit the remaining boundary condition:

$$V(0, y, z) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} C_{n,m} \sin(n\pi y/a) \sin(m\pi z/b) = V_0(y, z).$$

by choosing appropriate coefficients of  $C_{n,m}$ . To determine these constants, we multiply by  $\sin(n'\pi y/a) \sin(m'\pi z/b)$ , where  $n'$  and  $m'$  are arbitrary positive integers, and integrate:

$$\begin{aligned} & \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} C_{n,m} \int_0^a \sin(n\pi y/a) \sin(n'\pi y/a) dy \int_0^b \sin(m\pi z/b) \sin(m'\pi z/b) dz \\ &= \int_0^a \int_0^b V_0(y, z) \sin(n\pi y/a) \sin(m\pi z/b) dy dz. \end{aligned}$$

So:

$$C_{n,m} = \frac{4}{ab} \int_0^a \int_0^b V_0(y, z) \sin(n\pi y/a) \sin(m\pi z/b) dy dz.$$

We are almost at the end of our problem. For example, if the end of the tube is a conductor at constant potential  $V_0$ :

$$\begin{aligned} C_{n,m} &= \frac{4V_0}{ab} \int_0^a \sin(n\pi y/a) dy \int_0^b \sin(m\pi z/b) dz \\ &= \begin{cases} 0, & \text{if } n \text{ or } m \text{ is even,} \\ \frac{16V_0}{\pi^2 nm}, & \text{if } n \text{ and } m \text{ are odd.} \end{cases} \end{aligned}$$

In this case:

$$V(0, y, z) = \frac{16V_0}{\pi^2} \sum_{n,m=1,3,5,\dots}^{\infty} \frac{1}{nm} \exp\left(-\pi\sqrt{(n/a)^2 + (m/b)^2}x\right) \sin(n\pi y/a) \sin(m\pi z/b)$$

As the successive terms decrease rapidly, a reasonable approximation would be obtained by keeping only the first few terms ■

### 4.3.3 Spherical Coordinates

In the examples considered thus far, Cartesian coordinates were clearly appropriate, given the boundaries were planes. For round objects, however, spherical coordinates are more natural. In the spherical system, Laplace's equation becomes:

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial V}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial V}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 V}{\partial \phi^2} = 0. \quad (4.8)$$

We shall assume the problem has **azimuthal symmetry**, so that  $V$  is independent of  $\phi$  in that case, Eq. (4.8) reduces to:

$$\frac{\partial}{\partial r} \left( r^2 \frac{\partial^2 V}{\partial r^2} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial V}{\partial \theta} \right) = 0.$$

Similar to Cartesian coordinates, we look for solutions that are products:

$$V(r, \theta) = R(r) \Theta(\theta). \quad (4.9)$$

As before, we look for solutions that are products:

$$V(r, \theta) = R(r)\Theta(\theta). \quad (4.10)$$

Putting this into Eq. (4.9), and dividing by  $V$ , we get the following:

$$\frac{1}{R} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) + \frac{1}{\Theta \sin \theta} \frac{d}{d\theta} \left( \sin \theta \frac{d\Theta}{d\theta} \right) = 0. \quad (4.11)$$

Since the first term is dependent only on  $r$ , and the second only on  $\theta$ , it follows that each must be a constant:

$$\frac{1}{R} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) = I(I+1), \quad \text{and} \quad \frac{1}{\Theta \sin \theta} \frac{d}{d\theta} \left( \sin \theta \frac{d\Theta}{d\theta} \right) = -I(I+1). \quad (4.12)$$

Here  $I(I+1)$  is just a prettier way of writing the separation constant.<sup>30</sup> As always, separation of variables has converted a PDE in Eq. (4.9) into pair of ODEs Eq. (4.12).

<sup>30</sup>We will see its usefulness in just a minute.

To solve, lets start with the radial equation:

$$\frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) = I(I+1)R, \quad (4.13)$$

has the following general solution:

$$R(r) = Ar^I + \frac{B}{r^{I+1}}, \quad (4.14)$$

as we can see,  $A$  and  $B$  are the two (2) arbitrary constants to be expected in the solution of a 2<sup>nd</sup>-orderODE. But the angular equation,

$$\frac{d}{d\theta} \left( \sin \theta \frac{d\Theta}{d\theta} \right) = -I(I+1) \sin \theta \Theta, \quad (4.15)$$

is unfortunately, **NOT** so simple. The solutions are **Legendre polynomials** in the variable  $\cos \theta$ :

$$\Theta(\theta) = P_I(\cdot) \cos \theta. \quad (4.16)$$

$P_I(x)$  is most conveniently defined by the **Rodrigues formula**:

$$P_I(x) \equiv \frac{1}{2^I I!} \left( \frac{d}{dx} \right)^I (x^2 - 1)^I. \quad (4.17)$$



Table 4.1: Legendre Polynomials

Polynomial	Value
$P_0(x)$	1
$P_1(x)$	1
$P_2(x)$	$x$
$P_3(x)$	$(3x^2 - 1)/2$
$P_4(x)$	$(5x^3 - 3x)/2$
$P_5(x)$	$(35x^4 - 30x^2 + 3)/8$
$P_6(x)$	$(63x^5 - 70x^3 + 15x)/8$

The first few Legendre polynomials are listed in **Tbl.** 4.1 and the rest of the values can easily be generated using Eq. (4.17).

Notice that  $P_l(x)$  is<sup>31</sup> an  $l^{\text{th}}$  order polynomial in  $x$ ; it contains only even powers, if  $l$  is even, and odd powers, if  $l$  is odd. The factor in front ( $1/l!l!$ ) was chosen in order that

$$P_l(1) = 1. \quad (4.18)$$

The Rodrigues formula<sup>32</sup> obviously works only for non-negative integer values of  $l$ .

Moreover, it provides us with only one (1) solution. But Eq. (4.15) is 2<sup>nd</sup>- order, and it should possess two (2) independent solutions, for every value of  $l$ . It turns out that these “other solutions” blow up at  $\theta = 0$  and/or  $\theta = \pi$ , and are therefore unacceptable on physical grounds.

For instance, the second solution for  $l = 0$  is

$$\Theta(\theta) = \ln \left( \tan \frac{\theta}{2} \right). \quad (4.19)$$

Which as we all now is undefined. We can of course check for our self that this satisfies Eq. (4.15). In the case of azimuthal symmetry, then, the most general separable solution to Laplace’s equation, consistent with minimal physical requirements, is:<sup>33</sup>

$$V(r, \theta) = \left( Ar^l + \frac{B}{r^{l+1}} \right) P_l(\cos \theta).$$

As before, separation of variables gives an infinite set of solutions, one for each  $l$ . The general solution is the linear combination of separable solutions:

$$V(r, \theta) = \sum_{l=0}^{\infty} \left( A_l r^l + \frac{B_l}{r^{l+1}} \right) P_l(\cos \theta). \quad (4.20)$$

The following examples illustrate the power of this important result.

<sup>31</sup>As the name obviously suggests.



<sup>32</sup>Benjamin Olinde Rodrigues (1795 - 1851)

A French banker, mathematician, and social reformer. In mathematics include an overall constant Rodrigues is remembered in Eq.(4.16), as it can be for Rodrigues’ rotation absorbed into  $A$  and  $B$  at formula for vectors, the this stage, Rodrigues’ formula for the Legendre polynomials, and the Euler-Rodrigues parameters.

**Exercise 4.2 | Hollow Sphere Spherical Coordinates**

The potential  $V_0(\theta)$  is specified on the surface of a hollow sphere, of radius  $R$ . Find the potential inside the sphere.

**SOLUTION**

In this case,  $B_l = 0$  for all  $l$ —otherwise the potential would blow up at the origin. Thus,

$$V(r, \theta) = \sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta). \quad (4.21)$$

<sup>13</sup>In rare cases where the z axis is excluded, these "other solutions" do have to be considered.

At  $r = R$  this must match the specified function  $V_0(\theta)$ :

$$V(R, \theta) = \sum_{l=0}^{\infty} A_l R^l P_l(\cos \theta) = V_0(\theta). \quad (4.22)$$

Can this equation be satisfied, for an appropriate choice of coefficients  $A_l$ ? Yes: The Legendre polynomials (like the sines) constitute a complete set of functions, on the interval  $-1 \leq x \leq 1$  ( $0 \leq \theta \leq \pi$ ). How do we determine the constants? Again, by Fourier's trick, for the Legendre polynomials (like the sines) are orthogonal functions:  $\int_{-1}^1 P_l(x) P_{l'}(x) dx = \int_0^\pi P_l(\cos \theta) P_{l'}(\cos \theta) \sin \theta d\theta =$

$$\begin{cases} 0, & \text{if } l' \neq l, \\ \frac{2}{2l+1}, & \text{if } l' = l. \end{cases} \quad (4.23)$$

Thus, multiplying Eq. 3.67 by  $P_l(\cos \theta) \sin \theta$  and integrating, we have  $A_l R^l \frac{2}{2l+1} = \int_0^\pi V_0(\theta) P_l(\cos \theta) \sin \theta d\theta$ ,

$$A_l = \frac{2l+1}{2l} \int_0^\pi V_0(\theta) P_l(\cos \theta) \sin \theta d\theta. \quad (4.24)$$

Equation 3.66 is the solution to our problem, with the coefficients given by Eq. 3.69. It can be difficult to evaluate integrals of the form 3.69 analytically, and in practice it is often easier to solve Eq. 3.67 "by eyeball." For instance, suppose we are told that the potential on the sphere is  $V_0(\theta) = k \sin^2(\theta/2)$ . (3.70) where  $k$  is a constant. Using the half-angle formula, we rewrite this as  $V_0(\theta) = \frac{k}{2}(1 - \cos \theta) = \frac{k}{2}[P_0(\cos \theta) - P_1(\cos \theta)]$ .

Putting this into Eq. 3.67, we read off immediately that  $A_0 = k/2$ ,  $A_1 = -k/(2R)$ , and all other  $A_l$ 's vanish. Therefore,

$$V(r, \theta) = \frac{k}{2} \left[ r^0 P_0(\cos \theta) - \frac{r^1}{R} P_1(\cos \theta) \right] = \frac{k}{2} \left( 1 - \frac{r}{R} \cos \theta \right). \quad (4.25)$$

**Exercise 4.3 | Charge Density**

A specified charge density  $\sigma_0(\theta)$  is glued over the surface of a spherical shell of radius  $R$ . Find the resulting potential inside and outside the sphere.

**SOLUTION** You could, of course, do this by direct integration:

$$V = \frac{1}{4\pi\epsilon_0} \int \frac{\sigma_0}{\lambda} da,$$

but separation of variables is often easier. For the interior region, we have

$$V(r, \theta) = \sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta) \quad (r \leq R)$$

(no  $B_l$  terms—they blow up at the origin); in the exterior region

$$V(r, \theta) = \sum_{l=0}^{\infty} \frac{B_l}{r^{l+1}} P_l(\cos \theta) \quad (r \geq R)$$

(no  $A_l$  terms—they don't go to zero at infinity). These two functions must be joined together by the appropriate boundary conditions at the surface itself. First, the potential is *continuous* at  $r = R$  (Eq. 2.34):

$$\sum_{l=0}^{\infty} A_l R^l P_l(\cos \theta) = \sum_{l=0}^{\infty} [\infty] \frac{B_l}{R^{l+1}} P_l(\cos \theta).$$

It follows that the coefficients of like Legendre polynomials are equal:

$$B_l = A_l R^{2l+1}.$$

(To prove that formally, multiply both sides of Eq. 3.80 by  $P_{l'}(\cos \theta) \sin \theta$  and integrate from 0 to  $\pi$ , using the or-

thogonality relation 3.68.) Second, the radial derivative of  $V$  suffers a discontinuity at the surface (Eq. 2.36):

$$\left( \frac{\partial V_{\text{out}}}{\partial r} - \frac{\partial V_{\text{in}}}{\partial r} \right) \Big|_{r=R} = -\frac{1}{\epsilon_0} \sigma_0(\theta).$$

Thus

$$-\sum_{l=0}^{\infty} (l+1) \frac{B_l}{R^{l+2}} P_l(\cos \theta) - \sum_{l=0}^{\infty} [\infty] l A_l R^{l-1} P_l(\cos \theta) = -\frac{1}{\epsilon_0} \sigma_0(\theta),$$

or, using Eq. 3.81,

$$\sum_{l=0}^{\infty} (2l+1) A_l R^{l-1} P_l(\cos \theta) = \frac{1}{\epsilon_0} \sigma_0(\theta).$$

From here, the coefficients can be determined using Fourier's trick:

$$A_l = \frac{1}{2\epsilon_0 R^{l-1}} \int_0^\pi \sigma_0(\theta) P_l(\cos \theta) \sin \theta d\theta.$$

Equations 3.78 and 3.79 constitute the solution to our problem, with the coefficients given by Eqs. 3.81 and 3.84.

For instance, if

$$\sigma_0(\theta) = k \cos \theta = k P_1(\cos \theta),$$

for some constant  $k$ , then all the  $A_i$ 's are zero except for  $l = 1$ , and

$$A_1 = \frac{k}{2\epsilon_0} \int_0^\pi [P_1(\cos \theta)]^2 \sin \theta d\theta = \frac{k}{3\epsilon_0}.$$

The potential inside the sphere is therefore

$$V(r, \theta) = \frac{k}{3\epsilon_0} r \cos \theta \quad (r \leq R),$$

whereas outside the sphere

$$V(r, \theta) = \frac{kR^3}{3\epsilon_0} \frac{1}{r^2} \cos \theta \quad (r \geq R).$$

In particular, if  $\sigma_0(\theta)$  is the induced charge on a metal sphere in an external field  $E_0 \hat{z}$ , so that  $k = 3\epsilon_0 E_0$  (Eq. 3.77), then the potential inside is  $E_0 r \cos \theta = E_0 z$ , and the field is  $-E_0 \hat{z}$ —exactly right to cancel off the external field, as of course it should be. Outside the sphere the potential due to this surface charge is

$$E_0 \frac{R^3}{r^2} \cos \theta,$$

consistent with our conclusion in Ex. 3.8.

#### Exercise 4.4 Uncharged Metal Sphere

An uncharged metal sphere of radius  $R$  is placed in an otherwise uniform electric field  $E = E_0 \hat{z}$ . The field will push positive charge to the "northern" surface of the sphere, and-symmetrially-negative charge to the "southern" surface (Fig. 3.24). This induced charge, in turn, distorts the field in the neighborhood of the sphere. Find the potential in the region outside the sphere.

##### SOLUTION

The sphere is an equipotential—we may as well set it to zero. Then by symmetry the entire  $xy$  plane is at potential zero. This time, however,  $V$  does **not** go to zero at large  $z$ . In fact, far from the sphere the field is  $E_0 \hat{z}$ , and hence  $V \rightarrow -E_0 z + C$ .

«FIGURE»

Since  $V = 0$  in the equatorial plane, the constant  $C$  must be zero. Accordingly, the boundary conditions for this problem are (i)  $V = 0$  when  $r = R$ , (ii)  $V \rightarrow -E_0 r \cos \theta$  for  $r \gg R$ . (3.74) We must fit these boundary conditions with a function of the form 3.65. The first condition yields  $A_l R^l + \frac{B_l}{R^{l+1}} = 0$ ,

$$B_l = -A_l R^{2l+1}, \quad (4.26)$$

$$\text{so } V(r, \theta) = \sum_{l=0}^{\infty} A_l \left( r^l - \frac{R^{2l+1}}{r^{l+1}} \right) P_l(\cos \theta).$$

For  $r \gg R$ , the second term in parentheses is negligible, and therefore condition (ii) requires that  $\sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta) = -E_0 r \cos \theta$ .

Evidently only one term is present:  $l = 1$ . In fact, since  $P_1(\cos \theta) = \cos \theta$ , we can read off immediately  $A_1 = -E_0$ , all other  $A_i$ 's zero. Conclusion:

$$V(r, \theta) = -E_0 \left( r - \frac{R^3}{r^2} \right) \cos \theta. \quad (4.27)$$

The first term ( $-E_0 r \cos \theta$ ) is due to the external field; the contribution attributable to the induced charge is  $E_0 \frac{R^3}{r^2} \cos \theta$ . If you want to know the induced charge density, it can be calculated in the usual way:

$$\sigma(\theta) = -\epsilon_0 \frac{\partial V}{\partial r} \Big|_{r=R} = \epsilon_0 E_0 \left(1 + 2 \frac{R^3}{r^3}\right) \cos \theta \Big|_{r=R} = 3\epsilon_0 E_0 \cos \theta. \quad (3.77)$$

As expected, it is positive in the "northern" hemisphere ( $0 \leq \theta \leq \pi/2$ ) and negative in the "southern" ( $\pi/2 \leq \theta \leq \pi$ ). ■

**Example 3.9.** A specified charge density  $\sigma_0(\theta)$  is glued over the surface of a spherical shell of radius  $R$ . Find the resulting potential inside and outside the sphere. Solution You could, of course, do this by direct integration:  $V = \frac{1}{4\pi\epsilon_0} \int \frac{\sigma_0}{r} da$ , but separation of variables is often easier. For the interior region, we have

$$V(r, \theta) = \sum_{i=1}^{\infty} A_i r^i P_i(\cos \theta) \quad (r \leq R) \quad (4.28)$$

(no  $B_i$  terms-they blow up at the origin); in the exterior region

$$V(r, \theta) = \sum_{i=0}^{\infty} \frac{B_i}{r^{i+1}} P_i(\cos \theta) \quad (r \geq R) \quad (4.29)$$

(no  $A_i$  terms-they don't go to zero at infinity). These two functions must be joined together by the appropriate boundary conditions at the surface itself. First, the potential is **continuous** at  $r = R$  (Eq. 2.34):

$$\sum_{l=0}^{\infty} A_l R^l P_l(\cos \theta) = \sum_{l=0}^{\infty} \frac{B_l}{R^{l+1}} P_l(\cos \theta). \quad (4.30)$$

It follows that the coefficients of like Legendre polynomials are equal:

$$B_l = A_l R^{2l+1}. \quad (4.31)$$

(To prove that formally, multiply both sides of Eq. 3.80 by  $P_l(\cos \theta) \sin \theta$  and integrate from 0 to  $\pi$ , using the orthogonality relation 3.68.) Second, the radial derivative of  $V$  suffers a discontinuity at the surface (Eq. 2.36):

$$\left( \frac{\partial V_{\text{out}}}{\partial r} - \frac{\partial V_{\text{in}}}{\partial r} \right) \Big|_{r=R} = -\frac{1}{\epsilon_0} \sigma_0(\theta). \quad (4.32)$$

Thus  $-\sum_{l=0}^{\infty} (l+1) \frac{B_l}{R^{l+2}} P_l(\cos \theta) - \sum_{l=0}^{\infty} l A_l R^{l-1} P_l(\cos \theta) = -\frac{1}{\epsilon_0} \sigma_0(\theta)$ , or, using Eq. 3.81,

$$\sum_{k=0}^{\infty} (2k+1) A_k R^{2k-1} P_k(\cos \theta) = \frac{1}{\epsilon_0} \sigma_0(\theta). \quad (4.33)$$

From here, the coefficients can be determined using Fourier's trick:

$$A_l = \frac{1}{2\epsilon_0 R^{l-1}} \int_0^\pi \sigma_0(\theta) P_l(\cos \theta) \sin \theta d\theta. \quad (4.34)$$

Equations 3.78 and 3.79 constitute the solution to our problem, with the coefficients given by Eqs. 3.81 and 3.84. For instance, if

$$\sigma_0(\theta) = k \cos \theta = k P_1(\cos \theta), \quad (4.35)$$

for some constant  $k$ , then all the  $A_l$ 's are zero except for  $l = 1$ , and  $A_1 = \frac{k}{2\epsilon_0} \int_0^\pi [P_1(\cos \theta)]^2 \sin \theta d\theta = \frac{k}{3\epsilon_0}$ . The potential inside the sphere is therefore

$$V(r, \theta) = \frac{k}{3\epsilon_0} r \cos \theta \quad (r \leq R), \quad (4.36)$$

whereas outside the sphere

$$V(r, \theta) = \frac{kR^3}{3\epsilon_0} \frac{1}{r^2} \cos \theta \quad (r \geq R). \quad (4.37)$$

In particular, if  $\sigma_0(\theta)$  is the induced charge on a metal sphere in an external field  $E_0 \hat{z}$ , so that  $k = 3\epsilon_0 E_0$  (Eq. 3.77), then the potential inside is  $E_0 r \cos \theta = E_0 z$ , and the field is  $-E_0 \hat{z}$ -exactly right to cancel off the external field, as of course it should be. Outside the sphere the potential due to this surface charge is  $E_0 \frac{R^3}{r^2} \cos \theta$ , consistent with our conclusion in Ex. 3.8.

### Exercise 4.5 A Long Rectangular Tube

An infinitely long rectangular metal pipe (sides  $a$  and  $b$ ) is grounded, but one end at  $x = 0$ , is maintained at a specific potential  $V_0(y, z)$  as indicated above. Find the potential inside the pipe.

**SOLUTION** This is a 3D problem with the following Laplace equation:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0,$$

which are subject to the following boundary conditions given by the question either explicitly or implicitly:

- |                       |                             |                      |                |
|-----------------------|-----------------------------|----------------------|----------------|
| (i) $V = 0$           | when $y = 0$ ,              | (ii) $V = 0$         | when $y = a$ , |
| (iii) $V = 0$         | when $z = 0$ ,              | (iv) $V = 0$         | when $z = b$ , |
| (v) $V \rightarrow 0$ | as $x \rightarrow \infty$ , | (vi) $V = V_0(x, y)$ | when $x = 0$ . |

As usual, we are looking for solutions that are products using separation of variables:

$$V(x, y, z) = X(x) Y(y) Z(z)$$

Putting this into the Laplace equation presents to us:

$$\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} + \frac{1}{Z} \frac{d^2 Z}{dz^2} = 0.$$

Where once we isolate the variables, we can get the following relation:

$$\frac{1}{X} \frac{d^2 X}{dx^2} = C_1, \quad \frac{1}{Y} \frac{d^2 Y}{dy^2} = C_2, \quad \frac{1}{Z} \frac{d^2 Z}{dz^2} = C_3, \quad \text{with } C_1 + C_2 + C_3 = 0.$$

To keep this solution relatively short, We will assume  $C_1$  must be positive and  $C_2, C_3$  are negative. Setting  $C_2 = -k^2$  and  $C_3 = -l^2$ , we have  $C_1 = k^2 + l^2$ , and therefore:

$$\frac{d^2 X}{dx^2} = (k^2 + l^2) X, \quad \frac{d^2 Y}{dy^2} = -k^2 Y, \quad \frac{d^2 Z}{dz^2} = -l^2 Z$$

We have yet again turned a PDE into a series of ODEs. The solutions are:

$$X(x) = A \exp\left(\sqrt{k^2 + l^2} x\right) + B \left(-\sqrt{k^2 + l^2} x\right),$$

$$Y(y) = C \sin ky + D \cos ky,$$

$$Z(z) = E \sin lz + F \cos lz.$$

Boundary condition (v) implies  $A = 0$ , (i) gives  $D = 0$  and (iii) yields  $F = 0$ , whereas (ii) and (iv) require  $k = n\pi/a$  and  $l = m\pi/b$ , where  $n$  and  $m$  are positive integers. Combining the remaining constants, we are presented with:

$$V(x, y, z) = C \exp\left(-\pi \sqrt{(n/a)^2 + (m/b)^2} x\right) \sin(n\pi y/a) \sin(m\pi z/b).$$

This solution meets all the boundary condition except (vi). It contains two unspecified integers ( $m$  and  $n$ ) and the most general linear combination is a double sum:

$$V(x, y, z) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} C_{n,m} \exp\left(-\pi\sqrt{(n/a)^2 + (m/b)^2}x\right) \sin(n\pi y/a) \sin(m\pi z/b).$$

We hope to fit the remaining boundary condition:

$$V(0, y, z) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} C_{n,m} \sin(n\pi y/a) \sin(m\pi z/b) = V_0(y, z).$$

by choosing appropriate coefficients of  $C_{n,m}$ .

To determine these constants, we multiply by  $\sin(n'\pi y/a) \sin(m'\pi z/b)$ , where  $n'$  and  $m'$  are arbitrary positive integers, and integrate:

$$\begin{aligned} & \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} C_{n,m} \int_0^b \sin(n\pi y/a) \sin(n'\pi y/a) dy \int_0^b \sin(m\pi z/b) \sin(m'\pi z/b) dz \\ &= \int_0^a \int_0^b V_0(y, z) \sin(n\pi y/a) \sin(m\pi z/b) dy dz. \end{aligned}$$

So:

$$C_{n,m} = \frac{4}{ab} \int_0^a \int_0^b V_0(y, z) \sin(n\pi y/a) \sin(m\pi z/b) dy dz.$$

We are almost at the end of our problem. For example, if the end of the tube is a conductor at constant potential  $V_0$ :

$$\begin{aligned} C_{n,m} &= \frac{4V_0}{ab} \int_0^a \sin(n\pi y/a) dy \int_0^b \sin(m\pi z/b) dz \\ &= \begin{cases} 0, & \text{if } n \text{ or } m \text{ is even,} \\ \frac{16V_0}{\pi^2 nm}, & \text{if } n \text{ and } m \text{ are odd.} \end{cases} \end{aligned}$$

In this case:

$$V(x, y, z) = \frac{16V_0}{\pi^2} \sum_{n, m=1, 3, 5, \dots}^{\infty} \frac{1}{nm} \exp\left(-\pi\sqrt{(n/a)^2 + (m/b)^2}x\right) \sin(n\pi y/a) \sin(m\pi z/b) \quad \blacksquare$$

As the successive terms decrease rapidly, a reasonable approximation would be obtained by keeping only the first few terms.

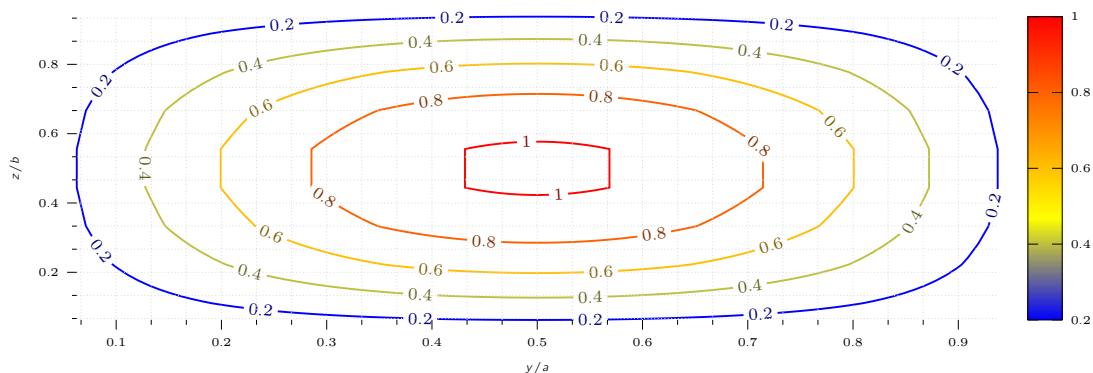


Figure 4.9

## 4.4 Multipole Expansion

### 4.4.1 Approximating Potentials at Large Distances

To start let's image we are at a point in space and we are trying to spot some charges. If we are very far away from a localised charge distribution, it may "look" like a point charge, and the potential is, to good approximation, would be:<sup>34</sup>

$$\left(\frac{1}{4\pi\epsilon_0}\right) \frac{Q}{r}$$

[V]

where  $Q$  is, of course, the total charge (C). We have often used this as a check on formulas for  $V$ .

*But what if  $Q$  is zero?*

We might think the potential is then approximately zero, and of course, we would be right, in a sense.<sup>35</sup> But we're looking for something a bit more informative than that.

<sup>34</sup>If we were to look at this a slightly humorous way, we basically state; *If it looks to small to me, It must be a tiny dot.*

<sup>35</sup>indeed, the potential at large  $r$  is pretty small even if  $Q$  is not zero

### Case Study: Estimating Field From Far Away

To get an idea of this method let's work on the following problem:

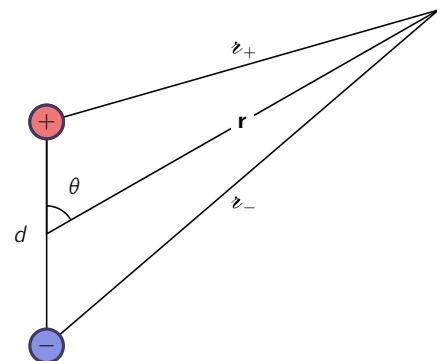
Let's have a system where a physical electric dipole consists of two (2) equal and opposite charges ( $\pm q$ ) separated by a distance  $d$ . Let's find the approximate potential ( $V(r)$ ) at points far from the dipole.

To start let's define our environment. Let  $z_-$  be the distance from  $-q$  and  $z_+$  the distance from  $+q$ , which we can see in **Fig. 4.10**. Then

$$V(r) = \frac{1}{4\pi\epsilon_0} \left( \frac{q}{z_+} - \frac{q}{z_-} \right), \quad [V]$$

and using the law of cosines we derive the following:

$$z_{\pm}^2 = r^2 + \left(\frac{d}{2}\right)^2 \mp rd \cos \theta = r^2 \left(1 \mp \frac{d}{r} \cos \theta + \frac{d^2}{4r^2}\right).$$



**Figure 4.10**  
The setup in question. Here we have two (2) charges and we are trying to estimate its value from far away.

We're interested in the scenario of  $r \gg d$ , so the third term can be assumed negligible, and the binomial expansion gives us:

$$\frac{1}{r_{\pm}} \cong \frac{1}{r} \left(1 \mp \frac{d}{r} \cos \theta\right)^{-1/2} \cong \frac{1}{r} \left(1 \pm \frac{d}{2r} \cos \theta\right).$$

Therefore:

$$\frac{1}{z_1} - \frac{1}{z_2} \cong \frac{d}{r^2} \cos \theta, \quad \therefore \quad V(r) \cong \frac{1}{4\pi\epsilon_0} \frac{qd \cos \theta}{r^2}. \quad (4.38)$$

As we can see from the aforementioned equation, the potential of a dipole decreases with a factor of  $r^2$  at large  $r$ ; as we might have anticipated, it falls off more rapidly than the potential of a point charge.

If we put together a pair of equal and opposite dipoles to make a quadrupole,<sup>36</sup> the potential goes with a factor of  $r^3$ ; for back-to-back quadrupoles<sup>37</sup>, it goes with a factor of  $r^4$ ; and so on. Figure 3.27 summarizes this hierarchy; for completeness let's include the electric monopole as well which decreases with a factor of  $r$ . ■

This short example we have conducted describes a very special charge configuration. We will now try to develop a systematic expansion for the potential of any localized charge distribution, in powers of  $1/r$ . Figure 3.28 defines the relevant variables; the potential at  $\mathbf{r}$  is given by

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{1}{\lambda} \rho(\mathbf{r}') d\tau'. \quad (4.39)$$

Using the law of cosines,

$$\epsilon^2 = r^2 + (r')^2 - 2rr' \cos \alpha = r^2 \left[ 1 + \left( \frac{r'}{r} \right)^2 - 2 \left( \frac{r'}{r} \right) \cos \alpha \right],$$

where  $\alpha$  is the angle between  $\mathbf{r}$  and  $\mathbf{r}'$ . Therefore we can write:

$$\epsilon = r\sqrt{1+\epsilon}, \quad \text{where} \quad \epsilon \equiv \left( \frac{r'}{r} \right) \left( \frac{r'}{r} - 2 \cos \alpha \right).$$

### «FIGURE»

For points **well outside** the charge distribution,  $\epsilon$  is much less than 1, and this invites a binomial expansion:

$$\frac{1}{\epsilon} = \frac{1}{r} (1 + \epsilon)^{-1/2} = \frac{1}{r} \left( 1 - \frac{1}{2}\epsilon + \frac{3}{8}\epsilon^2 - \frac{5}{16}\epsilon^3 + \dots \right), \quad (4.40)$$

or, in terms of  $r$ ,  $r'$ , and  $\alpha$ :

$$\begin{aligned} \frac{1}{\epsilon} &= \frac{1}{r} \left[ 1 - \frac{1}{2} \left( \frac{r'}{r} \right) \left( \frac{r'}{r} - 2 \cos \alpha \right) + \frac{3}{8} \left( \frac{r'}{r} \right)^2 \left( \frac{r'}{r} - 2 \cos \alpha \right)^2 \right] \\ &\quad - \frac{5}{16} \left( \frac{r'}{r} \right)^3 \left( \frac{r'}{r} - 2 \cos \alpha \right)^3 + \dots \\ &= \frac{1}{r} \left[ 1 + \left( \frac{r'}{r} \right) (\cos \alpha) + \left( \frac{r'}{r} \right)^2 \left( \frac{3 \cos^2 \alpha - 1}{2} \right) \right] \\ &\quad + \left( \frac{r'}{r} \right)^3 \left( \frac{5 \cos^3 \alpha - 3 \cos \alpha}{2} \right) + \dots \]. \end{aligned}$$

In the last step, we have collected together like powers of  $(r'/r)$ ; surprisingly, their coefficients<sup>38</sup> are **Legendre polynomials**. The remarkable result<sup>39</sup> is that

$$\frac{1}{\epsilon} = \frac{1}{r} \sum_{n=0}^{\infty} \left( \frac{r'}{r} \right)^n P_n(\cos \alpha). \quad (4.41)$$

<sup>36</sup>A major application of quadrupole configuration is its use as a **mass spectrometer**, which is the gold-standard when it comes to chemical analysis and identification.

<sup>38</sup>which are the terms inside round brackets.

<sup>39</sup>This suggests a second way of defining the Legendre polynomials, the first being Rodrigues' formula and  $1/\epsilon$  is called the **generating function** for Legendre polynomials.

Substituting this back into Eq. (4.39), and noting that  $r$  is a **constant**, as far as the integration is concerned, we can conclude the following expression

$$V(r) = \frac{1}{4\pi\epsilon_0} \sum_{n=0}^{\infty} \frac{1}{r^{(n+1)}} \int (r')^n P_n(\cos\alpha) \rho(r') d\tau' \quad (4.42)$$

or, written explicitly,

$$V(r) = \frac{1}{4\pi\epsilon_0} \left[ \frac{1}{r} \int \rho(r') d\tau' + \frac{1}{r^2} \int r' \cos\alpha \rho(r') d\tau' \right] \quad (4.43)$$

$$+ \frac{1}{r^3} \int (r')^2 \left( \frac{3}{2} \cos^2\alpha - \frac{1}{2} \right) \rho(r') d\tau' + \dots \quad (4.44)$$

This is the desired result—the **multipole expansion** of  $V$  in powers of  $1/r$ . The first term ( $n = 0$ ) is the monopole contribution (it goes like  $1/r$ ); the second ( $n = 1$ ) is the dipole (it goes like  $1/r^2$ ); the third is quadrupole; the fourth octopole; and so on. Remember that  $\alpha$  is the angle between  $\mathbf{r}$  and  $\mathbf{r}'$ , so the integrals depend on the direction to the field point. If you are interested in the potential along the  $z'$  axis (or putting it the other way around—if you orient your  $\mathbf{r}'$  coordinates so the  $z'$  axis lies along  $\mathbf{r}$ ), then  $\alpha$  is the usual polar angle  $\theta'$ . As it stands, Eq. 3.95 is exact, but it is useful primarily as an approximation scheme: the lowest nonzero term in the expansion provides the approximate potential at large  $r$ , and the successive terms tell us how to improve the approximation if greater precision is required.

## 4.4.2 Monopole and Dipole Terms

To put it simply, the multipole expansion is dominated<sup>40</sup> by the monopole term:

<sup>40</sup>when we assume  $r$  to be large

$$V_{\text{mon}}(r) = \frac{1}{4\pi\epsilon_0} \frac{Q}{r}, \quad (4.45)$$

where  $Q = \int \rho d\tau$  is the total charge of the configuration (C). This is just what we expect for the approximate potential at large distances from the charge. For a point charge placed at the origin,  $V_{\text{mon}}(r)$  is the **exact** potential and **NOT** merely a first approximation at large  $r$ .

In this case, all the higher multipoles vanish. If the total charge is zero, the dominant term in the potential will be the dipole (unless, of course, it **also** vanishes):

$$V_{\text{dip}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{1}{r^2} \int r' \cos\alpha \rho(r') d\tau'.$$

Since  $\alpha$  is the angle between  $\mathbf{r}'$  and  $\mathbf{r}$  (Fig. 3.28),

$$r' \cos\alpha = \hat{\mathbf{r}} \cdot \mathbf{r}'$$

and the dipole potential can be written more succinctly:

$$V_{\text{dip}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \int \mathbf{r}' \rho(r') d\tau'.$$

This integral (which does not depend on  $\mathbf{r}$ ) is called the **dipole moment** of the distribution:

$$\mathbf{p} \equiv \int \mathbf{r}' \rho(\mathbf{r}') d\tau', \quad (4.46)$$

and the dipole contribution to the potential simplifies to

$$V_{dip}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{\mathbf{p} \cdot \hat{\mathbf{r}}}{r^2}. \quad (4.47)$$

The dipole moment is determined by the geometry (size, shape, and density) of the charge distribution. Eq. (4.46) translates in the usual way for point, line, and surface charges. Therefore, the dipole moment of a collection of **point** charges is

$$\mathbf{p} = \sum_{i=1}^n q_i \mathbf{r}_i'. \quad (4.48)$$

For a **physical dipole**,<sup>41</sup> we write:

<sup>41</sup>equal and opposite charges,  $\pm q$ .

$$\mathbf{p} = qr'_+ - qr'_- = q(r'_+ - r'_-) = q\mathbf{d}, \quad (4.49)$$

where  $\mathbf{d}$  is the vector from the negative charge to the positive one (Fig. 3.29).

Is this consistent with what we got in Ex. 3.10? Yes: If you put Eq. 3.101 into Eq. 3.99, you recover Eq. 3.90. Notice, however, that this is only the **approximate** potential of the physical dipole—evidently there are higher multipole contributions. Of course, as you go farther and farther away,  $V_{dip}$  becomes a better and better approximation, since the higher terms die off more rapidly with increasing  $r$ . By the same token, at a fixed  $r$  the dipole approximation improves as you shrink the separation  $d$ . To construct a **perfect** (point) **dipole** whose potential is given exactly by Eq. 3.99, you'd have to let  $d$  approach zero. Unfortunately, you then lose the dipole term **too**, unless you simultaneously arrange for  $q$  to go to infinity! A **physical** dipole becomes a **pure** dipole, then, in the rather artificial limit  $d \rightarrow 0$ ,  $q \rightarrow \infty$ , with the product  $qd = p$  held fixed. When someone uses the word "dipole," you can't always tell whether they mean a physical dipole (with finite separation between the charges) or an **ideal** (point) dipole. If in doubt, assume that  $d$  is small enough (compared to  $r$ ) that you can safely apply Eq. 3.99. Dipole moments are vectors, and they add accordingly: if you have two dipoles,  $\mathbf{p}_1$  and  $\mathbf{p}_2$ , the total dipole moment is  $\mathbf{p}_1 + \mathbf{p}_2$ . For instance, with four charges at the corners of a square, as shown in Fig. 3.30, the net dipole moment is zero. You can see this by combining the charges in pairs (vertically,  $\downarrow + \uparrow = 0$ , or horizontally,  $\rightarrow + \leftarrow = 0$  or by adding up the four contributions individually, using Eq. 3.100. This is a **quadrupole**, as I indicated earlier, and its potential is dominated by the quadrupole term in the multipole expansion.

### 4.4.3 Origin of Coordinates in Multipole Expansions

We discussed previously, that a **point charge** at the origin constitutes a "pure" monopole.

If it is **NOT** at the origin, it's no longer a pure monopole.



For instance, the charge in Fig. 3.32 has a dipole moment  $\mathbf{p} = qd\hat{\mathbf{y}}$ , and a corresponding dipole term in its potential. The monopole potential  $(1/4\pi\epsilon_0)q/r$  is not quite correct for this configuration; rather, the exact potential is  $(1/4\pi\epsilon_0)q/\lambda$ . The multipole expansion is, remember, a series in inverse powers of  $r$  (the distance to the origin), and when we expand  $1/\nu$ , we get all powers, not just the first.

So moving the origin<sup>42</sup> can radically alter a multipole expansion. The **monopole moment**  $Q$  does not change, since the total charge is obviously independent of the coordinate system. (In Fig. 3.32, the monopole term was unaffected when we moved  $q$  away from the origin—it's just that it was no longer the whole story: a dipole term—and for that matter all higher poles—appeared as well.) Ordinarily, the dipole moment does change when you shift the origin, but there is an important exception: If the total charge is zero, then the dipole moment is independent of the choice of origin. For suppose we displace the origin by an amount  $\mathbf{a}$  (Fig. 3.33). The new dipole moment is then

$$\begin{aligned}\bar{\mathbf{p}} &= \int \bar{\mathbf{p}}' \rho(\mathbf{r}') d\tau' = \int (\mathbf{r}' - \mathbf{a}) \rho(\mathbf{r}') d\tau' \\ &= \int \mathbf{r}' \rho(\mathbf{r}') d\tau' - \mathbf{a} \int \rho(\mathbf{r}') d\tau' = \mathbf{p} - Q\mathbf{a}.\end{aligned}$$

In particular, if  $Q = 0$ , then  $\bar{\mathbf{p}} = \mathbf{p}$ . So if someone were for the dipole moment in Fig. 3.34(a),<sup>43</sup> we can answer with confidence “ $qd$ ”, but if we’re asked for the dipole moment in Fig. 3.34(b), the appropriate response would be:

*With respect to what origin?*

<sup>42</sup>Or, what amounts to the same thing, moving the charge

<sup>43</sup>Perhaps while you are commiserating while having lunch between lectures, perhaps?

#### 4.4.4 The Electric Field of a Dipole

So far we have worked only with potentials. Now let’s try to calculate the electric field of a **perfect** dipole. If we choose coordinates so that  $\mathbf{p}$  is at the origin and points in the  $\hat{\mathbf{z}}$  direction (Fig. 3.36), then using Eq. (4.47), the potential at  $r, \theta$  is:

$$V_{\text{dip}}(r, \theta) = \frac{1}{4\pi\epsilon_0} \frac{\hat{\mathbf{r}} \cdot \mathbf{p}}{r^2} = \frac{1}{4\pi\epsilon_0} \frac{p \cos \theta}{r^2}. \quad (4.50)$$

To get the field, we take the **negative gradient** of  $V$ :

$$\mathbf{E}_r = -\frac{\partial V}{\partial r} = \frac{1}{4\pi\epsilon_0} \frac{2p \cos \theta}{r^3}, \quad \mathbf{E}_\theta = -\frac{1}{r} \frac{\partial V}{\partial \theta} = \frac{1}{4\pi\epsilon_0} \frac{p \sin \theta}{r^3}, \quad \mathbf{E}_\phi = -\frac{1}{r \sin \theta} \frac{\partial V}{\partial \phi} = 0.$$

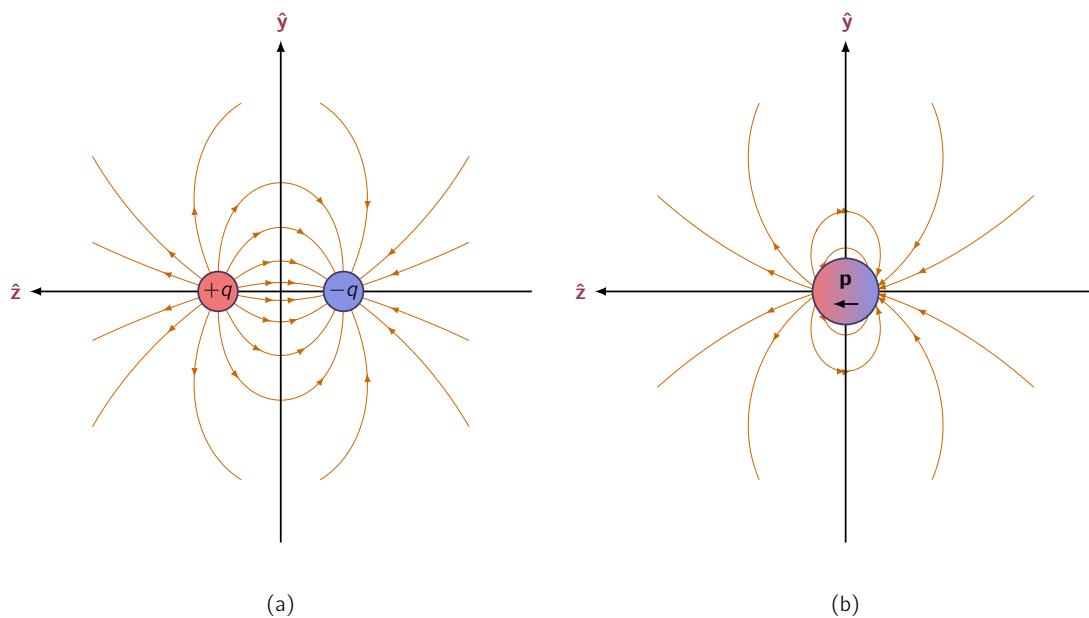
Which gives us the result we have been looking for:

$$\mathbf{E}_{\text{dip}}(r, \theta) = \frac{p}{4\pi\epsilon_0 r^3} (2 \cos \theta \hat{\mathbf{r}} + \sin \theta \hat{\theta}). \quad (4.51)$$

This formula makes explicit reference to a particular coordinate system<sup>44</sup> and assumes a particular orientation for  $\hat{\mathbf{p}}$  which is along  $\hat{\mathbf{z}}$  axis. It can be recast in a coordinate-free form, analogous to the potential in Eq. (4.47).

<sup>44</sup>In which case it is spherical

**Figure 4.11**  
A visual comparison between (a) physical description of a dipole and (b) a **pure** dipole.



Notice that the dipole field falls off as the inverse **cube** of  $r$ ; the monopole field ( $Q/4\pi\epsilon_0 r^2$ ) goes as the inverse square, of course. Quadrupole fields go like  $1/r^4$ , octopole like  $1/r^5$ , and so on. (This merely reflects the fact that monopole **potentials** fall off like  $1/r$ , dipole like  $1/r^2$ , quadrupole like  $1/r^3$ , and so on—the gradient introduces another factor of  $1/r$ .) Figure Fig. 4.11a shows the field lines of a "pure" dipole as described in Eq. (4.51).

For comparison, let's also compare the pure representation with the field lines for a "physical" dipole, also shown in Fig. 4.11b.

Please observe, how similar the two (2) pictures become if we blot out the central region; up close, however, they are entirely different. Only for points  $r \gg d$  does Eq. 3.103 represent a valid approximation to the field of a physical dipole.

As discussed earlier, this behaviour can be reached either by going to large  $r$  or by squeezing the charges very close together.<sup>45</sup>

<sup>45</sup>Even in the limit, there remains an infinitesimal region at the origin where the field of a physical dipole points in the "wrong" direction.



## Part III

# Magnetic Fields

### Part Contents

---

The experimental investigation by which Ampère established the law of the mechanical action between electric currents is one of the most brilliant achievements in science.

The whole, theory and experiment, seems as if it had leaped, full grown and full armed, from the brain of the "Newton of electricity". It is perfect in form, and unassailable in accuracy, and it is summed up in a formula from which all the phenomena may be deduced, and which must always remain the cardinal formula of electro-dynamics.

---

*(James Clerk Maxwell in A Treatise on Electricity and Magnetism (1873) quoted from 3rd edition (1892) Vol. 2, Ch. 3, p. 175.)*





## Part IV

# Electromagnetic Fields

## Part Contents

---

Glossary . . . . .	111
--------------------	-----

This velocity is so nearly that of light, that it seems we have strong reason to conclude that light itself is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws

---

(James Clerk Maxwell, in *A Dynamical Theory of the Electromagnetic Field* (1864)  
Introduction, p. 466.)



# Glossary

---

L

**LHS**      Left Hand Side

---

O

**ODE**      Ordinary Differential Equation

---

P

**PDE**      Partial Differential Equation

