Compendium of Knowledge
Volume I
Version 0.001

Compiled by: Antonius W. Torode[1]

1. Applied Research Laboratories - University Of Texas: Austin

Written in: LaTeX

Preface

This document is a compilation of scratch work, derivations, useful formulas, definitions, constants, and general information used for my own studies as a reference while furthering self education. These are my notes. It's purpose is to provide a complete 'compendium' per say of various mathematical and significant ideas used often. All the material in this document was either directly copied from one of the references listed at the end or derived from scratch. On occasion *typos may exist* due to human error but will be corrected when discovered.

The version number is updated every time the document is distributed, printed for distribution, or a major update is added. This ensures that there is no two copies with different information and similar version numbers. The latest update date is automatically set to the current date each time the document is edited. Please refrain from distributing this handbook without permission from the original author/compiler. This book is formatted for printing.

Topics Covered In This Book

- General Mathematics
- Calculus
- Linear Algebra
- Geometry
- Statistical Mechanics

- Astrophysics
- Electrodynamics
- Relativity
- Classical Mechanics
- Quantum Mechanics

The information in this book is in no way limited to the topics listed above. They serve as a simple guideline to what you will find within this document. For more information about this book or details about how to obtain your own copy please visit:

https://torodean.github.io/

"Scientific theories deal with concepts, not with reality. All theoretical results are derived from certain axioms by deductive logic. In physical sciences the theories are so formulated as to correspond in some useful sense to the real world, whatever that may mean. However, this correspondence is approximate, and the physical justification of all theoretical conclusions is based on some form of inductive reasoning." - Athanasios Papoulis (Probability, Random Variables, and Stochastic Processes book)

Disclaimer

This book contains formulas, definitions, and theorems that by nature are very precise. Due to this, some of the material in this book was taken directly from other sources such as but not limited to Wolfram Mathworld. This is only such in cases where a change in wording could cause ambiguities or loss of information quality. Following this, all sources used are listed in the references section and cited when used.

This page intentionally left blank.
(Yes, this is a contradiction.)

Contents

Constants and units

1.0.1: Physical Constants

Constant	Symbol	Value	Units
Speed of light in a vacuum	$c \equiv 1/\sqrt{\mu_0 \epsilon_0}$	2.99792458×10^8	m/s
Elementary charge	e	$1.602176565(35) \times 10^{-19}$	C
Gravitational constant	G	$6.67384(80) \times 10^{-11}$	$\mathrm{m^3 kg^{-1} s^{-2}}$
Avagadro's number	N_a	$6.02214129(27) \times 10^{23}$	$\mathrm{mol \cdot s^{-1}}$
Planck constant	h	$6.62606872(52) \times 10^{-34}$	J·s
		4.135668×10^{-15}	eV·s
	hc	1239.84	eV·nm
Reduced planck constant	$\hbar \equiv h/2\pi$	1.05×10^{-34}	J·s
Permittivity of the vacuum	ϵ_0	8.854×10^{-12}	$\mathrm{C^2 N^{-1} m^{-2}}$
Permeability of the vacuum	μ_0	$4\pi \times 10^{-7}$	$\mathrm{N/A^2}$
Permeability of the vacuum	μ_0	$4\pi \times 10^{-7}$	$\mathrm{N/A^2}$
Boltzmann constant	k_B	$1.38064852 \times 10^{-23}$	J/K
		8.61733×10^{-5}	eV/K
Stefan-Boltzmann constant	$\sigma_{\mathrm{B}} \equiv \frac{\pi^2 k_B^4}{60 \hbar^3 c^3}$	$5.670367(13) \times 10^{-8}$	$\mathrm{W \cdot m^{-2} K^{-4}}$
Thomson cross-section	σ_e	6.652×10^{-29}	m^2
The Bohr Magneton	$\mu_B \equiv \frac{e\hbar}{2m}$	5.788×10^{-5}	eV/T
		9.274×10^{-24}	$\mathrm{Am^2}$
Mass of an electron	m_e	$9.10938291(40) \times 10^{-31}$	kg
		510.9989	keV/c^2
Mass of a proton	m_p	$1.6726218 \times 10^{-27}$	kg
		938.27203	MeV/c^2
Mass of a neutron	m_n	$1.6749274 \times 10^{-27}$	kg
		939.56536	MeV/c^2
Unified amu	u	$1.660538782 \times 10^{-27}$	kg
		931.494028	$\mathrm{MeV/c^2}$

1.0.2: Stellar Data

Spectral Type	T_{eff} (K)	$M/M_\odot$	$L/L_\odot$	$R/R_\odot$	V_{mag}
O5	44,500	60	7.9×10^5	12	-5.7
B5	15,400	5.9	830	3.9	-1.2
A5	8,200	2.0	14	1.7	1.9
F5	6,440	1.4	3.2	1.3	3.4
G5	5,770	0.92	0.79	0.92	4.9
K5	4,350	0.67	0.15	0.72	6.7
M5	3,170	0.21	0.011	0.27	12.3

1.0.3: Astronomical Constants

Constant	Symbol	Value	Units
Mass of Earth	$M_{\oplus}$	5.974×10^{24}	kg
Mass of Sun	$M_{\odot}$	1.989×10^{30}	kg
Mass of Moon	$M_{\mathbb{C}}$	7.36×10^{22}	kg
Equatorial radius of Earth	$R_{\oplus}$	6.378×10^{6}	m
Equatorial radius of Sun	$R_{\odot}$	6.6955×10^{8}	m
Equatorial radius of Moon	$R_{\mathbb{C}}$	1.737×10^{6}	m
Mean density of Earth		5515	kg·m^{-3}
Mean density of Sun		1408	kg·m^{-3}
Mean density of Moon		3346	kg·m^{-3}
Earth-Moon distance		3.84×10^{8}	m
Earth-Sun distance		1.496×10^{11}	m
Luminosity of Sun	$L_{\odot}$	3.839×10^{26}	W
Effective temp. of Sun		5778	K
Hubble constant	H_0	70 ± 5	$\text{km·s}^{-1}\text{Mpc}^{-1}$
Parsec	pc	206264.81	AU
		3.0856776×10^{16}	m
		3.2615638	ly
Astronomical Unit	AU	1.496×10^{11}	m
Light year	ly	9.461×10^{15}	m
1 year on Earth	yr	365.25	days
		3.15576×10^{7}	s

1.0.4: Solar System

Planet	Symbol	Mass (kg)	Radius (m)	Sun-Distance (km)
Mercury	☿	3.285×10^{23}	2.44×10^{6}	5.791×10^{10}
Venus	♀	4.867×10^{24}	6.052×10^{6}	1.082×10^{11}
Mars	♂	6.39×10^{23}	3.390×10^{6}	2.279×10^{11}
Jupiter	♃	1.898×10^{27}	3.83×10^{11}	7.785×10^{11}
Saturn	♄	5.683×10^{26}	5.8232×10^{7}	1.429×10^{12}
Uranus	♅	8.681×10^{25}	2.5362×10^{7}	2.871×10^{12}
Neptune	♆	1.024×10^{26}	2.4622×10^{7}	4.498×10^{12}
Pluto	♇	1.309×10^{22}	1.187×10^{6}	5.906×10^{12}

1.0.5: Unit conversions

The International System of Units (SI) defines seven units of measure as a basic set from which all other SI units can be derived. These are [length](m), [time](s), [mass](kg), [electric current] ≡ [Ampere](A), [temperature](K), [luminous intensity](cd), [amount of substance](mol).

Unit Symbol	Unit	Equivalence
C	[Coulomb]	[Ampere][time]
N	[Newton]	$[\text{mass}][\text{length}][\text{time}]^{-2}$
P	[Pascal]	$[\text{mass}][\text{length}]^{-1}[\text{time}]^{-2}$
J	[Joule]	$[\text{mass}][\text{length}]^{2}[\text{time}]^{-2}$
W	[Watt]	$[\text{mass}][\text{length}]^{2}[\text{time}]^{-3}$
		$[\text{Ohm}][\text{Ampere}]^{2}$
		$[\text{Volt}]^{2}[\text{Ohm}]^{-1}$
V	[Volt]	$[\text{mass}][\text{length}]^{2}[\text{time}]^{-3}[\text{Ampere}]^{-1}$
Wb	[Weber]	$[\text{mass}][\text{length}]^{2}[\text{time}]^{-2}[\text{Ampere}]^{-1}$
T	[Tesla]	$[\text{mass}][\text{time}]^{-2}[\text{Ampere}]^{-1}$
H	[henry]	$[\text{mass}][\text{length}]^{2}[\text{time}]^{-2}[\text{Ampere}]^{-2}$
Ω	[Ohm]	$[\text{mass}][\text{length}]^{2}[\text{time}]^{-3}[\text{Ampere}]^{-2}$
F	[Farad]	$[\text{mass}]^{-1}[\text{length}]^{-2}[\text{time}]^{4}[\text{Ampere}]^{2}$
Hz	[Hertz]	$[\text{time}]^{-1}$

1.0.6: Number Sets ($i \equiv \sqrt{-1}$)

Symbol	Set	Symbol	Set
$\mathbb{R}$	Real numbers	$\emptyset$	{}
$\mathbb{N} \equiv \mathbb{N}_1$	$\{1,2,3,4,\dots\}$	$\mathbb{Z}$	$\{\dots,-2,1,0,1,2,\dots\}$
$\mathbb{Z}^+ \equiv \mathbb{N}_0$	$\{0,1,2,3,\dots\}$	$\mathbb{Z}^-$	$\{0,-1,-2,-3,-4,\dots\}$
$\mathbb{C}$	$\{x+iy \vert x,y \in \mathbb{R}\}$	$\mathbb{Q}$	$\{\frac{x}{y} \vert x,y \in \mathbb{Z}\}$
$\mathbb{I}$	$\{ix \vert x \in \mathbb{R}\}$	$\mathbb{U}$	Universal Set [a]
$\mathbb{A}$	Algebraic Numbers[b]	$\mathbb{T}$	Transcendental Numbers [c]

[a] Definition: The set containing all objects or elements and of which all other sets are subsets.
[b] Any number that is a solution to a polynomial equation with rational coefficients.
[c] Any number that is not an Algebraic Number.

1.0.7: Mathematical Notation

$\forall$	For all	$\exists$	There exists	$\because$	Because
$\in$	Is an element of	$\notin$	Is not an element of	$\therefore$	Therefore
$\implies$	Implies	$\iff$	Bi conditional	$\approx$	Approximately
$\longrightarrow$	Mapped to	$\not\subset$	Is not a subset of	$\ll$	Much smaller than
$\subset$	Is a subset of	$\subseteq$	Is a subset or equal to	$\gg$	Much greater than
$\propto$	Is proportional to	$\equiv$	Is equivalent to	$\cup/\cap$	Union/Intersection
$\perp$	Is perpendicular to	$\parallel$	Is parallel to	: or |	Such that

Geometry and Trigonometry

Geometry is a branch of mathematics that deals with the study of shapes, sizes, positions, and properties of space. It explores the relationships and properties of points, lines, angles, surfaces, and solids. In geometry, fundamental concepts include:

1. **Points**: Basic building blocks with no size or dimensions. They are represented by dots and are used to define other geometric elements.

2. **Lines**: Straight paths that extend infinitely in both directions. They are made up of an infinite number of points.

3. **Angles**: The measure of the rotation between two intersecting lines, rays, or line segments.

4. **Polygons**: Closed figures formed by connecting line segments to create shapes like triangles, quadrilaterals, pentagons, etc.

5. **Circles**: A set of points equidistant from a central point, forming a closed curve.

6. **Three-dimensional shapes**: Solids such as cubes, spheres, cylinders, pyramids, etc., with length, width, and height.

7. Higher-dimensional shapes or objects: In addition to the traditional two-dimensional and three-dimensional shapes, geometry also explores mathematical constructs extended to higher dimensions. These higher dimensions go beyond our familiar three-dimensional space and introduce concepts like 4D, 5D, and higher-dimensional shapes. For example, in 4D space, there could be hypercubes, 4D spheres, and other intriguing structures. While challenging to visualize, these higher-dimensional geometries play a crucial role in theoretical mathematics and various scientific fields, offering unique insights into the nature of space and dimensions beyond our immediate perception.

Geometry plays a significant role in various fields, including architecture, engineering, art, design, physics, and many other disciplines. It helps us understand the physical world and solve problems related to spatial relationships and measurements. Euclidean geometry, founded by the ancient Greek mathematician Euclid, is one of the most well-known and widely studied branches of geometry. However, there are also other types of geometries, such as non-Euclidean geometries, which explore different axioms and concepts, leading to intriguing and diverse mathematical systems.

2.1 Triangles

Definition 2.1.1: The Law Of Cosines

The Law of Cosines relates the side lengths of a triangle to the cosine of one of its angles. For a triangle with side lengths a, b, and c, and angle C opposite to side c, the Law of Cosines is given by:

$$c^2 = a^2 + b^2 - 2ab\cos C \qquad (2.1.1)$$

Notice from definition 2.1.1 that if the angle $C = 90° = \pi/2$, then the cosine term would evaluate to zero and we'd be left with **the Pythagorean theorem** $c^2 = a^2 + b^2$.

General Mathematics

Mathematics is a systematic field of study that deals with numbers, quantities, shapes, and patterns. It is often regarded as the language of science, as it provides a framework for analyzing and understanding various phenomena in the natural and social world. Mathematics involves a wide range of topics, including arithmetic, algebra, geometry, calculus, statistics, and more. It is based on rigorous logical reasoning and uses symbols and formulas to represent relationships and solve problems. Mathematics plays a crucial role in various fields, such as physics, engineering, economics, computer science, and many other disciplines, making it an essential tool for advancing knowledge and technology.

3.1   Coordinate Systems

Coordinate systems are mathematical tools used to represent points in space or on a surface using a set of numerical values. They provide a systematic way to describe the location of points relative to a reference point or axis. There are several types of coordinate systems, each with its unique set of parameters and applications.

3.1.1   Cartesian Coordinates

Cartesian coordinates, also known as **rectangular coordinates**, are a system of representing points in a two-dimensional or three-dimensional space using a set of perpendicular axes (See figure 3.1). Cartesian coordinates can be mathematically extended into higher dimensions.

Figure 3.1: 2-D and 3-D Cartesian Coordinate Systems with a single point located at one unit in each axes.

In two-dimensional Cartesian coordinates, a point is located by specifying its distance (magnitude) along the x-axis and the y-axis from the origin (0,0), where the x-axis is horizontal, and the y-axis is vertical. The coordinates of a point are written as (x, y), where x represents the horizontal distance and y represents the vertical distance. In three-dimensional Cartesian coordinates, a point is located by specifying its distance along the x-axis, y-axis, and z-axis from the origin (0, 0, 0). The coordinates of a point are written as (x, y, z), where x, y, and z represent the distances along the three axes. In higher n-dimensions, a similar formulation is used where a point is located by specifying its distance along n-axes (typically denoted by the unit vector $\hat{e}_i$) from the origin $\vec{0}$. The coordinates are $(x_0, x_1, ..., x_n)$, where x_i represents the distance along the respective n axes.

3.1.2 Polar Coordinates

Polar coordinates are a two-dimensional coordinate system (see figure 3.2)used to describe points in a plane. Unlike the traditional Cartesian coordinates (x, y), polar coordinates represent points using a distance (r) from the origin and an angle (θ) measured counterclockwise from a reference line (usually the positive x-axis). The distance r is always a non-negative value, representing the radial distance from the origin to the point. The polar coordinate system is especially useful when dealing with circular or radial patterns, as it simplifies calculations involving angles and distances. The conversion between polar coordinates and Cartesian coordinates can be done using trigonometric functions.

$$x = r\cos(\theta) \qquad\qquad y = r\sin(\theta) \tag{3.1.1}$$

$$r = \sqrt{x^2 + y^2} \qquad\qquad \theta = \arctan\left(\frac{y}{x}\right). \tag{3.1.2}$$

Figure 3.2: LEFT: The point (x, y) in Cartesian coordinates can be represented as (r, θ) in polar coordinates, where r is the distance from the origin to the point and θ is the angle measured counterclockwise from the positive x-axis. RIGHT: The point (x, y, z) in Cartesian coordinates can be represented as (r, θ, ϕ) in a modified spherical coordinate system, where r is the distance from the origin to the point, θ is the angle measured counterclockwise from the positive x-axis towards the y direction and ϕ is the angle from the z-axis.

3.1.3 Spherical Coordinates

Spherical coordinates (see figure 3.2) are a system of representing points in 3D space using three parameters: radial distance r, polar angle θ, and azimuthal angle ϕ. This system is particularly useful for describing points on a sphere or points in a 3D space relative to a fixed origin. The radial distance r represents the distance from the origin to the point. It is a non-negative value and measures how far the point is from the origin. It is important to note that polar angle and azimuthal angle can often be represented in different ways depending on the coordinate convention. Typically, polar angle represents the angle between the positive z-axis and the line connecting the origin to the point projected onto the xy-plane. It varies from 0° (on the positive z-axis) to 180° (on the negative z-axis). This 'polar' angle in essence would then connect both poles on a planet. I prefer to use the convention that matches 2-dimensional polar coordinates, as it removes some confusion when converting between the two - therefore θ is defined as the same angle as it would be in 2 dimensions and then ϕ is the angle of that 2-dimensional plane rotated about the z-axis. The conversion between spherical coordinates and Cartesian coordinates can be done using trigonometric functions. For the convention I prefer, and matching figure 3.2, we have

$$x = r\cos(\theta)\sin(\phi), \qquad y = r\sin(\theta), \qquad z = r\cos(\theta)\cos(\phi), \qquad r = \sqrt{x^2 + y^2 + z^2} \qquad (3.1.3)$$

These can be shown to be equivalent to the more commonly used spherical coordinates. To demonstrate this, I will denote the coordinates defined above with the subscript $_0$. For a more common convention (the physics one), imagine we modify our coordinates so that the x_0 axis is the y axis, The y_0 axis is the z axis, and the z_0 axis is the x axis. Then we further define θ as the angle from the z axis to our radial line, which is therefore $\theta_0 = \frac{\pi}{2} - \theta$ or $\theta = \frac{\pi}{2} - \theta_0$. In this change, r does not change as we are not changing our line or angles (just denoting them differently) giving us $r_0 \equiv r$, and ϕ also stays the same giving $\phi_0 \equiv \phi$. Plugging these into our coordinates above give us

$$x_0 = r_0\cos(\theta_0)\sin(\phi_0) \implies y = r\cos(\frac{\pi}{2} - \theta)\sin(\phi) = r\sin(\theta)\sin(\phi) \qquad (3.1.4)$$

$$y_0 = r_0\sin(\theta_0) \implies z = r\sin(\frac{\pi}{2} - \theta) = r\cos(\theta) \qquad (3.1.5)$$

$$z_0 = r_0\cos(\theta_0)\cos(\phi_0) \implies x = r\cos(\frac{\pi}{2} - \theta)\cos(\phi) = r\sin(\theta)\cos(\phi) \qquad (3.1.6)$$

$$r_0 = \sqrt{x_0^2 + y_0^2 + z_0^2} \implies r = \sqrt{x^2 + y^2 + z^2} \qquad (3.1.7)$$

This conversion gives us the more commonly used spherical coordinate convention

$$x = r\sin(\theta)\cos(\phi), \qquad y = r\sin(\theta)\sin(\phi), \qquad z = r\cos(\theta), \qquad r = \sqrt{x^2 + y^2 + z^2} \qquad (3.1.8)$$

Therefore we can see they are equivalent and can be used interchangeably.

3.1.4 Cylindrical Coordinates

Figure 3.3: The point (x, y, z) in Cartesian coordinates can be represented as (r, θ, z) in cylindrical coordinates, where r is the distance from the z axis to the point and θ is the angle measured counterclockwise from the positive x-axis.

Cylindrical coordinates (see figure 3.3) are a system used to represent points in three-dimensional space using three parameters: radial distance r, polar angle θ, and height z. In this system, the radial

distance measures the distance from the z-axis to the point in a plane perpendicular to the z-axis. The polar angle θ represents the azimuthal angle, measured counterclockwise from the positive x-axis to the projection of the point on the xy-plane. The height z indicates the distance of the point along the z-axis. Cylindrical coordinates are particularly useful for describing objects or phenomena with cylindrical symmetry. The conversion between cylindrical coordinates and Cartesian coordinates can be done using trigonometric functions.

$$x = r\cos(\theta), \qquad y = r\sin(\theta), \qquad z = z, \qquad r = \sqrt{x^2 + y^2} \tag{3.1.9}$$

3.1.5 Elliptic Coordinates

3.1.6 Elliptic Cylindrical Coordinates

3.2 Logarithms

Logarithms are fundamental mathematical functions that arise in various areas of mathematics, science, and engineering. They are used to solve exponential equations, simplify calculations involving large numbers, and analyze exponential growth and decay.

Definition 3.2.1: Logarithm

The **logarithm** of a positive number x with respect to a base b is defined as the exponent to which b must be raised to obtain x. This is expressed as:

$$\log_b(x) = y \quad \text{if and only if} \quad b^y = x \tag{3.2.1}$$

Two commonly used bases for logarithms are 10 and the mathematical constant e (approximately 2.71828). The logarithms with base 10 are called common logarithms, denoted as $\log_{10}(x)$ or simply $\log(x)$. The logarithms with base e are called natural logarithms, denoted as $\ln(x)$. Sometimes in physics or other natural science fields, the natural logarithm is implied and these will be used interchangeably.

Logarithms have a wide range of applications in various fields:

- **Exponential Growth and Decay**: Logarithms are used to model and analyze processes that involve exponential growth or decay, such as population growth, radioactive decay, and compound interest.

- **Scientific Notation**: Logarithms are used to express very large or very small numbers in a more compact and manageable form known as scientific notation.

- **Signal Processing**: Logarithms are used in signal processing to compress dynamic range and reduce computational complexity.

- **Complex Analysis**: Logarithms are used in complex analysis to define complex powers and extend the concept of logarithms to complex numbers.

3.2.1 Properties of Logarithms

Logarithms possess several important properties that make them useful in various mathematical manipulations:

- **Product Rule**: $\log_b(xy) = \log_b(x) + \log_b(y)$

Proof. Suppose $\log_b(xy) = z$, then $b^z = xy$ (by definition). Let $\log_b(x) = z_0$, and $\log_b(y) = z_1$. Taking these two expressions, we have (by definition)

$$\log_b(x) = z_0 \implies b^{z_0} = x \tag{3.2.2}$$
$$\log_b(y) = z_1 \implies b^{z_1} = y. \tag{3.2.3}$$

Multiplying x by y gives $xy = b^{z_0} b^{z_1} = b^{z_0 + z_1}$. But since $b^z = xy$, we have $b^z = b^{z_0 + z_1} \implies z = z_0 + z_1 = \log_b(x) + \log_b(y) = \log_b(xy)$. $\qquad\square$

- **Quotient Rule**: $\log_b\left(\frac{x}{y}\right) = \log_b(x) - \log_b(y)$

Proof. Suppose $\log_b\left(\frac{x}{y}\right) = z$, then $b^z = \frac{x}{y}$ (by definition). Let $\log_b(x) = z_0$, and $\log_b(y) = z_1$. Taking these two expressions, we have (by definition)

$$\log_b(x) = z_0 \implies b^{z_0} = x \tag{3.2.4}$$
$$\log_b(y) = z_1 \implies b^{z_1} = y. \tag{3.2.5}$$

Dividing x by y gives $\frac{x}{y} = \frac{b^{z_0}}{b^{z_1}} = b^{z_0 - z_1}$. But since $b^z = \frac{x}{y}$, we have $b^z = b^{z_0 - z_1} \implies z = z_0 - z_1 = \log_b(x) - \log_b(y) = \log_b\left(\frac{x}{y}\right)$. $\qquad\square$

- **Power Rule**: $\log_b(x^n) = n \cdot \log_b(x)$

Proof. Suppose $\log_b(x^n) = z$, then $b^z = x^n$ (by definition). Raising both sides by $\frac{1}{n}$ yields $b^{\frac{z}{n}} = x$. Let $n \log_b(x) = q$. Dividing both sides by n and applying the definition of logarithm gives $\log_b(x) = \frac{q}{n} \implies b^{\frac{q}{n}} = x$. Since $x = b^{\frac{z}{n}} = b^{\frac{q}{n}} \implies z = q \implies \log_b(x^n) = n \cdot \log_b(x)$ $\qquad\square$

- **Change of Base Formula**: $\log_b(x) = \frac{\log_c(x)}{\log_c(b)}$ (for any positive $c \neq 1$)

Proof. Let $\log_b(x) = z$, $\log_c(x) = z_0$, and $\log_c(b) = z_1$. By definition we then have $b^z = x$, $c^{z_0} = x$, and $c^{z_1} = b$. Raising b in this last expression to the power of z gives $b^z = c^{z z_1}$. Then it follows that $c^{z_0} = x = b^z = c^{z z_1}$. We can then relate the exponents $z z_1 = z_0 \implies z = \frac{z_0}{z_1}$. Using the definitions of z_0 and z_1, this gives $z = \frac{z_0}{z_1} = \frac{\log_c(x)}{\log_c(b)} = \log_b(x)$. $\qquad\square$

3.3 Conventions and basic Definitions

3.3.1 Complex Numbers

Definition 3.3.1: Complex Number

complex number is a number of the form $z = a + bi$, where $a, b \in \mathbb{R}$, and i is the imaginary unit with the property $i^2 = -1$.

> **Definition 3.3.2: Complex Conjugate**
>
> Suppose $z = a + bi$ is a **complex number**, where $a, b \in \mathbb{R}$ and i is the imaginary unit ($i^2 = -1$). The complex conjugate of z, typically denoted by $\bar{z}$ or z^*, is obtained by changing the sign of the imaginary part.
>
> $$z = a + bi \implies \bar{z} \equiv z^* = a - bi \qquad (3.3.1)$$

A few mathematical identities hold when combining complex numbers with their complex conjugates.

- The product of a complex number and its complex conjugate is always a real number:

$$z \cdot \bar{z} = (a + bi)(a - bi) = a^2 + b^2 \qquad (3.3.2)$$

- The sum of a complex number and its complex conjugate gives a real number with twice the real part:

$$z + \bar{z} = (a + bi) + (a - bi) = 2a \qquad (3.3.3)$$

- The difference between a complex number and its complex conjugate gives a purely imaginary number with twice the imaginary part:

$$z - \bar{z} = (a + bi) - (a - bi) = 2bi \qquad (3.3.4)$$

- The complex conjugate of a product is the product of the complex conjugates.

$$
\begin{aligned}
(z_1 \cdot z_2)^* &= [(a_1 + b_1 i)(a_2 + b_2 i)]^* & (3.3.5) \\
&= [a_1 a_2 + a_1 b_2 i + a_2 b_1 i - b_1 b_2]^* & (3.3.6) \\
&= a_1 a_2 - b_1 b_2 - (a_1 b_2 + a_2 b_1)i & (3.3.7) \\
&= (a_1 - b_1 i)(a_2 - b_2 i) & (3.3.8) \\
&= z_1^* \cdot z_2^* & (3.3.9)
\end{aligned}
$$

- The inverse of a non-zero complex conjugate number $z = a + bi$ is

$$\frac{1}{z} = \frac{1}{a + bi}\left(\frac{a + bi}{a + bi}\right) = \frac{a + bi}{a^2 + b^2} = \frac{a}{a^2 + b^2} + \frac{b}{a^2 + b^2}i \qquad (3.3.10)$$

3.3.2   Vectors

The **bra** and **ket** notation is commonly used to represent vectors and vector operations. A superscript "T" represents the **transposition** of a vector or a matrix. Vectors in 3 spacial dimensions, such as those in classical mechanics, are typically notated with an arrow overhead (i.e. $\vec{v}$). In other contexts, vectors are either represented by boldface script or by Dirac's **bra-ket** notation.

Definition 3.3.3: Dirac's Bra-ket Notation

Ket vectors are represented by column vectors, by vertically arranged tuples of scalers (or equivalently, by $n \times 1$ matrices).

$$\vec{x} \equiv \mathbf{x} \equiv |\mathbf{x}\rangle \equiv \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \left(x_1, x_2, \ldots, x_n\right)^T \tag{3.3.11}$$

A vector $\mathbf{x}^*$ with an asterisk symbol in its superscript denotes an element of the dual space. This is represented by the bra vector $\langle \mathbf{x}|$. Bra vectors represent row vectors - horizontally arranged tuples of scalars (i.e. $1 \times n$ matrices).

$$\mathbf{x}^* \equiv \langle \mathbf{x}| \equiv \left(x_1, x_2, \ldots, x_n\right) \tag{3.3.12}$$

The dot (inner or scaler) products between two vectors $\mathbf{x}$ and $\mathbf{y}$ in Euclidean space is denoted by $\langle \mathbf{x}|\mathbf{y}\rangle$.

The **dot product**, also known as the scalar product or inner product, is an operation defined on vectors in mathematics and physics. It takes two vectors of equal dimensionality and produces a single scalar value as the result. The dot product is commonly denoted using the dot symbol ($\cdot$), bra-ket notation, or by juxtaposing the vectors without any operator.

Definition 3.3.4: Dot Prodcuct

Given two vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, the dot product is defined as

$$\mathbf{x} \cdot \mathbf{y} \equiv \mathbf{x}^T \mathbf{y} \equiv \langle \mathbf{x}|\mathbf{y}\rangle \equiv |\mathbf{x}||\mathbf{y}|\cos(\theta) \equiv \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = \underbrace{x_i y_i}_{\text{Einstein Notation}} , \tag{3.3.13}$$

where θ is the smallest angle between the two vectors.

A few mathematical identities hold when dealing with dot products.

- **Commutativity**

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = y_1 x_1 + y_2 x_2 + \ldots + y_n x_n = \mathbf{y} \cdot \mathbf{x} \tag{3.3.14}$$

- Distributive Property

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \sum_{i=0}^{n} a_i(b_i + c_i)\hat{e}_i = \sum_{i=0}^{n} a_i b_i \hat{e}_i + \sum_{i=0}^{n} a_i c_i \hat{e}_i = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C} \tag{3.3.15}$$

- **Scalar Multiplication**

$$(k\mathbf{A}) \cdot \mathbf{B} = \sum_{i=0}^{n} (ka_i) b_i \hat{e}_i = k \sum_{i=0}^{n} a_i b_i \hat{e}_i = k(\mathbf{A} \cdot \mathbf{B}) = \sum_{i=0}^{n} a_i(kb_i)\hat{e}_i = \mathbf{A} \cdot (k\mathbf{B}) \tag{3.3.16}$$

- **Orthogonal** (**Perpendicular**) Vectors: If the dot product of two non-zero vectors is zero, they are orthogonal (perpendicular) to each other.

$$\mathbf{x} \cdot \mathbf{y} = 0 \implies |\mathbf{x}||\mathbf{y}| \cos(\theta) = 0 \implies \underbrace{\theta = \frac{\pi}{2} + \pi n, \text{ where } n \in \mathbb{N}}_{90° \text{ between vectors}} \tag{3.3.17}$$

- **Projection**: The dot product can be used to find the projection of one vector onto another. The projection of vector $\mathbf{A}$ onto vector $\mathbf{B}$ is given by

$$\text{proj}_{\mathbf{B}}(\mathbf{A}) = \frac{\mathbf{A} \cdot \mathbf{B}}{|\mathbf{B}|^2}\mathbf{B} \tag{3.3.18}$$

3.3.3 The Dot Product Theorem

Suppose we have two 3-dimensional vectors $\vec{r} = (r_x, r_y, r_z)$ and $\vec{s} = (s_x, s_y, s_z)$ in a Euclidean x, y, z coordinate system. The dot product is then $\vec{r} \cdot \vec{s} = r_x s_x + r_y s_y + r_z s_z = |\vec{r}||\vec{s}| \cos(\theta)$, where θ is the smallest angle between the two vectors. From this, the angle between these two 3 dimensional vectors can be determined as

$$\theta = \cos^{-1}\left(\frac{\vec{r} \cdot \vec{s}}{|\vec{r}||\vec{s}|}\right) \tag{3.3.19}$$

Consider two vectors $\vec{A} = A_x \hat{x} + A_y \hat{y}$ and $\vec{B} = B_x \hat{x} + B_y \hat{y}$ in a 2-dimensional plane. In polar coordinates this woulkd be $\vec{A} = A \cos(\theta_A)\hat{x} + A \sin(\theta_A)\hat{y}$ and $\vec{B} = B \cos(\theta_B)\hat{x} + B \sin(\theta_B)\hat{y}$, where $A = \sqrt{A_x^2 + A_y^2}$ and $B = \sqrt{B_x^2 + B_y^2}$ are the magnitudes of each vector. The definition of dot product according to Wolfram's Mathworld is simply

$$\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos(\theta) \tag{3.3.20}$$

, where θ is the smallest angle between the two vectors. Using the polar coordinates, the magnitudes of each vector are simply A and B and the angle between them is the differences of the angles $\theta_A - \theta_B$. Therefore the dot product is $\vec{A} \cdot \vec{B} = AB \cos(\theta_A - \theta_B)$. We can apply some trigonometric identities to show this is equal to the summation notation in Definition 3.3.4.

$$\begin{align}
\vec{A} \cdot \vec{B} &= AB \cos(\theta_A - \theta_B) \tag{3.3.21}\\
&= AB(\cos(\theta_A) \cos(\theta_B) + \sin(\theta_A) \sin(\theta_B)) \tag{3.3.22}\\
&= A \cos(\theta_A) B \cos(\theta_B) + A \sin(\theta_A) B \sin(\theta_B) \tag{3.3.23}\\
&= A_x B_x + A_y B_y \tag{3.3.24}
\end{align}$$

A question then arises of whether of not this holds for dimensions higher than two. Now, imagine rotating the 2 dimensional reference frame for $\vec{A}$ and $\vec{B}$ into a 3-dimensional reference frame (rotating the coordinate system). The vectors themselves would be mathematically identical, as the vectors themselves did not change, however given the reference frame changed, they would have different coordinates. Thus, we define the new coordinates of these vectors in the new frame as $\vec{A} \equiv \vec{X} = x_1 \hat{e}_1 + x_2 \hat{e}_2 + x_3 \hat{e}_3$ and $\vec{B} \equiv \vec{Y} = y_1 \hat{e}_1 + y_2 \hat{e}_2 + y_3 \hat{e}_3$. Since we have only rotated our coordinate frame and not changed the vectors, the angle between them $\theta_A - \theta_B$ remains unchanged. We can begin by trying a similar method to what we got above.

Converting our two vectors to spherical coordinates (now that we are in 3-dimensional we use spherical rather than polar) we have

$$\vec{X} = x_1\hat{e}_1 + x_2\hat{e}_2 + x_3\hat{e}_3 = X\sin(\theta_x)\cos(\phi_x)\hat{e}_1 + X\sin(\theta_x)\sin(\phi_x)\hat{e}_2 + X\cos(\theta_x)\hat{e}_3 \tag{3.3.25}$$

$$\vec{Y} = y_1\hat{e}_1 + y_2\hat{e}_2 + y_3\hat{e}_3 = Y\sin(\theta_y)\cos(\phi_y)\hat{e}_1 + Y\sin(\theta_y)\sin(\phi_y)\hat{e}_2 + Y\cos(\theta_y)\hat{e}_3, \tag{3.3.26}$$

where $X = \sqrt{x_1^2 + x_2^2 + x_3^2}$ and $Y = \sqrt{y_1^2 + y_2^2 + y_3^2}$. Assuming that definition 3.3.4 is correct, we can apply the summation form of the formula to calculate the dot product. In order to condense the formula a tad, I will preemptively divide both sides by the common magnitude factors that will exist in each term. Thus we have

$$\frac{\vec{X}\cdot\vec{Y}}{XY} \stackrel{?}{=} \sin(\theta_x)\cos(\phi_x)\sin(\theta_y)\cos(\phi_y) + \sin(\theta_x)\sin(\phi_x)\sin(\theta_y)\cos(\phi_y) + \cos(\theta_x)\cos(\theta_y) \tag{3.3.27}$$

Through some algebra, this simplifies to

$$\vec{X}\cdot\vec{Y} \stackrel{?}{=} XY\left[\sin(\theta_x)\sin(\theta_y)\cos(\phi_x - \phi_y) + \cos(\theta_x)\cos(\theta_y)\right] \tag{3.3.28}$$

Since the angle between the two vectors $\theta_A - \theta_B$ remained unchanged in our coordinate rotation and the magnitude of the vectors remained unchanged (that is, $A \equiv X$ and $B \equiv Y$), equation (3.3.28) should be equivalent to equation (3.3.21) giving us

$$\cos(\theta_A - \theta_B) \stackrel{?}{=} \left[\sin(\theta_x)\sin(\theta_y)\cos(\phi_x - \phi_y) + \cos(\theta_x)\cos(\theta_y)\right] \tag{3.3.29}$$

It is not obvious that this is an accurate equation. Although, if true, this gives a useful transformation between the angles between our two coordinate systems. Note that any 3-dimensional set of vectors can be transferred onto a 2-dimensional plane in a manner similar to this. We can try another way to demonstrate the relationship. Consider our vectors $\vec{X}$ and $\vec{Y}$. The vector from the tip of $\vec{X}$ to the tip of $\vec{Y}$ would be $\vec{X} - \vec{Y}$ (see figure 3.4).

Figure 3.4: Vector subtraction: $\vec{X} - \vec{Y}$ with angle θ as the smallest angle between $\vec{X}$ and $\vec{Y}$.

Let us further generalize these vectors into higher dimensions in a manner that we did previously by rotating our vectors into another higher n-dimension such that $\vec{X} = \sum_{i=1}^{n} x_i\hat{e}_i$ and $\vec{Y} = \sum_{i=1}^{n} y_i\hat{e}_i$. Note that the above relation in figure 3.4 between the two vectors holds for any dimension we rotate them into because the vectors themselves are not changing. With this in mind, we can take the magnitudes of the vectors and apply the law of cosines (see definition 2.1.1) to get

$$|\vec{X} - \vec{Y}|^2 = |\vec{X}|^2 + |\vec{Y}|^2 - 2|\vec{X}||\vec{Y}|\cos(\theta) \tag{3.3.30}$$

Since the magnitude is defined as $|A| = \sqrt{\sum_{i=1}^{n} a_i^2}$, this becomes

$$\sum_{i=1}^{n}(x_i - y_i)^2 = \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} y_i^2 - 2|\vec{X}||\vec{Y}|\cos(\theta) \tag{3.3.31}$$

We can then expand the left hand side of this to $\sum_{i=1}^{n}(x_i^2 - 2x_iy_i + y_i) = \sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} x_iy_i + \sum_{i=1}^{n} y_i^2$ and cancel similar terms to get

$$-2\sum_{i=1}^{n} x_iy_i = -2|\vec{X}||\vec{Y}|\cos(\theta) \tag{3.3.32}$$

Finally, dividing both sides by negative two and apply equation (3.3.20) we see that

$$\vec{X} \cdot \vec{Y} = \sum_{i=1}^{n} x_iy_i \tag{3.3.33}$$

3.4 Summation

Definition 3.4.1: Summation Notation

A sum is the result of addition. The symbol $\sum$ is used to represent the addition of a series of values. Let $..., x_{a-1}, x_a, x_{a+1}, ..., x_{b-1}, x_b, x_{b+1}, ...$ be a fixed series $\mathbb{S}$, where each term $x_i \in \mathbb{S}$, $b, a \in \mathbb{N}$ are index values. Then the sum of values in the series from a to b is given by

$$\sum_{i=a}^{b} x_i = x_a + x_{a+1} + ... + x_{b-1} + x_b \tag{3.4.1}$$

In the context of summing over all of the values of a series for some index i, it is often written in different ways which are typically equivalent.

$$\sum_{i\in\mathbb{S}} x_i \equiv \sum_{i} x_i \tag{3.4.2}$$

If the number of terms in the summation is infinite then it is called an infinite series.

Definition 3.4.2: Product Notation

A product is the result of multiplication. The symbol $\prod$ is used to represent the multiplication of a series of values. Let $..., x_{a-1}, x_a, x_{a+1}, ..., x_{b-1}, x_b, x_{b+1}, ...$ be a fixed series $\mathbb{S}$, where each term $x_i \in \mathbb{S}$, $b, a \in \mathbb{N}$ are index values. Then the product of values in the series from a to b is given by

$$\prod_{i=a}^{b} x_i = x_a \cdot x_{a+1} \cdot ... \cdot x_{b-1} \cdot x_b \tag{3.4.3}$$

In the context of multiplying over all of the values of a series for some index i, it is often written in different ways which are typically equivalent.

$$\prod_{i\in\mathbb{S}} x_i \equiv \prod_{i} x_i \tag{3.4.4}$$

If the number of terms in the summation is infinite then it is called an infinite series.

3.4.1   Pascal's Triangle

Pascal's Triangle is a mathematical arrangement of numbers that holds several interesting properties and has applications in various fields. Pascal's Triangle is constructed by starting with a single "1" at the top. Each subsequent row is created by placing "1"s on both ends, and each inner value is the sum of the two values directly above it.

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}
\tag{3.4.5}
$$

The numbers in Pascal's Triangle are also known as **binomial coefficients**. The entry in the nth row and kth column represents the coefficient of the term x^k in the expansion of $(x + 1)^n$. The values in Pascal's Triangle also correspond to combinations (binomial coefficients) that count the number of ways to choose k items from a set of n items without considering the order. Pascal's Triangle is symmetric along its central vertical axis. Each row is a palindrome. The sum of the numbers in each row is a power of 2 (see theorem 3.5.1). In number theory, Pascal's Triangle is involved in finding patterns and relationships in numbers, and it is often connected to the study of the binomial series and power series in calculus. Pascal's Triangle can be extended to include negative row indices and even more complex expansions.

3.5   Binomial Coefficients

Definition 3.5.1: Binomial Coefficient

The binomial coefficient, $\binom{n}{k}$, often denoted as "n choose k," is a mathematical concept used in combinatorics to represent the number of ways to choose k items from a set of n distinct items, regardless of the arrangement of the chosen elements.

$$
\binom{n}{k} = \frac{n!}{(n - k)!k!}
\tag{3.5.1}
$$

To begin with, a useful idea is to sum each term of the binomial coefficient which will be used later.

First, by definition of the binomial coefficient, we can write

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} \tag{3.5.2}$$

$$= \frac{(n-1)!n}{(n-k)!k!} \tag{3.5.3}$$

$$= \frac{(n-1)!n - k(n-1)! + k(n-1)!}{(n-k)!k!} \tag{3.5.4}$$

$$= \frac{(n-1)!(n-k)}{(n-k)!k!} + \frac{k(n-1)!}{(n-k)!k!} \tag{3.5.5}$$

$$= \frac{(n-1)!}{(n-1-k)!k!} + \frac{(n-1)!}{(n-k)!(k-1)!} \tag{3.5.6}$$

$$= \binom{n-1}{k} + \binom{n-1}{k-1} \tag{3.5.7}$$

$$\equiv \binom{m}{k} + \binom{m}{k-1}, \text{ with } m = n-1. \tag{3.5.8}$$

Observe for a moment that (3.5.7) can be expanding even further. By the same process of going from (3.5.2) to (3.5.7), we can say

$$\binom{n-1}{k-1} = \binom{n-2}{k-1} + \binom{n-2}{k-2}. \tag{3.5.9}$$

Thus, (3.5.7) becomes

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-2}{k-1} + \binom{n-2}{k-2}. \tag{3.5.10}$$

If we continue this pattern, we can see that we can write the binomial coefficients as a sum of binomial coefficients with incrementally decreasing numerators (i.e n!, (n-1)!, (n-2)!,...). This gives

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-2}{k-1} + \binom{n-3}{k-2} + \cdots + \binom{1}{k+2-n} + \binom{0}{k+1-n} \tag{3.5.11}$$

$$= \sum_{i=1}^{n} \binom{n-i}{k+1-i} \tag{3.5.12}$$

$$= \sum_{i=0}^{n-1} \binom{n-1-i}{k-i}. \tag{3.5.13}$$

Note that we stop the above sequence when the numerator of our factorial sequence has reached zero. If we were to continue the sequence, we would end up having negative factorials in our numerator which would make evaluating the binomial coefficient at that term and the following terms difficult. However, if we happen to have a smaller k than n, it may be such that we end up having negative k numbers. This is ok for now, as it will lead to complex infinities in the denominator of our binomial coefficient expressions hence making those terms zero. This will be demonstrated later. Now, suppose we claim the following theorem.

Theorem 3.5.1: Summing Over All binomial coefficients

For all $n \in \mathbb{N}$,

$$\sum_{i=0}^{n} \binom{n}{i} = \sum_{i=0}^{n} \frac{n!}{(n-i)!i!} = 2^n. \tag{3.5.14}$$

Proof. We can then do a proof by induction to prove this is in fact true for all $n, k \in \mathbb{N}$. First, we can see checking the base cases hold as when $n = 0$, we have $1 = 1$, when $n = 1$, we have $2 = 2$, and when $n = 3$ we have $4 = 4$. Next, let's assume for any $n = k$ that (1) holds true. Now, if we let $n = k+1$ we have from the left hand side of our expression,

$$\sum_{i=0}^{k+1} \binom{k+1}{i} = \binom{k+1}{0} + \binom{k+1}{1} + \cdots + \binom{k+1}{k} + \binom{k+1}{k+1} \tag{3.5.15}$$

$$= \binom{k}{0} + \binom{k}{-1} + \binom{k}{1} + \binom{k}{0} + \cdots + \binom{k}{k} + \binom{k}{k-1} + \binom{k}{k+1} + \binom{k}{k} \tag{3.5.16}$$

$$= 2\binom{k}{0} + 2\binom{k}{1} + \cdots + 2\binom{k}{k-1} + 2\binom{k}{k} \tag{3.5.17}$$

$$= 2\sum_{i=0}^{k} \binom{k}{i} \tag{3.5.18}$$

$$= 2(2^k) \tag{3.5.19}$$

$$= 2^{k+1}. \tag{3.5.20}$$

Therefore, we can see that for $n = k+1$ that equation (8) holds true, and thus we conclude by induction that (8) holds for all $n \in \mathbb{N}$. $\qquad\square$

Note that in the above proof, we made use of $\binom{k}{k+1} = \binom{k}{-1} = 0$. If we were to evaluate each of these using the definition of the binomial coefficient above we may notice a slight issue. Suppose we try to evaluate $\binom{n}{-1}$. Using the definition from (1), we would have

$$\binom{n}{-1} = \frac{n!}{(n-(-1))!(-1)!} \tag{3.5.21}$$

$$= \frac{n!}{(n+1)!(-1)!} \tag{3.5.22}$$

$$= \frac{n!}{(n)!(n+1)(-1)!} \tag{3.5.23}$$

$$= \frac{1}{(n+1)(-1)!}. \tag{3.5.24}$$

From above, we have a negative factorial in the denominator of our expression. Since this is not easily determined as a positive integer factorial would be, we will need to expand this using the Gamma function.

Definition 3.5.2: The Gamma function

The gamma function is a mathematical function that generalizes the concept of a factorial to a non-integer, complex number n. It is denoted by the Greek letter "Γ" (gamma) and defined as

$$\Gamma(n) = (n-1)! \equiv \int_{0}^{\infty} t^{n-1} e^{-t} \mathrm{dt}. \tag{3.5.25}$$

Using this gamma function with $n = 0$ gives

$$\Gamma(0) = \int_0^\infty t^{-1}e^{-t}\mathrm{dt}. \tag{3.5.26}$$

Since $\lim_{t \to 0^+} t^{-1}e^{-t} = \infty$, we can say the integral under the curve from 0 to ∞ will be divergent, and thus ∞. Therefore $\Gamma(0) \equiv \infty$. This allows us to write (18) as

$$\frac{1}{(n+1)(-1)!} = \frac{1}{(n+1)\Gamma(0)} = \lim_{x \to \infty} \frac{1}{x} = 0. \tag{3.5.27}$$

Thus, we can say that $\binom{n}{-1} = 0$. Similarly, by the same process, if we have $\binom{n}{n+1}$ we get

$$\binom{n}{n+1} = \frac{n!}{(n-(n+1))!(n+1)!} \tag{3.5.28}$$

$$= \frac{n!}{(-1)!(n)!(n+1)} \tag{3.5.29}$$

$$= \frac{1}{(-1)!(n+1)} \tag{3.5.30}$$

$$= 0 \tag{3.5.31}$$

Statistics and Probability

Probability is a fundamental concept in mathematics and statistics that measures the likelihood of an event occurring. It is represented as a value between 0 and 1, where 0 indicates the event is impossible, 1 denotes certainty, and values between 0 and 1 represent various degrees of likelihood. In simple terms, probability quantifies how probable or likely it is for an event to happen based on the total number of possible outcomes. It helps us make informed decisions, predict outcomes, and understand uncertainty in various real-world scenarios, such as games of chance, weather forecasts, and medical diagnoses. To understand probability on a mathematical level, some definitions of terminology is needed. '**Changes**' is a common word used, which is somewhat akin to a guess of the probability of something occuring.

A **state** (or outcome) is particular condition that something is in at a specific time. A **system** is an activity, experiment, process, or model with states or outcomes that are typically subject to uncertainty. A **sample space** of an system is the set of all possible states of a system. An **event** (also may be referred to as a trial or measurement) is any subset or collection of states contained in the sample space of a system. An event is a **simple event** if it consists of exactly one state and a **compound event** if it consistst of more than one state.

Definition 4.0.1: Probability (Axiomatic)

An **probability** p can be defined axiomatically with three postulates. The certain event E os the event that occurs in every trial. The union $A + B$ (of two events A and B) is the event that occurs when A or B or both occur. The intersection of AB of the events A and B is the event that occurs when both events A and B occur. The events A and B are mutually exclusive if the occurance of one of them excludes the occurance of the other.

1. The probability $P(s)$ of an event s is a positive number assigned to this event:

$$P(s) \geq 0 \tag{4.0.1}$$

2. The probability of the certain event is one:

$$P(E) = 1 \tag{4.0.2}$$

3. If the events A and B are mutually exclusive, then

$$P(A + B) = P(A) + P(B) \tag{4.0.3}$$

> **Definition 4.0.2: Probability (Relative Frequency)**
>
> An **probability** p can be defined as the asymptotic frequency of a system in the state s ($s \in \Omega$, where Ω is the sample space of the system) by the total number of occurrences of that state N_s in the limit of an infinite number of events N.
>
> $$p(s) = \lim_{N \to \infty} \frac{N_s}{N}, \text{ with } p(s) \in [0,1] \forall s \in \Omega \tag{4.0.4}$$
>
> For a system with n states, the total probabilities of all states must normalize to one.
>
> $$\sum_{i=0}^{n} p(i) = \sum_{s \in \Omega} p(i) = \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{n} N_i = 1 \tag{4.0.5}$$
>
> A **Bayesian probability** is defined as a person's knowledge of the outcome of a trial, based on the evidence at their disposal - often accompanied by an associated error. A **model probability** is an assumption or guess for the probability given the possibility of an infinite number of trials.

In a scenario with two possible outcomes, denoted as E_1 (Event 1) and E_2 (Event 2), each having its own probability, let's say P_1 and P_2 respectively (where $P_1 + P_2 = 1$), we can consider a series of n trials. The total number of possible sequences is $N = 2^n$ (see theorem 3.5.1) since there are only 2 possible events. The probability of obtaining Event 1 k times out of these n trials can be expressed as $P(k,n)$ and is given by the **Bernoulli probability** (also called the **binomial probability**)

$$P(k,n) = \lim_{N \to \infty} \frac{E_1}{N} = \frac{1}{2^n} \binom{n}{k} = \binom{n}{k} P_1^k P_2^{n-k} \tag{4.0.6}$$

Some more useful concepts for modeling a system are the mean (average) and standard deviation.

> **Definition 4.0.3: Arithmetic Mean**
>
> If $x_0, ..., x_N$ denotes N separate measurements of one quantity x, then we define the **arithmetic mean** (or average) $\bar{x}$ as
>
> $$\bar{x} = \frac{1}{N} \sum_{i=0}^{N} x_i \tag{4.0.7}$$

> **Definition 4.0.4: Standard Deviation**
>
> If $x_1, ..., x_N$ denotes N separate measurements of one quantity x, then we define the standard deviation σ_x as
>
> $$\sigma_x = \sqrt{\frac{1}{N-1} \sum_i (x_i - \bar{x})^2} \tag{4.0.8}$$

When dealing with **Bayesian probability**, there is typically an uncertainty associated with the measurements or functions.

Definition 4.0.5: Uncertainty

Suppose $x_0, ..., x_N$ denotes N separate measurements with associated uncertainties $\sigma_{x_0}, ..., \sigma_{x_N}$. If the measured values are used to compute some function $q(x_0, ..., x_N)$ and the uncertainties in $x_0, ..., x_N$ are independent and random, then the uncertainty in q is

$$\sigma_q = \sqrt{\sum_{i=0}^{N} \left(\frac{\partial q}{\partial x_i} \sigma_{x_i} \right)^2}. \tag{4.0.9}$$

It will always be true that

$$\sigma_q \leq \sum_{i=0}^{N} \left| \frac{\partial q}{\partial x_i} \right| \sigma_{x_i}. \tag{4.0.10}$$

Consider the uncertainty in the mean value. Definition 4.0.5 can be used to determine the uncertainty of $\bar{x}$. If we assume that all values of x_i are normally distributed, then each x_i value will have the same uncertainty, $\sigma_{x_i} \equiv \sigma_x \forall i \in \mathbb{N}$. The partial derivative of $\bar{x}$ with respect to any x_i value will simply be $\frac{\partial \bar{x}}{\partial x_i} = \frac{\partial}{\partial x_i} \frac{1}{N} \sum_{i=0}^{N} x_i = \frac{1}{N}$, since every term will be zero except the ith term. We then simply apply the definition to get

$$\sigma_{\bar{x}} = \sqrt{\sum_{i=0}^{N} \left(\frac{\partial \bar{x}}{\partial x_i} \sigma_x \right)^2} = \sqrt{\sigma_x^2 \sum_{i=0}^{N} \left(\frac{\partial \bar{x}}{\partial x_i} \right)^2} = \sqrt{\sigma_x^2 \sum_{i=0}^{N} \frac{1}{N^2}} = \sqrt{\sigma_x^2 \frac{1}{N}} = \frac{\sigma_x}{\sqrt{N}}. \tag{4.0.11}$$

Calculus

Calculus is a branch of mathematics that deals with the study of change and motion. It encompasses two main components: differentiation and integration.

- **Differentiation**: This involves finding the rate at which a quantity changes with respect to another variable. The derivative is a fundamental concept in calculus, representing the instantaneous rate of change of a function at a particular point. It helps analyze the behavior of functions, such as finding slopes of curves, determining maximum and minimum points, and understanding the concept of velocity and acceleration.

- **Integration**: Integration is the reverse process of differentiation. It involves calculating the accumulation or total amount of a changing quantity over an interval. The integral of a function represents the area under the curve of the function over a given range. It is used to solve problems involving areas, volumes, and quantities related to accumulation, such as calculating total distance traveled from a velocity function or finding the area of a region bounded by a curve.

Calculus has numerous real-world applications, ranging from physics and engineering to economics and biology. It provides essential tools for understanding how things change, predicting behavior, and solving complex problems that involve continuous change and motion. Developed independently by Isaac Newton and Gottfried Wilhelm Leibniz in the 17th century, calculus remains a fundamental and powerful branch of mathematics used in various scientific and practical domains.

5.1 Definitions and General Differential Equations

The derivative of a function represents an infinitesimal change in the function with respect to one of its variables [14].

Definition 5.1.1: Derivative

Let $f : \mathbb{R} \to \mathbb{R}$. Then the derivative of f with respect to a variable x is given by

$$\frac{d}{dx}f(x) \equiv f'(x) \equiv \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \equiv \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x}$$

Similarly, partial derivatives are defined as derivatives of a function of multiple variables when all but the variable of interest are held fixed during the differentiation [14].

Definition 5.1.2: Partial Derivative

Let $f : \mathbb{R}^z \to \mathbb{R}$, with $z \in \mathbb{N}$. Then the partial derivative of f with respect to a variable x_m is given by

$$\frac{\partial}{\partial x_m}f(x_1, \ldots, x_n) \equiv \lim_{\Delta x \to 0} \frac{f(x_1, \ldots, x_m + \Delta x, \ldots, x_n) - f(x_1, \ldots, x_m, \ldots, x_n)}{\Delta x}$$

A few rules from definition 5.1.1 can be defined.

> **Theorem 5.1.1: Distributive Property Of Derivatives**
>
> Let $g : \mathbb{R} \to \mathbb{R}$, and $h : \mathbb{R} \to \mathbb{R}$ be continuous functions with $f(x) = g(x) + h(x)$. Then the derivative of f with respect to the variable x is
>
> $$\frac{d}{dx}f(x) = \frac{d}{dx}\big(g(x) + h(x)\big) = \frac{d}{dx}g(x) + \frac{d}{dx}g(x).$$

Proof for theorem 5.1.1. Let $g : \mathbb{R} \to \mathbb{R}$, and $h : \mathbb{R} \to \mathbb{R}$ be continuous functions with $f(x) = g(x) + h(x)$. Then, by the definition of a derivative, we have

$$
\begin{aligned}
\frac{d}{dx}f(x) &= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(g+h)(x+\Delta x) - (g+h)(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{g(x+\Delta x) + h(x+\Delta x) - g(x) - h(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \left(\frac{g(x+\Delta x) - g(x)}{\Delta x} + \frac{h(x+\Delta x) - h(x)}{\Delta x} \right) \\
&= \lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x} + \lim_{\Delta x \to 0} \frac{h(x+\Delta x) - h(x)}{\Delta x} \\
&= \frac{d}{dx}g(x) + \frac{d}{dx}h(x).
\end{aligned}
$$

Thus, by simple manipulation we've shown

$$\frac{d}{dx}f(x) = \frac{d}{dx}g(x) + \frac{d}{dx}g(x).$$

$\square$

    If we want to take the derivative of a function Given a general polynomial function, we can give the following differentiation rule.

> **Theorem 5.1.2: Polynomial Derivatives**
>
> Let $f : \mathbb{R} \to \mathbb{R}$ be a function with $f(x) = c_n x^n + c_{n-1} x^{n-1} + \cdots + + C_2 x^2 + c_1 x + c_0$, where c_n is constant for all n. Then
>
> $$\frac{d}{dx}f(x) = f'(x) = nc_n x^{n-1} + (n-1)c_{n-1}x^{n-2} + \cdots + 2c_2 x + c_1$$

Proof for theorem 5.1.2. Let $f : \mathbb{R} \to \mathbb{R}$ be a function with $f(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + + C_2 x^2 + c_1 x + c_0$, where c_n is constant for all n. Let $g_n(x) = c_n x^n$. By theorem 1.1, we know that $f'(x) = g'_n(x) + g'_{n-1}(x) + \cdots + g_1(x) + g_0(x)$. Therefore, if we can prove $p(x) = Cx^n \implies p'(x) = nCx^n$, then this directly implies each term of $f(x)$ will have a derivative of equivalent form. Starting with p, we must first note that $p(x) = Cx^n$ and $p(x+h) = C(x+h)^n$. Inserting these into the limit definition will give us

$$p'(x) = \lim_{h \to 0} \frac{p(x+h) - p(x)}{h} = \lim_{h \to 0} \frac{C(x+h)^n - Cx^n}{h}.$$

As we can see, we can expand $(x + h)^n$ using the binomial theorem. This gives us

$$C \lim_{h \to 0} \frac{\sum_{k=0}^{n} \binom{n}{k} x^{n-k} h^k - x^n}{h}.$$

If we denote $a_k = \binom{n}{k} = \frac{n!}{k!(n-k)!}$, then expand the summation, we get

$$C \lim_{h \to 0} \frac{\sum_{k=0}^{n} a_k x^{n-k} h^k - x^n}{h} = C \lim_{h \to 0} \frac{a_0 x^n + a_1 x^{n-1} h + a_2 x^{n-2} h^2 + a_3 x^{n-3} h^3 + \cdots + a_k x^{n-k} h^k - x^n}{h}. \tag{5.1.1}$$

If we determine a_0 and a_1 we get

$$a_0 = \frac{n!}{0!n!} = 1 \tag{5.1.2}$$

$$a_1 = \frac{n!}{1!(n-1)!} = \frac{n!(n)}{(n-1)!(n)} = \frac{n!(n)}{n!} = n. \tag{5.1.3}$$

Inserting equations (1.2) and (1.3) into (1.1) yields

$$C \lim_{h \to 0} \frac{x^n + nx^{n-1} h + a_2 x^{n-2} h^2 + a_3 x^{n-3} h^3 + \cdots + a_k x^{n-k} h^k - x^n}{h}.$$

Then, we see that the first and last terms add to zero

$$C \lim_{h \to 0} \frac{nx^{n-1} h + a_2 x^{n-2} h^2 + a_3 x^{n-3} h^3 + \cdots + a_k x^{n-k} h^k}{h}.$$

From here, we simplify by dividing by h and take the limit as $h \to 0$

$$C \lim_{h \to 0} nx^{n-1} + a_2 x^{n-2} h + a_3 x^{n-3} h^2 + \cdots + a_k x^{n-k} h^{k-1} = Cnx^{n-1}.$$

Hence, we have determined that $p'(x) = Cnx^{n-1}$ which directly implies

$$\frac{d}{dx} f(x) = f'(x) = nc_n x^{n-1} + (n-1)c_{n-1} x^{n-2} + \cdots + 2c_2 x + c_1$$

$\square$

Consider the basic product rule of a differential equation. The product rule is given by

> **Theorem 5.1.3: Product Rule**
>
> Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(x)$ and $g(x)$ are continuous functions of the variable x. We can then say
>
> $$\frac{d}{dx} f(x)g(x) = g(x)\frac{d}{dx} f(x) + f(x)\frac{d}{dx} g(x) = f'(x)g(x) + f(x)g'(x)$$

Proof for theorem 5.1.3. Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(x)$ and $g(x)$ are continuous functions of the variable x. By the definition of the derivative, we have

$$
\begin{aligned}
\frac{d}{dx}f(x)g(x) &= \lim_{\Delta x \to 0} \frac{f(x+\Delta x)g(x+\Delta x) - f(x)g(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{f(x+\Delta x)g(x+\Delta x) - f(x+\Delta x)g(x) + f(x+\Delta x)g(x) - f(x)g(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{f(x+\Delta x)\big(g(x+\Delta x) - g(x)\big) + g(x)\big(f(x+\Delta x) - f(x)\big)}{\Delta x} \\
&= \lim_{\Delta x \to 0} f(x+\Delta x)\frac{g(x+\Delta x) - g(x)}{\Delta x} + \lim_{\Delta x \to 0} g(x)\frac{f(x+\Delta x) - f(x)}{\Delta x} \\
&= f(x)\frac{d}{dx}g(x) + g(x)\frac{d}{dx}f(x) \\
&= f'(x)g(x) + f(x)g'(x).
\end{aligned}
$$

$\square$

It is thus clear that given two continuous functions of x, say $f(x)$ and $g(x)$, the derivative follows as

$$
\frac{d}{dx}f(x)g(x) = g(x)\frac{d}{dx}f(x) + f(x)\frac{d}{dx}g(x) \tag{5.1.4}
$$

$$
= f(x)g'(x) + g(x)f'(x). \tag{5.1.5}
$$

Following this, the second derivative is given by

$$
\frac{d^2}{dx^2}f(x)g(x) = \frac{d}{dx}\left(g(x)\frac{d}{dx}f(x)\right) + \frac{d}{dx}\left(f(x)\frac{d}{dx}g(x)\right) \tag{5.1.6}
$$

$$
= \frac{d}{dx}f(x)\frac{d}{dx}g(x) + g(x)\frac{d^2}{dx^2}f(x) + f(x)\frac{d^2}{dx^2}g(x) + \frac{d}{dx}f(x)\frac{d}{dx}g(x) \tag{5.1.7}
$$

$$
= 2\frac{d}{dx}f(x)\frac{d}{dx}g(x) + g(x)\frac{d^2}{dx^2}f(x) + f(x)\frac{d^2}{dx^2}g(x) \tag{5.1.8}
$$

$$
= f(x)g''(x) + 2f'(x)g'(x) + f''(x)g(x). \tag{5.1.9}
$$

Similarly, the third derivative is given by

$$
\frac{d^3}{dx^3}f(x)g(x) = \frac{d}{dx}\left(2\frac{d}{dx}f(x)\frac{d}{dx}g(x) + g(x)\frac{d^2}{dx^2}f(x) + f(x)\frac{d^2}{dx^2}g(x)\right) \tag{5.1.10}
$$

$$
= \frac{d}{dx}\left(2\frac{d}{dx}f(x)\frac{d}{dx}g(x)\right) + \frac{d}{dx}\left(g(x)\frac{d^2}{dx^2}f(x)\right) + \frac{d}{dx}\left(f(x)\frac{d^2}{dx^2}g(x)\right). \tag{5.1.11}
$$

This derivative can be done in parts to help keep track of what is being done. Let us denote each term in the above sum as $h_1(x)$, $h_2(x)$ and $h_3(x)$ respectively. Then we have

$$
h_1(x) = \frac{d}{dx}\left(2\frac{d}{dx}f(x)\frac{d}{dx}g(x)\right) \tag{5.1.12}
$$

$$
h_2(x) = \frac{d}{dx}\left(g(x)\frac{d^2}{dx^2}f(x)\right) \tag{5.1.13}
$$

$$
h_3(x) = \frac{d}{dx}\left(f(x)\frac{d^2}{dx^2}g(x)\right). \tag{5.1.14}
$$

Simplifying these expressions then gives

$$h_1(x) = \left(2 \frac{d^2}{dx^2} f(x) \frac{d}{dx} g(x) + 2 \frac{d}{dx} f(x) \frac{d^2}{dx^2} g(x) \right) \tag{5.1.15}$$

$$h_2(x) = \left(\frac{d}{dx} g(x) \frac{d^2}{dx^2} f(x) + g(x) \frac{d^3}{dx^3} f(x) \right) \tag{5.1.16}$$

$$h_3(x) = \left(\frac{d}{dx} f(x) \frac{d^2}{dx^2} g(x) + f(x) \frac{d^3}{dx^3} g(x) \right). \tag{5.1.17}$$

Then, adding these three expressions together gives us

$$\frac{d^3}{dx^3} f(x) g(x) = f(x) \frac{d^3}{dx^3} g(x) + 3 \frac{d}{dx} f(x) \frac{d^2}{dx^2} g(x) + 3 \frac{d^2}{dx^2} f(x) \frac{d}{dx} g(x) + g(x) \frac{d^3}{dx^3} f(x) \tag{5.1.18}$$

$$= f(x) g'''(x) + 3 f'(x) g''(x) + 3 f''(x) g'(x) + f'''(x) g(x). \tag{5.1.19}$$

Using the product rule, a nice pattern seems to be emerging as we continually apply it to the same function. From equations (5.1.5), (5.1.9) and (5.1.19), we can see that it appears to be a similar pattern as when applying the binomial expansion theorem. As a reminder, the binomial expansion theorem is

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k \tag{5.1.20}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \tag{5.1.21}$$

If we observe the first three terms of this (with $n = 1, 2, 3$), we have

$$(a + b)^1 = a + b \tag{5.1.22}$$

$$(a + b)^2 = a^2 + 2ab + b^2 \tag{5.1.23}$$

$$(a + b)^3 = a^3 + 3a^2 b + 3ab^2 + b^3. \tag{5.1.24}$$

Comparing this to the three equations in questions allows us to notice interesting similarities. If we expand this to n we have

$$(a + b)^n = a^n b^0 + \binom{n}{1} a^{n-1} y^1 + \cdots + \binom{n}{n-1} ab^{k-1} + a^0 b^k. \tag{5.1.25}$$

Now, let $a = \frac{d}{dx} g(x)$ and $b = \frac{d}{dx} f(x)$. If we temporarily define[1] the following conditions

$$\left(\frac{d}{dx} h(x) \right)^n := \frac{d^n}{dx^n} h(x) \tag{5.1.26}$$

$$\frac{d}{dx} g(x) + \frac{d}{dx} f(x) := \frac{d}{dx} g(x) f(x), \tag{5.1.27}$$

then inserting a and b into equation (5.1.25) gives

$$\left[\frac{d}{dx} g(x) + \frac{d}{dx} f(x) \right]^n = \left(\frac{d}{dx} g(x) \right)^n \left(\frac{d}{dx} f(x) \right)^0 + \binom{n}{1} \left(\frac{d}{dx} g(x) \right)^{n-1} \left(\frac{d}{dx} f(x) \right)^1 + \cdots \tag{5.1.28}$$

$$\cdots + \binom{n}{n-1} \left(\frac{d}{dx} g(x) \right) \left(\frac{d}{dx} f(x) \right)^{n-1} + \left(\frac{d}{dx} g(x) \right)^0 \left(\frac{d}{dx} f(x) \right)^n. \tag{5.1.29}$$

---

[1]This is solely for an easy way of noticing the similarities in the formulas, these equations do not necessarily hold and are not used once the equations are expanded.

Simplifying the above expression gives us

$$\frac{d^n}{dx^n}g(x)f(x) = f(x)\frac{d^n}{dx^n}g(x) + \binom{n}{1}\frac{d^{n-1}}{dx^{n-1}}g(x)\frac{d}{dx}f(x) + \cdots \tag{5.1.30}$$

$$\cdots + \binom{n}{n-1}\frac{d}{dx}g(x)\frac{d^{n-1}}{dx^{n-1}}f(x) + g(x)\frac{d^n}{dx^n}f(x). \tag{5.1.31}$$

Therefore, we can see that with our re-definitions in equations (5.1.26) and (5.1.27), it appears that the nth derivative of something using the product rule can be expressed using the binomial expansion theorem. At least to $n = 3$ we can see from above that this is accurate. If we want to determine if this works for all values of n, we can try a proof by induction. Before going any further, since this may be a large amount of work for such a simple idea, let us quickly have a sanity check. We claim that the terms of the product rule can be expanded using the binomial theorem. It is important to note that when taking the product rule of an expression, you get two new expressions. Then, taking the product rule again will double those two into four, and then those four into eight, and so on. So we can see that each product rule doubles the number of terms we have. The binomial coefficient is used when we have like terms to combine them. If we were to add up each binomial coefficient in the sum, we should then see that each sum of coefficients is the same number of terms we would get as if we doubled our expression each time. If this were not so then our situation would have a flaw. In other words, it must hold true that

$$\sum_{i=0}^{n}\binom{n}{k} = 2^n. \tag{5.1.32}$$

This indeed is true. See theorem 3.5.1. Now, we know from equations (5.1.5), (5.1.9) and (5.1.19) that cases $n = 1, 2, 3$ all hold. Next, if we assume that for any $n = k$ that

$$\frac{d^k}{dx^k}g(x)f(x) = f(x)\frac{d^k}{dx^k}g(x) + \binom{k}{1}\frac{d^{k-1}}{dx^{k-1}}g(x)\frac{d}{dx}f(x) + \cdots$$

$$\cdots + \binom{k}{k-1}\frac{d}{dx}g(x)\frac{d^{k-1}}{dx^{k-1}}f(x) + g(x)\frac{d^k}{dx^k}f(x) \tag{5.1.33}$$

$$= \sum_{i=0}^{k}\binom{k}{i}\left[\frac{d^i}{dx^i}f(x)\right]\left[\frac{d^{k-i}}{dx^{k-i}}g(x)\right] \tag{5.1.34}$$

holds true, we can then determine if it holds true for $n = k + 1$. This gives

$$\frac{d^{k+1}}{dx^{k+1}}g(x)f(x) = f(x)\frac{d^{k+1}}{dx^{k+1}}g(x) + \binom{k+1}{1}\frac{d^k}{dx^k}g(x)\frac{d}{dx}f(x) + \cdots$$

$$\cdots + \binom{k+1}{k}\frac{d}{dx}g(x)\frac{d^k}{dx^k}f(x) + g(x)\frac{d^{k+1}}{dx^{k+1}}f(x) \tag{5.1.35}$$

$$= \sum_{i=0}^{k+1}\binom{k+1}{i}\left[\frac{d^i}{dx^i}f(x)\right]\left[\frac{d^{k+1-i}}{dx^{k+1-i}}g(x)\right]. \tag{5.1.36}$$

If we were to manually evaluate the $n = k + 1$ term of this expression we would have

$$\frac{d}{dx}\frac{d^k}{dx^k}g(x)f(x) = \frac{d}{dx}\sum_{i=0}^{k}\binom{k}{i}\left[\frac{d^i}{dx^i}f(x)\right]\left[\frac{d^{k-i}}{dx^{k-i}}g(x)\right]. \tag{5.1.37}$$

Since differentiating is distributive (by theorem 5.1.1), we can evaluate each term of this sum and compare it to (32) and (33). Doing this, we can use the product rule to first get

$$\frac{d^{k+1}}{dx^{k+1}}g(x)f(x) = \sum_{i=0}^{k}\binom{k}{i}\left[\frac{d^{i+1}}{dx^{i+1}}f(x)\right]\left[\frac{d^{k-i}}{dx^{k-i}}g(x)\right] + \binom{k}{i}\left[\frac{d^i}{dx^i}f(x)\right]\left[\frac{d^{k+1-i}}{dx^{k+1-i}}g(x)\right]. \tag{5.1.38}$$

From here, writing out the series explicitly allows us to combine terms in a useful matter. This would obviously be tedious to write out so let us denote each term by a_i, for any i^{th} term in the summation. Writing out a few of the terms then gives

$$a_0 = \binom{k}{0}\left[\frac{d^1}{dx^1}f(x)\right]\left[\frac{d^k}{dx^k}g(x)\right] + \binom{k}{0}f(x)\left[\frac{d^{k+1}}{dx^{k+1}}g(x)\right] \tag{5.1.39}$$

$$a_1 = \binom{k}{1}\left[\frac{d^2}{dx^2}f(x)\right]\left[\frac{d^{k-1}}{dx^{k-1}}g(x)\right] + \binom{k}{1}\left[\frac{d}{dx}f(x)\right]\left[\frac{d^k}{dx^k}g(x)\right] \tag{5.1.40}$$

$$a_2 = \binom{k}{2}\left[\frac{d^3}{dx^3}f(x)\right]\left[\frac{d^{k-2}}{dx^{k-2}}g(x)\right] + \binom{k}{2}\left[\frac{d^2}{dx^2}f(x)\right]\left[\frac{d^{k-1}}{dx^{k-1}}g(x)\right] \tag{5.1.41}$$

$$\vdots$$

$$a_{k-2} = \binom{k}{k-2}\left[\frac{d^{k-1}}{dx^{k-1}}f(x)\right]\left[\frac{d^2}{dx^2}g(x)\right] + \binom{k}{k-2}\left[\frac{d^{k-2}}{dx^{k-2}}f(x)\right]\left[\frac{d^3}{dx^3}g(x)\right] \tag{5.1.42}$$

$$a_{k-1} = \binom{k}{k-1}\left[\frac{d^k}{dx^k}f(x)\right]\left[\frac{d}{dx}g(x)\right] + \binom{k}{v}\left[\frac{d^{k-1}}{dx^{k-1}}f(x)\right]\left[\frac{d^2}{dx^2}g(x)\right] \tag{5.1.43}$$

$$a_k = \binom{k}{k}\left[\frac{d^{k+1}}{dx^{k+1}}f(x)\right]g(x) + \binom{k}{k}\left[\frac{d^k}{dx^k}f(x)\right]\left[\frac{d}{dx}g(x)\right]. \tag{5.1.44}$$

At this point, it becomes obvious as to why the $f^{(n)}(x)$ notation of a derivative was invented, and switching over will greatly help in our next step. Changing our notation gives us

$$a_0 = \binom{k}{0}f^{(1)}(x)g^{(k)}(x) + \binom{k}{0}f(x)g^{(k+1)}(x) \tag{5.1.45}$$

$$a_1 = \binom{k}{1}f^{(2)}(x)g^{(k-1)}(x) + \binom{k}{1}f^{(1)}(x)g^{(k)}(x) \tag{5.1.46}$$

$$a_2 = \binom{k}{2}f^{(3)}(x)g^{(k-2)}(x) + \binom{k}{2}f^{(2)}(x)g^{(k-1)}(x) \tag{5.1.47}$$

$$\vdots$$

$$a_{k-2} = \binom{k}{k-2}f^{(k-1)}(x)g^{(2)}(x) + \binom{k}{k-2}f^{(k-2)}(x)g^{(3)}(x) \tag{5.1.48}$$

$$a_{k-1} = \binom{k}{k-1}f^{(k)}(x)g^{(1)}(x) + \binom{k}{k-1}f^{(k-1)}(x)g^{(2)}(x) \tag{5.1.49}$$

$$a_k = \binom{k}{k}f^{(k+1)}(x)g(x) + \binom{k}{k}f^{(k)}(x)g^{(1)}(x). \tag{5.1.50}$$

Our end goal is to get this expression to be in the same form as in equation (5.1.34). If we first re-write equation (5.1.34) using $f^{(n)}(x)$ notation, we can easily see what our end goal is. This is

$$\sum_{i=0}^{k+1}\binom{k+1}{i}f^{(i)}(x)g^{(k+1-i)}(x). \tag{5.1.51}$$

The next step is then summing expressions (5.1.45) through (5.1.50), then determining if what we get is equivalent to equation (5.1.51). Let $\Upsilon = a_0 + a_1 + \cdots + a_{k-1} + a_k$, then we have

$$\Upsilon = \binom{k}{0} f(x) g^{(k+1)}(x) + \binom{k}{0} f^{(1)}(x) g^{(k)}(x) + \binom{k}{1} f^{(1)}(x) g^{(k)}(x) + \cdots$$

$$\cdots + \binom{k}{k} f^{(k)}(x) g^{(1)}(x) + \binom{k}{k-1} f^{(k)}(x) g^{(1)}(x) + \binom{k}{k} f^{(k+1)}(x) g(x) \tag{5.1.52}$$

$$= \binom{k}{0} f(x) g^{(k+1)}(x) + \left[\binom{k}{0} + \binom{k}{1}\right] f^{(1)}(x) g^{(k)}(x) + \cdots$$

$$\cdots + \left[\binom{k}{k-1} + \binom{k}{k}\right] f^{(k)}(x) g^{(1)}(x) + \binom{k}{k} f^{(k+1)}(x) g(x) \tag{5.1.53}$$

$$= \sum_{i=0}^{k+1} \left[\binom{k}{i-1} + \binom{k}{i}\right] f^{(i)}(x) g^{(k+1-i)}(x). \tag{5.1.54}$$

Using the definition of the binomial coefficient and the derivation in (7), if we let $k = m - 1 \implies m = k + 1$, we can re-write the binomial coefficient in equation (5.1.54) as follows.

$$\binom{k}{i-1} + \binom{k}{i} = \binom{m-1}{i-1} + \binom{m-1}{i} = \binom{m}{i} = \binom{k+1}{i}. \tag{5.1.55}$$

Thus, equation (5.1.54) becomes

$$\Upsilon = \sum_{i=0}^{k+1} \binom{k+1}{i} f^{(i)}(x) g^{(k+1-i)}(x). \tag{5.1.56}$$

Now, $\Upsilon = a_0 + a_1 + \cdots + a_{k-1} + a_k = (35)$ which we have shown is equivalent to equation (5.1.54). Therefore, we can conclude that (28) holds for all $n \in \mathbb{N}$, which allows us to tidy our work up as a theorem.

> ### Theorem 5.1.4: Repeating Product Rule
>
> Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be continuous and differentiable functions of the variable x. Then the mth derivative of $f(x)g(x)$ with respect to x is
>
> $$\frac{d^m}{dx^m} f(x) g(x) = \sum_{i=0}^{n} \binom{n}{i} \left[\frac{d^i}{dx^i} f(x)\right] \left[\frac{d^{n-i}}{dx^{n-i}} g(x)\right].$$

Consider a division of two functions rather than multiplication. The product rule can be used to derive what is referred to as the **quotient rule**. Begin by taking two arbitrary functions that are continuous and differentiable $f(x)$ and $g(x)$. What we want to find is $\frac{d}{dx} \frac{f(x)}{g(x)}$. We begin by applying the product rule from theorem 5.1.3, which gives

$$\frac{d}{dx} \frac{f(x)}{g(x)} = f(x) \left[\frac{d}{dx} \frac{1}{g(x)}\right] + \frac{1}{g(x)} \left[\frac{d}{dx} f(x)\right]. \tag{5.1.57}$$

Let $u(x) = \frac{1}{g(x)}$. We want to determine $u'(x)$ in order to simplify equation (5.1.57). From the definition of $u(x)$, we can then say $u(x)g(x) = 1$. The derivative of a constant value is just zero, so if we differentiate both sides (by yet again applying the product rule from theorem 5.1.3), we get

$$\frac{d}{dx} u(x) g(x) = \frac{d}{dx} 1 \implies u(x) g'(x) + u'(x) g(x) = 0. \tag{5.1.58}$$

Solving this for $u'(x)$ gives us $\frac{d}{dx}\frac{1}{g(x)}$ which we is then equal to

$$u'(x) = \frac{d}{dx}u(x) = \frac{d}{dx}\frac{1}{g(x)} = \frac{-u(x)g'(x)}{g(x)} = -\frac{g'(x)}{g(x)^2}. \tag{5.1.59}$$

By taking equation (5.1.59) and substituting it into equation (5.1.57), we get

$$\frac{d}{dx}\frac{f(x)}{g(x)} = f(x)\left[-\frac{g'(x)}{g(x)^2}\right] + \frac{1}{g(x)}f'(x) = -\frac{f(x)g'(x)}{g(x)^2} + \frac{g(x)f'(x)}{g(x)^2} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \tag{5.1.60}$$

Theorem 5.1.5: Quotient Rule

Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be continuous and differentiable functions of the variable x. Then the derivative of $\frac{f(x)}{g(x)}$ with respect to x is

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \tag{5.1.61}$$

A special case of this quotient rule is when the numerator is 1. This can be seen in equation (5.1.57). For simplicity, I will drop the "(x)" and assume functions with a single variable here. Taking consecutive derivatives of this give

$$\frac{d}{dx}\frac{1}{g} = \frac{-g'}{g^2} \tag{5.1.62}$$

$$\frac{d^2}{dx^2}\frac{1}{g} = \frac{d}{dx}\frac{-g'}{g^2} = -\frac{d}{dx}g'\frac{1}{g^2} = -\frac{g''}{g^2} - g'\left(\frac{-2gg'}{g^4}\right) = -\frac{g''}{g^2} + \frac{2g'^2}{g^3} \tag{5.1.63}$$

$$\frac{d^3}{dx^3}\frac{1}{g} = \frac{d}{dx}\left(-\frac{g''}{g^2} + \frac{2g'^2}{g^3}\right) = -\frac{g'''}{g^2} - g''\frac{d}{dx}\frac{1}{g^2} + \frac{4g'g''}{g^3} + 2g'^2\frac{d}{dx}\frac{1}{g^3}$$

$$= -\frac{g'''}{g^2} - g''\left(\frac{-2gg'}{g^4}\right) + \frac{4g'g''}{g^3} + 2g'^2\left(\frac{-3g^2g'}{g^6}\right)$$

$$= -\frac{g'''}{g^2} + \frac{2g'g''}{g^3} + \frac{4g'g''}{g^3} - \frac{6g'^3}{g^4}$$

$$= \frac{6g'g''}{g^3} - \frac{6g'^3}{g^4} - \frac{g'''}{g^2} \tag{5.1.64}$$

$$\frac{d^4}{dx^4}\frac{1}{g} = \frac{d}{dx}\left(\frac{6g'g''}{g^3} - \frac{6g'^3}{g^4} - \frac{g'''}{g^2}\right)$$

$$= 6g'g''\frac{d}{dx}\frac{1}{g^3} + \frac{1}{g^3}\frac{d}{dx}6g'g'' - 6g'^3\frac{d}{dx}\frac{1}{g^4} - \frac{1}{g^4}\frac{d}{dx}6g'^3 - g'''\frac{d}{dx}\frac{1}{g^2} - \frac{1}{g^2}\frac{d}{dx}g'''$$

$$= \frac{-18g'^2g''}{g^4} + \frac{6g'g''' + 6g''^2}{g^3} + \frac{24g'^4}{g^5} - \frac{18g'^2g''}{g^4} + \frac{2g'g'''}{g^3} - \frac{g^{(4)}}{g^2}$$

$$= \frac{-36g'^2g''}{g^4} + \frac{8g'g'''}{g^3} + \frac{6g''^2}{g^3} + \frac{24g'^4}{g^5} - \frac{g''''}{g^2} \tag{5.1.65}$$

$$\frac{d^5}{dx^5}\frac{1}{g} = \frac{-g^{(5)}}{g^2} - \frac{120g'^5}{g^6} + \frac{10g^{(4)}g'}{g^3} + \frac{20g'''g''}{g^3} - \frac{60g'''g'^2}{g^4} + \frac{240g'^3g''}{g^5} - \frac{90g'g''^2}{g^4}, \tag{5.1.66}$$

and so on. Notice on that last one I spared the algebra. to find a pattern in this sequence, a few things can be noted. First. The factor of $\left(\frac{1}{g}\right)^m$ is common in every term. The maximum m value here appears to be

equal to $m = n + 1$, where n is the number of derivatives performed. There also seem to be at least one term for each m value except for $m = 1$ (for example, in equation (5.1.62) there is only a g^2 term, with a maximum $m = 2$, but in equation (5.1.66), there are $m = 2, 3, 4, 5,$ and 6 terms). Next, there is a common occurrence of $\frac{d^k}{dx^k}g$ terms, where the highest k value appears to be $k = n$ in each derivative. There does not seem to be an immediate pattern with the coefficients. For each $\frac{d^n}{dx^n}$, there is also a $-\frac{g^{(n)}}{g^2}$ term.

Using these observations, we can begin to construct a potential formula to represent this pattern. It must be a sum of some kind, since it has summed terms. It must contain one term for each $\left(\frac{1}{g}\right)^m$ for all $1 \leq m \leq n+1$. Each term will have some coefficient, which will likely be dependent on m, n, or both m and n. And each term will have some product of $g^{(i)}$ terms. We can thus guess our formula to look something like

$$\frac{d^n}{dx^n}\frac{1}{g} \stackrel{?}{=} \sum_{m=0}^{n} \frac{C(n,m)}{g^{m+1}}G(g^{(i)}|i \leq n \in \mathbb{N}^+), \tag{5.1.67}$$

where $C(n,m)$ is some coefficient function dependent on n and m, and $G(g^{(i)}|i \leq n \in \mathbb{N}^+) \equiv G(g^{(1)}, g^{(2)}, ..., g^{(n)})$ is some function which is a product of the repeated derivatives of $g(x)$ in some combination. Some possible steps to determine this formula better would of course be to then determine G and C. We know the above equations involve deriving $g(x)$. With a careful and lucky observation, G can be determined. Consider this derivative for a moment $\frac{d}{dx}g(x) = g'(x)$. If we take consecutive derivatives, we simply get $\frac{d^i}{dx^i}g(x) = g^{(i)}(x)$. Now consider this derivative for a moment $\frac{d}{dx}g(x)^2 = 2g(x)g'(x)$. When there are multiple terms of our function, the consecutive derivatives require the product rule, which would give $\frac{d^2}{dx^2}g(x)^2 = 2g(x)g''(x) + 2g'(x)^2$. Careful observation will reveal that this corresponds to the derivatives in equation (5.1.63), where as the first derivative of $g(x)$ corresponds to the coefficients in equation (5.1.62). Taking a third derivative of $g(x)^2$ gives $\frac{d^3}{dx^3}g(x)^2 = 2g(x)g'''(x) + 6g'(x)g''(x)$. This only closely correlates to equation (5.1.64). If instead we take the third derivative of $g(x)^3$, we have $\frac{d^3}{dx^3}g(x)^3 = 3(g(x)^2g'''(x) + 2g'(x)^3 + 6g'(x)g''(x)$, which indeed correlates to the coefficients in equation (5.1.64). This pattern appears to repeat. The excess g terms in these will interfere with our g^{m+1} denominator in equation (5.1.67), and so we should modify this to be some function, say $q(x, n, m)$ (It has a dependency on m and n because it's value may depend on which term in the summation it is), and determine if it needs to be changed. The pattern of G reveals we can potentially make G a function of n (and more technically n and x since g is a function of x) to get our correct factors of $g^{(i)}$ in each term.

$$G(g^{(i)}|i \leq n \in \mathbb{N}^+) \stackrel{?}{\equiv} G(n,x) \stackrel{?}{=} \frac{\partial^n}{\partial x^n}g(x)^n. \tag{5.1.68}$$

Inserting this into equation (5.1.67) gives

$$\frac{d^n}{dx^n}\frac{1}{g} \stackrel{?}{=} \sum_{m=0}^{n} \frac{C(n,m)}{q(x,n,m)}\frac{\partial^n}{\partial x^n}g(x)^n. \tag{5.1.69}$$

To determine q, we can write out the first few expansions of G.

$$\frac{d}{dx}g(x) = g'(x) \tag{5.1.70}$$

$$\frac{d^2}{dx^2}g(x)^2 = 2g(x)g''(x) + 2g'(x)^2 \tag{5.1.71}$$

$$\frac{d^3}{dx^3}g(x)^3 = 3(g(x)^2g'''(x) + 2g'(x)^3 + 6g'(x)g''(x) \tag{5.1.72}$$

$$\frac{d^4}{dx^4}g(x)^4 = 4(6g'(x)^4 + g(x)^3g^{(4)}(x) + 9g(x)^2g''(x)^2 + 12g(x)^2g'(x)g^{(3)}(x) + 36g(x)g'(x)^2g''(x)) \tag{5.1.73}$$

Using this and beginning with $n = 1$, we have

$$\frac{d}{dx}\frac{1}{g} = \sum_{m=0}^{1} \frac{C(1,m)}{q(x,1,m)} \frac{\partial}{\partial x} g(x) = \frac{C(1,0)}{q(x,1,0)} g'(x) + \frac{C(1,1)}{q(x,1,1)} g'(x). \tag{5.1.74}$$

We can let $C(1,0) = 0$, which makes the first term vanish, $q(x,1,1) = g^2$, and $C(1,1) = -1$. For $n = 2$ we have

$$\frac{d^2}{dx^2}\frac{1}{g} \stackrel{?}{=} \sum_{m=0}^{2} \frac{C(2,m)}{q(x,2,m)} \frac{\partial^2}{\partial^2 x} g(x)^2 \tag{5.1.75}$$

$$= \left[\frac{C(2,0)}{q(x,2,0)} + \frac{C(2,1)}{q(x,2,1)} + \frac{C(2,m)}{q(x,2,2)} \right] \left(2g(x)g''(x) + 2g'(x)^2 \right) \tag{5.1.76}$$

$$= \frac{C(2,0)}{q(x,2,0)} \left(2gg'' + 2g'^2 \right) + \frac{C(2,1)}{q(x,2,2)} \left(2gg''(x) + 2g'^2 \right) + \frac{C(2,m)}{q(x,2,2)} \left(2gg'' + 2g'^2 \right). \tag{5.1.77}$$

From this we can compare coefficients to equation 5.1.63 and thus we have

$$2g \left[\frac{C(2,0)}{q(x,2,0)} + \frac{C(2,1)}{q(x,2,1)} + \frac{C(2,2)}{q(x,2,2)} \right] = \frac{-1}{g^2} \tag{5.1.78}$$

$$2 \left[\frac{C(2,0)}{q(x,2,0)} + \frac{C(2,1)}{q(x,2,1)} + \frac{C(2,2)}{q(x,2,2)} \right] = \frac{2}{g^3} \tag{5.1.79}$$

We can again make the first $m = 0$ term have a zero constant ($C(2,0 = 0)$) causing it to vanish, then these become

$$\frac{C(2,1)}{q(x,2,1)} + \frac{C(2,2)}{q(x,2,2)} = \frac{-1}{2g^3} \tag{5.1.80}$$

$$\frac{C(2,1)}{q(x,2,1)} + \frac{C(2,2)}{q(x,2,2)} = \frac{1}{g^3} \tag{5.1.81}$$

From this, $q(x,2,m) = \frac{1}{g^3}$, $C(2,1) + C(2,2) = \frac{-1}{2}$ and $C(2,1) + C(2,2) = 1$. {\color{red}This cannot be true and is in fact a contradiction. Therefore we must conclude that equation (5.1.69) is incorrect!}. One thing of noteworthiness is that in the above process, if the steps were repeated for more n values, there would always be each set of equations with the same coefficients leading to this contradiction. This is because G did not depend on anything other than n giving the same factor for each term of the sum. To remedy this, we can add an m dependence to G and allow the coefficient variables to cancel out the un-used terms. As long as we have the maximum n value, we will have some equation that gives the correct $g^{(i)}$ terms and then we can make the other unwanted terms cancel. Thus, we can try to modify G such that $G = \frac{\partial^n}{\partial x^n} g(x)^m$. This gives

$$\frac{d^n}{dx^n}\frac{1}{g} \stackrel{?}{=} \sum_{m=0}^{n} \frac{C(n,m)}{q(x,n,m)} \frac{\partial^n}{\partial x^n} g(x)^m. \tag{5.1.82}$$

We then try the same process again. Let $n = 1$. We have

$$\frac{d}{dx}\frac{1}{g} = \sum_{m=0}^{1} \frac{C(1,m)}{q(x,1,m)} \frac{\partial}{\partial x} g(x)^m = \frac{C(1,0)}{q(x,1,0)} \frac{\partial}{\partial x} 1 + \frac{C(1,1)}{q(x,1,1)} \frac{\partial}{\partial x} g(x) = \frac{C(1,1)}{q(x,1,1)} g'(x). \tag{5.1.83}$$

Notice that the initial term automatically canceled out this time leaving us with just one term. This will always be the case since the derivative of a constant is zero. Therefore we can conclude the summation can

instead start from $m = 1$. If we equate this to equation (5.1.62), we have $C(1, 1) = -1$ and $q(x, 1, 1) = g^2$. Now doing the same for $n = 2$ gives

$$\frac{d^2}{dx^2}\frac{1}{g} \stackrel{?}{=} \sum_{m=1}^{2} \frac{C(2, m)}{q(x, 2, m)} \frac{\partial^2}{\partial^2 x} g(x)^m \tag{5.1.84}$$

$$= \frac{C(2, 1)}{q(x, 2, 1)} \frac{\partial^2}{\partial^2 x} g(x) + \frac{C(2, 2)}{q(x, 2, 2)} \frac{\partial^2}{\partial^2 x} g(x)^2 \tag{5.1.85}$$

$$= \frac{C(2, 1)}{q(x, 2, 1)} g''(x) + \frac{C(2, 2)}{q(x, 2, 2)} \left(2g(x)g''(x) + 2g'(x)^2\right). \tag{5.1.86}$$

Pulling out our coefficients and equating this to equation (5.1.63) yields

$$\frac{C(2, 1)}{q(x, 2, 1)} + 2g\frac{C(2, 2)}{q(x, 2, 2)} = \frac{-1}{g^2} \tag{5.1.87}$$

$$2\frac{C(2, 2)}{q(x, 2, 2)} = \frac{2}{g^3}. \tag{5.1.88}$$

Notice how we have have equations that are more of an echelon-like[2] format. We can solve for $C(2, 2)$ and $q(x, 2, 2)$ using equation (5.1.88) giving $C(2, 2) = 1$ and $q(x, 2, 2) = g(x)^3$. Plugging these into equation (5.1.87) then gives $C(2, 1) = -3$ and $q(x, 2, 1) = g(x)^2$. Now doing the same for $n = 3$ gives

$$\frac{d^3}{dx^3}\frac{1}{g} \stackrel{?}{=} \sum_{m=1}^{3} \frac{C(3, m)}{q(x, 3, m)} \frac{\partial^3}{\partial^3 x} g(x)^m \tag{5.1.89}$$

$$= \frac{C(3, 1)}{q(x, 3, 1)} \frac{\partial^3}{\partial^3 x} g(x) + \frac{C(3, 2)}{q(x, 3, 2)} \frac{\partial^3}{\partial^3 x} g(x)^2 + \frac{C(3, 3)}{q(x, 3, 3)} \frac{\partial^3}{\partial^3 x} g(x)^3 \tag{5.1.90}$$

$$= \frac{C(3, 1)}{q(x, 3, 1)} g^{(3)} + \frac{C(3, 2)}{q(x, 3, 2)} \left(2gg^{(3)} + 6g'g''\right) + \frac{C(3, 3)}{q(x, 3, 3)} \left(3g^2 g^{(3)} + 6g'^3 + 18gg'g''\right). \tag{5.1.91}$$

Pulling out our coefficients and equating this to equation (5.1.64) yields

$$\frac{C(3, 1)}{q(x, 3, 1)} + 2g\frac{C(3, 2)}{q(x, 3, 2)} + 3g^2\frac{C(3, 3)}{q(x, 3, 3)} = \frac{-1}{g^2} \tag{5.1.92}$$

$$6\frac{C(3, 2)}{q(x, 3, 2)} + 18g\frac{C(3, 3)}{q(x, 3, 3)} = \frac{6}{g^3} \tag{5.1.93}$$

$$6\frac{C(3, 3)}{q(x, 3, 3)} = \frac{-6}{g^4}. \tag{5.1.94}$$

From this (and using a little bit of algebra) it's clear that $q(x, 3, 3) = g^4$, $q(x, 3, 2) = g^3$, $q(x, 3, 1) = g^2$, $C(3, 3) = -1$, $C(3, 2) = 4$, and $C(3, 1) = -6$. Now doing the same for $n = 3$ gives

$$\frac{d^4}{dx^4}\frac{1}{g} \stackrel{?}{=} \sum_{m=1}^{4} \frac{C(4, m)}{q(x, 4, m)} \frac{\partial^4}{\partial^4 x} g(x)^m \tag{5.1.95}$$

$$= \frac{C(4, 1)}{q(x, 4, 1)} \frac{\partial^4}{\partial^4 x} g(x) + \frac{C(4, 2)}{q(x, 4, 2)} \frac{\partial^4}{\partial^4 x} g(x)^2 + \frac{C(4, 3)}{q(x, 4, 3)} \frac{\partial^4}{\partial^4 x} g(x)^3 + \frac{C(4, 4)}{q(x, 4, 4)} \frac{\partial^4}{\partial^4 x} g(x)^4 \tag{5.1.96}$$

$$= \frac{C(4, 1)}{q(x, 4, 1)} g^{(4)} + \frac{C(4, 2)}{q(x, 4, 2)} \left(2gg^{(4)} + 6g''^2 + 8g^{(3)}g'\right) + \frac{C(4, 3)}{q(x, 4, 3)} \Big(3g^2 g^{(4)} + 18gg''^2$$

$$+ 24gg'g^{(3)} + 36g'^2 g''\Big) + \frac{C(4, 4)}{q(x, 4, 4)} \left(4g^3 g^{(4)} + 36g^2 g''^2 + 24g'^4 + 48g^2 g^{(3)}g' + 144gg'^2 g''\right) \tag{5.1.97}$$

---

[2]If we imagine the fraction as the variables, then this is echelon format. See definition 6.1.3.

Pulling out our coefficients and equating this to equation (5.1.65) yields

$$\frac{C(4,1)}{q(x,4,1)} + 2g\frac{C(4,2)}{q(x,4,2)} + 3g^2\frac{C(4,3)}{q(x,4,3)} + 4g^3\frac{C(4,4)}{q(x,4,4)} = \frac{-1}{g^2} \tag{5.1.98}$$

$$6\frac{C(4,2)}{q(x,4,2)} + 18g\frac{C(4,3)}{q(x,4,3)} + 36g^2\frac{C(4,4)}{q(x,4,4)} = \frac{6}{g^3} \tag{5.1.99}$$

$$8\frac{C(4,2)}{q(x,4,2)} + 24g\frac{C(4,3)}{q(x,4,3)} + 48g^2\frac{C(4,4)}{q(x,4,4)} = \frac{8}{g^3} \tag{5.1.100}$$

$$36\frac{C(4,3)}{q(x,4,3)} + 144g\frac{C(4,4)}{q(x,4,4)} = \frac{-36}{g^4} \tag{5.1.101}$$

$$24\frac{C(4,4)}{q(x,4,4)} = \frac{24}{g^5} \tag{5.1.102}$$

From this (and using a little bit of algebra) it's clear that $q(x,4,4) = g^5$, $q(x,4,3) = g^4$, $q(x,4,2) = g^3$, $q(x,4,1) = g^2$, $C(4,4) = 1$, $C(4,3) = -5$, $C(4,2) = 10$, and $C(4,1) = -10$. Fortunately, the same values satisfy both equations (5.1.99) and (5.1.100) - indicating that our equation (5.1.82) is more likely true.

It can already be seen that there is an obvious pattern for q. For all four n values tested, q has only been dependent on m. Furthermore, it can be observed that $q(x,n,m) \equiv q(x,m) = g^{m+1}$, which interestingly enough is the guess we had from equation (5.1.67). If accurate, the only thing left to do is determine the coefficient formula for $C(n,m)$. We can visualize $C(n,m)$ as a $n \times m$ matrix $\mathbf{C}$, where each row represents a single n value and each column represents a single m value, and insert the values we have determined above in an attempt to identify a pattern. Note we will start indexing both m and n at 1 and leave all undefined values at zero. Making all undefined values zero would make the coefficients zero thus canceling out the terms if they were to exist. Without taking the time to write out the work in LATEXfor the $n > 4$ cases, this gives

$$\mathbf{C} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ -3 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ -6 & 4 & -1 & 0 & 0 & 0 & 0 & \cdots \\ -10 & 10 & -5 & 1 & 0 & 0 & 0 & \cdots \\ -15 & 20 & -15 & 6 & -1 & 0 & 0 & \cdots \\ -21 & 35 & -35 & 21 & -7 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{5.1.103}$$

An immediate observation is that each column alternates from negative to positive. Therefore, since the columns represent changing m values, we can say there is likely a $(-1)^m$ term in $C(n,m)$. For $n = 1$, there appears to be a difference $C(1, m+1) = C(1,m) - (m+1)$ for each value. For $n = 2$ there isn't as obvious of a pattern. Let N_m represent the absolute value of each non-zero[3] m values from the columns of $\mathbf{C}$ as a set. Then we have

$$N_1 = \{1, 3, 6, 10, 15, 21, ...\} \tag{5.1.104}$$
$$N_2 = \{1, 4, 10, 20, 35, ...\} \tag{5.1.105}$$
$$N_3 = \{1, 5, 15, 35, ...\} \tag{5.1.106}$$
$$N_4 = \{1, 6, 21, ..\} \tag{5.1.107}$$
$$N_5 = \{1, 7, ...\} \tag{5.1.108}$$

---

[3]This is important. By taking only the non-zero values we are 'omitting' certain values of n from our pattern. This will need to be kept in mind throughout this process.

Our coefficients $C(n,m)$ are then given by $N_m(n)$. The first term of each set $N_i \forall i \in \mathbb{N}$ is one. The second term appears to be $N_{m+1}(2) = N_m(2) + N_m(1)$. the third term appears to be $N_{i+1}(3) = N_i(3) + N_i(2)$. The fourth term appears to be $N_{i+1}(4) = N_i(4) + N_i(3)$, and so on. This pattern is familiar and can be visualized using pascal's triangle (see section 3.4.1). Take the list of N_i sets and visually rotate them by $-\frac{\pi}{4}$. This gives a pascal-like chart of

$$
\begin{array}{ccccccc}
 & & & 1 & & & \\
 & & 1 & & 3 & & \\
 & 1 & & 4 & & 6 & \\
1 & & 5 & & 10 & & 10 \\
\end{array}
$$

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 3 & & & \\
 & & 1 & & 4 & & 6 & & \\
 & 1 & & 5 & & 10 & & 10 & \\
1 & & 6 & & 15 & & 20 & & 15 \\
\end{array}
$$

$$
1 \quad 7 \quad 21 \quad 35 \quad 35 \quad 21
$$

In the above triangle, starting from the top of the triangle heading down and to the left does *not* represent increasing n values (because we took only non-zero values of $C(n,m)$ when representing N_i), while heading down and to the left *does* represents increasing m values for our system. Since these only represent the $non-zero$ values of $C(n,m)$, the corresponding n values are in fact the horizontal rows in this triangle rather than diagonal rows[4]. This must be the case since $n \geq m$ also must be satisfied. The coefficients of Pascal's Triangle can be represented by the binomial coefficients $\binom{n}{m}$ and thus we can perhaps represent these values using this as a factor in $C(n,m)$. Writing each term in terms of its corresponding bonimial coefficient gives

$$
\begin{array}{ccccccccccc}
 & & & & & \frac{2}{2}\binom{1}{1} & & & & & \\
 & & & & \frac{3}{3}\binom{2}{2} & & \frac{3}{2}\binom{2}{1} & & & & \\
 & & & \frac{4}{4}\binom{3}{3} & & \frac{4}{3}\binom{3}{2} & & \frac{6}{3}\binom{3}{1} & & & \\
 & & \frac{5}{5}\binom{4}{4} & & \frac{5}{4}\binom{4}{3} & & \frac{5}{3}\binom{4}{2} & & \frac{5}{2}\binom{4}{1} & & \\
 & \frac{6}{6}\binom{5}{5} & & \frac{6}{5}\binom{5}{4} & & \frac{6}{4}\binom{5}{3} & & \frac{6}{3}\binom{5}{2} & & \frac{6}{2}\binom{5}{1} & \\
\frac{7}{7}\binom{6}{6} & & \frac{7}{6}\binom{6}{5} & & \frac{7}{3}\binom{6}{4} & & \frac{7}{4}\binom{6}{3} & & \frac{7}{5}\binom{6}{2} & & \frac{7}{2}\binom{6}{1} \\
\end{array}
$$

I wrote each term in the above triangle as a fraction to further demonstrate the obvious pattern. The coefficients multiplied by each binomial coefficient here is $\frac{n+1}{m+1}$. We know from previous steps that $C(n,m)$ has a $(-1)^m$ term, a $\binom{n}{m}$ term, and finally a $\frac{n+1}{m+1}$ term. Combining these gives $C(n,m) = (-1)^m \frac{n+1}{m+1}\binom{n}{m}$. Therefore our equation (5.1.82) becomes clear.

> **Lemma 5.1.1: repeating Differentiation of $\frac{1}{g(x)}$**
>
> Let $g : \mathbb{R} \to \mathbb{R}$ be a continuous and differential function over the variable x such that $g(x) \neq 0$ for all

---

[4]This can be seen in equation (5.1.103)

$x \in \mathbb{R}$. Then the nth derivative of $\frac{1}{g(x)}$ is

$$\frac{d^n}{dx^n}\frac{1}{g(x)} = \sum_{m=1}^{n}(-1)^m \frac{n+1}{m+1}\binom{n}{m}\frac{1}{g(x)^{m+1}}\frac{\partial^n}{\partial x^n}g(x)^m. \tag{5.1.109}$$

Proof for Lemma 5.1.1. Let $\partial_x^n \equiv \frac{\partial^n}{\partial x^n}$ and $g \equiv g(x)$ for simplicity. Since this equation involves a summation over n, we can do a proof by induction. From the above derivation of this equation, the first few cases have already been shown to hold true. Assume that for some $n = k$, that

$$\partial_x^k \frac{1}{g} = \sum_{m=1}^{k}(-1)^m \frac{k+1}{m+1}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^k g^m. \tag{5.1.110}$$

Consider $n = k+1$. This is simply the derivative with respect to x of equation 5.1.110.

$$\partial_x^{k+1}\frac{1}{g} = \partial_x \partial_x^k \frac{1}{g} \tag{5.1.111}$$

$$= \partial_x \sum_{m=1}^{k}(-1)^m \frac{k+1}{m+1}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^k g^m \tag{5.1.112}$$

$$= \sum_{m=1}^{k}(-1)^m \frac{k+1}{m+1}\binom{k}{m}\partial_x\left[\frac{1}{g^{m+1}}\partial_x^k g^m\right] \tag{5.1.113}$$

The product rule in theorem 5.1.3 can be applied to the derivative in equation (5.1.113).

$$\partial_x\left[\frac{1}{g^{m+1}}\partial_x^k g^m\right] = \frac{1}{g^{m+1}}\partial_x^{k+1}g^m - \frac{(m+1)g'}{g^{m+2}}\partial_x^k g^m \tag{5.1.114}$$

Plugging this back into equation (5.1.113) gives

$$\partial_x^{k+1}\frac{1}{g} = \sum_{m=1}^{k}(-1)^m \frac{k+1}{m+1}\binom{k}{m}\left[\frac{1}{g^{m+1}}\partial_x^{k+1}g^m - \frac{(m+1)g'(x)}{g^{m+2}}\partial_x^k g^m\right] \tag{5.1.115}$$

If lemma 5.1.1 is true, then equation (5.1.113) should be equal to what we get when we insert $n = k+1$ (call it D) which is

$$D = \sum_{m=1}^{k+1}(-1)^m \frac{k+2}{m+1}\binom{k+1}{m}\frac{1}{g^{m+1}}\partial_x^{k+1}g^m \tag{5.1.116}$$

Therefore, we simply need to show that equation (5.1.113) equals (5.1.116) for the $n = k+1$ case to hold. Analyzing the coefficients in equation (5.1.116), we can say

$$\frac{k+2}{m+1}\binom{k+1}{m} = \frac{(k+2)(k+1)!}{(m+1)m!(k-m+1)} = \frac{(k+2)(k+1)k!}{(m+1)m!(k-m)!(k-m+1)} = \frac{(k+2)(k+1)}{(m+1)(k-m+1)}\binom{k}{m} \tag{5.1.117}$$

Plugging this back in gives,

$$D = \sum_{m=1}^{k+1}(-1)^m \frac{(k+2)(k+1)}{(m+1)(k-m+1)}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^{k+1}g^m \tag{5.1.118}$$

Pulling out the $k+1$ term from the summation gives

$$D = \sum_{m=1}^{k} (-1)^m \frac{(k+2)(k+1)}{(m+1)(k-m+1)} \binom{k}{m} \frac{1}{g^{m+1}} \partial_x^{k+1} g^m + (-1)^{k+1} \frac{1}{g^{k+2}} \partial_x^{k+1} g^{k+1} \tag{5.1.119}$$

Looking at the coefficients of the sum again, $\frac{(k+2)(k+1)}{(m+1)(k-m+1)}\binom{k}{m}$, we can add some value to this to get $\frac{k+1}{m+1}\binom{k}{m}$. For each term of the sum, we can add some value and then subtract that same value and not modify the true value (simple arithmetic). Looking at the value X that would accomplish this gives

$$\frac{(k+2)(k+1)}{(m+1)(k-m+1)}\binom{k}{m} + X = \frac{k+1}{m+1}\binom{k}{m} \implies X = -\frac{k+1}{k-m+1}\binom{k}{m} \tag{5.1.120}$$

Using this, we can add a sum of X terms and subtract the same sum of X terms from equation (5.1.119) to get

$$D = \underbrace{\sum_{m=1}^{k}(-1)^m \frac{k+1}{m+1}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^{k+1}g^m}_{D_0} + \underbrace{\sum_{m=1}^{k}(-1)^m \frac{k+1}{k-m+1}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^{k+1}g^m + (-1)^{k+1}\frac{1}{g^{k+2}}\partial_x^{k+1}g^{k+1}}_{D_1}$$
$$\tag{5.1.121}$$

Since this is equivalent to equation (5.1.116), and we want to show that this equals equation (5.1.113), we can set them equal to each-other and find the condition that must hold for equality. We can do this because our inductive step is assumed to be true, therefore equation (5.1.115), which is simply the derivative of our inductive step, is true. In this case I will use equation (5.1.115) which is an alternate form of equation (5.1.113). Thus comparing equation (5.1.121) to $\partial_x^k \frac{1}{g}$ from equation (5.1.115), we have

$$D_1 = \sum_{m=1}^{k}(-1)^m \frac{k+1}{k-m+1}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^{k+1}g^m + (-1)^{k+1}\frac{1}{g^{k+2}}\partial_x^{k+1}g^{k+1} \tag{5.1.122}$$

$$= -\sum_{m=1}^{k}(-1)^m \frac{k+1}{m+1}\binom{k}{m}\frac{(m+1)g'}{g^{m+2}}\partial_x^k g^m \tag{5.1.123}$$

Simplifying this gives.

$$\sum_{m=1}^{k+1}(-1)^m \binom{k+1}{m}\frac{1}{g^{m+1}}\partial_x^{k+1}g^m = -\sum_{m=1}^{k}(-1)^m (k+1)\binom{k}{m}\frac{g'}{g^{m+2}}\partial_x^k g^m \tag{5.1.124}$$

Therefore if we can show this is true, it would suffice for our proof. Alternatively, this may just be an interesting and potentially useful result. Let $u_0 = \frac{1}{g^{m+1}}$ and $u_1 = \partial_x^k g^m$. Then it follows that $u_0' = -\frac{(m+1)g'}{g^{m+2}}$ and $u_1'(x) = \partial_x^{k+1}g^m$. Implementing the product rule from theorem 5.1.3, we have

$$\partial_x u_0 u_1 = u_0' u_1 + u_0 u_1' \implies u_0' u_1 = \partial_x u_0 u_1 - u_0 u_1' \tag{5.1.125}$$

$$-\frac{(m+1)g'}{g^{m+2}}\partial_x^k g^m = \partial_x \left(\frac{1}{g^{m+1}}\partial_x^k g^m \right) - \frac{1}{g^{m+1}}\partial_x^{k+1}g^m \tag{5.1.126}$$

Applying the product rule to equation (5.1.123), we have

$$\sum_{m=1}^{k}(-1)^m \frac{k+1}{k-m+1}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^{k+1}g^m + (-1)^{k+1}\frac{1}{g^{k+2}}\partial_x^{k+1}g^{k+1}$$

$$= \sum_{m=1}^{k}(-1)^m \frac{k+1}{m+1}\binom{k}{m}\left[\partial_x\left(\frac{1}{g^{m+1}}\partial_x^k g^m \right) - \frac{1}{g^{m+1}}\partial_x^{k+1}g^m \right] \tag{5.1.127}$$

simplifying this gives

$$\sum_{m=1}^{k}(-1)^m \frac{(k+1)(k+2)}{(m+1)(k-m+1)}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^{k+1}g^m + (-1)^{k+1}\frac{1}{g^{k+2}}\partial_x^{k+1}g^{k+1}$$

$$= \partial_x \sum_{m=1}^{k}(-1)^m \frac{k+1}{m+1}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^k g^m \qquad (5.1.128)$$

Looking at the summation coefficients once more we have

$$\frac{(k+1)(k+2)}{(m+1)(k-m+1)}\binom{k}{m} = \frac{(k+1)(k+2)k!}{(m+1)(k-m+1)m!(k-m)!} = \frac{(k+2)(k+1)!}{(m+1)m!(k-m+1)!} = \frac{k+2}{m+1}\binom{k+1}{m}$$

$$(5.1.129)$$

Plugging this back in to equation (5.1.128) gives

$$\sum_{m=1}^{k}(-1)^m \frac{k+2}{m+1}\binom{k+1}{m}\frac{1}{g^{m+1}}\partial_x^{k+1}g^m + (-1)^{k+1}\frac{1}{g^{k+2}}\partial_x^{k+1}g^{k+1} = \partial_x\sum_{m=1}^{k}(-1)^m\frac{k+1}{m+1}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^k g^m$$

$$(5.1.130)$$

Then finally, combining the terms on the left hand side of this equation gives

$$\sum_{m=1}^{k+1}(-1)^m \frac{k+2}{m+1}\binom{k+1}{m}\frac{1}{g^{m+1}}\partial_x^{k+1}g^m = \partial_x\sum_{m=1}^{k}(-1)^m\frac{k+1}{m+1}\binom{k}{m}\frac{1}{g^{m+1}}\partial_x^k g^m \qquad (5.1.131)$$

Therefore, for the condition mentioned earlier in equation (5.1.123) to hold true, equation (5.1.131) has to be true. Put another way, equation (5.1.113) equals (5.1.116), which is what we had to show to prove lemma 5.1.1.

Note. I'm not sure this proof is valid. I essentially say A is true, derive B from it (which must be true). Then analyzing some C, and assuming $C = B$, find that $C = B \implies C = B$, which is not saying anything!

❗ Section in Progress ❗

$\square$

5.2   Ordinary Linear Homogeneous Differential Equations

Consider the following linear homogeneous differential equation with constants A and B

$$\ddot{x} + 2A\dot{x} + Bx = 0. \qquad (5.2.1)$$

We can similarly write this as

$$\frac{d^2}{dt^2}x(t) + \frac{d}{dt}2Ax(t) + Bx(t) = 0. \qquad (5.2.2)$$

Next, we can simplify the expression by dividing both sides by $x(t)$. Note that this can only be done when $x(t) \neq 0$. We get

$$\frac{d^2}{dt^2} + \frac{d}{dt}2A + B = 0. \qquad (5.2.3)$$

Next, notice that the expression is of a quadratic form $af^2 + bf + c = 0$, with $f = \frac{d}{dt}$, $a = 1$, $b = 2A$, and $c = B$. Thus, using the quadratic formula we can say

$$f = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \iff \frac{d}{dt} = \frac{-2A \pm \sqrt{(2A)^2 - 4(B)}}{2}. \tag{5.2.4}$$

From here, we can simplify the expression to achieve

$$\frac{d}{dt} = \left(-A \pm \sqrt{A^2 - B} \right). \tag{5.2.5}$$

Notice that we have a differential on the left side of the expression, yet no function. To fix this, suppose we multiply both sides by $x(t)$, which is our original function of t in our differential equation. Then we get

$$\frac{d}{dt}x(t) = \left(-A \pm \sqrt{A^2 - B} \right)x(t). \tag{5.2.6}$$

In a case where we differentiate a function, and get the same function multiplied by a constant, we would have $x(t) = C_1 e^{kt}$ as a solution, where C_1 and k are constants. Note that taking $\frac{d}{dt}x(t)$ in this case yields $kC_1 e^{kt}$, which if matched to our differential yields $k = -A \pm \sqrt{A^2 - B}$. Therefore, we can say a solution to our differential equation will be of the form

$$x(t) = C_1 e^{(-A \pm \sqrt{A^2 - B})t}, \tag{5.2.7}$$

or if written more generally as two separate solutions $x_1(t)$ and $x_2(t)$, we have

$$x_1(t) = C_1 e^{(-A + \sqrt{A^2 - B})t} \tag{5.2.8}$$

$$x_2(t) = C_2 e^{(-A - \sqrt{A^2 - B})t} \tag{5.2.9}$$

Note that we must change the constant coefficients in front of each solution to be arbitrary related to each other because both possible cases are a unique solution and each case would be independent of the other. Finally, if we want to verify these solution, we can do so by differentiating twice and plugging our functions into our original differential equation. Doing this (while letting $k = -A + \sqrt{A^2 - B}$ and keeping our $x(t)$ in the general form of Ce^{kt}), we achieve

$$x(t) = Ce^{kt} \tag{5.2.10}$$

$$\frac{d}{dt}x(t) = (-A \pm \sqrt{A^2 - B})Ce^{kt} \tag{5.2.11}$$

$$\frac{d^2}{dt^2}x(t) = (-A \pm \sqrt{A^2 - B})^2 Ce^{kt} \tag{5.2.12}$$

$$= (A^2 \mp 2A\sqrt{A^2 - B} + (A^2 - B))Ce^{kt}. \tag{5.2.13}$$

Plugging these expressions into our differential equation gives us

$$(A^2 \mp 2A\sqrt{A^2 - B} + (A^2 - B))Ce^{kt} + 2A(-A \pm \sqrt{A^2 - B})Ce^{kt} + BCe^{kt} = 0. \tag{5.2.14}$$

As we can see, each term has a common factor of Ce^{kt}, so we can divide both sides by Ce^{kt} giving

$$(A^2 \mp 2A\sqrt{A^2 - B} + (A^2 - B)) + 2A(-A \pm \sqrt{A^2 - B}) + B = 0. \tag{5.2.15}$$

This removes all of the C constants which verifies that they can be arbitrary and the function will still work. From here, we can just apply algebraic manipulation to achieve

$$A^2 \mp 2A\sqrt{A^2 - B} + A^2 - B - 2A^2 \pm 2A\sqrt{A^2 - B} + B = 0. \tag{5.2.16}$$

Which we can clearly see is a true statement by canceling terms. Thus, we can say the following:

> **Theorem 5.2.1**
>
> Let A, B and C_n be constants and $x(t)$ be a function of t. Given any homogeneous ordinary differential equation of the form
>
> $$\frac{d^2}{dt^2}x(t) + \frac{d}{dt}2Ax(t) + Bx(t) = 0,$$
>
> There will exist a solution of the form
>
> $$x(t) = C_1 e^{(-A \pm \sqrt{A^2 - B})t}.$$

If we further the thought of this differential equation having more than one solution, we can easily understand why. First, consider an ordinary linear homogeneous differential equation given in terms of a function of t $(x(t))$ with constants C_n of the form

$$\frac{d^n}{dt^n}x(t) + \frac{d^{n-1}}{dt^{n-1}}C_{n-1}x(t) + \cdots + \frac{d}{dt}C_1 x(t) + C_0 x(t) = 0. \tag{5.2.17}$$

Assume there exists multiple solutions to this equation, say $x_1(t), x_2(t), \ldots, x_n(t)$ and let $x_g(t) = x_1(t) + x_2(t) + \cdots + x_n(t)$. Notice that differentiating $x_g(t)$ with respect to t can be written as (By Theorem 1.1)

$$\frac{d}{dt}x_g(t) = \frac{d}{dt}\big(x_1(t) + x_2(t) + \cdots + x_n(t)\big) \tag{5.2.18}$$

$$= \frac{d}{dt}x_1(t) + \frac{d}{dt}x_2(t) + \cdots + \frac{d}{dt}x_n(t). \tag{5.2.19}$$

Similarly, differentiating a second time gives

$$\frac{d^2}{dt^2}x_g(t) = \frac{d}{dt}\left(\frac{d}{dt}x_1(t) + \frac{d}{dt}x_2(t) + \cdots + \frac{d}{dt}x_n(t)\right) \tag{5.2.20}$$

$$= \frac{d^2}{dt^2}x_1(t) + \frac{d^2}{dt^2}x_2(t) + \cdots + \frac{d^2}{dt^2}x_n(t). \tag{5.2.21}$$

And so forth so that differentiating n times gives us

$$\frac{d^n}{dt^n}x_g(t) = \frac{d^n}{dt^n}x_1(t) + \frac{d^n}{dt^n}x_2(t) + \cdots + \frac{d^n}{dt^n}x_n(t). \tag{5.2.22}$$

Thus, looking at our original differential equation, since we initially claimed that $x_n(t)$ is a solution to our differential equation, we can say

$$\frac{d^n}{dt^n}x_n(t) + \frac{d^{n-1}}{dt^{n-1}}C_{n-1}x_n(t) + \cdots + \frac{d}{dt}C_1 x_n(t) + C_0 x_n(t) = 0 \tag{5.2.23}$$

$$\frac{d^n}{dt^n}x_{n-1}(t) + \frac{d^{n-1}}{dt^{n-1}}C_{n-1}x_{n-1}(t) + \cdots + \frac{d}{dt}C_1 x_{n-1}(t) + C_0 x_{n-1}(t) = 0 \tag{5.2.24}$$

$$\vdots$$

$$\frac{d^n}{dt^n}x_0(t) + \frac{d^{n-1}}{dt^{n-1}}C_{n-1}x_0(t) + \cdots + \frac{d}{dt}C_1 x_n(t) + C_0 x_0(t) = 0 \tag{5.2.25}$$

Furthermore, we can add these solutions together and factor the expression to get

$$\frac{d^n}{dt^n}\big(x_n(t) + \cdots x_0(t)\big) + \cdots + \frac{d}{dx}C_1\big(x_n(t) + \cdots x_0(t)\big) + C_0\big(x_n(t) + \cdots x_0(t)\big) = 0, \tag{5.2.26}$$

or simply

$$\frac{d^n}{dt^n}x_g(t) + \frac{d^{n-1}}{dt^{n-1}}C_{n-1}x_g(t) + \cdots + \frac{d}{dt}C_1x_g(t) + C_0x_g(t) = 0. \tag{5.2.27}$$

Thus, we have showed that a sum of solutions to this form of linear homogeneous differential equation is also a solution to the equation. This means we can say

---

Theorem 5.2.2

Let $x(t)$ be a function of t and C_n be constant for all n. Given any ordinary homogeneous differential equation of the form

$$\frac{d^n}{dt^n}x(t) + \frac{d^{n-1}}{dt^{n-1}}C_{n-1}x(t) + \cdots + \frac{d}{dt}C_1x(t) + C_0x(t) = 0,$$

if n solutions $x_n(t)$ exist, then another solution can be formed from summing any combination of those solutions and $x_g(t)$ will be the most general solution with

$$x_g(t) = \sum_{k=1}^{n} x_k(t).$$

5.3   Linear Ordinary Differential Equations

Now let us consider some differential equations that are not homogeneous and see if we can come up with a similar way of deriving a solution. Again, let $x(t)$ be some function of t with A and B as constants. In addition, let $g(t)$ be a function of t as well. Similar to earlier, consider the following differential equation

$$\frac{d^2}{dt^2}x(t) + \frac{d}{dt}2Ax(t) + Bx(t) = g(t) \tag{5.3.1}$$

In order for this equation to be satisfied, we must first notice that whatever function $g(t)$ is in, the left hand of the expression must also be that form.

5.3.1   n-Degree Polynomial Particular Functions

For example, let's consider $g(t) = at^2 + bt + c$, where $a, b,$ and c are constants. We can see then that if the equation is satisfied, we will have t^2, t and constant terms on the left side after differentiating $x(t)$ the appropriate number of times and plugging it in. Let us work through this example completely. Consider

$$\frac{d^2}{dt^2}x(t) + \frac{d}{dt}2Ax(t) + Bx(t) = at^2 + bt + c. \tag{5.3.2}$$

We can start by considering theorem 1.2 which tells us that the first derivative of $g(t)$ will no longer have any t^2 terms, and the second derivative will no longer have any t or t^2 terms. Therefore, by a careful observation, we can see that if we let $x(t)$ be a second degree polynomial, then $\frac{d^2}{dx^2}x(t)$ will be a constant polynomial and $\frac{d}{dx}x(t)$ will be a polynomial of degree 1. This allows us to see that the degree of the left side of this expression (given that $x(t)$ is a polynomial of degree 2) will be of degree 2 as well. Hence, by choosing $x(t) = C_1t^2 + C_2t + C_3$, we will end up getting a second degree polynomial equal to a second degree

polynomial, so choosing appropriate constants will allow this to be a solution. Let's demonstrate and check our answer by choosing

$$x(t) = C_1 t^2 + C_2 t + C_3. \tag{5.3.3}$$

From here, we have the appropriate derivatives as

$$\frac{d}{dx} x(t) = 2C_1 t + C_2 \tag{5.3.4}$$

$$\frac{d^2}{dx^2} x(t) = 2C_1. \tag{5.3.5}$$

Then, substituting this into our differential equation yields

$$2C_1 + 2A(2C_1 t + C_2) + B(C_1 t^2 + C_2 t + C_3) = at^2 + bt + c. \tag{5.3.6}$$

Consolidating terms then gives

$$(BC_1)t^2 + (BC_2 + 4AC_1)t + (2C_1 + 2AC_2 + BC_3) = at^2 + bt + c. \tag{5.3.7}$$

From here Since these must be equal, the coefficients in front of the t^2 terms must be equal, and so forth. This then gives us the following simple set of equations to solve:

$$BC_1 = a \tag{5.3.8}$$

$$BC_2 + 4AC_1 = b \tag{5.3.9}$$

$$2C_1 + 2AC_2 + BC_3 = c. \tag{5.3.10}$$

Solving for $C_1, C_2, and C_3$ then gives

$$C_1 = \frac{a}{B} \tag{5.3.11}$$

$$C_2 = \frac{b}{B} - \frac{4Aa}{B^2} = \frac{bB - 4Aa}{B^2} \tag{5.3.12}$$

$$C_3 = \frac{8A^2 a}{B^3} - \frac{2a}{B^2} - \frac{2Ab}{B^2} + \frac{c}{B} = \frac{8aA^2 - 2aB - 2AbB + B^2 c}{B^3}. \tag{5.3.13}$$

Therefore, a solution that would work for our differential is

$$x(t) = \frac{a}{B} t^2 + \frac{bB - 4Aa}{B^2} t + \frac{8aA^2 - 2aB - 2AbB + B^2 c}{B^3}. \tag{5.3.14}$$

It is simple enough to verify that this is indeed a solution. To begin, we can take the derivative twice, and plug the function in. Taking the first and second derivatives give us

$$\frac{d}{dx} x(t) = \frac{2a}{B} t + \frac{bB - 4Aa}{B^2} \tag{5.3.15}$$

$$\frac{d^2}{dx^2} x(t) = \frac{2a}{B}. \tag{5.3.16}$$

Then, plugging these into our equation gives

$$\frac{2a}{B} + 2A\left(\frac{2a}{B} t + \frac{bB - 4Aa}{B^2}\right) + B\left(\frac{a}{B} t^2 + \frac{bB - 4Aa}{B^2} t + \frac{8aA^2 - 2aB - 2AbB + B^2 c}{B^3}\right), \tag{5.3.17}$$

which should simplify to $at^2 + bt + c$. Foiling terms and canceling like terms gives

$$\frac{Ba}{B}t^2 + \frac{bB^2}{B^2}t + \frac{B^3c}{B^3} \equiv at^2 + bt + c. \tag{5.3.18}$$

Thus, we have verified that our solution is valid. Now by the same principles used in this above example, we can generalize the form of our solution. Let C_n and a_ℓ be constant for all n, ℓ. Consider any differential equation of the form

$$\frac{d^n}{dt^n}x(t) + \frac{d^{n-1}}{dt^{n-1}}C_{n-1}x(t) + \cdots + \frac{d}{dt}C_1x(t) + C_0x(t) = g(t) \tag{5.3.19}$$

$$g(t) = a_\ell t^\ell + a_{\ell-1}t^{\ell-1} + \cdots + a_1t + a_0. \tag{5.3.20}$$

By the same reasoning as before, the left side of the equation will equate to a polynomial of degree less than or equal to ℓ. This allows us to let $x(t)$ equal an arbitrary polynomial of degree ℓ similarly to above, and then we would be able to solve for the coefficients in order to make the equation a solution. Thus, let b_ℓ be constant for all ℓ and

$$x(t) = b_\ell t^\ell + b_{\ell-1}t^{\ell-1} + \cdots + b_4t^4 + b_3t^3 + b_2t^2 + b_1t + b_0. \tag{5.3.21}$$

We would then solve for our coefficients b_ℓ in the same manner as above so they are in terms of a_ℓ and C_n, for some n, ℓ. Let's look at what happens when we do this. First, let's differentiate $x(t)$ an appropriate number of times. This gives

$$\frac{d}{dt}x(t) = \ell b_\ell t^{\ell-1} + (\ell-1)b_{\ell-1}t^{\ell-2} + \cdots + 4b_4t^3 + 3b_3t^2 + 2b_2t + b_1 \tag{5.3.22}$$

$$\frac{d^2}{dt^2}x(t) = \ell(\ell-1)b_\ell t^{\ell-2} + (\ell-1)(\ell-2)b_{\ell-1}t^{\ell-3} + \cdots + 12b_4t^2 + 6b_3t + 2b_2 \tag{5.3.23}$$

$$\frac{d^3}{dt^3}x(t) = \ell(\ell-1)(\ell-2)b_\ell t^{\ell-3} + (\ell-1)(\ell-2)(\ell-3)b_{\ell-1}t^{\ell-4} + \cdots + 24b_4t + 6b_3 \tag{5.3.24}$$

$$\vdots$$

$$\frac{d^\ell}{dt^\ell}x(t) = \ell!b_\ell. \tag{5.3.25}$$

As we can see, a pattern begins to emerge that allows us to write the mth derivative as

$$\frac{d^m}{dt^m}x(t) = \frac{\ell!}{(\ell-m)!}b_\ell t^{\ell-m} + \frac{(\ell-1)!}{(\ell-m-1)!}b_{\ell-1}t^{\ell-m-1} + \cdots + (m+1)!b_{m+1}t + m!b_m. \tag{5.3.26}$$

Since ℓ is fixed as the degree of $g(t)$, we can write this m^{th} derivative expression as a sum of terms from above. This then gives us

$$\frac{d^m}{dx^m}x(t) = \sum_{k=0}^{\ell-m} \frac{(\ell-k)!}{(\ell-m-k)!}b_{\ell-k}t^{\ell-m-k}. \tag{5.3.27}$$

Thus, to formalize this into something useful, we can state it as a theorem.

Theorem 5.3.1

Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(t) \in \mathbb{P}_n$ of the form $f(t) = b_nt^n + b_{n-1}t^{n-1} + \cdots + +b_1t + b_0$. Then the mth

derivative of $f(t)$ is given by

$$\frac{d^m}{dx^m} f(t) = f^{(m)}(t) = \sum_{k=0}^{n-m} \frac{(n-k)!}{(n-m-k)!} b_{n-k} t^{n-m-k}.$$

What we have done above is find a solution to the differential equation that correlates to an expression on the right hand side of our equation. This solution is called a particular solution and is generally denoted with a subscript of p, such as $x_p(t)$. We have also shown that it is fairly simple to determine a particular solution when our differential equation equates to a polynomial. In some cases, the particular solution is not as simple to compute. Since $g(t)$ is a generic function of t, we know it can be in the form of any possible function. As it is seemingly impossible to determine the infinite possible particular solutions for all of these differential equations, we will try to generalize a few and make some rules for combining solutions.

5.3.2 Simple Exponential Particular Functions

An even easier case for particular solutions when we have exponential functions. Consider the case where $g(t) = ae^{kt}$, with a and k being constants. Then we have may have some differential of a form similar to

$$\frac{d^2}{dt^2} x(t) + \frac{d}{dt} 2Ax(t) + Bx(t) = ae^{kt}. \tag{5.3.28}$$

When given a homogeneous differential equation, it is sometimes possible to derive a solution in a completely analytical way. In some cases for particular solutions especially, it is easiest to make a keen observation, start with a possible solution, and verify from there. The most important observation here is that when an exponential has a derivative of the same form as the original exponential only with a different constant out front. Notice for example that $\frac{d}{dt} g(t) = ake^{kt}$ and $\frac{d^2}{dt^2} g(t) = ak^2 e^{kt}$. Thus, if $x(t)$ were to also be an exponential of the form be^{kt} (with b being a constant), we would have all terms on the left side of our expression in the form of a constant multiplied by e^{kt}. Therefore, let us choose $x(t) = be^{kt}$ and check to see if it could be a solution. First, we have

$$\frac{d}{dt} x(t) = bke^{kt} \tag{5.3.29}$$

$$\frac{d^2}{dt^2} x(t) = bk^2 e^{kt} \tag{5.3.30}$$

Then substituting this into our differential yields

$$bk^2 e^{kt} + 2ACke^{kt} + Bbke^{kt} = ae^{kt}. \tag{5.3.31}$$

From here, we can divide out the e^{kt}, factor and solve for our constant as follows

$$b(k^2 + 2Ak + B) = a \implies b = \frac{a}{k^2 + 2Ak + B}. \tag{5.3.32}$$

Hence, we have that the particular solution to our differential equation is

$$x_p(t) = \frac{ae^{kt}}{k^2 + 2Ak + B}. \tag{5.3.33}$$

Checking this solution is trivial and can be done simply by taking the first and second derivatives and plugging into our differential equation. The differential equation should then evaluate as true. As we can see, exponential particular solutions are very simple, and if we were to have something such as $g(t) = a_1 e^{kt} + a_2 e^{kt}$,

then we would solve the particular in exactly the same manner. This is because we would be able to factor and say $g(t) = (a_1 + a_2)e^{kt}$, with $a_1 + a_2$ being our constant in $g(t)$. Let us consider a more general case however. If for example we have $g(t) = a_\ell e^{k_\ell t} + a_{\ell-1} e^{k_{\ell-1} t} + \cdots + a_1 e^{k_1 t} + a_0 e^{k_0 t}$, where a_ℓ and k_ℓ are constant, we would use the same methods, but a combination of multiple solutions. As theorem 1.1 states, we can see that $g(t)$ is a combination of functions $h_i(t) = a_i e^{k_i t}$, and thus, the derivatives can be taken separately then added. A good way to represent this is

$$g(t) = \sum_{i=0}^{\ell} h_i(t) = \sum_{i=0}^{\ell} a_i e^{k_i t} \tag{5.3.34}$$

Consider as well the following differential equation again:

$$\frac{d^n}{dt^n} x(t) + \frac{d^{n-1}}{dt^{n-1}} C_{n-1} x(t) + \cdots + \frac{d}{dt} C_1 x(t) + C_0 x(t) = g(t). \tag{5.3.35}$$

Similarly, this can be written as

$$\frac{d^n}{dt^n} x(t) + \frac{d^{n-1}}{dt^{n-1}} C_{n-1} x(t) + \cdots + \frac{d}{dt} C_1 x(t) + C_0 x(t) = \sum_{i=0}^{\ell} a_i e^{k_i t}. \tag{5.3.36}$$

Let $g_\ell(t)$ denote the corresponding $b_\ell e^{k_\ell t}$ term. We know by the above example that if we were to choose $x_n(t) = b_n e^{k_n t}$, it would be a solution for $g_\ell(t) = a_\ell e^{k_\ell t}$. Similarly, if we were to choose $x(t) = x_0(t) + x_1(t) = b_0 e^{k_1 t} + b_0 e^{k_1 t}$, we can show that it is a solution to $g(t) = g_0(t) + g_1(t) = a_0 e^{k_0 t} + a_1 e^{k_1 t}$ using theorem 1.1. First we say

$$\frac{d^n}{dt^n} \big(x_0(t) + x_1(t) \big) + \cdots + \frac{d}{dt} C_1 \big(x_0(t) + x_1(t) \big) + C_0 \big(x_0(t) + x_1(t) \big) = g_0(t) + g_1(t). \tag{5.3.37}$$

This then becomes

$$\frac{d^n}{dt^n} x_0(t) + \frac{d^n}{dt^n} x_1(t) + \cdots + \frac{d}{dt} C_1 x_0(t) + \frac{d}{dt} C_1 x_1(t) + C_0 x_0(t) + C_0 x_1(t) = g_0(t) + g_1(t). \tag{5.3.38}$$

Then by grouping these we have

$$\left(\frac{d^n}{dt^n} x_0(t) + \cdots + \frac{d}{dt} C_1 x_0(t) + C_0 x_0(t) \right) + \left(\frac{d^n}{dt^n} x_1(t) + \cdots + \frac{d}{dt} C_1 x_1(t) + C_0 x_1(t) \right) = g_0(t) + g_1(t).$$

As we can see, this is a solution since $x_0(t)$ is a solution to

$$\frac{d^n}{dt^n} x_0(t) + \cdots + \frac{d}{dt} C_1 x_0(t) + C_0 x_0(t) = g_0(t) \tag{5.3.39}$$

and $x_1(t)$ is a solution to

$$\frac{d^n}{dt^n} x_1(t) + \cdots + \frac{d}{dt} C_1 x_1(t) + C_0 x_1(t) = g_1(t). \tag{5.3.40}$$

Similarly, this method will work for any number for terms in $g(t)$. Therefore, if we had

$$g(t) = \sum_{i=0}^{\ell} a_i e^{k_i t}, \tag{5.3.41}$$

We will have a solution of the form

$$x(t) = \sum_{i=0}^{\ell} b_i e^{k_i t}. \tag{5.3.42}$$

We can then verify this by plugging our above equations into the differential as follows:

$$\frac{d^n}{dt^n} \sum_{i=0}^{\ell} b_i e^{k_i t} + \frac{d^{n-1}}{dt^{n-1}} C_{n-1} \sum_{i=0}^{\ell} b_i e^{k_i t} + \cdots + \frac{d}{dt} C_1 \sum_{i=0}^{\ell} b_i e^{k_i t} x(t) + C_0 \sum_{i=0}^{\ell} b_i e^{k_i t} = \sum_{i=0}^{\ell} a_i e^{k_i t}. \tag{5.3.43}$$

Similar to theorem 4.1, we need a simple way of determining the mth derivative for $x(t)$. We can first express $x(t)$ as a sum of all of it's components

$$x(t) = b_\ell e^{k_\ell t} + b_{\ell-1} e^{k_{\ell-1} t} + \cdots + b_1 e^{k_1 t} + b_0 e^{k_0 t}. \tag{5.3.44}$$

From here, it is simple to see that the derivatives follow as

$$\frac{d}{dx} x(t) = b_\ell k_\ell e^{k_\ell t} + b_{\ell-1} k_{\ell-1} e^{k_{\ell-1} t} + \cdots + b_1 k_1 e^{k_1 t} + b_0 k_0 e^{k_0 t} \tag{5.3.45}$$

$$\frac{d^2}{dx^2} x(t) = b_\ell k_\ell^2 e^{k_\ell t} + b_{\ell-1} k_{\ell-1}^2 e^{k_{\ell-1} t} + \cdots + b_1 k_1^2 e^{k_1 t} + b_0 k_0^2 e^{k_0 t} \tag{5.3.46}$$

$$\vdots \tag{5.3.47}$$

$$\frac{d^n}{dx^n} x(t) = b_\ell k_\ell^n e^{k_\ell t} + b_{\ell-1} k_{\ell-1}^n e^{k_{\ell-1} t} + \cdots + b_1 k_1^n e^{k_1 t} + b_0 k_0^n e^{k_0 t}. \tag{5.3.48}$$

From here we can clearly see that

$$\frac{d^n}{dx^n} x(t) = \sum_{i=0}^{\ell} b_{\ell-i} k_{\ell-i}^n e^{k_{\ell-i} t}. \tag{5.3.49}$$

This can then be generalized for use as

Theorem 5.3.2

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function such that $f(t) = b_n e^{k_n t} + b_{n-a} e^{k_{n-1} t} + \cdots + b_1 e^{k_1 t} + b_0 e^{k_0 t}$, with b_n and k_n as constants. The mth derivative is then given by

$$\frac{d^m}{dx^m} f(t) = \frac{d^m}{dx^m} \sum_{i=0}^{n} b_i e^{k_i t} = \sum_{i=0}^{n} b_{n-i} k_{n-i}^m e^{k_{n-i} t}.$$

From here we can verify our differential equation by inserting the derivatives into our expression. Thus, equation (3.43) becomes

$$\sum_{i=0}^{\ell} b_{\ell-i} k_{\ell-i}^m e^{k_{\ell-i} t} + C_{n-1} \sum_{i=0}^{\ell} b_{\ell-i} k_{\ell-i}^{m-1} e^{k_{\ell-i} t} + \cdots + C_0 \sum_{i=0}^{\ell} b_{\ell-i} e^{k_{\ell-i} t} = \sum_{i=0}^{\ell} a_i e^{k_i t}. \tag{5.3.50}$$

Starting with the left hand side of the expression (let us denote it by $F(t)$), this can be shown to be equivalent to the right as follows:

$$F(t) = \sum_{i=0}^{\ell} \left[b_{\ell-i} k_{\ell-i}^m e^{k_{\ell-i}t} + C_{n-1} b_{\ell-i} k_{\ell-i}^{m-1} e^{k_{\ell-i}t} + \cdots + C_0 b_{\ell-i} e^{k_{\ell-i}t} \right] \tag{5.3.51}$$

$$= \sum_{i=0}^{\ell} \left[b_{\ell-i} e^{k_{\ell-i}t} \left(k_{\ell-i}^m + C_{n-1} k_{\ell-i}^{m-1} + \cdots + C_0 \right) \right]. \tag{5.3.52}$$

Since $b_{\ell-i}(k_{\ell-i}^m + C_{n-1} k_{\ell-i}^{m-1} + \cdots + C_0)$ is constant for all values of i; we can denote it by $C_\ell(i)$, since it is dependent on the value of i. We can then express $F(t)$ as

$$F(t) \equiv \sum_{i=0}^{\ell} C_\ell(i) e^{k_{\ell-i}t}. \tag{5.3.53}$$

If we write out the terms of this sum, we get

$$F(t) \equiv C_\ell(0) e^{k_\ell t} + C_\ell(1) e^{k_{\ell-1}} + \cdots + C_\ell(\ell-1) e^{k_1 t} + C_\ell(\ell) e^{k_0 t} \tag{5.3.54}$$

$$\equiv C_\ell(\ell) e^{k_0 t} + C_\ell(\ell-1) e^{k_1 t} + \cdots + C_\ell(1) e^{k_{\ell-1}} + C_\ell(0) e^{k_\ell t} \tag{5.3.55}$$

$$\equiv \sum_{i=0}^{\ell} C_\ell(\ell-i) e^{k_i t}. \tag{5.3.56}$$

Now, since $C_\ell(\ell-i)$ is still an arbitrary constant for all i, we can let $C_\ell(\ell-i) = a_i$ and then we see

$$F(t) \equiv \sum_{i=0}^{\ell} a_i e^{k_i t} \tag{5.3.57}$$

which verifies that $x(t)$ (equation 3.42) is a solution to our differential equation. From here, we can also solve for our constants in $x(t)$ in terms of the given constants C_n and a_n. First, we have (from equation 3.50)

$$\sum_{i=0}^{\ell} b_{\ell-i} \left(k_{\ell-i}^m + C_{n-1} k_{\ell-i}^{m-1} + \cdots + C_0 \right) e^{k_{\ell-i}t} = \sum_{i=0}^{\ell} a_i e^{k_i t}. \tag{5.3.58}$$

Writing the terms of this gives (let "$\rightsquigarrow$" denote "$+\cdots+$" to save some space)

$$b_\ell \left(k_\ell^m \rightsquigarrow C_0 \right) e^{k_\ell t} + b_{\ell-1} \left(k_{\ell-1}^m \rightsquigarrow C_0 \right) e^{k_{\ell-1}t} \rightsquigarrow b_1 \left(k_1^m \rightsquigarrow C_0 \right) e^{k_1 t} + b_0 \left(k_0^m \rightsquigarrow C_0 \right) e^{k_0 t} = a_0 e^{k_0 t} \rightsquigarrow a_\ell e^{k_\ell t}. \tag{5.3.59}$$

From this, we can then match up the corresponding $e^{k_n t}$ terms on both sides to solve for our constants. This gives us

$$b_\ell \left(k_\ell^m \rightsquigarrow C_0 \right) = a_\ell \tag{5.3.60}$$

$$b_{\ell-1} \left(k_{\ell-1}^m \rightsquigarrow C_0 \right) = a_{\ell-1} \tag{5.3.61}$$

$$\vdots \tag{5.3.62}$$

$$b_0 \left(k_0^m \rightsquigarrow C_0 \right) = a_0. \tag{5.3.63}$$

Thus, we notice that for all values of i,

$$b_i\left(k_i^m \rightsquigarrow C_0\right) = a_i \iff b_i = \frac{a_i}{k_i^m \rightsquigarrow C_0}. \tag{5.3.64}$$

Therefore, we can express our solution to our differential as

$$x(t) = \sum_{i=0}^{\ell} \frac{a_i e^{k_i t}}{k_i^m \rightsquigarrow C_0} = \sum_{i=0}^{\ell} \frac{a_i e^{k_i t}}{k_{\ell-i}^m + C_{n-1} k_{\ell-i}^{m-1} + \cdots + C_0} \tag{5.3.65}$$

Theorem 5.3.3

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous and differential function and let C_n, a_n, b_n and k_n be constant for all n. Given a differential equation of the form

$$\frac{d^n}{dt^n} f(t) + \frac{d^{n-1}}{dt^{n-1}} C_{n-1} f(t) + \cdots + \frac{d}{dt} C_1 f(t) + C_0 f(t) = \sum_{i=0}^{\ell} a_i e^{k_i t},$$

A solution of the following form exists:

$$f(t) = \sum_{i=0}^{\ell} b_i e^{k_i t} = \sum_{i=0}^{\ell} \frac{a_i e^{k_i t}}{k_{\ell-i}^n + C_{n-1} k_{\ell-i}^{n-1} + \cdots + C_0}$$

5.4 Integrals

The term "**integral**" encompasses several distinct concepts in mathematics. Its most prevalent meaning relates to the fundamental concept in calculus, where it involves the process of summing infinitesimally small components to determine the content of a continuous region. In calculus, an integral is a fundamental mathematical concept that can be interpreted as measuring area or extending the notion of area to more complex shapes. Alongside derivatives, integrals are the cornerstone of calculus. They are also referred to as **antiderivatives** or **primitives**. The act of evaluating an integral is known as integration (a term sometimes replaced by the more historical word "quadrature"). When an integral is approximated numerically, it is called numerical integration.

Definition 5.4.1: The Antiderivative

A function F is called an **antiderivative** of a function f on an interval I if $\frac{d}{dx} F(x) = f(x)$ for all $x \in I$.

Definition 5.4.2: Riemann Sum

Let f be defined on the closed interval $[a, b]$, and let Δ be a partition of $[a, b]$ given by $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$, where Δx_i is the width of the ith sub-interval. If $c_i \in [x_{i-1}, x_i]$, then the sum $\sum_{i=1}^{n} f(c_i) \Delta x_i$ is called a Riemann sum of f for the partition Δ.
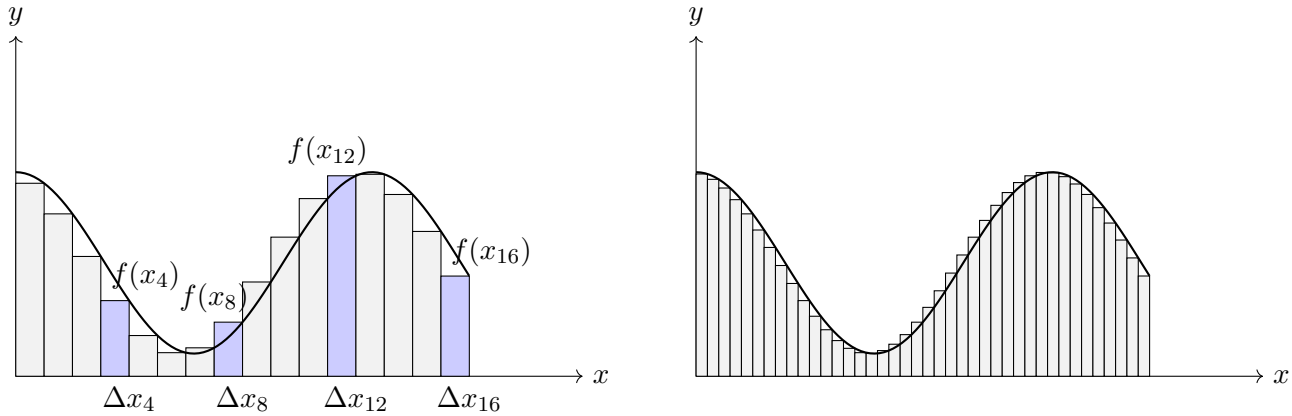
Figure 5.1: LEFT: A visualization of the Riemann Sums. The partition width is denoted by Δx_i and the height of the partitions is $f(x_i)$. RIGHT: Increasing the number of partitions and decreasing the width better approximates the area under the curve.

The largest width among the subintervals within a partition Δ is known as the partition's norm, represented by $||\Delta||$. If all subintervals have equal width, the partition is called regular, and its norm is denoted by $||\Delta|| = \Delta x = \frac{b-a}{n}$. The integral can be defined in the limit as this partition's norm approaches zero $||\Delta|| \to 0$. Therefore, there exists some $L \in \mathbf{R}$ such that for each $\epsilon > 0$, there exists a $\delta > 0$ so that for all partitions with $||\Delta|| < \delta$, we have

$$\lim_{||\Delta|| \to 0} \sum_{i=1}^{n} f(c_i)\Delta x_i = L \implies \left| L - \sum_{i=1}^{n} f(c_i)\Delta x_i \right| < \epsilon \tag{5.4.1}$$

This then give the definition of the definite integral.

> **Definition 5.4.3: The Definite Integral**
>
> IF f is defined on the closed interval $I = [a, b]$ and the limit of the Riemann sums over partitions Δ exists, then f is said to be integrable on I and the limit is denoted by the definite integral of f from a to b.
>
> $$\lim_{||\Delta|| \to 0} \sum_{i=1}^{n} f(c_i)\Delta x_i = \int_{a}^{b} f(x)dx \tag{5.4.2}$$

Integration By Parts

Integration by parts is a method used to find the integral of a product of two functions. It is derived from the product rule for differentiation and allows us to express the integral of a product in terms of simpler integrals.

> **Theorem 5.4.1: Integration By Parts**

Let f and g (or u and v) be differentiable functions for all x, then

$$\int f(x)\frac{d}{dx}g(x)\mathrm{dx} = f(x)g(x) - \int g(x)\frac{d}{dx}f(x)\mathrm{dx} \implies \int u\ dv = uv - \int v\ du \qquad (5.4.3)$$

Proof for theorem 5.4.1. Beginning with the product rule from theorem 5.1.3 we have

$$\frac{d}{dx}f(x)g(x) = g(x)\frac{d}{dx}f(x) + f(x)\frac{d}{dx}g(x) \qquad (5.4.4)$$

Integrating both sides with respect to x gives

$$\int \frac{d}{dx}f(x)g(x)\mathrm{dx} = f(x)g(x) = \int g(x)\frac{d}{dx}f(x)\mathrm{dx} + \int f(x)\frac{d}{dx}g(x)\mathrm{dx} \qquad (5.4.5)$$

This gives what is referred to as integration by parts.

$$\int f(x)\frac{d}{dx}g(x)\mathrm{dx} = f(x)g(x) - \int g(x)\frac{d}{dx}f(x)\mathrm{dx} = f(x)g(x) - \int g(x)f'(x)\mathrm{dx} \qquad (5.4.6)$$

Let $u = f(x)$ and $v = g(x)$, where dv and du represent the respective derivatives. Then this becomes $\int u\ dv = uv - \int v\ du$. $\qquad\square$

Linear Algebra

Linear algebra is a branch of mathematics that focuses on the study of vector spaces, linear transformations, and systems of linear equations. It provides a powerful framework for representing and solving problems involving linear relationships between variables. Key concepts in linear algebra include:

- **Vectors**: Vectors are mathematical objects that represent magnitude and direction. They can be represented as ordered lists of numbers and are used to describe quantities with both magnitude and direction, such as velocity and force.

- **Vector Spaces**: A vector space is a set of vectors that satisfy certain properties under addition and scalar multiplication. It forms the foundation of linear algebra and allows for the study of linear combinations and transformations.

- **Matrices**: Matrices are rectangular arrays of numbers or elements, organized into rows and columns. They are used to represent linear transformations and to solve systems of linear equations.

- **Linear Transformations**: Linear transformations are functions that preserve the structure of vector spaces. They map vectors from one vector space to another while maintaining properties like linearity and preservation of the origin.

- **Eigenvalues** and **Eigenvectors**: In linear algebra, eigenvalues and eigenvectors are associated with linear transformations. Eigenvectors are special vectors that remain in the same direction after a linear transformation, and eigenvalues represent how much the eigenvectors are scaled during the transformation.

Linear algebra finds applications in various fields, including physics, computer graphics, engineering, data science, and economics. It plays a fundamental role in solving systems of linear equations, understanding linear transformations, and providing tools to analyze complex systems with multiple variables and interactions. Moreover, it forms the basis for more advanced mathematical concepts and techniques in areas like optimization, machine learning, and numerical analysis.

6.1   Linear Systems

A linear system, in the context of mathematics and engineering, refers to a collection of linear equations involving multiple variables. Each equation in the system is linear, meaning that it represents a relationship where each variable is raised to the power of 1 and multiplied by a constant coefficient. The general form of a linear equation in one variable, x, is $mx + b = 0$, Where m and b are constants. In a linear system involving multiple variables, the equations have a similar form

$$a_0 x_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b, \tag{6.1.1}$$

where $a_0, ..., a_n$ are the coefficients, $x_0, ..., x_n$ are the variables, and b is a constant.

> **Definition 6.1.1: Linear Combination**
>
> A **linear combination** of $x_0, ..., x_n$ has the form
>
> $$a_0 x_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n,$$
>
> where the numbers $a_0, ..., a_n \in \mathbb{R}$ are the combination's coefficients. A linear equation in the variables $x_0, ..., x_n$ has the form $a_0 x_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = d$, where $d \in \mathbb{R}$ is the constant. An n-tuple $(s_0, s_1, ..., s_n) \in \mathbb{R}^n$ is a solution of, or satisfies, that equation if substituting the numbers $s_0, ..., s_n$ for the variables gives a true statement: $a_0 s_0 + a_1 s_1 + + a_n s_n = d$. A system of linear equations
>
> $$a_{0,0} x_0 + a_{0,1} x_1 + ... + a_{0,n} x_n = d_0$$
> $$a_{1,0} x_0 + a_{1,1} x_1 + ... + a_{1,n} x_n = d_1$$
> $$a_{2,0} x_0 + a_{2,1} x_1 + ... + a_{2,n} x_n = d_2$$
> $$\vdots$$
> $$a_{m,0} x_0 + a_{m,1} x_1 + ... + a_{2,n} x_n = d_m$$
>
> has the solution $(s_0, s_1, ..., s_n)$ if that n-tuple is a solution of all of the equations.

> **Theorem 6.1.1: Gauss's Method**
>
> If a linear system is changed to another by one of the following operations, then the two systems have the same set of solutions.
>
> 1. An equation is swapped with another.
>
> 2. An equation has both sides multiplied by a non-zero constant.
>
> 3. An equation is replaced by the sum of itself and a multiple of another.

Proof for theorem 6.1.1 item 1. The proof for this one is trivial. If we have a system of linear equations with a tuple that satisfies those equations, then by definition 6.1.1, all the equations hold true with that tuple. By swapping any of them, they will still hold true, and thus swapping them has no effect on the set of solutions. $\square$

Proof for theorem 6.1.1 item 2. The proof for this one is also trivial. Consider a system of linear equations with a tuple $(s_0, s_1, ..., s_n) \in \mathbb{R}^n$ that satisfies those solutions. Then by definition 6.1.1, for all $i \in \mathbb{N}$, we have a true statement $a_{i,0} x_0 + a_{i,1} x_1 + ... + a_{i,n} x_n = d_i$. Suppose we multiply both sides of this equation by some $p \in \mathbb{R}$, where $p \neq 0$. Then we have $p(a_{i,0} s_0 + a_{i,1} s_1 + ... + a_{i,n} s_n) = p d_i$. Since the tuple is a solution to the equation, there has to exist some $p d_i$ value to satisfies this equation call it d_i' (the prime representing our new linear system), giving us $a_{i,0}' s_0 + a_{i,1}' s_1 + ... + a_{i,n}' s_n) = p d_i = d_i'$, where a_i' are some new coefficients $p a_i$ that satisfy the equation and d_i' is some new constant. Our tuple will satisfy this equation, and since none of the other equations changed, it will also satisfy them. Therefore, our solution satisfies this new linear system. Since the solution was arbitrary, this implies that all solutions are also satisfied. $\square$

Proof for theorem 6.1.1 item 3. The proof for this one is also trivial. Consider a system of linear equations with a tuple $(s_0, s_1, ..., s_n) \in \mathbb{R}^n$ that satisfies those solutions. Let $p \in \mathbb{R}$. Take two equations in the system

$a_{i,0}x_0 + a_{i,1}x_1 + ... + a_{i,n}x_n = d_i$ and $a_{j,0}x_0 + a_{j,1}x_1 + ... + a_{j,n}x_n = d_j$. We can construct a new equation by adding the first to a multiple of the second, say $d_i + pd_j$. this gives $d_i + pd_j = a_{i,0}x_0 + a_{i,1}x_1 + ... + a_{i,n}x_n + p(a_{j,0}x_0 + a_{j,1}x_1 + ... + a_{j,n}x_n)$. Since our tuple is a solution to both of these equations, we know that the first term d_i and the second term pd_j (when our tuple is plugged into the variables) will satisfy the new expression. Thus, since this tuple is a solution to this equation, and since none of the other equations changed, it will also satisfy them. Therefore, our solution satisfies this new linear system. Since the solution was arbitrary, this implies that all solutions are also satisfied. □

Definition 6.1.2: Gaussian Operations

The three operations from theorem 6.1.1 are called **Gaussian operations** (or elementary reduction operations, or row operations). They are swapping, multiplying by a scalar, and row combination.

Definition 6.1.3: Echelon Form

In a system of equations, the initial variable with a non-zero coefficient within each row is referred to as the leading variable for that row. A system is said to be in **echelon form** when every leading variable, except for the first row's leading variable, is positioned to the right of the leading variable in the row above it. Additionally, any rows containing coefficients of all zeros are situated at the bottom.

In an echelon form linear system the variables that are not leading are **free**. An example of echelon form follows. Consider the equations $x - 4y + 2z = -1$, $7z + 3y = 10$, and $z = 10$. The echelon form of these is

$$1x - 4y + 2z = -1 \tag{6.1.2}$$
$$0x + 3y + 7z = 10 \tag{6.1.3}$$
$$0x + 0y + 1z = 2 \tag{6.1.4}$$

I added coefficients to each term in order to better demonstrate how the variables 'line up.' This form is analogous to representing data in matrix format. For these equations, a matrix representation would be written as

$$\begin{bmatrix} 1 & -4 & 2 \\ 0 & 3 & 7 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 10 \\ 2 \end{bmatrix} \qquad \text{or} \qquad \left(\begin{array}{ccc|c} 1 & -4 & 2 & -1 \\ 0 & 3 & 7 & 10 \\ 3 & 0 & 1 & 2 \end{array} \right) \tag{6.1.5}$$

Definition 6.1.4: Matrix

An $m \times n$ matrix is a rectangular array of numbers with m rows and n columns (2-dimensional). Each number in the matrix is an entry which has an index in each dimension. A matrix is typically denoted with an upper case letter. A matrix is typically a 2-dimensional object but can be mathematically extended into higher dimensions. for example, a $3x5$ matrix with elements a_{mn} is denoted

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{bmatrix}$$

Definition 6.1.5: Homogeneous Solutions

A linear equation is **homogeneous** if it has a constant of zero and thus can be written

$$a_0 x_0 + a_1 x_1 + \cdots + a_n x_n = 0 \qquad (6.1.6)$$

Lemma 6.1.1: Homogeneous solution

Any homogeneous linear system has a solution set of the form

$$\{c_1 \vec{\alpha}_1 + \cdots + + c_n \vec{\alpha}_n | c_1, \ldots c_n \in \mathbb{R}\},$$

where n is the number of free variables in an echelon form of the system.

Theorem 6.1.2: General Solution: Particular + Homogeneous

Any linear system has a solution set of the form

$$\{\vec{p} + c_1 \vec{\alpha}_1 + \cdots + + c_n \vec{\alpha}_n | c_1, \ldots c_n \in \mathbb{R}\},$$

where $\vec{p}$ is a particular solution and n is the number of free variables that the system has after a Gaussian reduction.

General Physics

Physics is the scientific study of the fundamental principles that govern the behavior of matter, energy, space, and time. It seeks to understand the natural world by formulating and testing mathematical models and theories based on observations and experimental evidence. Physics is a broad and diverse field that encompasses various sub-disciplines, including classical mechanics, electromagnetism, thermodynamics, quantum mechanics, and relativity, among others. Key areas of study in physics include:

- **Classical Mechanics**: This branch deals with the motion of macroscopic objects and is based on Newton's laws of motion and the concept of conservation of energy and momentum.

- **Electromagnetism**: This field explores the behavior of electric and magnetic fields and their interactions with charged particles, leading to the understanding of electricity, magnetism, and electromagnetic waves.

- **Thermodynamics**: It examines the relationships between heat, work, and energy, providing insights into the behavior of gases, liquids, and solids and the principles of heat engines and refrigeration.

- **Quantum Mechanics**: This revolutionary theory describes the behavior of particles at the atomic and subatomic levels, often challenging our classical intuitions and introducing concepts like wave-particle duality and quantum entanglement.

- **Relativity**: Special and General Relativity, proposed by Albert Einstein, deal with the nature of space, time, and gravity, especially at high speeds and in the presence of massive objects.

- **Nuclear Physics**: It focuses on the study of atomic nuclei, nuclear reactions, and nuclear forces, leading to applications in energy production and understanding the universe's early moments.

- **Particle Physics**: This field investigates the fundamental particles of the universe and their interactions, exploring the subatomic realm with particle accelerators and detectors.

- **Astrophysics**: The study of celestial objects and phenomena, such as stars, galaxies, black holes, and cosmic rays, to understand the structure and evolution of the universe.

  Condensed Matter Physics: It examines the properties of solid and liquid matter and studies phenomena like superconductivity, magnetism, and phase transitions.

Physics has been crucial in advancing technology and shaping our understanding of the universe, from everyday applications like electronics and telecommunications to groundbreaking discoveries about the fundamental nature of reality. It plays a central role in modern science and continues to be a driving force for innovation and progress in various fields. Physics stands as the foundational and all-encompassing science, profoundly shaping the advancement of all scientific disciplines.

> **💡 Interesting Note**
>
> A poet once said, "The whole universe is in a glass of wine." We will probably never know in what sense he meant that, for poets do not write to be understood. But it is true that if we look at a glass of wine closely enough we see the entire universe. There are the things of physics: the twisting liquid which evaporates depending on the wind and weather, the reflections in the glass, and our imagination adds the atoms. The glass is a distillation of the earth's rocks, and in its composition we see the secrets of the universe's age, and the evolution of stars. What strange array of chemicals are in the wine? How did they come to be? There are the ferments, the enzymes, the substrates, and the products. There in wine is found the great generalization: all life is fermentation. Nobody can discover the chemistry of wine without discovering, as did Louis Pasteur, the cause of much disease. How vivid is the claret, pressing its existence into the consciousness that watches it! If our small minds, for some convenience, divide this glass of wine, this universe, into parts—physics, biology, geology, astronomy, psychology, and so on—remember that nature does not know it! So let us put it all back together, not forgetting ultimately what it is for. Let it give us one more final pleasure: drink it and forget it all! [17]

7.1   Chemistry

Inorganic chemistry is the branch of chemistry that deals with the study of inorganic compounds, which are substances not primarily based on carbon-hydrogen (C-H) bonds. It explores the properties, structures, and behaviors of minerals, metals, gases, and other inorganic substances, playing a crucial role in understanding the fundamental elements and their interactions. The theory of chemistry and chemical reactions, was summarized to a large extent in the periodic chart of **Mendeleev**, also known as the **periodic table of elements**. All of these rules were eventually explained by quantum mechanics and thus theoretical chemistry is actually just physics. **Organic chemistry** is the chemistry of substances which are associated with living things. These are typically complex. This field is closely related to biology.

Statistical mechanics was a branch of physics developed alongside chemistry and has played an extremely important role in history. In any chemical situation, there are a large number of atoms all interacting. If all of these atoms could be analyzed at each collision and kept track of, the phenomenon of chemical reactions could be predicted. Statistical mechanics is the science of heat or thermodynamics which accomplishes this to some degree.

7.2   Biology

Biology is the scientific study of living organisms and their interactions with the environment. It encompasses a wide range of topics, from the molecular and cellular levels to the study of ecosystems and biodiversity. Biology seeks to understand the processes of life, including growth, reproduction, adaptation, and evolution, providing insights into the diversity and complexity of living systems. Biology is composed of many physical phenomenon and systems which can be studies and understood by physics. There are often very intriguing connections between physics and biology. The **conservation of energy** was in large part inspired by biology when Julius von Mayer demonstrated a connection with the amount of heat taken in and given out by a living creature. Another example is that of nerves which function as very find tubes with positive ions on the outside and negative ions on the inside - much like a capacitor that can be discharged to send signals.

Proteins are large, complex molecules composed of amino acids, essential for various biological functions. They play crucial roles in the structure, function, and regulation of cells and tissues, serving as enzymes, receptors, transporters, and antibodies. **Enzymes**, a specialized type of protein, act as biological

catalysts, speeding up chemical reactions within cells. They are highly specific and crucial for metabolic processes, allowing cells to carry out essential tasks such as digestion, energy production, and DNA replication. A significant achievement since the 1960s has been the precise determination of the atomic arrangement in certain proteins, comprising approximately fifty-six or sixty amino acids in sequence. These studies have located over a thousand atoms, including hydrogen atoms, forming intricate patterns in some proteins, such as **hemoglobin**. Despite this progress, the mystery remains as to why these patterns function as they do.

DNA, short for deoxyribonucleic acid, is a molecule that carries genetic information in all living organisms. It consists of a double helix structure made up of nucleotide units, each containing a sugar, phosphate group, and a nitrogenous base. DNA serves as the blueprint for the development, functioning, and inheritance of living beings, encoding the instructions required for the synthesis of proteins and controlling various cellular processes. The study of DNA's structure led to the remarkable discovery that it consists of a pair of twisted chains, with specific cross-links between them. These cross-links, represented by adenine, thymine, cytosine, and guanine, fit together in pairs. The pairs form strong bonds and only certain pairs will create these bonds. Therefore, when DNA is split, there is only one way to reproduce the other half in cell reproduction, ensuring the accurate replication of DNA during cell division and passing on specific instructions for protein synthesis to offspring.

No subject or field is making more progress on so many fronts at the present moment, than biology, and if we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms (see 12.1), and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms.

7.3　Astronomy

Astronomy is the scientific study of celestial objects, phenomena, and the universe as a whole. It encompasses the observation, analysis, and understanding of stars, planets, galaxies, and other cosmic structures, as well as the exploration of space and the fundamental principles governing the cosmos. Astronomy plays a crucial role in unraveling the mysteries of the universe, from studying the origins of galaxies to investigating the properties of distant exoplanets and the evolution of the cosmos over billions of years. Astronomy is older than physics and bridged the way for the beginning of physics through observations of the simplistic motion of starts and planets.

The most remarkable discovery in all of astronomy is that the stars are made of atoms of the same kind as those on the earth. With a **spectroscope**, we can analyze the frequencies of light waves and can therefore see the presence of atoms that are in different stars. Some chemicals were even discovered in stars before being discovered on earth.

One of the most impressive discoveries was the origin of energy in stars. The nuclear 'burning' of hydrogen supplies energy to stars. The stellar properties inside stars make them prime conditions for many nuclear reactions to occur and new elements to be created. The isotopic proportions in our own composition reveal the stellar origins of the elements, suggesting that we were "cooked" in stars and expelled through novae and supernovae explosions. Astronomy, closely linked to physics, offers valuable insights into these processes.

7.4　Geology

Geology, also called **earth sciences**, is the scientific study of the Earth's structure, composition, and processes that have shaped its history. It investigates the formation of rocks, minerals, and geological features, such as mountains, volcanoes, and earthquakes, as well as the evolution of landscapes and the interactions between the Earth's materials and natural forces. Geology plays a crucial role in understanding Earth's

past, present, and future, and it contributes to various fields, including resource exploration, environmental management, and hazard assessment. The question basic to geology is, what makes the earth the way it is?

Meteorology is the scientific study of the Earth's atmosphere and weather phenomena. It closely relates to physics because it involves the application of fundamental physical principles, such as thermodynamics, fluid dynamics, and radiation, to understand and predict atmospheric processes. By analyzing the behavior of air masses, temperature, humidity, pressure, and wind patterns, meteorologists use physics-based models and observations to forecast weather conditions and study atmospheric phenomena like hurricanes, tornadoes, and climate change. The interaction between meteorology and physics plays a crucial role in advancing our understanding of the Earth's atmosphere and its impact on weather patterns and climate dynamics.

While we have considerable knowledge about earthquake waves and the Earth's density distribution, physicists face challenges understanding matter's properties at the immense pressures within the Earth's core. Unlike conditions in stars, the complex mathematics involved in these extreme circumstances remains a hurdle to tackle. Additionally, determining the density and behavior of rocks at high pressure requires experimental investigations, as our theoretical understanding is limited.

7.5 Energy

The **conservation of energy** is believed to be a fundamental law governing all known natural phenomena that are known to date. It is believed to be an exact and unbroken principle, stating that a quantity known as **energy** remains constant amidst the various changes occurring in nature. This abstract idea is based on mathematical principles, representing a numerical quantity that remains unchanged throughout different events. It is not a description of a concrete mechanism but rather a remarkable fact: the calculated energy value remains the same after observing nature's transformations. Energy has a large number of different forms, and there is a formula for each one. These are: gravitational energy, kinetic energy, heat energy, elastic energy, electrical energy, chemical energy, radiant energy, nuclear energy, mass energy. If we total up the formulas for each of these contributions in a system, it will not change except for energy going in and out of the system.

In quantum mechanics it turns out that the conservation of energy is very closely related to another important property of the world, *things do not depend on the absolute time.*

7.5.1 Work

In physics, work is defined as the transfer of energy that results from the application of a force over a certain distance. When a force acts on an object and causes it to move in the direction of the force, work is done on the object. Work is a scalar quantity and is expressed in units of joules (J) in the International System of Units (SI).

Definition 7.5.1: Work

The work done on an object is defined as the force $\vec{F}$ applied to an object through some displacement $d\vec{r}$

$$W = \int \vec{F} \cdot d\vec{r} \tag{7.5.1}$$

7.5.2 Gravitational Potential Energy

The general name of energy which has to do with location relative to something else is called **potential energy**. In the case of objects relative to earth (or other bodies effected by gravitational energy), it is called **gravitational potential energy**.

Definition 7.5.2: Gravitational Potential Energy

ravitational potential energy E is a form of potential energy associated with the position of an object in a gravitational field. It represents the energy stored in an object due to its height h above a reference point. The higher an object is positioned above the reference point, the greater its gravitational potential energy.

$$E = mgh, \tag{7.5.2}$$

where m is the mass of the object and g is the acceleration due to gravity.

For static structures and systems, one can apply an imaginary motion to the system (even if it is not really moving or even movable) in order to apply the principal of conservation of energy. This approach is called the **principle of virtual work**.

7.5.3 Kinetic Energy

The principal of **motion** gives rise to another type of energy referred to as **kinetic energy**. Beginning with definition 7.5.1 and applying newtons second law (see equation 8.2.5), we have

$$W = \int \vec{F} \cdot d\vec{r} = \int m\vec{a} \cdot d\vec{r} = m \int \frac{d\vec{v}}{dt} \cdot d\vec{r} \tag{7.5.3}$$

Since $\vec{v} = \frac{d\vec{r}}{dt}$, we can say $\vec{v}dt = d\vec{r}$. We substitute this in to get

$$W = m \int \frac{d\vec{v}}{dt} \cdot \vec{v}dt = m \int \frac{d\vec{v}}{dt} \cdot \frac{d\vec{x}}{dt}dt \tag{7.5.4}$$

Let $u = v = \vec{v} = \frac{d\vec{x}}{dt}$, then $\frac{dv}{dt} = \frac{d^2\vec{x}}{dt^2} = \frac{d\vec{v}}{dt} \implies dv = \frac{d\vec{v}}{dt}dt = du$. This allows us to apply integration by parts (see theorem 5.4.1) giving

$$m \int \frac{d\vec{v}}{dt} \cdot \frac{d\vec{x}}{dt}dt = m\vec{v} \cdot \vec{v} - m \int \frac{d\vec{v}}{dt} \cdot \frac{d\vec{x}}{dt}dt \tag{7.5.5}$$

Solving for our original integral then gives

$$m \int \frac{d\vec{v}}{dt} \cdot \frac{d\vec{x}}{dt}dt = \frac{1}{2}m||\vec{v}||^2 \tag{7.5.6}$$

This is what is known as or termed the kinetic energy.

Definition 7.5.3: Kinetic Energy

inetic energy E is a fundamental concept in physics that refers to the energy possessed by an object due to its motion. It is the energy associated with the object's velocity and is dependent on both its mass m and the square of the magnitude of its velocity v.

$$E = \frac{1}{2}mv^2 = \frac{p^2}{2m} \tag{7.5.7}$$

---

when associated with relativity, there is a modification to this concept of kinetic energy. An object has energy from simply existing which is referred to as a mass energy which is combined with this kinetic energy of motion.

Classical Mechanics

Classical mechanics is a branch of physics that describes the motion of objects and systems under the influence of forces. It forms the foundation of mechanics before the advent of quantum mechanics and relativistic physics. Classical mechanics is based on Newton's laws of motion and the concept of conservation of energy and momentum. Key principles and concepts of classical mechanics include:

- **Newton's Laws of Motion**: Sir Isaac Newton formulated three fundamental laws that govern the motion of objects. The first law (Law of Inertia) states that an object at rest remains at rest, and an object in motion continues to move at a constant velocity unless acted upon by an external force. The second law describes how the acceleration of an object is directly proportional to the net force applied and inversely proportional to its mass. The third law states that for every action, there is an equal and opposite reaction.

- **Conservation of Energy**: The principle of energy conservation states that the total energy of an isolated system remains constant over time. Energy can transform from one form to another (e.g., kinetic energy to potential energy), but the total amount of energy remains unchanged.

- **Conservation of Momentum**: The principle of momentum conservation states that the total momentum of an isolated system remains constant, provided no external forces act on it. Momentum is the product of an object's mass and velocity.

- **Gravitation**: Classical mechanics includes the study of gravitational forces between objects, described by Newton's law of universal gravitation. This law explains how objects attract each other with a force proportional to their masses and inversely proportional to the square of the distance between them.

- **Harmonic Motion**: The study of harmonic motion involves oscillations and vibrations of systems, such as a pendulum or a mass-spring system. These motions follow simple harmonic motion equations and exhibit periodic behavior.

Classical mechanics provides accurate and practical predictions for a wide range of everyday scenarios and macroscopic systems. While it is highly effective in describing the behavior of objects at non-relativistic speeds, it becomes less accurate when dealing with extremely high speeds or microscopic particles, where quantum mechanics and relativistic physics are more appropriate. Nonetheless, classical mechanics remains a crucial and fundamental branch of physics, forming the basis for understanding the motion of everyday objects and engineering applications.

8.1   Gravity

The law of gravitation is that every object in the universe attracts every other object with a force which for any two bodies is proportional to the mass of each and varies inversely as the square of the distance between them. This is typically mathematically represented as

$$F = \frac{Gm_1m_2}{r^2} \tag{8.1.1}$$

While we can describe how gravity works mathematically, its underlying mechanism remains somewhat unknown. The absence of a known mechanism drives scientific inquiry, leading to further discoveries. No mechanism has been devised to "explain" gravity without simultaneously predicting the existence of

some other phenomenon. To a great precision, this force is exactly proportional to mass, and therefore fundamentally related to **inertia**.

? Questions

- Is this law a constant through time - perhaps the constant changes?

- Does this law somehow relate to the force of electrical attraction since they are both proportional to the inverse of the distance squared?

♀ Interesting Note

The relative strengths of electrical and gravitational interactions between two electrons is

$$\frac{\text{Gravitation Attraction}}{\text{Electrical Repulsion}} = \frac{1}{4.17 \times 10^{42}}. \tag{8.1.2}$$

If we compare the time it takes for light to go across a proton (10^{-24} seconds) to the age of the universe, which is approximately 2×10^{10} years, the result is 10^{-42}. They both have a similar number of zeros, leading to a potential idea that the gravitational constant may be linked to the age of the universe.

8.1.1   Keplar's Laws

Johannes Kepler was a German mathematician, astronomer, and astrologer who made significant contributions to our understanding of the solar system and planetary motion during the late 16th and early 17th centuries. Kepler is best known for formulating three fundamental laws of planetary motion, which laid the groundwork for modern astronomy. These laws were based on careful observations made by Tycho Brahe. **Kepler's laws** are

1. Each planet moves around the sun in an ellipse, with the sun at one focus.

2. The radius vector from the sun to the planet sweeps out equal areas in equal intervals of time.

3. The squares of the periods of any two planets are proportional to the cubes of the semi-major axes of their respective orbits: $T \propto a^{3/2}$.

Galileo was studying these laws when he discovered the principal of **inertia**-*if something is moving, with nothing touching it and completely undisturbed, it will go on forever, coasting at a uniform speed in a straight line.*

? Questions

- What is the cause of inertia?

- Does an objects inertia lower over time?

Newton further extended this by introducing the idea of a **force**. If a body is to change speed or direction, a force must be applied to it in the direction of the change. This gives the understanding that there must be a force acting on planets in order to keep them in rotation around the sun - rather than just flying off into space.

8.2 Motion

keywordMotion is the change in the position of an object relative to a reference point over time. It involves the displacement of an object from one location to another in a continuous manner. Motion can occur in various forms, such as **linear motion** (movement along a straight line), **circular motion** (movement in a circular path), **rotational motion** (movement around a fixed axis), or **oscillatory motion** (back-and-forth movement around a central position).

     Velocity is a vector quantity that describes the rate of change of an object's position with respect to time. It specifies both the speed and the direction of an object's motion. Velocity $\vec{v}$ is calculated by dividing the displacement (change in position) Δx of the object by the time taken to cover that displacement Δt. Velocity is expressed in units of distance per unit of time, such as meters per second (m/s) or kilometers per hour (km/h). If the velocity is constant over time, it is referred to as constant velocity, and if the velocity is changing, it is referred to as variable velocity. When the direction of an object's motion remains constant, the velocity is considered to be uniform; otherwise, it is non-uniform. Velocity plays a crucial role in describing and analyzing the motion of objects in physics and other scientific disciplines.

$$\vec{v}(t) = \lim_{\Delta t \to 0} \frac{\Delta \vec{x}(t)}{\Delta t} = \frac{d\vec{x}(t)}{dt} \implies \vec{x}(t) = \int \vec{v}(t)dt \tag{8.2.1}$$

     Acceleration is a vector quantity that describes the rate of change of an object's velocity with respect to time. It indicates how quickly an object's velocity is changing or how much the object is speeding up or slowing down. Acceleration is defined as the change in velocity divided by the change in time. Acceleration, like velocity, is a vector quantity, meaning it has both magnitude and direction. If an object is moving in a straight line, acceleration is in the same direction as the net force acting on the object. If the object is moving in a curved path, the acceleration is directed towards the center of the curvature. The units of acceleration are typically expressed as distance per unit time squared, such as meters per second squared (m/s^2) or centimeters per second squared (cm/s^2).

$$\vec{a}(t) = \lim_{\Delta t \to 0} \frac{\Delta \vec{v}(t)}{\Delta t} = \frac{d\vec{v}(t)}{dt} = \frac{d^2\vec{x}(t)}{dt^2} \implies \vec{v}(t) = \int \vec{a}(t)dt \tag{8.2.2}$$

     Jerk is a vector quantity that represents the rate of change of acceleration with respect to time. In other words, jerk measures how quickly an object's acceleration is changing or how rapidly the object's rate of change in velocity is varying. Jerk is a vector quantity, just like acceleration and velocity, meaning it has both magnitude and direction. However, jerk is less commonly used in everyday physics problems compared to acceleration, which is more frequently encountered. In most cases, acceleration is sufficient to describe the motion of objects, and jerk is not a significant factor in standard physics analysis. Nonetheless, jerk becomes relevant in certain specialized fields, such as robotics, aerospace engineering, and studies involving rapid changes in motion or acceleration profiles.

$$\vec{j}(t) = \lim_{\Delta t \to 0} \frac{\Delta \vec{a}(t)}{\Delta t} = \frac{d\vec{a}(t)}{dt} = \frac{d^2\vec{v}(t)}{dt^2} = \frac{d^3\vec{x}(t)}{dt^3} \implies \vec{a}(t) = \int \vec{j}(t)dt \tag{8.2.3}$$

8.2.1 Newton's Laws

Momentum $\vec{p}$ is a vector quantity that measures the quantity of motion possessed by an object. It is the product of an object's mass and its velocity and is a fundamental concept in classical mechanics. In this context, **Mass** is a quantitative measure of **inertia** (resistance to being accelerated). Since momentum is a vector, it has both magnitude and direction. The direction of the momentum vector is the same as the direction of the velocity vector. The SI unit of momentum is kilogram-meter per second (kg·m/s).

$$\vec{p} = m\vec{v} \tag{8.2.4}$$

Following Galileo's discovery of the principal of **inertia**, Newton set out to describe the rules for finding out how an object changes its **motion** when effected by something.

> **8.2.1: Newton's Laws Of Motion**
>
> 1. **The Law of Inertia**: An object at rest will remain at rest, and an object in motion will continue to move with a constant velocity in a straight line unless acted upon by an external force./footnoteThis is simply a restating of Galileo's described principal of inertia.
>
> 2. The time-rate-of-change of a quantity called momentum is proportional to the force.
>
> $$\vec{F} = \frac{d\vec{p}}{dt} = m\frac{d\vec{v}}{dt} = m\frac{d^2\vec{x}}{dt^2} = m\vec{a} \tag{8.2.5}$$
>
> The two forms of this law $\vec{F} = m\vec{a}$ and $\vec{F} = \frac{d\vec{p}}{dt}$ are *not* equivalent with respect to relativity.
>
> 3. For every action, there is an equal and opposite reaction.

Since relativity and quantum mechanics were formulated, it has been known that these laws are not universally valid and only a highly accurate approximation for classical physics scenarios.

> **Definition 8.2.1: Conservation of Momentum**
>
> If the net externat force F^{ext} on an N-particle system is zero, the system's total momentum P is constant.

8.2.2   Inertial Frames

An **inertial reference frame**, also known as an **inertial frame** of reference or simply an inertial frame, is a specific coordinate system or frame of reference in which Newton's first law of motion holds true. In an inertial reference frame, an object at rest remains at rest, and an object in motion with constant velocity continues to move in a straight line with the same velocity unless acted upon by an external force. Alternatively, an inertial reference frame is a coordinate system that is not accelerating and has no net force acting on it. In such a frame, objects move freely and uniformly in accordance with Newton's laws of motion without any additional forces or accelerations introduced by the frame itself. Newton's second law, given no net force on an object ($\vec{F} = 0$) gives an acceleration of zero ($a = 0$). This suggests that Newton's second law includes the first. However, this is only the case in an **inertial frame**. The laws of motion only hold in inertial frames and an **inertial frame** is defined as one where **the law of inertia** holds.

Electronics

Electronics is a multidisciplinary field that revolves around the study and manipulation of electrons and their behavior in different materials and devices. This field has revolutionized the way we communicate, compute, and interact with the world, giving rise to an unprecedented era of technological advancements. Electronics encompasses a vast range of topics, from the fundamental principles of semiconductor physics to the design and development of intricate integrated circuits, enabling the creation of devices that have become an integral part of our daily lives. Key Concepts in Electronics include the following:

- **Semiconductors**: Semiconductors are materials that have properties between conductors (like metals) and insulators. The behavior of electrons in semiconductors is crucial to the operation of electronic devices. By manipulating the conductivity of semiconductors, devices like transistors and diodes can be created.

- **Transistors**: Transistors are fundamental building blocks of modern electronics. They act as switches or amplifiers, controlling the flow of electrical current by varying a small input signal. Transistors are the core components of integrated circuits (ICs), which power computers, smartphones, and a myriad of other devices.

- **Integrated Circuits**: Integrated circuits (ICs) are miniaturized collections of electronic components, such as transistors, resistors, and capacitors, fabricated on a single semiconductor wafer. They range from simple logic gates to complex microprocessors, enabling the creation of highly compact and powerful devices.

- **Digital** vs. **Analog**: Electronics deals with both digital and analog signals. Digital electronics represent information using discrete values (0s and 1s), forming the basis of computers and digital communication systems. Analog electronics, on the other hand, work with continuous signals and are essential for tasks like audio processing.

- **Circuit Design**: Circuit design involves creating schematics and layouts for electronic circuits. Design considerations include selecting components, optimizing power consumption, managing heat dissipation, and ensuring reliable operation.

- **Power Electronics**: Power electronics focuses on the conversion and control of electrical power. This field is vital for technologies such as power supplies, motor drives, renewable energy systems, and electric vehicles.

- **Signal Processing**: Signal processing involves manipulating signals to extract, enhance, or encode information. It's used in applications like image and audio processing, telecommunications, and control systems.

- **Microcontrollers** and **Microprocessors**: Microcontrollers are small computers on a single chip, often used to control specific tasks in devices like household appliances and automotive systems. Microprocessors are the brains of computers and smartphones, executing complex instructions and tasks.

- **Printed Circuit Boards** (PCBs): PCBs provide a platform for connecting and mounting electronic components. They are found in virtually all electronic devices and play a crucial role in determining a device's functionality and reliability.

9.1   Basic Definitions and Notations

Definition 9.1.1: Conductors

An electrical **conductor** is a substance where the electrons are mobile.

Copper and Aluminum are excellent conductors and other metals such as iron and steel range from fair to good conductors. There is no such thing as a perfect conductor. A **superconductor** is a material that loses practically all of its resistance (but is still not perfect).

♥ Interesting Note
One of the best conductors at room temperature is pure elemental silver.

Definition 9.1.2: Insulators

An electrical **insulator** is a substance that prevents current from flowing.

Most gases are good electrical insulators. Glass, dry wood, paper, and plastics are examples.

♥ Interesting Note
Pure water is a good insulator but the slightest impurity can cause it to conduct current.

Definition 9.1.3: Resistors

Resistors are electrical components made to allow for the control of current flow. This is typically done by changing the size, shape, or adding impurities to a substance. Electrical resistance is measured in units called **ohms**.

Resistance is often references as per a unit length (ohms per foot, etc). resistance also converts electrical energy into heat, and therefore low resistance is generally desired in electrical systems. If there is a current flowing through a resistive material, there will always be a potential difference across the object. There is no such thing as a resistance free object or an object with a truly infinite resistance.

Definition 9.1.4: Semiconductor

A **semiconductor** is a substance that has a conductivity between an insulator and that of most metal conductors.

Semiconductors are typically made from adding impurities or changing the temperature conditions. Silicon is a notable semiconductor that is common in electronic circuits. Indium or antimony are often used to dope silicon, selenium or gallium to create semiconductors. A **hole** is a shortage of an electron, and the movement of holes is often used as reference rather than the movement of electrons. When the charge carriers are electrons, a semiconductor is N-type (because the electrons are negatively charged). When the majority of the charge carriers are holes, a semiconductor is called P-type (since the absence of the negative electron creates a positive charge).

> **Definition 9.1.5: Current**
>
> **Current** can be defined as the movement of charge carriers in a substance. Current is measured in terms of the number of charge carriers passing a single point in a second in units of **coulombs**[a] per second. Another name for this unit is an **Ampere**.
>
> ———————
> [a]A coulomb is a measure of charge and is equivalent to approximately 6.24×10^{18} charge carriers (electrons or holes).

> **Definition 9.1.6**
>
> defn:static electricity **Static electricity** is a stationary electric charge that is typically generated when two objects with different electrical charges come into contact and then separate, leaving one with an excess of electrons (making it negatively charged) and the other with a deficit of electrons (making it positively charged). Unlike current electricity, which flows through conductors in a continuous circuit, static electricity remains in one place until it is discharged.

> **Definition 9.1.7**
>
> defn:emg **Electromotive force** (EMF), is a concept in physics and electrical engineering that represents the electrical **potential difference** or voltage generated by a source such as a battery, generator, or a power supply. EMF is measured in **volts** (V) and represents the energy per unit charge provided by the source to move electric charges carriers through a circuit.

An EMF of one volt V, accross a resistance R of one ohm, will cause a current I of one ampere to flow. this is a classical relationship of electricity and referred to generally as **Ohm's Law**,

$$V = I \times R \tag{9.1.1}$$

> **Definition 9.1.8**
>
> defn:conductance **Conductance** is the reciprocal of resistance and is measured in **siemens** (1/ohms).

> **Definition 9.1.9**
>
> defn:power **Power** is the hear that results when a current flows through a resistance and is measured in **watts** (W), which represent electrical power.
>
> $$P = VI = I^2 R = \frac{V^2}{R} \tag{9.1.2}$$

There are many different units used to measure power and energy which typically depends on the context of the problem. A common unit used to measure energy by a physicist is the joule. One **joule** is the equivalent of one watt of power, dissipated for one second of time. Another unit is the **erg**, which is one ten-millionth of a joule. Most oftenly used to indicate cooling or heating capacity of air-conditioning units is

the **british thermal unit** (Btu), which is 1055 joules. The energy gained by a single electron in an electric field of 1 volt is 1.6×10^{-19} joules, which is called an **electron volt** (eV). The work needed to raise one pound by a distance of one foot (not including any friction) is 1.356 joules, and is called the **foot-pound** (ft-lb).

Definition 9.1.10

defn:ac **Alternating Current** (AC) is an electric current that reverses direction at regular intervals. In AC, the flow of electric charge periodically changes direction, typically oscillating back and forth in a sinusoidal (wave-like) pattern. The voltage and current in an AC circuit continuously alternate between positive and negative values, completing one full cycle within a defined period, usually measured in hertz (Hz).

In America, the frequency of utility electricity is 60 Hz (the current goes through a complete cycle every $\frac{1}{60}$ second. **Rectification** is a process in which AC is converted to DC (most commonly done using a **diode**.

Definition 9.1.11

defn:dc **Direct Current** (DC) is an electrical current that flows continuously in one direction without changing polarity. In a DC circuit, electric charge carriers move from the negative terminal of a power source (like a battery or generator) to the positive terminal in a steady, unidirectional flow.

9.2   Ciruits

In electronics, circuits can be connected in various ways, and one common configuration is the **series circuit**. In a series circuit, components are connected end-to-end, forming a single path for current to flow. In a series circuit, components are connected one after another in a linear fashion (see figure 9.1). The same current flows through each component in the circuit. The total resistance R_{total} in a series circuit is the sum of the individual resistances of the components.

$$R_{total} = R_1 + R_2 + R_3 + \cdots \tag{9.2.1}$$

Another common configuration is called a **parallel curcuit**. In a parallel circuit, components are connected in such a way that there are multiple paths for the electric current to flow (see figure 9.1). Components in a parallel circuit are connected side by side, forming separate branches for the current to travel. Each branch operates independently, and the voltage across each component in parallel is the same, which is equal to the source voltage. In a parallel circuit, the total resistance R_{total} is different from that in a series circuit. The reciprocal of the total resistance is the sum of the reciprocals of the individual resistances,

$$\frac{1}{R_{total}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots . \tag{9.2.2}$$

In a parallel circuit, the total current supplied by the source is divided among the different branches. Components with lower resistance receive more current, while those with higher resistance receive less. The sum of the currents in each branch equals the total current from the source, which is **Kirchhoff's Current Law**. Kirchhoff's Current Law states that the sum of currents entering any junction (or node) in an electrical
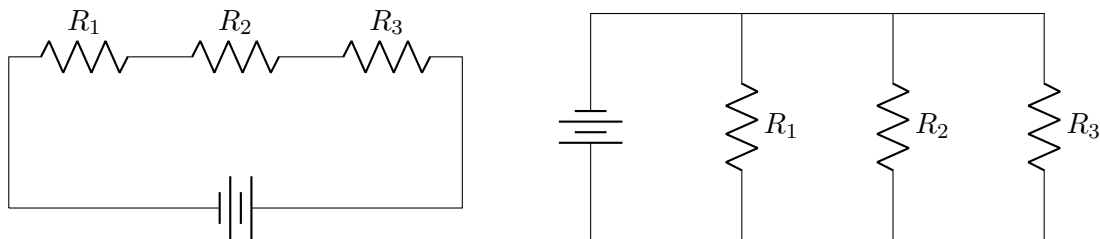
Figure 9.1: LEFT: A series circuit with three resistors and a batter. RIGHT: A parallel circuit with 3 resistors in parallel and a battery.

circuit is equal to the sum of currents leaving that junction. In other words, the total current at any point within a circuit is conserved.

$$\sum I_{in} = \sum I_{out} \tag{9.2.3}$$

Kirchhoff's Voltage Law states that the sum of the electromotive forces (EMFs) and the voltage drops in any closed loop within an electrical circuit (taking polarity into account) is equal to zero. In simpler terms, the total voltage around any closed loop in a circuit is conserved.

$$\sum V_{loop} = 0 \tag{9.2.4}$$

Acoustics

Acoustics is the scientific study of sound and its transmission, production, and effects. It is a branch of physics and engineering that explores the properties of sound waves, their interactions with various mediums, and their impact on our environment and human perception. Acoustics plays a crucial role in our everyday lives, from the design of concert halls and the development of audio technology to the understanding of natural phenomena like earthquakes and the behavior of underwater sound.

At its core, **acoustics** seeks to unravel the mysteries of sound, a complex and omnipresent phenomenon that shapes our auditory experiences and contributes to our understanding of the world. Acoustics is not just the study of auditory sound though. At its core, acoustics is the study of vibrations propogating through matter, which in some ways is the study of everything and plays a critical role in the universe and matter interactions.

10.1 Units

The fundamental unit of sound change is called the **decibel** (dB). Perceived levels of sound change according to a logarithmic relation of the actual power levels. Decibels are calculated relative to a ratio of change from power P_1 watts to a power P_2 watts. The change in decibels is then

$$dB = 10 \log_{10} \left(\frac{P_2}{P_1} \right) \tag{10.1.1}$$

10.2 Underwater Acoustics

Large bodies of water, such as the large ones covering Earth's surface do not allow for the propogation of radio and radar waves. Water, especially salt water is highly dissipative because it exhibits strong conductivity. This conductivity atteniates electromagnetic waves extremely rapidly, which limits their range and usefulness. Therefore, acoustic waves are the more practical way of carrying information underwater [21]. Acoustic waves have different transmission characteristics in water than they do in air. Some advantages in water are that they have a propogation speed or four to five times higher than in air, they reach higher levels, and they undergo less attentiation and therefore can propogate over larger distances[1]. Some typical constraints include higher ambient noise levels, and more echos. Many species utilize use of sound waves in oceans such as whales and cataceans.

---

[1]Sound propogation in the ocean can be observed at rangest of thousands of kilometers[21].

Optics

Optics is the branch of physics that focuses on the behavior and properties of light. It encompasses the study of how light interacts with various materials, how it propagates through space, and how it is perceived by our eyes. Optics plays a crucial role in understanding both the fundamental nature of light and its practical applications in fields such as astronomy, microscopy, telecommunications, and more. The exploration of optics has led to groundbreaking discoveries and technologies that have transformed the way we observe, communicate, and interact with the world around us. Key concepts in optics include the following:

- **Reflection**: Reflection occurs when light rays bounce off a surface. The angle of incidence (the angle at which light hits the surface) is equal to the angle of reflection (the angle at which light bounces off).

- **Refraction**: Refraction is the bending of light as it passes from one medium to another with a different optical density. This phenomenon is responsible for the way lenses focus light and how prisms disperse it into a spectrum of colors.

- **Dispersion**: Dispersion is the separation of white light into its component colors (spectrum) due to the varying speeds of different colors of light through a material. This is commonly observed in rainbows and in the phenomenon of chromatic aberration in lenses.

- **Diffraction**: Diffraction is the bending of light waves around obstacles and the spreading of light as it passes through small openings. This phenomenon is responsible for the characteristic patterns of light and dark bands observed when light encounters obstacles or slits.

- **Polarization**: Polarization refers to the orientation of light waves in a particular direction. Polarized light waves vibrate in a specific plane, and the study of polarization has applications in reducing glare, 3D movie technology, and various optical instruments.

- **Interference**: Interference occurs when two or more light waves combine and either reinforce (constructive interference) or cancel out (destructive interference) each other. This phenomenon is exploited in applications such as thin film coatings and optical coatings on lenses.

- **Optical Instruments**: Optics has given rise to a wide range of instruments that manipulate and utilize light. These include lenses, mirrors, telescopes, microscopes, cameras, projectors, lasers, and fiber-optic communication systems.

- **Wave-Particle Duality**: One of the intriguing aspects of optics is the wave-particle duality of light. Light can exhibit both wave-like and particle-like behaviors, depending on the context of observation. This duality is a fundamental concept in quantum physics.

- **Holography**: Holography is a technique that captures and reconstructs the complete three-dimensional information of an object using light interference patterns. It has applications in security, art, and data storage.

- **Optical Phenomena**: Optics is responsible for a multitude of natural and man-made phenomena, such as mirages, rainbows, the blue color of the sky, and the functioning of the human eye.

11.1   Snell's Law

Snell's Law, also known as the **Law Of Refraction**. is a physical formula that describes how light of other waves change direction when passing through different mediums. This law works because the light

takes the path for which the travel time from one medium to another is minimized (in time). Solving this problem gives **Snell's Law**. This law follows from **Fermat's principal of least time**.

11.1.1: Snell's Law (The Law Of Refraction)

Given two mediums, light traveling from medium 1 to medium 2 will refract at an angle given by the following relationship:

$$\frac{\sin(\theta_1)}{\sin(\theta_2)} = n = \frac{n_2}{n_1} = \frac{v_1}{v_2} \tag{11.1.1}$$

where n_1 and n_2 are the refractive indices and v_1 and v_2 are the phase velocities in the two media.
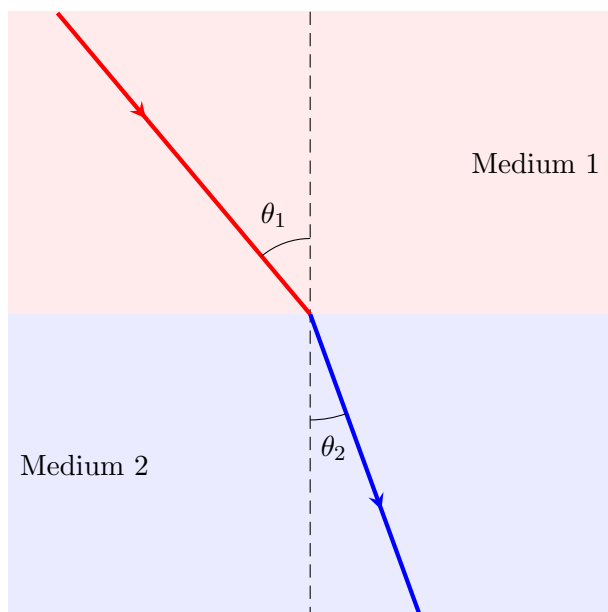
Statistical Mechanics & Thermodynamics

Statistical mechanics is a branch of physics that aims to explain the macroscopic properties of a system (such as temperature, pressure, and entropy) by understanding the behavior and interactions of its microscopic constituents at the atomic or molecular level. It bridges the gap between the microscopic world of particles and the macroscopic world we observe in everyday life. Key concepts and principles of statistical mechanics include:

- **Microstates** and **Macrostates**: In statistical mechanics, a system's microstate refers to the specific arrangement and energy distribution of its individual particles. A macrostate, on the other hand, represents the observable properties of the system, such as its temperature, pressure, and volume. The goal of statistical mechanics is to determine the probabilities of different microstates leading to a given macrostate.

- **Ensembles**: Statistical mechanics uses ensembles, which are collections of similar systems, to study statistical properties. Common ensembles include the microcanonical ensemble (isolated system with constant energy), canonical ensemble (system in thermal contact with a heat reservoir), and grand canonical ensemble (system with exchange of energy and particles with a heat reservoir).

- **Boltzmann Distribution**: The Boltzmann distribution relates the probabilities of different energy states to their corresponding energies and the system's temperature. It allows us to predict how particles are distributed among different energy levels in a system.

- **Entropy** and Entropy Maximization: Entropy is a fundamental concept in statistical mechanics, representing the measure of a system's disorder or randomness. The second law of thermodynamics states that isolated systems tend to evolve toward states of higher entropy. Statistical mechanics provides a statistical interpretation of entropy and explains the tendency of systems to maximize their entropy over time.

- **Statistical Thermodynamics**: Statistical mechanics connects with thermodynamics, relating the macroscopic thermodynamic properties (e.g., internal energy, temperature, and heat capacity) to the statistical properties of microscopic constituents. This connection allows us to derive thermodynamic quantities from the statistical behavior of particles.

Statistical mechanics has wide-ranging applications in various scientific disciplines, including physics, chemistry, biology, and materials science. It is essential for understanding phase transitions, chemical reactions, and the behavior of matter under different conditions. Additionally, statistical mechanics forms the basis for exploring complex systems, such as gases, liquids, and solids, and plays a crucial role in the development of many modern technologies.

12.1 The Atomic Hypothesis

The **atomic hypothesis** is a framework for how matter works on an atomic level.

> **❝ Quotation**
> "All things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another. [17]"

Many questions can arise from this.

Atoms are on the scale of Angstroms[1]. The jiggling motion between atoms and molecules is what's typically represented as heat. **Heat** is a form of energy that is transferred between objects or systems due to temperature differences.

? Interesting Note
For H_2O, the distance between the center of the hydrogen atom and the center of the oxygen atom is 0.957 Å, with a 105°3′ angle between the hydrogen atoms.

Definition 12.1.1: Pressure

Pressure is a fundamental concept in physics and thermodynamics. It is defined as the force F applied per unit area A on a surface. Mathematically, pressure P is expressed as

$$P = \frac{F}{A}. \tag{12.1.1}$$

Pressure is a scalar quantity and is typically measured in units of **pascals** (Pa) in the International System of Units (SI).

In everyday life, pressure is commonly encountered in various contexts, such as atmospheric pressure (the pressure exerted by the Earth's atmosphere), fluid pressure (in liquids and gases), and contact pressure (when an object is in contact with a surface). When a gas is slowly compressed, temperature increases and visa versa. In solids, as opposed to liquids, the atoms will arrange themselves into an array called a **crystalline array**. The positions are dependent on the other atoms within the array. The atoms still jiggle and bibraye and have an associated heat. They are vibrating 'in place' rather than bouncing around with a somewhat free motion that would occur in liquids or a gas. **Melting** is the process by which the vibrations cause the atoms to shake apart from the crystalline array. At absolute zero temperature, the vibrations reach a minimum allowed state that atoms can have - but not zero.

---

[1]An angstrom, symbolized as "Å", is a unit of length used primarily in the fields of chemistry, physics, and nanotechnology to measure atomic and molecular distances, as well as the size of atoms and molecules. One angstrom is equal to 0.1 nanometers or 10^{-10} meters, representing a very small scale typically used to describe atomic and molecular dimensions.

🔆 Interesting Note

Helium is a special case for atoms. Helium does not freeze at absolute zero into a solid without the help of applied pressures.

Air consists almost entirely of nitrogen, oxygen, and water vapor. The remaining elements are a small amount of CO_2, Argon, Neon, Helium, Methane (CH_4), Krypton, etc. **Evaporation** occurs when air molecules hit water and 'chip away' the H_2O molecules. When the molecules or atoms in air hit against hte surface of water, some of the energy and heat is transferred to the water molecules. This causes the vibration of those molecules to increase. In the case of those molecules having enough energy to break away from the rest of the water, the molecules become water vapor. This process essentially goes in both directions continuously until there is no water left as a liquid. The air around the water will also bounce into the water molecules and mix with it - creating gassed water. This is what allows for bubbles to exist. Given a small sealed container, this process would continue indefinitely. The higher energy atoms escape which in turn creates a cooling effect on the liquid itself because the overall temperature would be lowered. This is why blowing on something how will cool it - because the atoms in the air will interact and 'shed' the higher energy atoms from it or gain heat upon their interaction.

Ions are atoms that have either gained or lost electrons. In a salt crystal, there are chlorine ions (one extra electron) and sodium ions (one fewer electron). The ions stick together by **electrical attraction**. The electrons have a negative charge and the protons within the nucleus of the atoms has a positive charge. Because these ions have opposite charges, they attract one another.

? Questions

- What gives matter its charge?

- Why are protons positively charged, electrons negatively charged, and neutrons neutral charged?

- Why is charge intrinsically a binary feature of reality? Can there be partial charges?

Nature is doing the same thing during a **chemical process** and a **physical process**. These are when atomic partners are rearranged. Atoms prefer very specific partners and molecule formations which is a peculiar part of nature. One example is when hydrogen gas (H_2) and oxygen gas (O_2) react together in the presence of a spark or heat to produce water vapor (H_2O). The reaction is exothermic, meaning it releases energy in the form of heat. During the reaction, the hydrogen molecules (H_2) break apart into individual hydrogen atoms, and the oxygen molecules (O_2) break apart into individual oxygen atoms. Then, the hydrogen atoms combine with the oxygen atoms to form water molecules (H_2O). The resulting water vapor is in the gaseous state, and it can condense into liquid water or freeze into ice depending on the temperature and pressure conditions.

> **? Questions**
>
> - Can/should heat be thought of as a wave or field?
>
> - What median does heat travel through between atoms?

kinetic energy is the energy of motion. When **burning** occurs, some atoms form easily with little energy, while others snap together with tremendous commotion - releasing energy to everything around them. **Flames** are when this energy is so enormous that light is generated rather than just heat.

Every substance is some type of arrangement of atoms. A **molecule** is simple an arrangement of atoms. The concept of a molecule is only approximate though and exists for a certain class of substances. Sometimes there is a natural grouping in a substances structure. **Organic chemistry** is essentially detective work that guesses and checks and uses mixtures of chemicals and atoms in order to determine where atoms are located - it is often incredibly accurate.

When determining if atoms really exist, ideas of how to detect them arise. One such idea is referred to as **Brownian motion**, which is the motion of a larger object due to the interactions with the smaller objects around it which are moving. For example, a basketball in a swimming pool will move as the water molecules around it collide with it. This is one method of detecting the presence of atoms.

Astrophysics

Astrophysics is a branch of astronomy that deals with the study of celestial objects, phenomena, and the physical processes that govern the universe. It combines principles of physics with observational data to understand the behavior, composition, and evolution of stars, galaxies, planets, and other cosmic entities. Key areas of study in astrophysics include:

- **Stellar Astrophysics**: This field focuses on the properties and life cycles of stars. It examines stellar formation, nuclear fusion processes that power stars, their structure, and the various stages of stellar evolution, including the final fate of stars as supernovae, white dwarfs, neutron stars, or black holes.

- **Galactic Astrophysics**: Galactic astrophysics explores the structure, dynamics, and evolution of galaxies, which are vast collections of stars, gas, dust, and dark matter. It investigates the formation of galaxies, their interactions, and the supermassive black holes at their centers.

- **Cosmology**: Cosmology studies the large-scale properties and evolution of the universe as a whole. It addresses questions about the universe's origin, its expansion, the distribution of galaxies, dark matter, dark energy, and the ultimate fate of the cosmos.

- **Exoplanets and Planetary Systems**: This area examines planets outside our solar system (exoplanets) and investigates planetary systems' formation and dynamics. It aims to find potentially habitable exoplanets and understand the diversity of planetary systems in the universe.

- **High-Energy Astrophysics**: High-energy astrophysics focuses on cosmic phenomena that involve extreme conditions, such as black hole accretion disks, active galactic nuclei, gamma-ray bursts, and cosmic rays. It explores high-energy emissions and their impact on the surrounding space.

- **Astrobiology**: Astrobiology is an interdisciplinary field that combines aspects of astronomy, biology, and chemistry to study the potential for life beyond Earth. It seeks to understand the conditions required for life to exist on other planets and moons in our solar system and beyond.

Astrophysicists use various observational and theoretical tools, such as telescopes, space missions, computer simulations, and mathematical models, to unravel the mysteries of the cosmos. Their research not only expands our knowledge of the universe's workings but also addresses fundamental questions about our place in the cosmos and the potential existence of life elsewhere in the universe.

Electromagnetism

Electromagnetism is a branch of physics that deals with the study of electric and magnetic fields and their interactions with charged particles and currents. It unifies two important phenomena: **electricity** and **magnetism**, which were initially considered separate until the 19th century. Key concepts and principles of electromagnetism include:

- **Electric Fields**: Electric fields are regions around charged particles or objects where electric forces can influence other charged particles. They are represented by vectors, indicating the direction and strength of the force experienced by a test charge placed in the field.

- **Magnetic Fields**: Magnetic fields arise due to moving charges or currents. They also have a vector nature and affect magnetic objects, such as magnets and magnetic materials. Magnetic fields are created in closed loops and can be visualized using field lines.

- **Electromagnetic Waves**: Electromagnetic waves are a form of energy propagation resulting from oscillating electric and magnetic fields. These waves do not require a medium for transmission and travel at the speed of light. Light itself is an electromagnetic wave.

- **Maxwell's Equations**: These are a set of four fundamental equations that form the basis of classical electromagnetism. They describe how electric and magnetic fields are generated by charges and currents and how these fields interact with one another.

- **Electromagnetic Induction**: Electromagnetic induction is the process where a changing magnetic field induces an electric field, leading to the generation of an electric current in a conductor. This phenomenon is the foundation of electric generators and transformers.

- **Electromagnetic Force**: Electric and magnetic forces are fundamental forces in nature. Charged particles experience electric forces due to their electric charges, while moving charges experience magnetic forces in the presence of magnetic fields.

Electromagnetism has extensive applications in modern technology and everyday life. It is the basis for the operation of electrical circuits, power generation, telecommunications, and numerous electrical devices. Moreover, it plays a vital role in the understanding of light and optics, electromagnetic radiation, and the behavior of charged particles in magnetic fields, making it one of the fundamental pillars of modern physics. The study of electromagnetism is crucial for advancing technology and our understanding of the natural world.

14.1   The Electromagnetic Field

The **electromagnetic field** is a vector field that pervades all of space and consists of the components of an **electric field** and a **magnetic field**. The electromagnetic field can carry waves. The general name of these waves is **electromagnetic waves**. These oscillatory waves can have various frequencies which is one of the most important feature. The only thing that is really different from one wave to another is the frequency of oscillation. When charge is oscillated and shaken, there are different effects that occur which are all determined by the frequency. See **??** for some common frequencies - referred to as the **electromagnetic spectrum**. At higher frequencies, these waves behave more so like particles than actual waves, which was explained by quantum mechanics around 1920.

EM Range	Frequency (Hz)	Uses
Radio Waves	$< 3 \times 10^{11}$	Radio and TV broadcasting, wireless communication (Wi-Fi, Bluetooth), radar systems, radio astronomy
Microwaves	$3 \times 10^{11} - 10^{13}$	Microwave ovens, satellite communication, radar systems, wireless data transmission (e.g., 5G networks)
Infrared (IR)	$10^{13} - 10^{14}$	Remote controls, heat sensors, infrared imaging (night vision cameras), IR communication, thermal imaging
Visible Light	$4 - 7.5 \times 10^{14}$	Human vision, optical communication (fiber optics), photography, displays (LCD, LED), lasers
Ultraviolet (UV)	$10^{15} - 10^{16}$	Sterilization (UV-C), fluorescent lights, sun tanning, UV imaging, forensic analysis
X-rays	$10^{17} - 10^{20}$	Medical imaging (X-ray radiography and CT scans), security screening, material analysis, X-ray diffraction
Gamma Rays	$10^{20} - 10^{24}$	Medical imaging (gamma camera), radiation therapy in medicine, nuclear industry, gamma-ray astronomy

Table 14.1: Typical EM Frequencies (approximate) and their uses.

Quantum Mechanics

Quantum mechanics is a fundamental branch of physics that describes the behavior of matter and energy at the smallest scales, such as subatomic particles and photons. It provides a unique and revolutionary framework for understanding the peculiar and counterintuitive behavior of particles at the quantum level. Key concepts and principles of quantum mechanics include:

- **Wave-Particle Duality**: One of the central tenets of quantum mechanics is the wave-particle duality. It states that particles, such as electrons and photons, exhibit both particle-like and wave-like characteristics. They can be described by wave functions, which represent probabilities of finding a particle at different locations.

- **Quantization of Energy**: Quantum mechanics introduced the concept of quantized energy levels, where energy levels of particles are restricted to discrete values rather than continuous values. This is exemplified in the energy levels of electrons in an atom, resulting in the discrete emission and absorption of photons.

- **Uncertainty Principle**: The Heisenberg uncertainty principle states that it is impossible to simultaneously know both the position and momentum of a particle with absolute precision. The more accurately one quantity is known, the less precisely the other can be determined. This fundamental limitation is inherent in quantum mechanics.

- **Quantum Superposition**: Quantum systems can exist in a state of superposition, where they are in multiple states simultaneously. For example, an electron can exist in a superposition of spin-up and spin-down states until measured, at which point it collapses into one definite state.

- **Quantum Entanglement**: Quantum entanglement is a phenomenon where the properties of two or more particles become correlated in such a way that the state of one particle is directly related to the state of another, regardless of distance. This has profound implications for quantum information and potential applications in quantum computing.

  Quantum Mechanics and Measurement: The process of measurement in quantum mechanics is non-deterministic. Upon measurement, the system's wave function collapses to a specific state corresponding to the observed measurement outcome, introducing inherent randomness into quantum events.

Quantum mechanics has revolutionized our understanding of the subatomic world and is the foundation for modern technologies such as transistors, lasers, and MRI machines. While it has proven to be highly successful in describing the behavior of particles on a small scale, it also challenges our classical intuition and raises profound philosophical questions about the nature of reality and our perception of the universe.

15.1   Introduction

The quantum theory first arose from studies of radiation. Many historical developements lead to the formulation of quantum mechanics, which include **Planck's radiation law**, the **Einstein-Debye theory of specific heats**, the **Bohr atom**, **de Broglie's matter waves**, and others. In the years leading up to 1920, the picture of space as 3-dimensional with time being separate was changed by Einstein. Following this, the rules for motions of particles was found to be incorrect. The mechanical rules for inertia and forces established by Newton was found to be wrong[1]. Around 1920, **quantum mechanics** was discovered which

---

[1]Or at least only accurate as an approximation in certain cases.

explained how things on small scale behave nothing like things on the larger scales. According to quantum mechanics, the things on a small scale behave so unnaturally, that it is hard to understand outside of an analytic and mathematical framework. **Quantum mechanics** is the description of the behavior of matter and light in all its details and, in particular, of the happenings on an atomic scale.

When thinking of atoms from the atomic hypothesis (outlined in section 12.1), we can think of some mysterious questions and paradoxes that seem to arise.

> **? Questions**
>
> - If the atoms are made out of plus and minus charges, why don't the minus charges simply sit on top of the plus charges (they attract each other) and get so close as to completely cancel them out?
>
> - Why are atoms so big?
>
> - Why is the nucleus at the center with the electrons around it?

An atom has a diameter on the scale of about 10^{-8} cm. Within the atom, the nucleus only has a diameter on the scale of 10^{-13}, which is much smaller! The Heisenberg uncertainty principal $\Delta x \Delta p \geq \hbar/2$ was a rule proposed by Werner Karl Heisenberg to answer these questions above. Take an electron for example. Why is it that the electron does not fall into the nucleus of an atom? The uncertainty principal states that if we know the position of the particle, the momentum would have to be very large - the electron would have a very high kinetic energy. This energy would cause the electron to break away from the nucleus. This is also why there is still a little 'jiggle' in the atoms at absolute zero. They need to keep moving in order to follow this rule.

> **? Questions**
>
> - How does the nucleus fit into this rule?
>
> - Do the particles within the nucleus that are so closely attracted obey this rule? Perhaps this is why there is so much energy within a nucleus.
>
> - Where does the energy come from in these systems that keep matter moving?

Another interesting change that quantum physics brought about to the philosophy of physics is that *it is not possible to predict exactly what will happen in any circumstance.* This is the statistical nature of quantum mechanics. According to quantum physics, it is fundamentally impossible to make a precise (*exactly what will happen*) prediction of nature. One of the consequences of quantum mechanics is that things which we used to consider as waves also behave like particles, and particles behave like waves; in fact everything behaves the same way. There is no distinction between a wave and a particle. Quantum mechanics therefore unifies the idea of the field and its waves with the idea of particles into one. When the frequency is low, the field aspect of a phenomenon is more evident. As frequencies increase, the particle aspects of a phenomenon become more evident within equipment used in experimentation. The **photon**[2] is a new particle introduced alongside the electron, neutron, and proton. The new view of the interaction of electrons and photons that is electromagnetic theory, but with everything quantum-mechanically correct, is called **quantum electrodynamics**.

---

[2]A photon in some contexts can be thought of as just a 'packet of light'.

15.1.1 The Stern-Gerlach Experiment

Begin with silver (Ag) atoms heated in an oven. The oven has a small hole where some of the silver atoms can escape, creating a beam of silver atoms. The beam goes through a collimator and is then subjected to an inhomogeneous magnetic field created by a oaur of pole pieces (one of which has a very sharp edge). A sikver aton has 47 electrons, where 46 form a spherically symmetrical electron cloud with no angular momentum. The silver atom as a whole, has an intrinsic (as opposed to orbital) angular momentum due solely to the spin (ignoring the nuclear spin) angular momentum of the single 47th (5s) electron. The nucleus of the atom is $\approx 2 \times 10^5$ times heavier than the electron. Therefore, the magnetic moment $\vec{\mu}$ of the atom equal to the spin magnetic moment of the 47th electron, and thus proportional to the electron spin $\vec{S}$, giving $\vec{\mu} \propto \vec{S}$, where the proportionality factor is $\frac{e}{m_e c}$ to an accuracy of about 0.2%[20].

The interaction energy of the magnetic moment with the magnetic field is $-\vec{\mu} \cdot \vec{B}$ and the z-component of the force experienced by the atom is

$$F_z = \frac{\partial}{\partial z}(\vec{\mu} \cdot \vec{B}) \simeq \mu_z \frac{\partial B_z}{\partial z}. \tag{15.1.1}$$

When $\mu_z > 0$ ($S_z < 0$), atoms will experience an downward force. Similarly, when $\mu_z < 0$ ($S_z > 0$), atoms will experience an upward force. The beam is then expected to be split according to the values of μ_z. In the sense of a classical spinning object, this setup would expect all values of μ_z to be possible between $|\vec{\mu}|$ and $-|\vec{\mu}|$, since the atoms in the oven are randomly oriented. What was experimentally observed, however, was that there was only two 'spots' observed, corresponding to one 'up' (S_z^+) and one 'down'(S_z^-) orientation, splitting the original silver beam into two distinct components. These components are multiples of the fundamental units of angular momentum which are $S_z = \frac{\hbar}{2}$ and $S_z = -\frac{\hbar}{2}$.

Sequential Stern-Gerlach Experiments

Consider a sequential Stern-Gerlach experiment where the silver beam passes through multiple Stern-Gerlach apparatuses (SGA) in sequence. The first setup is by blocking the S_z^- component from the first SGA and allow the remaining S_z^+ component to enter a second SGA. In this case, only the S_z^+ component leaves the second SGA. Another setup follows where the first SGA remains the same, while allowing only the S_z^+ through, but the second SGA now has an inhomogeneous magnetic field in the x direction rather than the z direction. In this case, the S_z^+ beam entering the second SGA is split into two components, S_x^+ and S_x^-. Yet

another setup follows where we use three SGAs. The first SGA will be setup as the others, where the silver bean is split into S_z^+ and S_z^-. The second SGA takes only the S_z^+ beam using an inhomogeneous magnetic field in the x direction and splits the beam into a S_x^+ and S_x^- beam. The third SGA takes only the S_x^+ beam and uses a magnetic field in the z direction (the same as the first SGA). In this case, the final output beam of the third SGA has both S_z^+ and S_z^- components.

This example of sequential Stern-Gerlach experiments is often used to illustrate that in quantum mechanics, you cannot determine both S_z and S_x simultaneously. This is typically understood in that any measurement in one apparatus negates any previously obtained information measured. This is a well regarded result and well accepted to not be due to the incompetence of experimentalists, but rather an inherent microscopic phenomena of matter itself.

In classical mechanics, where angular momentum is defined as $\vec{L} = I\vec{\omega}$, there is no difficulty in determining the various components of ω_x, ω_y, and ω_z simultaneously by observing how fasts in object spins and in what direction. This effect is inherent to matter on a small scale and is a quantum mechanical property.

15.2   Conventions and Notation

The double-slit experiment is a fundamental experiment in quantum physics that demonstrates the wave-particle duality of particles, such as electrons and photons. It involves a barrier with two narrow slits through which particles can pass before hitting a screen. When particles are sent through the double slits one by one, they create an interference pattern on the screen, similar to the pattern created by waves. This suggests that particles exhibit wave-like behavior and interfere with themselves, producing areas of constructive and destructive interference. However, when the particles are observed or measured to determine which slit they pass through, the interference pattern disappears, and particles behave like individual particles, creating two distinct bands on the screen. This can be summarized using complex numbers and probabilities.

A probability density function P describing an event in an ideal experiment is given by the square of the absolute value of the probability amplitude $|A|^2$, where A is a complex number. When an event can occur in several alternate ways, the probability amplitude for hte event is the sum of the probability amplitudes for each way considered separately, giving interference. In mathematical terms this gives $A = A_1 + A_2$ with $P = |A_1 + A_2|^2$. When conducting an experiment capable of discerning between different alternatives, the probability of an event occurring is the sum of the probabilities associated with each alternative. Consequently, any interference or overlap between the alternatives is eliminated and you get $P = P_1 + P_2$.

Definition 15.2.1: Expectation Value

If $P(x)$ denotes a probability density function of x, then the **expectation Value** (or average) of x, denoted $\langle x \rangle$ as

$$\langle x \rangle = \int_{-\infty}^{\infty} x P(x) dx \tag{15.2.1}$$

For discrete values of x_i each with a correlated probability P_i, then the expectation value is

$$\langle x \rangle = \sum_i x_i P_i \tag{15.2.2}$$

A physical state is represented by a **state vector** in a complex vector space called a **ket**. This state is postulated to contain the complete information about the physical state of a system. An **observable**

(momentum, spin components, etc) is represented by an **operator** within the vector space. The operators act on a ket from the left. In quantum physics, operators play a fundamental role in describing the behavior of quantum systems and the mathematical representation of physical observables. Quantum operators are mathematical entities that represent physical properties, transformations, and measurements within the framework of quantum mechanics. They are used to describe the behavior of quantum states, interactions, and measurements in a mathematical and abstract way. For example, consider the Hamiltonian operator $\hat{\mathbf{H}}$ acted on a ket $|x\rangle$, represented by $\hat{\mathbf{H}}|x\rangle$. This is not the operator 'times' the ket, but rather the operator acting on the ket. In this example, the hamiltonian operator in classical mechanics is defined by $\hat{\mathbf{H}} = \frac{\hbar}{2m}\nabla^2 + V(\vec{r,t})$. This would act on $|x\rangle$ by taking the partial differentials of $|x\rangle$ as found in the first term of $\hat{\mathbf{H}}$.

Relativity

Relativity refers to two groundbreaking theories developed by Albert Einstein: **Special Relativity** and **General Relativity**. Both theories revolutionized our understanding of the universe and how it functions, especially at high speeds and in the presence of strong gravitational fields.

- **Special Relativity**: Introduced in 1905, Special Relativity deals with the behavior of objects moving at constant velocities, especially near the speed of light. It is based on two postulates: the principle of relativity (laws of physics are the same for all inertial observers) and the constancy of the speed of light in a vacuum. Key principles of Special Relativity include:

  - **Time Dilation**: Moving clocks appear to run slower compared to stationary clocks from the perspective of an observer at rest.
  - **Length Contraction**: Moving objects appear shorter in the direction of motion relative to a stationary observer.
  - **Mass-Energy Equivalence**: $E_0 = m_0 c^2$, where E_0 is the rest energy, m_0 is the rest mass, and c is the speed of light. This famous equation shows that mass and energy are interchangeable.

- **General Relativity**: Formulated in 1915, General Relativity is a theory of gravity that describes the curvature of spacetime caused by the presence of mass and energy. Unlike Newtonian gravity, which considers gravity as an attractive force between masses, General Relativity attributes gravity to the curvature of spacetime caused by massive objects. Key principles of General Relativity include:

  - **Curved Spacetime**: Massive objects like stars and planets curve the fabric of spacetime around them, and other objects move along the curved paths in response to this curvature.
  - **Gravitational Time Dilation**: Clocks in stronger gravitational fields (e.g., near massive objects) run slower compared to clocks in weaker fields.
  - **Gravitational Waves**: General Relativity predicts the existence of gravitational waves, ripples in spacetime caused by violent cosmic events.

Both Special and General Relativity have been confirmed through numerous experiments and observations, and they have far-reaching implications for our understanding of the cosmos. Special Relativity's effects become significant at high speeds, near the speed of light, while General Relativity is essential for describing gravity and the behavior of massive objects, such as stars, galaxies, and black holes. These theories have not only revolutionized fundamental physics but have also influenced technology, astronomy, and our perception of the nature of space, time, and the universe.

16.1   Time

What is time? This is a big question and there are many different levels of definitions that can be given in varying contexts. One definition, given by the Oxford Languages Dictionary follows

> **❝ Quotation**
> "The indefinite continued progress of existence and events in the past, present, and future regarded as a whole."

This definition is not very useful in quantifying time in a useful manner. **Time** is a fundamental concept in physics and philosophy that represents the sequence of events, changes, or movements in the

universe. It is an abstract dimension in which events appear to occur in a chronological order, including the past, present, and future. Time appears as a continuous and one-directional flow, and it is often measured and quantified in relation to periodic or repetitive phenomena, such as the rotation of celestial bodies or the oscillations of certain particles. In modern physics, time is intricately connected to the fabric of spacetime, and its nature is further explored in theories like the theory of relativity and quantum mechanics. One can think of time in reference to something with regular periodicity, such as the rotation of the earth. We can just say that we base our definition of time on the repetition of some apparently periodic event. Of course many questions can follow.

> **? Questions**
>
> - How do we regard time as existing in the past or future (or even the now)?
>
> - How is time different for different observers?
>
> - Does everyone perceive time similarly?
>
> - Can we measure time accurately without relying on periodic events like the rotation of the earth?
>
> - Is time an objective reality or a human construct?
>
> - How does time interact with other fundamental aspects of the universe, such as space and matter?
>
> - Are there alternative ways to conceptualize time that differ from our current understanding?
>
> - Is time one-directional and can be be stopped, slowed down, or reversed?

For measurements of time, typically a periodic event is used and counted, such as a watch running at a stable **frequency**[1] For most day-to day events, this is simple and easy to do and we can get smaller units of time by dividing our periodic events into smaller equal units of time. For very long periods of time, atomic half-lives can be used to deduce time intervals, such as the age of objects. This typically relies on assumptions and assuming that the known methods and laws are the same in the past as they are now. For very small or consistent units of time, electronic oscillators or atomic clocks can be used. And example of an atomic clock is a cesium atomic. A **cesium atomic clock** is a highly accurate and precise timekeeping device that operates based on the resonant frequency of cesium atoms. It uses the transition between two energy levels in cesium atoms to generate a stable and constant microwave signal. By counting the number of microwave oscillations, cesium atomic clocks can measure time with incredible accuracy, making them essential tools for applications in scientific research, navigation, and communication systems.

> **? Questions**
>
> - Is there a smallest unit of possible time?

The relativity of space and time imply that the time measurements have a minimum error (similar to the common Heisenbery uncertainty principal) given by $\Delta t \geq \frac{\hbar}{2\Delta E}$.

---

[1] note that a frequency is not easily definable without the concept of time.

16.2   Distances

Distance refers to the numerical measurement of the physical separation between two points or objects in space. It is a fundamental concept in physics and is typically expressed in units such as meters, kilometers, feet, or miles. Distance is a scalar quantity, meaning it only has magnitude and no direction, and it plays a crucial role in various scientific and everyday contexts, including navigation, transportation, and spatial analysis. From a mathematical sense, distance can typically be defined and/or determined exactly. In reality, this is not the case. There are generally errors associated with measured, deduced, or calculated distances for physical matter.

> **? Questions**
>
> - Are distances an absolute measurement?

Triangulation is a geometric method used to determine the distance or position of a point by forming triangles with known sides or angles. In the context of surveying and navigation, triangulation involves measuring the angles from two known points to an unknown point and using trigonometric principles to calculate its distance or location. This technique is widely employed in various fields, such as cartography, geodesy, astronomy, and GPS (Global Positioning System) to accurately determine positions and distances on the Earth's surface or in three-dimensional space. Triangulation is commonly used to determind the distance that planets, meteors, satellites, stars and other objects are from an observer. Other methods of determining distances include using electromagnetic radiation and calculating the time associated with its travel to deduce distances. there are many unique methods for deducing both large and small distances.

Very small objects cannot be seen (objects that are smaller than the wavelength of visibuel light (about 5×10^{-7} meters). **Electron microscopes** can be used to 'photograph' objects down to about 10^{-8} meters. At even smaller scales, x-rays can be used to determine/deduce even smaller scales as low as 10^{-10} meters by looking at the pattern of scattering off of atoms in a crystal. At the atomic nucleus scale, we can deduce sizes by sending high energy particles through materials and calculating how many do not make it through - deducing a cross section of a nucleus. The radius of a nucleus can be determined by

$$\pi r^2 = \sigma = \frac{A}{N} \frac{n_1 - n_2}{n_1}, \tag{16.2.1}$$

where A is the area of a slab of material containing N atoms, σ is the cross-sectional area of the nucleus, n_1 is the number of particles sent into the material with a beam and n_2 is the number that passed out the other side. Using this equation in conjunction with experiments, the nuclei of atoms are found to be on the scale of 10^{-15} meters, which is called the **fermi**, in honor of Enrico Fermi (1901–1954). According to the Heisenberg uncertainty principal, there is no possible way to perfectly and precicely measure distances or times assiciated between particles of matter and the error in measurements of positions of objects must be at least as large as $\Delta x \geq \frac{\hbar}{2\Delta p}$.

16.3   Gravity

Most forces in nature are represented by fields defined on **spacetime** (i.e. the electromagnetic field). General Relativity instead states that gravity is inherent in spacetime itself and thus what is experienced as **gravity** is simply a manifestation of the **curvature of spacetime** [22]. General relativity is a theory of gravity.

> **? Questions**
>
> - Can gravity be understood in terms of fields?
>
> - Can the other forces understood as fields be understood as some sort of geometric curvature?
>
> - Isn't a geometric curvature just another form of a field (mathematically)?

16.4 Special Relativity

16.4.1 Time Dilation and Length Contraction

To begin, we start with Einsteins postulates which are given as

1. The laws of physics must always hold in all inertial reference frames.

2. The speed of light in a vacuum c is the same in all reference frames.

From this, we can explore the implications and effects of these assumptions. First, consider a frame[2] of reference $\mathcal{F}'$. Suppose we set up two mirrors in this frame that are facing each other and have a beam of light traveling back and forth between them. Let these mirrors be aligned such that the beam of light is traveling vertically and assume we are in a vacuum. If we set $t = 0$ as the point when the light beam hits the bottom mirror, then the time it takes for the light to travel to the top mirror can be determined simply by the distance formula

$$v = \frac{\Delta x}{\Delta t} \implies \Delta t = \frac{\Delta x}{v}. \tag{16.4.1}$$

Since we are in a vacuum, the speed of our light is simply c and thus we have $\Delta t' = \frac{\Delta x'}{c}$.

Now consider an inertial frame $\mathcal{F}$ such that the first frame is moving horizontally relative to it. The light will remain traveling between these two mirrors, only the observer in frame $\mathcal{F}$ will see the light traveling both horizontally and vertically (since $\mathcal{F}'$ is moving horizontally as seen in $\mathcal{F}$). Say the speed that $\mathcal{F}'$ is moving as seen from $\mathcal{F}$ is v. Assuming no acceleration, since light must travel at a constant speed in our vacuum, the distance the light travels from the bottom mirror to the top mirror as observed in $\mathcal{F}$ is longer than that of $\mathcal{F}'$. The vertical distance between the two plates is still $\Delta x'$, but now there is also a horizontal distance. Suppose it takes Δt for the light to hit the top mirror in $\mathcal{F}$, then the horizontal distance that the light traveled is given by $\Delta x_h = v\Delta t$ and so the total distance the light traveled in $\mathcal{F}$ is

$$\Delta x = \sqrt{(\Delta x')^2 + (\Delta x_h)^2} = \sqrt{(\Delta x')^2 + (v\Delta t)^2}. \tag{16.4.2}$$

Now, by definition, $\Delta x = c\Delta t$ and so if we combine this with (16.4.1) and (16.4.2) we have the equation for time dilation which is

$$(c\Delta t)^2 = (c\Delta t')^2 + (v\Delta t)^2 \implies \boxed{\Delta t = \frac{\Delta t'}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}}. \tag{16.4.3}$$

Now, we can take (16.4.1) and use it to relate $\Delta x'$ and Δx. First, notice that $\Delta x' = c\Delta t'$ and $\Delta x = c\Delta t$. Solving for c and setting these equal to each other gives

$$\frac{\Delta x'}{\Delta t'} = \frac{\Delta x}{\Delta t} \implies \frac{\Delta x'}{\Delta x} = \frac{\Delta t}{\Delta t'}. \tag{16.4.4}$$

---

[2]Note that $\mathcal{F}'$ is the "proper frame".

Combining this with (16.4.3) will then give us the relationship for length contraction between reference frames which is

$$\frac{\Delta x'}{\Delta x} = \frac{\Delta t}{\Delta t'} \implies \Delta x = \boxed{\Delta x' \sqrt{1 - \left(\frac{v}{c}\right)^2}}. \tag{16.4.5}$$

The factor of $\sqrt{1 - \left(\frac{v}{c}\right)^2}$ appears often within relativity and is generally denoted γ^{-1}. Therefore these equations can also be written as $\Delta t = \gamma \Delta t'$ and $\Delta x = \frac{\Delta x'}{\gamma}$.

16.4.2   Lorentz Coordinate Transformations

Consider again two frames, $\mathcal{F}$ and $\mathcal{F}'$. Suppose we have $\mathcal{F}'$ moving at some speed v as observed in the $\mathcal{F}$ frame in the $\hat{x}$ direction. Suppose there is a bar moving along with frame $\mathcal{F}'$ with length L. After some time, the bar will have moved a distance of vt and thus the new location will be at $x = L + vt \implies L = x - vt$. Now, by the length contraction we derived previously, this length can be related to the length in it's frame by $L' = \frac{L}{\gamma}$ and so we will have $L' = \gamma(x - vt)$. Now, say at $t' = 0$ that $x' = x$, and so we can set the position of the end of the bar at $x' = L'$ and thus since in the prime frame the position of the bar does not change it remains that $x' = L'$ after some time t and so we have $x' = \gamma(x - vt)$. Notice that if there is no velocity ($v = 0$), this becomes $x' = x$ and so our case, the respective y and z components become $y = y'$ and $z = z'$. These are all of the position transformations contained within the Lorentz coordinate Transformations.

$$\boxed{x' = \gamma(x - vt) \qquad\qquad y = y' \qquad\qquad z = z'} \tag{16.4.6}$$

We can now take the above position change in the x direction and derive the Lorentz transformation for time. Consider some change in position while moving at constant velocity in the prime frame

$$\Delta x' = x'_f - x'_i = \gamma(x_f - vt_f) - \gamma(x_i - vt_i) = \gamma(\Delta x - v\Delta t). \tag{16.4.7}$$

Dividing this by $\Delta t'$ and applying (16.4.4) yields

$$\frac{\Delta x'}{\Delta t'} = \frac{\gamma}{\Delta t'}(\Delta x - v\Delta t) = \frac{\Delta x}{\Delta t} \implies \Delta t' = \gamma\left(\Delta t - \frac{v\Delta t^2}{\Delta x}\right). \tag{16.4.8}$$

Since $\frac{Deltax}{\Delta t} = c$ (from our derivations of time dilation), we can square this and use some manipulation to get $\frac{Deltax^2}{\Delta t^2} = c^2 \implies v\Delta t^2 = \frac{v\Delta x^2}{c^2}$. Plugging this into (16.4.8) gives us

$$\Delta t' = \gamma\left(\Delta t - \frac{v\Delta x}{c^2}\right) \implies t'_f - t'_i = \gamma\left(t_f - t_i - \frac{v(x_f - x_i)}{c^2}\right) \tag{16.4.9}$$

$$\implies t'_f - t'_i = \gamma\left(t_f - \frac{vx_f}{c^2}\right) - \gamma\left(t_i - \frac{vx_i}{c^2}\right). \tag{16.4.10}$$

Examining this result allows us to clearly see a solution to any arbitrary time transformation that will be consistent with the above formulations and that result is the Lorentz time transformation. This is given by

$$\boxed{t' = \gamma\left(t - \frac{vx}{c^2}\right)}. \tag{16.4.11}$$

16.5   General Relativity

General relativity is a theory of gravity. Free particles move along paths of 'shortest possible distance', or geodesics. Geodesics is a profound feature of general relativity [22].

> **? Questions**
>
> - Is this related to **Maupertuis' Principle** and **Hamilton's Principal**?

Quantum Field Theory

Quantum Field Theory (QFT) is a theoretical framework that combines the principles of quantum mechanics with special relativity to describe the behavior of particles as quantized fields. It is one of the fundamental theories in modern theoretical physics and forms the basis for understanding particle interactions at both the microscopic and cosmological scales. Key aspects and principles of Quantum Field Theory include:

- **Quantized Fields**: In QFT, particles are described not as individual discrete entities but as quantized fields that pervade all of space and time. Each particle type corresponds to a specific field, and particles are represented as excitations or quanta of these fields.

- **Lagrangian Formalism**: QFT employs a Lagrangian formalism to construct the equations of motion and interactions for the fields. The Lagrangian describes the dynamics and symmetries of the system and allows for the derivation of fundamental equations, such as the equations of motion and scattering amplitudes.

- **Creation and Annihilation Operators**: QFT introduces creation and annihilation operators to describe particle creation and annihilation processes. These operators act on the field states to generate or destroy particles and provide a mathematical framework for quantized states.

- **Feynman Diagrams**: Feynman diagrams are a visual tool used in QFT to represent particle interactions and scattering processes. They provide a pictorial representation of complex particle interactions and are instrumental in calculating scattering amplitudes.

- **Renormalization**: Similar to Quantum Electrodynamics (QED), QFT encounters infinities in certain calculations. Renormalization is a method to remove these infinities and obtain finite, meaningful predictions.

- **Gauge Theories**: QFT includes gauge theories, which describe interactions involving force-carrying particles (gauge bosons). Examples of gauge theories include Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD).

- **Quantum Electrodynamics** (QED): QED is a specific example of Quantum Field Theory that describes the electromagnetic force and its interactions with charged particles through quantized electromagnetic fields and photons.

Quantum Field Theory has proven to be highly successful and is an essential framework for describing and understanding the behavior of elementary particles and their interactions. It is the foundation for the Standard Model of particle physics, which provides a comprehensive description of the known elementary particles and their interactions, except for gravity. Additionally, QFT plays a crucial role in studying fundamental questions about the nature of matter, energy, and the universe at the most fundamental level.

Quantum Electrodynamics

Quantum Electrodynamics (QED) is a quantum field theory that describes the electromagnetic force and its interactions with charged particles, such as electrons and photons. It is considered one of the most successful and accurate scientific theories ever developed and is a cornerstone of modern particle physics. Key aspects and principles of Quantum Electrodynamics include:

- **Quantization of Electromagnetic Fields**: QED treats the electromagnetic field as a quantum field, where photons are the quantized particles representing the discrete packets of electromagnetic energy.

- **Feynman Diagrams**: Feynman diagrams are a powerful tool used in QED to visualize and calculate particle interactions. They depict the exchange of photons between charged particles, enabling precise predictions of scattering processes and particle interactions.

- **Renormalization**: QED encounters infinities in certain calculations due to the self-interactions of charged particles with their own electromagnetic fields. Renormalization is a technique used to remove these infinities, allowing meaningful and accurate predictions.

- **Quantum Loops**: In QED, particles can interact with each other through quantum loops, where virtual particles (e.g., virtual photons) briefly pop into existence and influence the interactions between charged particles.

- **Gauge Invariance**: QED exhibits gauge invariance, meaning that different mathematical representations of the theory lead to physically equivalent results. This property ensures that observable quantities are independent of the choice of mathematical description.

- **Electron Self-Energy**: QED accounts for the electron's self-energy, which arises from its interaction with its own electromagnetic field. This self-energy correction leads to subtle shifts in electron properties, such as its mass and magnetic moment.

Quantum Electrodynamics has been extensively tested through precise experiments and is one of the most accurate and well-validated theories in physics. It successfully explains a wide range of phenomena, including electromagnetic interactions between charged particles, the behavior of light and photons, and the fine structure of atomic spectra. Moreover, QED has been instrumental in the development of other quantum field theories, such as Quantum Chromodynamics (QCD) and the electroweak theory, which describe the strong and weak nuclear forces.

18.1   Introduction

In this one theory we have the basic rules for all ordinary phenomena except for gravitation and nuclear processes. At the present time no exceptions are found to the quantum-electrodynamic laws outside the nucleus, and there we do not know whether there is an exception because we simply do not know what is going on in the nucleus. This theory predicted a lot of interesting things, such as very high energy photons, gamma rays, etc. Another remarkable prediction is that of antiparticles. These are particles of the same mass, but opposite charge. For example, the opposite of the electron is the positron and these two coming together would annihilate each-other with the emission of electromagnetic radiation. The generalization of this, that for each particle there is an antiparticle, turns out to be true.

Just as electrical interactions can be connected to a photon, Hideki Yukawa suggested that the forces between neutrons and protons also have a field of some kind, and that when this field jiggles, it behaves

like a particle. The characteristics were able to be deduced for this particle and experimentation was done. Around 1947, other particles were found, the π-**meson** (or **pion**). After this, more particles were discovered in cosmic rays and elsewhere which are distinct. We do not today understand these various particles as different aspects of the same thing, and the fact that we have so many unconnected particles is a representation of the fact that we have so much unconnected information without a good theory. In an attempt to classify and categorize these new particles similar to how elements are categorized in the periodic table, a new number (similar in concept to the electric charge) can be assigned to each particle, called its "**strangeness**," S. This number is conserved, like the electric charge, in reactions which take place by nuclear forces.

Particle	Mass (MeV)	Charge (e)	Strangeness	Category
Proton	938.272	+1	0	Baryon
Neutron	939.566	0	0	Baryon
Pion (π^+)	139.570	+1	0	Meson
Pion (π^-)	139.570	−1	0	Meson
Kaon (K^+)	493.677	+1	±1	Meson
Kaon (K^-)	493.677	−1	±1	Meson
Lambda (Λ^0)	1115.683	0	−1	Baryon
Sigma (Σ^+)	1189.37	+1	−1	Baryon
Sigma (Σ^-)	1197.45	−1	−1	Baryon
Xi (Ξ^0)	1314.86	0	−2	Baryon
Xi (Ξ^-)	1321.71	−1	−2	Baryon
Omega (Ω^-)	1672.45	−1	−3	Baryon
Electron	0.511	−1		Lepton
Neutrino	$< 2.2 \times 10^{-9}$	0		Lepton

Table 18.1: Classification of Particle Strangeness

All particles which are together with the neutrons and protons are called **baryons**, and the following ones exist: There is a "lambda," with a mass of 1115 MeV, and three others, called sigmas, minus, neutral, and plus, with several masses almost the same. In addition to the baryons the other particles which are involved in the nuclear interaction are called **mesons**. There are first the **pions**, which come in three varieties, positive, negative, and neutral; they form another multiplet. We have also found some new things called K-mesons, and they occur as a doublet, K^+ and K^-. Also, every particle either has an antiparticle or is its own antiparticle. There are then **leptons**, which are particles which do not interact strongly in nuclei, have nothing to do with a nuclear interaction, and do not have a strong interaction. Finally, we have two other particles which do not interact strongly with the nuclear ones: one is a photon, and perhaps, if the field of gravity also has a quantum-mechanical analog (a quantum theory of gravitation has not yet been worked out), then there will be a particle, a **graviton**, which will have zero mass[1].

---

[1]The term "zero mass" in this context refers to the particles rest energy. If a particle has "zero mass" then it can never be found at rest. A photon for example, will always travel at a speed near c.

Quantum Chromodynamics

Quantum Chromodynamics (QCD) is a fundamental theory in particle physics that describes the strong nuclear force, which is responsible for holding quarks together to form protons, neutrons, and other hadrons. QCD is a quantum field theory and is an essential component of the Standard Model, which describes the known elementary particles and their interactions. Key aspects and principles of Quantum Chromodynamics include:

- **Quarks** and **Gluons**: QCD postulates the existence of quarks, which are elementary particles that come in six flavors (up, down, charm, strange, top, and bottom). Quarks are bound together by exchanging gluons, which are the force-carrying particles of the strong nuclear force.

- **Color Charge**: In QCD, quarks carry a property called "color charge," which is analogous to electric charge in electromagnetism but comes in three types: red, green, and blue (plus their corresponding anticolors). Gluons, which also carry color charge, mediate the strong force between quarks.

- **Asymptotic Freedom and Confinement**: QCD exhibits two remarkable phenomena. At high energies or short distances, the strong force weakens, a property known as "asymptotic freedom." This allows for precise calculations in perturbation theory. However, at low energies or long distances, the force becomes strong, and quarks are confined within hadrons, making isolated quarks inaccessible in nature.

- **Lattice QCD**: Because QCD becomes strongly coupled at low energies, direct calculations become challenging. Lattice QCD is a numerical approach that uses a discrete space-time lattice to perform non-perturbative calculations of QCD phenomena.

- **Hadronization** and **Hadron** Structure: QCD governs the process of hadronization, where quarks and gluons combine to form color-neutral hadrons (e.g., mesons and baryons). QCD also provides insights into the structure and properties of hadrons, such as their masses and decay rates.

- **Strong Interactions** in Particle Colliders: QCD is crucial in understanding the strong interactions observed in high-energy particle colliders, where quarks and gluons are produced in energetic collisions. The study of these interactions helps test the predictions of QCD and explore new physics.

Quantum Chromodynamics is a fundamental theory in the Standard Model, working in conjunction with Quantum Electrodynamics (QED) and the Electroweak Theory to describe the behavior of elementary particles and their interactions. The study of QCD has led to significant advances in our understanding of the strong force, the nature of matter, and the dynamics of quarks and gluons within hadrons. QCD is also a subject of ongoing research and remains a key area of exploration in particle physics.

Stochastic Electrodynamics

Stochastic Electrodynamics (SED) is a theoretical framework that attempts to provide an alternative interpretation of **quantum mechanics** by combining classical **electromagnetism** with elements of quantum theory. It proposes that the underlying dynamics of particles and fields can be explained in terms of deterministic classical physics, but with the addition of stochastic (random) processes. Key points of Stochastic Electrodynamics include:

- **Deterministic Dynamics**: SED assumes that the behavior of particles and fields is governed by deterministic classical equations, specifically Maxwell's equations of electromagnetism. In this view, particles follow definite trajectories, and their behavior is predictable.

- **Stochastic Element**: Despite the deterministic dynamics, SED introduces a stochastic (random) element into the framework. This stochastic component accounts for the probabilistic nature of quantum phenomena, such as the wave-particle duality and probabilistic measurement outcomes.

- **Zitterbewegung**: An important concept in SED is "zitterbewegung," which refers to the rapid trembling or oscillation of a particle's motion around its classical trajectory. This trembling motion introduces the randomness required by quantum mechanics.

- **Zero-Point Field**: SED incorporates the concept of the zero-point field, a field of energy that exists even in the absence of any particles. This field interacts with particles and contributes to their motion and behavior.

- **Wave Function Collapse**: SED provides an interpretation of wave function collapse in quantum measurements. When a measurement occurs, the zitterbewegung of the particle undergoes a collapse due to the interaction with the zero-point field, leading to the observed probabilistic outcomes.

- **Correlation** and **Non-Locality**: SED suggests that the apparent non-local correlations observed in certain quantum experiments arise from the underlying dynamics of particles and fields, rather than from mysterious instantaneous actions at a distance.

- Critiques and Challenges: While SED presents an interesting attempt to bridge classical and quantum physics, it faces challenges in explaining all aspects of quantum phenomena, such as entanglement and the violation of Bell inequalities. Critics argue that SED may not fully account for the experimental results observed in quantum mechanics.

It's important to note that Stochastic Electrodynamics is a controversial and minority interpretation of quantum mechanics. Most mainstream interpretations of quantum theory, such as the Copenhagen interpretation and many-worlds interpretation, rely on fundamentally probabilistic principles. SED seeks to provide a deterministic foundation for quantum behavior, but its success in explaining the full range of quantum phenomena remains a subject of debate among physicists and researchers.

Geometrodynamics

Geometrodynamics is a term used in theoretical physics, particularly in the context of general relativity, to refer to a mathematical framework that describes the dynamics of spacetime geometry itself. It's a concept within the broader field of gravitational physics that seeks to explain the curvature of spacetime and how it changes over time as influenced by the presence of matter and energy.

- **General Relativity**: Geometrodynamics is closely related to Albert Einstein's theory of general relativity, which describes gravity as the curvature of spacetime caused by the presence of mass and energy. In this framework, massive objects cause spacetime to curve, and the curvature influences the paths that objects follow through space.

- **Space** and **Time** as Dynamic Entities: In geometrodynamics, both space and time are considered to be dynamic entities that can change and evolve. Instead of treating spacetime as a fixed background, as in Newtonian physics, geometrodynamics recognizes that the curvature of spacetime itself is influenced by the distribution of matter and energy.

- **Einstein Field Equations**: The dynamics of spacetime curvature are described by the Einstein field equations, a set of highly complex and nonlinear differential equations that relate the curvature of spacetime to the distribution of matter and energy within it.

- **Canonical Formulation**: Geometrodynamics can be expressed using various mathematical formulations. One of the most well-known approaches is the canonical formulation, which involves expressing the Einstein field equations in terms of variables like metric components and their conjugate momenta.

- **Quantum Geometrodynamics**: There have been attempts to apply the principles of quantum mechanics to geometrodynamics, leading to the development of quantum geometrodynamics or quantum gravity theories. These theories aim to describe the behavior of spacetime at very small scales and under extreme conditions where classical general relativity breaks down.

- Challenges: Geometrodynamics and its associated attempts to develop a quantum theory of gravity face significant challenges, such as reconciling general relativity with quantum mechanics, resolving the singularity problem at the centers of black holes, and providing a consistent description of spacetime at the Planck scale.

Geometrodynamics represents a unique and profound way of understanding the nature of gravity and the structure of the universe. However, due to its complexity and the challenges it presents, research in this area is ongoing, and finding a complete and satisfactory theory that unifies gravity with the other fundamental forces of nature remains one of the outstanding problems in theoretical physics.

References

[1] Bauer, W., and Gary D. Westfall. University Physics with Modern Physics. 2nd ed. Vol. 2. New York, NY: McGraw-Hill Companies, 2014. Print.

[2] http://www.dictionary.com

[3] Griffiths, David J. Introduction to Electrodynamics. 4th ed. N.p.: Pearson India Education Services Pvt, 2015. Print.

[4] Griffiths, David J. Introduction to Quantum Mechanics. 2nd ed. Harlow: Pearson, 2014. Print.

[5] Haberman, R. "Applied Partial Differential Equations: with Fourier series and Boundary Value Problems." 4th ed. Upper Saddle River, New Jersey. 2004. Print.

[6] Judson T. W. "Abstract Algebra: Theory and Applications." Stephen F. Austin State University. August 5, 2017.

[7] Reitz, John Richard. Foundations of Electromagnetic Theory. 4th ed. N.p.: Addison-Wesley, 1992. Print.

[8] Kittel, Charles, and Herbert Kroemer. Thermal Physics. San Francisco: W.H. Freeman, 1980. Print.

[9] Ross, Kenneth A. Elementary Analysis: The Theory of Calculus 2nd edition.

[10] Ryden, Barbara Sue., and Bradley M. Peterson. Foundations of Astrophysics. San Francisco: Addison-Wesley, 2010. Print.

[11] "Table of Isotopic Masses and Natural Abundances." (n.d.): n. pag. NC State University. Web. "https://www.ncsu.edu/chemistry/msf/pdf/IsotopicMass_NaturalAbundance.pdf".

[12] Taylor, John R. Classical Mechanics. Sausalito, CA: U Science, 2005. Print.

[13] "Wolfram|Alpha: Computational Knowledge Engine." N.p., n.d. Web. 2016 x

[14] "Wolfram MathWorld: The Web's Most Extensive Mathematics Resource." Wolfram MathWorld. N.p., n.d. Web.

[15] Zee, Dennis G. A first Course in Differential Equations with Modeling Applications." Edition 10. Book.

[16] Taylor, John R. (John Robert), 1939. An introduction to error analysis I John R. Taylor.-2nd ed.

[17] The Feynman Lectures on Physics

[18] IUPAC Periodic Table of the Elements and Isotopes. https://applets.kcvs.ca/IPTEI/IPTEI.html

[19] Table Of Nuclides. Nuclear Data Center at KAERI. https://atom.kaeri.re.kr/nuchart/

[20] Sakurai, J. J. (Jun John), 1933-1982. Modern quantum mechanics. - 2nd ed. I J.J. Sakurai, Jim Napolitano. p. cm. ISBN 978-0-8053-829 1-4 (alk. paper)

[21] Xavier Lirton, An Introduction to Underwater Acoustics: Principals and Applications, Praxis Publishing, 2004.

[22] Sean M. Carroll, An Introduction to General Relativity: Spacetime and Geometry,

Index