Topics on Fundamental Science

D. T. McGuiness, Ph.D

M.Sc Electrodynamics

Lecture Book

2025.WS

Electrodynamics
Lecture Book

D. T. McGuiness, Ph.D
MCI

Current version is 2025.WS.

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 LuaTEX v1.22.0, and all editing were done using GNU Emacs v30.1 using AUCTEX 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. McGuines, Ph.D. (dtm@mci4me.at).

Table of Contents

List of Figures

List of Tables

List of Examples

List of Theorems

Part I

Prologue

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

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extremely small, sizes comparable to that of atoms, then the theory of relativity fails for different reasons, and is superseded by quantum mechanics.[2]

Finally, for objects which are both very fast and very small,[3] a mechanics which combines both relativity and quantum principles is needed:

> 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 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.[4]

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

[1] as in speeds comparable to the speed of light.

[2] Developed by Bohr, Schrodinger, Heisenberg, and many others, in the 1920's, mostly.

[3] As is common in modern particle physics.

[4] which can be extended to other domains if needed.

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

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

[1] Elementary particle responsible for making **protons** and **neutron**.
[2] Elementary particle acting as exchange particle for the **strong force** between **quarks**.
[3] Elementary particle affected by the **weak force** but not by the **strong force**.
[4] Elementary particle that is a **quantum** of the **electromagnetic** field, such as light and radio waves.
[5] 1 femtometre ($1\,\text{fm} = 1 \times 10^{-15}\,\text{m}$). i.e., the gold nucleus radius is approx. 8,45 fm.
[6] Known as the **fine structure constant.**

1.1.1 Four Kinds of Forces

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[5] that it is only in scale of huge mass concentrations, for example the earth and the sun, that we ever notice it at all.

[5] 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 1×10^{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.

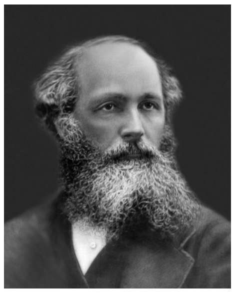
There is, of course, a classical theory of gravity and a relativistic one, but no entirely satisfactory quantum mechanical theory of gravity has been constructed. At the present time there is a very successful theory for the weak interactions, and a strikingly attractive candidate[6] for the strong interactions.

[6]called chromodynamics.

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.[7]



[7]**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.

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, Ampére 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.

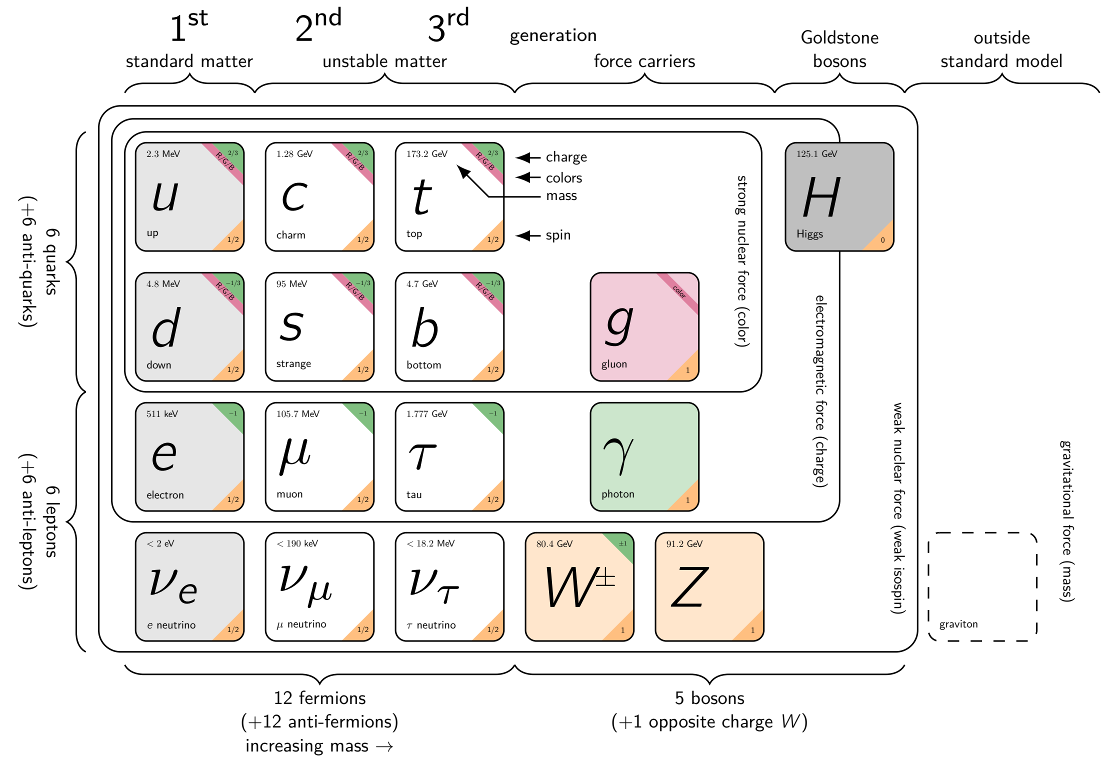


Figure 1.1: The Standard Model of particle physics is the theory describing three of the four known fundamental forces (electromagnetic, weak and strong interactions - excluding gravity) in the universe and classifying all known elementary particles. It was developed in stages throughout the latter half 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.

> 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.[8]

[8] 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.

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.

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[9] us to regard the fields as independent dynamical entities in their own right, every bit as real as atoms or baseballs.

[9] if not compels

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.[10] 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.

[10] 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. 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.[11]

[11] It is never $7.392e$, or $1/2e$.

Chapter 2

Vector Calculus

Table of Contents

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 we're not 13 km from where we set out, which is only 7. We need a set of mathematics principles to describe quantities like this, 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** ($\boldsymbol{A}$, $\boldsymbol{B}$, and so on) for vectors and ordinary type for scalars. The magnitude of a vector $\boldsymbol{A}$ is written $|\boldsymbol{A}|$ 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 $\boldsymbol{A}$ $(-\boldsymbol{A})$ is a vector with the same magnitude as $\boldsymbol{A}$ 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 $\boldsymbol{B}$ at the head of $\boldsymbol{A}$. Their sum, $\boldsymbol{A}+\boldsymbol{B}$, is the vector from the tail of $\boldsymbol{A}$ to the head of $\boldsymbol{B}$. Addition is *commutative*:

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

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*:

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

To subtract a vector, add its opposite

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

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*:

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

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

$$\boldsymbol{A}\cdot\boldsymbol{B}\equiv AB\cos\theta \tag{2.1}$$

where θ is the angle they form when placed tail-to-tail.

> $\boldsymbol{A}\cdot\boldsymbol{B}$ is itself a scalar.[1]

[1] which is why its alternative name is **scalar product**.

The dot product is *commutative*,

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

and *distributive*,

$$\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{A}\cdot\boldsymbol{B}+\boldsymbol{A}\cdot\boldsymbol{C}, \tag{2.2}$$

Geometrically, $\boldsymbol{A} \cdot \boldsymbol{B}$ is the product of A times the projection of $\boldsymbol{B}$ along $\boldsymbol{A}$.[2] If the two vectors are parallel, then $\boldsymbol{A} \cdot \boldsymbol{B} = AB$. In particular, for any vector $\boldsymbol{A}$,

[2]Or the product of B times the projection of $\boldsymbol{A}$ along $\boldsymbol{B}$.

$$\boldsymbol{A} \cdot \boldsymbol{A} = A^2 \tag{2.3}$$

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

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

$$\boldsymbol{A} \times \boldsymbol{B} \equiv AB \sin\theta\, \hat{\boldsymbol{n}} \tag{2.4}$$

where $\hat{\boldsymbol{n}}$ is a **unit vector** (vector of magnitude 1) pointing perpendicular to the plane of $\boldsymbol{A}$ and $\boldsymbol{B}$. 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 $\hat{\boldsymbol{n}}$.

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

The cross product is *distributive*,

$$\boldsymbol{A} \times (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} \times \boldsymbol{B}) + (\boldsymbol{A} \times \boldsymbol{C}) \tag{2.5}$$

but **NOT** commutative:

$$(\boldsymbol{B} \times \boldsymbol{A}) = -(\boldsymbol{A} \times \boldsymbol{B}) \tag{2.6}$$

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

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

for any vector $\boldsymbol{A}$.

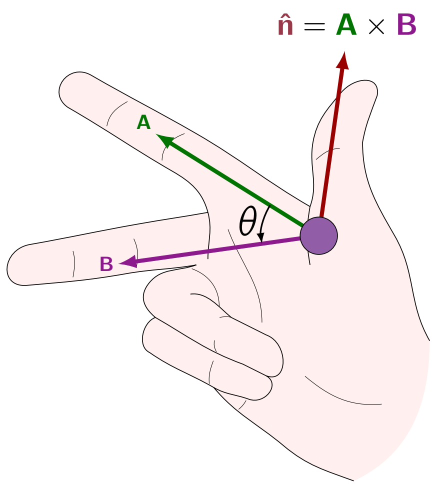


Figure 2.1: A mnemonic, used to define the orientation of axes in three-dimensional space and to determine the direction of the cross product of two vectors, as well as to establish the direction of the force on a current-carrying conductor in a magnetic field.

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 $\boldsymbol{A}$ can be expanded in terms of these **basis vectors**:

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

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

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

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

The operation rules are as follows:

i. To add vectors, add like components.

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

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

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

Accordingly,

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

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

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

Similarly the following relations can be derived:

$$\begin{aligned}\hat{x} \times \hat{x} = \hat{y} \times \hat{y} = \hat{z} \times \hat{z} = 0,\\ \hat{x} \times \hat{y} = -\hat{y} \times \hat{x} = \hat{z},\\ \hat{y} \times \hat{z} = -\hat{z} \times \hat{y} = \hat{x},\\ \hat{z} \times \hat{x} = -\hat{x} \times \hat{z} = \hat{y}.\end{aligned}$$

Therefore,

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

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

$$\boldsymbol{A} \times \boldsymbol{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \tag{2.7}$$

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

2.1.3 Triple Products

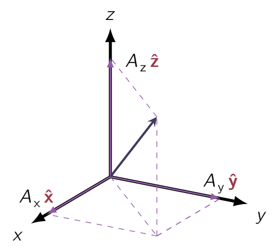


Figure 2.2: The decomposition of the vector $\boldsymbol{A}$ into its components.

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

Scalar Triple Product Evidently,

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

for they all correspond to the same value.

> The **alphabetical** order is preserved.

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

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

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

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

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

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

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

> **Information**: Cross Product of Two Scalars
>
> $(\boldsymbol{A} \cdot \boldsymbol{B}) \times \boldsymbol{C}$ 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:

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

Please observe that the following triple product

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

is an entirely different vector.[3] 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,

$$(\boldsymbol{A} \times \boldsymbol{B}) \cdot (\boldsymbol{C} \times \boldsymbol{D}) = (\boldsymbol{A} \cdot \boldsymbol{C})\,(\boldsymbol{B} \cdot \boldsymbol{D}) - (\boldsymbol{A} \cdot \boldsymbol{D})\,(\boldsymbol{B} \cdot \boldsymbol{C})\,,$$
$$\boldsymbol{A} \times \big[\boldsymbol{B} \times (\boldsymbol{C} \times \boldsymbol{D})\big] = \boldsymbol{B}\,\big[\boldsymbol{A} \cdot (\boldsymbol{C} \times \boldsymbol{D})\big] - (\boldsymbol{A} \cdot \boldsymbol{B})\,(\boldsymbol{C} \times \boldsymbol{D})\,.$$

[3] To reiterate cross-products are not associative.

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**:

$$\boldsymbol{r} \equiv x\,\hat{x} + y\,\hat{y} + z\,\hat{z}$$

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

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

is the distance from the origin, and

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

is a unit vector pointing radially outward. The **infinitesimal displacement vector**, from (x, y, z) to $(x + \mathrm{d}x, y + \mathrm{d}y, z + \mathrm{d}z)$, is

$$\mathrm{d}\boldsymbol{l} = (\,\mathrm{d}x)\ \hat{x} + (\,\mathrm{d}y)\ \hat{y} + (\,\mathrm{d}z)\ \hat{z}.$$

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

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

field point ($\boldsymbol{r}$), 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:

$$\boldsymbol{\mathscr{r}} \equiv \boldsymbol{r} - \boldsymbol{r}'$$

Its magnitude is:

$$\mathscr{r} = \left|\boldsymbol{r} - \boldsymbol{r}'\right|$$

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

$$\hat{\boldsymbol{\mathscr{r}}} = \frac{\hat{\boldsymbol{\mathscr{r}}}}{\mathscr{r}} = \frac{\boldsymbol{r} - \boldsymbol{r}'}{\left|\boldsymbol{r} - \boldsymbol{r}'\right|}$$

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

$$\boldsymbol{\mathscr{r}} = \left(x - x'\right)\ \hat{x} + \left(y - y'\right)\ \hat{y} + \left(z - z'\right)\ \hat{z}$$
$$\mathscr{r} = \sqrt{\left(x - x'\right)^2 + \left(y - y'\right)^2 + \left(z - z'\right)^2}$$
$$\hat{\boldsymbol{\mathscr{r}}} = \frac{\left(x - x'\right)\ \hat{x} + \left(y - y'\right)\ \hat{y} + \left(z - z'\right)\ \hat{z}}{\sqrt{\left(x - x'\right)^2 + \left(y - y'\right)^2 + \left(z - z'\right)^2}}$$

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?

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

$$\mathrm{d}f = \left(\frac{df}{dt}\right) \mathrm{d}x$$

If we increment x by an infinitesimal amount $\mathrm{d}x$, then f changes by an amount $\mathrm{d}f$.[4]

[4] Here we can think the derivative as the proportionality factor.

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

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.

In fact, the question of "How fast does T vary?" can have an infinite number of answers, one for

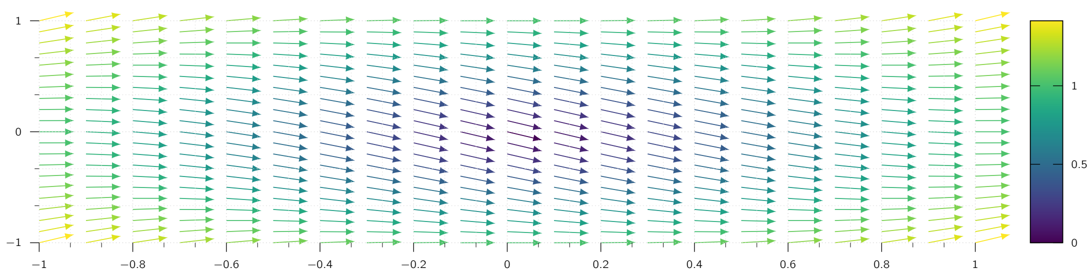


Figure 2.3: An example of a gradient field. Here the field itself is plotted by using arrows to designate the direction of the gradient. To explain the magnitude of the gradient it is either shown by the length of the arrow or by imposing colour on to the plot, which the latter has been used here. Extending 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.

each direction we might choose to explore. Fortunately, the problem is not as bad as it looks. A theorem on partial derivatives states:

$$\mathrm{d}T = \left(\frac{\partial T}{\partial x}\right)\,\mathrm{d}x + \left(\frac{\partial T}{\partial y}\right)\,\mathrm{d}y + \left(\frac{\partial T}{\partial z}\right)\,\mathrm{d}z$$

This tells us how T changes when we alter all three (3) variables by the infinitesimal amounts of $\mathrm{d}x$, $\mathrm{d}y$, $\mathrm{d}z$. We can write the aforementioned equation as a dot product:

$$\mathrm{d}T = \left(\frac{dT}{dx}\,\hat{x} + \frac{dT}{dy}\,\hat{y} + \frac{dT}{dz}\,\hat{z}\right)\cdot\big((\mathrm{d}x)\,\hat{x} + (\mathrm{d}y)\,\hat{y} + (\mathrm{d}z)\,\hat{z}\big) = (\nabla T)\cdot(\mathrm{d}\boldsymbol{l}),$$

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 ∇T is a **vector quantity**, with three (3) components.[5]

[5] This is the generalised derivative we have been looking for.

Exercise 2.1 Finding Vector Components - I

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

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

SOLUTION The solution is as follows:

$$\begin{aligned}
\boldsymbol{v} &= (5-3)\,\hat{x} + (-2-2)\,\hat{y} + (2-0)\,\hat{z} = (2)\,\hat{x} + (-4)\,\hat{y} + (2)\,\hat{z}, \\
|\boldsymbol{v}| &= \sqrt{(2)^2 + (-4)^2 + (2)^2} = 2\sqrt{6}. \\
\hat{\boldsymbol{v}} &= \frac{\boldsymbol{v}}{|\boldsymbol{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} \quad \blacksquare \\
\boldsymbol{v} &= (-4-1)\,\hat{x} + (-4-1)\,\hat{y} + (-4-1)\,\hat{z} = (-5)\,\hat{x} + (-5)\,\hat{y} + (-5)\,\hat{z}, \\
|\boldsymbol{v}| &= \sqrt{(-5)^2 + (-5)^2 + (-5)^2} = 5\sqrt{3}. \\
\hat{\boldsymbol{v}} &= \frac{\boldsymbol{v}}{|\boldsymbol{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} \quad \blacksquare \\
\boldsymbol{v} &= (0-1)\,\hat{x} + (0-0)\,\hat{y} + (6.2-1.2)\,\hat{z} = (-1)\,\hat{x} + (0)\,\hat{y} + (5)\,\hat{z}, \\
|\boldsymbol{v}| &= \sqrt{(-1)^2 + (0)^2 + (5)^2} = \sqrt{26}. \\
\hat{\boldsymbol{v}} &= \frac{\boldsymbol{v}}{|\boldsymbol{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} \quad \blacksquare \\
\boldsymbol{v} &= (0-2)\,\hat{x} + \big(4-(-2)\big)\,\hat{y} + (6-0)\,\hat{z} = (-2)\,\hat{x} + (6)\,\hat{y} + (6)\,\hat{z}, \\
|\boldsymbol{v}| &= \sqrt{(-2)^2 + (6)^2 + (6)^2} = 2\sqrt{19}. \\
\hat{\boldsymbol{v}} &= \frac{\boldsymbol{v}}{|\boldsymbol{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} \quad \blacksquare \\
\boldsymbol{v} &= (-4-4)\,\hat{x} + (-3-3)\,\hat{y} + (2-2)\,\hat{z} = (-8)\,\hat{x} + (-6)\,\hat{y} + (0)\,\hat{z}, \\
|\boldsymbol{v}| &= \sqrt{(-8)^2 + (-6)^2 + (0)^2} = 10. \\
\hat{\boldsymbol{v}} &= \frac{\boldsymbol{v}}{|\boldsymbol{v}|} = \frac{(-6)\,\hat{x} + (-8)\,\hat{y} + (0)\,\hat{z}}{10} = \left(-\frac{3}{5}\right)\hat{x} + \left(-\frac{4}{5}\right)\hat{y} + (0)\,\hat{z} \quad \blacksquare \\
\boldsymbol{v} &= (6-0)\,\hat{x} + (8-0)\,\hat{y} + (10-0)\,\hat{z} = (6)\,\hat{x} + (8)\,\hat{y} + (10)\,\hat{z},
\end{aligned}$$

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

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

Exercise 2.2 Finding Vector Components - II

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

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

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

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

2.2.3 The Del Operator

The gradient has the formal appearance of a vector, (∇),[6] multiplying a scalar T:

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

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

$$\nabla = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{y}\frac{\partial}{\partial z} \tag{2.8}$$

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

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

[6]The History of Nabla

The operator (∇), 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 $(\triangleleft)$, 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.

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

With this qualification, though, ∇ 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 ∇. Now, an ordinary vector $\boldsymbol{A}$ can multiply in three (3) ways:

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

Correspondingly, there are three ways the operator ∇ can act:

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

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

Divergence From the definition of ∇ we construct the divergence as follows:

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

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

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

[7] i.e., diverges

For example, a vector function in **Fig.** 2.4 has a large (positive) divergence,[8] two functions in **Fig.** 2.4

[8] if the arrows pointed in, it would be a negative divergence

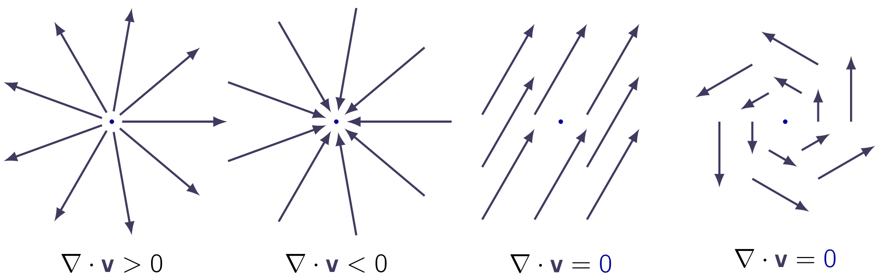


Figure 2.4: Visual description of the divergence operation.

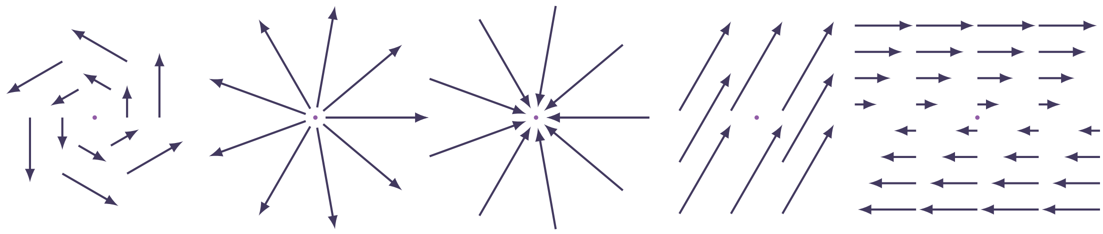


Figure 2.5: Different behaviours of the curl operation.

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 ∇ we construct the curl:

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

As with divergence, the name curl is also well chosen, for $\nabla \times \boldsymbol{v}$ is a measure of how much $\boldsymbol{v}$ 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 $\hat{z}$ 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 he curl ($\nabla \times$) of the following functions.

$$\boldsymbol{v} = (y)\ \hat{x} + \left(2x^2\right)\ \hat{y} + (0)\ \hat{z}, \qquad \boldsymbol{v} = \left(y^n\right)\ \hat{x} + \left(z^n\right)\ \hat{y} + \left(x^n\right)\ \hat{z},$$
$$\boldsymbol{v} = (\sin y)\ \hat{x} + (\cos z)\ \hat{y} + (-\tan x)\ \hat{z}, \qquad \boldsymbol{v} = \left(x^2 - z\right)\ \hat{x} + \left(xe^z\right)\ \hat{y} + (xy)\ \hat{z}.$$

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

$$f(x, y, z) = (y)\ \hat{x} + \left(2x^2\right)\ \hat{y} + (0)\ \hat{z},$$
$$\nabla \times f = (0)\ \hat{x} + (0)\ \hat{y} + (-1 + 4x)\ \hat{z}.$$

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

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,

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

and

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

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:

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

and two ways to make a vector:

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

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

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

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

It is also possible to formulate three quotient rules:

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

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 $\nabla\cdot\boldsymbol{v}$ by applying ∇ twice, we can construct five (5) types of 2$^{\text{nd}}$ derivatives.

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

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

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

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

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

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

This object, which we write as $\nabla^2 T$ for short, is called the Laplacian[9] 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, $\nabla^2 \boldsymbol{v}$. By this we mean a **vector** quantity whose x-component is the Laplacian of v_x, and so on.

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

This is nothing more than a convenient extension of the meaning of ∇^2.

Exercise 2.4 The Laplacian of a Vector

Calculate the Laplacian of the following functions:

$$(i)\quad T_a = x_2 + 3xy + 3z + 4, \qquad (ii)\quad T_b = \sin x \sin y \sin z,$$
$$(iii)\quad T_c = e^{-5x}\sin 4y \cos 3z, \qquad (iv)\quad \boldsymbol{v} = \left(x^2\right)\,\hat{x} + \left(3xz^2\right)\,\hat{y} + (-2xz)\,\hat{z}.$$

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

(i) $\dfrac{\partial^2 T_a}{\partial x^2} = 2;\ \dfrac{\partial^2 T_a}{\partial y^2} = 0;\ \dfrac{\partial^2 T_a}{\partial z^2} = 0 \quad \rightarrow \quad \nabla^2 T_a = 2 \quad \blacksquare$

(ii) $\dfrac{\partial^2 T_b}{\partial x^2} = \dfrac{\partial^2 T_b}{\partial y^2} = \dfrac{\partial^2 T_b}{\partial z^2} = -3T_b \quad \rightarrow \quad \nabla^2 T_b = -3T_b = 3\sin x \sin y \sin z \quad \blacksquare$

(iii) $\dfrac{\partial^2 T_c}{\partial x^2} = 25T_c;$

$\dfrac{\partial^2 T_c}{\partial y^2} = -16T_c;\quad \dfrac{\partial^2 T_c}{\partial z^2} = -9T_c \quad \rightarrow \quad \nabla^2 T_c = 0 \quad \blacksquare$

(iii) $\dfrac{\partial^2 v_x}{\partial x^2} = 2;\ \dfrac{\partial^2 v_x}{\partial y^2} = 0;\ \dfrac{\partial^2 v_x}{\partial z^2} = 0 \quad \rightarrow \quad \nabla^2 v_x = 0,$

[9] A differential operator given by the divergence of the gradient of a scalar function on Euclidean space. It is usually denoted by the symbols $\nabla \cdot \nabla$, ∇^2 or Δ. 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 (17491827)*, 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.

$$\frac{\partial^2 v_y}{\partial x^2} = 0; \ \frac{\partial^2 v_y}{\partial y^2} = 0; \ \frac{\partial^2 v_y}{\partial z^2} = 6 \quad \rightarrow \quad \nabla^2 v_y = 6x,$$
$$\frac{\partial^2 v_z}{\partial x^2} = 0; \ \frac{\partial^2 v_z}{\partial y^2} = 0; \ \frac{\partial^2 v_z}{\partial z^2} = 0 \quad \rightarrow \quad \nabla^2 v_z = 0,$$
$$\nabla^2 \boldsymbol{v} = 2\,\hat{x} + 6x\,\hat{y} \qquad \blacksquare$$

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 $\nabla(\nabla \cdot \boldsymbol{v})$ is **NOT** the same as the Laplacian of a vector:

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

Divergence of a Curl

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

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

Curl of a Curl

As we can check from the definition of ∇:

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

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:[10]

1. Laplacian,

2. Gradient of divergence.

[10] It is possible to work out 3rd 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:

$$\int_a^b \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{l}$$

where $\boldsymbol{v}$ is a vector function, $d\boldsymbol{l}$ 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,[11]

[11] i.e., if $\boldsymbol{b} = \boldsymbol{a}$.

We put a circle on the integral sign:

$$\oint \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{l}$$

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

> A good example of a line integral is the work done by a force $\boldsymbol{F}$:
>
> $$W = \int \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{l}$$

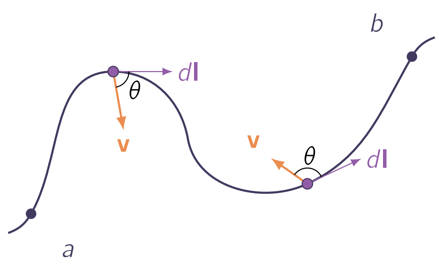


Figure 2.6: The method in which line integral is calculated. At each point the dot product of the vector ($\boldsymbol{v}$) is taken with the length vector (d$\boldsymbol{l}$) which is always tangential to the point in which the integration is taken.

Ordinarily, the value of a line integral depends critically on the path taken from $\boldsymbol{a}$ to $\boldsymbol{b}$, 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 $\boldsymbol{F} = (x)\ \hat{x} + (z)\ \hat{y} + (y)\ \hat{z}$. Find the flow along the helix $\boldsymbol{\ell}(t) = (\cos t)\ \hat{x} + (\sin t)\ \hat{y} + (t)\ \hat{z}$ with a range of $0 \leq t \leq \pi/2$.

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

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

and then find $d\boldsymbol{\ell}/dt$:

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

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

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

Which makes,

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

Exercise 2.6 Field Circulation

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

SOLUTION On the circle, $\boldsymbol{F} = (x - y)\,\hat{x} + (x)\,\hat{y} + (0)\,\hat{z} = (\cos t - \sin t)\,\hat{x} + (\cos t)\,\hat{y} + (0)\,\hat{z}$ and

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

Then

$$\boldsymbol{F} \cdot \frac{d\boldsymbol{\ell}}{dt} = -\sin t\cos t + \underbrace{\sin^2 t + \cos^2 t}_{1},$$

Gives.

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

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

$$\int_{\mathcal{S}} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{a}$$

where $\boldsymbol{v}$ is a vector function, and the integral is over a specified surface $\mathcal{S}$. Here $\mathrm{d}\boldsymbol{a}$ 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.

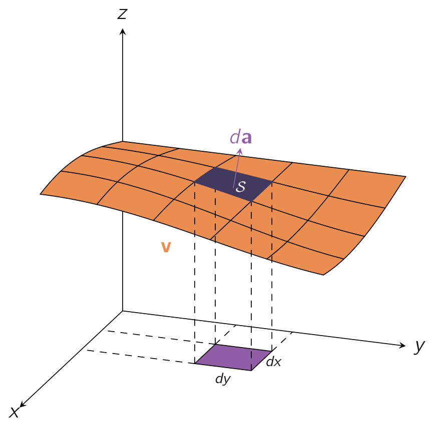


Figure 2.7: An visual description of the surface integral.

If the surface is closed,[12] we put a circle on the integral sign:

$$\oint \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{a}$$

[12] imagine it forming a balloon.

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

As an example, if $\boldsymbol{v}$ describes the flow of a fluid,[13] then $\int \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{a}$ represents the total mass per unit time passing through the surface.

[13] 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:

$$\int_{\mathcal{V}} T \, \mathrm{d}\tau$$

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

$$\mathrm{d}\tau = \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$$

As an example, if T is the density of a substance,[14] then the volume integral would give the total mass.

[14] This might vary from point to point.

Occasionally we shall encounter volume integrals of vector functions:

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

> 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:

$$\int_0^1 \int_x^{2x} (x+y)^2 \, dy \, dx, \qquad \int_0^1 \int_y^{\sqrt{y}} (1-2xy) \, dx \, dy,$$

$$\int_0^3 \int_x^3 \cosh(x+y) \, dy \, dx, \qquad \int_0^1 \int_0^{y^3} \exp y^4 \, dx \, dy.$$

SOLUTION

The solution to integrations are as follows:

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

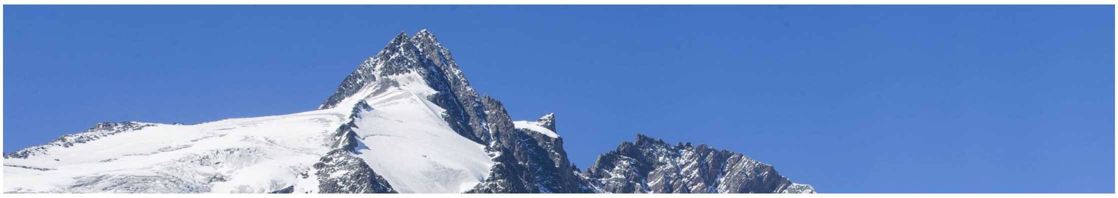


Figure 2.8: To measure the height of a mountain, it doesn't matter what way we take, as long as we know the base and the top, we will know the height.

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**)

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

In vector calculus there are three (3) species of derivative[15] 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.

[15] 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 **a**, move a small distance $\mathrm{d}\boldsymbol{l}_1$. The function T will change by an amount:

$$\mathrm{d}T = (\nabla T) \cdot \mathrm{d}\boldsymbol{l}_1$$

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

$$\mathrm{d}T = (\nabla T) \cdot \mathrm{d}\boldsymbol{l}_2$$

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

Theory 2.2: Gradient Theorem

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

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

Similar to "ordinary" fundamental theorem, this theorem says the integral[16] of a derivative, which here the gradient, is given by the value of the function at the boundaries which are $\boldsymbol{a}$ and $\boldsymbol{b}$ respectively.

[16] 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 or the options we take, we should get the same answer either way.[17]

[17] 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:

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

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).

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

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 $\boldsymbol{v}$ represents the flow of an incompressible fluid, then the flux $\boldsymbol{v}$ 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 $\boldsymbol{F} = (x)\ \hat{x} + (y)\ \hat{y} + (z)\ \hat{z}$ over the sphere $x^2 + y^2 + z^2 = a^2$

SOLUTION The outer unit normal to S, calculated from the gradient of $f\{\cdot\}\ x,\ y,\ z = x^2 + y^2 + z^2 - a^2$, is:

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

Therefore:

$$(\boldsymbol{F} \cdot \hat{n})\ \mathrm{d}a = \frac{x^2 + y^2 + z^2}{a}\ \mathrm{d}a = \frac{a^2}{a}\ \mathrm{d}a = a\,\mathrm{d}a.$$

This in turn gives us:

$$\iint_S (\boldsymbol{F} \cdot \hat{n})\ \mathrm{d}a = \iint_S a\,\mathrm{d}a = a \iint_S \mathrm{d}a = a\left(4\pi a^2\right) = 4\pi a^3. \qquad \text{Area of S is 4ßa}^2$$

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

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

So,

$$\iiint_V (\nabla \cdot \boldsymbol{v})\ d\tau = \iiint_V 3\ d\tau = 3\left(\frac{4}{3}\pi a^3\right) = 4\pi a^3 \quad \blacksquare$$

Exercise 2.9 Divergence Theorem of an Octant of a Sphere

Check the divergence theorem for the function:

$$\boldsymbol{v} = \left(r^2 \cos\theta\right)\ \hat{r} + \left(r^2 \cos\phi\right)\ \hat{\theta} + \left(-r^2 \cos\theta \sin\phi\right)\ \hat{\phi}.$$

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:

$$\underbrace{\iiint_{\mathcal{V}} (\nabla \cdot \boldsymbol{v})\ dV}_{\text{Divergence integral}} = \underbrace{\iint_{\mathcal{S}} \boldsymbol{v} \cdot \boldsymbol{n}\,da}_{\text{Outward flux}}.$$

First solve the left hand side of the equation:

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

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

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 $xyz \to \boldsymbol{a}_1$, and $xz \to \boldsymbol{a}_2$, $yz \to \boldsymbol{a}_3$ and $xy \to \boldsymbol{a}_4$. These are

$$\begin{aligned}
d\boldsymbol{a}_1 &= \hat{r}\, dl_\theta\, dl_\phi = \hat{r}\, R^2 \sin\theta\, d\phi\, d\theta, & d\boldsymbol{a}_2 &= dl_r\, dl_\theta = -\hat{\phi}\, r\, dr\, d\theta, \\
d\boldsymbol{a}_3 &= \hat{\phi}\, dl_r\, dl_\theta = \hat{\phi}\, r\, dr\, d\theta, & d\boldsymbol{a}_4 &= dl_r\, dl_\phi = \hat{\theta}\, r\, dr\, d\theta. \quad \left(\theta = \pi/2\right)
\end{aligned}$$

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

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

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

Time to do some integration.

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

Final touches and cleaning up,

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

2.3.3 The Fundamental Theorem for Curls

The fundamental theorem for curls, also known as **Stokes' theorem**, states:[18]

Theory 2.5: Stokes' Theorem

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

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

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 $\boldsymbol{v}$. 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.

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

[18]**Sir George Gabriel Stokes, 1st Baronet,** *(1819 - 1903)*

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 NavierStokes 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?[19]

[19] 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 $\mathrm{d}\boldsymbol{a}$ point?

For a closed surface,[20] $\mathrm{d}\boldsymbol{a}$ 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 $\mathrm{d}\boldsymbol{a}$.

[20] 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 $\int(\nabla \times \boldsymbol{v}) \cdot d\boldsymbol{a}$ is equal to the line integral of $\boldsymbol{v}$ around the boundary, and the latter makes no reference to the specific surface we choose.

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

Preposition II $\oint(\nabla \times \boldsymbol{v}) \cdot d\boldsymbol{a} = 0$ 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 $z = 4$.

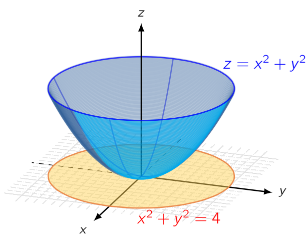


Figure 2.9: The paraboloid of "Surface Area of an Implicit Surface".

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 $F(x, y, z) = x^2 + y^2 - z = 0$, ans R is the disk $x^2 + y^2 \leq 4$ in the xy-plane.

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

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

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

In the region R, the area is defined to be $\mathrm{d}A = \mathrm{d}x\,\mathrm{d}y$. Therefore:

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

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

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

SOLUTION The start by calculating the counter-clockwise circulation around C using the following parametrisation:

$$\boldsymbol{\ell}(\theta) = (3\cos\theta)\ \hat{x} + (3\sin\theta)\ \hat{y} + (0)\ \hat{z}, \quad \text{where} \quad 0 \leq \theta \leq 2\pi.$$

Using this we can calculate the counter-clockwise circulation.

$$\begin{aligned} d\boldsymbol{\ell} &= (-3\sin\theta\, d\theta)\ \hat{x} + (3\cos\theta\, d\theta)\ \hat{y} + (0)\ \hat{z}, \\ \boldsymbol{F} &= (y)\ \hat{x} + (-x)\ \hat{y} + (0)\ \hat{z} \\ &= (3\sin\theta)\ \hat{x} + (-3\cos\theta)\ \hat{y} + (0)\ \hat{z}, \\ \boldsymbol{F} \cdot d\boldsymbol{\ell} &= -9\sin^2\theta\, d\theta - 9\cos^2\theta\, d\theta = -9\, d\theta, \end{aligned}$$

$$\oint_C \boldsymbol{F} \cdot d\boldsymbol{\ell} = \int_0^{2\pi} -9\, d\theta = -18\pi.$$

For the curl of integral we have:

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

$$= (0-0)\ \hat{x} + (0-0)\ \hat{y} + (-1-1)\ \hat{z} = -2\,\hat{z}$$

$$\hat{n} = \frac{\nabla S}{|\nabla S|} = \frac{(x)\ \hat{x} + (y)\ \hat{y} + (z)\ \hat{z}}{\sqrt{x^2 + y^2 + z^2}} = \frac{(x)\ \hat{x} + (y)\ \hat{y} + (z)\ \hat{z}}{3} \qquad \text{Unit normal}$$

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

$$\begin{aligned} da &= \frac{|\nabla S|}{|\nabla S \cdot \hat{z}|}\, dA \\ &= \frac{\left|(2x)\ \hat{x} + (2y)\ \hat{y} + (2z)\ \hat{z}\right|}{2z} \\ &= \frac{2\overbrace{\sqrt{x^2 + y^2 + z^2}}^{a}}{2z} \\ &= \frac{3}{z}\, dA, \end{aligned}$$

$$\nabla \times \boldsymbol{F} \cdot \boldsymbol{n}\, da = -\frac{2z}{3}\frac{3}{z}\, dA = -2\, dA$$

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

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

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

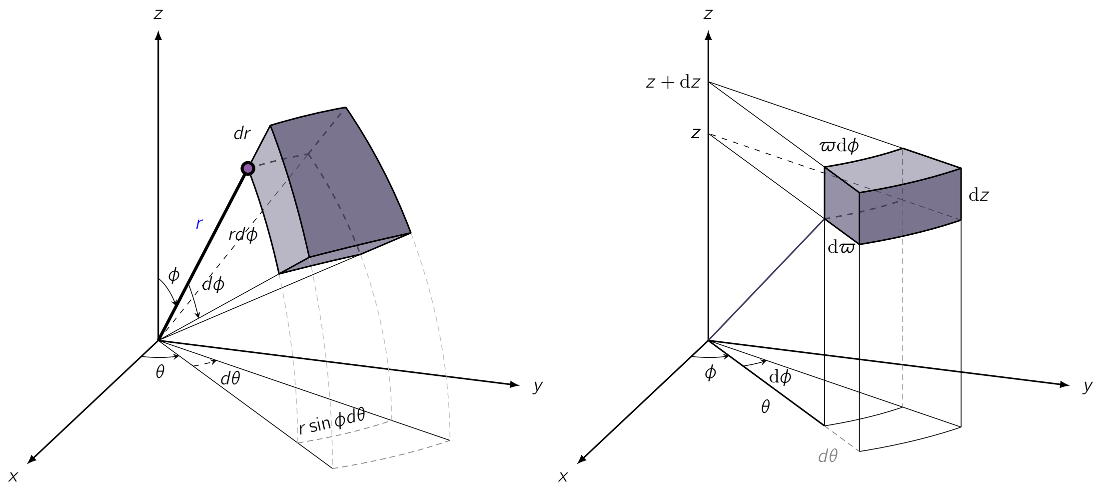


Figure 2.10: The two types of coordinate systems in question. (a) Spherical coordinate system (b) Spherical coordinate system.

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, θ, ϕ); r is the distance from the origin, θ is called the **polar angle**, and ϕ is the **azimuthal angle**. Their relation to Cartesian coordinates can be read from **Fig.** 2.10a.

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

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,[21] similar to $\hat{x}$, $\hat{y}$, $\hat{z}$, and any vector $\boldsymbol{A}$ can be expressed in terms of them, in the usual way:

[21] meaning mutually perpendicular.

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

Here, A_r, A_θ, A_ϕ are the radial, polar, and azimuthal components of vector $\boldsymbol{A}$. In terms of the Cartesian unit vectors:

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

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

$$\mathrm{d}l_r = \mathrm{d}r$$

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

$$\mathrm{d}l_\theta = r\,\mathrm{d}\theta$$

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

$$\mathrm{d}l_\phi = r \sin\theta \, \mathrm{d}\phi$$

Thus the general infinitesimal displacement $\mathrm{d}\boldsymbol{l}$ is:

$$\mathrm{d}\boldsymbol{l} = (\mathrm{d}r)\ \hat{r} + (r\,\mathrm{d}\theta)\ \hat{\theta} + (r\sin\theta)\ \hat{\phi}$$

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

$$\mathrm{d}\tau = \mathrm{d}l_r\, \mathrm{d}l_\theta\, \mathrm{d}l_\phi = r^2 \sin\theta \, \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{d}\phi.$$

> It is not possible to give a general expression for surface elements $\mathrm{d}\boldsymbol{a}$, 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 θ and ϕ change:

$$\mathrm{d}\boldsymbol{a}_1 = \mathrm{d}l_\theta\, \mathrm{d}l_\phi\, \hat{r} = r^2 \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi\, \hat{r}$$

On the other hand, if the surface lies in the xy plane, making θ is constant, while r and ϕ vary:

$$\mathrm{d}\boldsymbol{a}_2 = \mathrm{d}l_r\, \mathrm{d}l_\phi\, \hat{\theta} = r\, \mathrm{d}r\, \mathrm{d}\phi\, \hat{\theta}$$

Finally: r ranges from 0 to ∞, ϕ from 0 to 2π, and θ from 0 to π.

Exercise 2.12 **The Volume of a Sphere**

Find the volume of a sphere of radius R.

SOLUTION The derivation is as follows:

$$V = \int \mathrm{d}\tau$$

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

2.4.2 Cylindrical Coordinates

The cylindrical coordinates (s, ϕ, z) of a point P are defined in **Fig.** 2.10b. Observe that ϕ 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:

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

The unit vectors are:

$$\hat{\boldsymbol{s}} = \cos\phi\,\hat{\boldsymbol{x}} + \sin\phi\,\hat{\boldsymbol{y}}$$
$$\hat{\boldsymbol{\phi}} = -\sin\phi\,\hat{\boldsymbol{x}} + \cos\phi\,\hat{\boldsymbol{y}}$$
$$\hat{\boldsymbol{z}} = \hat{\boldsymbol{z}}$$

The infinitesimal displacements are

$$\mathrm{d}l_{\mathrm{s}} = \mathrm{d}s \qquad \mathrm{d}l_{\phi} = s\,\mathrm{d}\phi, \qquad \mathrm{d}l_{\mathrm{z}} = \mathrm{d}z$$

which makes:

$$\mathrm{d}\boldsymbol{l} = \mathrm{d}s\,\hat{\boldsymbol{s}} + s\,\mathrm{d}\phi\,\hat{\boldsymbol{phi}} + \mathrm{d}z\,\hat{\boldsymbol{z}}.$$

and the volume element is

$$\mathrm{d}\tau' = s\,\mathrm{d}s\,\mathrm{d}\phi\,\mathrm{d}z$$

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

2.5 Dirac Delta Function

2.5.1 A Mathematical Anomaly

Consider the following vector function:

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

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

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

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

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

But the volume integral ($\int \nabla \cdot \boldsymbol{v} \,\mathrm{d}\tau$), 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 $\boldsymbol{v}$ **blows up**. It is true that $\nabla \cdot \boldsymbol{v} = 0$ 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

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

for any non-zero vector and the origin.

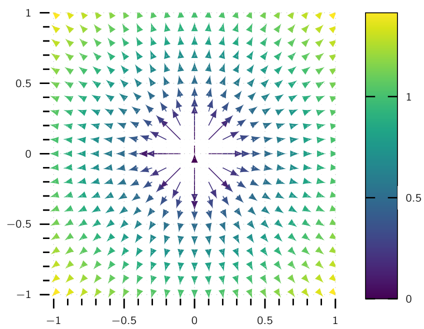


Figure 2.11: The vector plot of the "divergence problem".

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

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.

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.

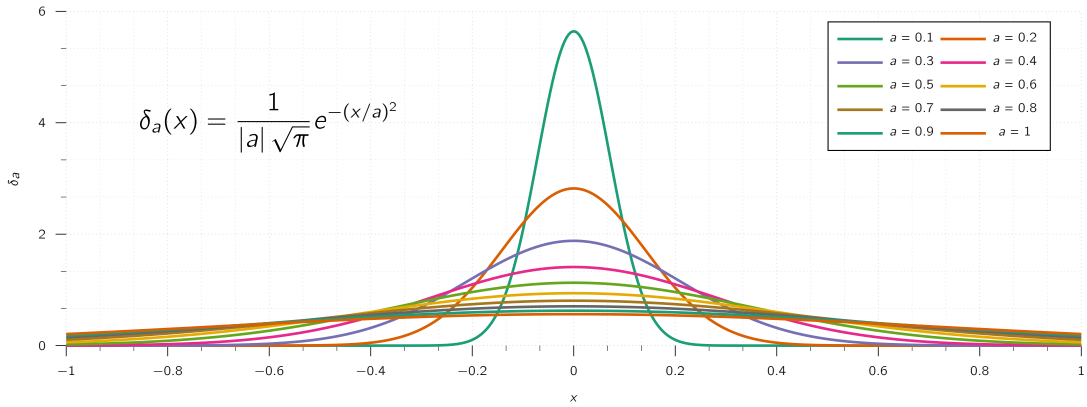


Figure 2.12: A visual representation of a 1D Dirac Delta Function. Think of it as a distribution function being squeezed to an infinitely small width.

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:

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

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.[22]

[22] 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.

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

$$f(x)\delta(x) = f(0)\delta(x). \tag{2.11}$$

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.

Particularly

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

Under an integral, then, the delta function "picks out" the value of $f(x)$ at $x = 0$.[23] Of course, we

[23] 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.

can shift the spike from $x = 0$ to some other point, $x = a$:

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

which turns Eq. (2.11) into:

$$f(x)\delta(x-a) = f(a)\delta(x-a), \tag{2.14}$$

and Eq. (2.12) generalises to:

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

While δ itself is not a proper function, integrals over δ 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:

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

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

Exercise 2.13 A Simple Dirac Integral

Evaluate the following integral:

$$\int_0^3 x^3\delta(x-2)\,\mathrm{d}x$$

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:

$$\int_2^6 \left(3x^2 - 2x - 1\right)\delta(x-3)\,dx, \tag{xvii}$$

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

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

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

SOLUTION The solution are as follows:

(a) $3\left(3^2\right) - 2(3) - 1 = 27 - 6 - 1 = 20$ ■

(b) $\cos\pi = -1$ ■

(c) 0 ■

(d) $\ln(-2+3) = \ln 1 = 0$ ■

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:

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

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

$$\int\limits_{\text{all space}} \delta^3(\boldsymbol{r}) d\tau = \iiint\limits_{-\infty}^{\infty} \delta(x)\delta(y)\delta(z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 1$$

And, the general form is:

$$\int\limits_{\text{all space}} f(\boldsymbol{r})\delta^3(\boldsymbol{r}-\boldsymbol{a})\,\mathrm{d}\tau = f(\boldsymbol{a}). \tag{2.22}$$

As in the 1D case, integration with δ 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 $\hat{r}/r^2$ 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

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

Or in a more general fashion:

$$\nabla \cdot \left(\frac{\hat{\boldsymbol{\mathscr{r}}}}{\mathscr{r}^2}\right) = 4\pi\delta^3(\hat{\boldsymbol{\mathscr{r}}}) \tag{2.23}$$

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

2.6 Vector Field Theory

2.6.1 Helmholtz Theorem

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

As $\boldsymbol{E}$ and $\boldsymbol{B}$ 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 $\boldsymbol{E}$ and $\boldsymbol{B}$.

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 $\boldsymbol{F}$. If the divergence of $\boldsymbol{F}$ is a specified scalar function D,

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

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

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

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

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

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

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

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

Some examples are:

$$\begin{aligned}
\boldsymbol{F} &= 0, \\
\boldsymbol{F} &= (y)\ \hat{x} + (zx)\ \hat{y} + (xy)\ \hat{z}, \\
\boldsymbol{F} &= (\sin x \cosh y)\ \hat{x} + (-\cos x \sinh y)\ \hat{y} + (.)\ \hat{z}
\end{aligned}$$

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**[24] guarantees the field is uniquely determined by its divergence and curl.

[24]**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 Bernoulliamong othersfor 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 ($\boldsymbol{F}$) vanishes everywhere, then $\boldsymbol{F}$ can be written as the **gradient of a scalar potential** (V):

$$\nabla \times \boldsymbol{F} = 0 \Longleftrightarrow \boldsymbol{F} = -\nabla V$$

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. $\nabla \times \boldsymbol{F} = 0$ everywhere,

ii. $\int_a^b \boldsymbol{F} \cdot d\boldsymbol{l}$ is independent of path, for any given end points,

iii. $\oint \boldsymbol{F} \cdot d\boldsymbol{l} = 0$ for any closed loop,

iv. $\boldsymbol{F}$ is the gradient of some scalar function: $\boldsymbol{F} = -\nabla V$.

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

$$\nabla \cdot \boldsymbol{F} = 0 \Longleftrightarrow \boldsymbol{F} = \nabla \times \boldsymbol{A}$$

That's the main conclusion of the following theorem:

Theory 2.7: Zero Divergence Fields

The following conditions are **equivalent**:

i. $\nabla \cdot \boldsymbol{F} = 0$ everywhere.

ii. $f \boldsymbol{F} \cdot d\boldsymbol{a}$ is independent of surface, for any given boundary line.

iii. $f \boldsymbol{F} \cdot d\boldsymbol{a} = 0$ for any closed surface.

iv. $\boldsymbol{F}$ is the curl of some vector function: $\boldsymbol{F} = \nabla \times \boldsymbol{A}$.

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

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

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

Part II

Electric Fields

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))

Part III

Magnetic Fields

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

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.)